Mobile Information Systems

Big Data-Based Decision Making for Complex Information Systems

Lead Guest Editor: Jiafu Su
Guest Editors: Xuefeng Zhang, Aijun Liu, and Stavros Sindakis

Big Data-Based Decision Making for Complex Information Systems

Big Data-Based Decision Making for Complex Information Systems

Lead Guest Editor: Jiafu Su
Guest Editors: Xuefeng Zhang, Aijun Liu, and Stavros Sindakis

Contents

Contents

Contents

Contents

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9894243, 1 page
https://doi.org/10.1155/2023/9894243

Retraction

Retracted: Application and Analysis of Hypergraph Association Rule Redundancy Algorithm in Data Mining

Mobile Information Systems

Received 26 September 2023; Accepted 26 September 2023; Published 27 September 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] H. Pang and L. Zhou, "Application and Analysis of Hypergraph Association Rule Redundancy Algorithm in Data Mining," *Mobile Information Systems*, vol. 2022, Article ID 1193586, 11 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9871982, 1 page
https://doi.org/10.1155/2023/9871982

Retraction

Retracted: Fusion Analysis of Chinese Painting Color Teaching and Intelligent Image Color Processing Technology

Mobile Information Systems

Received 26 September 2023; Accepted 26 September 2023; Published 27 September 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] N. Zhao, "Fusion Analysis of Chinese Painting Color Teaching and Intelligent Image Color Processing Technology," *Mobile Information Systems*, vol. 2022, Article ID 3906277, 9 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9838372, 1 page
https://doi.org/10.1155/2023/9838372

Retraction

Retracted: Prediction of Long-term Variation of Ocean Circulation Based on Grey Correlation Clustering

Mobile Information Systems

Received 26 September 2023; Accepted 26 September 2023; Published 27 September 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] N. Ma and G. Yang, "Prediction of Long-term Variation of Ocean Circulation Based on Grey Correlation Clustering," *Mobile Information Systems*, vol. 2022, Article ID 4791054, 8 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9868629, 1 page
https://doi.org/10.1155/2023/9868629

Retraction

Retracted: Predicament and Thinking of College Students' Employment and Entrepreneurship under the Background of Supply-Side Reform

Mobile Information Systems

Received 22 August 2023; Accepted 22 August 2023; Published 23 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] M. Yu, "Predicament and Thinking of College Students' Employment and Entrepreneurship under the Background of Supply-Side Reform," *Mobile Information Systems*, vol. 2022, Article ID 1495123, 10 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9858357, 1 page
https://doi.org/10.1155/2023/9858357

Retraction

Retracted: Psychological Analysis of Athletes during Basketball Games from the Perspective of Deep Learning

Mobile Information Systems

Received 22 August 2023; Accepted 22 August 2023; Published 23 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Q. Meng, "Psychological Analysis of Athletes during Basketball Games from the Perspective of Deep Learning," *Mobile Information Systems*, vol. 2022, Article ID 4319437, 9 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9813079, 1 page
https://doi.org/10.1155/2023/9813079

Retraction

Retracted: Interior Soft Decoration Product Design Based on 3D Modeling and Image Processing Technology

Mobile Information Systems

Received 22 August 2023; Accepted 22 August 2023; Published 23 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Jin, Y. Chen, and J. Qu, "Interior Soft Decoration Product Design Based on 3D Modeling and Image processing Technology," *Mobile Information Systems*, vol. 2022, Article ID 9095614, 7 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9805461, 1 page
https://doi.org/10.1155/2023/9805461

Retraction

Retracted: The Construction of Green Building Integrated Evaluation System Based on BIM Technology

Mobile Information Systems

Received 22 August 2023; Accepted 22 August 2023; Published 23 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] H. Li and C. Wang, "The Construction of Green Building Integrated Evaluation System Based on BIM Technology," *Mobile Information Systems*, vol. 2022, Article ID 5906827, 12 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9796507, 1 page
https://doi.org/10.1155/2023/9796507

Retraction

Retracted: Exploring the Application of Online Teaching Platform in the PBL Teaching Mode of Film and Television Courses for Teacher Trainees

Mobile Information Systems

Received 22 August 2023; Accepted 22 August 2023; Published 23 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Peng and M. Yuan, "Exploring the Application of Online Teaching Platform in the PBL Teaching Mode of Film and Television Courses for Teacher Trainees," *Mobile Information Systems*, vol. 2022, Article ID 2289981, 13 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9790252, 1 page
https://doi.org/10.1155/2023/9790252

Retraction

Retracted: Integration and Recommendation of Multimedia Network-Assisted English Instructional Resources Based on Association Rules Mining

Mobile Information Systems

Received 22 August 2023; Accepted 22 August 2023; Published 23 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] L. Liu, "Integration and Recommendation of Multimedia Network-Assisted English Instructional Resources Based on Association Rules Mining," *Mobile Information Systems*, vol. 2022, Article ID 8806525, 10 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9783613, 1 page
https://doi.org/10.1155/2023/9783613

Retraction

Retracted: Hierarchical Strategies for Building Multiethnic Communities from the Perspective of Data Mining Analysis

Mobile Information Systems

Received 22 August 2023; Accepted 22 August 2023; Published 23 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] S. Qin, "Hierarchical Strategies for Building Multiethnic Communities from the Perspective of Data Mining Analysis," *Mobile Information Systems*, vol. 2022, Article ID 7506323, 10 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9765805, 1 page
https://doi.org/10.1155/2023/9765805

Retraction

Retracted: The Biomechanical Analysis of Jumping Difficulty Movement in National Traditional Sports

Mobile Information Systems

Received 22 August 2023; Accepted 22 August 2023; Published 23 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] L. Zhang and W. Yang, "The Biomechanical Analysis of Jumping Difficulty Movement in National Traditional Sports," *Mobile Information Systems*, vol. 2022, Article ID 7679116, 8 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9756026, 1 page
https://doi.org/10.1155/2023/9756026

Retraction

Retracted: Design of Intelligent Speech Translation System Based on Deep Learning

Mobile Information Systems

Received 22 August 2023; Accepted 22 August 2023; Published 23 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Tan, "Design of Intelligent Speech Translation System Based on Deep Learning," *Mobile Information Systems*, vol. 2022, Article ID 2463812, 7 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9842050, 1 page
https://doi.org/10.1155/2023/9842050

Retraction

Retracted: Exploring the Entrepreneurship Training Mode of Medical Students in Beijing-Tianjin-Hebei Universities under the Strategic Background of "Healthy China" from the Perspective of Mixed Big Data

Mobile Information Systems

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] L. Gong, C. Wang, and J. Bai, "Exploring the Entrepreneurship Training Mode of Medical Students in Beijing-Tianjin-Hebei Universities under the Strategic Background of "Healthy China" from the Perspective of Mixed Big Data," *Mobile Information Systems*, vol. 2022, Article ID 8177760, 17 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9796124, 1 page
https://doi.org/10.1155/2023/9796124

Retraction

Retracted: Ideology in Sino-Foreign Cooperative Education: The Application of Big Data Mining Technology in the Work Conducted by Counselors

Mobile Information Systems

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Z. Shu, "Ideology in Sino-Foreign Cooperative Education: The Application of Big Data Mining Technology in the Work Conducted by Counselors," *Mobile Information Systems*, vol. 2022, Article ID 4062281, 12 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9754217, 1 page
https://doi.org/10.1155/2023/9754217

Retraction

Retracted: Scientific Research Management Helping the Development of Regional Cultural Industry from the Perspective of Artificial Intelligence

Mobile Information Systems

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Wang and X. Hu, "Scientific Research Management Helping the Development of Regional Cultural Industry from the Perspective of Artificial Intelligence," *Mobile Information Systems*, vol. 2022, Article ID 1032081, 14 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 4450347, 8 pages
https://doi.org/10.1155/2023/4450347

Research Article

Automatic Scoring Model of English Interpretation Based on Semantic Scoring

Hua Ma

School of Foreign Languages, Chizhou University, Chizhou 247000, Anhui, China

Correspondence should be addressed to Hua Ma; 20151001086@m.scnu.edu.cn

Received 7 August 2022; Revised 11 September 2022; Accepted 23 September 2022; Published 17 April 2023

Academic Editor: Jiafu Su

In order to explore how English interpreting can achieve automatic scoring, the author proposes an automatic scoring model for English interpreting based on semantic scoring. This method recommends key technical problems and solutions based on information represented by semantic scoring, and explores the research on how interpreting can achieve automatic scoring of oral examinations. Research has shown that, the automatic scoring of English interpretation based on semantic scoring is faster than traditional methods, and the efficiency is improved by about 75%. However, the current automatic scoring of interpreters faces huge challenges. It needs to be tested and improved in more teaching, learning, and testing practice. The automatic scoring of interpretation should consider multiple dimensions such as semantic accuracy, content integrity, expressive fluency, and language authenticity.

1. Introduction

Smart education is the general trend of education development in the Internet era, and computers have become an important tool to assist learning [1, 2]. In the context of English language training and automated assessments, automatic question scoring systems that define candidate answer content such as reading questions and follow-up questions based on recent developments have reached a practical level. There is very little research. The interpretation test is a comprehensive test of foreign language application ability, including foreign language thinking ability and language organization ability. The research and development of an effective automatic scoring system for Chinese-English translation not only provides a platform for students to practice translation but also facilitates teachers' lectures and reduces the pressure of teachers' teaching and scoring. Based on this, by analyzing the scoring requirements of the interpreting test, focusing on the semantic scoring method at the content level of interpreting, a multiparameter Chinese-English sentence-level interpreting automatic scoring model is established as the basis for building an application system.

Taking the semantic scoring of spoken Chinese to English questions as the research focus, this paper introduces a semantic scoring model integrating long-short-term memory neural network and self-attention mechanism, which can be applied to keyword scoring and sentence semantic scoring [3]. The scoring principle of the model is as follows: firstly extract word and sentence features and represent them in a vectorized form, then use a bidirectional long-short-term memory neural network to optimize the feature vector, and then use the self-attention mechanism to obtain the semantic features of words or sentences, and finally the semantic score is calculated by a simple neural network. Experiments show that, compared with the stretchable recursive autoencoder-based semantic scoring model that performs better in semantic scoring, this model has better results in sentence semantic scoring. The average consistency rate between the sentence semantic scoring results and the original scores reached 55%.

Chinese-English translation quality evaluation has been one of the hotspots in the field of automatic Chinese-English translation quality evaluation in recent years. In terms of automatic spoken language scoring, most of the current research focuses on assessing spoken language at the level of pronunciation quality, such as reading questions and follow-up questions [4]. English reading questions were scored using the most probable linear regression and most probable north probability algorithms with moderate results. However, there is still a lack of effective evaluation strategies for research on question types related to textual content, such as explanatory questions and repetitive questions (keywords, main content of sentences). Although some scholars have carried out relevant research, the actual results of large-scale oral test scores are very limited [5, 6].

Therefore, we provide an automated Chinese-English quality translation method. To evaluate translation quality, we choose three main parameters: semantic keywords, sentence similarity, and speaking ability. In sentence-level Chinese-English translation, the translation of keywords must be meaningful, and the general meaning of Chinese-English sentences must also be accurate. As a spoken language translator, the fluency parameter is also very important, and fluency also reflects the overall level of the translator's spoken language. In the Chinese-English translation question and answer scoring research, researchers generally focus on assessing the accuracy of the Chinese-English translation and the respondents' comprehension of the entire sentence. This is also the main reason for choosing the previous three evaluation parameters. Since many Chinese-English translation questions are the main types of Chinese-English translation questions, automatic scoring of Chinese-English translation questions has practical uses. The framework of the automatic scoring system for Chinese-English spoken translation is shown in Figure 1.

2. Literature Review

Rajagede et al. said that in the 1960s, people began to study controlled automatic assessment of oral quality in the form of university research projects [7]. The world's first large-scale computer-assisted language learning system PLATO system is a programmable learning system for automatic teaching. It was developed in 1959 by the University of Illinois and its business partner, Control Data. Its appearance has greatly promoted the application of computer in foreign language learning. The second-generation CALL system, represented by ALLP of the Massachusetts Institute of Technology, studies the application of computers in the field of education. Liu said that after the 1990s, the research on the third-generation CALL system paid more attention to the application of human-computer interaction and multimedia in language learning [8]. The Stanford International Consulting Institute (formerly known as the Stanford Research Institute) has a research and development group focused on speech research. The VILTS (Voice Interactive Language Training system) system developed by the group is used to test students' intonation and pronunciation fluency. The system uses a posterior probability algorithm and a log-likelihood algorithm to calculate the speaker's pronunciation accuracy, while using the duration score to characterize the speaker's pronunciation fluency. Carnegie Mellon University has designed a special automatic scoring system for the SET-10 oral English test. Xia, L. believes that the system can achieve good results in judging the spoken English proficiency of non-native speakers, but the system does not automatically score the types of open-ended questions [9]. The SCILL algorithm and the simplified posterior probability algorithm jointly developed by the University of Cambridge and the Massachusetts Institute of Technology calculate the pronunciation accuracy. The simplified posterior probability algorithm greatly shortens the calculation time, as a result, system performance has been improved.

Zhang believes that there are not many related research studies on the oral quality assessment of semicontrolled topics in the United States, and the main research representatives are the TOEFL test system and the Pearson Academic English test system [10]. Peng Research on the exploratory use of support vector machines and classification and regression tree algorithms for question and answer scoring methods for the TOEFL test [11]. The study not only found that vector machines have the advantages of quantitative analysis but also found that the classification and regression tree algorithm is very effective in mining the underlying laws of data. The TOEFL test system uses the multiple linear regression method to integrate the four characteristic scoring parameters of intonation, grammar, fluency, and vocabulary diversity to calculate the test taker's score. The scoring system grades the TOEFL test's six test question types in turn. The Pearson Academic Test System, developed by Pearson, selects intonation, fluency, sentence proficiency, and vocabulary as characteristic scoring parameters. Different from the TOEFL test system, the system does not distinguish between question types and scores, and directly calculates the scores of the candidates' four scoring characteristics based on the candidates' answers.

In the 1970s, American researchers began to study translation quality. Translation is a theory put forward by western countries to evaluate the quality of translation. It emphasizes that the translator must fully express the emotion, goal, and meaning contained in the text in the translator's language on the basis of understanding the original text language. After the 1980s, Western scholars tried to quantify the quality of explanations through empirical research, trying to find scientific variables and proportions to evaluate the quality of explanations. Ban and Translators conducted a survey of translators' expectations and found that the most important indicator for translators to measure translation quality is content consistency, followed by translation coherence, translation completeness, and grammatical features correctness [12]. Yuan believes that this study lays the foundation for an empirical study of interpretation quality assessment [13]. Gillier et al. proposed that, at major international conferences such as medicine and law, the audience's views on the interpretation of quality assessment were summarized and case studies were conducted, and it was found that there were differences in the interpretation of the quality assessment standards between audiences and translators. Ismagilov and I interviewed members of the International Conference of Translators, and the study found that audiences and translators place a high value on the accuracy and clarity of translated content [14].

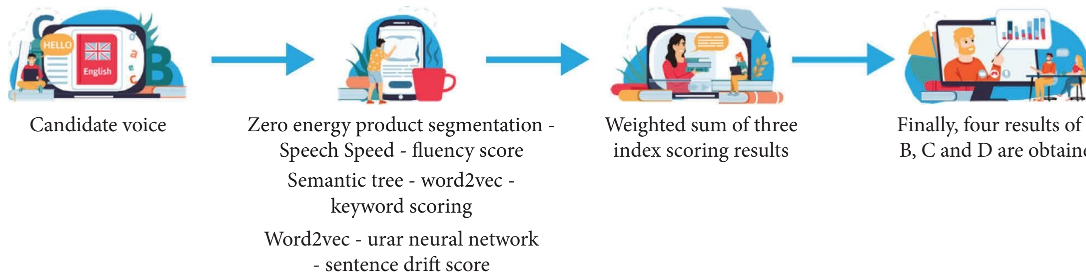


Figure 1: The framework of the automatic scoring system for Chinese-English spoken translation.

Compared with content, their requirements for expressions such as voice and speech speed are not high. Qin et al. introduced a semantic scoring model combining short-term and short-term memory neural networks and self-attention mechanism, which can be applied to keyword scoring and sentence semantic scoring. The scoring principle of the model is as follows: first, extract the features of words and sentences and express them in the form of vectorization, then optimize the feature vector using the two-way short-term memory neural network, and then use the self-attention mechanism to obtain the semantic features of words or sentences. Finally, a simple neural network is used to calculate the semantic score. The experimental results show that the average correlation between the model and the original score is 0.444, compared with the semantic scoring model based on the scalable recursive automatic encoder. The minimum coincidence rate with the original score is 95%. The highest consistency rate with neighboring countries is 74%. The automatic scoring model of Japanese interpretation based on semantic scoring has proved to be practical and has achieved good results [15].

The interpreting test and interpreting ability assessment can not only grasp the interpreting level of students, but also evaluate the teaching quality of teachers. It can be seen that the interpreting test and interpreting ability evaluation can provide an important basis for teaching improvement and have a certain guiding role. From the research status of interpreting quality assessment, it can be seen that information communication has become particularly important in interpreting assessment. In the scoring rules, it is rare to see scoring points that emphasize voice intonation, and more emphasis is placed on the integrity of information expression and the accuracy of information transmission.

3. Methods

3.1. Model Evaluation Indicators. An automatic scoring system is basically a computer model that scores on a rater's answer sheet, and the difference between the system scoring results and the manual scoring results reflects the performance of the automatic scoring system. To create an automated scoring system, you first need a standard set of human scoring data. This data are also known as the raw score of the test taker's answer. The aim of this study is to make the machine scoring results as close as possible to the candidate's initial score. We can evaluate system performance based on the correlation and consistency between the automatic scoring system scoring results and the initial scoring.

Correlation: correlation is an important metric for evaluating system scoring performance, which is used to measure the similarity between machine scores and initial scores in a linear sense. The Pearson correlation coefficient is used to measure the similarity between the machine score and the original score, and the calculation formula is shown in equation (1) as follows:

$$\rho\,\text{human, machine} = \frac{\sum_{n=1}^{N}\left[(Sn-\overline{S})\times(SRn-\overline{SR})\right]}{\sum_{n}^{N}(Sn-\overline{S})^{2}\times(SRn-SR)^{2}}. \quad (1)$$

Convenience: the initial score consistency assessment and automatic scoring model have two parameters: the consistency level and the adjacent stability level, based on the initial score and an explicit distribution of the automatic scoring model for different scores [16]. The fitness is the ratio of the number of samples at the same level to the total number of samples, that is, the formula for calculating the ratio of the number of samples with a full score of S to the number of samples N in the second formula as follows:

$$\text{consistency} = \frac{\mathrm{S}}{\mathrm{N}}\times 100\%. \quad (2)$$

The adjacent consistency rate refers to the ratio of the number of samples whose machine score differs from the original score by one level (less than or equal to 0.5 points) to the total number of samples N, and it can usually be used as an effective indicator for comparing the degree of consistency between the two; the calculation formula is as shown in equation (3) as follows:

$$\text{Adj}_{\text{consistency}} = \frac{S_{+0.5}}{N}\times 100\%. \quad (3)$$

The LSTM storage unit is mainly composed of memory cells, forgetting gates, input gates, and output gates. The forgetting gate is used to screen old cell information and update the current memory cells according to the memory cell candidate information generated by the input gate. The sigmoid activation function in the forget gate processes the input information and outputs a value between [0, 1] [17]. The output value indicates that the old storage unit information is stored, the output value 0 indicates that all the information in the old storage unit is forgotten, and the output value 1 indicates that all the information in the old storage unit is stored. The information state calculation formula of the forget gate is shown in equations (4) and (5) as follows:

$$f_t = \sigma\left(W_f \bullet [H_{t-1}, X_t] + b_f\right), \tag{4}$$

$$C_t = f_t * C_{t-1} + i_t * \overline{C}_t. \tag{5}$$

The input gate determines what new information can be added to the current memory cell. Each LSTM unit input includes the output Ht-x of the previous unit and the new information input Xr. The sigmoid activation function in the input gate processes the input information and outputs a value between [0, 1]. The output value represents the state in which the current information needs to be retained. Then, use the tanh function to generate new memory cell candidate information. The calculation formulas of the two functions are shown in equations (6) and (7), respectively, as follows:

$$i_t = \sigma\left(W_i \bullet [H_{t-1}, X_t] + b_i\right), \tag{6}$$

$$\overline{C}_t = \tanh\left(W_C \bullet [H_{t-1}, X_t] + b_C\right). \tag{7}$$

The output gate determines the output of the current cell information state [18]. As with the previous two gate designs, the sigmoid function is used to process the output result of the input gate and output a value between [0, 1]. The output value is multiplied by the tanh function value of the updated memory cell to obtain the final output result Ht. The calculation formula for the output gate is shown in equations (8) and (9) as follows:

$$o_t = \sigma\left(W_o \bullet [H_{t-1}, X_t] + b_o\right), \tag{8}$$

$$H_t = o_t * \tanh\left(C_t\right). \tag{9}$$

It can be seen from the LSTM unit structure diagram and calculation formula that the memory cell C is propagated through a simple linear transformation in the LSTM network, so it can remain in the LSTM model for a long time. By adding forgetting gates, input gates, output gates, and memory cells to the neural unit to screen memory information, the LSTM unit uses memory cells C to retain long-term memory, and the hidden layer H to retain short-term memory, realizing the processing and learning of long sequence data [19].

The sigmoid function can map real numbers to the interval (0, 1), but not centered at zero. In the case that the feature difference is relatively complex or the difference is not particularly large, the sigmoid function is better for text classification. The biggest disadvantage of the sigmoid function is that it is easy to cause the problem of gradient disappearance when backpropagating. The sigmoid function formula is shown in equation (10) as follows:

$$f(z) = \frac{1}{1 + \exp(-z)}. \tag{10}$$

3.2. Examination Interpreting Scoring Criteria. It can be seen from the scoring standard of interpreting that interpreting emphasizes the accuracy of information transmission and the fluency of expressing information. In the exam, grammar and pronunciation and intonation are not tested [20]. Therefore, the information transfer is divided into keyword score and sentence semantic score, and the score parameter at the phonetic level selects fluency. Combined with the opinions of the interpreting teachers, the key words, sentence semantics and pronunciation fluency are determined as the characteristic scoring parameters of the Chinese-English interpreting automatic scoring system, as shown in Figure 2. In order to facilitate the analysis of the experimental results, the original scores of candidates can be divided into four grades, as shown in Table 1.

3.3. Conduct Experiments. The recording of candidates for the interpreting and listening exam in a certain examination room was used as the experimental data set. Select the first 1–5 questions in Volume A and Volume B, a total of 10 Chinese-to-English sentence translation questions as the research object. In the original data, there are 328 candidates in Volume A, that is, each question has 328 recorded data, a total of 328 $*$ 5 = 1640 data; Volume B has a total of 334 candidates, that is, each question has 334 recorded data, a total of 334 $*$ 5 = 1670 pieces of data. Since the graders are graded according to the candidates' recordings, the quality of the recordings has a great influence on the grades [21]. In order to reduce the influence of this factor of recording quality, we screen the original recordings. We excluded recordings with no sound or loud ambient noise. In order to reduce the subjective influence of manual scoring, we select the data with the scoring results less than or equal to 0.5 points by two raters for the experiment. After screening, the experimental data are shown in Table 2.

The experimental data are generally divided into two parts, one part is used for modeling and the other part is used to test the model hypothesis. The dataset used to create the model is called the training package, and the dataset used to test the accuracy of the model's assumptions is called the test package. The average sample size of the Chinese-English sentence interpretation for the self-study test is 272, which is a small sample of machine learning data. If the data are simply divided into a test set and a data set, it is difficult to make full use of the data of the few-sample data set. This simple approach to data distribution makes it difficult to accurately assess model predictability. K-cross-validation is a common method for evaluating the predictability of a few-shot model. K-cross-validation refers to randomly dividing the experimental data into K subsets, using each subset as a test set in turn, and combining the remaining subsets into a training set. After the model is trained K times, the average of all test sets is taken as the final calculation result [22]. This method reduces the influence of a single test set and training set division method on the prediction results by calculating the average value of the model prediction performance of each subset. After repeated testing, the author uses 3-fold cross-validation to evaluate the model scoring accuracy, and the basic steps are as follows:

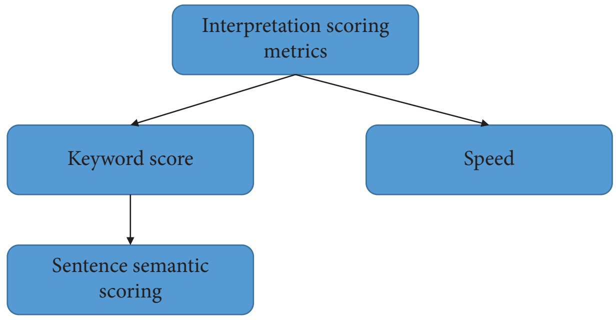


Figure 2: Scoring parameters of the automatic scoring model for interpretation.

Table 1: Scoring criteria and ratings.

Score	Grade	Grading criteria
[1, 2, 5]	A	The key information is accurate, the language is fluent, the vocabulary is used correctly, and the sentences are correct and popular.
[1, 5]	B	The key information is relatively accurate, the language expression is relatively fluent, the vocabulary application is relatively obtained, and the sentence generality is generally grasped.
[1, 0, 5]	C	The key information is not accurate enough, the language expression is insufficient, the sentences are biased, and the overall understanding is average.
[0, 5, 0]	D	Key information is incorrect or irrelevant, language is not fluent, sentences are highly distorted, and general comprehension is poor.

Table 2: Experimental data.

Volume A topics	Number of recordings	Volume B topics	Number of recordings
A_1	270	B_1	278
A_2	281	B_2	261
A_3	273	B_3	253
A_4	289	B_4	275
A_5	284	B_5	262

(1) Divide the data set into 3 subsets with basically the same amount of data,

(2) Use the first subset as the test set, and use the union of the remaining two subsets as the training set,

(3) Use the training set data to train the model, and use the test set data to verify the predictive ability of the model,

(4) Repeat steps 2-3, and take the remaining subset as the test set in turn.

In a small sample data set, if the training set contains the vast majority of sample data, theoretically the trained model can learn more data features [23]. However, at this time, the sample data in the test set will be relatively small, and the evaluation results are prone to large fluctuations, and the reliability will be reduced. If the test set contains more sample data and the sample data of the training set becomes relatively small, the model may not be able to learn the effective features of the data, thus reducing the credibility of the evaluation results. A common practice is to use about 2/3–4/5 of the samples as training data, and the remaining 1/5 to 1/3 of the samples as test data. Therefore, the division of the data set for each subject in this experiment follows the principle of training set: testing set = 7:3 for the experiment. The experimental data in this study has a total of 10 sentence translation scoring questions, and the average number of samples per sentence is 272, which is a small sample experiment in deep learning. In order to reduce the randomness of the experimental results and improve the reliability of the experiment, during the experiment, 10 questions were modeled and tested, and 3-fold cross-validation was used, that is, each question was tested 3 times, and the last 3 times were taken, the average value of the experiment is used as the experimental result of each question. Three experiments are carried out for each question, and Table 3 shows the data set of each experiment.

4. Results and Analysis

Sentence semantic score is not only to test the overall understanding of the candidate's sentence but also to reflect the candidate's ability to express the sentence. In the manual scoring, there is no specific item of sentence semantics, but the scorer will score the overall situation of the candidate's translation. The rater can directly tell by listening to the recording whether the examinee is speaking a complete sentence or a string of keywords. However, computers cannot easily make such judgments [24]. At the semantic

Table 3: Experimental dataset.

Topic	Number of training sets	Number of test sets
A_1	180	91
A_2	187	94
A_3	182	90
A_4	193	96
A_5	189	95
B_1	185	92
B_2	174	88
B_3	177	84
B_4	183	92
B_5	174	87

level, sentences usually consist of keywords and common words. Keywords are generally words that can affect the meaning of a sentence, consisting of nouns, verbs, and adjectives with specific meanings. The universal word is not decisive for understanding the meaning of the whole sentence, but it is an essential part of the sentence, such as prefixes, conjunctions, and sentences. The author compares the sentence semantic scoring model based on the stretchable recurrent autoencoder neural network with the proposed BiLSTM-AM-based semantic scoring model and analyzes the advantages and disadvantages of the two models in sentence semantic analysis.

Stretchable recursive autoencoders for semantic detection: this model is based on a recurrent neural network and improved with autoencoder, which can extract effective features of sentence generality [25]. Recurrent neural network processes sentence sequence information through tree structure; the basic process is to combine the input sentences according to the order of their network nodes to generate parent nodes, and then process the newly generated parent nodes and other child nodes as input again. The model recurses from bottom to top until all child nodes are integrated, and the characteristics of the last root node are obtained. Moreover, this feature can be thought of as a feature extraction representation for all input nodes. The autoencoder is divided into an encoding layer and a decoding layer. The former obtains another representation of the input data by encoding and compressing the features of the input data. The latter restores the original input by decoding. If the recovery result of the decoding layer is very close to the input data properties, it is assumed that the encoded function can represent an approximation of the input. The URAE-based sentence semantic scoring model studies sentence semantic properties through encoded latent layer neurons. The repetitive neural network of the autoencoder, combined with URAE, reconstructs the function compressed into the parent node to form the child node, filtering out the nature of the error measurement method between the atomic node and the reconstructed child node. If the error is too large, it means that the effect of the adjacent node merging is not good, and it is necessary to continuously traverse and optimize to combine the two adjacent nodes with the smallest reconstruction error.

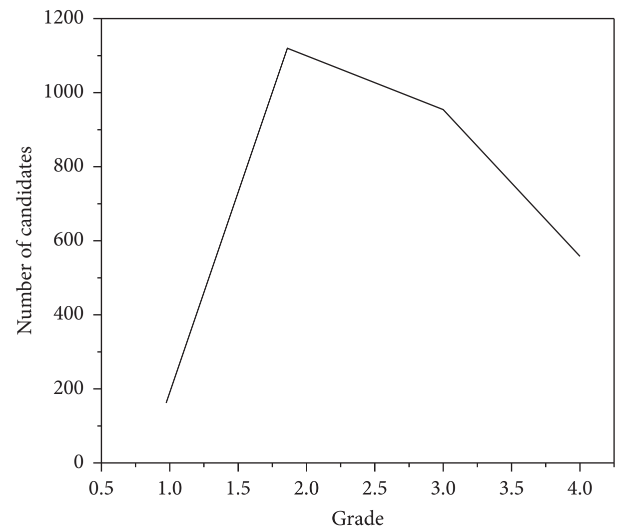


Figure 3: Distribution of manual scoring levels.

To minimize the effect of fictitious subjectivity on grades, the authors used data from answer sheets that were graded independently at multiple points, with the average of the two teachers' scores as the initial score. To confirm the validity of the initial score and the accuracy of the test data, we confirmed the accuracy of the machine scoring model by subtracting the speech data from over 0.5 points and comparing the difference in scores. Two-point scorer 1, scorer 2, or the histogram of the first score follows a normal distribution, with a high number of candidates with moderate scores, and a small number of candidates with excellent and unsuccessful scores. Research shows that, in general, the distribution of student grades follows a normal distribution, which demonstrates the accuracy of student assessments across grades and the reliability of initial grades. Figure 3 shows the distribution of the two scorers and the original score.

In order to test the superiority of the random memory algorithm in melting point, the control variable method was used to conduct comparative experiments under the condition that the speech rate scoring method, the keyword scoring method, and the sentence meaning scoring method were compatible. Stochastic memory algorithms are replaced by linear regression forecasting methods for integral supply. First, we compare the score distribution of the scores of the two models with the score distribution of the original scores. Whether the random memory algorithm or the linear regression prediction algorithm is used to combine the scores, it can be seen that the final score of the two-point model conforms to the normal distribution law. From the four-level distribution of ABCD, the automatic scoring model using the random forest algorithm for score fusion is closer to the original score distribution.

Original scoring protocols is compared using the automated scoring model scoring results and different scoring methods. The correlation between the scores of the two models and the initial score is more than 90%, while the score of the automatic scoring model using the random memory algorithm is generally 100% adjacent to the initial score. In terms of average transaction performance, the automatic scoring model using the random memory algorithm reached 77.4%, and the automatic scoring model using the linear regression prediction method reached 55.5%. The automatic model using the random forest algorithm has an average agreement rate of 21.9% higher than using the linear regression method.

5. Conclusion

It is undeniable that doubts about the automatic scoring system have never stopped since it appeared in the field of language testing. People may not believe that AI technology can be used for automatic scoring of interpretation, and think that it cannot make correct judgments on the content of interpretation. We believe that, from the perspective of formative assessment, applying AI technology in low-risk teaching assessment or learning diagnosis to help interpreting teachers and students, still a viable evaluation option.

Because there is a certain error in the ASR results, the final score prediction model will also be affected to varying degrees due to differences in recording quality. In the future, with the technical development of the automatic scoring system for interpretation, the speech features of the expression dimension of the CSE interpretation scale can be further added to the algorithm model, which can not only effectively avoid recording quality problems, but also help to give more detailed speech suggestions. In the future, deep learning algorithms will be used to directly model raw speech and textual information, enabling automatic scoring systems to better learn how to utilize textual semantic features for interpretation score prediction. In terms of teaching use, we will continue to enrich descriptions and feedback scenarios. The automatic grading system for interpreting can also provide interpreting students with more personalized and detailed feedback (such as specific error examples and correction suggestions) in the future, establish milestones in the development of interpreting abilities, form the overall trend of learning and personal portraits of students' learning, and teach teachers and students test to provide more reference.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] A. Barwise, M. E. Yeow, and D. K. Partain, "The premise and development of check in—check-in for exchange of clinical and key information to enhance palliative care discussions for patients with limited English proficiency," *American Journal of Hospice and Palliative Medicine*, vol. 38, no. 6, pp. 533–538, 2021.

[2] J. Zhang and F. Wang, "A better medical interpreting service: interpreter's roles and strategies under goffman's participation framework," *International Journal of Translation Interpretation and Applied Linguistics*, vol. 3, no. 1, pp. 1–14, 2021.

[3] Y. S. Kashikar, "The clinical images in the works of jhumpa lahiri: an approach in medical humanities," *College English*, vol. 17, no. 4, pp. 2456–8104, 2020.

[4] N. Voievodina, "Plurilingual education today: a perspective of teaching intercultural competence in German after English to student interpreters/translators," *ARS LINGUODIDACTICAE*, vol. 5, pp. 45–54, 2020.

[5] M. Leather, G. Fewings, and S. Porter, "Outdoor education: the romantic origins at the university of st mark and st john," *History of Education Review*, vol. 49, no. 1, pp. 85–100, 2020.

[6] A. M. El-Zawawy, "Simultaneous interpretation of complex structures from English into Arabic," *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics*, vol. 35, no. 1, pp. 30–64, 2022.

[7] R. A. Rajagede, "Improving automatic essay scoring for Indonesian language using simpler model and richer feature," *Kinetik Game Technology Information System Computer Network Computing Electronics and Control*, vol. 6, no. 1, pp. 11–18, 2021.

[8] J. Liu, L. Lin, and X. Liang, "Intelligent system of English composition scoring model based on improved machine learning algorithm," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, pp. 2397–2407, 2021.

[9] L. Xia, D. Luo, J. Liu, M. Guan, Z. Zhang, and A. Gong, "Attention-based two-layer long short-term memory model for automatic essay scoring," *Journal of Shenzhen University Science and Engineering*, vol. 37, no. 6, pp. 559–566, 2020.

[10] Z. Yuan, "Interactive intelligent teaching and automatic composition scoring system based on linear regression machine learning algorithm," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, pp. 2069–2081, 2021.

[11] H. Peng and Q. Li, "Research on the automatic extraction method of web data objects based on deep learning," *Intelligent Automation and Soft Computing*, vol. 26, no. 3, pp. 609–616, 2020.

[12] H. Ban and J. Ning, "Design of English automatic translation system based on machine intelligent translation and secure internet of things," *Mobile Information Systems*, vol. 2021, no. 7639, pp. 1–8, Article ID 8670739, 2021.

[13] Z. Yuan, C. Jin, and Z. Chen, "Research on language analysis of English translation system based on fuzzy algorithm," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 4, pp. 6039–6047, 2021.

[14] I. I. Ismagilov, A. I. Sabirova, D. V. Kataseva, and A. S. Katasev, "Collection scoring models development and research based on the deductor analytical platform," *Nexo Revista Científica*, vol. 33, no. 2, pp. 608–615, 2021.

[15] W. Qin, "Automatic scoring model of Japanese interpretation based on semantic scoring," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 4, pp. 5809–5819, 2020.

[16] Q. Xiong, "Research on English spoken semantic recognition machine learning model based on neural network and statistics fusion," *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 6, pp. 7341–7350, 2020.

[17] J. Xu and C. Yi, "The scoring mechanism of players after game based on cluster regression analysis model," *Mathematical Problems in Engineering*, vol. 2021, no. 3, pp. 1–7, Article ID 5524076, 2021.

[18] Y. Wang, “Implications of blended teaching based on theory of semantic wave for teaching English writing in high school,” *Journal of Higher Education Research*, vol. 3, no. 2, pp. 166–168, 2022.

[19] L. Duan, K. Yang, and R. Lang, “Research on Automatic Recognition of Casting Defects Based on Deep Learning,” *IEEE Access*, vol. 9, pp. 12209–12216, 2020.

[20] X. Li, F. Xu, X. Lyu et al., “Dual attention deep fusion semantic segmentation networks of large-scale satellite remote-sensing images,” *International Journal of Remote Sensing*, vol. 42, no. 9, pp. 3583–3610, 2021.

[21] J. Chen and S. Ou, “Research on the construction of the semantic model for Chinese ancient architectures based on architectural narratives,” *The Electronic Library*, vol. 38, no. 4, pp. 769–784, 2020.

[22] D. Selva, B. Nagaraj, D. Pelusi, R. Arunkumar, and A. Nair, “Intelligent network intrusion prevention feature collection and classification algorithms,” *Algorithms*, vol. 14, no. 8, p. 224, 2021.

[23] J. Hu, Y. M. Kang, Y. H. Chen, X. Liu, X. Li, and Q. Liu, “Analysis of aerosol optical depth variation characteristics for 10 years in urumqi based on MODIS_C006,” *Huan jing ke xue Huanjing kexue*, vol. 39, no. 8, pp. 3563–3570, 2018.

[24] R. Huang, S. Zhang, W. Zhang, and X. Yang, “Progress of zinc oxide-based nanocomposites in the textile industry,” *IET Collaborative Intelligent Manufacturing*, vol. 3, no. 3, pp. 281–289, 2021.

[25] Q. Zhang, “Relay vibration protection simulation experimental platform based on signal reconstruction of MATLAB software,” *Nonlinear Engineering*, vol. 10, no. 1, pp. 461–468, 2021.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 5008465, 16 pages
https://doi.org/10.1155/2022/5008465

Research Article

"Dimension Reduction: Feature Subset" Method for Selecting the Best Index Combination in Reputation Evaluation of Crowdsourcing Participants

Yanrong Huang,[1] **Zhan Zheng**,[2] **and Bo Wei**[3]

[1]*College of Economics & Management, Zhejiang University of Water Resource and Electric Power, Hangzhou 310018, China*
[2]*School of Media & Communication, Wuhan Textile University, Wuhan 430073, China*
[3]*School of Computer Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China*

Correspondence should be addressed to Zhan Zheng; czheng@wtu.edu.cn

Received 6 July 2022; Revised 3 September 2022; Accepted 14 September 2022; Published 12 October 2022

Academic Editor: Xuefeng Zhang

An effective reputation evaluation mechanism is an essential guarantee for the crowdsourcing mode's healthy, orderly, and rapid development. Aiming at the problems of unsound reputation evaluation mechanism, single reputation evaluation index, and poor discrimination ability of crowdsourcing platforms a "dimension reduction feature subset" method for selecting the best reputation evaluation index combination of crowdsourcing participants is proposed. This method first selects the best dimensionality reduction method by empirical method, then uses the classifier as the evaluation function of feature selection, and uses the sequential backward selection strategy (SBS) to select the feature subset and reputation evaluation algorithm with the best classification performance. The experimental results show that the reputation evaluation method of crowdsourcing participants based on ReliefF-SVM has the best performance in terms of accuracy, F1 measure, and stability and can select a comprehensive, objective, and effective evaluation index combination to distinguish the reputation status of crowdsourcing participants.

1. Introduction

Crowdsourcing is an online activity in which the task publisher convenes the crowdsourcing participants to complete the task through the crowdsourcing platform. It is an effective way to solve problems from a long distance [1]. Crowdsourcing participants provide knowledge, wisdom, experience, and skills through the crowdsourcing platform, participate in tasks, and receive remuneration. The task publisher publicly convenes the participants of crowdsourcing through the crowdsourcing platform to participate in the completion of the task and pay compensation. Crowdsourcing platform is the intermediary and bridge between task publishers and task participants [2]. Crowdsourcing is an open innovation form that can pool talents from all fields to participate in technological innovation and value creation, stimulate the innovation vitality of skills, provide valuable achievements [3, 4], and help enterprises solve the problems they face.

Hundreds of crowdsourcing platforms such as Zhu Bajie, Yipin Weike, Upwork, and Freelancer have emerged at home and abroad. Most domestic and foreign crowdsourcing platforms use transaction amounts and positive feedback to evaluate the reputation of crowdsourcing participants [1, 2]. The reputation evaluation index is single, and the reputation evaluation mechanism is not sound, so it is not easy to give comprehensive feedback on the reputation of crowdsourcing participants.

As the reputation evaluation mechanism of the crowdsourcing platform is not sound, the information on both sides of the transaction is asymmetric, and traders' speculative psychology [5] leads to frequent violations of crowdsourcing participants. Crowdsourcing participants ask for added remuneration, submit plagiarized results, provide the same result to participate in multiple tasks, are unable to complete tasks, fail to complete tasks as required, blackmail task publishers for evaluation, fail to provide follow-up maintenance services, maliciously lower prices to rob customers, conduct false transactions, and guide offline transactions and other violations. Some crowdsourcing platforms in China, such as Zhu Bajie, have taken ex-postpunishment measures such as restricting trading and closing accounts for illegal traders, but the effect is insignificant.

When the reputation evaluation mechanism of the crowdsourcing platform is challenging to regulate and restrict the behavior of participants effectively, a large number of crowdsourcing participants with a poor reputation will disrupt the trading order at low prices, and the crowdsourcing participants with a good reputation will leave the market one after another, resulting in the "Lemon Effect," causing the collapse of the crowdsourcing market. Whether we can select the key indicators for the reputation evaluation of crowdsourcing participants and establish an effective reputation evaluation mechanism is directly related to the healthy, orderly, and rapid development of crowdsourcing activities [6, 7].

However, the academic research on the reputation of e-commerce platform participants primarily concentrates on physical commodity trading and financial services, which provides a reference for constructing a reputation evaluation mechanism for a crowdsourcing platform. The research on reputation evaluation of crowdsourcing platforms mainly focuses on reputation evaluation methods, ignoring the discussion of evaluation indicators. The premise of practical evaluation is whether to select evaluation indicators that can significantly distinguish the reputation of crowdsourcing participants. Therefore, the vital issue is removing the indicators that have little impact on the reputation of crowdsourcing participants from the numerous evaluation indicators and selecting the best combination that can significantly distinguish the reputation of crowdsourcing participants.

The selection of the best indicator combination for the reputation evaluation of crowdsourcing participants is mainly a research method problem. How to rely on the Internet's open and real-time participation environment, make use of the large amount of behavior data generated by the online transactions of crowdsourcing participants, and use machine learning technology to select the best indicator combination for the reputation evaluation of crowdsourcing participants? Based on research paradigm III of "computational, experimental paradigm of selection behavior research" in [7], that is, following the idea of "raising questions, data collection, data analysis, and theoretical conclusion," two major problems need to be solved in this study. First, find the data dimensionality reduction method and find the indicators related to the reputation evaluation of crowdsourcing participants from a large number of index data. Second, find the method to obtain the best feature subset, establish the correlation between the optimal index combination and the machine learning algorithm, and find the best index combination that can significantly distinguish the reputation of crowdsourcing participants.

2. Literature Review

In the context of the Internet, the reputation evaluation of network platform participants has attracted extensive attention from the academic community and has become one of the frontier issues in e-commerce. Scholars have made rich achievements in the research of e-commerce platforms, which provides an essential theoretical basis for the reputation evaluation of crowdsourcing participants.

2.1. Research on Selection Method of Reputation Evaluation Index. The selection method of the reputation evaluation index is divided into single index selection and index combination selection.

The single index selection method selects a better group of indexes to build an index system according to the discrimination ability of a single index. Jiang et al. [8] analyzed that the number of frauds has a great impact on the reputation of platform traders. Yan et al. [9] proposed to measure the credibility of participants through active factors and historical factors. Liu et al. [10] concluded through literature analysis that professional level, work speed, work attitude, smooth communication, after-sales service, and innovation are the key factors to evaluating reputation. Zhang et al. [11] proposed to take interest and ability as the reputation evaluation index of crowdsourcing participants. Wang et al. [12] proposed that service quality and user score are important indicators affecting the reputation of crowdsourcing participants.

Although the index combination selected by the single index selection method can improve the discrimination ability of the model, it cannot guarantee that the selected index combination has the most robust discrimination ability. To make up for this shortcoming, scholars put forward the selection method of the index group. Jiang et al. [13] used the SBS method to screen P2P platform users' reputation evaluation indicators. Zhang et al. [14] used, CHAID, C5.0, and CART, three decision tree models to screen farmers' reputation evaluation indicators. Li et al. [15] extracted key indicators affecting personal credit using the Sparse Bayesian model. Wei et al. [16] proposed an optimal feature subset classification algorithm for the self-tuning particle swarm optimization algorithm. Zhang et al. [17] used the SVM model optimized based on the firefly algorithm to study the financial evaluation indexes of the supply chain of small- and medium-sized enterprises. Zhao et al. [18] mined key default characteristics of farmers based on the least significant difference method. Zhang et al. [19] used elastic network regression to select a credit characteristic index and determine listed enterprises' credit evaluation index systems in China's A-share market.

At present, the research on the selection of index combinations mainly focuses on the financial field, and there is less research on service e-commerce. The research results of the reputation evaluation index are difficult to be transplanted to the evaluation of the reputation of crowdsourcing participants.

2.2. Research on Reputation Evaluation Method. In recent years, the reputation evaluation method has attracted the attention of academia and has achieved a series of research results. At present, scholars' research mainly focuses on game theory and mathematical models. The research on reputation evaluation using the game theory method includes the following. Zhan et al. [20] established a multi-objective Stackelberg game model under incomplete information, trying to solve the problem of maintaining the reputation of the cross-border e-commerce platform when the reputation maintainer faces one or more damages. Quan et al. [21] discussed a reputation evaluation mechanism considering tolerance based on game theory. Wang et al. [22] established a reputation update method for free-riding and false data, which improved the credibility of the system. Lu et al. [23] designed a new optimal rating protocol based on game theory. Zhu et al. [24] designed an incentive mechanism based on game theory to make candidate nodes more willing to give honest reputation verification results.

The research of using a mathematical model to evaluate reputation includes the following. Yan et al. [9] improved the mean model, considered the activity factor and historical factor in the model, and proposed the worker reputation model based on activity. Bhattacharjee et al. [25] put forward the QNQ reputation model, using a regression method to take quantity and quality as judgment indicators to distinguish honest, selfish, and fraudulent users. Lu et al. [26] incorporated the comment text into the reputation evaluation and constructed the reputation calculation model. Sun et al. [27] studied the vehicle crowdsourcing service that provided users with real-time traffic feedback information and calculated participants' reputation values through trust propagation and feedback similarity. Huang et al. [1] proposed a multidimensional weighted cumulative reputation calculation model, MWCRM.

Scholars have begun to explore the use of machine learning technology for reputation evaluation in recent years. For example, Al Quadri et al. [28] used machine learning methods to predict the reliability of consumers from their data. Rantanen et al. [29] constructed an index system from the dimensions of quality, reliability, responsibility, success, pleasure, and innovation. They used a convolutional neural network (CNN) to classify and evaluate online corporate reputation. Yang et al. [30] used a semi-SVM model to measure the reputation of online service providers. Wang et al. [12] proposed an HMRep reputation evaluation method that could effectively resist malicious evaluation and improve calculation accuracy. In addition, in recent years, scholars have studied issues related to reputation management. Yu et al. [31] proposed an adaptive fog blockchain reputation storage method to improve the security and effectiveness of reputation management systems. Kealeboga [32] used the life cycle theory combined with the geographical location information of participants to judge the quality, reliability, and credibility of the dataset of reputation contribution data of crowdsourcing participants.

In conclusion, the research on crowdsourcing reputation evaluation mechanisms has made great progress. However, the method of game theory shows a strong explanatory power of social phenomena through deductive reasoning. However, the assumption of the ideal subject of game theory often deviates from the actual phenomena. Mathematical model methods need to make assumptions or judgments first, and the prediction results will be expressed as one or a group of functional relations, but they fail to combine the participant behavior generated by the crowdsourcing platform with massive amounts of data for reputation evaluation.

In the context of big data, how to use the massive data generated by the crowdsourcing platform and consider the relationship between the selection of reputation evaluation indicators and evaluation methods to evaluate the reputation of crowdsourcing participants remains to be discussed. How to objectively and accurately identify the key index combination of reputation evaluation from many indicators and put forward the best evaluation index combination that is objective and can significantly distinguish crowdsourced reputation is a question

Based on the analysis of literature, to start from the reputation characteristics of crowdsourcing participants, firstly, by systematically combining the relevant literature and learning from the reputation evaluation indicators of a crowdsourcing platform, we preliminarily select the reputation evaluation indicators of outstanding crowdsourcing participants. Secondly, using empirical research methods, select the best data dimensionality reduction method. Thirdly, the combination of the search strategy and machine learning algorithm is used to establish the association between the optimal index combination and the optimal machine learning algorithm, mine the combination effect of different data dimensionality reduction methods and machine learning algorithms, and put forward the best index combination selection method.

3. Research Procedures

Based on the computational experimental paradigm of choice behavior research, we put forward theoretical conclusions based on experiments, which mainly include three steps.

First is data acquisition and preprocessing. Considering the characteristics of reputation, following the selection principle of evaluation indicators, analyzing the reputation evaluation indicators of China's mainstream crowdsourcing platforms, systematically combining the relevant literature, and preliminarily screening 28 reputation evaluation indicators of crowdsourcing participants [33], we collect reputation evaluation index data from the Zhu Bajie crowdsourcing platform and preprocess the data.

Secondly, the best data dimensionality reduction method is selected. To find the most suitable method for dimensionality reduction of reputation evaluation indicators of crowdsourcing participants, four methods, ReliefF, mean impact value (MIV), principal component analysis (PCA), and linear discriminant analysis (LDA), are selected for comparative analysis. Select PCA method and ReliefF to calculate the index weight, sort according to the weight of the evaluation index, judge the data noise, and delete the

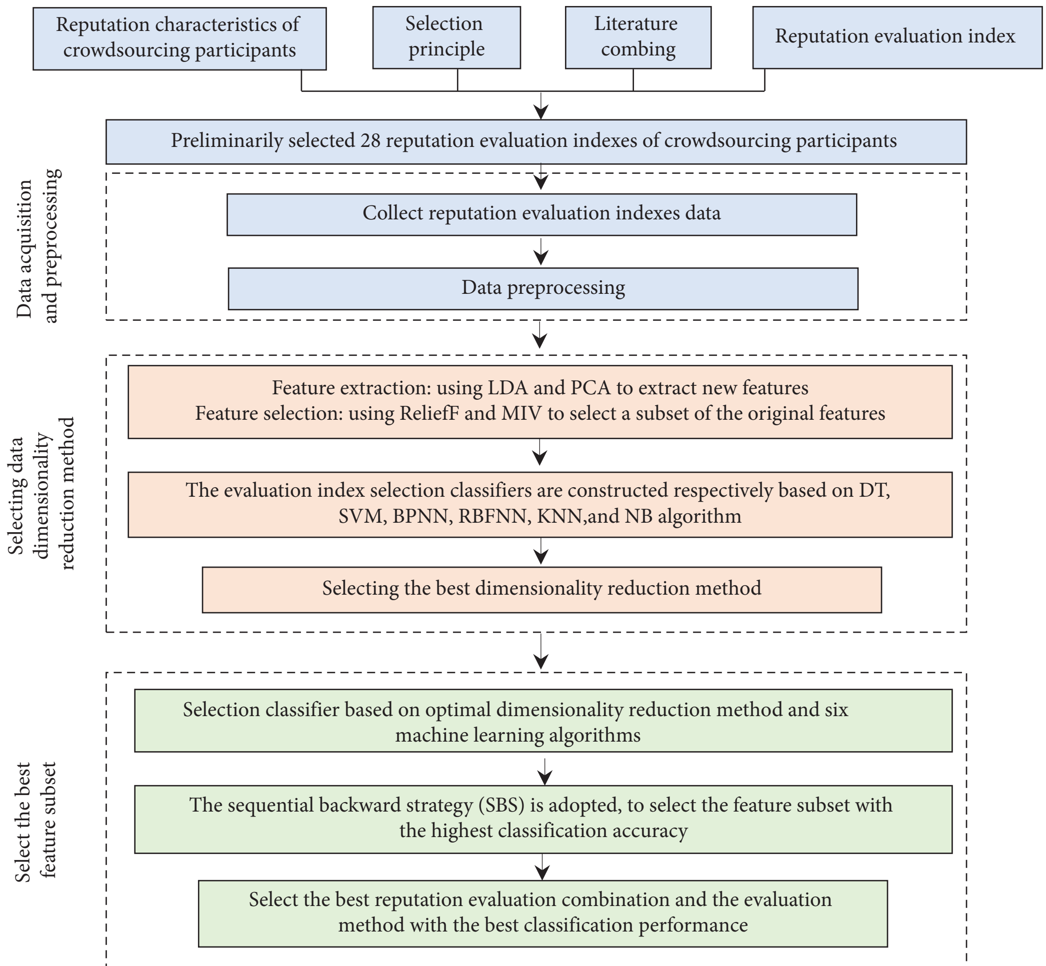


Figure 1: Selection steps of the best evaluation index combination of the reputation of crowdsourcing participants.

redundant index with weak reputation discrimination ability and small impact on crowdsourcing participants.

Data mining is the combined effect of a data dimensionality reduction method and a machine learning algorithm. Six commonly used machine learning classification algorithms are selected to cross construct 24 classifiers with four data dimensionality reduction methods. Taking the reputation evaluation dataset of crowdsourcing participants on the Zhu Bajie platform as an example, classification and prediction are carried out. Six machine learning algorithms include Decision Tree (DT), BP Neural Network (BPNN), RBF Neural Network (RBFNN), Support Vector Machine (SVM), K-Nearest Neighbor Classifier (KNN), and Naive Bayes (NB). Taking Zhu Bajie dataset as an example, the data dimensionality reduction method with the best overall classification effect of the six algorithms is selected.

Thirdly, select the best combination of indicators. Based on the selected best data dimensionality reduction method and six machine learning algorithms, a reputation evaluation index selection classifier for crowdsourcing participants is constructed. The sequential backward selection strategy (SBS) is adopted, and the classifier is used as the evaluation function to delete the indicators that have the least impact on the reputation discrimination ability of crowdsourcing participants one by one until the optimal number of indicators is reached, to determine the feature subset with the highest classification accuracy.

The confusion matrix is used to compare the accuracy, recall, and F1 value of the classifier. Using statistical significance test methods, including Friedman test, Kruskal-Wallis test, and dispersion degree, it compares the classification accuracy and stability of the classifier constructed by different data dimensionality reduction methods and machine learning algorithms, selects the classifier with the best performance, and puts forward the best index combination; the specific steps are shown in Figure 1.

4. Data Acquisition and Preprocessing

We select the Zhu Bajie platform to collect data, mainly based on three reasons: first, in China, the market share of the Zhu Bajie platform is the highest, reaching 50%, and the transaction is the most active, which has practical research significance. Secondly, the Zhu Bajie platform is one of the most widely studied crowdsourcing platforms and has been selected by many scholars as an important platform for data collection and research [4, 9, 26, 27]. Third, the Zhu Bajie platform provides the transaction data of crowdsourcing participants, the published list of illegal crowdsourcing participants, violation records, and reputation-related data such as work attitude, work speed, and completion quality, which is suitable for this study.

Based on the reputation evaluation index system of crowdsourcing participants established in [2], the evaluation indexes are collected. Collect 28 reputation evaluation index variables of crowdsourcing participants on the Zhu Bajie platform from December 1, 2006, to January 31, 2019, including the city of crowdsourcing participants, store type, margin deposit amount, years of opening, transaction amount in recent three months, transaction times of three months, number of employers served, transaction activity, number of refunds in this month, refund rate in this month, number of refunds in three months, refund rate in three months, comprehensive scoring, task completion quality, work speed, work attitude, employer repurchase rate, positive feedback rate, the number of positive feedback, the number of medium feedback, the number of negative feedback, number of employer recommendation, the number of employers did not choose, the number of employer nonrecommend, growth scoring, number of punishments, number of punishments in three months, and credibility frozen after reporting.

Through data collection, a total of 4357 samples of crowdsourcing participants were obtained. Each sample of data contains 28 index variables. After data preprocessing, 3298 valid samples are finally obtained. Through manual marking, the reputation level of crowdsourcing participants is marked as good, medium, and poor.

5. Selection of Data Dimensionality Reduction Methods

The massive data generated by the crowdsourcing platform provides information for the reputation evaluation of crowdsourcing participants, but the high dimension of variables will increase the operation cost. It is necessary to eliminate redundant indicators unrelated to the reputation evaluation and having no significant impact and determine the best data dimensionality reduction method. It selects four data dimensionality reduction methods and six machine learning algorithms for empirical research. Take the participants of the Zhu Bajie platform as the research object, collect the reputation evaluation index data, and propose the evaluation index selection method based on the experiment.

5.1. Existing Data Dimensionality Reduction Methods. Data dimensionality reduction methods mainly include feature selection and feature extraction. The goal of both is to reduce the dimension of index variables, but there are differences in the ways.

Feature extraction is to get new features by mixing the original features. For example, the Mean Impact Value (MIV) method reflects the influence of output variables on input variables through the change of the weight matrix in the neural network [33]. For example, the ReliefF algorithm makes a feature extraction method, which gives different weights to different index variables and removes the index variables with insignificant evaluation results. The ReliefF method has the characteristics of a simple algorithm and high operational efficiency.

The specific calculation process of the ReliefF algorithm is as follows: randomly extract the sample subset p from the dataset, then extract s nearest neighbor samples from p similar sample set and a different sample set, respectively, calculate the feature weight w, and update the correlation between features and categories in turn. Sort the features according to the weight, set the threshold, and select the feature subset; then, there are the following.

Input: p samples and corresponding characteristic attributes.

Output: feature weight vector w.

Initialize $w = 0$.

Weight value w calculation: the formula is

$$w[I] = w[I] - \frac{\sum_{j=1}^{s} \operatorname{diff}\left(I, P_i, H_j\right)}{\text{ms}} + \sum_{c \neq \text{class}\left(P_i\right)} \left[\frac{p(C)}{1 - p\left(\text{class}\left(P_i\right)\right)} \sum_{j=1}^{s} \operatorname{diff} \frac{\left(I, P_i, M_j(C)\right)}{\text{MS}}\right], \tag{1}$$

where m is the number of samples, $\sum_{j=1}^{s} \operatorname{diff}(I, p_i, H_j)$ calculate the distance between two sample examples about feature I, $M_j(C)$ is the jth nearest neighbor sample of different classes, and $\text{class}(P_i)$ is the category of the sample P_i. The ReliefF algorithm has the advantages of easy expansion, high computational efficiency, and strong stability. It can process noisy data and quickly process large datasets.

5.2. Common Machine Learning Algorithms

5.2.1. Support Vector Machine. Support vector machine (SVM) was first proposed by Cortes and Vapnik in 1995. SVM constructs the maximum interval classifier by dividing multiple features and draws lessons from the concepts of maximum interval hyperplane [34], the inner product of kernel function as feature space, use of kernel function, and sparsity. Vapnik Chervonenkis (VC) theory proves the existence of a risk VC boundary.

It is a method to classify linear and nonlinear data. For the linear separable problem, suppose the training set is $T = \{(x_1, y_1), \cdots, (x_n, y_n)\}$, among $x_i \in R^P, y_i \in \{-1, +1\}$, $i = 1, 2, \cdots, n$. Since the problem is linearly separable, there is

$$\left\{x\colon f\left(x_i\right) = w^T x + w_0 = 0\right\}. \tag{2}$$

The positive and negative sample points in the dataset are completely and correctly divided on both sides of the hyperplane. That is, for all sample points of $y_i = +1$, we have $f(x_i) > 0$, and for all samples of $y_i = -1$, we have $f(x_i) < 0$. The decision function can be constructed:

$$G(x) = \operatorname{sgn}(f(x)) = \operatorname{sgn}\left(w^T x + w_0\right). \tag{3}$$

For any hyperplane $L = \{x\colon f(x) = w^T x + w_0\}$, let $x_1$and x_2 be two points on the hyperplane; then, $w^T(x_1 - x_2) = 0$. So, $w^* = w/\|w\|$ is the unit vector of the hyperplane. Let x_0 be a point on the hyperplane; then, $w^T x + w_0 = 0$。 Therefore, the distance from any point x to the hyperplane is

$$\begin{aligned} S(L,x) &= \left|w^{*T}(x, x_0)\right| = \frac{1}{\|w\|}\left|w^T(x - x_0)\right| \\ &= \frac{1}{\|w\|}\left|w^T x + w_0\right| = \frac{f(x)}{\|w\|}. \end{aligned} \tag{4}$$

Divide all sample points in the hyperplane; that is, all the sample points of $y_i = +1$ have $f(x_i) > 0$; for all the samples with $y_i = -1$, we have $f(x_i) < 0$. Therefore,

$$S(L,x) = \frac{1}{\|w\|}\left(y_i f(x_i)\right) = \frac{1}{\|w\|} y_i\left(w^T x_i + w_0\right). \tag{5}$$

SVM has advantages in modeling complex nonlinear boundaries. It can control the model by maximizing the edge of the decision boundary. It has good stability, generalization characteristics, and a unique global optimal solution.

5.2.2. BP Neural Network. Backpropagation neural network (BPNN) is a multilayer feedforward neural network [35]. The input tuple of BPNN is weighted by the input layer and then given to the middle layer [36]. The output of the middle layer unit can be input to another middle layer. The number of middle layers is arbitrary, and usually, only one layer is used in practice. The weighted output of the last middle layer is used as the input of the unit constituting the output layer, and the output layer publishes the network prediction of a given tuple.

5.2.3. Radial Basis Function Neural Network. The radial basis function neural network (RBFNN) is a neural network structure proposed by Broomhead and Lowe in 1988. It is a three-layer feedforward network with a single hidden layer. The RBF neural network reduces the weight update link of error feedback, applies radial basis function as an excitation function in the hidden layer to fit the nonlinearity of the data set, and has the characteristics of simple training and fast learning convergence.

5.2.4. Decision Tree. The decision tree (DT) method came into being in the 1960s. It is a learning system built by Quinlan et al. when modeling human concepts. It is a method of decision-making based on the tree structure. The decision tree starts from the root node and tests a certain feature of instance x; according to the test results, the instance is assigned to the child node until it reaches the leaf node, and the class to which the leaf node belongs, that is, the label y of instance x, is predicted.

5.2.5. K-Nearest Neighbor Classifier. The K-nearest neighbor classifier (KNN) is a simple machine learning algorithm. This method assumes that adjacent points have the same attributes and "proximity" is generally measured by distance, such as Euclidean distance.

5.2.6. Naive Bayesian. Naive Bayesian (NB) is a statistical classification method. This method creates the probability distribution of features on each class label, which is characterized by combining a priori probability and a posteriori probability; that is, it avoids the subjective bias of using only a priori probability and the overfitting phenomenon of using sample information alone.

5.3. Preliminary Selection of Evaluation Indicators. This study uses empirical methods. Firstly, the filtering method is used to preliminarily select the 28 index variables in the dataset in the ZhuBajie platform and preliminarily eliminate the variables with irrelevant redundancy, weak discrimination ability, and insignificant impact on the reputation evaluation, to reduce the dimension of the variables.

Two feature selection methods, ReliefF and Mean Influence Value (MIV), and two feature extraction methods, principal component analysis (PCA) and linear discriminant analysis (LDA), were selected to compare the dimensionality reduction effect. Among the two-dimensionality reduction methods, ReliefF and PCA are selected to focus on the selection results of indicators in the first stage.

5.3.1. Selection Results of Phase I Indicators of ReliefF. The sample subset is selected randomly from the original data sample set, and then, the nearest neighbor samples are selected from a similar sample set of the sample subset. Each feature weight value is calculated and updated in turn. Repeat the above process many times to obtain the weight of features, arrange the features in descending order according to their feature weight value, and select part of the feature set by giving a threshold. That is, when the feature weight value is greater than the given threshold, the feature is used to form a new feature subset. If it is less than the given threshold, the feature is removed. Relevant calculations are carried out on MATLAB2016b, and the sorting results of feature weights are shown in Figure 2 and Table 1.

As shown in Table 1, the most significant indicator affecting the reputation evaluation of crowdsourcing participants is the praise rate, followed by the number of three-month penalties and the three-month refund rate, with weights of 0.054, 0.037, and 0.029, respectively. The evaluation index with negative index feature weight means that the information reflected by this feature increases the data noise and affects the accuracy of classification evaluation, which should be eliminated. Therefore, seven evaluation indexes, including store type, number of transactions in three months, employer served, three-month transaction volume, the employer not selected, total good evaluation, and employer recommendation, should be deleted. Eliminate redundant indicators with weak discrimination ability and little impact on reputation evaluation. The weight of the

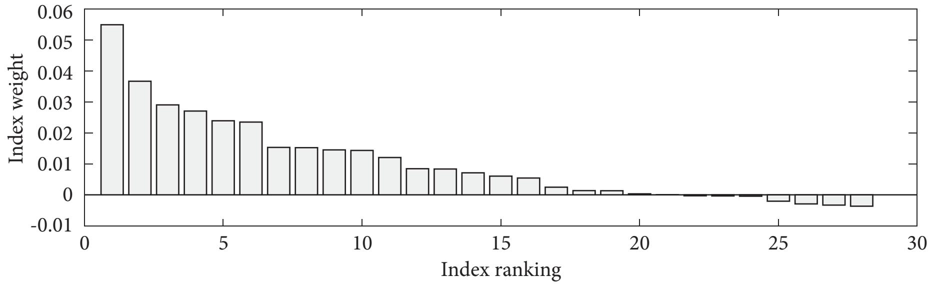


FIGURE 2: Index variable weight table after dimension reduction using ReliefF.

TABLE 1: Weight of index variables after dimension reduction using ReliefF.

Rank	Index variable	Weight
1	x_1 positive feedback rate	0.054
2	x_2 number of punishments in three months	0.037
3	x_3 refund rate in three months	0.029
4	x_4 refund rate in this month	0.027
5	x_5 credibility frozen after reporting	0.023
6	x_6 transaction activity	0.023
7	x_7 work attitude	0.015
8	x_8 task completion quality	0.015
9	x_9 work speed	0.014
10	x_{10} years of opening	0.014
11	x_{11} comprehensive scoring	0.012
12	x_{12} number of employer nonrecommend	0.008
13	x_{13} margin deposit amount	0.008
14	x_{14} number of negative feedback	0.007
15	x_{15} growth scoring	0.006
16	x_{16} number of medium feedback	0.0055
17	x_{17} number of punishments	0.037
18	x_{18} employer repurchase rate	0.001
19	x_{19} number of refunds in this month	0.001
20	x_{20} number of refunds in three months	0.0003
21	x_{21} city	0.00001
22	x_{22} number of employer recommendation	−0.002
23	x_{23} number of positive feedback	−0.0029
24	x_{24} number of employers did not choose	−0.003
25	x_{25} transaction amount in recent three months	−0.0003
26	x_{26} number of employers served	−0.0037
27	x_{27} transaction times of three months	−0.0004
28	x_{28} store type	−0.0004

city is 0.00001, which shall be eliminated. According to the first-stage index selection results of ReliefF, 20 of the 28 indexes, are retained.

5.3.2. Selection Results of Indicators in the First Stage of the PCA. The principal component analysis is a commonly used feature extraction method. It extracts new feature variables from the original feature variables, condenses the original many variables into a few variables, and retains the original information as much as possible while reducing the number of feature variables. The principal component analysis first determines whether the characteristic variables are suitable for PCA analysis by the Bartlett sphericity test and KMO test. Using SPSS25.0 software for the Bartlett sphericity test, the observed value of the statistics is 67924.04, and the probability P value is close to 0. The KMO value is 0.77, indicating that the original variable is suitable for principal component analysis. The principal component is extracted by the PCA method, and the total variance of the original variables explained by the principal component is shown in Table 1.

In Table 2, the first column is the principal components extracted by the PCA method, which are numbered according to the variance contribution rate. The eigenvalue of the first principal component is 5.484, which explains 19.587% variables. The eigenvalues of the first 10 principal components are greater than 1, which explains 74.738% variables, indicating that the first 10 factors retain most of the original information.

To sum up, the feature selection method ReliefF is used to rank the importance of the evaluation index system, then judge the data noise, and eliminate the redundant indexes with weak reputation discrimination ability and little impact on the crowdsourcing participants. According to the ranking results of the evaluation indicators, 8 indicators are removed from the original 28 indicator variables and 20 indicators are retained. Using the feature extraction method to calculate the total variance of principal components, the eigenvalues of the first 10 of the 28 indicators are greater than 1, indicating that the original feature information is less lost when 10 principal components are extracted.

5.4. Select the Best Data Dimensionality Reduction Method. After the preliminary screening of evaluation indicators through empirical methods, they use the combination method of data dimensionality reduction and machine learning algorithm and then empirically finds the best data dimensionality reduction method. Four data dimensionality reduction methods, including ReliefF, MIV, PCA, and LDA, are selected, together with DT, BPNN, RBFNN, SVM, KNN, and NB six machine learning algorithms cross construct 24 kinds of crowdsourcing participants' reputation evaluation index selection classifier, mine the combined effect of data dimensionality reduction method and machine learning algorithm, and select the best data dimensionality reduction method.

Taking the dataset of the Zhu Bajie platform as an example, the performance of 24 crowdsourcing participant evaluation indexes' selection classifiers constructed cross is verified. Through ten-fold cross-validation and the

Table 2: Total variance of original variables explained by principal components.

Principal component	Initial eigenvalue			Extract the sum of squares of loads		
	Total	Variance contribution rate (%)	Cumulative percentage (%)	Feature value	Variance contribution rate (%)	Cumulative percentage (%)
1	5.484	19.587	19.587	5.484	19.587	19.587
2	4.699	16.781	36.369	4.699	16.781	36.369
3	1.983	7.084	43.452	1.983	7.084	43.452
4	1.636	5.842	49.295	1.636	5.842	49.295
5	1.577	5.634	54.929	1.577	5.634	54.929
6	1.243	4.439	59.367	1.243	4.439	59.367
7	1.218	4.352	63.719	1.218	4.352	63.719
8	1.046	3.737	67.456	1.046	3.737	67.456
9	1.031	3.683	71.139	1.031	3.683	71.139
10	**1.008**	**3.600**	**74.738**	**1.008**	**3.600**	**74.738**
11	0.967	3.454	78.192			
12	0.946	3.377	81.569			
13	0.858	3.063	84.632			
14	0.738	2.635	87.268			
15	0.680	2.428	89.696			
16	0.528	1.887	91.583			
17	0.485	1.733	93.316			
18	0.357	1.276	94.592			
19	0.320	1.143	95.735			
20	0.303	1.083	96.818			

Note. List the principal components with weight ranking 1–20. The eigenvalues of the 10th principal components are greater than 1 in bold.

Friedman test, the performances of different classifiers are compared, and the best data dimensionality reduction method is selected.

The dataset is divided into the training set and test set, and the data proportions are 70% and 30%. We select the classification regression algorithm CART to classify the whole sample and set the input layer of BPNN as 20 input variables, 20 hidden layer neurons, and 3 output neurons; the learning rate value is 0.1, the training required accuracy value is 0.00001, and the maximum training time is 100. We use Newrbe function to create RBFNN and Gaussian function as kernel function, by observing the change of accuracy of KNN classifier, to set the nearest neighbor value to 8.

Based on the ranking results of the reputation evaluation index weights of crowdsourcing participants in the first stage, to avoid losing important evaluation indexes, ReliefF is used to filter the indexes with data noise and weak discrimination ability, select the first 20 evaluation indexes, and calculate the average ten-fold cross-verification accuracy of the 24 selection classifiers constructed, as shown in Table 3.

It is shown that the average classification accuracy of ReliefF-BPNN ten-fold cross-validation is the highest, which is 0.91, followed by LDA-SVM, which is 0.908, and PCA-DT, which is the lowest, which is 0.817. After using the ReliefF method for data dimensionality reduction, the evaluation index selection classifier constructed based on DT, BPNN, RBFNN, and NB algorithms has the highest accuracy. After using the LDA method for data dimensionality reduction, the evaluation index selection classifier constructed based on the SVM and KNN algorithms has the highest accuracy.

Table 3: Accuracy of ten-fold cross-validation of the reputation evaluation index selection classifier.

Model	DT	SVM	BPNN	RBFNN	NB	KNN	Rank
ReliefF	**0.899**	0.901	**0.910**	**0.856**	**0.878**	0.883	**3.67**
MIV	0.875	0.887	0.886	0.828	0.819	0.872	1.67
PCA	0.817	0.844	0.855	0.849	0.846	0.854	1.5
LDA	0.881	**0.908**	0.905	0.848	0.871	**0.893**	3.17

The best result for each classifier is in bold.

The effects of different dimensionality reduction methods are verified by the Fredman test. The average rank results of the four dimensionality reduction methods ReliefF, MIV, PCA, and LDA are shown in Table 3. The observed value of the Friedman test statistic is 12.2, and the asymptotic significance is 0.007, which indicates that the four feature dimensionality reduction methods have significant differences.

The average rank of the four dimensionality reduction methods ReliefF, MIV, PCA, and LDA are 3.67, 1.67, 1.5, and 3.17, respectively. The ReliefF method has the best dimensionality reduction effect, and the LDA method can also achieve a good classification effect. ReliefF evaluates the correlation and redundancy of features by calculating the correlation statistics of the corresponding features of adjacent samples between the same class and different classes, which is suitable for multiclassification problems. The LDA method removes irrelevant information in the data by projecting the dataset to a lower dimension to achieve the effect of dimension reduction.

In this case, the overall dimensionality reduction effect of the classifier is the best when using the ReliefF method. The classification performance of MIV and PCA is significantly lower than that of ReliefF and LDA, and PCA is the worst. When using the PCA dimensionality reduction method, the evaluation index selection classifier based on DT, SVM, BPNN, and KNN shows the worst performance.

In terms of training time, the DT classifier has the fastest average training speed, followed by the NB classifier. The average training time of the RBFNN classifier is 5.13 seconds, which is the longest among the six algorithms selected. Secondly, the average training time of the BPNN classifier is 3.72, which shows that the operation cost of this method is high. Among the four dimensionality reduction methods, the average training time of LDA and ReliefF is 1.85 seconds and 2.03 seconds, respectively, and the average training time of MIV is 4.47, which takes the longest time. The average training time of PAC is 1.26 seconds, which is the shortest.

After using the ReliefF method to reduce the dimension of data, the accuracy of classifier selection based on the evaluation indexes of DT, SVM, BPNN, RBFNN, and NB algorithms is concentrated in the range of 0.86–0.93 and fluctuates gently. The highest single classification accuracy of ReliefF-BPNN is 0.922. The classification accuracy of relief RBFNN fluctuates in the range of 0.83–0.88, and the performance of the classifier is significantly lower than that of the other five classifiers. The ten-fold cross-validation results of the evaluation index selection classifier of crowdsourcing participants are shown in Figure 3.

After data dimensionality reduction using the MIV method, the accuracy of classifier selection based on evaluation indexes of DT, SVM, BPNN, and KNN algorithms is concentrated in the range of 0.86–0.92, with gentle fluctuation and high classification stability. MIV-NB has the lowest single classification accuracy of 0.775 and the highest of 0.895, with a difference of 0.12. Among all the classifiers, the classification stability is the worst.

After using the LDA method to reduce the dimension of data, the accuracy of classifier selection based on SVM and BPNN algorithms is concentrated in the range of 0.89–0.92, and the accuracy of classifier selection based on the Nb algorithm fluctuates between 0.83 – 0.87.

After data dimensionality reduction using the PCA method, the accuracy of classifier selection based on the evaluation indexes of SVM, BPNN, RBFNN, and KNN algorithms is concentrated in the range of 0.83–0.88, and the overall classification accuracy is lower than that of other dimensionality reduction methods. The fluctuation range of PCA-NB accuracy is the largest, and the maximum amplitude is 0.065, indicating that the performance of the classifier is the most unstable. The accuracy of PCA-DT fluctuates between 0.8 – 0.85, and the classification accuracy is the lowest among all methods.

Through the analysis and comparison of four data dimensionality reduction methods and six machine learning algorithms, the performance of 24 crowdsourcing participants' reputation evaluation index selection classifiers was cross constructed; the research shows that the classification effect of using the ReliefF method to select evaluation indexes is the best, which is better than the LDA, PAC, and MIV methods. Among them, the dimensionality reduction effect of PCA and MIV methods is significantly lower than that of ReliefF and LDA dimensionality reduction methods. In terms of training time, the PCA method has the fastest training speed and the MIV dimensionality reduction method has the slowest training speed. The training speed of the LDA method is slightly higher than that of the ReliefF dimension reduction method. Compared with the LDA dimensionality reduction method, ReliefF adopts the feature selection method, which does not change the original index and is more explanatory. Therefore, we select the best data dimensionality reduction method, ReliefF.

6. Optimal Feature Subset Selection

Based on the empirical selection of the best data dimensionality reduction method, ReliefF, a classifier based on six machine learning algorithms, DT, BPNN, SVM, KNN, and NB, is constructed. The prediction performance of the classifier is used as the standard to evaluate the combination of indicators. The Sequential Backward Selection (SBS) strategy is used to adjust the selected evaluation indexes, and the prediction performance of the classifier is used as the standard to evaluate the combination of evaluation indexes. Through the empirical analysis of data, the selection method of the best feature subset is proposed to select the best feature subset.

6.1. Selection Method of the Best Evaluation Index Combination Based on ReliefF. The selection method of the best evaluation index combination based on ReliefF is shown in Figure 4.

Suppose there are n samples in dataset D, and m features are preliminarily selected after deleting redundant variables in the first stage. M_i represents the feature subset with i features in the dataset, where $1 < i \leq m$; A_i represents the accuracy of classification prediction by the classifier when the dataset has i features, and Amax represents the highest classification accuracy of the classifier. 70% of the samples in dataset D are divided into training sets to train classifiers; 30% are divided into test sets, which calculate the accuracy A_i of the test set when the number of features is i.

The sequential backward selection strategy (SBS) is adopted to delete a feature one by one from the set containing m features. Calculate the accuracy A_i of the six classifiers selected, compare the accuracy of classifiers with different feature numbers, and gradually remove redundant variables. A_{max} with the highest accuracy of each classifier is selected, and then, the best feature subset max of the classifier is determined. The steps are shown in Figure 4.

Based on the selected 20 evaluation indicators, we gradually delete an indicator that has little impact on the reputation discrimination ability of crowdsourcing participants until the number of indicators reaches the specified number. By adjusting the number of selected evaluation indicators, we evaluate the accuracy of classifiers in the different numbers of evaluation indicators, select the selection method and feature subset of evaluation indicators with the highest classification accuracy, and then determine the reputation evaluation index system.

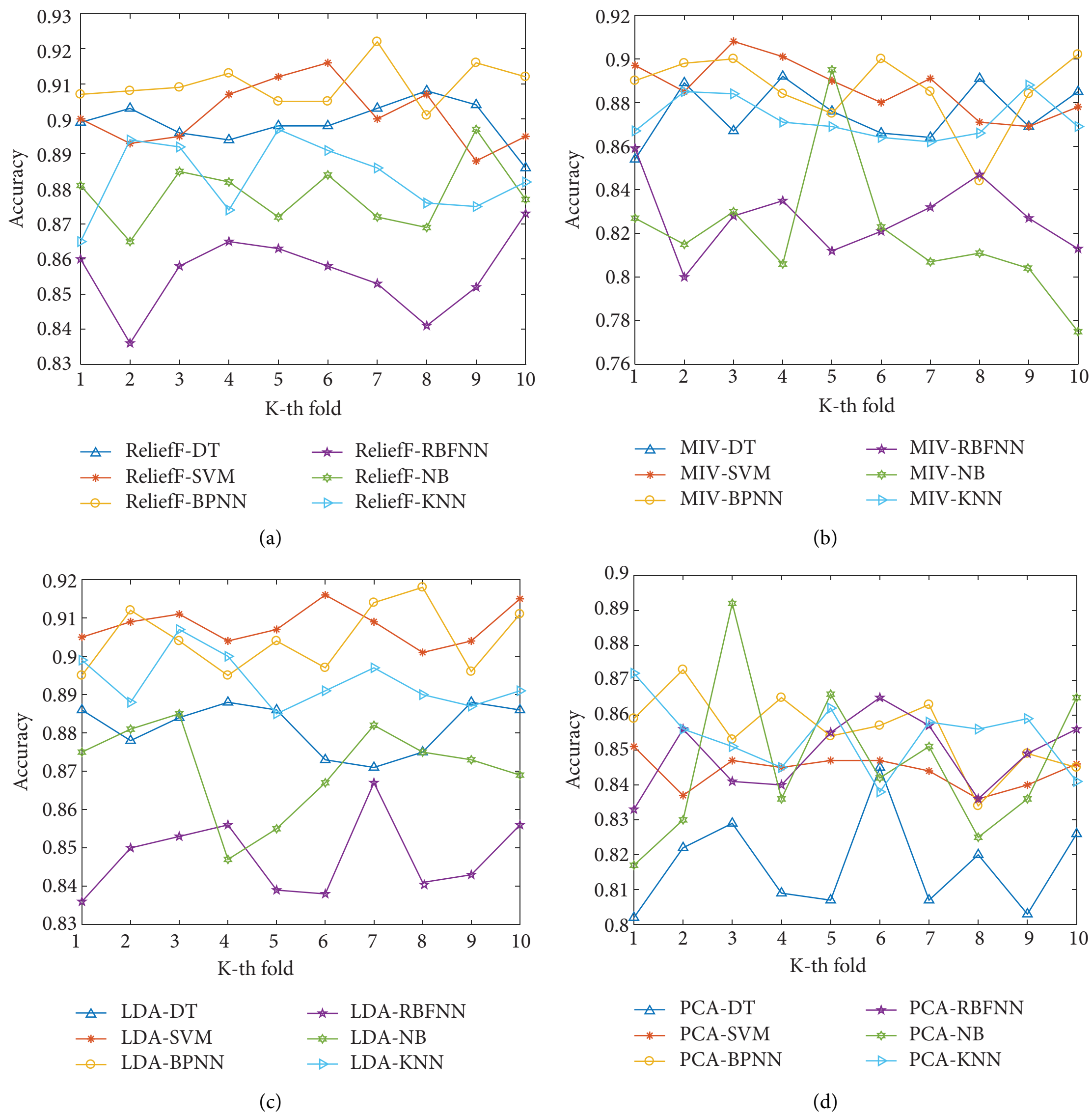


FIGURE 3: Ten-fold cross-validation diagram of the evaluation index selection classifier of crowdsourcing participants.

6.2. Select the Best Combination of Reputation Evaluation Indicators

6.2.1. Selection Results of the Best Evaluation Indicators. The accuracy of the classifier selected by the evaluation index changes with the number of selected features. The experimental results show that, after using ReliefF dimensionality reduction, the crowdsourcing participant evaluation index based on six different machine learning algorithms selects the classifier, and the classification accuracy is the highest when selecting the best feature subset. The optimal feature number of ReliefF-KNN is 8, the optimal feature number of ReliefF-DT, ReliefF-SVM, and ReliefF-RBFNN algorithms is 9, Relief-BPNN is 12, and relief NB is 13. When the number of features is different, the accuracy results of ten-fold cross-validation of the classifier selected by the evaluation index are shown in Table 4 and Figure 5.

As shown in Figure 5, the accuracy of the six classifiers when selecting the best feature subset is higher than that when selecting 20 index variables as the feature subset. When the classifier selects the best feature subset, the number of evaluation indexes decreases, but the accuracy of the classifier is improved, among which the accuracy of ReliefF-RBFNN is improved the most. The research shows that the classifier has the highest accuracy when the best feature subset is selected; adding new redundancy indicators or reducing the indicators of the best feature subset will lead to the decline of classification accuracy.

Combined with the selection results of evaluation indicators in the first and second stages, the weight ranking of the selected reputation evaluation indicators of crowdsourcing participants is shown in Table 5.

6.2.2. Analysis of Evaluation Index Selection Results. The accuracy of the six classifiers was tested by the Kruskal–Wallis test when the reputation evaluation index variable was set to 20 and when the best evaluation index

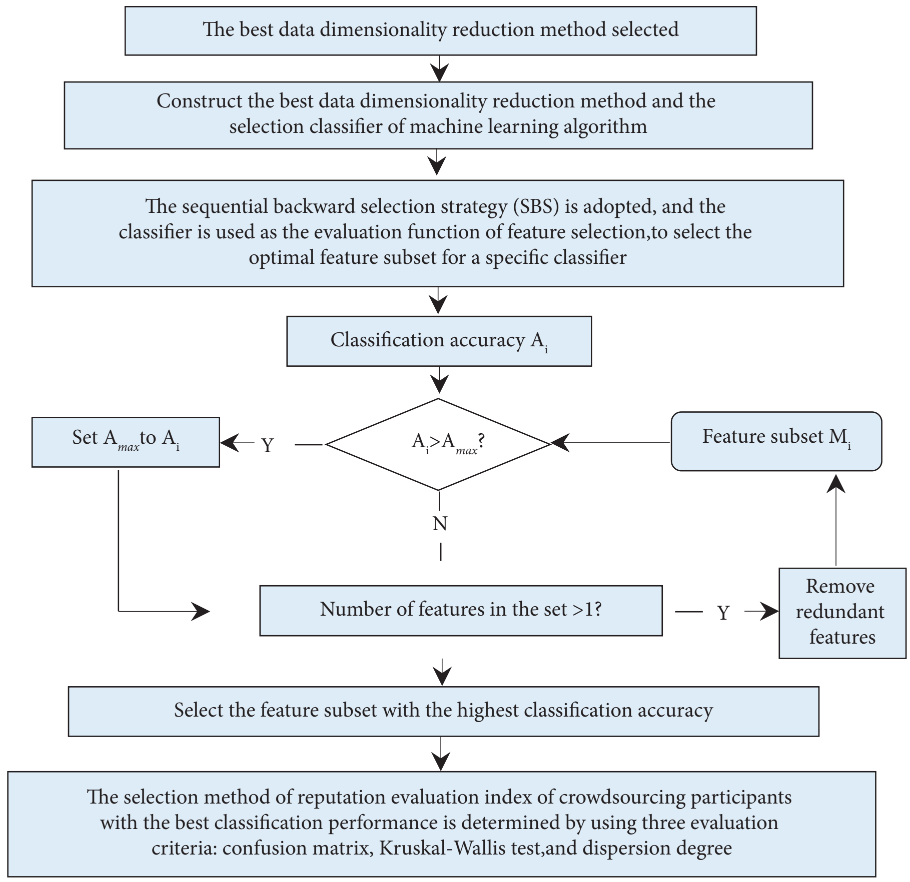


FIGURE 4: Selection method of the best evaluation index combination based on ReliefF.

variable was selected. The average rank of classifier accuracy when the best subset is represented by the blue column; the yellow column indicates the average rank of classifier accuracy when the number of selected evaluation indicators is 20. The Kruskal–Wallis test results of classifier accuracy selected by reputation evaluation indicators are shown in Figure 6.

The test results show that the K-W statistic is 92.661, and the probability P value is close to 0. After selecting the best feature subset, the performance of the six reputation evaluation index selection classifiers is improved to varying degrees, and the classification performance of ReliefF-RBFNN is the most significant. After the selection of evaluation indicators, the highest average rank of ReliefF-SVM is 105.6 and the second and third are ReliefF-BPNN and ReliefF-DT, respectively, and the average rank is 102 and 81.65, respectively. The classification performance of the ReliefF-SVM algorithm is the best.

6.3. Comparative Evaluation of Selection Methods

6.3.1. Confusion Matrix Analysis. The evaluation results of the confusion matrix of the classifier selected by the reputation evaluation index of crowdsourcing participants are shown in Table 6.

In terms of precision, the first type precision rate of ReliefF-SVM is 0.991, and the first-type precision rate of ReliefF-BPNN and ReliefF-RBFNN is also more than 0.98. The second type of ReliefF-NB has the highest precision rate of 0.534, indicating that the classifier has the strongest ability to distinguish crowdsourcing participants with a medium reputation. The third category of ReliefF-DT has the highest precision rate of 0.736 and ReliefF-NB has the lowest precision rate of 0.515, with a difference of 0.221, indicating that ReliefF-DT performs best for crowdsourcing participants with a poor reputation, while ReliefF-NB performs worst.

In terms of recall rate, it is particularly important, for it reflects the misclassification costs, especially, the recall rate of the third category. The first kind of recall rate of ReliefF-NB is the highest, which is 0.953, followed by ReliefF-DT, which is 0.95. The second kind of recall rate of ReliefF-SVM is the highest, which is 0.561, followed by ReliefF-BPNN, which is 0.488. The highest recall rate of the third category is ReliefF-SVM with 0.857, indicating that the selection method has the strongest ability to distinguish crowdsourcing participants with a poor reputation, and the misclassification cost of the third category is the lowest.

When evaluating classifiers, it is difficult to comprehensively evaluate the function of classifiers only by their precision rate or recall rate. Therefore, the F1 measure value is often used to comprehensively evaluate classifiers combined with precision rate and recall rate. The

Table 4: Accuracy of ten-fold cross-validation of feature subset selected by classifier.

Feature subset	ReliefF-DT	ReliefF-SVM	ReliefF-BPNN	ReliefF-RBFNN	ReliefF-NB	ReliefF-KNN
5	0.9036	0.9104	0.9050	0.8930	0.8592	0.8835
6	0.9043	0.9131	0.9087	0.8967	0.8643	0.8826
7	0.9026	0.9112	0.9085	0.8974	0.8665	0.8875
8	0.9031	0.9132	0.9088	0.8981	0.8788	**0.8958**
9	**0.9055**	**0.9147**	0.9086	**0.8987**	0.8843	0.8934
10	0.9012	0.9117	0.9099	0.8938	0.8874	0.8897
11	0.8993	0.9120	0.9126	0.8920	0.8869	0.8905
12	0.8998	0.9111	**0.9132**	0.8860	0.8856	0.8894
13	0.9034	0.9127	0.9108	0.8850	**0.8890**	0.8895
14	0.9032	0.9140	0.9109	0.8800	0.8818	0.8872
15	0.9012	0.9104	0.9100	0.8740	0.8832	0.8870
16	0.8997	0.9112	0.9130	0.8710	0.8791	0.8926
17	0.8987	0.9137	0.9100	0.8670	0.8760	0.8900
18	0.9000	0.9089	0.9110	0.8630	0.8781	0.8870
19	0.8977	0.9052	0.9130	0.8570	0.8782	0.8848
20	0.8989	0.9010	0.9100	0.8560	0.8781	0.8830

The best result for each classifier is in bold.

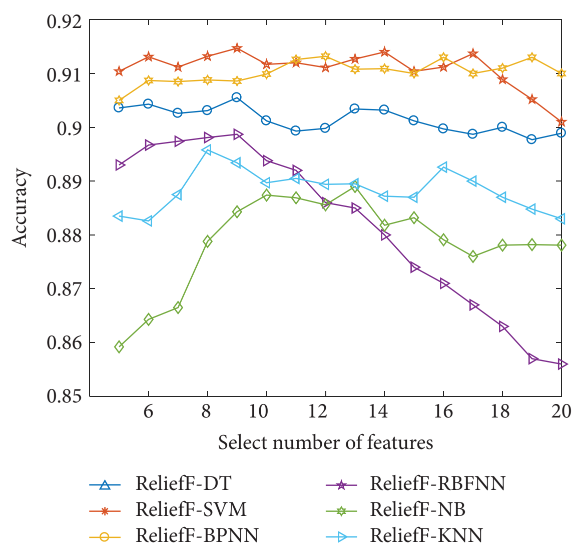


Figure 5: Ten-fold cross-validation of the feature subset selected by the classifier.

Table 5: Selection results of reputation evaluation indicators of crowdsourcing participants.

Rank	Index variable	Weight
1	x_1 positive feedback rate	0.054
2	x_2 number of punishments in three months	0.037
3	x_3 refund rate in three months	0.029
4	x_4 refund rate in this month	0.027
5	x_5 credibility frozen after reporting	0.023
6	x_6 transaction activity	0.023
7	x_7 work attitude	0.015
8	x_8 task completion quality	0.015
9	x_9 work speed	0.014
10	x_{10} years of opening	0.014
11	x_{11} comprehensive scoring	0.012
12	x_{12} number of employer nonrecommend	0.008
13	x_{13} margin deposit amount	0.008

maximum F1 measure of ReliefF-DT is 0.927, and the maximum F1 measure of ReliefF-SVM is 0.693 and 0.885, respectively. On the whole, ReliefF-SVM has good performance in accuracy and recall, high classification precision, and strong stability.

6.3.2. Dispersion degree. The robustness of the classifier is verified by the discrete degree of classification accuracy. The accuracy of the classifier is verified by ten-fold cross-validation, and the maximum value, minimum value, standard deviation, skewness coefficient, and kurtosis coefficient are calculated, as shown in Table 7.

In Table 7, the kurtosis coefficient of ReliefF-RBFNN accuracy is −0.014, indicating that the distribution form is relatively symmetrical and close to normal distribution. The kurtosis of the six evaluation index selection methods is negative, indicating that the data distribution is more

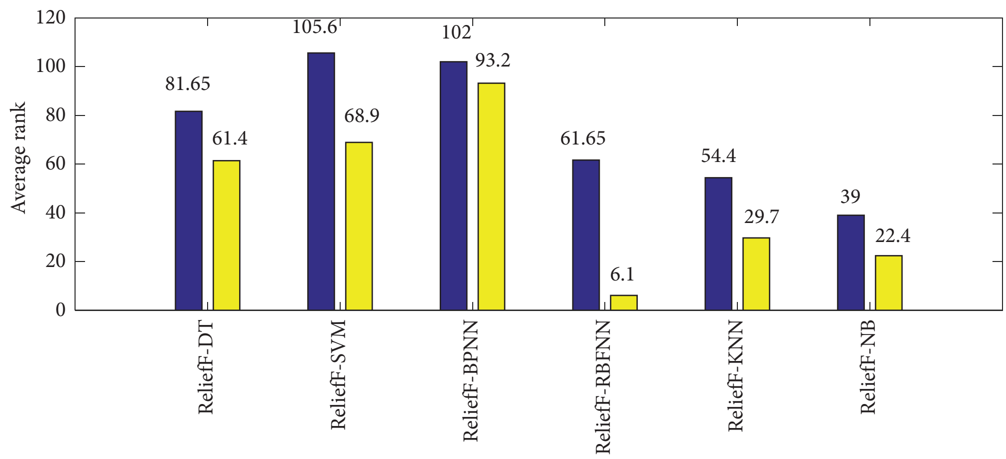


FIGURE 6: Kruskal–Wallis test results of accuracy of the classifier selected by the evaluation index.

TABLE 6: Evaluation results of crowdsourcing participants' reputation: evaluation index selection classifier confusion matrix.

Data dimensionality reduction method	Machine learning technology	Number of indicators	Accuracy	Precision rate			Recall rate			F1-measure		
				P_1	P_2	P_3	R_1	R_2	R_3	$F1_1$	$F1_2$	$F1_3$
ReliefF	DT	9	0.9055	0.964	0.431	**0.736**	0.950	0.477	0.799	**0.927**	0.624	0.849
	SVM	9	0.9147	**0.991**	0.291	0.675	0.934	**0.561**	**0.857**	0.924	**0.693**	**0.885**
	BPNN	12	0.9132	0.987	0.321	0.672	0.938	0.488	0.838	0.925	0.611	0.873
	RBFNN	9	0.8987	0.980	0.363	0.555	0.936	0.432	0.761	0.917	0.560	0.822
	KNN	8	0.8958	0.957	0.375	0.729	0.942	0.412	0.815	0.919	0.561	0.853
	NB	13	0.8890	0.951	**0.534**	0.515	**0.953**	0.416	0.756	0.920	0.567	0.814

TABLE 7: Accuracy and dispersion degree of the selection classifier of the reputation evaluation index of crowdsourcing participants.

Selection classifier	Minimum value	Maximum value	Mean value	Standard deviation	Skewness coefficient	Kurtosis coefficient
ReliefF-DT	0.898	0.915	0.906	0.006	0.406	−0.581
ReliefF-SVM	**0.906**	**0.922**	**0.915**	**0.005**	−0.485	−0.553
ReliefF-BPNN	0.905	**0.922**	0.913	0.006	0.576	**−0.282**
ReliefF-RBFNN	0.886	0.911	0.899	0.008	**−0.014**	−1.040
ReliefF-KNN	0.882	0.911	0.896	0.009	0.031	−0.335
ReliefF-NB	0.879	0.905	0.889	0.009	0.544	−0.474

The best result of the classifier is shown in bold under different measures.

gentle than the standard distribution. The minimum standard deviation of ReliefF-SVM is 0.005, and the maximum and minimum accuracy are the best levels of all methods. The discrete trend of accuracy from the center value is small, and the better the representation of mean value to data. It is shown that the ReliefF-SVM method has the best performance.

7. Conclusion

The credibility evaluation of crowdsourcing participants is a crucial issue for the rapid and healthy development of crowdsourcing. This study aims at the problems of a single reputation evaluation index, poor discrimination ability, and simple evaluation method of the participants of the crowdsourcing platform and studies the selection method of the reputation evaluation index of the participants of the crowdsourcing platform [37].

7.1. Main Contributions. This study has two main contributions.

First, a "dimension reduction-feature selection" method for selecting the optimal index combination is proposed. Firstly, the process selects the best data dimensionality reduction method from ReliefF, mean impact value (MIV), linear discriminant analysis (LDA), and principal component analysis (PCA). The sequential backward selection strategy (SBS) is adopted, and the accuracy of the classifier is used as the evaluation function of feature selection. The best feature subset is selected by evaluating the accuracy of the classifier with different feature numbers. An evaluation index selection method based on ReliefF-SVM is proposed, which has the best performance in accuracy, F1 measures, and stability.

Second, the best index combination of crowdsourcing participants' reputation evaluation is proposed. The feature subset with the best classification performance is selected using the proposed method (ReliefF-SVM) of selecting the best evaluation index. The research results show that the best combination of the chosen crowdsourcing participants' reputation evaluation indicators is: positive feedback rate, number of punishments, refund rate in three month, refund rate in this month, credibility frozen after reporting, transaction activity, work attitude, task completion quality, and work speed.

This study discusses the selection of reputation evaluation indicators of crowdsourcing participants for the first time. The proposed crowdsourcing participant reputation evaluation index combination makes up for the problem that the crowdsourcing platform evaluation index is single and cannot feedback the reputation status of the crowdsourcing participants. It solves the problem that a single evaluation index has a significant impact on the reputation status in the evaluation process, but the built-in index combination has a weak ability to distinguish the reputation status.

7.2. Main Conclusions

7.2.1. When Using the ReliefF Method to Reduce the Dimension of Data, the Classification Performance of the Six Machine Learning Algorithms Is the Best. Four dimensionality reduction methods of ReliefF, ReliefF, mean impact value (MIV), linear discriminant analysis (LDA), and principal component analysis (PCA), are selected, and six machine learning algorithms of decision tree (DT), BP Neural Network (BPNN), RBF neural network (RBFNN), support vector machine (SVM), K-nearest neighbor classifier (KNN), and Naive Bayes (NB) are used to cross construct 24 kinds of reputation evaluation index selection classifiers for crowdsourcing participants. Through the tenfold cross-validation and Friedman test, the average accuracy of classifiers when different dimensionality reduction methods are used is compared. The research shows that the accuracy of selecting classifiers based on reputation evaluation indicators constructed by other dimensionality reduction methods is significantly different. When using the ReliefF method to reduce the dimension of data, the average rank of classifier accuracy is the highest, which is better than the LDA, PAC, and MIV and has the best effect on dimension reduction. The feature selection method adopted by ReliefF does not change the original index and is more explanatory than LDA and PAC feature extraction methods.

7.2.2. The Reputation Evaluation Index Selection Method (ReliefF-SVM) of Crowdsourcing Participants Based on ReliefF Feature Selection Can Select Evaluation Indexes That Comprehensively, Objectively, and Effectively Identify the Reputation Status of Crowdsourcing Participants. The best number of features of the ReliefF-KNN classifier is 8, the best number of features of ReliefF-DT, ReliefF-SVM, and ReliefF-RBFNN classifier is 9, and the best number of features of ReliefF-BPNN and ReliefF-NB are 12 and 13, respectively. The experimental results show that the classifier selected by the six crowdsourcing participants' reputation evaluation indicators based on ReliefF feature selection has the highest classification accuracy when choosing the best feature subset. Adding new redundancy indicators or reducing the indicators of the best feature subset will lead to the decline of the accuracy of the classifier.

The selected ReliefF feature selection method and six machine learning algorithms are used to build a reputation evaluation index selection classifier for crowdsourcing participants. The results are analyzed and compared by the Kruskal–Wallis test, confusion matrix, and dispersion degree. The results show that ReliefF-SVM has the highest accuracy of 0.906, the highest F1 measure values of the second and third categories are 0.693 and 0.885, respectively, and the minimum standard deviation is 0.005. The experimental results show that the ReliefF-SVM classifier is excellent in accuracy; F1 measures value and stability and has stronger robustness and promotion value.

7.3. Future Research Directions. Future research directions include the following: how to further explore the classification problem of the new sample reputation of crowdsourcing participants by improving the usability of classifiers, reducing resource consumption, and improving prediction ability? In the machine learning algorithm, how to constantly update and adapt the new data, reduce the amount of computation required to train the classifier repeatedly to learn new and old data, and improve the performance and accuracy of the classifier.

Based on the best evaluation index combination and evaluation method research, how to combine the reputation of crowdsourcing participants with the recommendation of crowdsourcing tasks is studied. According to the reputation of crowdsourcing participants, how to achieve task recommendation under multidimensional constraints, how to recommend task selection sequences for crowdsourcing participants through matching algorithms, and how to improve the transaction success rate need further discussion.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China, under Grant 72101235, the Zhejiang Federation of Social Sciences, under Grant 2023B040, the State Scholarship Fund, under Grant 202108330330, Scientific Research Foundation of Zhejiang University of Water Resources and Electric Power, under Grant xky2022051, the Scientific Research Projects of Zhejiang Education

Department, under Grant Y202045307, the Soft Science Project of Zhejiang Provincial Science and Technology Department under Grant 2022C35057, and the General Research Project of Humanities and Social Sciences of the Ministry of Education, under Grant 21YJC790062.

References

[1] Y. R. Huang and M. Chen, "Multidimensional reputation evaluation model for crowdsourcing participants based on big data," in *Proceedings of the IEEE Int. Conf. on High Performance Big Data and Intelligent Systems*, pp. 41–46, Shenzhen, China, May 2019.

[2] Y. R. Huang and M. Chen, "Improve reputation evaluation of crowdsourcing participants using multidimensional index and machine learning techniques," *IEEE Access*, vol. 7, no. 1, pp. 118055–118067, 2019.

[3] J. Howe, "The rise of crowdsourcing," *Wired Magazine*, vol. 14, no. 6, pp. 1–5, 2006.

[4] Y. R. Huang, R. Wang, B. Huang, B. Wei, S. L. Zheng, and M. Chen, "Sentiment classification of crowdsourcing participants' reviews text based on LDA topic model," *IEEE Access*, vol. 9, pp. 108131–108143, 2021.

[5] L. N. Hao, "Research on enterprise marketing innovation strategy based on crowdsourcing mode information sharing under the sharing economy," *Journal of Hebei University of Economics and Trade*, vol. 18, no. 04, pp. 54–60, 2018.

[6] J. F. Li and Y. Y. Zhou, "Information disclosure and market concentration of it crowdsourcing service providers," *Journal of Systems Management*, vol. 30, no. 1, pp. 88–93, 2021.

[7] X. Han, M. X. Xue, and W. H. Song, "Role of reputation in sustainable performance of online crowdsourcing vendors: an explanation from transaction cost theory," *Frontiers in Psychology*, vol. 12, pp. 756134–756214, Nov. 2021.

[8] W. J. Jiang, Y. S. Xu, and H. Guo, "Dynamic trust calculation model and reputation management mechanism of online transactions," *Chinese Science: Information Science*, vol. 44, no. 9, pp. 1084–1101, 2014.

[9] J. Yan, S. P. Ku, and C. Yu, "Reputation model of crowdsourcing workers based on activity degree," *Computer Applications*, vol. 37, no. 7, pp. 2039–2043, 2017.

[10] J. F. Liu and X. Y. Jiang, "Construction of evaluation system for it service providers in crowdsourcing Market," *Journal of system management*, vol. 28, no. 5, pp. 857–863, 2019.

[11] X. F. Zhang and J. F. Su, "A combined fuzzy DEMATEL and TOPSIS approach for estimating participants in knowledge-intensive crowdsourcing," *Computers & Industrial Engineering*, vol. 137, p. 106085, 2019.

[12] M. Wang, G. L. Wang, Y. Zhang, and Z. Li, "A high-reliability multifaceted reputation evaluation mechanism for online services," *IEEE Transactions on Services Computing*, vol. 12, no. 6, pp. 836–850, 2019.

[13] C. Q. Jiang, Z. Wang, R. Y. Wang, and Y. Ding, "Loan default prediction by combining soft information extracted from descriptive text in online peer-to-peer lending," *Annals of Operations Research*, vol. 266, no. 1-2, pp. 511–529, 2018.

[14] Y. J. Zhang, G. T. Chi, and Z. Zhang, "Decision tree for credit scoring and discovery of significant features: an empirical analysis based on Chinese microfinance for farmers," *Filomat*, vol. 32, no. 5, pp. 1513–1521, 2018.

[15] T. Y. Li, H. J. Wang, J. Wu, Z. L. Zhang, and C. J. Tang, "Sparse Bayesian learning for credit risk evaluation," *Journal of Computer Applications*, vol. 33, no. 11, pp. 3094–3096, 2013.

[16] B. Wei, X. Wang, X. Xia, M. F. Jiang, Z. H. Ding, and Y. R. Huang, "Novel self-adjusted particle swarm optimization algorithm for feature selection," *Computing*, vol. 103, pp. 1569–1597, 2021.

[17] H. Zhang, Y. X. Shi, X. R. Yang, and R. L. Zhou, "A firefly algorithm modified support vector machine for the credit risk assessment of supply chain finance," *Research in International Business and Finance*, vol. 58, pp. 101482–101521, 2021.

[18] Z. C. Zhao, G. T. Chi, and X. P. Bai, "Mining farmers' key default characteristics based on the least significant difference method," *System Engineering Theory and Practice*, vol. 40, pp. 2339–2351, 2020.

[19] T. Zhang and G. T. Chi, "Default discrimination model based on optimal combination of credit characteristics -- a case study of Chinese A-share listed companies," *System engineering theory and Practice*, vol. 40, no. 10, pp. 2546–2562, 2020.

[20] H. L. Zhan, W. Yan, and B. Song, "Online game of cross-border e-commerce platform reputation based on incomplete information," *Practice and Understanding Of Mathematics*, vol. 51, pp. 57–69, 2021.

[21] J. Quan, C. X. Tang, Y. W. Zhou, X. J. Wang, and J. B. Yang, "Reputation evaluation with tolerance and reputation-dependent imitation on cooperation in spatial public goods game," *Chaos, Solitons & Fractals*, vol. 131, pp. 1–9, 2020.

[22] Y. J. Wang, Z. P. Cai, X. R. Tong, Q. X. Pan, Y. Gao, and G. S. Yin, "Online incentive mechanism of reputation based mobile crowdsourcing system," *Computer application*, vol. 36, no. 8, pp. 2121–2127, 2016.

[23] J. F. Lu, C. B. Tang, X. Li, and Q. Wu, "Designing socially-optimal rating protocols for crowdsourcing contest dilemma," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 6, pp. 1330–1344, 2017.

[24] X. Y. Zhu, Y. Li, L. Fang, and P. Chen, "An improved proof-of-trust consensus algorithm for credible crowdsourcing blockchain services," *IEEE Access*, vol. 8, pp. 102177–102187, 2020.

[25] S. Bhattacharjee, N. Ghosh, V. K. Shah, and S. K. Das, "QnQ: a reputation model to secure mobile crowdsourcing applications from incentive losses," in *Proceedings of the Communications and Network Security*, pp. 1–9, Las Vegas, NV, USA, October 2017.

[26] X. Y. Lu, Q. Lu, M. M. Huang, and Z. Q. Li, "Construction of contractor reputation evaluation model under crowdsourcing mode based on emotional tendency," *Statistics & Decisions*, vol. 34, no. 17, pp. 177–180, 2018.

[27] L. Sun, Q. Yang, X. Chen, and Z. X. Chen, "RC-chain: reputation-based crowdsourcing blockchain for vehicular networks," *Journal of Network and Computer Applications*, vol. 176, pp. 102956–103017, 2021.

[28] A. Alqwadri, M. Azzeh, and F. Almasalha, "Application of machine learning for online reputation systems," *International Journal of Automation and Computing*, vol. 18, no. 3, pp. 492–502, 2021.

[29] A. Rantanen, J. Salminen, F. Ginter, and B. J. Jansen, "Classifying online corporate reputation with machine learning: a study in the banking domain," *Internet Research*, vol. 30, pp. 45–66, 2020.

[30] T. D. Yang, X. D. Fu, L. Liu, K. Yue, L. J. Liu, and Y. Feng, "Online service reputation measurement based on multi-dimensional evaluation information," *Minicomputer System*, vol. 39, no. 12, pp. 2625–2631, 2018.

[31] Y. Yu, S. M. Liu, L. Guo et al., "CrowdR-FBC: a distributed fog-blockchains for mobile crowdsourcing reputation

management," *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 8722–8735, 2020.
[32] M. Kealeboga, "Volunteer reputation determination in crowdsourcing projects using latent class analysis," *Transactions in GIS*, vol. 25, no. 2, pp. 1–17, 2020.
[33] G. W. Dombi, P. Nandi, J. M. Saxe, A. M. Ledgerwood, and C. E. Lucas, "Prediction of rib fracture injury outcome by an artificial neural network," *The Journal of Trauma, Injury, Infection, and Critical Care*, vol. 39, no. 5, pp. 915–921, 1995.
[34] T. Cover, J. Thomas, and J. Wiley, "Elements of information theory," *Tsinghua University Press*, vol. 12, pp. 73–75, 2003.
[35] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.
[36] D. West, "Neural network credit scoring models," *Computers & Operations Research*, vol. 27, no. 11-12, pp. 1131–1152, 2000.
[37] Z. Zheng, X. S. Xu, W. Zhao, D. Chen, and L. Yang, "Computational experiment paradigm of choice behavior research -- Based on complex scientific management methodology," *Technology Economics*, vol. 37, no. 1, pp. 84–92, 2018.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 1425403, 12 pages
https://doi.org/10.1155/2022/1425403

Research Article

Dynamical Systems of Differential Equations Based on Information Technology: Effects of Integral Step Size on Bifurcation and Chaos Control of Discrete Hindmarsh–Rose Models

Xueli Chen

Mathematics Teaching and Research Office of Public Basic Education Department, Henan Vocational University of Science and Technology, Zhoukou 466000, Henan, China

Correspondence should be addressed to Xueli Chen; chenxueli_2010@163.com

Received 15 August 2022; Revised 6 September 2022; Accepted 16 September 2022; Published 12 October 2022

Academic Editor: Jiafu Su

Based on the differential equation dynamical system of information technology, this article analyzes how the integral step affects the bifurcation and chaos control of the discrete Hindmarsh–Rose model and its effect. By introducing the advantages of information technology in information management and information processing to the application of the differential equation dynamical system, the stability of the differential equation dynamical system model can be guaranteed. The integration step size is an important factor that affects the accuracy of the study results, and therefore, this paper understands how it affects the 3D discrete Hindmarsh–Rose model by choosing the appropriate step size.

1. Introduction

As a new technology, information technology (IT) has unique advantages in information management and information processing. It is a general description of various technologies, including computers, communications, sensors, and intelligence. The application and application field of information technology are very broad. It can design and develop software, and install information system to identify, transmit, and process information and data through computer computing and intelligent communication technology. It has broad market development prospects. As an important kind of equation in modern mathematics, differential equations can provide reference for researchers in solving practical problems in various fields. Dynamical systems of differential equations have abundant application information and broad research prospects in geography, mathematics, astronomy, and pedagogy. Therefore, the study of the differential equation dynamical system not only has a strong theoretical value, but also has strong practical significance. When information technology is combined with the research prospect of the differential equation dynamical system, it is very beneficial to the accuracy of the model construction of the differential equation dynamical system [1].

In 1952, Alan Lloyd and Andrew Huxley proposed a 3d model to study the discrete processes associated with the potential for conflict in the giant octopus axis. Many scholars have introduced various types of theories and methods to improve this model in an attempt to reduce its complexity and to obtain a simplified model without changing the model dynamics. The 3D time dispersion is carried out step by step according to the swamp integral method, and the local and global branches of the Roth model have serious consequences. By changing the integration step, the system exhibits complex dynamic behavior [2].

In this article, the influence of integral step size on bifurcation and chaos control of the discrete Hindmarsh–Rose model is studied under the premise of the differential equation dynamical system based on information technology, and the theoretical and practical experience obtained is expected to provide help for subsequent workers to understand bifurcation and chaos control of the discrete Hindmarsh–Rose model. Understand how the change in the integral step has an impact on the bifurcation and chaos control of the 3D discrete Hindmarsh–Rose model.

2. Research Background

In the field of academic specialization, information technology is defined as the visual operation of tools capable of retrieving, transferring, processing, storing, and applying information. From another point of view, information technology includes all technologies, methods, and systems related to the production, processing, storage, transmission, communication, and application of information, as well as related equipment and tools. In terms of nature, information technology can be divided into hard and soft [3]. Hard information technology mainly refers to the labor technology, labor tools, and the knowledge level of workers related to information, while soft information technology mainly refers to the management methods, management systems, and strategies to deal with problems related to information. The nature division of information technology is shown in Figure 1.

The information technology industry has developed rapidly. The statistics of the information technology industry in recent years are shown in Figure 2.

It can be observed from Figure 2 that the market value of the information technology industry developed from 456.5 billion yuan in 2015 to 718.1 billion yuan in 2021, with an increase of 261.6 billion yuan.

Per capita generated business income of China's software and information technology service industry from 2014 to 2021 is shown in Figure 3.

As can be observed from Figure 3, the per capita generated business income of China's software and information technology service industry was 1.158 million Yuan in 2021, an increase of 8.6% compared with 2014.

The growth rate of the electronic information technology industry in the past six years is shown in Figure 4.

It can be seen from Figure 4 that the growth rate of the electronic information technology industry is all positive. Although the growth rate has declined, it still shows a trend of continuous growth, and there is still a certain space for development in the future.

The realization carrier of the information technology industry is computer. The equipment sales of microcomputers in the last year are shown in Figure 5.

As can be seen from Figure 5, the sales of microcomputer equipment in China have shown rapid growth in the recent year, which also indicates that the development of the information technology industry is promising and there is a huge market potential for its development.

Based on the role generated by information technology in terms of human activities, its role can be summarized into two major aspects: mechanization, and information extraction and processing. Mechanization can be seen as the increase in the level of mechanization brought about by the recent changes in the level of production, while information extraction and processing has brought about changes in human society in terms of communication.

Differential equations are an important equation in modern mathematics that provide researchers with guidance in solving practical problems in various fields. There are two main types of differential equations: partial differential equations and ordinary differential equations. Partial differential equations have made a major breakthrough and have very broad application in the probability theory, and ordinary differential equations have a comparative advantage in higher mathematics for algebraic problems [4]. Dynamical systems, as a special discipline to study the order of evolution of a certain system with time, can be introduced to the application of differential equations and can be studied as an important subhead in the field of modern dynamics. Mapping iterations and discrete motions with differential as the core are an integral part of modern dynamics research. Differential equation dynamical systems contain rich information of application and broad research prospects in the fields and disciplines of geography, mathematics, astronomy, and education. Therefore, the study of differential equation dynamical systems has not only a strong theoretical value but also a strong practical significance [5].

In recent years, in the theoretical study of differential equation dynamical systems, researchers have exhausted their efforts and created many important research results in this field. For nonlinear differential equations, the German mathematician Black and the American mathematician Martos derived a model for the amplitude equation of nonlinear differential equations based on the characteristics of the solutions of quadratic nonlinear differential equations and in conjunction with Veda's theorem. Subsequently, they used the integral step principle and McLaughlin's formula to obtain a model for the amplitude equation of the cubic nonlinear differential equation, and applied the model to the Euler equation through the sensing function of information technology to study the bifurcation and chaos control effects in the model [6].

3. Materials and Methods

3.1. Differential Equation Dynamical System

3.1.1. Concept of the Differential Equation Dynamical System. The differential equation dynamical system has attracted the attention of scholars both at home and abroad and is the approximate solution of the differential equation. The parameter deviation effect caused by the linear differential constant term given by the unknown parameters affects the accuracy of the differential equation dynamical system to the model construction, leading to the approximation of the divergence and convergence effect of the dynamical system [7]. In this article, two types of linear differential equations based on dynamical systems under the convergence effect of

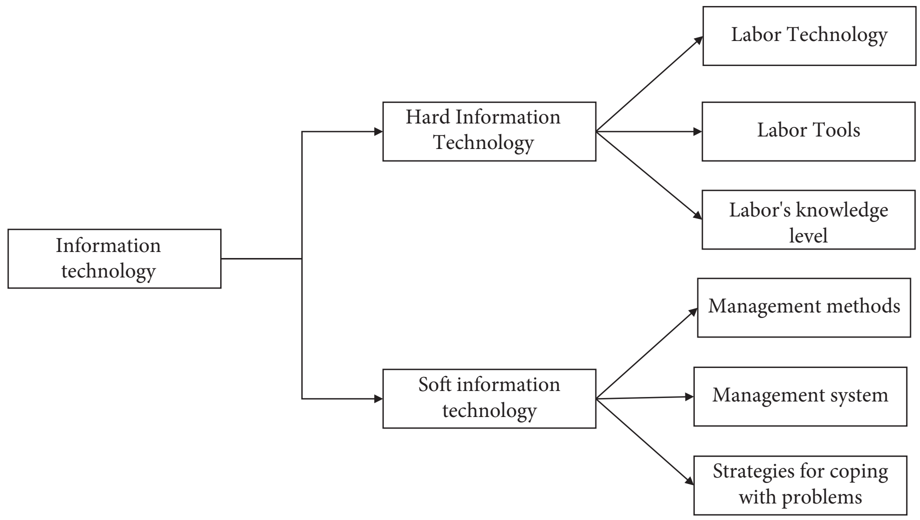


FIGURE 1: The nature of information technology.

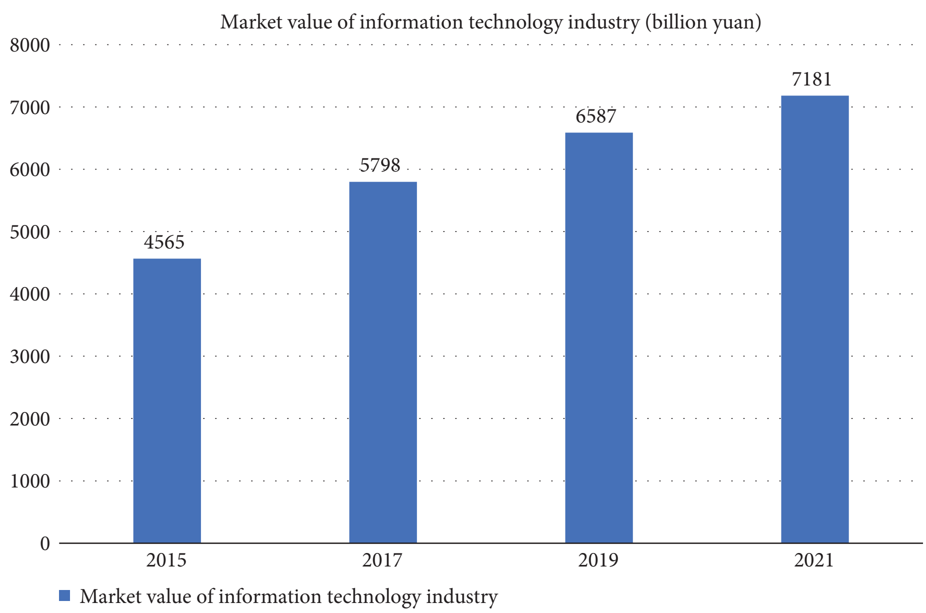


FIGURE 2: Market capitalization of the information technology industry.

equations are derived on the basis of Laplace's theorem. The first class is the nonlinear differential equations with the parameter terms reorganized by the convergence effect, as shown in the following equation:

$$du = \left[\xi u + \varepsilon^2 \zeta u + B(u, u)\right]dt + \varepsilon^{\delta/3} dW, \quad (1)$$

where the stochastic process = (): [0,] × [0,] ⟶ R, A is a self-accompanied nonpositive operator, 2L is a small perturbation term, a bilinear operator, and a finite-dimensional Gaussian process. The parameters are sufficiently small positive numbers, that is, 0 < ≪ 1.

Another class is the linear differential equations with parameter terms undergoing divergence effects, as shown in the following equation:

$$du = \left[\frac{1}{\varepsilon^2}(\triangle + I)u + B(u, u)\right]dt + dW, \quad (2)$$

where Δ is the Laplace operator, denoting the constant operator, such that the self-concurrent nonpositive operator A in Assumption 1.3.1: = 12(Δ+). The same, positive numbers with sufficiently small parameters, are bilinear operators and are degenerate Gaussian processes.

When 12's converge to zero, equations (1) and (2) satisfy the conditions for the operation of an effective dynamical system in the limit. And the de-parameterized ordinary differential equations can also satisfy the conditions for the dynamical system operation. This shows that both linear differential equations and nonlinear differential equations can satisfy the operation of the dynamical system after being

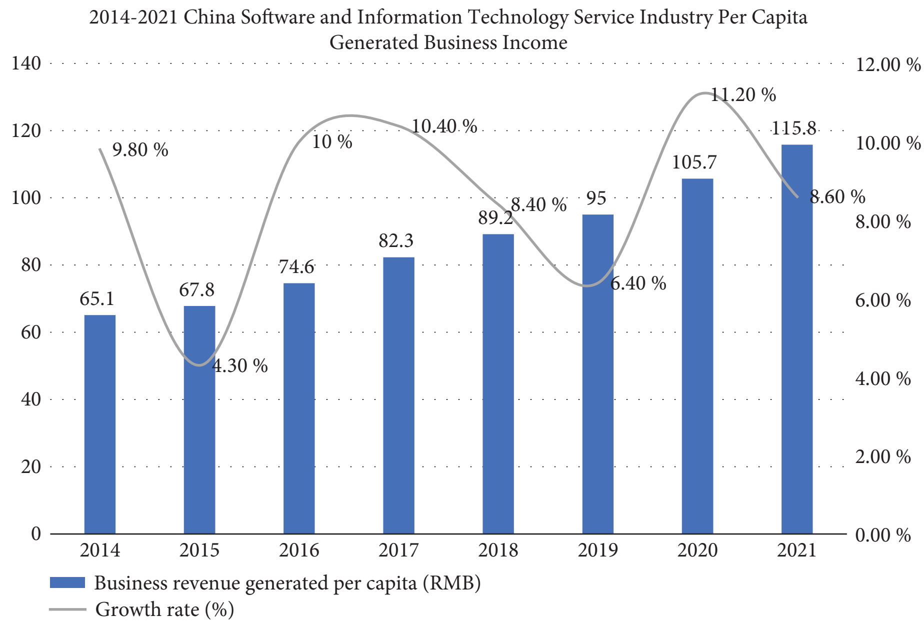


Figure 3: Per capita business revenue generated by China's software and information technology service industry, 2014–2021.

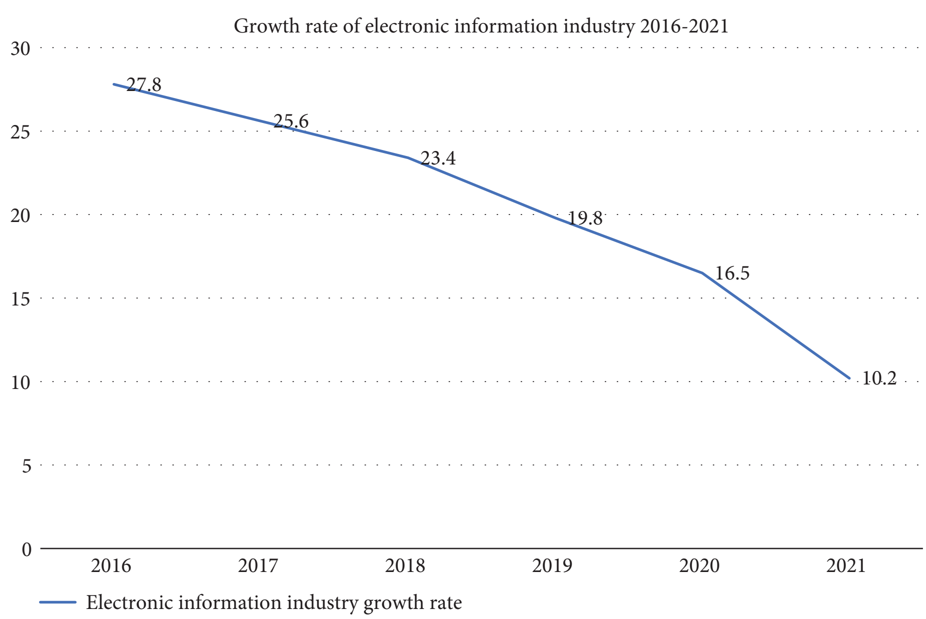


Figure 4: Growth rate of the electronic information technology industry.

reorganized by the denoise treatment and given the convergence conditions and approximation forms.

3.1.2. Denoising of the Differential Equation Dynamical System. Since the differential equation dynamical system after the denoising process can cause formal deviations after experiencing the effects of additive noise driving, it may have an impact on the final research results [8]. In this article, we decompose the initial stochastic part into a finite space generated by operator A and a non-nuclear space by a proper division of the time variables. The resulting simple equations describing the nuclear evolution process, that is, the efficient dynamical system, are derived. In addition, the time stop method is usually applied to differential equations in the infinity region. Many examples from physics show that this technique overcomes the difficulty of defining differential equations for infinite regions in the absence of the central manifold theory. Also, this time-stopping method is applicable to finite-area differential equations with arbitrary deviations in the absence of the central form theory [9].

The different degree of denoising also affects the effectiveness of the dynamical system to different degrees. If the

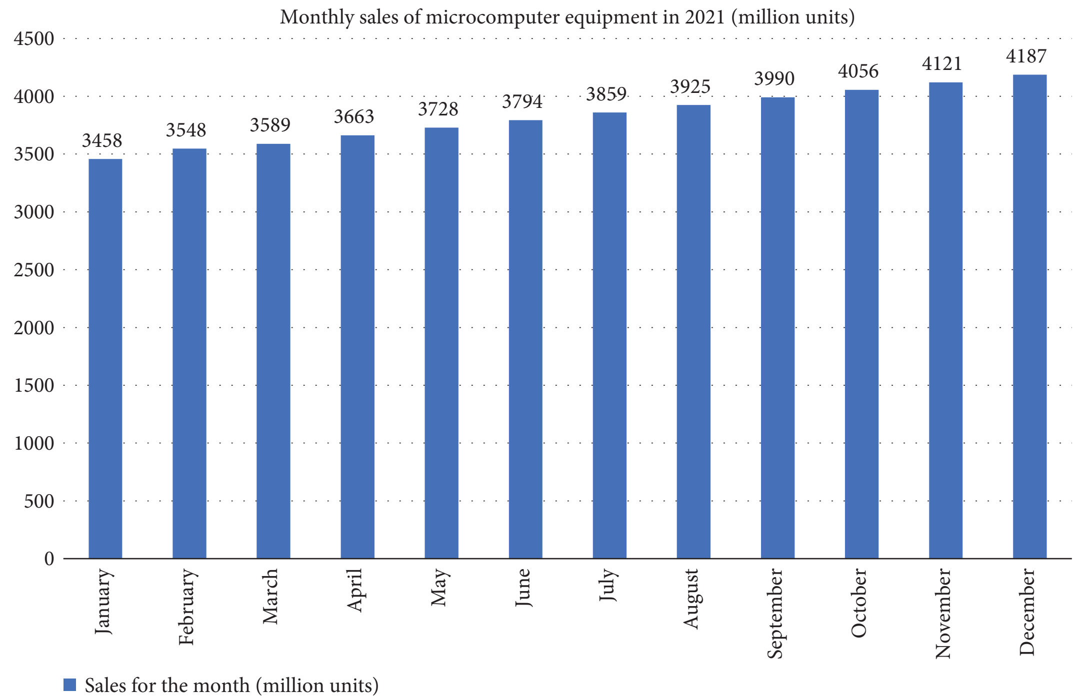


Figure 5: Equipment sales of microcomputers in the recent year.

spatial and temporal scales are in the same stratum as the degree of denoising, then the convergence effect of the nonlinear differential equations after the denoising process is received in the model, producing constant term differential equations containing more parametric terms. In addition, the chi-square linear differential equations, after the deconvolution process, cause deviations from the parameter terms in the equations, interfere with the interactions between the linear equations, affect the element replacement in the non-nuclear space, affect the way of reaction diffusion in the differential equation dynamical system, and have a degenerative effect on the dynamical system. For the first type of parameter term after the convergence effect reorganization of the nonlinear differential equations, out of getting the constant terms in the nonkernel space generated in the dynamical system under the play of nonlinear interaction, the effectiveness of the dynamical system can be evaluated by using equation (1) to arrange the elements in the nonkernel space according to the random principle. For another class of linear differential equations with parameter terms undergoing divergence effects, the dynamical system containing only the elements inside the nuclear space is obtained by using the constant terms in the non-nuclear space during the element replacement process [10].

3.2. Selection of the Integration Step Size. The computational speed of the integral depends not only on the length of the integral, but also on the total number of integrals. The calculations involved in each integration step are mainly related to the integration method used and are reflected in the structure and complexity of the derived functions and the upper differential equations required for each integration step. For most cases, the integration step for differential equations with the highest order, complex structure of the derivative function, and high accuracy requirements should be simulated numerically using the indefinite integration method. For the choice of integration method, it is necessary to make a reasonable choice according to the actual situation; otherwise, it will affect the determination of the integration step and the effect of its role. There are usually three choices of integration methods, which are RW method, Amder evaluation method, and gear correction method [11]. The conditions of applicability of the three methods are shown in Figure 6.

The determination of the integration step size is an important factor that affects the accuracy of the study results. If the step size is too large, it will bring a large truncation error and affect the process and effect of discrete Hindmarsh–Rose model bifurcation and chaos control. If the integration step is too small, it will lead to too many errors caused by the accumulation of the number of calculations, which will eventually increase the total error and affect the normal operation of the discrete Hindmarsh–Rose model bifurcation and chaos control during the calculation. After the integration method is determined, the choice of the integration step is particularly critical. For variables with large variations and fast change rates, it is necessary to use higher strata of the calculation method and to select a step size with a small range of magnitude [12]. For the selection of integration steps relevant to model bifurcation and chaos control, there are selection formulas specifically to determine change of steps, as shown in the following equation:

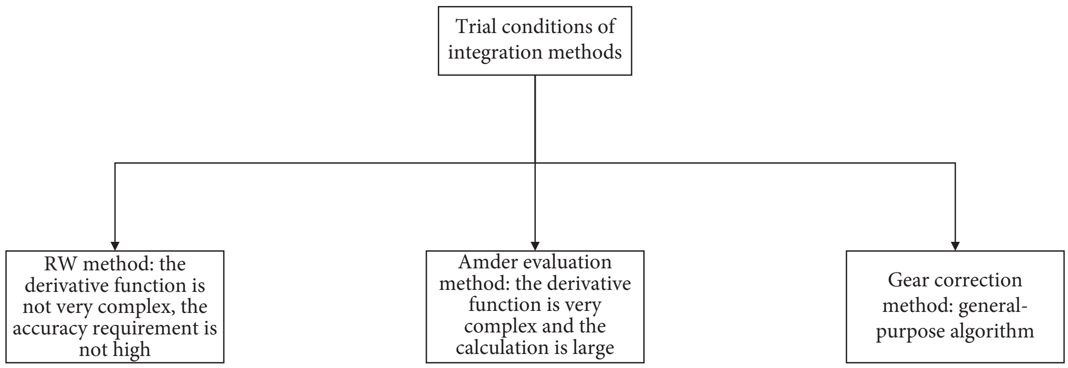


Figure 6: Conditions of applicability of the three methods.

$$\begin{cases} \dfrac{1}{20}T_{\min} \leq h \leq \dfrac{1}{5}T_{\min} \\ \dfrac{1}{20w_c} \leq h \leq \dfrac{1}{5w_c} \\ h \leq t_n/40 \end{cases}. \quad (3)$$

In equation (3), the meaning represented by Tn is the accumulation time of the step integration of the function in the discrete process; ωc denotes the integration frequency; and $t_{\min}$ denotes the minimum time constant of each part of the integration without long integration.

Usually, the integration step in the selected range is difficult to be accurately estimated, and its performance indicators are in a relatively vague range. The minimum time constant $t_{\min}$ corresponds $_{\text{to}}$ the extreme value point, which is limited to the starting form, and the largest role in the whole selection process is played by the extreme value point of the real axis. The calculation of the integration step in the invariant state is chosen according to the criteria of the starting phase, which reduces the occurrence of errors and time wastage caused by the use of too small integration steps in the later phases [13]. In general, there are three strategies for varying the step size: the first one is to vary the step size in segments, that is, by dividing the whole integration process into several segments and choosing a different step size for each segment. The second one is to make a prediction and appropriate adjustment of the next integration step with reference to the sum of errors resulting from the accumulation of each integration step. The third one is to reduce the error accumulation caused by too many integration steps and to select the largest step size without affecting the accuracy of the result of each integration step. By selecting the appropriate step size, the whole integration process is shown in Figure 7.

The above method of automatic step size adjustment has both disadvantages and advantages. The disadvantage is that it leads to an increase in local computation, which not only consumes a lot of time, but also may lead to an increase in the overall error. The advantage is that it solves the conflict between computational accuracy and computational volume in general and finds the maximum balance between the two, avoiding that the inappropriate selection of integration step size affects the later application results of discrete Hindmarsh–Rose model bifurcation and chaotic control. In the process of determining the integration not length, there are more problems related to the initial values of the system of ordinary differential equations, and solving their analytical expressions is not feasible in the general method, so it is necessary to use the numerical integration method to find the approximate solutions of the ordinary differential equations. Especially for small parameter differential equations, the choice of step size has been a complex and important issue for solving differential equations with small values of the highest-order derivative term. How the step size can be reasonably chosen is crucial to derive the influence of the integral step size on the process and effect of studying bifurcation and chaos control of the discrete Hindmarsh–Rose model [14].

3.3. Types of Bifurcation. Branching theory is the mathematical study of mass or topological variations in galaxy populations, such as the solutions of integral curves and differential equations for families of vector fields. It is usually used in the mathematical study of dynamical systems. If the value of the system parameters (branching parameters) changes smoothly, the system behavior suddenly changes "qualitatively" or topologically. There are two main types of bifurcations: local bifurcations and global bifurcations [15]. Each of the two main types contains several subtypes, as shown in Figure 8.

The analytical entry point of local bifurcation is related to the spatial variation of the equilibrium point of the discrete model in constructing the dynamical system when the parameters cross the critical values. Local bifurcation occurs when changes are made in the location of the equilibrium points due to changes in the parameters [16]. In the discrete Hindmarsh–Rose model system, the equilibrium point at zero can be called the bifurcation point, where the bifurcation parameters are placed at the bifurcation point, and the topological changes are restricted by the position of the bifurcation point to change the original change trajectory, producing a torsional effect so that the equilibrium point is restricted to the minimum neighborhood of the bifurcation point, a phenomenon

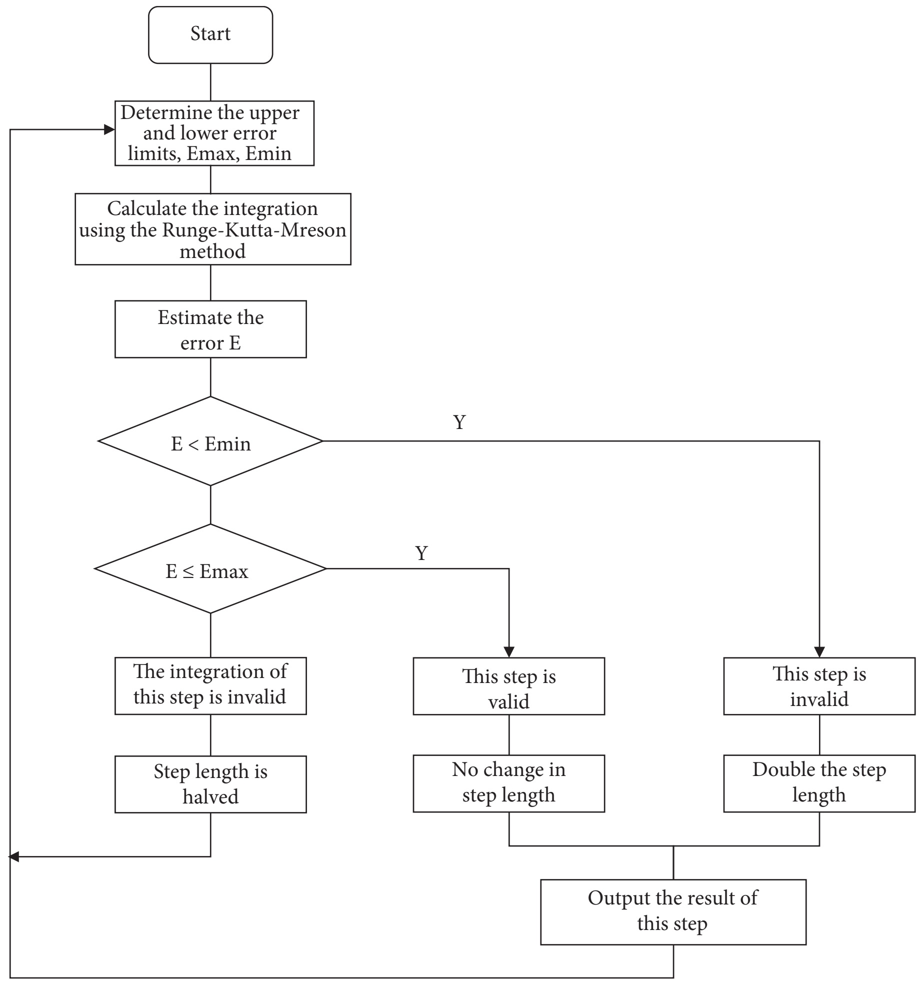


Figure 7: Integration process.

known in academia as local bifurcation. There are five types of local bifurcations, namely, folding bifurcation, interactive bifurcation, cluster bifurcation, periodic bifurcation, and Hopf bifurcation.

Folded bifurcation is often found in continuous systems. According to the multidimensional nature of space, not all equilibrium points are in a stable state, and the spatial position of equilibrium points can be replaced at any time due to the mutation and hysteresis effects, resulting in local changes of discrete model bifurcation points. Interactive bifurcation has strong special features in local bifurcation. First of all, it is characterized by the fact that the real part passes through zero values and there exists at least one immobile point where the parameter taking all values takes the equilibrium point without changing. However, if the position of the parameter changes, it should not be understood that the point changes from another fixed point [17]. In other words, there is stability in the collision process with inertial point exchange and unstable mass exchange. A manifold is an unusual type of bifurcation that implies the expansion of the system from one fixed point to several fixed points. Periodic bifurcation is the result of a change in the behavior of the system due to a small change in the system parameters, producing a cyclic iterative effect. Periodic bifurcation contains several periodic phases that possess a stable temporal pattern, and the magnitude of its periodicity is related to the degree of change of the parameter values. When the parameter values change twice as much as the original degree, the system produces twice as many iterations, then a double periodic phase is generated; that is, the period after the change is twice the original period. The Hopf bifurcation, as a replacement of the stability of the system and the critical point at which the cycle solution changes, is also a special type of local bifurcation. When the system parameters change at the position of the critical value point, the complex conjugate eigenvalues that maintain the normal operation of the system shift in the equilibrium point when passing through the imaginary axis of the complex plane, and will gradually lose its stability and produce limit loops from it.

The occurrence of global bifurcation is restricted to a small area range compared to local bifurcation, which is wider than that of local bifurcation. It usually occurs when large permanent blocks of the system collide or overlap with the equilibrium point of the system. It cannot be analyzed exclusively on the basis of the stability

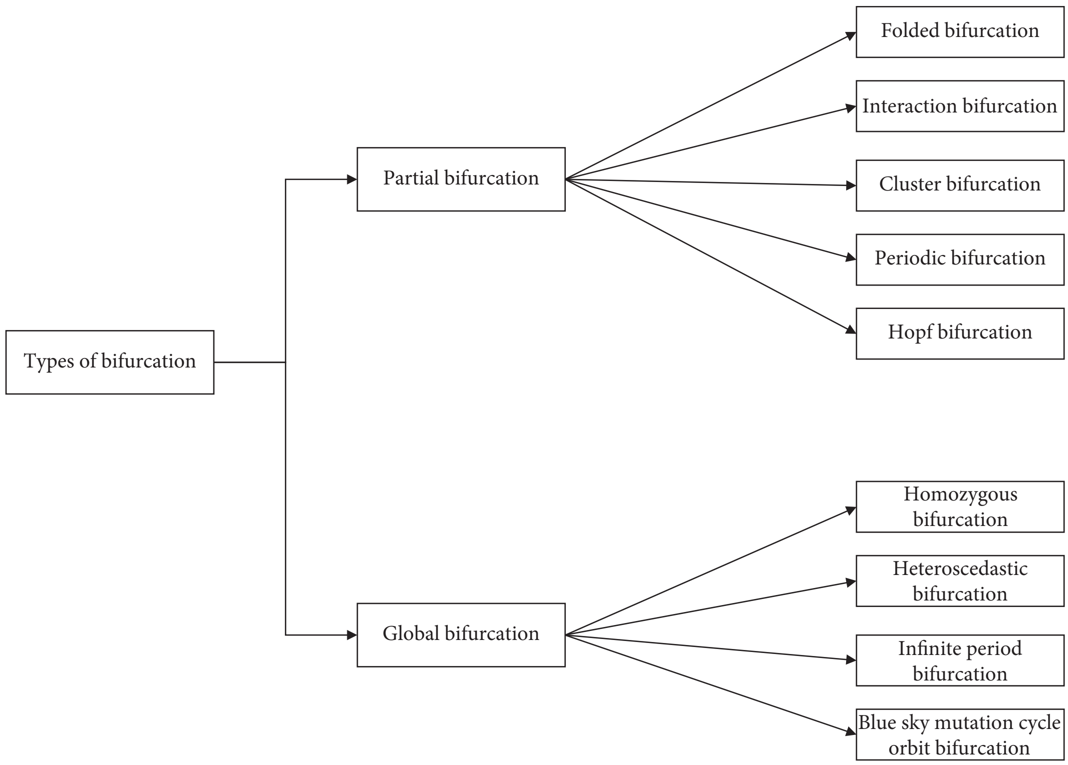


FIGURE 8: Types of bifurcation.

of the equilibrium point (stationary point). Global branches are called global branches because they occur in larger regions than microscopic areas like the partial branches of the authority. There are several types of global offices.

(1) Nighttime separation is a global phenomenon that occurs frequently in collisions of circular trajectories with saddle surfaces. A double rendezvous is produced when a periodic orbit collides with a saddle. For small values, there are saddle points and limit rings. As the bifurcation parameter increases, the boundary reaches the saddle point and produces an infinite orbit. (2) Crossing branches are global branches where the boundary ring collides with two or more saddles. The subclinical branches are divided into resonance branches and side branches. These two branches lead to changes in the stability of the cycle. The stability of the ring changes if the eigenvalues of the equilibrium points in the ring satisfy the requirements of the algebra. This is usually accompanied by the appearance or disappearance of circular paths. If the alignment point of the horizontal eigenvalues passes the zero value, a horizontal transmission cycle is generated. This can also lead to changes in subclinical resistance. (3) Branches with stable nodes and unstable saddles on the boundary are called infinite cycle branches. Infinite cycle bifurcation is a global bifurcation that can occur when two immobile points appear on the boundary ring. When the parameter limit approaches the threshold value, the oscillation speed decreases and the cycle approaches infinity, and the cycle infinity appears at that threshold value. When the threshold is exceeded, two consecutive fixed points appear in the limit loop, generating an oscillation error and forming two saddle points. In other words, it describes the possible behavior of the differential equation stabilizer by changing the equation. The duration and length of the trace is usually infinite, but it remains confined to a part of the phase space and remains stable until the bifurcation. In other words, the trace disappears into the blue sky [18].

Although the chaos theory is deeply connected to the knowledge of physics, it is actually classified as a branch of mathematics. As an interdisciplinary theory, the chaos theory represents the existence of dynamical system settings with highly connected parameters to the initial conditions in a multilatitude space of apparent randomness [19]. Just as the butterfly flaps its wings to bring a Hurricane in the butterfly effect, in a linear system with a range of parameters, subtle differences in the initial conditions present in the chaos control theory can lead to widely varying results in the operation of the dynamical system, which creates a large disturbance in the future predicted behavior and moreover makes the long-term prediction of the system behavior very difficult or even impossible. Even for very deterministic systems, the butterfly effect can occur. The initial conditions of the system already determine the long-term behavior of the dynamical system, and its behavior is not disturbed by other factors, a phenomenon defined as the chaos effect. The chaos theory has been used to varying degrees in mathematics, chemistry, physics, and biology, and since the first chaos model was proposed in the 1960s, the discussion of the chaos problem has been limited to theoretical aspects, but not to practice and application. With the expansion of the field and the deepening of the research problems, the study of the chaos theory has increased the application of its research results to human services. Due to the complexity of the chaos theory, the control of chaotic systems is very challenging. However, with the efforts of people, various strategic approaches to control chaotic systems have been discovered [20].

4. Results and Discussion

4.1. Effect of Integral Step Size on Chaotic Control of the Discrete Hindmarsh–Rose Model. In the above material, it can be recognized that the selection of the integration step size affects the accuracy of the results on discrete studies. It is important to avoid both the large truncation error caused by too large a step size, which affects the process and effect of discrete Hindmarsh–Rose model bifurcation and chaos control. It is also necessary to avoid the accumulation of errors caused by too many calculations due to too small integration steps, which eventually increases the total error and affects the normal operation of discrete Hindmarsh–Rose model bifurcation and chaos control.

For the methods of chaos control, academics are mainly divided into two major types: the feedback control method and the minimum energy control method.

American scholars Ott, Grebogi, and Yorke proposed a feedback control method to control chaos with parameter perturbation, that is, using the characteristics of parameter perturbation to make the position of immobile points move relatively to correct the unstable operation of the system state; Hunt introduced a new form of the feedback control method based on the research results of the above three scholars, and this method has extremely superiority. On the premise of not affecting the operation of the periodic orbit, the gain can be adjusted by a small perturbation, so that the change of the trajectory of the high periodic orbit can be controlled within a reasonable range. Since then, the feedback control method has been used by scientists from all over the world as an important method to study the chaos control problem. The feedback control method has been continuously improved and upgraded, and different types of feedback control methods have been introduced, including linear feedback control method, internal force feedback control method, automatic feedback control method, and dispersive feedback control method.

The minimum energy control method basically assumes the theoretical premise that in a normally operating system, the conditions for the emergence of the most stable state of the system always occur when the system energy is minimal. This method has a wide range of use, and both continuous systems and discrete systems can be more appropriate use. Although it has a wide range of applications, this method also has some limitations. For example, the system energy in chaotic systems is difficult to control within the applicable range, and it is difficult to obtain the energy function. Therefore, this method has long been less applied to the study of discrete model dynamical systems.

The scope of the study of discrete Hindmarsh–Rose models mainly includes bifurcation studies and chaotic control. The vast majority of the models are smooth systems, and the study of nonsmooth systems is still in the development field, and there are no clear indicators of development. The study of Filipov systems is a relatively new area of research in the discrete Hindmarsh–Rose model. In most cases, the convex Filipov method can be used to convert the Filipov system describing the differential equations of the right side breaks into a differential [107]. However, Filipov systems have not been investigated based on neuronal models. In neurons, the energy of the membrane potential fluctuates between depolarization (up) and hyperpolarization (down) due to the external excitation or spontaneous movement of ions in the ion channel. This upward movement is called dynamics. The signal propagates along the axis of the neuron to the end of the axis, where it connects with other neurons at the synapse. Thus, action potentials play an important role in intercellular communication. Action potentials can also be referred to as nerve impulses or peak discharges.

4.2. Effect of Integration Step on Bifurcation of the Discrete Hindmarsh–Rose Model. The formal composition of the discrete Hindmarsh–Rose model can be expressed as

$$\begin{cases} \overline{x}(t) = y(t) - ax(t)^3 + bx(t)^2 + I \\ \overline{y}(t) = c - dx(t)^2 - y(t) \end{cases}. \tag{4}$$

In the model composition, the slow variable is a quantity that plays a relatively small role and is largely negligible. In the above Hindmarsh–Rose model, the parameters represent a fixed value where $a = 1$, $b = 2$, $c = 3$, and $d = 3$. In this model, the switching control parameters can be set to b and I. Before $x(t)$ reaches the critical value MT, b and I have the maximum values $b$max and Imax, respectively, and when XT reaches the critical value MT or is in the outer domain of the critical value, b and I have the minimum values $b$min and Imin, respectively, which decreases the mode potential energy $x(t)$. For the discrete Hindmarsh–Rose model, previous studies have focused on the theoretical and numerical aspects of bifurcation and chaos control, including the critical value control of the discrete model. In practice, the discrete Hindmarsh–Rose model, excluding its own convergence effect, is rarely used as an important function in operation. Among the models, the impulse action can be topologically obstructive due to the discontinuity of the process and the difficulty of estimating the parameters. During the determination of the auxiliary peak discharge structure model, the minimum current of the membrane potential energy is usually set at 25 mA, and the constant c represents the bifurcation value of the membrane potential energy. By the above setting, the decomposition loop of the system, which is discontinuous reset bifurcation, is transformed into continuous periodic bifurcation. Combined with the experience of the above work, the new discrete Hindmarsh–Rose model subjected to the improvement yields the following equation, as shown in the following equation:

$$\left.\begin{aligned}\frac{\mathrm{d}x(t)}{\mathrm{d}t} &= y(t) - ax(t)^3 + bx(t)^2 + I \\ \frac{\mathrm{d}y(t)}{\mathrm{d}t} &= c - dx(t)^2 - y(t)\end{aligned}\right\} x(t) < MT$$

$$\left.\begin{aligned}x(t^+) &= \left(1 - \frac{\zeta x(t)}{x(t)+q}\right) \\ y(t^+) &= y(t)\end{aligned}\right\} x(t) = MT \quad . \tag{5}$$

In the new discrete model, I represents the input DC and p takes values in the range $1 < p < 2$ and q denotes the constant term. $0 < px(t)/(x(t)+q) < 1$ means that the membrane potential energy after resetting is not higher than MT. Considering that the model undergoes abrupt change effects in its state near the critical value, the system periodicity of the pulse action is problematic when compared to the conditions for the occurrence of the pulse action, which can be transformed into the mapped immobile point problem with the addition of the above conditions. Without considering the periodic solution, bifurcation and chaotic control are inevitable focal points in dynamical systems that deserve to be noticed. The path of individual behavior bifurcation to chaotic control transformation is referred to as doubly periodic bifurcation in the academic definition division. For example, if the original period is represented by k, then the doubled period is represented by $2k$, and the quadratic being period is based on the doubled period and the period can be represented by $4k$. Thus, from period k to period $2k$ and then to period $4k$, after many times of multiplication, the doubled period bifurcation has occurred in the fluid after the iterative effect, which sets the air flow rate as a variable parameter and derives the doubled period bifurcation in the neuron model. The range of the multiplicative period bifurcation is very wide, and it can appear in both neuron models and discrete Hindmarsh–Rose models with diversified structures, and it is applied to the study of the behavior of dynamical systems as an important type of bifurcation. The emergence of multiplicative bifurcation is inseparable from the nonlinear term, based on the previous theoretical results; however, according to the present study, multiplicative bifurcation caused by nonlinear impulse action has not yet appeared.

Although the impulse function in the discrete Hindmarsh–Rose model is defined by the academic community as a mapping process that is not connected, the impulse theory is used in this study for the accuracy of the results. For the same system, although the parameter values set are the same, the difference in the selection of the integration step can also affect the effect of the model dynamic transformation play. The characteristics of bifurcation control as an important control method can be corrected using different mechanisms of feedback control methods. One of the most frequently occurring methods is the nonlinear dynamic feedback control method. The target of bifurcation emergence is not only related to the conversion rate of the intrinsic bifurcation and the point of time parameter, but also related to the change of the parameters of the nonlinear differential equations that emerge from the bifurcation process. The effect of the integration step on the bifurcation of the discrete Hindmarsh–Rose model implies that when the bifurcation process is in an unstable state, the bifurcation stability can be controlled by properly controlling the integration step to generate a limit loop during the bifurcation process to delay the onset of the intrinsic bifurcation and to change the amplitude and interparameter range of the bifurcation point. Compared with the bifurcation control, the inverse control of bifurcation uses appropriate limit-loop amplitude regulation without changing the parameters, and a class of bifurcation with feedback control occurs at a suitable reservation position to realize the integral step for discrete Hindmarsh–Rose model bifurcation and chaos control.

Based on the purpose of the inverse control system to achieve the local balance of bifurcation through the change of parameters on the normal bifurcation of the system output, this article introduces a local bifurcation method in which a pair of real values in the form of complex conjugate complexes is passed through the imaginary axis without changing the function of the dynamical system, and the operation process of the dynamical system still has continuity and stability, and the limit ring in the form of small amplitude is still generated near the immobile point bifurcation type. This type of bifurcation is a common type of bifurcation, and it is not difficult to see it in systems such as chemistry, physics, ergonomics, engineering cost, Internet of Things, and computer networks. The discrete Hindmarsh–Rose model after being affected by a change in the integral step size is shown in the following equation:

$$\begin{cases}\overline{x}(t) = y(t) - ax(t)^3 + bx(t)^2 - \int z(t) + I \\ \overline{y}(t) = c - dx(t)^2 - y(t) \\ \overline{z}(t) = r\left(s\left(x(t) - x_r\right) - z(t)\right)\end{cases}, \tag{6}$$

where $x(t)$ is the membrane potential energy; $y(t)$ and $z(t)$ are the recovery variables; a, b, c, d, f, r, s, and I are the system parameters; and $x\gamma$ is the rest potential energy value of the system. In the present case, the explicit criterion for bifurcation and chaos control of the Hindmarsh–Rose model has not been reasonably formulated. The bifurcation and chaos control of the model can be regarded as a way to change the control of limit loops or nonlinear oscillations, and the new bifurcation solution that appears here is the optimal solution obtained after the integration step is regulated, which can be used as a new type of bifurcation solution to balance the parameter composition inside the discrete model system. This also demonstrates that the step size of the integral has a large impact on the bifurcation and chaotic control of the discrete Hindmarsh–Rose model, and provides a reference for the study of chaotic control behavior.

5. Conclusion

The subject of this article is the effect of the integral step size on the bifurcation and chaos control of the discrete

Hindmarsh–Rose model under the information technology-based differential equation dynamical system. In the introductory part of the paper, a brief introduction to information technology and differential equation dynamical systems is given, followed by the introduction of the concept and meaning of the discrete Hindmarsh–Rose model. In the research background, a detailed introduction to the recent developments of information technology and differential equation dynamical systems is further given to set a good background premise for the research content of this article. The article will analyze how the integral step size affects the discrete Hindmarsh–Rose model bifurcation and chaotic control and its effect based on the differential equation dynamical system of information technology. The advantages of information technology in information management and information processing are introduced into the application of differential equation dynamical systems, which can ensure the stability of differential equation dynamical system models. The determination of the integration step size is an important factor affecting the accuracy of the research results. Too large or too small a step size can affect the process and results of discrete model bifurcation and chaos control; therefore, in this article, we understand how the change in the integration step size affects the three-dimensional discrete Hindmarsh–Rose model bifurcation and chaos control through the selection of a suitable integration step size.

In the Materials and Methods section, the differential equation dynamical system is denoised to obtain an optimized differential equation dynamical system that overcomes the shortcomings of the original differential equation dynamical system. And an overview of the chaos control theory is given. Finally, in the Results and Discussion section, the effect of the integration step on the chaotic control of the discrete Hindmarsh–Rose model and its effect on the bifurcation of the discrete Hindmarsh–Rose model are discussed, and it is hoped that the theoretical and practical experiences obtained can provide subsequent workers with the understanding of the bifurcation and chaotic control of the discrete Hindmarsh–Rose model. It is hoped that the theoretical and practical experience gained will help subsequent workers in their understanding of discrete Hindmarsh–Rose model bifurcation and chaotic control.

In view of the limitation of research time and personal ability, there is still some incompleteness in this article. There are still many questions about the effect of the integration step on the bifurcation and chaos control of the discrete Hindmarsh–Rose model that need to be further explored and discovered.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] B. Martin, S. Scindia Pallavi, and T. Sanket, "Qualitative results for nonlinear integro-dynamic equations via integral inequalities," *Qualitative Theory of Dynamical Systems*, vol. 21, no. 4, 2022.

[2] M. M. Metwali, "Solvability in weighted L1-spaces for the m-product of integral equations and model of the dynamics of the capillary rise," *Journal of Mathematical Analysis and Applications*, vol. 515, no. 2, p. 126461, 2022.

[3] J. Anthony and A. May Dave, "An efficient partial-differential-equation-based method to compute pressure boundary conditions in regional geodynamic models," *Solid Earth*, vol. 13, no. 6, 2022.

[4] W. Rui, "Separation method of semi-fixed variables together with dynamical system method for solving nonlinear time-fractional PDEs with higher-order terms," *Nonlinear Dynamics*, vol. 109, no. 2, pp. 943–961, 2022.

[5] C. Han, L. Li, G. Su, and T. Sun, "Dynamical behaviors of a k-order fuzzy difference equation," *Open Mathematics*, vol. 20, no. 1, pp. 391–403, 2022.

[6] R. Arezou and A. Zakieh, "Numerical approach for solving two dimensional fractal-fractional PDEs using peridynamic method," *International Journal of Computer Mathematics*, vol. 99, no. 3, 2022.

[7] B. Lakshmi and Z. Fabio, "Monotonicity of the period function and chaotic dynamics in a class of singular ODEs," *Journal of Mathematical Analysis and Applications*, vol. 507, no. 2, 2022.

[8] T. H. Thai, N. A. Dai, and P. T. Anh, "Global dynamics of some system of second-order difference equations," *ELECTRONIC RESEARCH ARCHIVE*, vol. 29, no. 6, p. 4159, 2021.

[9] M. V. Nikolaev, U. Dieckmann, and A. A. Nikitin, "Application of special function spaces to the study of nonlinear integral equations arising in equilibrium spatial logistic dynamics," *Doklady Mathematics*, vol. 104, no. 1, pp. 188–192, 2021.

[10] O. Mitsuhiro and K. Inoue, "Modeling viral dynamics in SARS-CoV-2 infection based on differential equations and numerical analysis," *Heliyon*, vol. 7, no. 10, 2021.

[11] O. Zaid, V. Suat Erturk, P. Kumar, and V. Govindaraj, "Dynamics of generalized Caputo type delay fractional differential equations using a modified Predictor-Corrector scheme," *Physica Scripta*, vol. 96, no. 12, 2021.

[12] T. J. Sego, J. O. Aponte-Serrano, J. F. Gianlupi, and J. A. Glazier, "Generation of multicellular spatiotemporal models of population dynamics from ordinary differential equations, with applications in viral infection," *BMC Biology*, vol. 19, no. 1, p. 196, 2021.

[13] T. A. N. G. Shengbing, K. Fujimoto, and I. Maruta, "Learning dynamic systems using Gaussian process regression with analytic ordinary differential equations as prior information: regular section," *IEICE - Transactions on Info and Systems*, vol. 104, 2021.

[14] R. Guo, Q. Huang, and Q. Liu, "Some new oscillation criteria of even-order quasi-linear delay differential equations with neutral term," *Mathematics*, vol. 9, no. 17, p. 2074, 2021.

[15] J. Xu and T. Caraballo, "Dynamics of stochastic nonlocal partial differential equations," *The European Physical Journal Plus*, vol. 136, no. 8, p. 849, 2021.

[16] H. Yacine, A. Allam, and B. Zineb, "Dynamical behavior of a P-dimensional system of nonlinear difference equations," *Mathematica Slovaca*, vol. 71, no. 4, 2021.

[17] O. Alexey, P. V. Isabel, P. Gleb, and T. Mirco, "CLUE: exact maximal reduction of kinetic models by constrained lumping of differential equations," *Bioinformatics*, vol. 37, no. 12, 2021.

[18] T. S. Doan and P. E. Kloeden, "Semi-dynamical systems generated by autonomous caputo fractional differential equations," *Vietnam Journal of Mathematics*, vol. 49, no. 4, pp. 1305–1315, 2021.

[19] A. Moch Fandi, S. K. Adji, S. Novriana, and G. Iman, "Dynamics of bank's balance sheet: a system of deterministic and stochastic differential equations approach," *International Journal of Mathematics and Computer Science*, vol. 16, no. 3, 2021.

[20] C. Zhang and D. Yang, "Nonlocal stochastic functional differential equations driven by G-Brownian motion and mean random dynamical systems," *Mathematical Methods in the Applied Sciences*, vol. 43, no. 12, 2020.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 4636485, 9 pages
https://doi.org/10.1155/2022/4636485

Research Article

Artificial Intelligence-Assisted Decision-Making Method for Legal Judgment Based on Deep Neural Network

Wenqing Ma

Gansu University of Political Science and Law, Gansu, Lanzhou 730070, China

Correspondence should be addressed to Wenqing Ma; mwq6824@163.com

Received 19 July 2022; Revised 9 September 2022; Accepted 26 September 2022; Published 11 October 2022

Academic Editor: Xuefeng Zhang

With the arrival of the third revolution of artificial intelligence, the applications of artificial intelligence in the fields of automatic driving, image recognition, smart home, machine translation, medical services, e-sports, and so on can be seen everywhere, and topics about artificial intelligence are constantly emerging. Since 2017, the discussion on artificial intelligence in the field of law has become more and more active. In this context, the application of artificial intelligence in the field of legal judgment and the hypothetical system based on this technology in court judgment has also become the object of discussion from time to time. In this paper, based on the artificial intelligence decision-making method of the deep neural network, aiming at the three subtasks of legal judgment prediction, namely, crime prediction, law recommendation, and sentence prediction, a multi-task judgment prediction model BERT12multi and a sentence interval prediction model BERT-Text CNN are proposed, which improve the prediction accuracy and adopt the knowledge distillation strategy to compress the model parameters and improve the reasoning speed of the judgment model. Experiments on the CAIL2018 data set show that the performance of the deep neural network model in crime prediction and law recommendation tasks can be significantly improved by adopting the pre training model adaptive training, grouping focus loss, and gradient confrontation training strategies. Using a step-by-step sentence prediction strategy can realize the weight sharing of pre training model and make use of the prediction results of charges and laws in sentence prediction. The recall training-prediction strategy can avoid error accumulation and improve the accuracy of sentence prediction. By integrating the artificial intelligence decision-making method, the case reasoning speed can be greatly improved, the highest compressible model volume can be about 11% of the original one, and the reasoning speed can be increased by about 8 times. At the same time, performance close to that of the deep neural model can be obtained, which is superior to other legal decision prediction models based on word embedding.

1. Introduction

With the advent of the Internet and the big data era, the development of all walks of life is more and more combined with computer science. As the most cutting-edge technology of computer science, artificial intelligence is also frequently combined with the industry in the era of network development. In recent years, the combination of the judicial field and artificial intelligence technology is generally favored by the outside world. The combination of the two is an innovative exploration of the application of artificial intelligence technology, which is conducive to improving the level of judicial reform. The reason why it is generally optimistic is that the application of artificial intelligence technology improves the efficiency of judicial adjudication, and further saves manpower and time [1]. However, due to the limitation of the current level of information network technology, there are still many problems in the application of artificial intelligence technology in the field of legal judgment. Legal judgment prediction is the most typical application of artificial intelligence technology, especially natural language processing methods in the judicial field. The legal judgment prediction task generally includes sub tasks such as crime prediction, relevant law prediction, criminal sentence prediction, and so on. Through the study of legal materials such as adjudicated historical cases and published judgment documents, the prediction model is constructed by using machine learning or a deep learning algorithm to infer the judgment results.

Aiming at crime prediction, legal recommendation, and judicial decision prediction, this paper proposes a hybrid deep neural network model HAC (hybrid attention and CNN model) for long text classification. And Deep Pyramid Convolutional Neural Networks. The model's prediction of crime and the recommended F1-Score (mean of Micro-F1 and Macro-F1) for the relevant statutes are 85% and 87%, respectively. For the prediction of the sentence, due to the differences in regions, ages, courts, judges, and defendants' attitudes, it will become more difficult to predict judicial decisions [2]. The model has excellent predictive performance and generalization ability and can adapt well to these differences. At the same time, the output results of the model's crime prediction and legal recommendation are added to the input of the judicial decision prediction task, and the classification method is used to predict the judicial decision, which further improves the effect of the model. Finally, in the sentence prediction task F1-Score was more than 77% and obtained excellent results predicted by the CAIL2018 judicial ruling.

It is undeniable that the application of artificial intelligence technology has further improved legal efficiency and saved legal resources, but there are still many drawbacks in terms of the application status of artificial intelligence. On the one hand, we have high expectations for the promotion of artificial intelligence technology; on the other hand, we are troubled by the threat theory of artificial intelligence. Whether the bottleneck of artificial intelligence technology in the application of legal judgment will hinder further integration and development in the legal field poses a severe challenge to both the theoretical circle and the expert group [3]. Therefore, scholars and specialized institutions are committed to judicial reform, aiming at realizing a networked and intelligent legal adjudication system by studying the combination of judicial and artificial intelligence technology. The innovation of this paper is the following: First, by analyzing the specific practice and theoretical research of artificial intelligence in legal judgments at home and abroad, we will gain insight into the loopholes existing in artificial intelligence in legal judgments in our country, and put forward relevant suggestions for the problems that arise. Second, the status and role of artificial intelligence in the application of legal judgments are demonstrated from three aspects: legislation, judiciary, and court system; evidence. By demonstrating the rationality of the application of artificial intelligence in the judicial field of our country, we will discuss the development direction of artificial intelligence technology in the field of law in China in the future, and gradually improve the new development model of the combination of law and computer science in my country.

This article is organized into seven chapters. The first chapter is the introduction part. This part analyzes the application of artificial intelligence in the field of legal judgment and the current situation of the hypothetical system based on the practice of this technology in court judgments summarizes the reasons for the problems and compresses new models to improve the judgment model. Inference speed. The second chapter mainly summarizes the relevant literature, summarizes the advantages and disadvantages, and puts forward the research ideas of this paper. The third chapter introduces the application characteristics of artificial intelligence in legal adjudication in detail. Chapter 4 discusses pre trained models based on deep neural networks. The fifth chapter expounds on the application of artificial intelligence to assist decision-making in legal judgments, which is enough to prove the positive significance of the application of artificial intelligence technology to judicial construction. The sixth chapter is the part of the experimental results, which analyzes the research results in detail and the outlook based on the results. The seventh chapter is the conclusion, summarizing the significance of the research.

2. Related Work

The most far-reaching impact of training word vector models with neural networks is the Neural Network Language Model (NNLM) proposed by Chen et al. The NNLM model is based on the n-gram language model and uses a three-layer linear neural network. 1 word is used as input to predict the nth word [4]. This method of word vector training lays the groundwork for the technical direction of word vector training, and later generations of research on word vector training methods are mostly inspired by this. In their published papers, Hong et al. proposed the idea of using word vectors to solve various tasks such as part-of-speech tagging and named entity recognition in natural language processing. The open-source system SENNA [5]. According to the idea of transfer learning and the use of Transformer structure, researchers proposed a large-scale pre training model represented by Schemmer et al. [6]. The pre training model generally adopts a multi-layer Transformer structure to perform unsupervised pre training on a large corpus. In practical applications, the pre training model has the defects of low computational efficiency and large resource consumption. The researchers mainly use strategies such as knowledge distillation to compress the volume of the pre training model. Rastogi et al. proposed a Text CNN text classification model based on a one-dimensional convolution structure, which first maps the text into vectors and then uses one-dimensional convolutions of different sizes to capture the local semantic information of the text and capture the local semantic information of the text through pooling. Important feature information is input to the classifier to obtain the probability distribution of the labels [7]. Gerards and Borgesius proposed a text classifier based on a deep convolutional structure. By increasing the network depth, the vector representation of the text is refined for classification [8]. The patient knowledge distillation strategy (Patient Knowledge Distillation, PKD) proposed by Wu and School for pre training models such as BERT, extracts rich information through the deep structure of the teacher network and adds supervision from the middle layer of the teacher model in the distillation process, which improves the depth. The performance of the network model [9]. Qian et al. designed a dual feedback mechanism with multi-view forward prediction and backward verification to match the numerical collocations in the text, aiming at the numerical unit keywords in the legal text, such as the content of alcohol, the weight of drugs, and the amount of theft. It improves the ability of the model to capture the matching information of numbers and keywords [10]. Solum used a keyword extraction algorithm to mine judgment documents to extract the keywords of crimes, integrated the deep learning model, and proposed the MTL-Fusion model. It effectively improves the model's ability to distinguish easily confusing charges [11]. Walker et al. applied the deep learning text classification model HAN and DPCNN to legal judgment prediction, integrated and improved the model for judicial intelligence applications, and proposed a legal judgment prediction model HAC (Hybrid attention and CNN model) [12]. Applying the pre training model to the legal judgment prediction task can further improve the performance of the judgment prediction model and improve the reliability and accuracy of the prediction. Brennan-Marquez and Henderson proposed a BERT-based HIER-BERT model to handle very long case texts [13].

3. Application Characteristics of Artificial Intelligence in Legal Adjudication

3.1. Electronicization of Legal Information. In recent years, the electronic application of data and information has brought technological innovation to the court. The rise of the "smart court" is the technical product of the combination of artificial intelligence and the court system. The combination of artificial intelligence and judicial technology makes judicial judgment have the characteristics of electronic information. Through the data collection and analysis of laws and cases, artificial intelligence can process judicial data more scientifically and accurately [14]. The exertion of this feature allows the judge to find the necessary theoretical basis at any time according to the circumstances and needs of the case. Through intelligent analysis and statistical data, it not only saves a lot of time for investigation and data search but also prevents information omission. The application of this technology also provides convenience for the smooth development of judicial adjudication, such as intelligent identification of the parties, and retrieval of all relevant information. Finally, the application of artificial intelligence makes the office of judicial adjudication intelligent and further liberates manpower [15].

3.2. Intelligence to Prevent Judgment Risks. Artificial intelligence can prevent the judgment trap of "similar cases with the same judgment" to the greatest extent. Through the judge's ability to analyze massive case files using artificial intelligence technology, it can help judges identify the differences between massive cases and check the differences between similar cases, preventing the judgment trap of "similar cases with the same judgment." Secondly, artificial intelligence can prevent data loopholes through the use of artificial intelligence technology to supervise judicial personnel, intelligently monitor the process of legal judgment, monitor the illegal behavior of judges in real-time, and prevent judicial personnel from operating under the shadows to affect judicial justice and cause unnecessary legal judgments, anthropogenic risk [16]. In this way, judicial corruption can be effectively suppressed and the risk of legal judgment can be reduced, as shown in Figure 1.

4. Pre Training Model Based on Deep Neural Network

4.1. Pretrained Model Encoder. In this paper, the BERT pre training model is used as the encoder of the model to enhance the performance of the model, and further adaptive training is carried out on the CAIL2018 data set to enhance the performance of the pre training model. The subtasks of crime prediction and law recommendation are classified by multi-labels, two subtasks are jointly trained in one model, and each category of crime and law is classified by binary [17]. The categories of crimes and laws are grouped according to the number of samples, and the focus loss of grouping is used to improve the classification ability of the model for unbalanced samples. In the fine-tuning of the pre training model, gradient confrontation training is adopted to further improve the performance of the pre training model. The algorithm flow is shown in Figure 2.

For the legal judgment prediction task, due to the complex content of the judgment document, the defendant's behavior differences, the items involved and other information will directly affect the judgment results, and there are high requirements for the information extraction ability of the model [18]. And in the task of accusation prediction and related law prediction, the number of samples of some categories is scarce. If the model of traditional word vector combined with neural network structure is used, it is difficult to accurately understand its semantic information only through a small number of samples in the training set. Using the pre training model can effectively enhance the performance of the model, and the pre training model has strong generalization, which is conducive to the model's understanding of some rare categories of samples [19].

In the current research, the encoder or decoder in the Transformer structure is often applied independently to extract the features in text or sequence data. For example, the Bert and GPT prep models adopt the transformer multilayer encoder and decoder structure as a hidden layer, as shown in Figure 3.

The Transformer structure abandons the convolutional neural network structure and the recurrent neural network structure and adopts the self-attention mechanism to extract the information in the sequence data [20]. First, map the sequence into three sets of vectors of key (Key), value (Value), and query (Query) through three sets of neural networks with different weights, and then map the query vector of each unit in the sequence to the key vector of all units in the sequence. The scaling dot product is calculated, and the attention score is calculated through the softmax function mapping. The final encoded hidden representation at a certain position is jointly determined by the attention score and the value vector, namely,

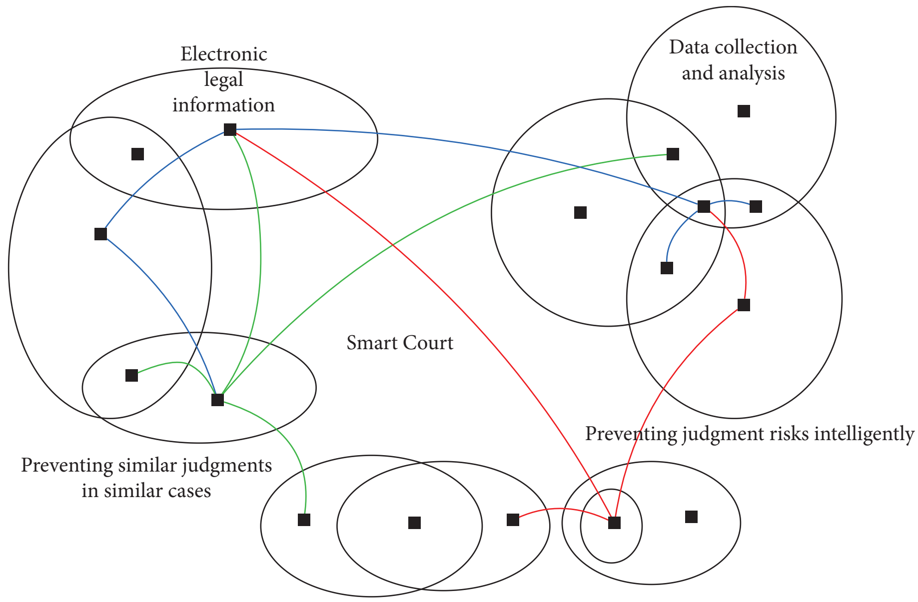


Figure 1: Application features.

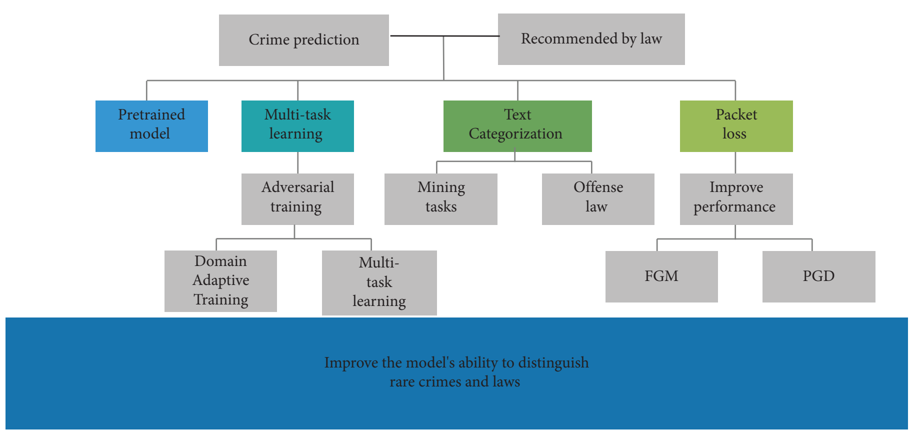


Figure 2: Algorithm flow chart.

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK}{\sqrt{d_k}}\right)V. \quad (1)$$

In practical use, the Transformer structure adopts a multi head attention mechanism similar to the multi convolution kernel structure in a convolution neural network, which further improves the nonlinearity of the model and enables the model to pay attention to multi-level semantic information, namely,

$$\begin{aligned} \text{head}_i &= \text{Attention}(QW_i, KW_i, VW_i), \\ \text{MultiHead}(Q, K, V) &= \text{Concat}(\text{head}_1, \text{head}_2, \cdots, \text{head}_h). \end{aligned} \quad (2)$$

Compared with the mainstream recurrent neural network structures such as (RNN, GRU, and LSTM) used for text or sequence data modeling, the Transformer structure does not have the risk of gradient disappearance, has a stronger ability to capture long-distance dependencies, and can be calculated in parallel, which can effectively improve the efficiency of training and reasoning, but the calculation of the self-attention mechanism in the Transformer has a high space complexity and requires high memory resources in the calculation.

4.2. Multi-Task and Multi-Label Text Classification Model. For the task of crime prediction and law recommendation, in studies, only one crime or law regulation was often considered, and it was regarded as a multi-category text classification task. In practice, the trial of a case often involves several related laws, and the defendant has the situation of combined punishment for several crimes. Therefore, the task of crime prediction and law recommendation is treated as a multi-label text classification task, and all crimes and law categories are classified in binary, so as to calculate the probability that each judgment document belongs to each category.

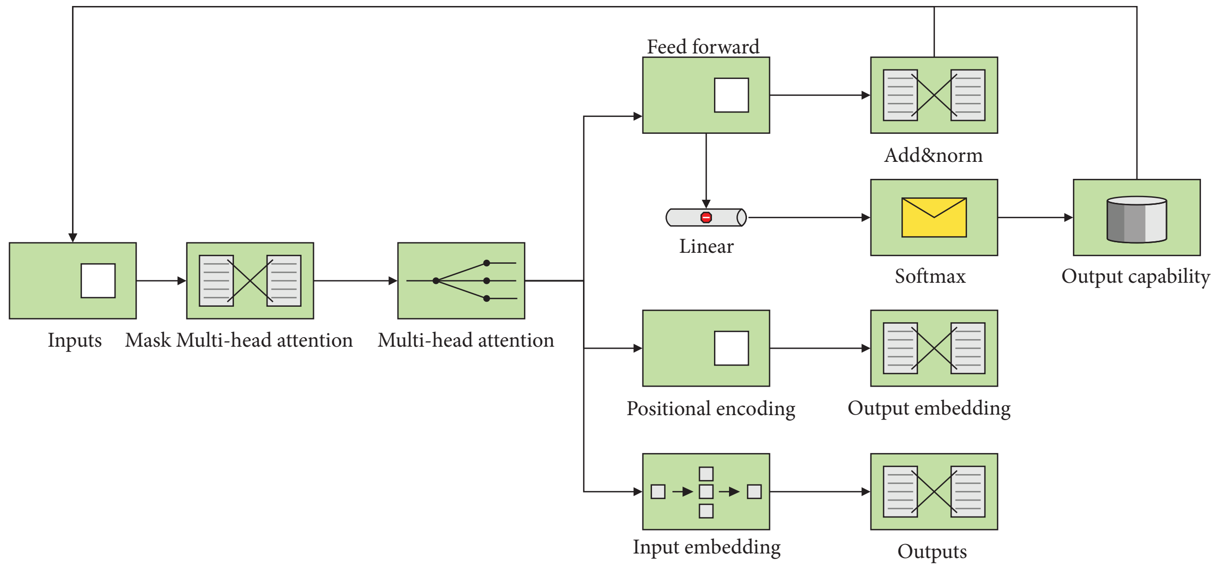


Figure 3: Schematic diagram of the Transformer structure.

Completes or truncates the preprocessed case text into a uniform length, converts it into an index sequence matrix according to BERT's pre training vocabulary, inputs BERT models in batches, and uses the output vector corresponding to the special symbol position of (CLS) after coding by BERT model as the sentence vector of the input text to input into the classifier. After dropping, the text sentence vector is mapped into the corresponding category by the classifier, namely,

$$\begin{aligned} h_i &= \mathrm{BERT}(x_i, \theta), \\ y_i &= \delta(\mathrm{Dropout}(W \cdot h_i)). \end{aligned} \tag{3}$$

Here, h_i represents the output vector of the sample x_i input BERT model at the corresponding position of the (CLS) mark; W is the weight matrix of the classifier; $\delta(x)$ represents the sigmoid activation function, and y_i represents the class probability predicted by the model for sample x_i. This multi-task learning model with 12 hidden layers is named $\mathrm{BERT}_{12\mathrm{multi}}$ to distinguish it from the student model that uses fewer layers of Transformer encoding layers in knowledge distillation.

In the training of the neural network, for binary or multi-label classification problems, the binary cross-entropy (Cross Entropy Loss) is generally used as the loss function, namely,

$$\begin{aligned} L_{\mathrm{CE}}(x_i) &= -\log y_i, y_i = 1, \\ L_{\mathrm{CE}}(x_i) &= -\log((1 - y_i). \end{aligned} \tag{4}$$

The two-category cross-entropy loss function has no tendency to positive and negative samples. When the negative samples in the data set are far more than the positive samples, the ordinary cross-entropy loss function will make the classification results tend to be negative samples. Moreover, for multilabel or multiclassification, the problem is that the uneven distribution of samples in different categories will cause the classification results to tend to the categories with more samples so that the recall rate of the categories with fewer samples can be lower than that of the categories with more samples, which affects the final classification results.

4.3. Adversarial Training. In the field of computer vision, a certain amount of noise interference is often added to the picture samples of the training set to carry out a certain degree of anti disturbance training, enhance the ability of the model to resist disturbance, reduce the dependence of the model on the original training set, make the model more robust, have higher anti disturbance ability, and also play the role of expanding data to a certain extent. Image data itself can be regarded as a matrix, and each pixel can be regarded as a group of vectors. However, for text data, because the character representation is discrete, it is impossible to directly add disturbance in the field of computer vision. In natural language processing tasks, it is often necessary to vectorize the discrete character representation through the embedding layer, so the gradient confrontation training can be realized by disturbing the embedded layer according to the gradient direction of the embedded layer. Confrontation training can generally be expressed in the following format:

$$\min E_{(x,y)\sim D}\left[\max_{\Delta x \in \Omega} L(x + \Delta x, y, \theta)\right]. \tag{5}$$

Among them, D represents the training set, x represents the input of a single sample, Δx represents the applied perturbation, L represents the loss function of the sample x, and Ω represents the perturbation space. Common adversarial training methods in natural language processing are FGM (Fast Gradient Method) and PDG (Projected Gradient Descent).

The FGM adversarial training strategy is as follows:

The sample x mapped by the embedding layer, through forwarding propagation, calculates the loss function, and obtains the gradient g of x:

$$g = \Delta_x L(\theta, x, y). \tag{6}$$

Calculate the perturbation radius r_{adv} according to the perturbation range ε and the gradient g of the embedding layer:

$$r_{\text{adv}} = \varepsilon \cdot \frac{g}{|g|_2}. \tag{7}$$

Compared with the FGM confrontation method, which directly increases the disturbance to the boundary of the disturbance radius, the PGD confrontation method realizes the confrontation process step by step and gradually adjusts the disturbance. When the disturbance range exceeds the disturbance space, it reduces the disturbance momentum and looks for the best disturbance range.

5. Legal Decision-Making AI-Assisted Decision-Making Applications

Now, in relatively developed cities, intelligent trial systems have been specially developed to assist in the trial of judicial cases. For example, the "trial-centered litigation service software" system developed by a city's high court, the system uses intelligent scoring and induction to conduct case trials, shortening the trial time. This system converts litigation materials into digital materials by scanning and systematically entering them, which is convenient for digital extraction in judicial practice. It provides basic information such as intelligent comparison and verification of evidence for judicial adjudication by automatically pushing cases of the same nature, which provides convenience for judicial case handling. There is also the "intelligent trial-assisted sentencing adjudication system" of a provincial intermediate court. Through further analysis of previous cases, the court's rate of re sentencing and retrial initiation has shown a downward trend, as shown in Figure 4. It is enough to illustrate the positive significance of the application of artificial intelligence technology to judicial construction.

The specific parameters and settings of the whole artificial intelligence design decision RESNET algorithm include the input layer, residual module, batch normalization layer, pooling layer, and activation function. The input layer takes the preprocessed construction machinery modeling image metadata as the data set, including the preprocessed image, semantic label, and scoring data. The final data of the convolution layer is converted into a 13-layer fully connected network output. After the operation of the first layer of the convolution layer, it is divided into three main residual modules. The main diameter is extracted through three convolution layers. The purpose of the two-path fusion is to fuse the deep and shallow features of the design image data and semantic data. In this way, we can get a better feature classification effect. As shown in Figure 5, the design decision accuracy curve can be used to judge the quality of the model. It can be seen that with the increase in the number of iterations, the design decision accuracy of the RESNET artificial intelligence decision model gradually increases. When the number of iterations is about 160, the decision accuracy tends to be stable.

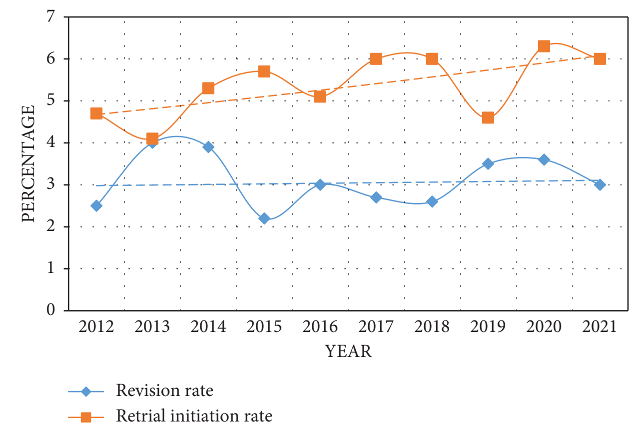


Figure 4: The retrial initiation rate and the rate of revision of the judgment in the provincial intermediate court.

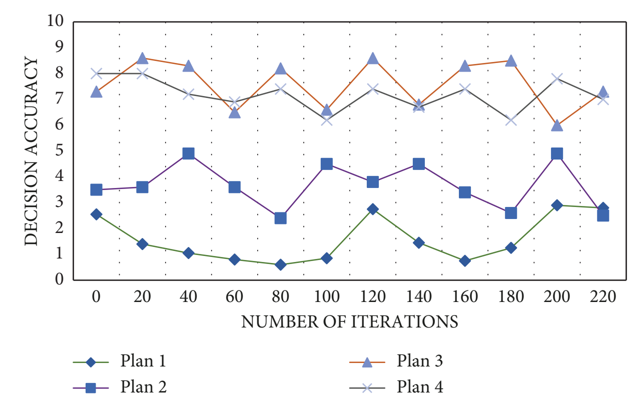


Figure 5: Accuracy comparison.

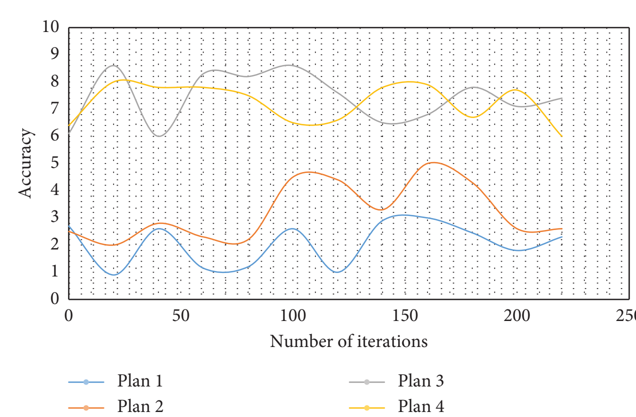


Figure 6: Comparison of decision satisfaction.

Figure 6 shows that the three modeling image words of science and technology, sensibility, and integrity are used as the semantic vocabulary labels of product design, and the three design scheme renderings are used as the decision accuracy curve comparison chart after image input. After the corresponding data training process, the design decision accuracy of the RESNET artificial intelligence decision model is further improved and shows a stable trend.

The input of the deep neural network generally uses the embedding method to convert the sparse one-hot encoding into a dense vector representation. At the same time, studies have shown that the combined use of multiple different types of embeddings can effectively improve the ability of the model. In our model, three embedding connections (concatenates) are used: a 256-dimensional word embedding followed by training, a pre trained 100-dimensional word vector, and a pre trained 200-dimensional word vector. The pre training uses the C language version of Word2Vec released by Google and performs unsupervised training on the case description and the original text of the facts of the data set provided by the sponsor. The training results are shown in Figure 7.

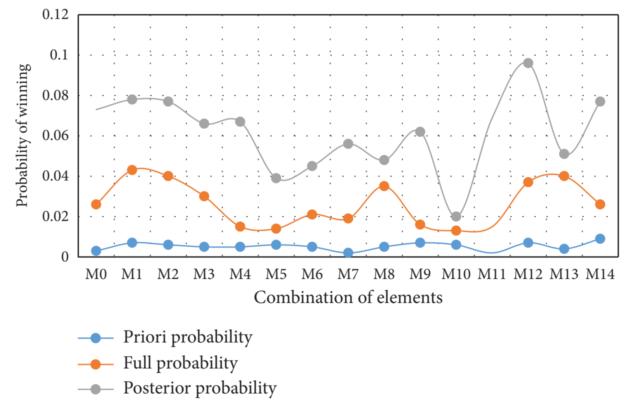


Figure 7: Prediction results of winning probability of experimental data points and similar data points.

6. Experimental Results

Referring to the existing research of text classification technology, the tf-idf + svm machine learning model and the mainstream deep learning text classification model based on word embedding text CNN model, Han model, and DPCNN model are used as comparison models. After word segmentation of the case text, the fixed sentence length is 300 words, word2vec and glove word embedding algorithms are used, and the fixed word embedding dimension is 200. Word vectors are trained on all data sets respectively. It is compared with the open source word vector with a dimension of 200 obtained by directional skip Gram (DSG) algorithm in a large corpus. According to the performance of each model in different word vectors, word2vec word vectors are finally used in the text CNN model and Han model, and open source word vectors are used in the DPCNN model. For the two tasks of crime prediction and legal recommendation, each comparison model is modeled independently and jointly, and the best results of the model are recorded.

First, the comparison models such as TF-IDF + SVM and TextCNN are compared with the BERT12multi which uses the BERT pre training model as the encoder. At the same time, ablation experiments are performed on the effects of joint training and adaptive training, as shown in Table 1.

Experiments are carried out on the effect of grouping focus loss and the effect of gradient adversarial training. Considering the computational resource consumption of establishing independent models for the two tasks, the comparison models in the experiment all use the same joint training method as BERT12multi for multi-task learning. Because BERT12multi using adaptive training has better results in the experiment, BERT12multi in the following experiments adopts adaptive training. The experimental results are shown in Table 2 and Figure 8:

As shown in Table 2 and Figure 8, each model using grouped focus loss has improved on the Macro-F1 index, indicating that grouped focus loss can effectively improve the performance of each high model in the unbalanced data set of category samples and reduce the judgment error for small sample categories without significantly increasing the amount of super parameters. The bert12multi model perturbed by FGM and PGD improves the F1 score of the two tasks by about 0.001 and 0.003 respectively, indicating that the gradient confrontation training strategy can improve the generalization of the text classification model, and then improve the accuracy of decision prediction.

With the rapid development of artificial intelligence technology, the main development trends for assistant decision-making systems in the future are as follows:

(1) Assistant decision-making under man-machine fusion. Due to the physiological limitations of human perception and decision-making ability, human beings cannot control future wars. The traditional method of controlling the whole battlefield factors by human beings has been unable to cope with the rapidly changing high-intensity future battlefield environment. Human soldiers will have to gradually jump out of the battle chain and cooperate with machine intelligence as planners, administrators, and commanders in most cases to control the process of the whole war, so as to promote the war from the war between humans to a new form of man-machine collaborative warfare, including precise perception, natural interaction in the environment, man-machine collaborative perception, and man-machine fusion computing.

(2) Evolution of human-machine intelligence for decision-making. Machine intelligence may overturn the traditional human war decision logic. The new game confrontation strategy based on deep reinforcement learning, big data, and supercomputing shows a new data-driven optimal game strategy generation method and shows a strong game confrontation ability and self-learning evolution ability. In the future, the learning of evolutionary war decision logic driven by the mixture of knowledge and data will comprehensively replace the traditional military game decision logic based solely on operations research. The evolving new artificial intelligence technology has become a disruptive force in future military game decision-making. Mechanized war thinking will be replaced by new intelligent war thinking. The machine intelligence of non human traditional war methods has greatly extended human

Table 1: Comparison of evaluation indicators of each model

Mission name	Crime prediction task			Statutory recommendations		
Model name	Micro-F1	Macro-F1	F1-score	Micro-F1	Macro-F1	F1-score
SVM for task1	0.815	0.648	0.792	—	0.592	0.683
SVM for task2	—	—	—	0.773	—	—
Text CNN for task1	0.841	0.751	0.813	0.848	0.721	0.785
Text CNN for task2	0.875	0.753	0.814	—	0.721	0.753
Text CNN multi	—	—	—	0.85	—	—
HAN for task1	0.878	0.733	0.822	0.841	0.737	0.783
HAN for task2	0.987	0.764	0.819	—	0.739	0.791
HAN multi	—	—	—	0.842	—	—

Table 2: Comparison of evaluation indicators of each model 2.

Mission name		Crime prediction task			Statutory recommendations	
Model name	Loss function	Micro-F1	Macro-F1	F1-Score	Micro-F1	Macro-F1
Text CNN	Cross entropy	0.875	0.753	0.814	0.85	0.721
HAN	Cross entropy	0.879	0.764	0.822	0.842	0.739
DPCNN	Focal loss	0.882	0.766	0.824	0.858	0.739
BERT12multi	Focal loss	0.884	0.782	0.822	0.847	0.758

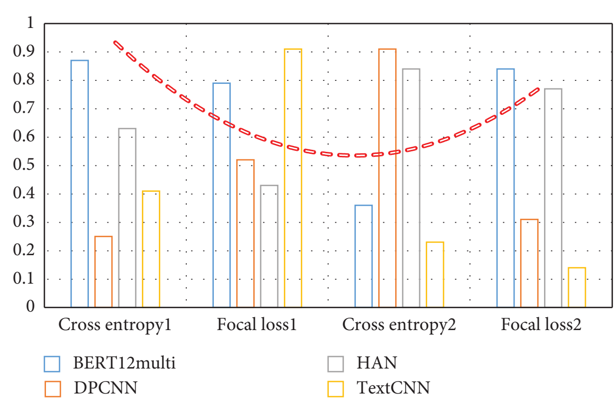


Figure 8: F1-score comparison of each model under different loss functions.

intelligence, which is the latest cognition of human beings to explore the laws of war and enhance intelligence by human-computer mixing.

7. Conclusions

With the progress of the times, the series connection of information and intelligent network pulse is the requirement of the times of legal development. The application of artificial intelligence in legal judgment is a progressive exploration of the combination of computer science and law. The application of artificial intelligence technology further improves the efficiency of legal case handling, promotes the standardization of the legal process, and strengthens the construction of legal data. The application of artificial intelligence in legal judgment also provides a reference for solving problems and improves the ability of dispute resolution. The practice has proved that the combination of the legal field and artificial intelligence technology is the choice of the times, and the better the development, the more prosperous the legal civilization can be created. However, from the perspective of the long-term development of the Internet and big data, the exploration of artificial intelligence in legal judgment still needs to overcome many difficulties, for example, strengthening the research of artificial intelligence in cognitive reasoning and solving technical logic obstacles. By incorporating artificial intelligence decision-making methods, this paper can greatly improve the speed of case reasoning, with the highest compressible model volume reaching about 11% of the original, and the reasoning speed being increased by about 8 times. At the same time, performance close to that of the deep neural model is obtained, which is superior to other prediction models of legal judgment based on word embedding. It is far from enough to have such research results. It takes a long time to explore and accumulate practical experience in order to establish a complete artificial intelligent legal judgment theory.

Data Availability

All data used in this study are presented in the manuscript.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] H. Karimi and J. Tang, "Decision boundary of deep neural networks: challenges and opportunities," *WSDM '20: The Thirteenth ACM International Conference on Web Search and Data Mining*, vol. 29, no. 11, p. 389, 2020.

[2] Y. Guan, J. Zhao, and Y. Qiu, "Melodic phrase segmentation by deep neural networks," *Social Science*, vol. 20, no. 1, p. 20, 2018.

[3] M. Pirhooshyaran and L. V. Snyder, "Simultaneous decision making for stochastic multi-echelon inventory optimization with deep neural networks as decision makers," 2020, https://arxiv.org/abs/2006.05608.

[4] L. Chen, J. W. Guo, and W. U. Jian-Gong, "Research on artificial intelligence assisted decision-making algorithms for lawyers based on legal-computing theory," *Journal of East China Normal University*, vol. 28, no. 23, p. 928, 2019.

[5] H. Hong, X. Lin, and K. Tang, "Artificial-intelligence assisted decision making: a statistical framework," *Social Science Electronic Publishing*, vol. 39, no. 2, p. 488, 2020.

[6] M. Schemmer, P. Hemmer, and M. Nitsche, "A meta-analysis on the utility of explainable artificial intelligence in human-AI decision-making," *Journal of Artificial Intelligence*, vol. 30, no. 4, p. 29, 2022.

[7] C. Rastogi, Y. Zhang, and D. Wei, "Deciding fast and slow: the role of cognitive biases in AI-assisted decision-making," *Computer deep learning research*, vol. 29, no. 23, p. 98, 2020.

[8] J. Gerards and F. Z. Borgesius, "Protected grounds and the system of non-discrimination law in the context of algorithmic decision-making and artificial intelligence," *Social Science Electronic Publishing*, vol. 249, no. 23, p. 48, 2019.

[9] X. Wu and L. School, "The possibility and limits of modeling judicial decision-making in artificial intelligence," *Zhejiang Social Sciences*, vol. 3, no. 12, p. 48, 2017.

[10] D. J. Qian, S. U. Hang, and S. O. Law, "The change of China's legislative decision-making mode under the perspective of artificial intelligence," *Journal of Hunan University of Science &Technology (Social Science Edition)*, vol. 2e, no. 2, pp. 18–28, 2018.

[11] L. B. Solum, "Legal personhood for artificial intelligences," *Legal Information Retrieval and Applications*, vol. 38, no. 4, p. 98, 2019.

[12] V. R. Walker, N. Carie, and C. C. Dewitt, "A framework for the extraction and modeling of fact-finding reasoning from legal decisions," *Artificial Intelligence and Law*, vol. 19, no. 2, p. 49, 2021.

[13] K. Brennan-Marquez and S. Henderson, "Artificial intelligence and role-reversible judgment," *Journal of Criminal Law and Criminology*, vol. 20, no. 2, p. 423, 2019.

[14] D. H. Berman and C. D. Hafner, "The potential of artificial intelligence to help solve the crisis in our legal system," *Communications of the ACM*, vol. 32, no. 8, pp. 928–938, 1989.

[15] H. P. Olsen, J. L. Slosser, and T. T. Hildebrandt, "What's in the box? The legal requirement of explainability in computationally aided decision-making in public administration," *Social Science Electronic Publishing*, vol. 309, no. 33, pp. 128–210, 2019.

[16] K. Hawkins, "On legal decision-making," *Washington and Lee Law Review*, vol. 23, no. 39, p. 477, 2017.

[17] D. Merkl and E. Schweighofer, "The exploration of legal text corpora with hierarchical neural networks," *Guide tour in public international law*, vol. 34e, no. 23, p. 384, 2017.

[18] A. Preece, R. Ashelford, and H. Armstrong, "Hows and whys of artificial intelligence for public sector decisions: explanation and evaluation," *Public affairs management*, vol. 43, no. 34, p. 280, 2018.

[19] N. E. Sondak and V. K. Sondak, "Neural networks and artificial intelligence," *ACM*, vol. 2, no. 7, pp. 241–245, 2017.

[20] J. Zeleznikow and A. Stranieri, "The split-up system: integrating neural networks and rule-based reasoning in the legal domain," *The fifth international conference of Springer-Verlag*, vol. 32, no. 89, p. 990, 2020.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9894243, 1 page
https://doi.org/10.1155/2023/9894243

Retraction

Retracted: Application and Analysis of Hypergraph Association Rule Redundancy Algorithm in Data Mining

Mobile Information Systems

Received 26 September 2023; Accepted 26 September 2023; Published 27 September 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] H. Pang and L. Zhou, "Application and Analysis of Hypergraph Association Rule Redundancy Algorithm in Data Mining," *Mobile Information Systems*, vol. 2022, Article ID 1193586, 11 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 1193586, 11 pages
https://doi.org/10.1155/2022/1193586

Research Article

Application and Analysis of Hypergraph Association Rule Redundancy Algorithm in Data Mining

Huanli Pang[1] and Lianzhe Zhou[2]

[1]*School of International Education, Changchun University of Technology, Changchun 130012, Jilin, China*
[2]*School of Computer Science and Engineering, Changchun University of Technology, Changchun 130012, Jilin, China*

Correspondence should be addressed to Lianzhe Zhou; zhoulianzhe@ccut.edu.cn

Received 28 July 2022; Revised 17 September 2022; Accepted 24 September 2022; Published 11 October 2022

Academic Editor: Xuefeng Zhang

In order to realize "from individual data research to data system research" and "from passive data verification to active discovery," this study proposes a hypergraph-based association rule redundancy processing algorithm in data mining. This study introduces the concepts of hypergraph and system, explores the establishment of hypergraph on a three-dimensional matrix model, and adopts a new hyperedge definition method according to the characteristics of big data and the concept of the system, which improves the ability to deal with problems; the association rules are transformed into a directed hypergraph, and the adjacency matrix is redefined. The detection of redundancy and loops is transformed into the processing of connected blocks and cycles in the hypergraph. The experimental results show that two UCI datasets were selected, namely, the balloons dataset and the shuttle landing control dataset, in which the minimum support and minimum confidence of the balloons dataset are both 5%. The dataset has 4 attributes, and 18 association rules are obtained through the Aprior algorithm. Although the running time of the coevolution algorithm is slightly longer than that of the other two global optimization algorithms, the running time is completely within the acceptable range. Moreover, due to the effective introduction of the idea of coevolution, compared with the use of the other two algorithms for association rule mining, it not only has a better mining quality but also has a significant advantage in the ability to jump out of the local optimal solution, realizing the search of high-quality association rules in high-dimensional datasets. *Conclusion*. This model provides a new idea and method for the redundant processing of association rules.

1. Introduction

The Internet of things, cloud computing, and other information technologies are updating day by day and constantly integrated with the human world, economy, politics, military, scientific research, life, and other fields. The speed of data generation is rapid, and the amount of data is increasing day by day, giving birth to a huge amount of data [1]. Data visualization aims to express data clearly and effectively through a graphical representation and uses visualization to find data connections that are not easy to observe in the original data. Information visualization is conducive to enhancing users' understanding of high-dimensional and large-scale data and plays an important role in association rule mining, recognition, and understanding. As an important method of knowledge discovery and pattern recognition, association rule mining aims to find valuable relationships in the form of if then. Association rule visualization is an indispensable subset of association rule research. Its main goal is to display data and help users with insight into the results of association rule mining.

Data mining can explore the hidden rules in data and give full play to the value of data. Association rule mining can extract potential and valuable frequent patterns or relationships between attributes from data [2]. Text can clearly and intuitively show frequent patterns and related relationships, but due to the limited cognitive ability of users, the value of association rule mining cannot be fully reflected. Hypergraphs are widely used in many fields of information science. Previous information visualization and visual analysis techniques mainly focused on the simple binary information between data objects. However, the research found that multiple relationships can more natural express the internal relationships and patterns hidden in information. A hypergraph is a generalization of ordinary graphs in a topological structure, and it can intuitively show multivariate relationships [3]. This also provides strong conditions and theoretical support for the visualization of association rules. The hypergraph model combines the advantages of hypergraph and directed graph and can be used for a visual representation of association rules. Nodes in the graph represent data items, and edges represent association relationships. The support and confidence of rules can be expressed by different values and colors.

2. Literature Review

Khan et al. proposed a big data entity recognition algorithm based on a graph, which maps high-dimensional data relationships in the graph, where the edge represents some data relationships, and the weight on the edge represents the near degree of association between items. This method avoids the explicit calculation of the degree of association between high-dimensional data and has made corresponding progress in the reduction of high-dimensional data [4]. Skuratovskii et al. proposed the concept of neighborhood knowledge granularity from the perspective of granular computing to evaluate the granulation ability for high-dimensional data features and combined it with neighborhood dependency as a heuristic function for data attribute reduction [5]. Baert et al. analyzed the concept of information granularity and granularity division for the granularity selection of characteristic attributes of high-dimensional data systems, accurately reflected the roughness of data dimensions in decision-making systems, and made up for the defect of reduction based on domain attributes only when high-dimensional data are granulated in a big data environment [6]. Shekhawat et al. proposed to define neighborhood complementary entropy and neighborhood complementary conditional entropy through analytical simulation and replacement of information particles, so as to obtain nonmonotonic high-dimensional data attribute granulation and nonmonotonic high-dimensional data attribute reduction. In the research process on high-dimensional data features, the above three algorithms ensure data value, accurately capture data features, and reduce data complexity. Therefore, how to preprocess high-dimensional data through granulation and accurately capture their data characteristics is a hot issue in high-dimensional data mining [7]. Elmanakhly et al. proposed the load balancing strategy of high-low frequency division and grouped the nodes evenly by estimating the number of tasks, to avoid the problems of data skew and overload [8]. Liu et al. proposed a TBLB algorithm, which combines node energy and node degree to form a load balance tree for path selection according to the path performance evaluation factor. The formation of the balance tree effectively balances the node load and greatly improves the node energy consumption [9]. Xiao et al. proposed the mrpropost algorithm, which gets the f-list of frequent 1-itemsets after the first MapReduce task is executed and constructs the PPC tree to mine the frequent itemsets of multiple computing nodes distributed on it. This process does not need to save the PPC tree in memory, which can not only quickly calculate the itemset support but also reduce the time and space consumption of the algorithm [10].

In this study, a retarget-based hypergraph analysis based on a three-dimensional matrix model is used for project data analysis. The dataset measures the performance and feasibility of the model and the data mining algorithm according to it.

3. Research Methods

3.1. Hypergraph. The Knowledgebase is a knowledge base that uses semantic research to gather data from multiple sources to improve research efficiency. Knowing Atlas is an art form that contains many objects and elements in the real world and their relationships and is used to represent all objects and their relationships in the real world [11]. As shown in Table 1, the knowledge map can be divided into layer structure and data layer of the logical architecture phase.

Although the knowledge map is widely used, the representation methods based on triples often oversimplify the complexity of the data stored in the knowledge map; especially, for hyper-relational data connecting two or more entities, the loss of high-order structure information will lead to the limitation of knowledge hypergraph representation and reasoning ability. Relevant work has proved that, in Freebase, more than 33.3% of entities and 61% of relationships cannot be represented by binary relationships. A knowledge hypergraph is a special kind of heterogeneous graph. In order to understand the characteristics of a knowledge hypergraph more clearly, we first study the representation of a heterogeneous hypergraph. According to its relevance to knowledge hypergraph, the representation method of knowledge hypergraph is further studied. Finally, a three-tier architecture of the knowledge hypergraph is proposed, which can effectively improve the reasoning ability and efficiency of the knowledge hypergraph. The definition, characteristics, and main tasks of hypergraph and correlation graph are shown in Table 2. Where, V refers to the number of node types and E refers to the number of relationship types.

3.2. Redundancy Rule Detection. Let the association rules R_1: $X \longrightarrow Y$ and R_2: $A \longrightarrow B$. If $(A \cup B) \subseteq (X \cup Y)$ and $X \subseteq A$ are satisfied, the total number of redundant rules is $(3^{|Y|} - 2^{|Y|} - 1)$, where $|Y|$ is the number of items contained in the itemset Y.

Theorem 1. *The theorem proves that, under the existing evaluation criteria, there will be a large number of redundant rules that can be deleted in the mining association rules, and it theoretically analyzes the total number of redundant rules [12, 13].*

Table 1: Logical structure of knowledge map.

Logical structure level	Primary coverage	Example
Mode layer	Data model of knowledge class	Concept and relationship
Data layer	Specific data information	Fact triplet

Table 2: Definition of hypergraph and related graph.

Graph classification	Definition	Characteristic	Task
Isomorphic graph	$H = (V, E)$, where V represents the node set and E represents the relationship set	$\|TV\| = 1, \|TE\| = 1$, its node and relationship type are single, which can only represent binary relationship data	Link prediction, node classification, and clustering
Heterogeneous graph	$H = (V, E)$, where V represents the node set and E represents the relationship set	$\|TV\| + \|TE\| > 2$, which has more than one node and relationship, and can only represent binary relationship data	Link prediction, node classification, and clustering
Heterogeneous hypergraph	$H = (V, E)$, where V represents the node set and E represents the relationship set; $e \in E, e(v_1, v_2, \cdots, v_n) \subseteq V$ for any superedge	$\|TV\| > 1$, its node types are diverse, which can represent hyper-relational data, and the relationship usually has no semantics	Link prediction, node classification, and clustering
Knowledge hypergraph	$H = (V, E)$, where V represents the entity set and E represents the hyper-relational set; the basic unit is the hyper-relational fact facta = (e, 1, 2,..., n)	$\|TV\| > 1, \|TE\| > 1$, with various types of entities and relationships; it can represent hyper-relational data and rich relational semantics	Link prediction and hyper-relational fact judgment

Definition 1. (association rule redundancy). Redundant rules can generally be divided into two forms: one is dependent rules; that is, if the conclusion of rule X_i is the same as that of rule X_j and while the premise of X_i is the sufficient condition of premise X_j, then X_j is redundant, and repeated rules can be regarded as a special case of dependent rules [14]. The second is the repeated path rule. If there are selectors X_i and X_j in the rule base and there are at least two paths between X_i and X_j, it can be determined that there are redundant rules.

Dependent rules can be represented by rules (1) and (2):

$$X_2 \longrightarrow X_4, \tag{1}$$

$$X_2 X_3 \longrightarrow X_4. \tag{2}$$

It can be seen from rules (1) and (2) that the subsequent items of the two rules are the same, and the previous item has an intersection, so we think that rule (2) is a redundant rule; then, delete rule (2) and retain rule (1); that is, retain the party with fewer children in the previous item, in which rules (1) and (2) become dependent rules.

Repeated path rules can be represented by rules (3) and (4):

$$X_1 \longrightarrow X_2 X_3 \longrightarrow X_4, \tag{3}$$

$$X_1 \longrightarrow X_5 \longrightarrow X_4. \tag{4}$$

According to rules (3) and (4), there are two paths from X_1 to X_4. We think that the path is repeated, and delete one of them.

3.3. Directed Hypergraph Representation of Association Rules. In a directed hypergraph, the directed hyperedge $e \in E$ is defined as an ordered pair composed of head node $H(e)$ and tail node $T(e)$, and $H(e)$ and $T(e)$ are subsets of vertex set V; that is, it can be composed of a set of multiple vertices. This feature is conducive to the representation of association rules as a directed hypergraph [15]. According to the correspondence between the head node $H(e)$ and the subsequent term of the association rule and the tail node $T(e)$ and the previous term of the association rule, each association rule can be uniquely represented as a super edge in the directed hypergraph.

The form of association rules obtained in practice is $x_1 x_2 \ldots x_m \Rightarrow y_1 y_2 \ldots y_n$; that is, the first item Ante(R) is a set composed of multiple items, and the second item Cons(R) also contains multiple items. We define the rule that the latter item contains only one item as a simple rule and the rule that the latter item contains multiple items as a composite rule. This project defines a directed super edge to represent an association rule. The front term of each association rule corresponds to the head node of a directed hypergraph, and the rear term of the association rule corresponds to the tail node of the same directed hypergraph. There are multiple head nodes and tail nodes for each directed super edge, so the composite rule is successfully represented.

This study adopts a spanning tree-based classification method to remove association rule redundancy. This is a new redundancy check method for association rules, which can effectively check the redundant rules, subordinate rules, and duplicate path rules. Since the adjacency matrix of the directed hypergraph is mainly used in simple graphs and the directed hypergraph we want to use here has composite points, which makes the composite rules only represented by the directed hypergraph, the adjacency matrix must be redefined. A spatial database is integrated with spatial relational data and object-relational data to realize a database of spatial data. The generation process of a spatial database includes the logical structure design of the database and the integrated storage of spatial data. Among them, the logical structure design of the database uses the classic E-R (entity connection) diagram to describe the real geographical world, and the number of paths between layers is proportional to the number of data attribute features. The specific design is shown in Figure 1.

3.4. Graphic Representation and Processing of Redundancy Rules. The adjacency matrix of a hypergraph expression completely defines the relationships of the vertices in figures. The adjacency matrix of an expressed hypergraph based on organizational rules describes the interrelationships of the objects of organizational policies. Retrieval can be accomplished according to the information hypergraphs according to the definition of redundancy rules in Definition 1 and related items of the diagram.

Definition 2. (road). Let graph $G = (V(G), E(G))$; its path is a finite nonempty sequence $W = v_0 e_1 v_1 e_2 \cdots e_k v_k$, that is, the alternating sequence of vertices and edges, where $e_i \in E(G)$ and $v_j \in V(G)$, where e_i is associated with v_{i-1} and v_i, respectively, $1 \le i \le k, 0 \le j \le k$, which is recorded as (v_0, v_k) path, called vertex v_0, v_k are the starting point and end point of path W, respectively, $v_1, v_2, \ldots, v_{k-1}$ is the inner vertex of path W, and k is the length of W. If $e_1, e_2, \ldots, e_k$ in path W is different from each other, it is called trace. If $v_0, v_1, v_2, \ldots, v_{k-1}, v_k$ in trace W is different from each other, it is called a path. If the starting point and ending points of a path (trace and road) are the same, it is called a closed path (closed trace and closed circuit). Closed trace is also called circle [16].

From the definition in graph theory, we found that, to realize the processing of redundant rules in a directed hypergraph based on association rules, it can be transformed into discovering connected blocks in the hypergraph and transforming it into a spanning tree [17]. Because each edge in the directed hypergraph represents an association rule and when the connected graph becomes a spanning tree, the edge needs to be deleted, this edge is the redundant rule in the association rule. Reduction of redundant rules:

(1) If $E_i \cap E_j \neq 0$, called the associated super edge, then there is the following formula:

$$\left| E_i \cap E_j \right| = 1. \tag{5}$$

(2) If condition (1) is true and $|E_i| \neq 2$, there must be the following formula:

$$\left| E_j \right| \neq 2. \tag{6}$$

Then, hypergraph $H = (X, E)$ has a spanning hyper tree T.

Lemma 1. *Let the hypergraph H have n vertices, m hyperedges, and c connected branches. If and only if the following equation $\sum_{i=1}^{m} (|e_i| - 1) = n - c$ exists, H does not contain a superloop [18].*

The bipartite graph corresponding to hypergraph $H = (V, E)$ refers to the vertex set V and hyperedge set E of h as vertex sets. When $v_i \in e_i$ in H, vertex v_i and vertex e_i in the bipartite graph are edge connected. The bipartite graph corresponding to H is represented by G<H>. Figures 2(a) and 2(b) show a hypergraph and its corresponding bipartite graph.

We get the number of connected blocks contained in the directed hypergraph and the location of the connected blocks where each point is located. On this basis, we must perform spanning tree processing on each connected block. Delete redundant rules by obtaining the spanning tree.

3.4.1. Algorithm to Get Spanning Tree. At present, there are generally two methods to find the spanning tree of a connected graph: the ring-breaking method and the ring-avoiding method. The so-called loop breaking method is to break all loops in a connected graph, and the remaining connected graph without loops is a spanning tree of the original graph. This algorithm is called the "loop breaking method." Take an arbitrary edge e_1 in graph G, find an edge e_2 that does not form a loop with e_1, and then find an edge e_3 that does not form a loop with $\{e_2, e_3\}$. This continues until the process cannot be carried out. At this time, the obtained graph G is a spanning tree. This algorithm is called the "circle avoiding method." According to the meaning of the hypergraph we generated, each hyperedge represents an association rule. So, obviously, we should use the broken circle method.

Input the adjacency matrix of the connected block and get an adjacency matrix of the spanning tree. By restoring the adjacency matrix of the spanning tree, we can eliminate the redundancy of the association rules. In practice, we find that there is often more than one spanning tree of a connected graph. At the same time, this will present a problem; that is, in the obtained rules, the rules that people are interested in and think are important may be deleted. In order to solve this problem, we give a certain weight to the more important association rules. Reflected in the graph is to give weight to each edge. Combined with the current algorithm, we give a smaller weight to the edge corresponding to the more important association rules and a larger weight to the unimportant and uninterested association rules and then use the prim algorithm to calculate the minimum spanning tree.

Definition 3. (minimum spanning tree). In figure $H = (V, E)$, (u, v) represents the edge-connecting vertex u and vertex v, that is, $(u, v) \in E$, and $W(u, v)$represents the weight of this edge. If there is a subset of Tas E, that is, $T \subseteq E$ and acyclic graph, making $W(T)$ of $W(T) = \sum_{(u,v) \in T} W(u, v)$ minimum, this T is called the minimum spanning tree of H. The minimum spanning tree is actually an abbreviation of the minimum weight spanning tree.

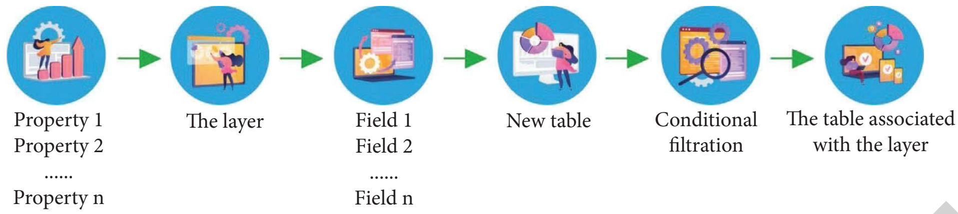


FIGURE 1: E-R model between layers.

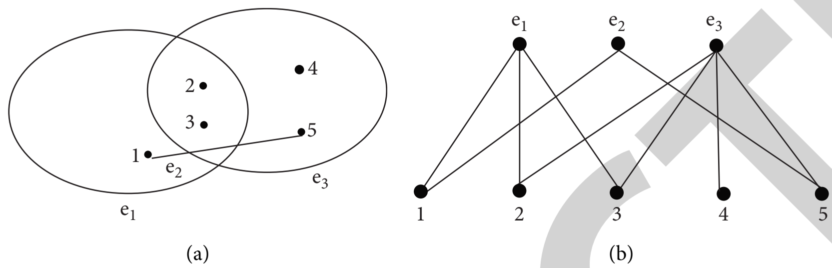


FIGURE 2: Hypergraph and its corresponding bipartite graph. (a) Hypergraph. (b) Divided into two pictures.

3.4.2. Basic Idea of Prim Algorithm. Starting from a vertex of the connected graph $H = (V, E)$, select the edge (u_0, v) with the smallest weight associated with it and add its vertices to the vertex set U of the spanning tree. In each subsequent step, select the edge (u, v) with the smallest weight from the edges where one vertex is in U and the other vertex is not in U and add its vertices to the set U. In this way, until all vertices in the graph are added to the vertex set u of the spanning tree, a minimum spanning tree is obtained.

Through practice, it has been found that the edge set of the minimum spanning tree is sometimes different. We introduce weight to deal with the minimum spanning tree, which plays a corresponding protective role in the preservation of important and interesting rules.

3.5. Algorithm Flow. Figure 3 shows the outline of the tree spanning plan in this study to eliminate the recurrence of the organizational policy. The special steps are as follows.

(1) Analyze test data, create aggregation rules, use hypergraphs to represent participatory rules, and revise and obtain its integers

(2) The preprocessed adjacency matrix is obtained by subtracting the algorithm [16, 19]

(3) Unspanned and linked spanning trees are obtained by the spanning tree algorithm

(4) The adjacency matrix of the surrounding tree is reconstructed by the organizational law, and its final completion is possible

This study proposes an algorithm to remove subordinate rules by redefining the adjacency matrix. Each association rule is defined as an edge of a directed hypergraph. According to the previous section, the redefining adjacency matrix is obtained. The columns of the matrix represent the subsequent terms of the association rule. The flowchart of the algorithm is shown in Figure 4. After this algorithm, all the subordinate rules in the redundant rules can be deleted, and the preprocessed adjacency matrix can be obtained.

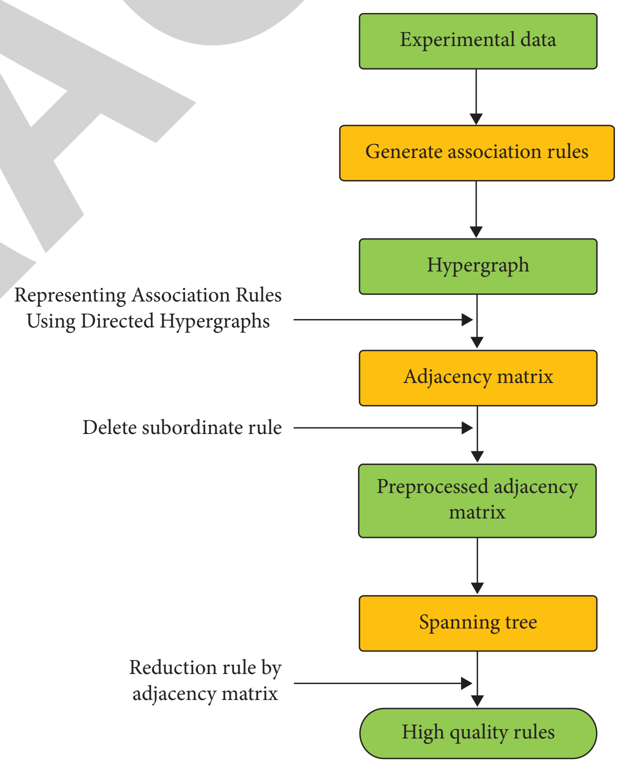


FIGURE 3: Flowchart of association rule redundancy detection.

4. Result Analysis

4.1. Verification Results. In this study, the spanning tree-based classification method to eliminate the redundancy of connection policies consists of three modules: the redefine adjacency matrix module, the delete dependency rule

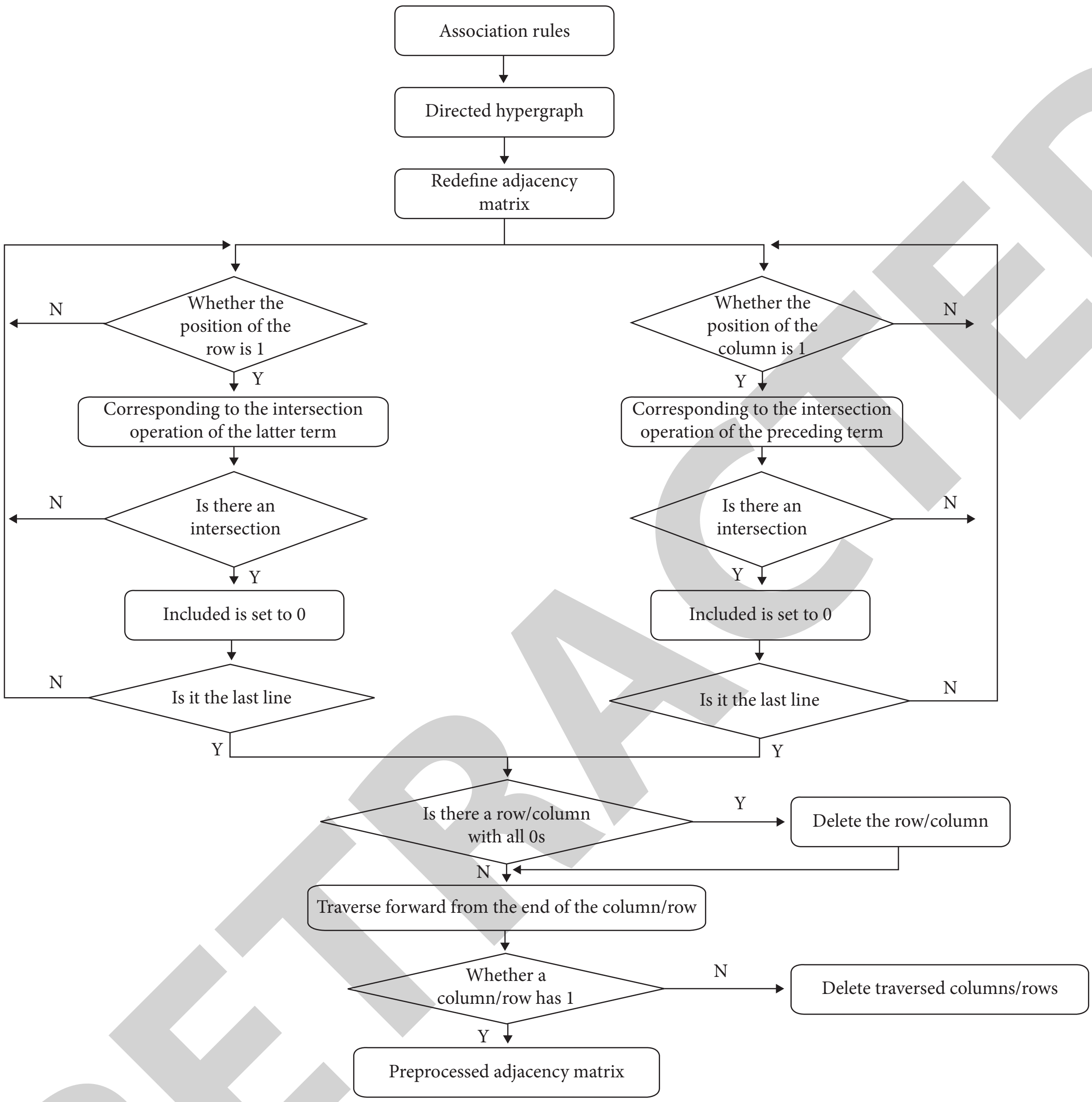


Figure 4: Flowchart of removing dependent rules.

module, and the spanning tree module. The first two modules are used in VB programming, while the spanning tree module is used in MATLAB. The special points are shown in Figure 5.

After removing the coding algorithm and the spanning tree algorithm, we obtain a tree spanning without connection. The special procedure of the spanning tree algorithm and hypergraph instructions is shown in Figure 6. It can be seen from Figure 6 that the preprocessed adjacency matrix (the result of the dependency removal algorithm) yields the total deleted adjacency matrix. The redundant, on the right-hand side, shows the variation of the hypergraph indicator before and after the spanning tree algorithm [17, 20].

Through the method introduced in this study, the redundant rules are removed accurately and quickly in both datasets. The specific results are shown in Table 3.

4.2. Experiment. The experimental data studied in this study came from the data obtained from the special task project of Humanities and Social Sciences Research of the Ministry of Education, "Research on building a scientific and complete network culture construction and management system of colleges and universities," the special project of moral education innovation and development of a city, "analysis of the influence factors and validity of social environment on young students," and the project of moral education innovation and development

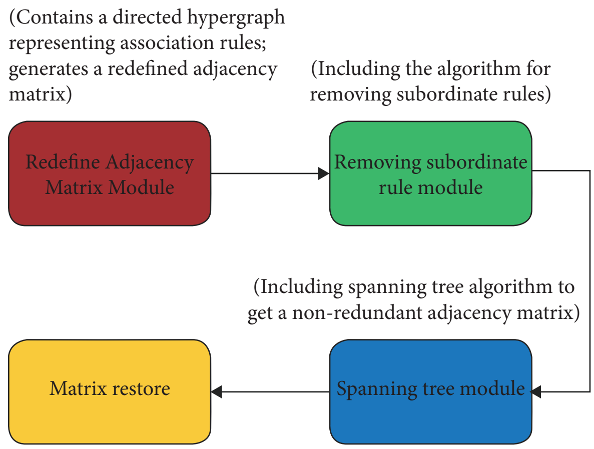


Figure 5: Module composition of classification redundancy removal method.

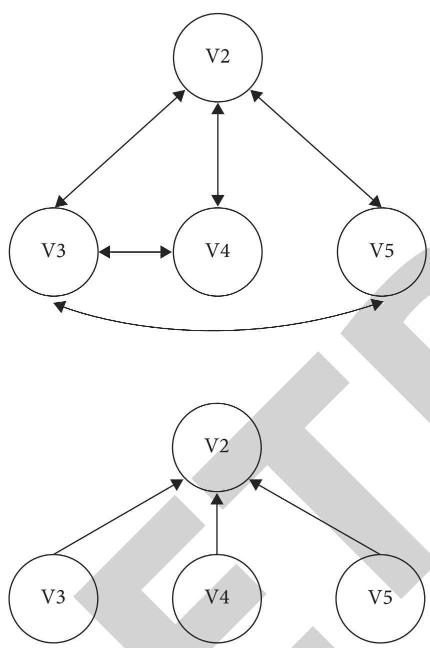


Figure 6: Matlab simulation and its directed hypergraph.

of a city, "large-scale special research on contemporary universities in the Internet environment." The purpose of this project is to understand whether the current Internet environment impacts contemporary college students' life, learning, ideology, especially their outlooks on life, world outlook, and values, and strive to determine the major influencing factors that affect young students, to provide decision support for constructing network culture in colleges and universities [21]. The data consisted of 63 questions and 30143 valid records. The specific questionnaires include the following: A1~A7 are basic personal information, T1~T23 involve college students' habits of surfing the Internet (online time, online purpose, habits of going to social networking sites, views on hot online events, etc.), and T24~T29 are college students' political attitude and learning attitude. We have preprocessed this data, established a three-dimensional matrix model, and done a lot of data analysis. Some examples of data analysis are listed below.

4.2.1. Basic Statistical Analysis (Sample). T1: how long do you spend online every day?

As shown in Figure 7, there are 29933 cases of effective data in this part, and 0.7% of this item is missing. For the problem of young students' online time, 3608 students (11.97%) spend less than 1 hour online every day, 8986 students (29.81%) spend between 1-2 hours online every day, 11735 students (38.93%) spend between 2–4 hours online every day, 4370 students (14.5%) spend between 4–8 hours online every day, and 1234 students (4.09%) spend more than 8 hours online every day.

According to the test statistics, $\chi^2 = 12177.003$, $p = 0 < 0.05$, reaching a significant level, indicating that there is a significant difference in the number of times the five options in "daily online time" are checked by the sample.

4.2.2. Crosstab Analysis (Example). This case studies whether there is a significant difference in the percentage of choices the five options in "daily online time" among young students of different genders. The statistical results are shown in Figure 8. According to the Chi-square test statistics, the Pearson Chi-square value is 2.368, the degree of freedom is 3, and the significance probability value $p = 0 < 0.05$, reached a significant level, indicating that there is a significant difference between the percentages of at least one choice time of young students of different genders in the five T1 options.

4.2.3. Logistic Regression Analysis (Sample). In this case, t26.1 "contemporary young students should take realizing the great rejuvenation of the Chinese nation as their own responsibility" as the dependent variable, A1, A4, A5, A6, A7, and T4 "browsing the content of interest online" as the independent variables for logistic regression analysis.

First, we test the likelihood ratio of each independent variable. If $p = 0 < 0.05$, it means that the independent variable has statistical significance for the corresponding variable. From Table 4, we can see that A7 is "whether it is the only child," and the p value of D option "science and technology trends" and I option "job hunting and employment" in T4 is greater than 0.05, indicating that A7, T4.D science and technology trends, and T4.I job hunting and employment have no influence on whether young students agree to "realize the great rejuvenation of the Chinese nation as their own responsibility."

4.2.4. Association Rule Analysis (Example). In this example, the basic information (A1~A7) and T1~T23 are used as the antecedents of association rules, and T24~T29 are used as the antecedents of association rules for association rule analysis. We set the support of association rules to 70% and the confidence to 90%, expecting to get high-quality rules and reduce the number of rules. Some rules we obtained are shown in Table 5.

We take the basic information in A1–A7 and T1–T29 in the questionnaire as the antecedents of association rules and T1–T29 as the antecedents of association rules for

Table 3: Results of removing redundant rules.

	Balloons	Shuttle landing control
Total association rules	18	15
Subordinate rule	8	9
Duplicate path rule	7	2
Number of redundant rules removed	15	11
Number of remaining association rules	3	4

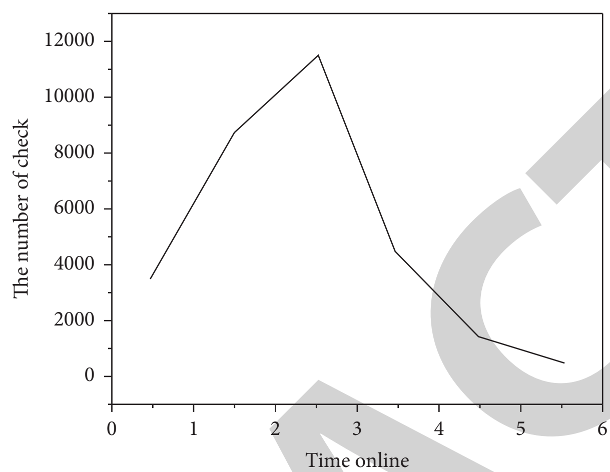


Figure 7: Sample of basic statistical analysis.

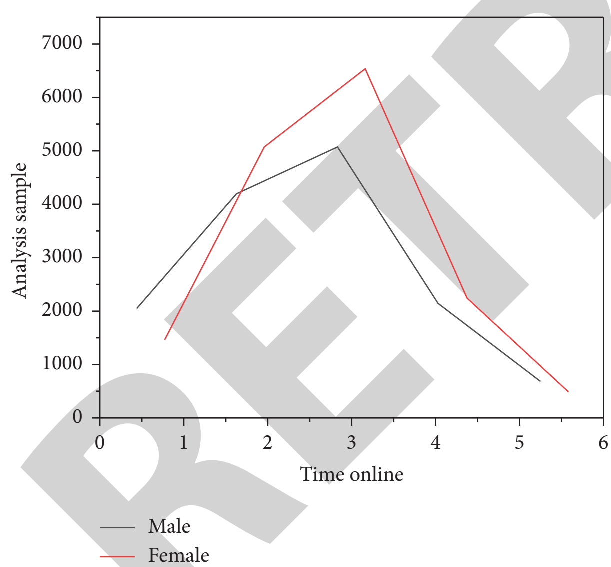


Figure 8: Example of Crosstab analysis.

association rule analysis. In order to get higher quality rules and reduce the number of redundant and loop rules as much as possible, because there are many rules, we set the support and confidence as high as possible. We set the support at 75% and the confidence at 85%. Taking some association rules in the results as examples, the specific explanations are as follows:

(1) Young students who agree that “China must develop a low-carbon, green economy and take the path of sustainable development” and agree that “honesty, trustworthiness and doing what one says is the bottom line that everyone should abide by” the main topic of chatting with people online does not love “(support 78.436% and confidence 89.599%). (2) Young students who agree that “filial piety to parents and respect for teachers are natural” and agree that “honesty and trustworthiness and doing what one says are the bottom lines that everyone should abide by” will also agree with the view that “tolerance is a virtue” (82.258% support and 97.1% confidence).

Through the interpretation of the mined association rules, we can find the important relationships and laws contained therein to provide a practical basis for enriching the theoretical research of young students’ education and promoting the cultural education of young students.

It can be seen from Figure 9 and Table 6 that although the running time of the coevolutionary algorithm is slightly longer than that of the other two global optimization algorithms, the running time is completely within an acceptable range. Moreover, due to the effective introduction of the idea of coevolution, compared with the use of the other two algorithms for association rule mining, it not only has better mining quality but also has a significant advantage in the ability to jump out of the local optimal solution, realizing the search of high-quality association rules in high-dimensional datasets.

After obtaining the association rules, we use the algorithm introduced above for redundancy processing. The experimental results are shown in Table 7.

As shown in Table 7, it can be seen that the use of hypergraph-based redundancy and loop detection methods can reduce the release of redundancy and loop policy,

Table 4: Example of logistic regression analysis.

Influence factor	p value
A1 gender	0
A4 nationality	0
A5 political outlook	0
A6 home location	0
A7 is it an only child	0.628
T4.A current affairs and politics	0
T4.B economic finance	0
T4.C education and training	0
T4.D technology trends	0.146
T4.E game entertainment	0.006
T4.F sports events	0.002
T4.G emotional marriage	0.007
T4.H culture and art	0.016
T4.I job hunting and employment	0.171

Table 5: Example of association rule analysis.

Consequent	The aforesaid	Support (%)	Confidence (%)
T28.4 = agree	T20 talk about news or interesting things that happen around you = *Y*	70.899	91.371
T28.3 = agree	T20 talk about news or interesting things that happen around you = *Y*	70.899	91.353
T28.3 = agree	T11 classmates and friends around = *Y*	70.258	90.452
T28.4 = agree	T11 classmates and friends around = *Y*	70.258	90.339
T28.3 = agree	T15 seek visual stimulation = *N* and T23 pretend to be someone else and make trouble = *N* and nationality = han nationality	71.721	90.222
T28.4 = agree	T15 seek visual stimulation = *N* and T23 pretend to be someone else and make trouble = *N* and nationality = han nationality	71.721	90.148
T28.3 = agree	T15 seek visual stimulation = *N* and nationality = han nationality and T12 I have experienced it, but I will never believe it again = *N*	75.948	90.076
T28.4 = agree	T23 pretend to be someone else and make trouble = *N* and nationality = han nationality and T12 I have experienced it, but I will never believe it again = *N*	77.142	90.074
T28.3 = agree	T23 pretend to be someone else and make trouble = *N* and nationality = han nationality and T12 I have experienced it, but I will never believe it again = *N*	77.142	90.049
T28.4 = agree	Political outlook = masses and T12 I have experienced it, but I will never believe it again = *N* and T12 I hope online love can come to me = *N*	70.235	90.019
T28.4 = agree	T15 seek visual stimulation = *N* and nationality = han nationality and T12 I have experienced it, but I will never believe it again = *N*	75.948	90.01

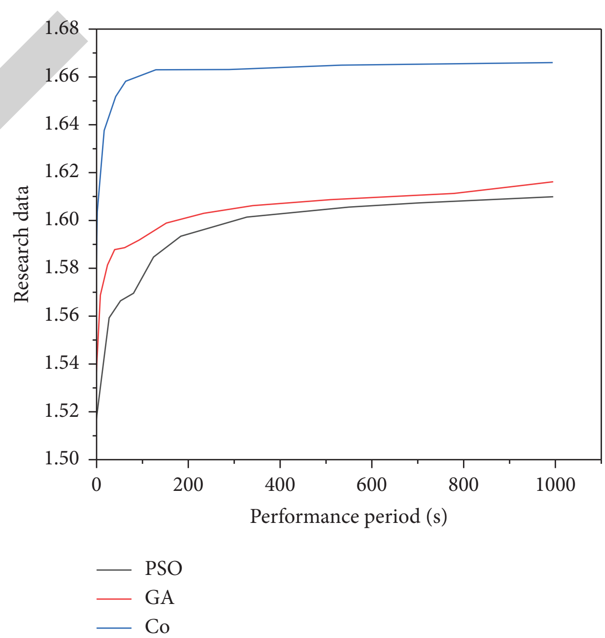


Figure 9: Comparison diagram of Co, GA, and PSO applied to research database.

Table 6: Running time of Co, GA, and PSO applied to the research database.

Algorithm	Running time (s)
Co	5946.13
GA	4601.43
PSO	3153.56

Table 7: Test results of dataset.

Dataset	Number of rules to be detected	Number of redundant rules	Number of rules after processing
Set A	23	6	17
Set B	7	1	6
Set C	98	34	64
Set D	9	2	7

without wasting the capital mining, and facilitate user selection and implementation.

5. Conclusion

In the context of big data, the task of exploring the organization's policy redundancy and improving the efficiency and effectiveness of the organization's policy mining is more pressing, and it has been increased to become the gold research and a key tool in the organization of the mining law industry. In this study, a wood spanning partition method is adopted to eliminate the redundancy of organizational policies. By redefining the adjacency matrix and its directed hypergraph, it uses the adjacency matrix to reflect the relationship between the association rule items to be detected and uses the adjacency matrix to find the connected blocks. By running the program, it obtains the number of connected blocks contained in the directed hypergraph and the location of the connected blocks where each point is located. On this basis, the adjacency matrix of the spanning tree is obtained using the ring breaking method of the spanning tree of the connected graph, and a certain weight is given to avoid the possibility of deleting important rules, to obtain the minimum spanning tree. In this way, we can remove redundant rules by checking dependent rules and repeated path rules.

Data Availability

The labeled dataset used to support the findings of this study can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by Project of Jilin Provincial Department of Science and Technology (20180201129GX).

References

[1] Y. Wen, J. Liu, W. Dou, X. Xu, B. Cao, and J. Chen, "Scheduling workflows with privacy protection constraints for big data applications on cloud," *Future Generation Computer Systems*, vol. 108, pp. 1084–1091, 2020.
[2] H. Xiao, S. Ali, Z. Zhang, M. S. Sarfraz, F. Zhang, and M. Faisal, "Big data, extracting insights, comprehension, and analytics in cardiology: an overview," *Journal of Healthcare Engineering*, vol. 2021, no. 7541, pp. 1–14, 2021.
[3] X. Lv and M. Li, "Application and research of the intelligent management system based on internet of things technology in the era of big data," *Mobile Information Systems*, vol. 2021, no. 16, pp. 1–6, 2021.
[4] B. Khan, R. Naseem, M. A. Shah et al., "Software defect prediction for healthcare big data: an empirical evaluation of machine learning techniques," *Journal of Healthcare Engineering*, vol. 2021, no. 2, pp. 1–16, 2021.
[5] R. Skuratovskii, A. Bazarna, and E. Osadhyy, "Analysis of speech mel scale and its classification as big data by parameterized knn," *Artificial Intelligence*, vol. 26, no. 1, pp. 42–57, 2021.
[6] L. Baert, E. Chérière, C. Sainvitu, I. Lepot, A. Nouvellon, and V. Leonardon, "Aerodynamic optimization of the low-pressure turbine module: exploiting surrogate models in a high-dimensional design space," *Journal of Turbomachinery*, vol. 142, no. 3, pp. 031005–031005.11, 2020.
[7] S. S. Shekhawat, H. Sharma, S. Kumar, A. Nayyar, and B. Qureshi, "Bssa: binary salp swarm algorithm with hybrid data transformation for feature selection," *IEEE Access*, vol. 9, no. 99, pp. 14867–14882, 2021.
[8] D. A. Elmanakhly, M. M. Saleh, and E. A. Rashed, "An improved equilibrium optimizer algorithm for features selection: methods and analysis," *IEEE Access*, vol. 9, pp. 120309–120327, 2021.
[9] S. Liu and K. Zhang, "Under-sampling and feature selection algorithms for s2smlp," *IEEE Access*, vol. 8, pp. 191803–191814, 2020.
[10] Y. Xiao, X. Li, B. Liu et al., "Multi-view support vector ordinal regression with data uncertainty," *Information Sciences*, vol. 589, pp. 516–530, 2022.
[11] M. Lin, X. Chen, Z. Wang, D. Wang, and J. L. Zhang, "Global profiling and identification of bile acids by multi-dimensional data mining to reveal a way of eliminating abnormal bile acids," *Analytica Chimica Acta*, vol. 1132, pp. 74–82, 2020.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 3102768, 8 pages
https://doi.org/10.1155/2022/3102768

Research Article

Financial Management Performance Evaluation of Institutions on the Basis of Intuitive Fuzzy Information

Qingqing Jiang[1] and Zhuming Zhao[2]

[1]*Haidu College of Qingdao Agricultural University, Yantai 265200, China*
[2]*Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao 266071, Shandong, China*

Correspondence should be addressed to Zhuming Zhao; zhaozm@ysfri.ac.cn

Received 22 August 2022; Revised 17 September 2022; Accepted 24 September 2022; Published 9 October 2022

Academic Editor: Xuefeng Zhang

As government reform continues to progress, institutions are also involved in this wave of reform. To be specific, there have been changes in both the structure and scope of operations of institutions. The continuous improvement of the socialist market system has further stimulated the reform of financial management. Nowadays, unlike enterprises, the financial management system of institutions is backward and the subjective nature of financial management is poor, which is far from meeting the needs of reform and development. As a result, in order to enhance the efficiency of the use of funds in institutions, financial management should adapt to the needs of reform and development and follow the trend of reform. In addition, the financial management system should adapt to the development of the times and the concept of financial management should keep pace with the times. Furthermore, with the reform of China's political system and the rapid development of the market economy, there are more and more uncertainties in the management of the daily affairs of the grassroots institutions, which brings more and more pressure on the management of local financial affairs. As a result, in order to improve the efficiency of grassroots financial management and effectively adjust the financial management activities of government and departments at all levels, the transition from manual income and expenditure management to information technology has become an inevitable trend. To perform public duties, the important material basis of business units is fixed assets. In recent years, with the increasing proportion of fixed assets in the total assets of the society, the problems of fixed asset management in institutions have become more and more obvious, such as focusing on the management of budgetary funds but neglecting the management of fixed assets, irregular basic work, disconnection between business and financial management, unbalanced allocation, low efficiency in use, and low activation rate of fixed assets, as well as even loss of fixed assets. Financial management performance evaluation has achieved some research and practical results. As an extended field, performance evaluation of state-owned asset management is gradually being emphasized by theoretical and practical circles. As a public service function of institutions, it is necessary to have a material basis to perform. Performance evaluation of state-owned asset management in institutions can improve the level of asset management, achieve higher management performance, and ultimately realize the purpose of providing high-quality public services needed by society. As a result, this study takes this as the starting point to establish a performance evaluation model around the financial performance management funds of institutions.

1. Introduction

The nature of institutions is not for profit, but to serve the society, and the scope of services covers many aspects, such as science, education, culture, and health [1]. As a result, in China, the main social functions and institutions are business units. As shown in Figure 1, China's institutions are divided into administrative institutions and operational institutions. However, nowadays, most of the institutions are still funded by the state [2]. Therefore, the attributes of the institutions are owned by the state, and the institutions themselves are less involved in business activities. In other words, this nature somehow determines that the financial management of the institutions will be based on budget rather than operating income [3]. In the actual survey, it can be found that institutions are often easy to ignore the

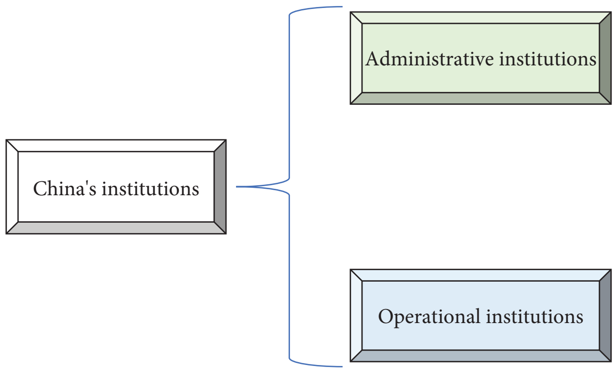


Figure 1: Component of China's institutions.

economic benefits of the institution [4]. In other words, the financial management work is not rigorous or the management activities are too one-sided, with accounting as the main focus and ignoring overall management [5]. At the same time, unit leaders do not pay much attention to financial work and focus on fulfilling social responsibilities. In fact, this is the main reason why the state spends a lot of money, but the institution feels the lack of funds and cannot obtain social benefits well [6].

In recent years, due to the influence of national policies and changes in relevant government departments, institutions have begun to pay attention to internal management and reform their systems [7, 8]. In this context, China's institutions have started to go out into the society and participate in the market competition and are no longer in a noncompetitive state under the protection of the government. Changing the original management concept and establishing a better management system, as well as adapting to the new situation of financial management will be the key issues to be solved by domestic institutions for a long time [9]. Nowadays, the main target of spending money on the activities of institutions is the society. As a result, they can fully perform their role as state institutions [10]. In fact, the purpose of financial management activities is to better integrate budgeting with actual asset management [11]. For example, in the area of public procurement, it is necessary to strictly control unnecessary purchases, which can speed up the operation of funds and improve their utilization. At the same time, risk awareness can be strengthened and a sound performance management system can be developed according to the actual situation of the unit. For the purpose of financial management, Chinese scholars have also put forward many feasible suggestions [12]. On the one hand, it is necessary to strengthen the accounting team. On the other hand, it is quite necessary to increase internal audits [13]. As a result, in order to improve the efficiency of the use of state funds in social welfare activities and to improve business performance, institutions should first improve their financial management theories [14]. Furthermore, they should apply advanced management concepts in order to social welfare activities as well as improve their practical skills [15].

With the political reform of our society and the rapid development of the market economy, there are more and more uncertainties in the management of the daily affairs of the grassroots institutions. At the same time, this has put more and more pressure on the management of local financial affairs [16]. As a result, in order to improve the efficiency of grassroots financial management and effectively adjust the financial management activities of the government and departments at all levels, the transition from manual income and expenditure management to information technology has become an inevitable trend [17]. However, traditional financial management relies heavily on manual processes. While recording budgets and accounts, accounting and auditing staff at all levels are needed to maintain an ever-increasing number of paper tickets [18]. The drawbacks are obvious: the workload is heavy, the preparation time is long, and the content is single and not comprehensive. From the perspective of the finance department alone, the budget must be consolidated and organized by the finance department after it has been submitted by all departments concerned [19]. Nevertheless, the budgeting methods used vary from department to department. In fact, coupled with the lagging nature of the work at the grassroots level, this can add to the uncertainty of the workload as finance is burdened with increasingly heavy adjustments as well as the need to constantly coordinate budget submissions across departments [20].

In fact, our institutions mainly provide social services to the public. By performing social functions, they can ensure that the needs of public utilities such as education, culture, and health are effectively met [21]. Therefore, the development of institutions is closely related to the management of state-owned assets. At the same time, the development of such units cannot be achieved without state-owned assets [22]. Looking at the current situation of state-owned asset management, we can see that there are still some problems, such as an inadequate system, unclear evaluation subjects and objects, imperfect evaluation indicators, and insufficient application of evaluation results [23]. As a result, how to strengthen the management of state-owned assets, evaluate the efficiency and effectiveness of the use of state-owned assets, and ensure that institutions can fulfill their social responsibilities more efficiently is the main research issue of this study [24]. In view of the shortcomings revealed in the management of state-owned assets, the application of a performance evaluation system for state-owned asset management to institutions can effectively solve these problems. It can promote asset management in institutions and improve the efficiency of using state-owned assets [25]. In addition, it can realize the combination of budget management and asset management, which is also a necessary way to deepen the government accounting reform.

Institutions cannot operate and develop without the support of state-owned assets. After all, only with sufficient state-owned assets can institutions fulfill their responsibilities, provide various public services to the public and society, meet the development needs of education, health, and culture, and ensure the smooth promotion of various public programs [26]. Strict management and control of state-owned assets can ensure that financial distribution meets the requirements of fairness and justice and that social order can be effectively maintained. However, at present, there are many problems in the management of state-owned assets in China [27]. The reasons for these problems are that, on the one hand, the units themselves do not strictly follow the requirements of the new system and, on the other hand, it is expected that the problems of state-owned asset management will identify the imperfections of the current management system in China and propose some strategies and methods to solve them. In addition, there are only a few studies on the performance evaluation of state-owned asset management in public institutions, and most of the studies focus on the overall performance of administrative institutions, so it is expected that the study will add to the existing literature [28].

From the perspective of practical work, the establishment of a financial performance management evaluation system is a guideline for practical work [29]. Some problems were faced in practice, such as focusing on the management of budgetary funds but neglecting the management of fixed assets, high duplication rate of purchase, irregular basic work, inadequate specific systems, low efficiency of use, and low rate of asset revitalization, etc., which limit the development of fixed asset management in institutions. In addition, the new government accounting system has put forward new requirements for the classification, accounting, measurement, approval, and reporting of fixed assets. Faced with these problems and requirements, it is quite necessary to establish scientific and reasonable performance evaluation indexes from the perspective of performance evaluation, which can provide comprehensive feedback data [30]. Furthermore, it can make the fixed assets use the maximum efficiency, thus improving the level of fixed asset management in institutions and promoting the great development of fixed asset management in institutions.

Finance is a crucial department of a unit, and the normal operation of the unit cannot be achieved without the management of finance. In essence, financial management is value management. Initially, financial management was created with business activities, including financial activities and financial relations, and is a financial relations approach. The social attributes of business units, the industries they belong to, and the specific social environment are different, which directly determines the different ways of operating funds, but the final expression is the same, namely, financial management. Therefore, financial management is internal to the unit. It is also the basis and core of business unit management and is the pillar of business unit management. It is essential to study the current situation of financial management in institutions. To be specific, it is relatively important to analyze the current financial situations of public organizations at home and abroad and to learn from advanced management models and management concepts. After all, the introduction of positive management models as well as concepts should consider the national conditions of China. In other words, the financial management model needs to be adapted to the nature of local organizations in China. Ultimately, this can significantly improve the efficiency of capital utilization and business performance of public welfare organizations. At the same time, the reform of the financial management of public institutions can be a catalyst and a reference for the reform of other public organizations. This can promote the development of financial theory, which has great practical significance for individuals as well as institutions themselves.

2. Intuitive Fuzzy Information

2.1. Expression Form of Fuzzy Evaluation Information. In the process of comprehensive multi-indicator evaluation with complex fuzzy evaluation information, the evaluator should give the evaluation information value based on his own knowledge and experience and the objective situation of the evaluation object, no matter if it is a single person evaluation or group evaluation. In fact, the evaluation information is a complex fuzzy set based on intuitionistic fuzzy sets and hesitant fuzzy sets. With the introduction of intuitionistic fuzzy sets, Pythagorean fuzzy sets, and hesitant fuzzy sets, scholars have been expanding the existing fuzzy sets. As a result, more and more complex fuzzy evaluation information expressions are proposed in terms of the composition of affiliation and nonaffiliation functions and the composition elements of fuzzy sets, which are suitable for modern comprehensive evaluation processes. In addition, the expressions of subordinate and non-subordinate functions in terms of the interval, triangular, and trapezoidal numbers can reflect the fuzzy degree in the process of multi-indicator evaluation in a more detailed way. This makes it easier for the evaluator to provide comprehensive evaluation information on alternative solutions. As a result, the correlation between the indicators in the comprehensive evaluation process should not be neglected, and the existing solutions are mainly based on the Bonferroni mean as well as the Choquet integral. The process of the comprehensive evaluation method based on intuitive fuzzy information is shown in Figure 2.

Hamacher operations are a good alternative to algebraic products and algebraic sums. However, most of the existing fuzzy aggregation operators in fuzzy environments are based on algebraic operations. In this study, we use the Hamacher operator to aggregate probabilistic interval-valued hesitant Pythagorean fuzzy evaluation information and propose the probabilistic interval-valued hesitant Pythagorean fuzzy Hamacher operator rule to make the evaluation information aggregation more flexible. A new aggregation operator, the probabilistic interval-valued Pythagorean hesitant fuzzy Hamacher Choquet integral geometric mean operator, is also proposed, and the related properties of the operator are discussed. The comprehensive evaluation process based on intuition and hesitant fuzzy information is illustrated in Figure 3.

2.2. Trapezoidal Intuitionistic Fuzzy Evaluation Information. In social life, things are changing rapidly, and it is increasingly difficult for people to accurately grasp their development direction. This makes the shortcomings of the traditional multi-indicator comprehensive evaluation method come to the fore. In fact, the complexity and uncertainty of comprehensive evaluation are becoming more

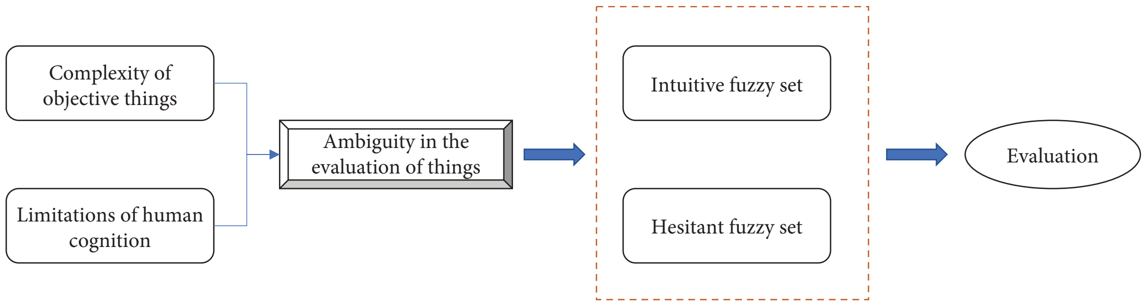


Figure 2: Evaluation method based on intuitive fuzzy information.

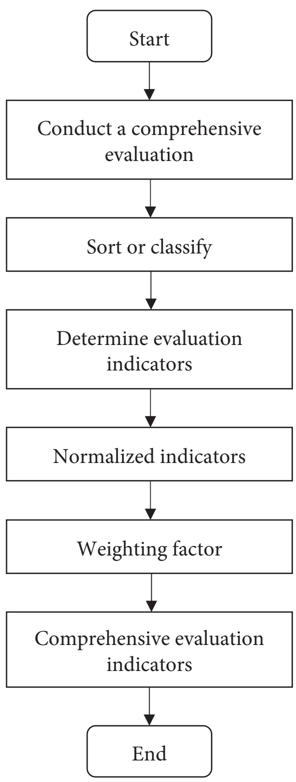


Figure 3: Comprehensive evaluation process based on intuition and hesitant fuzzy information.

and more obvious. As a result, the real-valued form of comprehensive evaluation information is replaced by fuzzy sets and intuitionistic fuzzy sets. Fuzzy comprehensive evaluation information reflects the process of the evaluator's brain and has a complete evaluation dimension of approval, disapproval, and abstention for the evaluation object.

Let $A = \{a_1, a_2, \ldots, a_n\}$ be a nonempty set; then, the fuzzy set is defined as follows:

$$F = \{\langle a, \mu_F(a)\rangle \mid a \in A\}, \quad (1)$$

where $\mu_F(a)$ refers to the affiliation of the element a in A belonging to F.

In addition to approval and disapproval, the intuitionistic fuzzy set can better fit the psychological activities of the evaluator. To be specific, it can express neither a favorable nor an unfavorable attitude toward the evaluation object according to its own perception, which ensures a more scientific and comprehensive evaluation process. However, with the rapid development of the social economy, the subordinate degree function of the intuitionistic fuzzy set with exact value and the non-subordinate degree function have large errors, which cannot well represent the complexity as well as the uncertainty of the comprehensive evaluation process of multiple indicators. As a result, the interval value contains more evaluation information than the exact value.

Obviously, the domains of both intuitionistic fuzzy sets and interval intuitionistic fuzzy sets are discrete sets. Its limitation is that it can only broadly represent the subordination and non-subordination functions of fuzzy sets. Therefore, it cannot express fuzziness and hesitancy at the same time. As a generalization of interval intuitionistic fuzzy sets, trapezoidal intuitionistic fuzzy sets can better portray the complexity of the comprehensive evaluation process. The trapezoidal intuitionistic fuzzy set is defined as follows:

$$\mu_{\tilde{z}}(a) = \frac{(a - z)\omega_{\tilde{z}}}{y - z}, \quad (2)$$

where $\omega_{\tilde{z}}$ represents the maximum degree of affiliation.

In addition to this, the trapezoidal intuitionistic fuzzy number can be represented in Figure 4.

3. Financial Management Performance Evaluation of Institutions

3.1. Demand Analysis. The main functions of the system can be basically determined through the demand research on the current situation of financial management of institutions. Figure 5 depicts the main use case of the system. To be specific, the whole use case consists of login management, user management, payroll management, budget management, account management, announcement management, and other modules.

The main purpose of login management is to complete the user login. The system is designed in such a way that only users authorized by the super user have the right to log in and view the system records of the corresponding department. In addition, each user can modify their personal information after logging in. Users can only perform other business operations if they have successfully logged in. The specific login management use case is shown in Figure 6.

Among them, the use case description of user login through the web side is shown in Table 1.

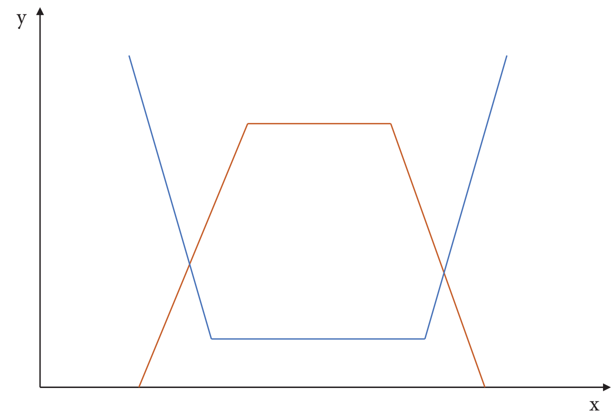


Figure 4: Trapezoidal intuitionistic fuzzy number.

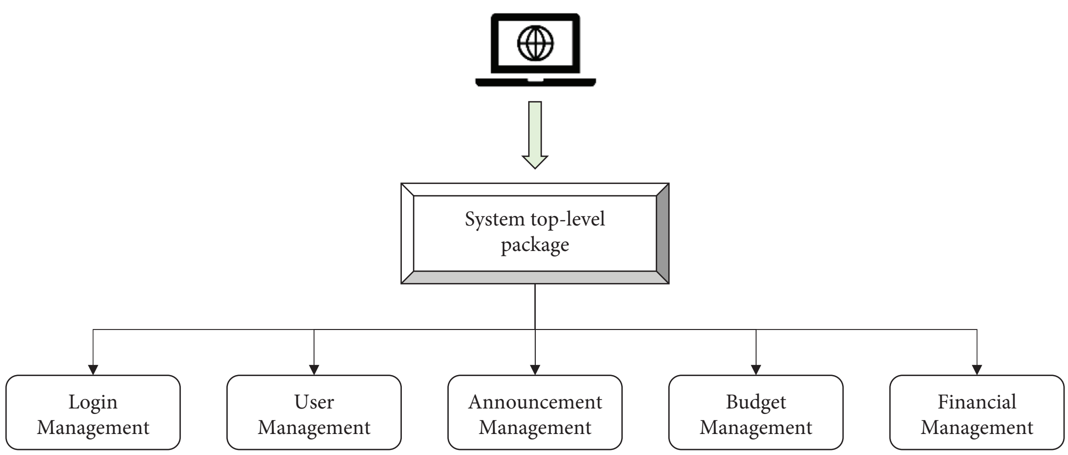


Figure 5: Main use case of the system.

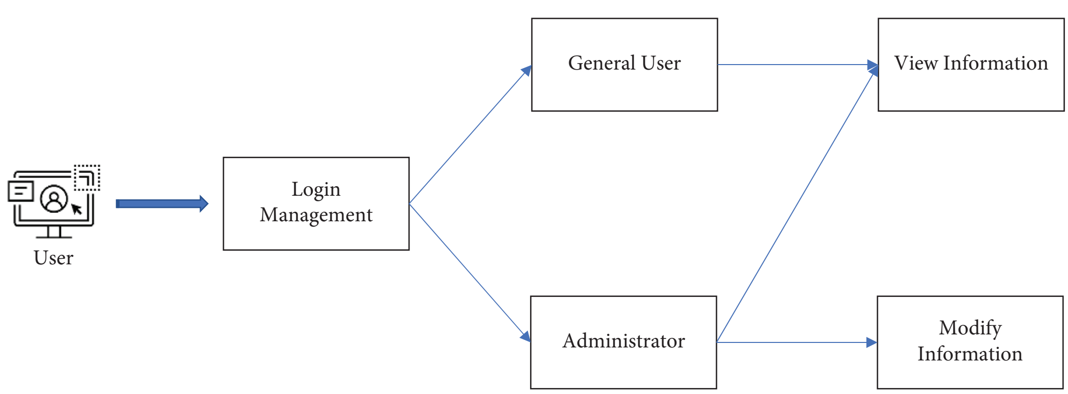


Figure 6: Specific login management use case.

User management is mainly used to maintain the information of the system users. The system is designed to allow employees from different departments to log in to the system with certain rights. The management of users can be divided into functions such as adding, modifying, querying, and deleting. In particular, users can be authorized to operate during the process of adding and modifying users. The example of user management is shown in Figure 7.

3.2. Current Status of Institution's Asset Structure. Some institutions are engaged in public services, and their assets are nonoperating assets. In addition, some institutions have been restructured to operate according to the laws of the market, and their assets have been converted to operating assets. As our country grows stronger and stronger, the volume of institutions has also grown rapidly. The number of state-owned assets held by business units has also

Table 1: User login system description.

Objective	Access to the system
Participants	Administrators and regular users
Prerequisites	User is already registered
Postcondition	Successfully enter the main interface

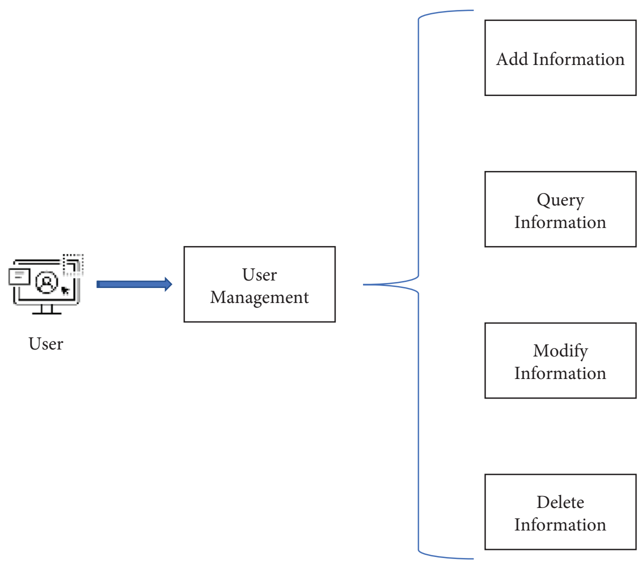


Figure 7: Example of user management.

continued to increase. These assets include real estate and buildings, as well as office equipment, vehicles, and instruments. Table 2 shows the asset structure of China's institutions from 2016 to 2021.

The management of state-owned assets by business units is included in the financial management system. Therefore, the reform of the relevant system is covered by the reform of the financial system. At present, the hierarchical management of state-owned assets in institutions is also determined by the current financial management system. Since the establishment of the State-owned Assets Management Department in 2004, the state-owned assets of institutions have been managed in a unified manner. At the same time, governments at all levels have also established special management departments. After more than a decade of development, the management system has achieved remarkable results. In addition, there are also documents that clarify that state-owned assets of institutions in China have been unified under state ownership and that institutions are managed by the government at all levels, while state-owned assets are occupied and used by each institution.

In fact, the business unit adds assets through procurement or redeployment. However, the initial source of funds must be the budget. The process of fund movement is to have a budget first and then financial expenditure and finally become state-owned assets. Therefore, the core aspects of financial management include budget management, financial expenditure management, and management of state-owned assets of institutions. As a result, the main contents of the current budget performance management system in China are formed.

Table 2: Asset structure of China's institutions from 2016 to 2021 (trillion yuan).

Year	Total asset	Total current asset	Net asset
2016	13.32	4.23	10.97
2017	14.79	4.56	9.45
2018	15.85	4.97	8.89
2019	16.78	5.03	8.27
2020	17.74	5.21	7.46
2021	18.85	5.87	6.75

Although China's system for evaluating the performance of state-owned assets of institutions has clarified some requirements and evaluation contents, the system as a whole is not standardized enough, which is also the most prominent problem at present. On the one hand, there are no special laws and regulations for the current performance evaluation activities, and there is no unified legal regulation for the specific operation of each state-owned unit. In addition, there are no unified operation rules, so it is difficult for each unit to standardize its behavior when carrying out work, and the operation methods adopted by each unit vary greatly. Although the existing Interim Measures for the Management of State-owned Assets of Public Institutions have set forth requirements for performance management, the system support and implementation measures are not reflected in the document. Among the existing relevant documents, a few of them specify the design direction of the relevant indicators. Therefore, in general, there is a gap in the laws and regulations on the performance evaluation of state-owned assets of public institutions in China. In other words, the existing supporting documents only contain individual provisions of some institutional documents, but there is no set of scientifically complete and operationally tested institutional systems.

3.3. Financial Performance Evaluation Index System. In order to revitalize state-owned assets and promote the development of the local economy, Chinese institutions have chosen a market-based model for managing state-owned assets to enhance the development of institutions. In the management process, performance evaluation tools should be introduced to assess the management of state-owned assets and guide institutions to adopt standardized management behaviors. In addition, government finance departments at all levels need to be involved in the management process. The design of the indicator system should be based on the principles of comprehensiveness and systematization. The successful implementation of this principle can complement the limitations of the system. Secondly, it is important to follow the principles of scientific, which can make the performance evaluation more accurate and scientific to meet the management needs. Only in this way can the indicator system ensure accurate results and lead to an overall improvement in asset management. Then, the principle of both qualitative and quantitative analysis should be used, and quantitative analysis of indicators should be conducted when selecting evaluation indicators to make them easier to operate.

For factors that cannot be analyzed quantitatively but are mandatory to evaluate the effect, a comprehensive qualitative analysis is also required. Based on the principle of quantitative calculation and qualitative requirements, it is necessary to make performance evaluation completer and more systematic. Finally, the principle of the hierarchy should be followed. This is a requirement that the unit should be realistically adjusted to meet the specific needs of the unit based on the purpose of the performance evaluation.

4. Conclusion

Common approaches to managing state assets include intermediate management, macromanagement, and micromanagement. In practice, a combination of these approaches is required. In addition, it is necessary to strengthen the internal supervision of the institutions. The finance department should carefully review all the economic operations of the institutions. The asset management department is mainly responsible for managing maintenance and asset acquisition. At the same time, they should manage and inventory assets in a timely manner to ensure that the accounts and cards are consistent. The audit department is responsible for auditing the use of state-owned assets in business units, rectifying, analyzing, and giving feedback on the problems found in a timely manner, and holding the responsible persons responsible for serious problems accountable. At the same time, external supervision should be strengthened.

The performance management of state-owned assets in administrative institutions is a complete framework covering goal setting, performance mechanism establishment, performance evaluation, and an improvement plan proposal. The fundamental purpose of performance evaluation is to promote the improvement of the quality, level, and efficiency of state-owned asset management in public institutions. The fundamental purpose of performance evaluation is to promote the improvement of the quality, level, and efficiency of state-owned asset management in public institutions, and the subsequent summary and improvement plan are more important than performance evaluation.

However, the research content of this study is preliminary, and somewhat objective, and one-sided. Therefore, there is a need to pay attention to the theoretical guidance measures related to the reform of the classification of institutions in the future. Only by applying financial management norms to the internal responsibilities of units can we promote the normal, perfect, and smooth reform transition of operating institutions from the source.

Data Availability

The labeled dataset used to support the findings of this study can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by Special project of "traditional culture and economic and social development" in Shandong Province, project name: Research on landscape image construction of traditional villages in Shandong Province Based on Cultural Representation Theory, subject no. zc202011219, topic ranking: 3/5.

References

[1] D. A. Webber, "State divestment and tuition at public institutions," *Economics of Education Review*, vol. 60, pp. 1–4, 2017.

[2] M. Bahamón, P. Cruz, and D. Campillo, "Curricular integration as a strategy to strengthen the educational process in public institutions in COVID-19 times," *Revista Inclusiones*, pp. 233–241, 2021.

[3] A. Ferede, G. D. Kibret, Y. Million, M. M. Simeneh, Y. A. Belay, and D. Hailemariam, "Magnitude of turnover intention and associated factors among health professionals working in public health institutions of North Shoa Zone, Amhara Region, Ethiopia," *BioMed Research International*, vol. 2018, Article ID 3165379, 9 pages, 2018.

[4] R. Falcone, E. Colì, S. Felletti, A. Sapienza, C. Castelfranchi, and F. Paglieri, "All we need is trust: how the COVID-19 outbreak reconfigured trust in Italian public institutions," *Frontiers in Psychology*, vol. 11, Article ID 561747, 2020.

[5] G. M. Mbouthieu Teumta, L. L. Niba, N. T. Ncheuveu et al., "An institution-based assessment of students' hand washing behavior," *BioMed Research International*, vol. 2019, Article ID 7178645, 7 pages, 2019.

[6] N. A. Khan, Q. Gao, and M. Abid, "Public institutions' capacities regarding climate change adaptation and risk management support in agriculture: the case of Punjab Province, Pakistan," *Scientific Reports*, vol. 10, no. 1, Article ID 14111, 2020.

[7] M. T. Coe, K. M. Evans, K. Gasic, and D. Main, "Plant breeding capacity in US public institutions," *Crop Science*, vol. 60, no. 5, pp. 2373–2385, 2020.

[8] A. M. Ward and J. Forker, "Financial management effectiveness and board gender diversity in member-governed, community financial institutions," *Journal of Business Ethics*, vol. 141, no. 2, pp. 351–366, 2017.

[9] M. Zhai, "Risk prediction and response strategies in corporate financial management based on optimized BP neural network," *Complexity*, vol. 2021, Article ID 9973377, 10 pages, 2021.

[10] D. Purwita and Y. Murdayanti, "Measuring financial management in islamic educational institutions in Indonesia," *Journal of Chemical Information and Modeling*, vol. 53, no. 9, pp. 1689–1699, 2019.

[11] G. Matei and I. M. Alexandru, "Financial management of public institutions," *Finance: Challenges of the Future*, vol. 18, no. 20, pp. 34–44, 2018.

[12] X. Liu, "Accounting and financial management cost accounting integrating rough set knowledge recognition algorithm," *Discrete Dynamics in Nature and Society*, vol. 2022, Article ID 9286252, 9 pages, 2022.

[13] M. W. Mabonga and E. N. Kimani, "Financial management practices and financial performance of microfinance institutions in Bungoma County, Kenya," *International Academic Journal of Economics and Finance*, vol. 2, no. 3, pp. 335–347, 2017.

[14] S. Mishchenko, S. Naumenkova, V. Mishchenko, and D. Dorofeiev, "Innovation risk management in financial institutions," *Investment Management and Financial Innovations*, vol. 18, no. 1, pp. 190–202, 2021.
[15] E. W. Gitee, H. K. Mburu, and M. A. Muhavani, "Determinants of Effective financial management in Catholic Church managed health institutions in Kenya," *African Journal of Emerging Issues*, vol. 1, no. 9, pp. 48–68, 2019.
[16] M. A. Terlizzi, A. L. Albertin, and H. RdO. C. de Moraes, "IT benefits management in financial institutions: practices and barriers," *International Journal of Project Management*, vol. 35, no. 5, pp. 763–782, 2017.
[17] A. A Shehu, P. Teru, and B. Musa, "Impact of public financial management reforms on expenditure control in federal tertiary institutions in adamawa state," *International Journal of Accounting and Finance*, vol. 9, no. 2, pp. 104–119, 2020.
[18] C. J. Godlewski and B. Sanditov, "Financial institutions network and the certification value of bank loans," *Financial Management*, vol. 47, no. 2, pp. 253–283, 2018.
[19] P. M. Wangai and J. N. Mungai, "Financial management practices and loan performance of micro finance institutions in Starehe Constituency, Nairobi County, Kenya," *International Academic Journal of Economics and Finance*, vol. 3, no. 3, pp. 117–135, 2019.
[20] S. K. Shrivastava, "Accounting practices and financial management in higher educational institutions," *International Journal of Research in Social Sciences*, vol. 8, no. 5, pp. 478–486, 2018.
[21] N. Kraus, K. Kraus, and V. Osetskyi, "New quality of financial institutions and business management," *Baltic Journal of Economic Studies*, vol. 6, no. 1, pp. 59–66, 2020.
[22] M. Wijesiri, J. Yaron, and M. Meoli, "Assessing the financial and outreach efficiency of microfinance institutions: do age and size matter?" *Journal of Multinational Financial Management*, vol. 40, pp. 63–76, 2017.
[23] X. G. Yue, Y. Cao, N. Duarte, X. F. Shao, and O. Manta, "Social and financial inclusion through nonbanking institutions: a model for rural Romania," *Journal of Risk and Financial Management*, vol. 12, no. 4, p. 166, 2019.
[24] N. N. S. Ekayani, "Implementation of good corporate governance and regulation of the performance of micro financial institutions in village," *Journal of Advanced Research in Dynamical and Control Systems*, vol. 12, no. 7, p. 1, 2020.
[25] S. Mohd, "A study on the performance of microfinance institutions in India," *International Academic Journal of Accounting and Financial Management*, vol. 5, no. 4, pp. 116–128, 2018.
[26] A. Hamidov, U. Kasymov, A. Salokhiddinov, and M. Khamidov, "How can intentionality and path dependence explain change in water-management institutions in Uzbekistan?" *International Journal of the Commons*, vol. 14, no. 1, pp. 16–29, 2020.
[27] P. M. Guest, "Risk management in financial institutions: a replication," *The Journal of Finance*, vol. 76, no. 5, pp. 2689–2707, 2021.
[28] S. Khawar and A. Sarwar, "Financial literacy and financial behavior with the mediating effect of family financial socialization in the financial institutions of Lahore, Pakistan," *Future Business Journal*, vol. 7, no. 1, pp. 27–11, 2021.
[29] S. Ibrahim, G. Diibuzie, and M. Abubakari, "The impact of internal control systems on financial performance: the case of health institutions in upper west region of Ghana," *International Journal of Academic Research in Business and Social Sciences*, vol. 7, no. 4, pp. 684–696, 2017.
[30] G. Abara and T. Banti, "Role of financial institutions in the growth of micro and small enterprises in Assosa Zone," *Research Journal of Finance and Accounting*, vol. 8, no. 3, pp. 36–40, 2017.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 4542513, 10 pages
https://doi.org/10.1155/2022/4542513

Research Article

The Application of Real-Time Video Processing Technology in English Culture Teaching

Qing Xu

Changsha University of Science and Technology, Changsha, Hunan 410000, China

Correspondence should be addressed to Qing Xu; amydarling20220922@163.com

Received 10 August 2022; Revised 13 September 2022; Accepted 22 September 2022; Published 9 October 2022

Academic Editor: Jiafu Su

In order to improve the effect of English culture teaching, this paper combines the real-time video image processing technology to construct an English culture teaching system. Moreover, this paper selects the image processing based on octree structure with the advantages of full 3D modeling and being updatable, flexible, and compact by comparison. At the same time, this paper analyzes the existing Gaussian-based inverse sensor model that can be used for monocular vision and improves the measurement inverse sensor model for stereo vision published by Andert. In addition, this paper deeply analyzes the influence of key parameters of the improved measurement inverse sensor model and realizes the real-time octree occupancy grid map of the improved measurement inverse sensor model. Finally, after improving the real-time video processing technology, this paper evaluates the application effect of the method proposed in this paper in English culture teaching. The experimental research results show that the real-time video processing technology proposed in this paper has a good effect in English culture teaching, and it can be applied to English culture teaching to improve the effect of English culture teaching.

1. Introduction

Language is not only the carrier of culture but also a part of culture. For example, many people immigrate to a new country and still retain their old habits, speaking their first language. This is very common in the US, even though they are very comfortable with the new locale. The reason for this is the desire of these migrants to preserve their cultural heritage, which includes not only their traditions and customs but also their own language. This is also seen in many Jewish communities. In particular, it is manifested in some elderly people who often speak Yiddish. The reason for this is that it is seen as an integral part of Jewish culture. Language is a special part of culture [1].

It develops with the change and development of human society. From another perspective [2], without language, culture will not exist. Language is the formation and development of culture. The formation and development of culture promotes the development of language. Culture emerges with the appearance of language. There are many similarities [3]. However, humans created language, while animals did not. Primitive humans have cultural categories, like religion, morality, customs, and so on, while animals cannot [4]. So, we say that language teaching is essentially cultural teaching. In the actual teaching process, it seems that as long as we listen and speak, with training in reading and writing, we can understand English and be able to communicate with others in English. But in fact, people often misunderstand each other because they do not understand the cultural background of the language [5].

There is an inseparable “natural” relationship between language and cultural practitioners. Language is the carrier of culture, and culture is the connotation of language [6]. Although language and culture are inseparable and cultural factors are ubiquitous in language teaching, in actual teaching, it may be influenced by traditions and habits, ignoring cultural factors in language use or exaggerating literary teaching. In foreign language teaching, cultural teaching takes over [7]. It can be seen that the relationship between cultural teaching and language teaching deserves a thorough discussion and research from theory to practice. A large number of research results and practice have proved that the relationship between them has the characteristics of synchronization, complementarity, and compatibility [8].

A key issue in the implementation of cultural introduction and cultural teaching in foreign language teaching is the quality of language teachers. The language teaching activities are mainly completed by teaching and students. In such a large environment of cultivating students' cross-cultural communication ability, teachers must change their teaching concepts in teaching and change the traditional teaching method of teachers' teaching and students' passive acceptance of teaching methods. Teachers guide and help students to collect information and consciously absorb the nutrition of Chinese and Western cultures [9]. Therefore, the traditional status of teachers, "teaching teaching and solving doubts" has also been given a new connotation. Teachers have changed from "sages at the front desk" to "guiding guides." At the same time, in order to strengthen teachers' understanding of the dual culture, the requirements for teachers have also been improved [10]. At the same time, it is also necessary to have the ability to accurately express their own culture in a foreign language. Only with these two abilities can you be competent in teaching, play an inspiring and guiding role in the teaching process, fully mobilize students' enthusiasm for learning, effectively organize classroom teaching centered on learning, discover problems, solve problems, and improve students' humanistic quality. Only with full-scale, two-way interaction can the language learning process become a process of cultural learning and communication and a live language practice process. The introduction of cross-literature, mathematics, and inclusiveness in English teaching will help students to open up their horizons, develop their ideas, and improve their comprehensive quality [11].

Intercultural communication is the process of communication and interaction between people from different cultural backgrounds. In the exchange, the two sides understand each other and express their own ideas and cultures, which is also a process of equal interaction between local culture and foreign culture. To build a bridge of cross-cultural communication, learners must have bicultural and bilingual abilities [12]. In the research of cultural teaching in English teaching in colleges and universities, many scholars have proposed the teaching mode of cross-cultural communication ability from the perspective of foreign language teaching [13]. All of these teaching modes take the cultivation of cultural awareness and the inclusiveness of different cultures as the main content of cultivating intercultural communication ability [14]. It is necessary to have a solid mother tongue culture foundation; especially in the era of foreign culture penetration, only when students have a profound cultural background in their own language can they have a clear sense of identity with their mother tongue culture and be tolerant of different cultures. Incorporating it into the existing cultural teaching of students, mother tongue culture has become an indispensable and important factor in college English teaching because college English teaching has corresponding topics or scenarios; if college English teaching cannot find the similarities and differences between the two cultures, you cannot effectively acquire cross-cultural skills. The two-way flow of language and culture requires that our language learning cannot move forward alone [15].

In order to improve the effect of English culture teaching, this paper combines the real-time video image processing technology to construct an English culture teaching system to improve the cultural communication in the English teaching process.

2. Real-Time Video Processing Algorithms

2.1. Design and Implementation of Occupancy Grid Map Generation Algorithm Based on Octree. The octree hierarchical data structure represents a three-dimensional environment. The octree data structure shown in Figure 1 consists of nodes.

By applying an ideal sensor model and update equations, sensor measurements integrate the finest resolution or leaf nodes into the teaching image. When the measurement unit is updated to be occupied, the update equation calculates the probability of leaf nodes as

$$\begin{gathered}\left(\frac{p(m_i \mid z_{1:t})}{1-p(m_i \mid z_{1:t})}\right)=\left(\frac{p(m_i \mid z_t)}{1-p(m_i \mid z_t)}\right)\left(\frac{p(m_i \mid z_{1:t-1})}{1-p(m_i \mid z_{1:t-1})}\right)\left(\frac{1-p(m_i)}{p(m_i)}\right),\\ L(m_i \mid z_{1:t})=L(m_i \mid z_t)+L(m_i \mid z_{1:\,t-1}).\end{gathered}\tag{1}$$

At the same time,

$$L(m_i)=\log\left[\frac{p(m_i)}{1-p(m_i)}\right].\tag{2}$$

Among them,

$$L(m_1)=\log(1)=0.\tag{3}$$

In the ideal case, the system computes the posterior distribution over the teaching image and constructs the pose estimate x from start to time t to get

$$p(m_i \mid z_{1:t}, x_{1:t}).\tag{4}$$

For convenience, we assume that all sensor measurements z are converted to world coordinates and that merging poses to measurements z eliminates x.

$$p(m \mid z_{1:t}).\tag{5}$$

In general, the occupancy rate can be used to calculate the probability of a cell being idle. For convenience, as in formula (6), the problem of calculating the posterior distribution on the teaching image is a dimensional problem.

$$p(m_i \mid z_{1:t}).\tag{6}$$

It is assumed that the instructional image units are statistically independent of each other. In other words, adjacent cells are not allowed to have dependencies, but the posterior teaching image is taken as the product of its edges, as shown in the following formula.

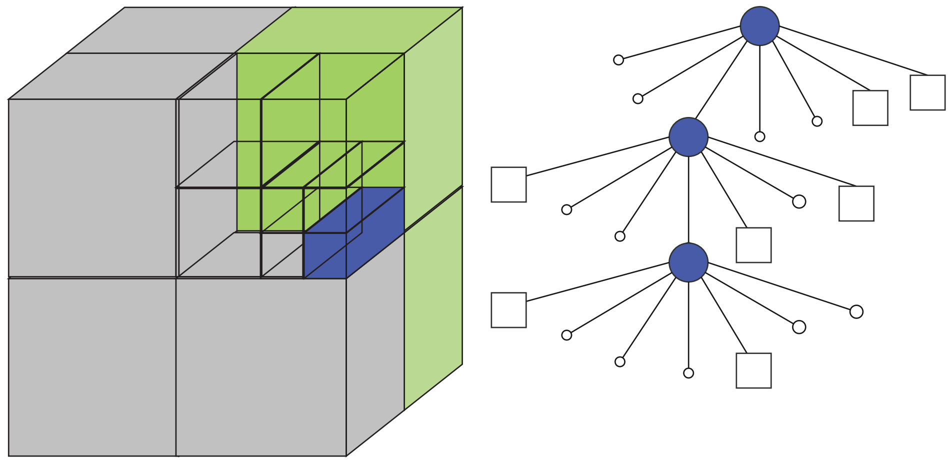


Figure 1: Visual representation of octree voxels. (a) Voxel presentation. (b) Tree presentation.

$$p(m_1, m_2 \cdots, m_N \mid z_{1:t}) = \prod_{i}^{N} p(m_i \mid z_{1:t}). \tag{7}$$

Binary Bayesian filters are well suited for this problem when considering the static environment of the binary state and the conditional independence of measurements over time for a given m_i:

$$\begin{aligned} p(m_i \mid z_{1:t}) &= \frac{p(z_t \mid m_i, z_{1:t-1}) p(m_i \mid z_{1:t-1})}{p(z_t \mid z_{1:t-1})} \\ &= \frac{p(z_t \mid m_i) p(m_i \mid z_{1:t-1})}{p(z_t \mid z_{1:t-1})}. \end{aligned} \tag{8}$$

Bayes' rule is applied to $(z_t \mid m_i)$:

$$p(z_t \mid m_i) = \frac{p(m_i \mid z_t) p(z_t) p(m_i \mid z_{1:t1})}{p(m_i) p(z_t \mid z_{1:t1})}. \tag{9}$$

Likewise, the complementary probability of a cell being idle is

$$p(\overline{m_i} \mid z_{1:t}) = \frac{p(\overline{m_i} \mid z_t) p(z_t) p(\overline{m_i} \mid z_{1:t-1})}{p(\overline{m_i}) p(z_t \mid z_{1:t-1})}. \tag{10}$$

To eliminate some tedious calculations, equation (9) can be divided by its supplementary equation (10) to get the following formula:

$$\frac{p(m_i \mid z_{1:t})}{p(\overline{m_i} \mid z_{1:t})} = \frac{p(m_i \mid z_t) p(m_i \mid z_{1:t-1}) p(\overline{m_i})}{p(\overline{m_i} \mid z_t) p(\overline{m_i} \mid z_{1:t-1}) p(m_i)}. \tag{11}$$

To avoid probabilistic numerical instabilities near 0 or 1, the log-difference ratio can be used to express the occupancy, as shown in formula (12). The probability can be easily recovered as shown in formula (13).

$$L(m_i \mid z_{1:t}) = \log\left(\frac{p(m_i \mid z_{1:t})}{1 - p(m_i \mid z_{1:t})}\right), \tag{12}$$

$$p(m_i \mid z_{1:t}) = 1 - \left(\frac{1}{1 + e^{L(m_i \mid z_{1:t})}}\right). \tag{13}$$

Using the log-difference ratio, the multiplication of probabilities is transformed into the addition of the log-difference ratio, reducing computational cost and complexity. When it is applied to formula (11), formula (14) is obtained.

$$\begin{aligned} \log\left(\frac{p(m_i \mid z_{1:t})}{p(\overline{m_i} \mid z_{1:t})}\right) &= \log\left(\frac{p(m_i \mid z_t)}{p(\overline{m_i} \mid z_t)}\right) + \log\left(\frac{p(m_i \mid z_{1:t-1})}{p(\overline{m_i} \mid z_{1:t-1})}\right) \\ &\quad - \log\left(\frac{p(m_i)}{p(\overline{m_i})}\right). \end{aligned} \tag{14}$$

Formula (12) simplifies to

$$L(m_i \mid z_{1:t}) = L(m_i \mid z_t) + L(m_i \mid z_{1:t-1}) - L(m_i). \tag{15}$$

This is the basis of the inverse sensor model, and it updates the instructional image by integrating new information occupied by the unit without updating the entire instructional image.

Through the above derivation, especially formula (12), the remainder of this paper assumes that the occupancy probability of a unit is easy to convert between probability and log-difference ratio.

$$L(m_i \mid z_{1:t}) = \max\left(\min\left(L(m_i \mid z_t) + L(m_i \mid z_{1:t-1}) - L(m_i)\right), l_{\max}, l_{\min}\right). \tag{16}$$

Update formula (13), respectively. This restriction strategy has two main advantages. First, it intuitively limits the number of updates required to change the cell state, allowing the instructional image to change rapidly according to the

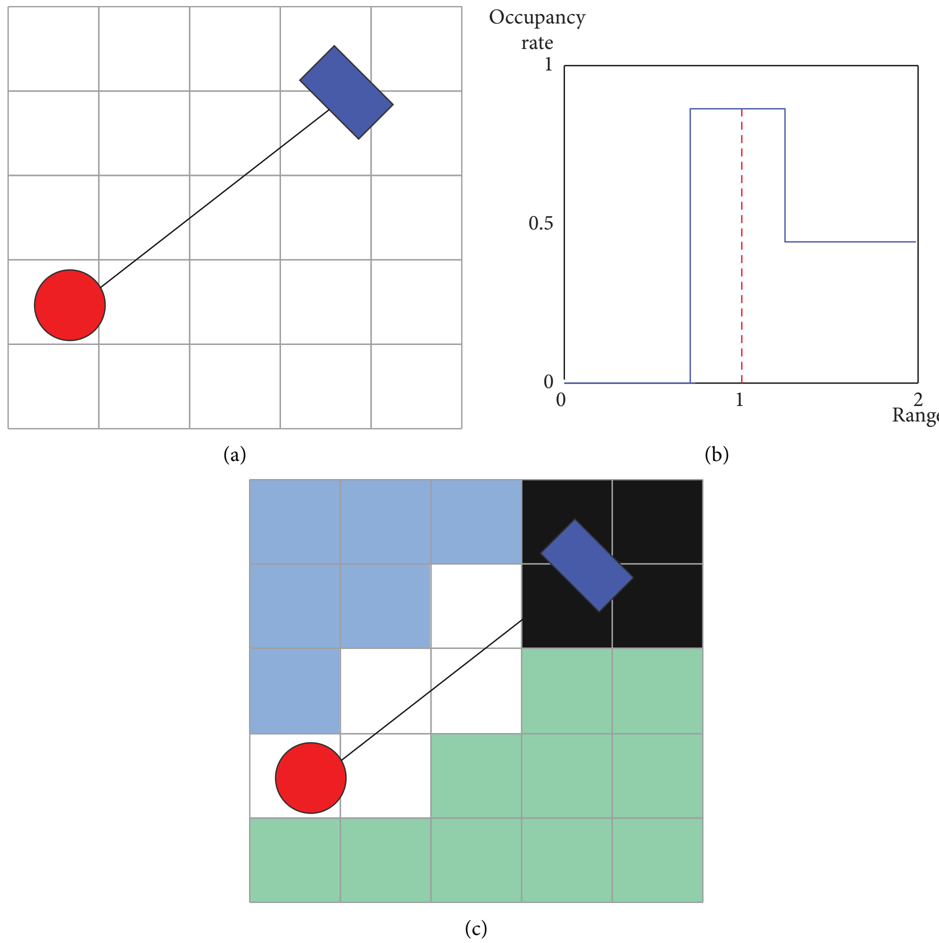


Figure 2: The ray casting diagram in two-dimensional space. (a) Light measurement. (b) Ideal model. (c) Occupation grid diagram.

changing environment. Second, it allows adjacent cells to be compressed with values close to the occupancy boundary.

The appearance of a traditional sensory device and the distance that exists in the world is always intellectual, as shown in Figure 2(a). Sensor noise should be considered when designing an ISM. The general equation for ISM is

$$p\left(m_i \mid z_{1:t}\right). \tag{17}$$

When an ideal sensor model is used to update an initially unknown teaching image for all cells, as shown in Figure 2(b), the occupancy probability of each unit between the sensor and the actual measurement unit is $p(m_i|z_t) = 0$. In addition, the probability that each sensor does not observe other units is 1. The updated grid map thus generated is shown in Figure 2(c).

It is also possible that the light missed an obstacle that only partially occupies the cell, as shown in Figure 3. Figures 3(a)–3(c) correspond to Figures 2(a)–2(c), respectively.

2.2. Gaussian Reverse Octree Occupancy Raster Map Generation. The Gaussian inverse sensor model is derived by convolving. As shown in Figure 4, the ideal sensor model, its measurement range is $z = 2\,\mathrm{m}$, and the buffer range around the measurement value z is $L = 1$, which can be expressed as formula (19). Hard probabilities are assigned $p(m_i)$ to each cell i, namely,

$$g(r) = \begin{cases} 0, & \text{if } r < z - \dfrac{L}{2}, \\ 1, & \text{if } z, \\ 0.5, & \text{otherwise for } r > 0. \end{cases} \tag{18}$$

As shown in Figure 5, the probability distribution function (PDF) of Gaussian noise is used to simulate the noise measured by the sensor. The standard probability density function of a Gaussian distribution is given by following equation:

$$f(r) = \frac{1}{\sigma\sqrt{2\pi}} e^{1\,(r-0)^2/2\sigma^2}. \tag{19}$$

The variance σ_z represents the uncertainty of the measured value z.

The Gaussian probability distribution function (PDF) of the normal distribution is shown in Figure 6.

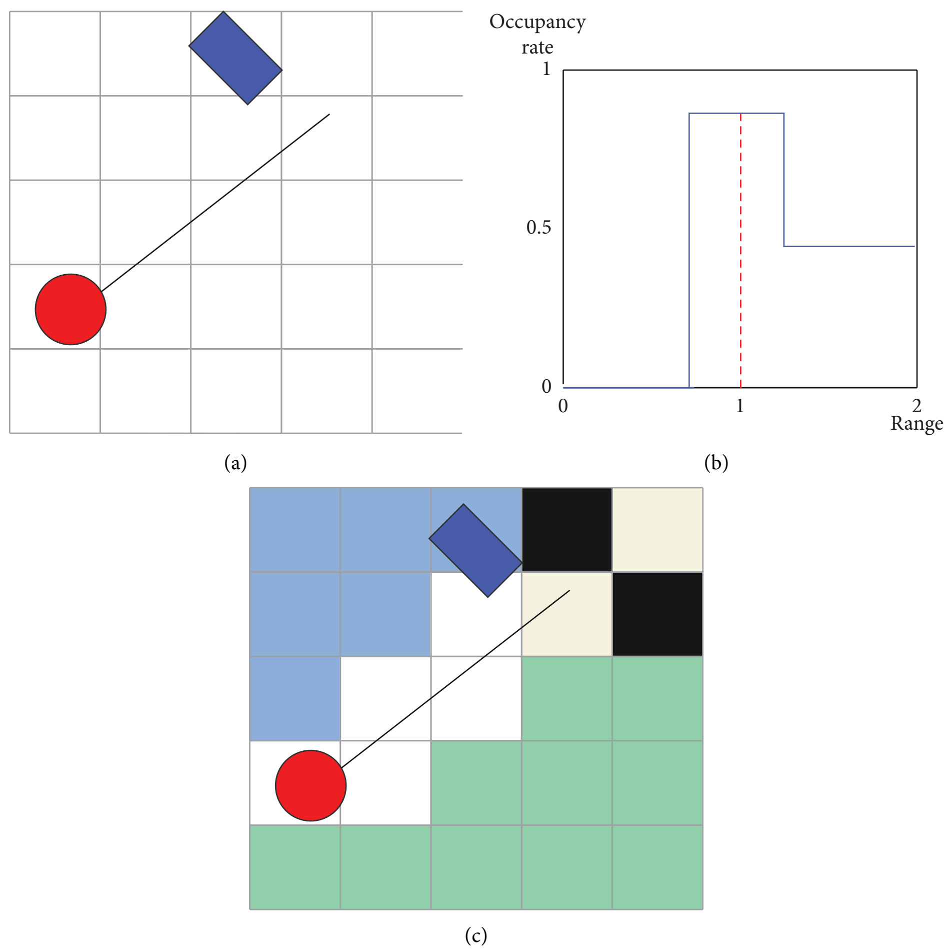


FIGURE 3: The ray casting diagram in the cell occupied by the obstacle part. (a) Light measurement. (b) Ideal model. (c) Occupation grid diagram.

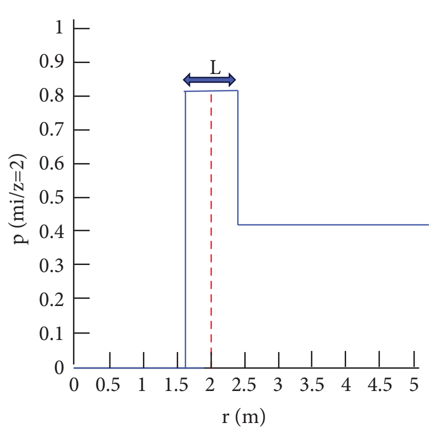


FIGURE 4: Ideal reverse sensor model diagram.

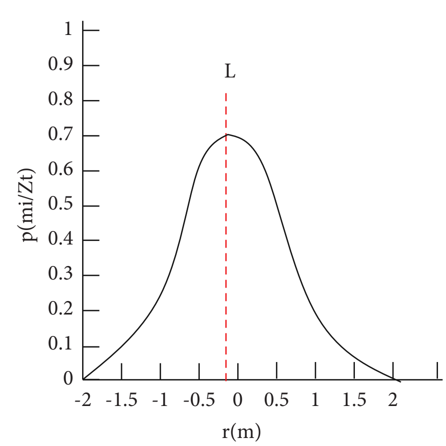


FIGURE 5: Diagram of Gaussian probability distribution function with normal distribution $\sigma_z = 0.5$.

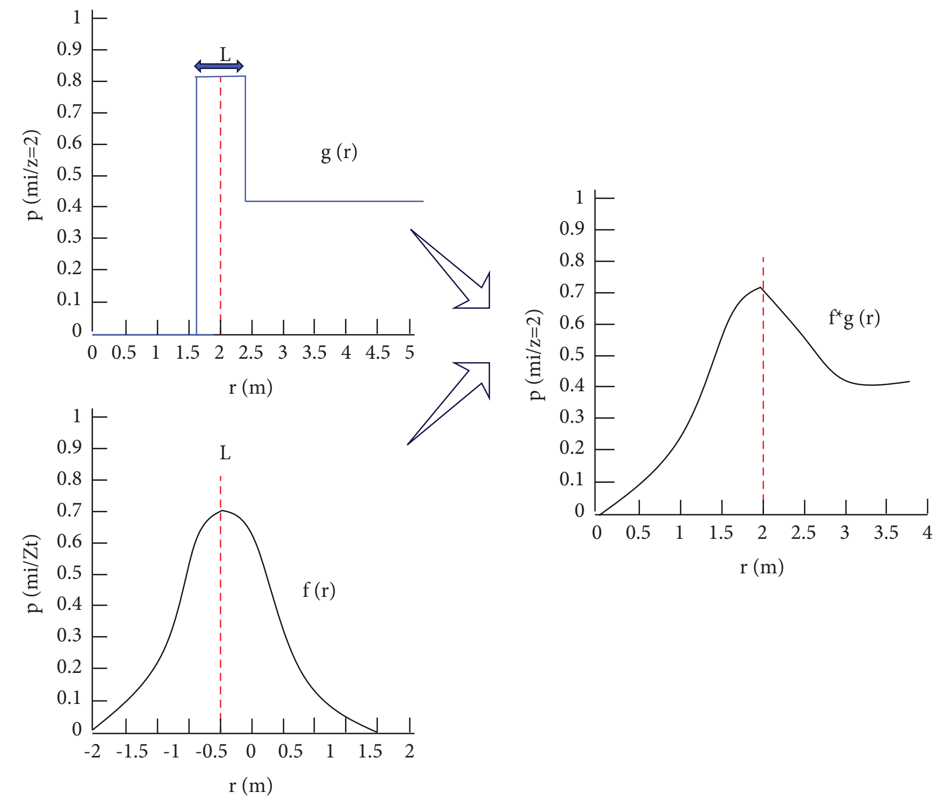


Figure 6: Convolution diagram of ideal sensor model and Gaussian probability distribution function.

$$\begin{aligned}(f * g)(r) &= 0F\left(-\infty, z-\frac{L}{2}\right) + 1F\left(z-\frac{L}{2}, z+\frac{L}{2}\right) + 0.5F\left(z+\frac{L}{2}, r\right)\\ &= -\frac{1}{2}erf\left(\frac{r-z-L/2}{\sigma\sqrt{2}}\right) + \frac{1}{2}erf\left(\frac{r-z+L/2}{\sigma\sqrt{2}}\right) - \frac{1}{4}erf\left(\frac{-z}{\sigma\sqrt{2}}\right) + \frac{1}{4}erf\left(\frac{r-z-L/2}{\sigma\sqrt{2}}\right)\\ &= -\frac{1}{4}erf\left(\frac{r-z-L/2}{\sigma\sqrt{2}}\right) + \frac{1}{2}erf\left(\frac{r-z+L/2}{\sigma\sqrt{2}}\right) - \frac{1}{4}erf\left(\frac{-z}{\sigma\sqrt{2}}\right).\end{aligned} \tag{20}$$

As shown in Figure 7, the Gaussian inverse sensor model describes that the evaluation along the measurement range tends to 0 until the actual distance to the obstacle is closer. Uncertainty causes the probability value to gradually increase. For cells close to the actual measurement peak, the probability value tends to 1, and then for cells beyond the measurement range, the probability value gradually decreases to 0.5.

Through the derivation and analysis of the reverse sensor model, it can be known that the main factors affecting the generated probability calculation and the shape of the curve are the buffer L, the standard deviation σ, and the measurement range z.

First, a description of each component of the model is given.

$$p\left(x_c, y_c, z_c\right)^T, l_p = \sqrt{x_c^2 + y_c^2 + z_c^2}. \tag{21}$$

Moreover, there is $l_{\min} = z_{c\min} > 0$. The uncertainty Δl_p of the distance l_p is related to the depth and can be described as

$$\Delta l_p = \Delta z_c \frac{l_p}{z_c}. \tag{22}$$

d is the parallax distance between the corresponding pixels.

$$z_c = \frac{bf}{d}. \tag{23}$$

Since the given image coordinates are imprecise and have uncertainty Δd, the distance measurement z_c given by

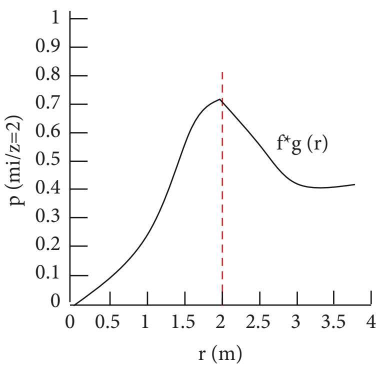


FIGURE 7: Gaussian noise reverse sensor model diagram.

the disparity function has an error Δz_c called the distance resolution, which is given by

$$\Delta z_c = \frac{z_c^2}{bf} \cdot \Delta d. \tag{24}$$

Similar to the Gaussian ISM, the ideal sensor model of Andert's ISM is given by

$$P_{occ}(r) = \begin{cases} P_{\min}, & \text{if } 0 < r < l_p, \\ 0.5, & \text{if } r > l_p, \end{cases} \tag{25}$$

where r is the measurement range.

Andert's sensor model is shown in the following formula:

$$P(s(r)) = P_{occ}(r) + \left(\frac{k}{\Delta l_p \sqrt{2\pi}} + 0.5 - P_{occ}(r)\right) e^{-1/2\left(r - l_p/\Delta l_p\right)^2}. \tag{26}$$

The effective factor ensures that farther measurements have less effect on the instructional image, while closer measurements have a greater effect on the instructional image. The combination of the effective factor, scaling, and quadratic exponential function produces an asymmetric probability curve around the distance measure. When the measurement distance reaches the maximum range, the probability curve gradually tends to 0.5.

According to the improvement of this paper, it is adapted to monocular vision. The final ISM formula is

$$P(s(r)) = P_{occ}(r) + \left(\frac{k}{\sigma_{zt} \sqrt{2\pi}} + 0.5 - P_{occ}(r)\right) e^{-1/2\left(r - l_p/\sigma_{zt}\right)^2}, \tag{27}$$

where σ_{zt} is the uncertainty of the LSD-SLAM inverse depth measurement. This formula determines the occupancy probability of a unit in a monocular vision system and differs from stereo vision in that the distance-related uncertainty Δl_p is replaced by the uncertainty σ_{zt} of the LSD-SLAM inverse depth measurement.

3. Simulation Test

Whether the real-time video processing technology proposed in the second part can promote the effect of English culture teaching in English culture teaching needs to be verified by experiments. First, the effect of the algorithm model in this paper is verified.

As shown in Figure 8, where $z = 1$ m and $\sigma = 0.1$, larger values of L expand and increase the height of the peak, and smaller values shrink and lower the peak. The value of L is usually the diagonal width of the cells in the occupied raster image. If the value is too large, it is defined that the occupancy of more surrounding cells is actually measured along the measurement ray.

The standard deviation σ, when applied to the ISM, represents the uncertainty in the occupancy prediction. Intuitively, smaller σ corresponds to higher and narrower peaks, and larger σ corresponds to lower and wider peaks, as shown in Figure 9, where $z = \text{lm}$ and $L = 0.4$.

LSD-SLAM does not have the same degree of distance-dependent uncertainty as conventional fixed baseline stereo matching and is therefore not considered in the final model. As shown in Figure 10, measurements were taken at $z = 1$ m, 2 m, and 3 m and $L = 0.4$, showing the ISM probabilities for different measurement distances. It can be seen that the Gaussian ISM overshoots within the peak probability range of the occupancy prediction, and the probability of the measured value in the range is slightly lower than the actual measured value.

Figure 11 shows the resulting probabilities for different measured distances independent of distance uncertainty. The unit size is 0.1 m. For cells before measurement, the cells are divided into free spaces and the probability is assigned the minimum value. In addition to this, the units after the measurement range tend to be 0.5, indicating that they are unknown.

This section will discuss in detail the impact of the key parameters of the improved measurement ISM model, especially the effective factor k and the standard deviation σ (measurement uncertainty). It will show the shape of the resulting probability curve when depth-dependent uncertainty is considered.

3.1. Effective Factor k. Figure 12 shows the effect of the effective factor k on the probability of results based on depth measurements. The unit size is 0.1 m, the standard deviation $\sigma = 0.1$, and the measurement distance $z = 3$ m. The effective factor acts as a scale factor for the measurement, allowing measurements at close range to have a greater impact on the instructional image than measurements at a distance. Since it is unclear how to choose this factor, it is still necessary to find a suitable effective factor to provide the best results for all measurements of monocular systems.

As expected, the standard deviation σ is very similar to the Gaussian ISM. As shown in Figure 13, the effective factor $k = 0.1$, and the measurement distance $z = 3$ m. The larger σ, the lower and wider the peak, and the smaller σ, the narrower and higher the peak. However, the improved measured ISM has a significantly narrower curve than the Gaussian ISM, making its prediction of occupancy more specific.

For the improved measurement inverse sensor model, the effect of distance-dependent uncertainty is shown in Figure 14. The unit size is 0.1 m, the effective factor $k = 0.1$,

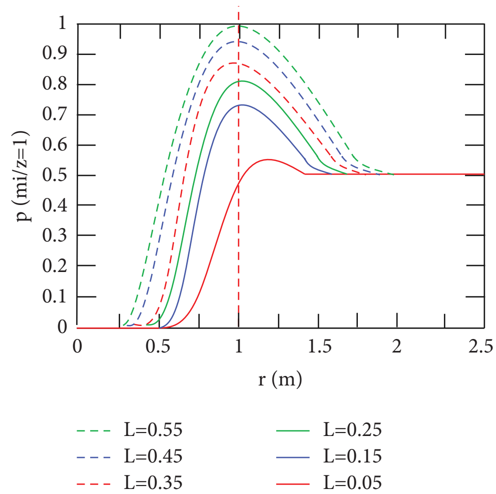


FIGURE 8: Effect diagram of different L values on the Gaussian inverse sensor model.

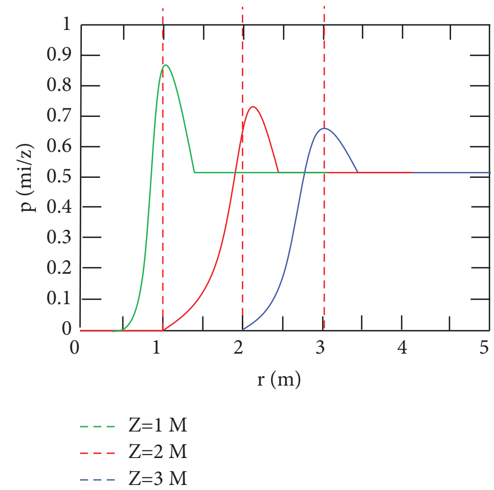


FIGURE 10: Effect diagram of different z values on the Gaussian inverse sensor model.

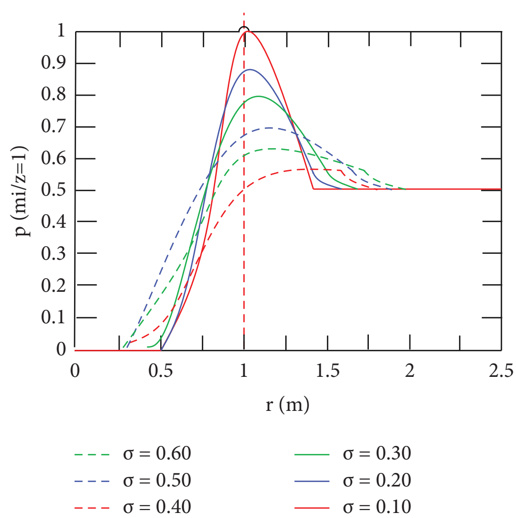


FIGURE 9: Effect diagram of different σ values on the Gaussian inverse sensor model.

$b = 0.3$ m, $f = 700$ pixels, and the parallax $d = 0.5$. In addition, it should be noted that compared to Gaussian ISM, the improved measurement ISM predicts the peak occupancy within the actual measurement range, while the Gaussian ISM predicts the occupancy after the actual measurement.

The above system is applied to English culture teaching, and the evaluation is carried out through the teaching, and the results shown in Table 1 are obtained.

Language and culture are closely linked. It is impossible to teach a language without culture, which is the necessary context for language use. The target language cannot be mastered well without cultural knowledge related to the target language. Therefore, learners should be as close as possible to native speakers of the target language, to understand their way of life, way of thinking, and behavior. Cultural teaching provides the necessary connections between language and society and culture and real life. Only by understanding the culture of a society can they fully understand the way of thinking, behavior, and speech of native speakers, so as to truly master the target language. In this sense, gaining information about foreign cultures is as important as language learning itself. Cultural studies provide learners with insight into the target culture, providing the opportunity to learn about other countries, people, and ways of life without direct experience. People may be more tolerant of other cultures than being narrow-minded. Cultural studies can broaden people's horizons and develop their personalities.

The learner's affective factors require imparting of cultural knowledge. The learner is the main body of learning, and their cognitive and affective factors play an important role in learning. Emotional factors include motivation, attitude, and personality traits. Motivation can be divided into extrinsic and intrinsic motivation. Intrinsic motivation refers to the student's current goals and objectives, and extrinsic motivation refers to artificial incentives, that is, "bribes of pleasure." All these joys motivate learners to study hard. However, not all learners can obtain such happiness because not all learners can achieve good results in language learning. Therefore, it is necessary to explore another kind of joy that every language learner can derive from language learning, and it is the teacher's responsibility to help them get the motivation from learning itself rather than from getting good grades in exams. The joy of learning a foreign language can come from the culture conveyed by the language itself. Language is the carrier of culture. If learners can not only learn about the language system itself but also gain some cultural knowledge, they will definitely have fun exploring new things. English reflects the beliefs, values, and lifestyles of native English speakers, which may differ from Chinese. If teachers combine cultural teaching with language teaching, students can have fun in understanding the similarities and differences between the culture of the mother tongue and the culture of the target language. Needless to

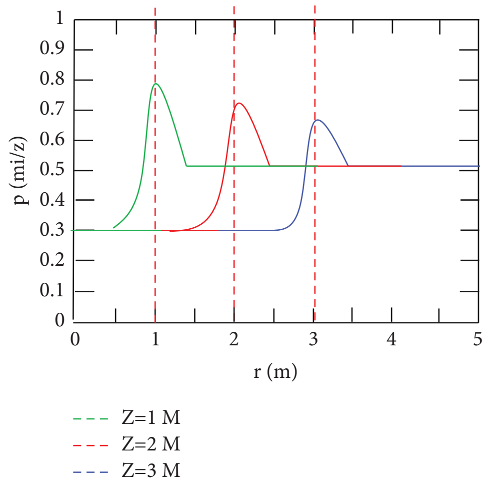


Figure 11: Diagram of an improved measurement inverse sensor model without depth-dependent uncertainty.

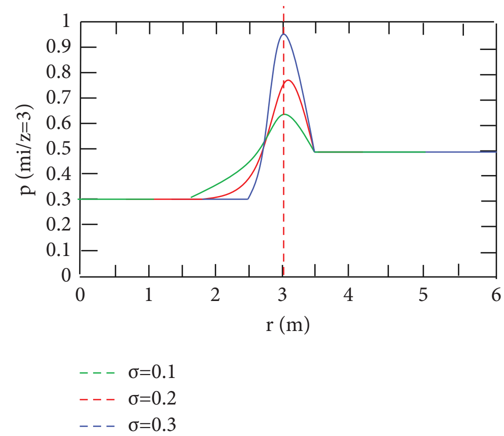


Figure 13: Effect of different σ values on the improved measurement inverse sensor model.

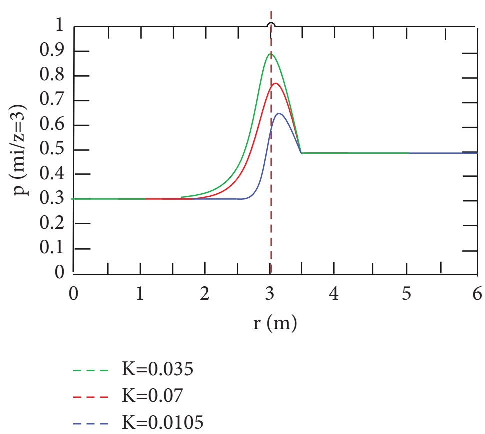


Figure 12: Effect of different k values on the improved measurement inverse sensor model.

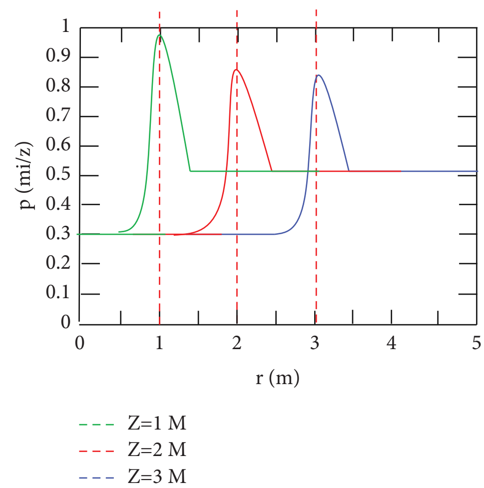


Figure 14: Effect of different z values on the improved measurement inverse sensor model.

Table 1: The application effect of real-time video processing technology in English culture teaching.

Number	Teaching evaluation	Number	Teaching evaluation	Number	Teaching evaluation
1	85.00	18	85.22	35	82.25
2	82.09	19	83.20	36	84.84
3	85.26	20	87.18	37	86.03
4	86.45	21	83.76	38	86.20
5	87.47	22	85.56	39	85.82
6	84.39	23	82.14	40	83.84
7	86.94	24	87.07	41	87.34
8	82.02	25	84.77	42	85.79
9	82.01	26	82.85	43	83.74
10	87.46	27	87.58	44	87.83
11	83.97	28	81.54	45	82.32
12	85.04	29	82.93	46	85.08
13	83.25	30	83.62	47	83.49
14	83.48	31	86.96	48	81.05
15	86.74	32	83.25	49	81.72
16	81.99	33	81.55	50	85.51
17	82.38	34	85.86	51	85.23

say, cultural teaching is an important supplement to language teaching, which enables students to experience fun in the process of learning a lot of language materials. The study of the target language culture can improve and enrich students' understanding of the native language culture and can enable students to gain pleasure from learning a large number of language materials. Obviously, learning the target language culture can enhance students' understanding of the world, cultivate multicultural awareness, and increase their interest in foreign language learning.

It can be seen from the above research that the real-time video processing technology proposed in this paper has a good effect in English culture teaching, and it can be applied to English culture teaching to improve the effect of English culture teaching.

4. Conclusion

Intercultural communication is an exchange process of equal bilateral interaction between different cultures. Mother tongue culture is an integral part of it, a springboard, starting point, and bridge for communication. In developing intercultural communication skills, the mother tongue culture is the basis for understanding different cultures and improving sensitivity and tolerance and critical acceptance of different cultures. Moreover, the introduction of culture should also follow scientific principles, that is, avoid subjective arbitrariness and partial generalization and strive to be accurate, comprehensive, and objective. In addition, cultural teaching must be combined with language teaching, so as to be natural and not rigid. In the process of lesson preparation, classroom teaching should be designed according to the development of students' abilities, and specific teaching methods should be explored. In order to improve the effect of English culture teaching, this paper combines the real-time video image processing technology to construct an English culture teaching system. The experimental research results show that the real-time video processing technology proposed in this paper has a good effect in English culture teaching, and it can be applied to English culture teaching to improve the effect of English culture teaching.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This research was supported by 2020 Hunan Education Science "13th Five-Year Plan" Project (project no. XJK20CGD048) and Research Project of Teaching Reform in Colleges and Universities of Hunan Province in 2020 (project no. HNJG-2020-1055).

References

[1] H. Salehi, "Effects of using instructional video games on teaching English vocabulary to iranian pre-intermediate efl learners," *International Journal of Learning and Change*, vol. 9, no. 2, p. 111, 2017.

[2] A. Zhao and W. Liu, "Design and implementation of English autonomous listening system based on web database," *Revista de la Facultad de Ingenieria*, vol. 32, no. 11, pp. 247–253, 2017.

[3] P. J. B. Tan and M. H. Hsu, "Designing a system for English evaluation and teaching devices: a pzb and tam model analysis," *Eurasia Journal of Mathematics, Science and Technology Education*, vol. 14, no. 6, pp. 2107–2119, 2018.

[4] S. Qi, S. Li, and J. Zhang, "Designing a teaching assistant system for physical education using web technology," *Mobile Information Systems*, vol. 2021, no. 6, pp. 1–11, 2021.

[5] C. Wen, "Study on the role of multimedia video in oral English communicative teaching," *Boletin Tecnico/Technical Bulletin*, vol. 55, no. 18, pp. 698–704, 2017.

[6] А. Федорова, A. Fedorova, А. Быстрова, and A. Bystrova, "Video project as a method of teaching a foreign language in the department of linguistics," *Scientific Research and Development Socio-Humanitarian Research and Technology*, vol. 8, no. 3, pp. 68–74, 2019.

[7] Y. Du and T. Zhao, "Network teaching technology based on big data mining and information fusion," *Security and Communication Networks*, vol. 2021, no. 9, pp. 1–9, 2021.

[8] J. Lounek and Josef, "Developing creativity in teaching digital video courses at the faculty of informatics and management," *Procedia - Social and Behavioral Sciences*, vol. 186, no. 3, pp. 283–287, 2015.

[9] L. F. Lin, "The impact of video-based materials on Chinese-speaking learners' English text comprehension," *English Language Teaching*, vol. 9, no. 10, pp. 1–11, 2016.

[10] L. E. Gustilo, M. R. C. Lapinid, J. S. Barrot, M. K. L. Gabinete, C. P. Magno, and A. J. M. Osman, "Effects of knowledge channel videos on the achievement of students in various learning environments," *Advanced Science Letters*, vol. 21, no. 7, pp. 2276–2280, 2015.

[11] M. Ando and M. Ueno, "Effect analysis of pointer presentations on multimedia e-learning materials based on dual channel model," *Educational Technology Research*, vol. 34, no. 6, pp. 59–73, 2017.

[12] S. M. Juquiana, J. Carlo Padua, H. Ilao, J. Simangan, and L. Esteban Elipane, "Teaching of angle of elevation by integrating angle measurement in ancient egyptian mathematics," *International Journal of Engineering & Technology*, vol. 7, no. 4.1, pp. 76–84, 2018.

[13] H. Madokoro, Y. Ishioka, S. Takahashi, K. Sato, and N. Shimoi, "Visual saliency based multiple objects segmentation and its parallel implementation for real-time vision processing," *Computer Science and Information Technology*, vol. 3, no. 5, pp. 187–197, 2015.

[14] Y. Lv, B. Li, and B. Niu, "A model research of basketball teaching and training aided system in colleges and universities based on video image processing technology," *Revista de la Facultad de Ingenieria*, vol. 32, no. 14, pp. 795–800, 2017.

[15] R. Rajarathinam and R. Latha, "Analysis on video retrieval using speech and text for content-based information," *Middle-East Journal of Scientific Research*, vol. 23, no. 6, pp. 370–376, 2015.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 8021288, 15 pages
https://doi.org/10.1155/2022/8021288

Research Article

FPGA-Based Real-Time Embedded Fish Embryo Detection System

Mengqi Wang,[1] **Guofu Feng**,[1] **Ming Chen**,[1] **Ruijuan Ye**,[1] **and Yaohui Wang**[2]

[1]*Key Laboratory of Fisheries Information, Ministry of Agriculture and Rural Affairs, Shanghai Ocean University, Hucheng Ring Road 999, Shanghai 201306, China*
[2]*Nantong Long Yang Aquatic Company, Haian 226600, Jiangsu, China*

Correspondence should be addressed to Guofu Feng; gffeng@shou.edu.cn

Received 17 August 2022; Revised 6 September 2022; Accepted 19 September 2022; Published 7 October 2022

Academic Editor: Jiafu Su

In today's aquaculture industry, the period of fish embryos needs to be detected in real time because different periods of fish embryos require different environments for their cultivation. The paper proposes an FPGA-based real-time embedded fish embryo detection system to solve the problems of today's fish embryo period detection, which consumes a lot of human resources and has low accuracy. Based on the selection of YOLOv3 as the basic model, this paper combines model optimization and hardware acceleration to make the whole system achieve the effect of real-time detection. With only a 4.23% reduction in the mAP value of the model, the inference speed of the model was improved by nearly 150%. Finally, we achieved 34.1 fps of real-time fish embryo detection on a 100 MHz FPGA chip, while having 82.49% mAP, and the whole embedded system only consumed 2.524w while running. The results show that the FPGA-based real-time fish embryo detection system proposed in this paper is fully capable of meeting the needs of today's aquaculture embedded devices. It is beneficial to promote the transformation of fish embryo analysis from traditional manual methods to AI-based embedded devices and facilitate the rapid analysis of fish embryos.

1. Introduction

Fish embryonic development has always been an extremely critical aspect of aquaculture. It is crucial for the continuation of rare fish species, high yield breeding, and sustainable development of fishery resources [1]. Since a mismatched culture environment can have a significant impact on embryo development, mainly in terms of survival rate and malformation rate, real-time fish embryo detection results are needed to adjust the corresponding culture environment [2–4]. Because fish embryos are very small before they are fully developed, the traditional method is to manually operate a microscope to observe the growth of fish embryos [1]. The approach has many problems, such as the difficulty of deploying equipment in fish culture greenhouses, the experience gap between analysts that affects the accuracy of identification, and the difficulty of analyzing directly in the culture environment. Therefore, automation and miniaturization of equipment are the main research directions for fish embryo detection in the future [5, 6].

It is difficult to achieve good results in aquaculture by relying only on embryo analysis by human eyes, so it is clearly not feasible to spend a lot of human resources to perform embryo analysis. Artificial intelligence would be a good solution to solve the problem in traditional methods. In 2017, Ishaq et al. demonstrate the potential of a deep learning approach for accurate high-throughput classification of whole-body zebrafish [7]. In 2019, Rauf et al. develop a framework based on improved CNN to classify the fish species [8]. In 2021, Naderi et al. develop CNN-based framework for cardiac function assessment in embryonic zebrafish from heart beating videos [9]. With the widespread use of artificial intelligence in cytology, bioinformatics, and medicine [10–12], embryo detection in aquaculture and academia is gradually experimenting with the approach. Qian et al. [13] established a microinjection system for detecting zebrafish embryos using CNN. However, the method is more suitable for deployment in the laboratory and cannot be directly applied to aquaculture. Xu et al. [14] built a Drosophila embryo detection framework using multifeature fusion (MFF) CNN on GPU. Dirvanauskas et al. [15] used CNN to classify embryonic development. The above studies are CPU based for both classification and detection tasks for CNNs and use GPUs for acceleration. In environments such as greenhouses for aquaculture, the approach may make it difficult to deploy equipment directly to the field due to various uncertainties. The conventional CPU + GPU architecture is difficult to fully meet the needs of aquaculture due to various instrument failures in environments such as greenhouses due to temperature and humidity.

Recently, field programmable gate arrays (FPGAs) have attracted increasing attention from researchers because of their excellent performance, high energy efficiency, fast development cycles, and strong ability of refactoring [16–18]. Although GPUs outperform FPGAs in terms of inference efficiency, dedicated devices made of FPGAs significantly outperform GPUs in terms of power consumption. Such characteristics make FPGAs a better choice for embedded devices [19–21]. In addition, FPGAs can design camera devices directly into the programable logic (PL), allowing FPGAs to process video streams directly with near-no latency [22]. And a series of modern chips represented by ZYNQ make up for the weak logic processing power by embedding ARM processors in FPGAs, fully meeting the needs of deep learning in terms of flexibility [17, 23].

2. Related Work

To the best of our knowledge, we are not aware of any implementation of embryo detection on FPGA.

Nowadays, there are indeed many researchers who have achieved some results in the intersection of FPGA and other disciplines. Zhao et al. [24] built an FPGA-based underwater image recognition system. The system successfully deploys a convolutional neural network on Xilinx's ZYNQ7035 chip. Wei et al. [25] complete an FPGA-based remote sensing image classification system. Luo and Chen [26] built an FPGA-based defect detection system for additive manufacturing, and they used a binarized neural network (BNN) to turn the computational part of the neural network into the most suitable FPGA bit operation. Al Koutayni et al. [27] built the first CNN-based gesture recognition system on ZYNQ. However, the above approach does not take into account the optimization of the model itself, and since the system itself only deals with simple classification problems, it is only necessary to deploy the model to the FPGA. However, if the model structure or the data type of the model can be changed to directly reduce the computation of the model itself, the inference speed of the model will be significantly improved.

To solve the above problems, we propose an FPGA-based embedded system to achieve real-time detection of fish embryos. The embedded system uses a combination of hardware and software to improve the inference speed. We first performed channel pruning, layer pruning, and INT4 quantize on the pretrained model. After pruning, the model is retrained to fine-tune it. In quantization, the TQT method is used to increase the inference speed of the model with as little loss of accuracy as possible. After completing the optimization of the model, we designed an efficient hardware architecture on ZYNQ. The architecture uses FPGAs to implement the computationally intensive part and an ARM CPU to implement the logic control part. By deploying the optimized model to this structure, we have accomplished the detection of fish embryos in real time.

3. Software Design Process

3.1. YOLOv3-Based Fish Embryo Detection. The target detection model chosen for the paper is the YOLOv3 model proposed by Redmon [28]. YOLOv3 was chosen as the final model for our implementation of embryo detection. Because compared to YOLOv4 and YOLOv5, the accuracy of YOLOv3 is not as good as these two models, but YOLOv3 gets the simplicity of the network model and higher inference speed with the cost of some accuracy on the data set relevant to this application. For embryo detection tasks with a small number of classifications and large of image similarity, the accuracy of YOLOv3 is also fully capable of handling the corresponding needs. Faster R-CNN has higher accuracy but slower inference speed. MobilenetV3 has faster inference speed but poorer performance for small target detection. The YOLO series, which combines inference speed and accuracy, is more suitable for deployment on FPGA platforms to accomplish the goal of real-time embryo detection than the previously mentioned models [22, 29, 30].

Figure 1 illustrates the network structure of YOLOv3. YOLOv3 adds a BN layer and a LeakyReLU layer [31] after each convolutional layer in the network. This structure is called DBL (DarknetConv2D + BN + LeakyReLU). The DBL structure is the basic component of the YOLOv3 network. As seen in Figure 1, two DBLs form the Res structure by direct connection and hop-level connection. Five res structures and one DBL structure form the backbone of YOLOv3. The input image size of YOLOv3 is $416 \times 416 \times 3$. There are three outputs in traditional YOLOv3 with feature map depths of 255 and an image edge ratio of 13 : 26 : 52. The depth of each output feature map is based on the bounding box of the YOLO algorithm.

3.2. YOLOv3 Optimization Strategy for FPGAs. Since deploying the entire YOLOv3 model directly to the FPGA would put a huge strain on the memory space and computational resources of the entire system. YOLOv3 has 52 convolutional layers. Direct deployment without any optimization of the model will make the network inference time much higher than the expected time. Thus, in the paper, model pruning and model quantization of YOLOv3 are chosen to reduce the calculation of the network. The process of YOLOv3 model optimization is given in Figure 2. Following the training of the initial neural network, the model will be pruned. After the pruning of the model, the accuracy lost by the pruned model is compensated by fine-tuning operations. In this paper, we choose to use the original data set to retrain the pruned model again. Then, fine-tuning operations are performed to compensate for the lost

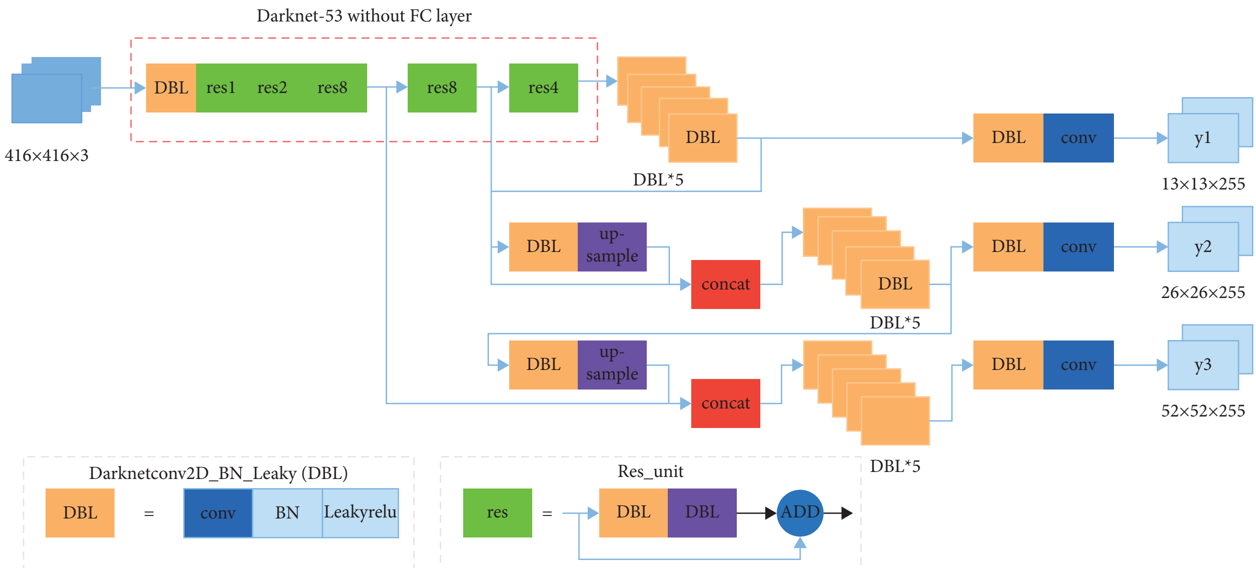


Figure 1: Network structure of YOLOv3.

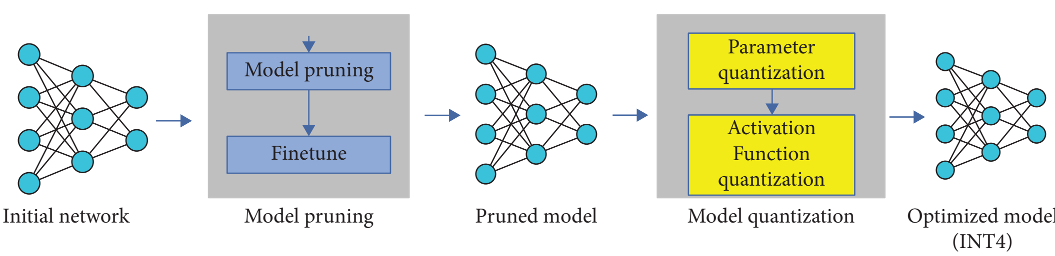


Figure 2: Model optimization process.

accuracy of the model after pruning. The approach taken in this paper is to use the original data set to train the pruned model again. Finally, the whole model is quantified into the INT4 model. The quantized YOLOv3 model will be more suitable for computing on FPGAs.

3.2.1. Pruning of YOLOv3. The pruning operation for convolutional neural networks is classified into layer pruning and channel pruning. The pruning of the network model can reduce the computation of the model and improve the inference speed of the model. Traditional pruning is mostly based on the size of the model's weights. This is not suitable for YOLOv3, which has an extremely large number of weight parameters. In order to achieve better effect, a pruning method based on BN layer scaling factor is used in this paper. Since more than 90% of the operations of the convolutional neural network are done in the convolutional layer, the importance of the convolutional layer in the network should be considered when pruning the model. In YOLOv3, the basic unit is the DBL structure, in which the batch normalization (BN) layer plays the role of accelerating the convergence of the network and controlling the occurrence of the overfitting phenomenon. In the DBL structure, the BN layer is after the convolution layer and before the activation function LeakyReLU. Its calculation formula is as follows:

$$x_{\text{out}} = \frac{\gamma\left(x_{conv} - \mu\right)}{\sqrt{\sigma^2} + 0.000001} + \beta. \tag{1}$$

In formula (1), μ is the mean value of all input x in an input batch; γ is called the scaling factor; σ^2 is the variance; β is the bias; and $x_{\text{conv in channel } c}$ is the computed result of the convolution layer output in channel c of the input, which can be expressed as

$$x_{\text{conv in channel } c} = \sum_{i=0}^{n} \left(x_{ic} * w_{ic}\right). \tag{2}$$

`In equation (2), w_{ic} is the weight required in the convolution operation and is stored in the corresponding weight file. As seen from equation (1), the result of the convolutional layer outputs is directly affected the scaling factor γ, and the magnitude of γ can be understood as the importance of the convolutional layer output result in the next round of operations. The larger the scaling factor γ is, the more important its corresponding channel is in the overall DBL structure. In order to make the distribution of the scaling factor γ in the DBL structure more suitable for the pruning operation, we add the scaling factor γ to the loss function of the model and perform sparse training again on the already trained model. The formula is as follows:

$$\text{Loss} = \text{Loss}(w) + \lambda \sum_{\gamma\in\Gamma} \text{Reg}_g(\gamma). \tag{3}$$

In formula (3), $\text{Loss}(w)$ is the original loss function of YOLOv3; $\text{Reg}_g(\gamma)$ is the sparse regular term; Γ is the set of γ; and λ is the sparsity factor. In this paper, stochastic gradient descent (SGD) optimization is used, and the gradient calculation process can be written as

$$\nabla\gamma^* = \nabla_\gamma \lambda \sum_{\gamma\in\Gamma} \text{Reg}_g(\gamma). \tag{4}$$

Assuming that the learning rate is η, the parameters in equation (4) can be updated as follows:

$$\gamma' = \gamma - \eta\nabla\gamma^*, = \gamma - \eta\lambda\nabla_\gamma \sum_{\gamma\in\Gamma} \text{Reg}_g(\gamma). \tag{5}$$

From equation (5), it is clear that calculating the amount of variation of the γ parameter only requires calculating the partial derivatives of the regular terms and multiplying them by the corresponding sparsity factor $\eta\lambda$ throughout the training process of the network. After the sparse training, channels with relatively small absolute values of γ, i.e., channels with γ parameters close to 0, are removed to perform channel pruning on the model. In this paper, channels with scaling factor γ less than 10^{-3} are removed.

Figure 3 shows the specific operation of channel pruning. In the convolution operation from layer j to layer k, channel 2 and channel 4 in the convolution layer j will be removed because their value of the scaling factor γ is smaller to other channels and less than 10^{-3}.

After the channel pruning, the YOLOv3 still needs to be layer pruned due to insufficient compression of the model. Since the smallest unit in the YOLOv3 structure is the DBL structure, the value of all γ on the BN layer can also reflect the importance of the DBL structure in the whole network. The γ of all remaining channels in each layer is averaged, and the average value is used as a criterion to rank the calculation importance of all DBL structures at the time of deletion. A pruning rate will be set to remove unimportant DBL structures in the network. As shown in Figure 4, layer pruning was performed to cull the model DBL structure based on the mean value of γ. DBL2 and DBLM with smaller mean values of γ were removed.

After the whole pruning operation is completed, a narrower, more compact model is obtained, which needs to be fine-tuned since the accuracy of this model will necessarily decrease. In this paper, the original data set is used as the fine-tuned data set to train the pruned model for another 50 epochs with the original loss function and 1% learning rate. After fine-tuning, the pruned model will be quantified.

3.2.2. Quantification of the Pruned YOLOv3. Traditional quantization methods used in FPGAs generally choose to convert the entire floating-point model into an 8 bit fixed-point model (INT8). This type of quantification method generally discards the decimal values directly in the quantification process [32]. However, such an approach leads to a significant decrease in the accuracy of the network model. To minimize the loss of model accuracy due to quantization while increasing the speed of inference as much as possible, the trained quantization threshold (TQT) method is used in this paper to transform the already pruned single-precision floating-point model (FP32) into a 4 bit fixed-point model (INT4) during the training process [33], where the quantization function $q(x; s)$ can be written as

$$q(x; s) = \text{clip}\left(\text{p}\lfloor\frac{x}{s}\rfloor; n, p\right) \cdot s. \tag{6}$$

In formula (6), x is the value to be quantified, and s is called the quantization scale factor. When x is a signed number, $n = -2^{b-1}$, $p = 2^{b-1} - 1$, and $s = 2^{\lceil\log_2 t\rceil}/2^{b-1}$. b is the bit width of quantification results. t is a trainable parameter, which is used as a quantization threshold in equation (6). It is usually set to the maximum of the weights. $clip\ (a; b, c)$ function is used to restrict a to the middle of b, c. Let a equal c when a is greater than c, and let a equal b when a is less than b, $\lfloor x\rceil$, i.e., rounding the input x. Since the quantization scale factor s is a power of 2, it is more friendly to hardware, and the pruning operation in the previous section has removed some of the outliers and promoted a tighter distribution of weights and activation functions, which makes the quantized results more accurate.

As mentioned before, the threshold t is a learnable parameter in training, and $\log_2 t$ is used to control the range of the parameter after quantization. The gradient of t can be obtained by the chain rule. The gradient is calculated as follows:

$$\nabla_{(\log_2 t)} q(x; s) \coloneqq s \cdot \ln 2 \cdot \begin{cases} \left\lceil \frac{x}{s} \right\rfloor - \frac{x}{s}, & n \le \left\lceil \frac{x}{s} \right\rfloor \le p, \\ n, & \left\lceil \frac{x}{s} \right\rfloor < n, \\ p, & \left\lceil \frac{x}{s} \right\rfloor > p. \end{cases} \tag{7}$$

Then, the gradient of x can be obtained as

$$\nabla_x q(x; s) \coloneqq \begin{cases} 1, & n \le \left\lceil \frac{x}{s} \right\rfloor \le p, \\ 0, & \text{otherwise}, \end{cases} \tag{8}$$

where we define the gradient of the integrable function $\lceil x\rfloor$ as

$$\frac{\partial}{\partial x}\lceil x\rfloor = \frac{\partial}{\partial x}\lceil x\rfloor = 1. \tag{9}$$

In this paper, Algorithm 1 is used to quantify the entire network during the quantization training.

4. Hardware Design Process

4.1. General Architecture. In the paper, a heterogeneous architecture is used to implement the whole system. The architecture can be divided into processing system (PS) and programable logic (PL). The PS part consists of a CPU that can be used as a controller and an off-chip memory that is

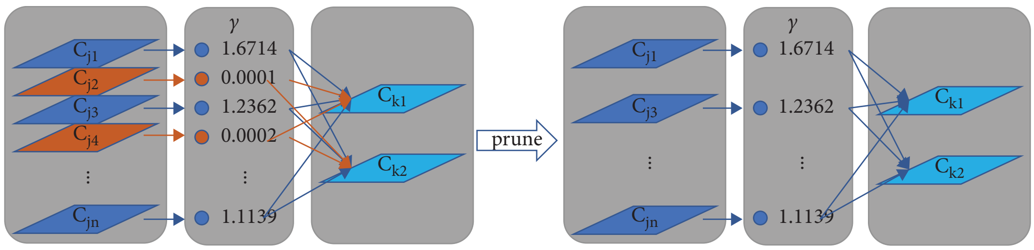


Figure 3: Channel pruning.

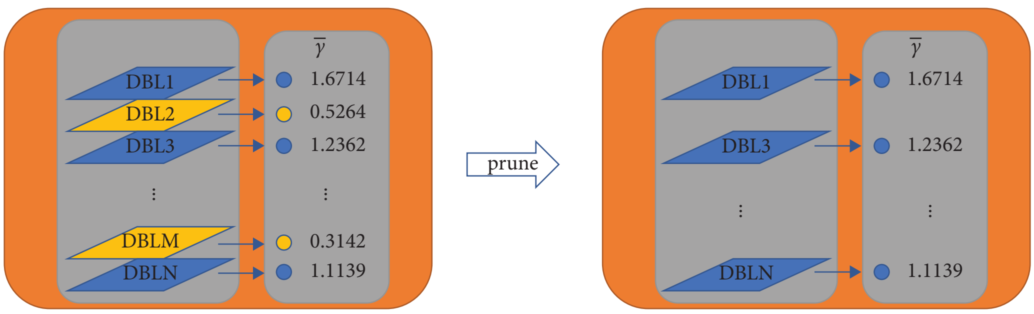


Figure 4: Layer pruning.

Full precision inputs, weights and bias, outputs: $X, W,$ bias
Bit width: for inputs and weights, $b = 4$, for bias, $b = 8$
Output: Y, predict number: Label
(1) Calculate q(x), q(w), q(bias) according to equation (6)
(2) Y = Forward(q(x), q(w), q(bias))
(3) Calculate losses and regularize all learnable parameters: L = loss (Y, Label)
(4) According to Equation (8), calculate gradient: $\nabla_X L = \partial L/\partial q(x) * \partial q(x)/\partial X, \partial q(x)/\partial X$
(5) Use the Adam optimizer [31] to update a portion of the full precision parameters according to Equation (9); W = Backward (W, Adam, $\nabla_X L$)
(6) Returning to 1.

Algorithm 1: Quantization algorithm for each layer.

used to store data; the PS is responsible for the scheduling and storage of the entire system. The PL part consists of several custom IP cores implemented on FPGAs, each of which is responsible for implementing a network layer function; the PL part is responsible for the computational part of the system. An entire system is implemented on Xilinx ZYNQ chip. Figure 5 illustrates the overall architecture of the system.

The processor is an ARM processor embedded on the chip. DDR is off-chip memory. DMA indicates that the data transfer method between memory and CPU and FPGA is direct memory access. The AXI bus is used for data and control signal transfer. The CNN accelerator is a neural network computing part composed of multiple IP cores. To transfer large amounts of data, the off-chip DDR is connected to the CNN accelerator via the AXI4-STREAM bus. The system uses a DMA controller internally to take care of address reading and writing. DMA0 divides the separate read and write data ports, each of which is 64 bits, and is responsible for transferring the feature maps and weights that the CNN accelerator needs to compute, as well as for writing the results of the CNN accelerator's computation to memory. Since DMA1 is only responsible for transferring the results of the previously obtained convolutional layer operations, DMA1 is only responsible for moving data from the main memory to the slave memory, which is located inside the CNN accelerator and is used for the layer-hopping accumulation operations inside the accelerator. The on-chip processor coordinates system operation via the AXI4-Lite bus. The AXI bus between the processor and the CNN accelerator can also be used to transfer parameter types, and since the model in this paper has been quantized in the previous section, only INT types and bool types are transferred.

4.2. CNN Accelerator

4.2.1. Overall Architecture of the CNN Accelerator. The internal design of the CNN accelerator is shown in Figure 6. To enable the full functionality of the network structure in YOLOv3, the IP cores in the CNN accelerator contain: convolutional layer IP, UpSampling layer IP, YOLO layer IP,

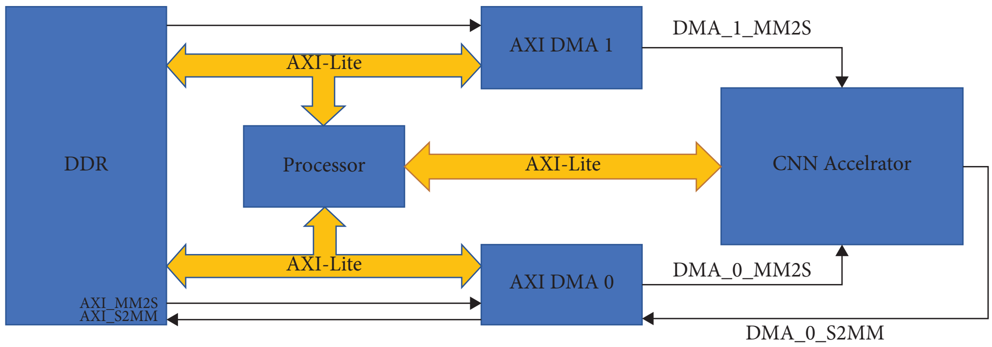


Figure 5: System architecture.

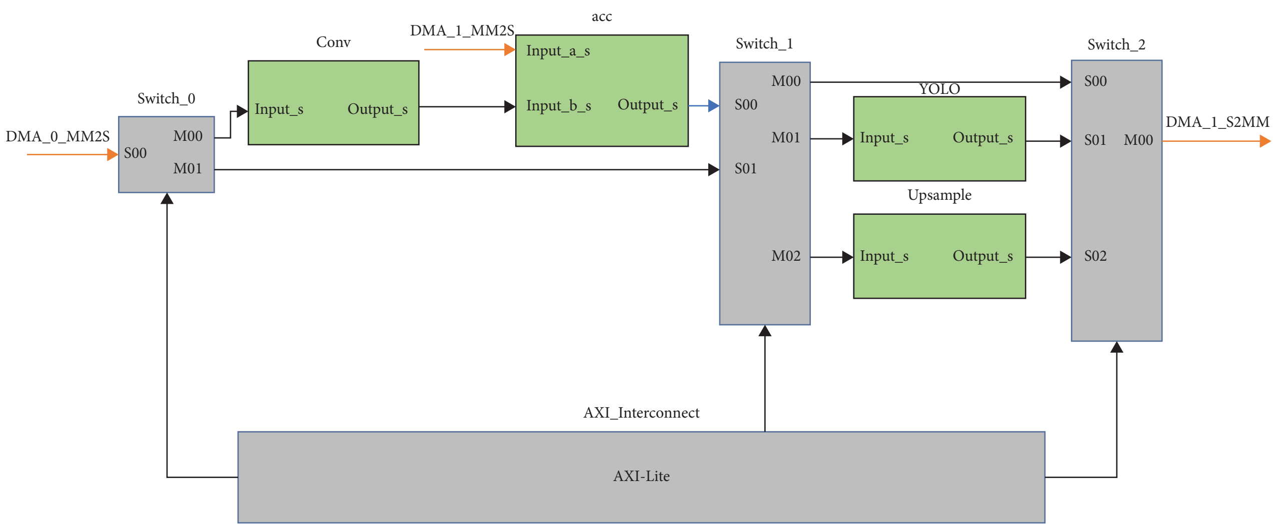


Figure 6: CNN accelerator design.

and acc IP. First, the convolutional layer IP core obtains the weights and input feature maps from DDR by DMA and outputs the output feature maps to the acc IP after convolutional operation. The acc IP implements the calculation function of activation and hopping function. Switches 0 and 1 control whether the DMA0 input flows into the convolutional IP or into other IP cores. Also, switches 1 and 2 determine which layer the output of the convolution operation should be placed on, and these network layers can be YOLO layers, Up-Sampling layers, or direct outputs without any network layers. Finally, the output of the whole process is saved in the main memory by DMA. Among them, since the detection system designed by FPGA only uses the inference part of YOLOv3, this paper chooses to directly merge the BN layer in the convolution operation. Therefore, the IP core of the convolutional layer in this paper is the same as the DBL structure in the YOLOv3 model. Since the minimum structure in YOLOv3 is the DBL structure, such a design can satisfy the whole YOLOv3 deployment on FPGAs.

4.2.2. Convolutional Layer IP Core. Figure 7 shows the internal design of the convolutional layer IP. The interface of the convolutional layer IP is divided into a data port and a control port. The data port is responsible for transmitting the weights, input feature maps, and output feature maps of the data stream. Before calling the convolution kernel IP, the control logic determines whether the input is a feature map or a weight file based on the parameters about the input type transmitted by the C++ program on the PS. After judging, if it is a feature map, it will be stored to buffer 1. If it is a weight file, it will be saved directly into the internal memory in the convolution kernel. If it is a feature map, the corresponding output will be obtained after convolution performed by multiple convolution kernel modules.

Finally, the result is output to the on-chip DDR in the form of AXI-Stream. According to the flow in Figure 5, the loop continues between the off-chip memory and the FPGA until the whole network operation is finished.

Since the convolution kernel in YOLOv3 needs to input multiple rows of data at the same time in one operation, and the convolution kernel slides in a left-to-right, top-to-bottom direction. In this paper, we choose to access the input in one convolution operation as a line buffer. The convolutional kernel modules in Figure 7 are all internally cached with the line buffer as shown in Figure 8. During the top-down sliding of the convolution kernel, two rows of data are reused for each per-row sliding. For each input image, an array of shift registers with the same column and the same row as the size of the convolution kernel is constructed. Whenever the register

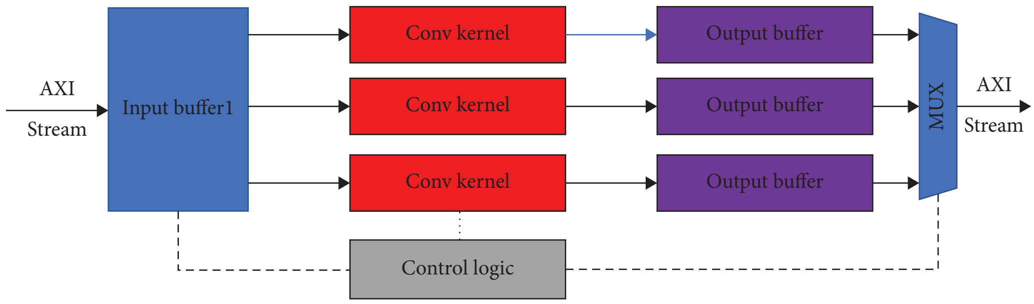


Figure 7: Design of convolutional IP.

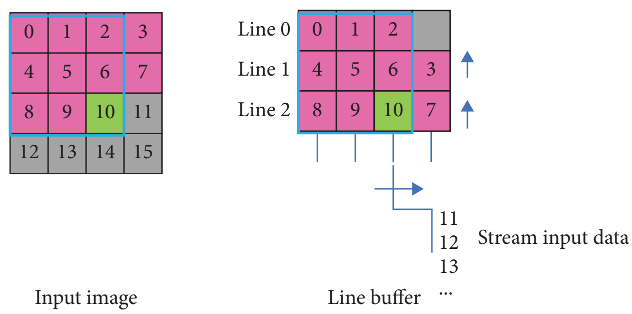


Figure 8: Data allocation of line buffer.

gets an input, the data previously stored in the corresponding column of the input are shifted upward, and then the new number is filled in the vacant position, so that each time the convolution kernel is shifted by row, only the data with the same size of the convolution kernel width need to be filled. Since the size of the convolution kernel in YOLOv3 is only 3×3, and the maximum column size is 418 (the initial input size is 416×416, but the input size will change to 418×418 because of the zero-padding operation during the convolution operation), the maximum row buffer created for each convolution operation is 3×418.

Figure 9 shows the structure inside the convolution kernel module. For a feature map with N number of input channels, N row buffers are created inside the module, and each row buffer corresponds to one channel of the feature map. The sliding window corresponds to the shift window mechanism used in the convolution process to save the 3×3 matrix in the row buffer that needs to be convolved to calculate the 3×3 matrix. C represents the convolutional unit, and the weights and corresponding biases of the convolutional matrix are stored in C. When the sliding window part acquires the matrix to be computed, N convolutional units are computed in parallel. The SUM module is responsible for adding up all the results of the computation. The output results are stored in the output buffer and written to memory at once after the entire feature map is computed.

The internal structure of the convolutional unit C is shown in Figure 10. The convolution unit C expands the input 3×3 feature map matrix and the weight matrix separately. The corresponding multiplications are calculated by nine multipliers in parallel, and the sum of all multiplications is calculated by eight address. The last adder is used to add the bias to the calculated result, and the final value obtained is the result value of one convolution operation.

4.2.3. *The acc IP Core.*

The acc IP core implements the residual part of YOLOv3 and the activation function calculation after the convolution layer. Figure 11 shows the specific design of the acc layer. Since the residual part needs to use the result of the previous convolutional layer operation, it needs to use two input ports, and the output port is still one. The control logic first determines if there are two input streams, if AXI stream b is empty, the accumulation module outputs AXI stream b directly without any operation. If AXI stream b is not empty, it outputs input a and input b after accumulation operation. Since all convolutional layers in YOLOv3 go through the activation function calculation, the output of the accumulation module goes directly to the activation module for the activation function calculation, and the output is still output in the form of AXI-Stream.

4.2.4. *Upsampling Layer IP Core.*

The operation of the upsampling layer is the reverse operation of the pooling layer. For each input data, the upsampling layer in YOLOv3 is expanded into a window of size 2×2, where the upper left corner of the window is the original data and the data in the rest of the window are the same as the original data. The buffers are still line buffer, but for each input, it is first stored inside the line buffer, and then the portion without input of the window is filled. So the maximum line buffer size created by the upper sampling layer IP core is 418×2.

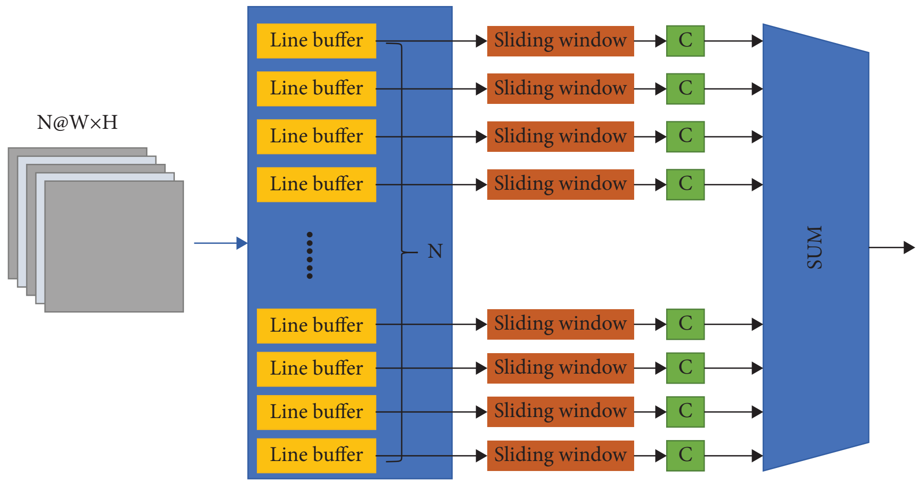


FIGURE 9: Structure of conv kernel.

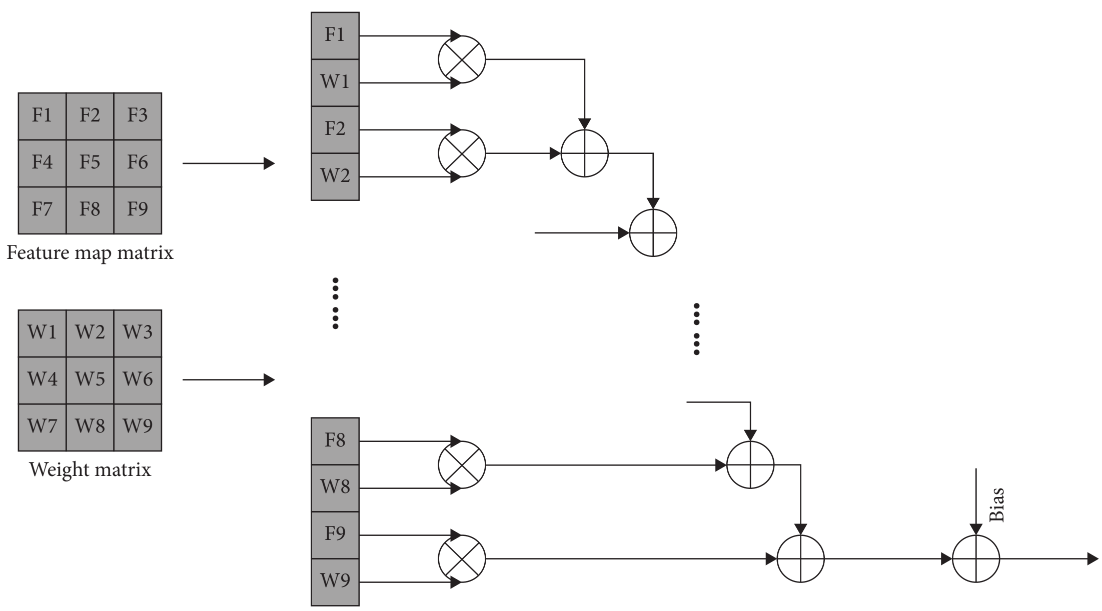


FIGURE 10: Structure of convolutional calculation unit.

4.2.5. YOLO Layer IP Core. The YOLO layer is used to transform the output feature maps derived from the entire YOLOv3 network into target detection result for different bounding box sizes. The YOLO layer transforms the input into a $W \times H \times C$ size input, that is, the original image is divided into $W \times H$ grids, where the channel C is equal to $(4 + 1 + \text{Class}) \times \text{Numbers}$. Class and Numbers represent the total number of classes predicted by the network and the number of objects predicted by each grid. In this paper, $\text{Class} = 6$, $\text{Numbers} = 3$, $W = H = 13, 26, 52$.

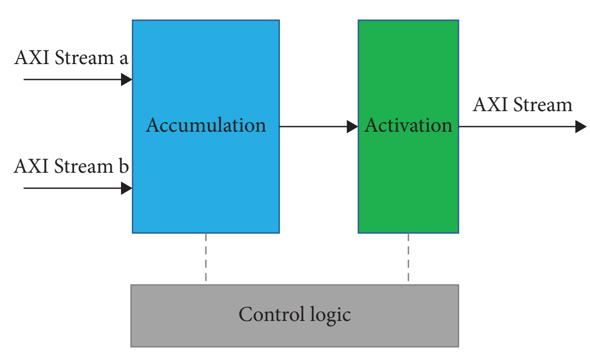


FIGURE 11: Design of acc IP.

5. Experiments and Results

In this paper, we use ZYNQ7020 to complete the implementation of YOLOv3 acceleration system, and the chip model is 7Z020-2CLG484I. HDMI driver is used to output the detection results on a $1920 * 1080$ resolution screen in real time. For image acquisition, we used XSP-1CA electron microscope to capture the lens video with a frame rate of 60 FPS. The YOLOv3 network was trained and optimized on an NVIDIA GeForce RTX 2080Ti graphics card.

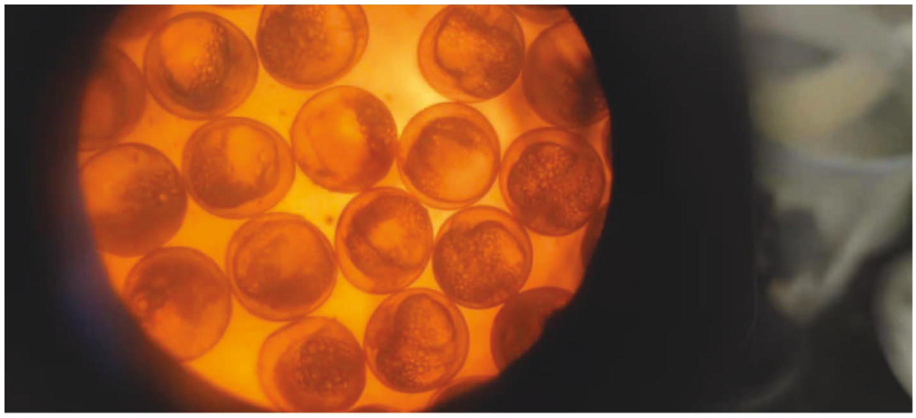


Figure 12: Original image under the microscope.

5.1. Fish Embryo Data set. Since this paper aims at accurate real-time detection of fish embryos, a large number of images of fish embryos are needed to pretrain the network. Therefore, in this paper, images of fish embryos were collected in batches using an electronic microscope in pools 1–10 of phase 2 of the intensive puffer fish farm in Zhongyang, Nantong, Jiangsu Province (120.9°E, 32.6°N). Figure 12 shows a frame from the microscope acquisition video.

As can be seen, the features of the embryo in the figure are more difficult to be extracted directly by the network model and also more difficult to be recognized directly in the system, so it is not advisable to use the YOLOv3 model directly to train the whole image captured by the electron microscope. In this paper, we chose to extract individual embryo images using electron microscope magnification and use LabelImg to annotate each target embryo and make a data set in VOC format. In this paper, we choose to extract individual embryo images after zooming in and use LabelImg to mark each target embryo with annotations (Annotations) to produce a data set in VOC format. A total of 5710 images were collected and divided into training set, test set, and validation set in the ratio of 8 : 1 : 1. Among them, 6 types of embryos are included, respectively, dead, gastrul, neurula, caryolytes, ocular, and olfactory. The individual embryo images in the data set are shown in Table 1. In fact, in the results of embryo detection, the detection accuracy of dead embryos is required to be the highest. In this paper, more than twice as many dead embryos were selected as the training set in the selection of data set for ensuring the accuracy of dead embryos in the detection. Figure 13 shows the embryo images of 6 periods in our data set.

5.2. Embryo Data Set in Different Target Detection Algorithms. As shown in Table 2, in the YOLO family of algorithms, the mAP value of YOLOv3 is smaller than that of YOLOv4 and YOLOv5. But its inference speed is about twice that of YOLOv4 and four times that of YOLOv5. Since the mAP value reflects the degree of overlap between the detection frame and the label, a small loss is acceptable for this application. mAP value and inference speed of Fast-RCNN are both inferior to YOLOv3. MobileNetV3 has the best performance in terms of inference speed, but its accuracy loss is too large. In summary, we choose YOLOv3 as our porting model.

5.3. Basic Training of the YOLOv3 Model. For basic training of the model, we pretrained the model using a 20-class VOC2007 data set in order to enhance the generalization ability of the network model and to speed up the convergence of the model. In basic training and pretraining, the model is saved once for every 100 epochs trained, the learning rate is 0.01, the batch size is 32, the pretraining is 50 epochs, and the basic training is 500 epochs. The pretrained obtained model is trained using the data of embryos. The loss curves of the training process are shown in Figure 14(a). Figure 14(b) shows the loss curves of the training process without pretraining and directly training the model using the embryo data set. Figure 14(b) shows the loss curves of the training process for the model trained directly using the embryo data set after no pretraining.

In Figure 14(a), the model was trained until 500 epochs without convergence. In Figure 14(b), the model has completed convergence by the 300th epoch of training, so pretraining clearly improves the convergence rate of model training. The loss values and mAP values after the model training are completed in both cases and are given in Table 3.

As shown in Table 3, the mAP values and loss values obtained from pretraining are all superior compared to the model without pretraining. Thus, this paper uses VOC2007 data for pretraining before basic training. The trained model achieves 86.72% mAP and 1.922 loss values on the test data set. This indicates that the model has far surpassed the accuracy of traditional manual identification methods for six embryonic periods: dead, gastrul, neurula, caryolytes, ocular, and olfactory. However, the trained model only has an inference speed of 12 fps on the GPU, which is obviously not sufficient for the task of real-time detection and thus requires optimization of the model.

5.4. Model Pruning of the YOLOv3 Model

5.4.1. Channel Pruning. After completing the training of YOLOv3, the optimal model from the base training is sparsely trained according to Equation (1) to facilitate

Table 1: Fish embryo data set.

	Dead	Gastrul	Neurula	Caryolytes	Ocular	Olfactory
Train	1528	896	912	712	814	848
Test	191	112	114	89	101	106
Validation	191	112	114	89	101	106

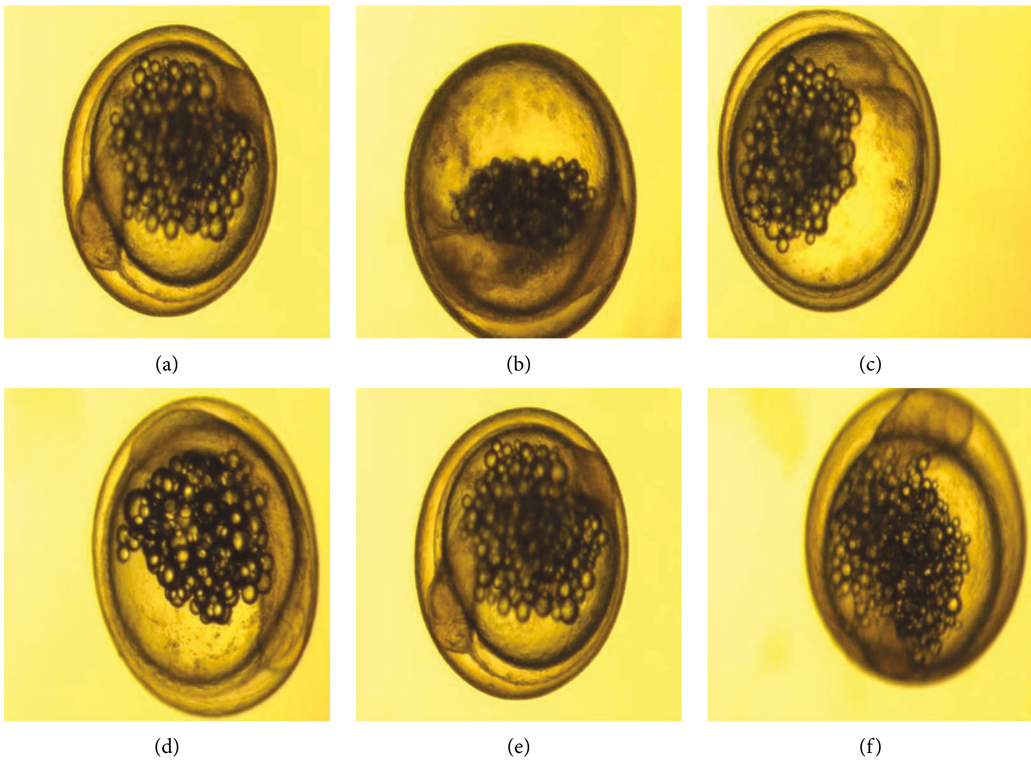


Figure 13: Embryo image: (a) dead embryo, (b) gastrul embryo, (c) neurula embryo, (d) embryo, (e) ocular embryo, and (f) olfactory embryo.

Table 2: Performance of several models trained by the embryo data set.

Network	mAP (%)	FPS
YOLOv3	79.93	12.1
YOLOv4	81.22	6.8
YOLOv5	84.51	3.1
Fast-RCNN	76.14	6.4
MobileNetV3	54.33	14.2

channel pruning and layer pruning of the model. The distribution of the corresponding γ parameters in the BN layer during sparse training and basic training is shown in Figure 15, and all four plots in Figure 15 are histograms. The vertical axis in Figures 15(a) and 15(b) plots the number of rounds of model training, the horizontal axis indicates the value of the γ parameter, and the height of the histogram represents the number of γ parameters at this size. As shown in Figure 15(a), in the basic training, as the number of training rounds increases, the distribution of γ parameters still resembles a normal distribution, with the vast majority of parameters in the range of 0.5–1.5. This distribution makes it possible that most of the channels will still remain after channel pruning, and it will make the mean values of γ parameters required for subsequent layer pruning almost all the same, making the process of layer pruning of the model difficult.

Figure 15(b) shows the changes in the distribution of γ parameters during sparse training, in which most of the parameters gradually decrease to close to 0 as the number of training epochs gradually increases, as a way to filter out the channels that can be pruned. Plots (c) and (d) represent the distribution of γ parameters in each DBL structure in the model after basic training and sparse training, respectively. The vertical axis indicates the number of layers in which the specific DBL in its model is located, and the horizontal axis is the size of the γ parameter, and its histogram height indicates the number of γ parameters at this size. After the base training, the parameter distribution of the model is very heterogeneous, and most of the parameter distributions are in the range of 0.5–1.5, making it difficult to perform channel pruning for each layer. In contrast, after sparse training, most of the parameters are close to with 0. This distribution is more convenient for channel pruning. After sparse training, all channels with γ parameters less than 10^{-3} are removed in this paper.

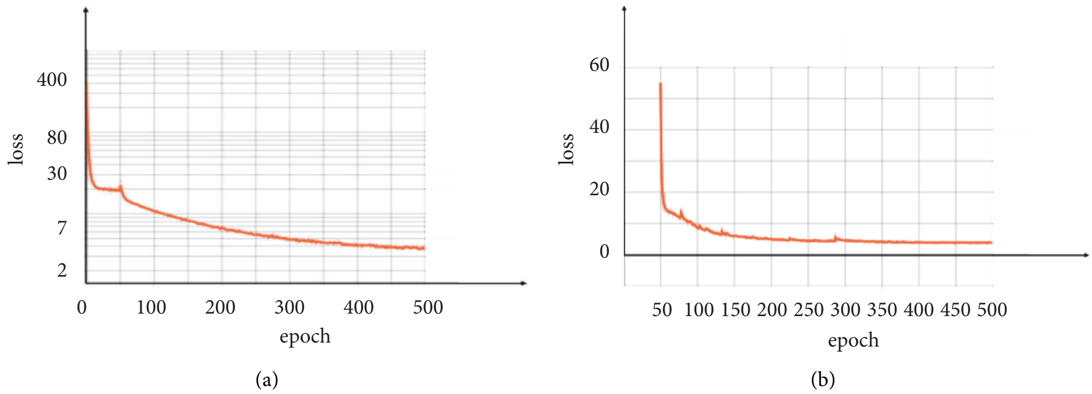


FIGURE 14: Loss curve in training. (a) Loss without pretraining. (b) Loss with pretraining.

5.4.2. Layer Pruning. After channel pruning, the mean values of all the DBL structural γ parameters in the model need to be calculated and layer pruning is performed. Fine-tuning is performed by secondary training after layer pruning. Since the model needs to be deployed to hardware for practical application in this paper, it is necessary to increase the inference speed of the model as much as possible on the basis of ensuring the accuracy of the model. In this paper, we start with 10% pruning rate in increments of 10% and gradually increase the pruning rate. The relationship between mAP value, output video FPS and pruning rate is shown in Figure 16. When the pruning rate of the model exceeds 60%, the mAP value decreases dramatically, and this change cannot be corrected by retraining. Because the model eventually needs to be deployed to the hardware, a balance point of inference speed and accuracy is chosen, and the pruning rate chosen in this paper is 50%. That is, the network layers with the last 50% of the γ parameter size ranking are deleted. It is equivalent to subtracting 26 DBLs (i.e., subtracting 26 convolutional layers) to reduce the model by about half of the parameters and computational effort.

Since there was a drop in model accuracy after pruning, this paper chooses to fine-tune the model using the original data set as the fine-tuning data set. After retraining 50 epochs with the original training set at 1%, the mAP value of the model on the test set rebounded to 86.14%, where the model size was reduced to 110.4 MB and the video inference speed on GPU could reach 17.4 FPS.

5.5. Model Quantization of the YOLOv3 Model. After the model pruning, the 32 bit floating-point parameters of the model were chosen to be quantized into INT4 and INT8 models all through Algorithm 1. The quantization results are shown in Table 4.

The data in Table 4 are the individual performance metrics of the model quantized by Algorithm 1 for INT8 quantization and INT4 quantization, respectively, the mAP value, the model size, and the FPS at inference. The data in Table 4 are the individual performance metrics of the model quantized by Algorithm 1 for INT8 quantization and INT4 quantization, respectively, the mAP value, the model size, and the FPS at inference. It can be seen that the quantization method of Algorithm 1 can significantly improve the inference speed of the model while ensuring little change in model accuracy. The INT8 quantized model achieves 83.52% mAP and 33.9 fps on the GPU. The INT4 quantized model achieves 82.49% mAP on the GPU and 50.3 fps on the GPU, which is a 32.9 fps improvement over the pruned model. After quantization by INT4 and INT8 of Algorithm 1, the two quantization methods only have a difference of about 1.03% in mAP. In fact, in the actual detection process, about 1% of mAP may make 1 out of 1000 embryos misclassified, which has a negligible impact on the detection accuracy of the actual system application, but the network model after INT4 quantization is nearly 17 fps higher in GPU inference speed of processing. In terms of model size, INT4 quantization can reduce the model size to 1/8 of the original size, making it easier to store the model. Since this paper has the requirement of real-time detection for the system deploying YOLOv3 on FPGA, all the YOLOv3 models are quantized into 4 bit fixed points.

5.6. FPGA Implementation of Optimized YOLOv3. After finishing the optimization of the whole YOLOv3 model, the weight parameters of the optimized model are stored into the DDR at the PS side of ZYNQ in advance. According to the content in Section 3.2, we package the IP core in Vivado and connect it using AXI bus protocol to complete the hardware implementation of the YOLOv3 model. Figure 17 shows the specific design of the YOLOv3 gas pedal in Vivado corresponding to Figure 6. The hardware acceleration is accomplished by splitting each YOLOv3 network layer into its corresponding network structure and passing in weight parameters at the time of use. In which, as designed in Chapter 3, the convolutional layer passes through switch 1 and selectively enters the subsequent yolo layer or upsample layer, where the acc IP core after the conv IP core is used to handle the accumulation operation of the results output from the two different network layers shown in Figure 7 and to implement the activation function LeakyRelu.

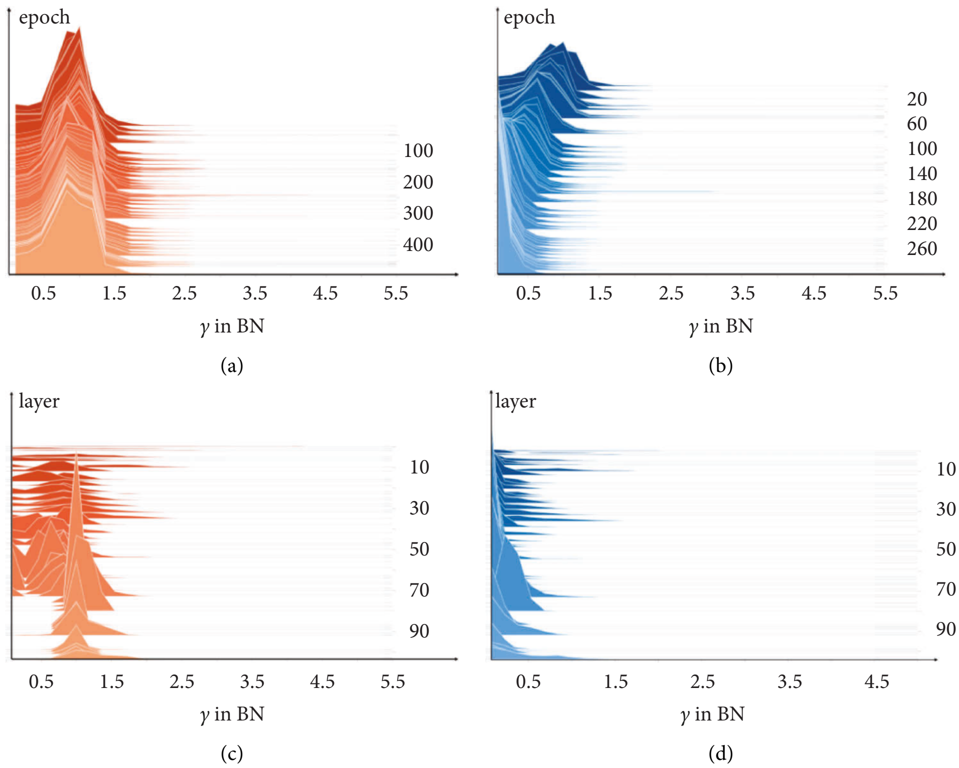


FIGURE 15: γ parameter distribution. (a) Distribution of γ parameters during basic training, (b) distribution of γ parameters during sparse training, (c) distribution of all γ parameters of the whole model after basic training is completed, and (d) distribution of all γ parameters of the whole model after sparse training is completed.

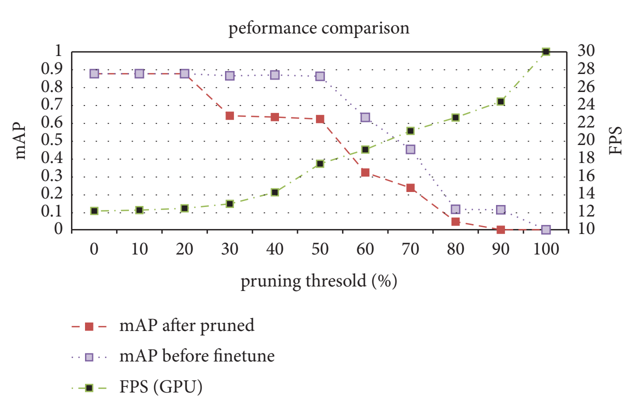


FIGURE 16: mAP and FPS in different model pruning thresholds.

TABLE 3: mAP and loss comparison.

	Loss	mAP (%)
Pretrained model	1.922	86.72
Without pretrained model	3.4368	79.93

Table 5 shows the resource utilization of the YOLOv3 gas pedal system built in this paper on the FPGA platform. In terms of resource utilization, the resource utilization of LUT, BRAM, DSP, and FF in this paper is 51.78%, 82.27%, 66.07%, and 47.75%, respectively. The LUT, DSP, and FF parts are used to process numerical and logical operations. The BRAM is used to store feature maps and weight inputs. As can be seen, the YOLOv3 accelerator in this paper is built using almost all the resources on the chip in order to achieve the maximum inference speed. Such a hardware architecture can complete the inference process of YOLOv3 more quickly.

The performance of the detection model on different types of embedded devices is given in Table 6, where we use CUDA10.1 with Tensorflow2.4 to accelerate the inference speed of the YOLOv3 optimized model on RTX-2080Ti and ARM, and the ARM architecture uses the Cortex-A9 dual-core ARM currently used by most

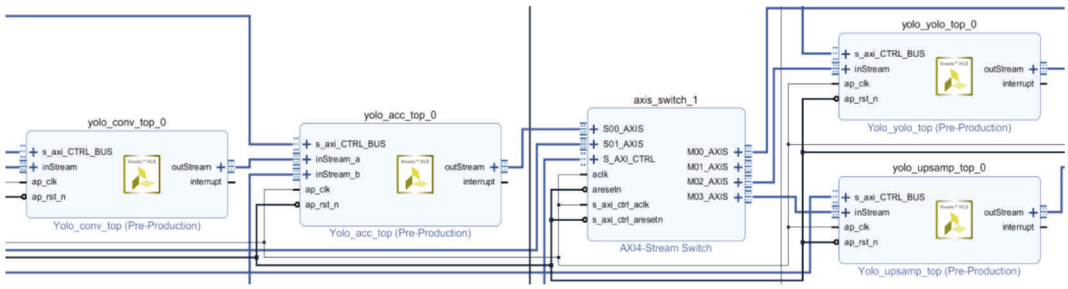


Figure 17: YOLOv3 accelerator implementation in Vivado.

Table 4: Indicators of the quantization model.

	mAP (%)	Model size (MB)	FPS on GPU
Pruned model	86.14	110.4	17.4
INT8 quantization	83.52	24.9	33.9
INT4 quantization	82.49	12.8	50.3

Table 5: Performance of the embryo detection model on different types of embedded devices.

	LUT	BRAM	DSP	FF
Used	27548	50801	181	50801
Available	53200	106400	220	106400
Utilization	51.78%	66.07%	82.27%	47.75%

Table 6: Performance in different embedded devices.

Device	Power (w)	FPS	Processing time per image (ms)
RTX-2080Ti	230	50.3	7.9
ARM Cortex-A9 dual-core	1.6	2.7	246
ZYNQ-7020	2.524	34.1	12.1

embedded systems. FPGA hardware acceleration is used to improve the model inference speed. Since the optimized model is used on all devices in this paper for comparison on the same test set, there will be no gap in the mAP value reflecting the accuracy of the model itself. The FPGA accelerated system built on ZYNQ in this paper can reach about 5 times the inference speed of pure CPU inference, although the power consumption is slightly higher than that of the inference system with ARM for the model. With 24 frames as the real-time detection standard, this system steadily achieves the effect of real-time inference. Compared with GPU, this paper lacks about 50% of the image inference speed. However, in terms of application, the power consumption of GPU reaches nearly 92 times of the power consumption in this paper compared to this paper, and it is impossible to meet this power consumption demand in actual aquaculture deployment. For embedded deployment, the system in this paper not only saves cost compared to traditional GPU + CPU and pure CPU for aquaculture but also satisfies basically all fish embryo detection needs at the same time.

5.7. Result. For each input embryo image acquired from the microscope, we scaled it to the input size of 416×416 required by the optimized YOLOv3 model. After input to the optimized YOLOv3 model, based on the obtained 3 feature map results, the nonmaximal value suppression algorithm was used on the PS side to eliminate the duplicate detection results and get the most accurate results. Some sample embryo detection images are shown in Figure 18.

The corresponding FPGA on-chip real-time detection system effect and system structure are shown in Figure 19, the whole real-time detection system consists of three parts, namely ZYNQ chip responsible for CPU control and YOLOv3 acceleration, electronic microscope responsible for embryo data acquisition, and HDMI display responsible for image output.

The experimental results show that the system is able to acquire the results of embryo detection correctly and in real time. The inference speed of 34.1 fps on average is achieved with 2.524w of power consumption at 100 MHz FPGA operating frequency, while 82.49% of mAP value can be achieved in detection accuracy. Compared with the

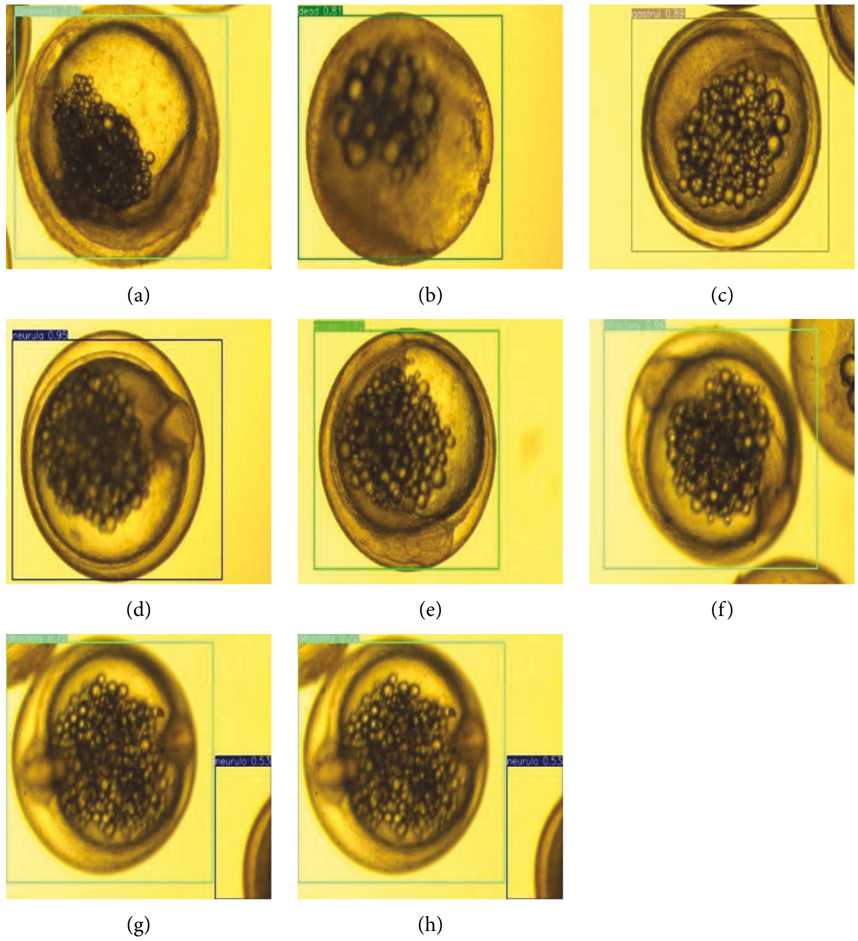


Figure 18: Embryo detection by the optimized model. (a) Caryolites, (b) dead, (c) gastrul, (d) neurula, (e) ocular, and (f–h) olfactory.

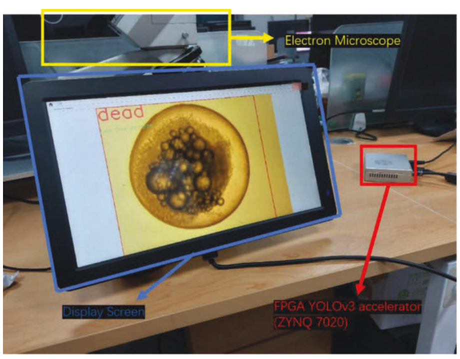


Figure 19: Real-time fish embryo detection system.

original model deployed directly with CPU + GPU, the mAP value is only 4.23% lower, but the real-time detection can be completed with nearly 1/90th of the power consumption.

6. Conclusion

In the paper, an FPGA-based system for fish embryo detection is proposed. Based on using YOLOv3 as the basic model, the model itself is optimized by combining a model pruning method based on scaling factors and a quantization method with trainable thresholds, and the optimized model is deployed on an FPGA platform to realize a real-time fish embryo detection system. Compared with the direct use of the original YOLOv3 model, a better compromise is achieved in terms of inference speed and inference accuracy, and the on-chip computing resources are utilized to the maximum. This FPGA-based fish embryo detection system fully satisfies the current demand for embedded devices in fish farming and facilitates the shift from the use of high-volume experienced staff to artificially intelligent embedded devices for embryo detection and rapid analysis of embryos in farming. The system has been applied in the embryo farming of Zhongyang Pufferfish Estate in Nantong, Jiangsu Province. It has achieved 34.1 fps with 82.49% mAP value for real-time accurate embryo detection at 2.524w power consumption.

Data Availability

All data used in this study are included in the manuscript.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This research was funded by Guangdong Key Areas R&D Program Projects, under grant no. 2021B0202070001.

References

[1] J. Chen, S. Fangfang, J. Zhang, and F. Yongjie, "Zhao Daoquan," *Jiangsu Agricultural Science*, vol. 49, p. 45, 2021.
[2] M. Ashaf-Ud-Doulah, S. M. M. Islam, M. M. Zahangir, M. S. Islam, C. Brown, and M. Shahjahan, "Increased water temperature interrupts embryonic and larval development of Indian major carp rohu Labeo rohita," *Aquaculture International*, vol. 29, no. 2, pp. 711–722, 2021.
[3] L. Ma, R. Dessiatoun, J. Shi, and W. R. Jeffery, "Incremental temperature changes for maximal breeding and spawning in Astyanax mexicanus" Jove-Journal of Visualized Experiments," 2021.
[4] S. E. Cordova-de la Cruz, G. Martinez-Bautista, E. S. Pena-Marin et al., "Morphological and cardiac alterations after crude oil exposure in the early-life stages of the tropical gar (Atractosteus tropicus)," *Environmental Science and Pollution Research*, vol. 29, no. 15, pp. 22281–22292, 2021.
[5] R. L. Naylor, R. W. Hardy, A. H. Buschmann et al., "A 20-year retrospective review of global aquaculture," *Nature*, vol. 591, no. 7851, pp. 551–563, 2021.
[6] J. Clarke, "Live imaging of development in fish embryos," *Seminars in Cell & Developmental Biology*, vol. 20, no. 8, pp. 942–946, 2009.
[7] O. Ishaq, S. K. Sadanandan, and C. Wahlby, "Deep fish," *SLAS Discovery*, vol. 22, no. 1, pp. 102–107, 2017.
[8] H. T. Rauf, M. I. U. Lali, S. Zahoor, S. Z. H. Shah, A. U. Rehman, and S. A. C. Bukhari, "Visual features based automated identification of fish species using deep convolutional neural networks," *Computers and Electronics in Agriculture*, vol. 167, p. 105075, 2019.
[9] A. M. Naderi, H. Bu, J. Su et al., "Deep learning-based framework for cardiac function assessment in embryonic zebrafish from heart beating videos," *Computers in Biology and Medicine*, vol. 135, Article ID 104565, 2021.
[10] P. Dey, "The emerging role of deep learning in cytology," *Cytopathology*, vol. 32, no. 2, pp. 154–160, 2021.
[11] B. H. Tang, Z. X. Pan, K. Yin, and A. Khateeb, "Recent advances of deep learning in bioinformatics and computational biology," *Frontiers in Genetics*, vol. 10, p. 214, 2019.
[12] Y. Li, Z. G. Luo, N. Y. Guan et al., *Progress in Biochemistry and Biophysics*, vol. 43, p. 472, 2016.
[13] C. Qian, M. Tong, X. Yu, and S. Zhuang, "CNN-based visual processing approach for biological sample microinjection systems," *Neurocomputing*, vol. 459, pp. 70–80, 2021.
[14] Q. Xu, Z. Wang, F. Wang, and Y. Gong, "Multi-feature fusion CNNs for Drosophila embryo of interest detection," *Physica A: Statistical Mechanics and Its Applications*, vol. 531, p. 121808, 2019.
[15] D. Dirvanauskas, R. Maskeliunas, V. Raudonis, and R. Damasevicius, "Embryo development stage prediction algorithm for automated time lapse incubators," *Computer Methods and Programs in Biomedicine*, vol. 177, pp. 161–174, 2019.
[16] I. Kuon, R. Tessier, and J. Rose, "FPGA architecture: survey and challenges," *Foundations and Trends® in Electronic Design Automation*, vol. 2, pp. 135–253, 2007.
[17] M. Xia, Z. Huang, L. Tian et al., "SparkNoC: an energy-efficiency FPGA-based accelerator using optimized lightweight CNN for edge computing," *Journal of Systems Architecture*, vol. 115, Article ID 101991, 2021.
[18] N. Y. Khanday and S. A. Sofi, "Taxonomy, state-of-the-art, challenges and applications of visual understanding: a review," *Computer Science Review*, vol. 40, Article ID 100374, 2021.
[19] A. G. Blaiech, K. Ben Khalifa, C. Valderrama, M. A. C. Fernandes, and M. H. Bedoui, "A survey and taxonomy of FPGA-based deep learning accelerators," *Journal of Systems Architecture*, vol. 98, pp. 331–345, 2019.
[20] T. Li, T. Zhang, G. Yu, J. Song, and J. Fan, "Minimizing temperature and energy of real-time applications with precedence constraints on heterogeneous MPSoC systems," *Journal of Systems Architecture*, vol. 98, pp. 79–91, 2019.
[21] V. R. S. Mani, A. Saravanaselvan, and N. Arumugam, "Performance comparison of CNN, QNN and BNN deep neural networks for real-time object detection using ZYNQ FPGA node," *Microelectronics Journal*, vol. 119, Article ID 105319, 2022.
[22] V. Sharma and R. N. Mir, "A comprehensive and systematic look up into deep learning based object detection techniques: a review," *Computer Science Review*, vol. 38, Article ID 100301, 2020.
[23] I. Damaj, S. K. Al Khatib, T. Naous, W. Lawand, Z. Z. Abdelrazzak, and H. T. Mouftah, "Journal of King Saud University - Computer and Information Sciences," *Open access journal*, vol. 12, 2021.
[24] M. Zhao, C. Hu, F. Wei, K. Wang, C. Wang, and Y. Jiang, "Real-time underwater image recognition with FPGA embedded system for convolutional neural network," *Sensors*, vol. 19, no. 2, p. 350, 2019.
[25] X. Wei, W. Liu, L. Chen, L. Ma, H. Chen, and Y. Zhuang, "FPGA-based hybrid-type implementation of quantized neural networks for remote sensing applications," *Sensors*, vol. 19, no. 4, p. 924, 2019.
[26] Y. Luo and Y. Chen, "FPGA-based acceleration on additive manufacturing defects inspection," *Sensors*, vol. 21, no. 6, p. 2123, 2021.
[27] M. R. Al Koutayni, V. Rybalkin, J. Malik et al., "Real-time energy efficient hand pose estimation: a case study," *Sensors*, vol. 20, no. 10, p. 2828, 2020.
[28] A. F. J. Redmon, "YOLOv3: An Incremental Improvement", https://arxiv.org/pdf/1804.02767.pdf, 2018.
[29] K. H. S. Ren, R. B. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks" https://arxiv.org/abs/1506.01497, 2015.
[30] A. G. H. Z. C. K. W. W. A. Adam, 2017 "MobileNets: efficient convolutional neural networks for mobile vision applications" https://arxiv.org/abs/1704.04861.
[31] E. A. Smirnov, D. M. Timoshenko, and S. N. Andrianov, "Comparison of regularization methods for ImageNet classification with deep convolutional neural networks," *AASRI Procedia*, vol. 6, pp. 89–94, 2014.
[32] C. Wu, V. Fresse, B. Suffran, and H. Konik, "Accelerating DNNs from local to virtualized FPGA in the Cloud: a survey of trends," *Journal of Systems Architecture*, vol. 119, p. 102257, 2021.
[33] A. G. Sambhav R Jain, M. Wu, and C. Dick, 2019, "Trained quantization thresholds for accurate and efficient fixed-point inference of deep neural networks" https://arxiv.org/abs/1903.08066.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 5623067, 11 pages
https://doi.org/10.1155/2022/5623067

Research Article

Motivation of Student Management Subjects in Universities Based on Network Intelligence

Like Xin,[1] Feng Wang,[2] Xiaodong Sun,[3] Guoqing Ding,[4] and Hui Wang [5]

[1] *College of Animal Science and Technology, Qingdao Agricultural University, Qingdao 266109, China*
[2] *Student Affairs Division, Qingdao Agricultural University, Qingdao 266109, China*
[3] *Science and Technology Division, Qingdao Agricultural University, Qingdao 266109, China*
[4] *College of Veterinary Medicine, Qingdao Agricultural University, Qingdao 266109, China*
[5] *College of Economics and Management, Qingdao Agricultural University, Qingdao 266109, China*

Correspondence should be addressed to Hui Wang; huieeking@qau.edu.cn

Received 2 August 2022; Revised 29 August 2022; Accepted 5 September 2022; Published 6 October 2022

Academic Editor: Jiafu Su

With the intelligent development of the society and the development of the big data era, the network intelligence applications in various fields have deeply penetrated them, and colleges and universities, as the place where information technology is most widely used, have been profoundly affected by the network intelligence. In this context, besides the traditional use of digitalization, informatization, and networking to assist in managing students' studies and life, the work of the main body of student management in colleges and universities is also gradually developing towards intelligence. In the research of this study, through the application and development of network intelligence, the intelligent research of college student management is carried out by using the technical advantages of massive data access, so as to improve the motivation of college student management subjects through intelligence, make students feel the orderly and reliable humanized management while improving the efficiency of student management, and provide sufficient technological guarantee for college management.

1. Introduction

The level of technology in student management directly determines the quality of future talent training, and the move towards intelligent student management is both a challenge and an opportunity [1]. Intelligent management will further improve efficiency and allow students to enjoy the convenience of technology and feel the orderly stability and humanity of management, which will bring benefits to their physical and mental health, the overall development of personal qualities, and innovative thinking. At present, the problems of intelligent student management work under the vision of big data are as follows: the data information management system is not yet fully unified, and there is the phenomenon of information silo [2]; the intelligent data monitoring is not strict enough; there are problems such as low level of quantitative analysis and utilization and the lack of high-level technical personnel, which seriously affect and restrict the process of intelligent student management in colleges and universities.

The development of intelligence is both an opportunity for the external community to promote student management by virtue of intelligent research and development and human-machine collaboration and a challenge for internal student management by the strategic goal of intelligence and increased student self-awareness. The advantage of intelligent student management is the use of artificial intelligence to avoid the error rate of manual labor and thus improve efficiency [3]. The development of intelligent student management under the Internet is an effective result of data mining and analysis, but it is also a change in the way of life and thinking. Since intelligent technology is a new and general technology, there are hardware facilities and soft resources and other conditions restrictions, as well as the establishment of the relevant student management system, all of which have posed challenges to the intelligent development of student management to a certain extent.

Based on the concept of "application-driven, connected everything," the intelligent network provides more initiative, adaptability, and diversity for enterprise business applications. It solves the horizontal connection problem by using "intelligent perception and life" to build a broad panoptic connection system. The "network reorganization and flexible expansion" approach is used to build a flexible network from the bottom up.

2. State of the Art

2.1. Intelligent Technology. Intelligence is the eternal pursuit of information technology development, and the main way to achieve this pursuit is to develop artificial intelligence technology [4]. In the past 60 years, since the birth of artificial intelligence technology, although it has experienced three ups and two downs, it has made great achievements. 1959–1976 was the stage based on artificial knowledge representation and symbolic processing, which produced expert systems with important applications in some fields [5]; 1976–2007 was the stage based on statistical learning and knowledge self-representation, which produced various neural network systems [6]; in recent years, the stage based on environmental self-adaptation, self-gaming, self-evolution, and self-learning has started. The research based on environmental self-adaptation, self-gaming, self-evolution, and self-learning, which started in recent years, is forming a new stage of AI development, meta-learning or methodological learning stage, which constitutes a new generation of AI. The new generation of AI mainly includes big data intelligence, population intelligence, cross-media intelligence, human-machine hybrid augmented intelligence, and brain-like intelligence.

The world's major industrial countries have planned ahead. In the late 80s, smart manufacturing was proposed, and the world began to carry out a lot of research on it, starting with smart manufacturing technology, and later, with the development of the economy and the changing needs of society, people began to conduct in-depth research on smart manufacturing technology. In the world's major industrial powers, the implementation of "smart manufacturing" is an important means to revitalize the manufacturing industry. Since the financial crisis, the governments of the United States, Germany, and Japan, and relevant experts have proposed the development of intelligent manufacturing to revitalize the manufacturing industry. The United States officially launched the "Advanced Manufacturing Partnership" in June 2001, which includes industrial robots.

Intelligence reflects the quality attributes of an information product [7]. When we say an information product is intelligent, we usually mean that this product can accomplish what only an intelligent person can accomplish or has reached a level that only humans can achieve. Intelligence generally includes perception ability, memory and thinking ability, learning and self-adaptive ability, and behavioral decision-making ability [8]. Therefore, intelligence can usually be defined as making objects with sensitive and accurate perception functions, correct thinking and judgment function, adaptive learning function, effective execution function, etc. Network management is an inevitable product of the development of computer networks. Early networks were mainly local area networks, while the emergence of the Internet broke the geographical limitation of networks and the cross-regional wide area networks developed rapidly, at which time network management was no longer limited to ensuring the transmission of files, but to guaranteeing the normal operation of network objects connected to the network, as well as monitoring the operational performance of the network and optimizing the topology of the network. As IT business changes become more and more challenging [9], the work in the field of information technology becomes more and more complex, to manage a mixed network of various architectures, how network administrators set oil painting equipment and network configuration to give full play of the network system to the advantages, and with it, the emergence of intelligent network management system can take the pulse of the network, view the network connection relationship of the whole network, monitor the possible problems of various network devices in real time, detect where the bottlenecks of network performance are, and automatically handle or remotely repair them.

2.2. Problems of Intelligent Student Management. First of all, there is a lack of corresponding supporting intelligent development systems. Intelligent management of college students in the background of big data is in the initial stage, lacking practical national policy documents, and the rules and regulations of school administrators in this regard are also relatively old, not using the convenient and effective technology of intelligent management to drive the overall development of college students; meanwhile, there is the phenomenon of loose management and unknown rewards and punishments in colleges and universities, which leads to the learning process of students. The polarization is serious; the excellent students are always excellent, but the generally unconscious students are self-indulgent and uncontrolled, which is against the original purpose of educating people in colleges and universities.

Secondly, the Internet of Things is a new development direction and breakthrough on the basis of the Internet. Smart Earth refers to the full application of the new generation of IT and Internet technologies to all industries, and then, through supercomputers and cloud computing, it enables human beings to work and live in a more refined and dynamic way, thus enhancing the level of wisdom in the world. The concept of "Smart Earth" [10] further elaborates that the Internet of Things is interconnected, as shown in Figure 1. Through the intelligent sensing information network system of radio frequency identification, infrared sensing, positioning system, scanner, and other sensing facilities, the Internet of Things carries out network communication and information exchange, realizing the interconnection of dialogue and communication between things,

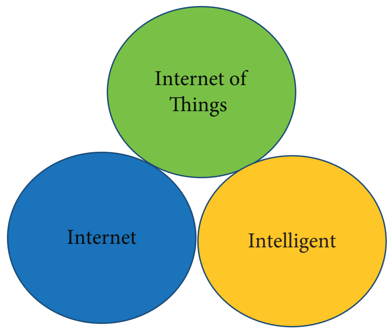


Figure 1: Three elements of smart Earth.

thus making the digital information era to the intelligent transition and upgrading. Colleges and universities are the concentrated place of talent training and cutting-edge technology application implementation and should strengthen the capital investment of IoT infrastructure [11], such as the increase of information collection by GPS module of student card in the construction of student management system, the research of information storage and transmission by the server on the Internet, and the construction of a platform for school management personnel or parents to inquire about student location information through the Internet.

Finally, as a group of higher information use, college students are affected by various fragmented information, such as all kinds of knowledge, microblogging, WeChat, public number on the network, various short articles, short videos, and other information of varying quality; students generally develop a kind of reading habit; it is difficult to form a systematic scientific knowledge system. In addition, because fragmented information tends to simplify complex issues, students can only recognize some appearances of things, but do not know their principles and inner essence and the relevance of other things. The development of informationization of student management directly affects whether students in colleges and universities in intelligent management can use data and information in high quality, fully grasp the opportunities of intelligent development, truly analyze the principles, backgrounds, and applications behind the phenomena, convert them into knowledge systems when searching and processing information, and then realize resource sharing through secondary fine processing [12].

2.3. Intelligent Construction Framework. Intelligent campus construction is a long-term, complex, and huge system project, which not only has a large implementation time span but also involves various organizational departments of the university, and must have a corresponding management structure and supporting policies to ensure smooth implementation and real benefits. The framework of intelligent campus construction is to form a digital teaching environment, research environment, and living environment with teachers and students as the core and around the three elements of an intelligent campus: resources, management, and services, relying on a data center and campus service center application system. Intelligent campus construction is a long-term construction process, involving research, overall planning, system construction, and maintenance improvement, and is a process of continuous improvement and enhancement, as shown in Figure 2.

According to the network condition of the campus network, the overall solution of campus network security should be established based on PKI security structure, and the overall authentication, transmission encryption, information encryption, data integrity protection, VPN technology, and a variety of security technologies such as firewall and secure proxy server should be an integrated solution [13], as shown in Table 1. The establishment of a PKI-based system can provide a foundation for network information security, but to truly realize the security of the whole system, it is necessary to gradually transform the existing application system so that authentication, identification, and audit are integrated into the details of each application.

PKI is a public key system. It uses the characteristics of public key algorithms to establish a system for certificate issuance, management, and use to ensure and achieve authentication, information encryption, data integrity, and denial prevention. The architecture, implementation methods, and communication protocols of PKI systems are all possible.

MALL security is only e-mail security; e-mail is obtained by attackers or tampering with e-mail, virus e-mail, spam, etc., which are seriously endangering the normal use of e-mail and even cause serious damage to computers and networks.

2.4. Intelligent Implementation Basis. The application of intelligence on the campus mainly responds to the following four aspects: first, intelligent monitoring system; second, intelligent classroom; third, one-card system; fourth, perimeter protection system. The campus computer network system is the operation platform of the "digital campus" and is an indispensable support environment for campus teaching [14], research and management systems, network multimedia teaching systems, campus cards, and other application systems. By adopting advanced and mature network technologies, such as 10-gigabit ethernet link backbone technology and sFlow traffic monitoring technology, we can build a broadband IP campus network with a high technical starting point and safe and reliable, manageable and maintainable, scalable, and providing comprehensive and distinctive services with certain service quality assurance.

According to the understanding of the security requirements of the school campus network and application system, we have established the following technical general framework for the specific design and implementation consideration of the network system project. (1) From the vertical consideration, information security is divided into network infrastructure, intelligent trust and authorization service layer, and security application layer; network infrastructure and intelligent trust and authorization service layer form the basis of the security application layer [15]. (2)

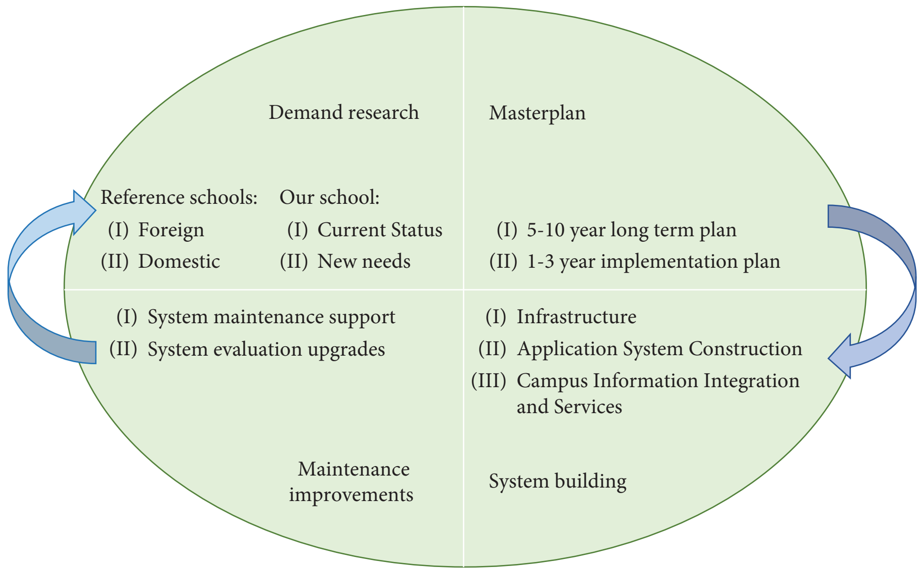


Figure 2: Campus intelligent construction process.

Table 1: Application authentication system.

Serial number	Targeted questions	Purpose of the application authentication system
1	Security isolation of each local area network internal application and information distribution website	Add a security proxy server and implement access control based on the authentication system
2	Security of remote endpoints accessed via dial-up	Identity and permission-based access control
3	MAIL security	Centralized public and private key management system for easy deployment of e-mail authentication and encryption systems
4	Database security	A flexible multilevel and multiscope security database system based on an authentication system can be easily established
5	Host or network-based data encryption and decryption technology	Centralized key management and distribution system to improve the management intensity of data encryption and decryption

From the product and technology, information security not only involves specific information network equipment but also mainly involves information security strategy and management; it reflects the modern information security concept of technology as a supplement and management as the main focus. The specific implementation must follow the overall design, distribution implementation, and continuous improvement of the information security concept. According to the network model, generally speaking, the needs for network security are shown in Table 2.

In addition, for the specific characteristics of the campus network, the specific targeted contents of the campus intelligent network are shown in Table 3. With the list of security problems and proposed measures for the current campus network, the construction of a perfect authentication system based on the PKI system will greatly enhance the security capability of the whole network, and a network security system can be fundamentally established.

3. Methodology

3.1. Measures for Intelligent Development of Student Management. First of all, the intelligent management of students in colleges and universities must form a responsive system in order to achieve long-term sustainable development. On the one hand, the state should provide a responsive system guarantee from the macropolicy to help the intelligence and encourage universities to implement intelligent management, and on the other hand, student management personnel of universities should keep pace with the times, actively synchronize with the mainstream technology development of the society, and be ready to adjust and improve student management. In addition, each university should highlight its individuality and create different intelligent management systems according to its characteristics [16]. For example, students' management in liberal arts colleges can use intelligence to sense the talents and specialties of each student at any time, while science colleges focus on highlighting students' active participation in the construction of intelligent laboratories and intelligent student management, so as to truly customize the intelligent management of college students under the vision of big data.

Secondly, since China lags behind the international average in certain core technology fields of big data and artificial intelligence, as the cultivation place of high-end talents, colleges and universities should introduce high-end technical talents extensively, especially the teachers who meet the development needs of today's intelligent era. A team of experienced backbone teachers can make the management of college students more scientific, reasonable,

Table 2: Network model protocols and technologies.

OSI seven-layer protocol	TCP/IP protocol	Security requirements and key technologies
Application layer Representation layer Session layer	Process layer	Application confidentiality requirements: the main technologies are SHTTP, PGP, SMIM, development-specific protocols, etc.
Transport layer	Transport layer	Confidentiality of the transmitted data: the main technologies are SSL protocol and public key-based authentication and symmetric key encryption technology, etc.
Network layer	Network layer	Access control requirements: the main technologies are packet filtering, IPSEC protocol, VPN, etc.

Table 3: Network security targeting table.

Serial number	Targeted questions	Recommended measures
1	Network access control	Integrated use of firewall, VLAN, and VPN technologies to achieve website access control within and between LANs
2	Security of remote endpoints accessed via dial-up	Deployment of link encryption or IP encryption or VPN technology
3	Storage security	Deploy file encryption system on key machines
4	Security of the operating system	Security hardening of NT servers and new installations are recommended to use secure operating systems
5	Antivirus	Deployment of hierarchical network antivirus workstations and client software
6	MAIL security	Deploying an e-mail encryption system

and efficient, and a team of excellent young teachers can convey new vitality and provide new directions for the management mode of the school [17]. Meanwhile, major universities should pay attention to recruiting talents in intelligent management and do scientific research and projects in intelligent management; firstly, pilot in our university to accumulate experience, and then, other universities gradually implement it with the actual situation and their own schooling characteristics, so as to continuously develop technology and truth in practice.

Third, the construction work of the Internet of things in colleges and universities should be unified planning, gradually building or introducing the successful Internet of things construction experience of other colleges and universities, especially schools with the Internet of things majors should do a good job of advanced and open construction of the IoT blubber laboratory [18], which can keep in touch with the development of society and facilitate upgrading also to student management in the case of retaining privacy real-time monitoring, tracking, and positioning. Universities actively exchange experiences of IoT construction, share the efficacy and convenience brought by successful IoT construction to intelligent university student management, and send relevant experts to guide how to implement relevant equipment systems for IoT construction if necessary, enrich data sources, and provide scenarios for intelligent applications to be landed.

Finally, to gradually build a unified comprehensive information platform for student management, college administrators should gradually integrate the scattered and fragmented information resources of students' daily life, continuously analyze and dig deep into the potential motives behind these data and information, and provide powerful data support for those concerned to make decisions.

3.2. General Structure of Intelligent System. The construction goal of the campus intelligent weak power system project is to establish an intelligent campus with advanced technology, comprehensive functions, and easy operation and can cover all functional areas [19]. The architecture of the weak electrical system implementation focuses on the following areas: campus security system, emergency broadcast system, campus computers, network system, and integrated cabling system. As shown in Figure 3, the whole safe campus monitoring system mainly includes the following parts: school security monitoring center, front-end video monitoring system, and communication transmission system. (1) Security monitoring center: the security monitoring center is the general control room of the system, which is the control and management center of the whole monitoring system; (2) video monitoring system: front-end video monitoring points can be divided into school entrance video monitoring points, channel video monitoring points, and large activity area video monitoring points according to the site conditions; (3) communication transmission system: the project goal is to establish a technologically advanced, comprehensive, and easy to operate system which can cover all functional areas of the intelligent campus, such as channel video monitoring points and large event area video monitoring points [20]. The process included in edge detection using the Canny operator is as follows.

In the first step, the Canny operator low pass filters the image with the first-order derivative of a two-dimensional Gaussian function, which is given by

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right). \quad (1)$$

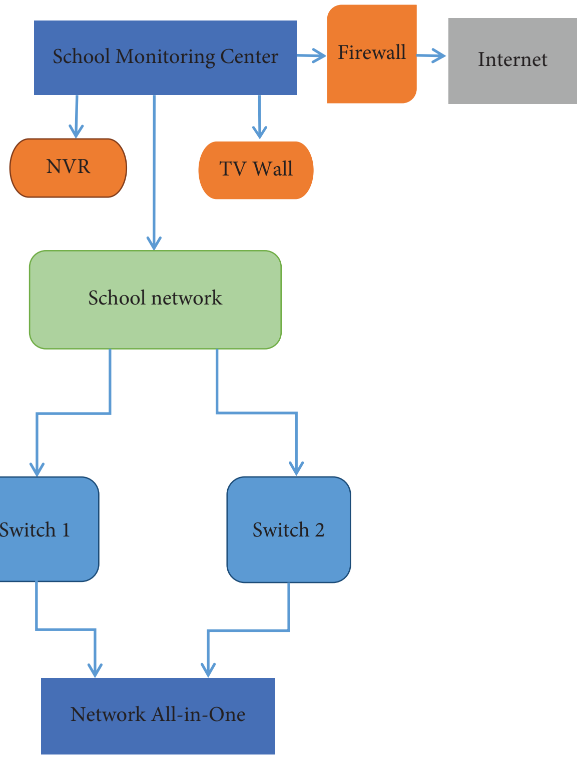


Figure 3: Overall system architecture.

The second step calculates the amplitude and direction of the gradient. Using the divisibility of the Gaussian function, the two filter convolution templates of G are decomposed into two one-dimensional determinant filters, which are then convolved with the image $f(x, y)$ separately to obtain the output:

$$\begin{aligned} E_{x/y} &= \frac{\partial G}{\partial x/y} * f(x, y), \\ A(x, y) &= \sqrt{E_X^2(x, y) + E_y^2(x, y)}, \\ \partial(x, y) &= \arctan\left[\frac{E_y(x + y)}{E_x(x, y)}\right]. \end{aligned} \tag{2}$$

Then, $A(x, y)$ and $\partial(x, y)$ reflect the magnitude of the gradient amplitude and the direction angle at the point (x, y) on the image, respectively.

3.3. Intelligent System Hardware Facilities. In order to further strengthen the security of the school, the school area is divided into entrances and exits, perimeters, and channels in a total of three areas, as shown in Table 4. The core layer is located in the security center server room, responsible for the entire network video surveillance service data traffic exchange, and can be connected to the remote network through the security policy. The core switching area mainly deals with the convergence point data traffic exchange; the switch requirements should be able to all IP packets for wire-speed nonblocking forwarding, while being able to work for a long time without failure. For each network, a hard disk recorder is configured with a 21.5-inch LCD monitor for real-time display of images of each monitoring point, which can be a multiscreen split display or cyclic switching display. Campus intelligent system hardware facilities include many subsystems, such as a security prevention system, multimedia teaching system, computer room project, communication engineering system, and weak electricity pipeline project.

Table 4: Intelligent monitoring configuration.

Precautionary area	Coverage locations
Entrance and exit	School entrance Main entrances and exits of teaching and office areas
Perimeter	School campus perimeter
Passage	Teaching and office area main corridors and passages

The campus integrated wiring system should follow the principles of practicality, flexibility, modularity, expandability, standardization, and reliability. Each room is designed with one network and one conversation, and considering line redundancy, two network cables are laid in each room. In order to meet the requirements of the structured cabling system, according to the above design principles, a super category 5 structured cabling system should be used, with a single set of a wiring system that

Table 5: Integrated wiring standards.

EIA/TIA-568	Civil building cable standards
EIA/TIA-569	Communication access and space standards for civil buildings
IEEE 802.3	Bus-based Ethernet LAN standard
IEEE 802.5	Loop LAN standard
ANSI FDDI	Fiber distributed data interface high-speed local area network standard
TPDDI	Copper wire distributed data interface high-speed LAN standard
ATM	Asynchronous transfer mode standard
RS232 RS422	Asynchronous and synchronous transmission standards

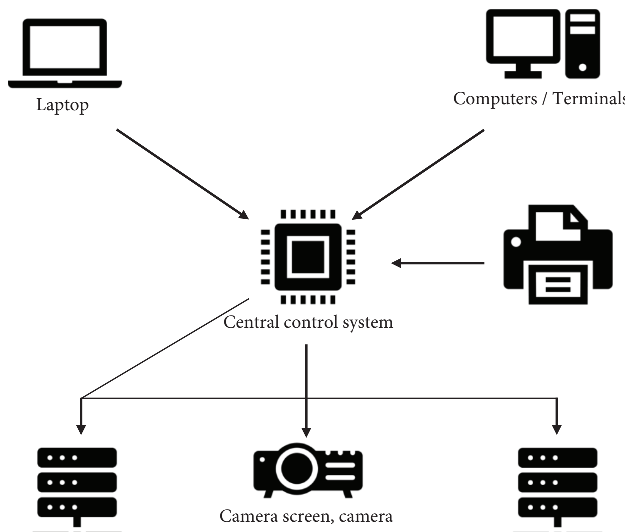


Figure 4: Intelligent student teaching system.

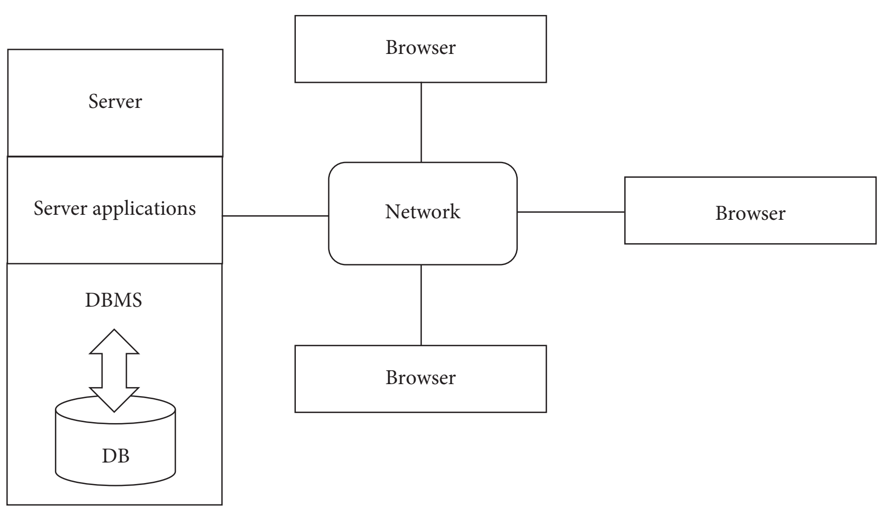


Figure 5: System development B/S model.

integrates the entire communication, including the wiring needed for voice, data, image, monitoring, and other equipment, as shown in Table 5.

3.4. Application of Intelligent System in Campus Teaching. The introduction of computer technology and communication technology into all aspects of teaching, thus triggering a major

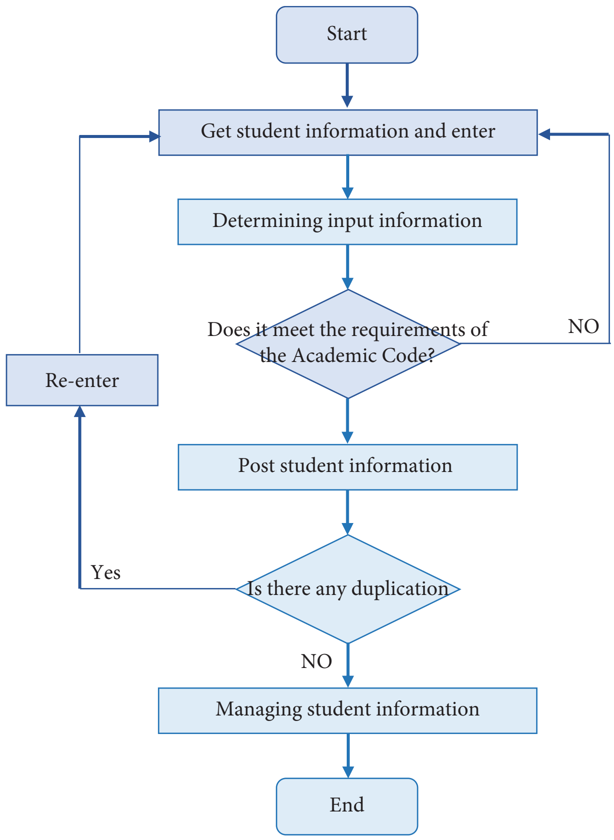


Figure 6: Flowchart of student information management.

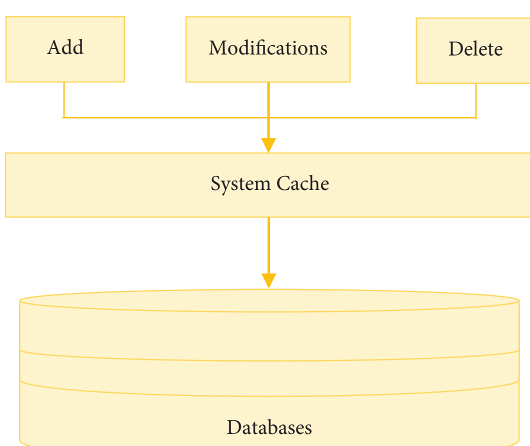


Figure 7: Flowchart of scholarship management.

innovation in teaching methods, teaching methods, and teaching tools, plays an immeasurable role in improving the quality of teaching and promoting the development of modernization of education. In this program design, multimedia teaching, courseware teaching, seminars, and large screen display functions are fully realized and convenient to access the Internet so that various images, audio, and video are perfectly displayed, as shown in Figure 4.

4. Result Analysis and Discussion

4.1. Student Management System Design. The basic architecture of intelligent student management information

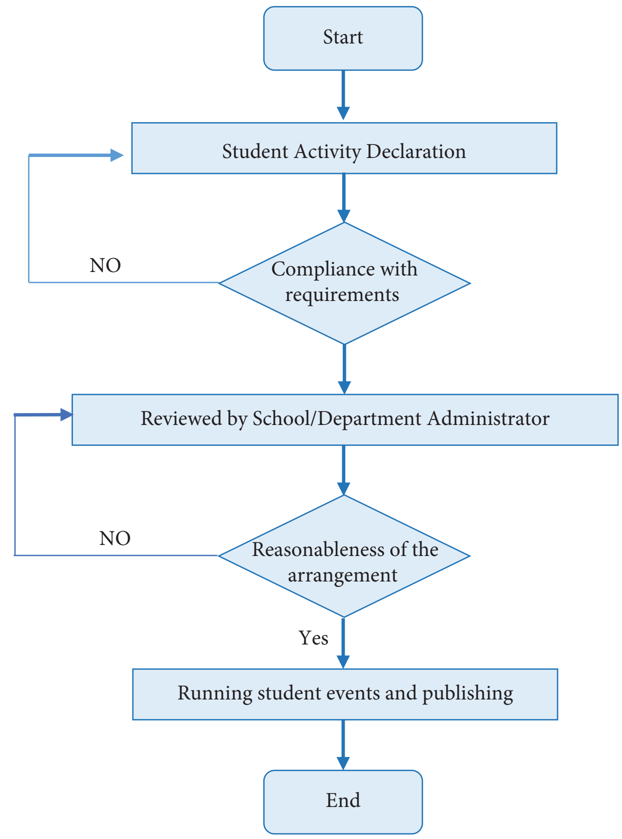


Figure 8: Flowchart of student activity management.

system: regarding the development of the system, an intelligent student information management system adopts the language as the core development language of the system [21]. In terms of development model selection, MVC three-layer development model is mainly adopted. Presentation layer (UI) mainly refers to the interface that interacts with the user. It is used to receive data input from users and display the data needed by users after processing. Business logic layer (BLL) is the bridge between the UI layer and DAO layer and realizes the business logic. Data access layer (Dao) deals with the database, mainly realizes adding, deleting, changing, and checking data, submits the data stored in the database to the business layer, and saves the data processed by the business layer to the database. As for the framework template of the front page design, easy-ui is mainly used. As for the backend database, MySQL is mainly used. As for the development framework, SSM framework is mainly used. The system development model is shown in Figure 5. In terms of functionality, Spring can achieve the integration of all the project's frameworks. In terms of transaction control, Spring can achieve the function of managing beans. Spring MVC framework: SpringMVC is a module provided by the Spring framework. The MVC three-tier development model is mainly to achieve the separation of the data layer, business layer, and representation layer so that it simplifies the development of Web applications, but in terms of core development implementation, MVC layered development model is mainly request-driven to achieve the purpose of data control forwarding.

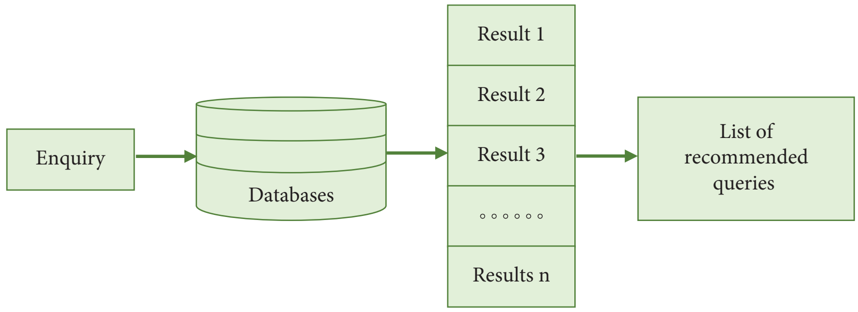


Figure 9: System query recommendation process.

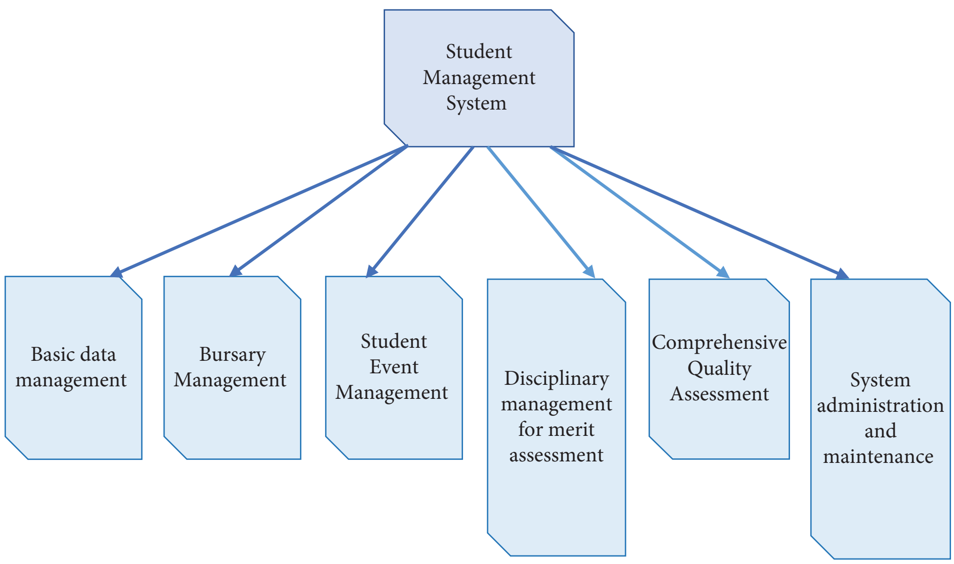


Figure 10: Composition of the student management system.

Basic data management is mainly to view the basic information of the students under their control, to modify and improve the student information [22]; you can query all student information, the information can be modified, and you can view the details of the student, as shown in Figure 6.

Scholarship and grant management business are divided into two parts, declaration and approval, each corresponding to students and counselors and departmental administrators at all levels, to declare and approve students' scholarships and grants; with Ajax technology to achieve rapid modification of scholarships and grants, dynamic refresh effect, the process of scholarship and grant management is as shown in Figure 7.

Student activity management business is for declaring and approving students' activities, each corresponding to students and counselors, departmental administrators at all levels, to manage students' activities at school and design the main processes of student activity management, as represented in Figure 8.

To manage the login account permissions in the whole system, only the school-level administrator has the permission to use this module, system log management; you can turn on and off log monitoring and monitor all users of the system operating system at any time; the process of querying the recommended system log information is shown in Figure 9.

4.2. Student Management System Functions. Student information management system is a management software developed for a large amount of work processing work in the school student office, mainly for school student information management; the overall task is to realize the systematization, scientific, standardization, and automation of student information relations; its main task is to use cell phones and computers for daily management of various student information. Analysis of the need to develop a student management information system to a large extent: the student management information system should be a comprehensive system that dynamically manages all data and information of students in a networked and informative way, as shown in Figure 10.

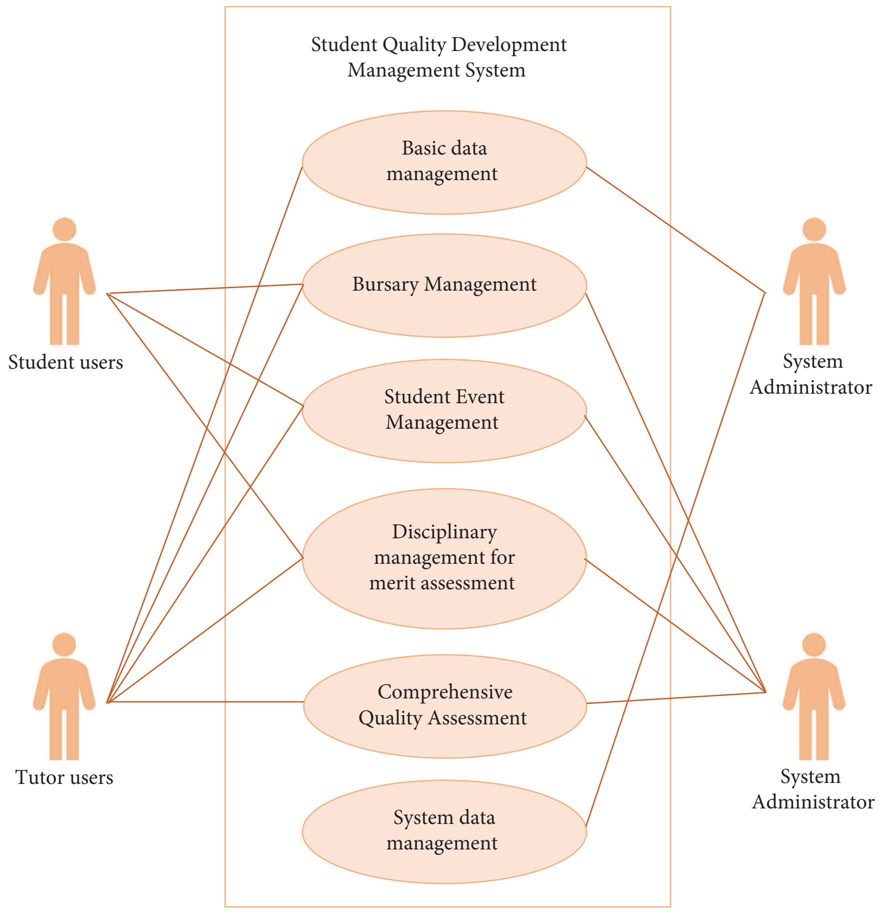


Figure 11: Student management system use case.

Meanwhile, in this intelligent student management system, according to the research results, we divide the users into four categories, such as student users, counselor users, system administrators, and department administrators. Each user role exercises different operation rights in the system, and the detailed functional requirements of users are shown in Figure 11.

The student management information system can ensure the following goals. (1) The system will ensure the practicality, reliability, advancement, and applicability of the system and strive to operate concisely and simplify the operation process. (2) The system will achieve classification and grading management of each submodule, dynamic management of the database, and classification of data to prevent data confusion. (3) The system is capable of simple query and compound query according to different conditions selected by the user, and the system development focuses on user experience. (4) The system is capable of classifying and summarizing the results of the user query, realizing the export function, and finally realizing the printing of reports. (5) The system focuses on data security and is capable of backing up the data and recovering the data deleted by mistake. (6) The import and export function of the data.

5. Conclusion

The research and application of an intelligent student information management system are mainly based on the actual needs of work and future expansion needs; the main architecture is mainly based on a three-layer framework, separating the representation layer, business layer, and data layer, and the development method of the prototype can improve the security and confidentiality of data in the process of network environment transmission, further enhancing the transparency of student information management and informationization. The level of student information management is further improved. The student management information system can better balance the supply and demand of student information resources in colleges and universities, make the management of student information and the selection, award, and punishment of financial aid in colleges and universities more scientific, fair, and efficient, and finally realize the process of accurate and efficient management of student information. It can integrate the data and information of various departments effectively to the maximum extent, which largely meets the needs of student information intelligent management of student work managers.

The informatization and intelligent construction of colleges and universities is an important reflection of the comprehensive strength of schools, so it is necessary to improve the understanding of the informatization of schools and improve the management of school teaching work. Colleges and universities should improve the intelligent construction of students' campus life and create an intelligent one-card of the campus, strengthen the intelligent technology training for teachers, and make continuous improvements to the school's faculty. We strive to create a college campus network with informationization, science, and intelligence and realize the integration of intelligent campus construction in colleges and universities, so as to facilitate the learning and life of college teachers and students.

Data Availability

The labeled dataset used to support the findings of this study can be obtained from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

The research was supported by Social Science Planning Research Project of Shandong Province (21CSDJ28), Qingdao Philosophy and Social Science Planning Project (QDSKL2101220), Humanities and Social Science project of Shandong Province (2022), and Continuing Education Research Project of Qingdao Agricultural University (2020).

References

[1] W. Li, "A preliminary investigation on the path of constructing the ecological environment of higher vocational innovation education," *Joint Journal of Tianjin Vocational Colleges and Universities*, vol. 19, no. 12, pp. 8–12+26, 2020.
[2] A. Melani, M. Bertogna, and V. Bonifaci, "Etc. Schedulability analysis of conditional parallel task graphs in multicore systems," *IEEE Transactions on Computers. Feb.*vol. 66, no. 2, pp. 339–353, 2021.
[3] M. Pricopi and T. Mitra, "Task scheduling on adaptive multicore," *IEEE Transactions on Computers*, vol. 63, no. 10, pp. 2590–2603, 2014.
[4] L. I. Jian-Feng and P. Jian, "Task scheduling algorithm based on improved geneticalgorithm in cloud computing environment," *Journal of Computer Applications*, vol. 31, no. 1, pp. 184–186, 2021.
[5] T ChiaIlg, J DouglaS, and V. K. Gurbani, "IN services for conveed (Intemet) telephony," *IEEE Communications Magazine*, vol. 38, no. 6, pp. l08–l115, 2020.
[6] C. A. Licciardi and A. Qandreetto, "An architecture for IN-intemet hybrid sevices," *Computer NetworI(S:The Intemational JoumaI of Computer and Telecommunications Networking*, vol. 35, no. 5, pp. 537–549, 2018.
[7] D. J. Watts, "The "new" science of networks," *Annual Reviewing Sociology*, vol. 30, pp. 241–270, 2022.
[8] M. O. Ball, "Complexity of network reliability computations," *Networks*, vol. 10, no. 2, pp. 153–165, 1980.
[9] J. F. Li and S. H. Peng, "An improved genetic algorithm-based task scheduling algorithm for cloud computing environment," *Computer Applications*, vol. 18, no. 1, pp. 184–186, 2019.
[10] R. Li and R. Kang, "A Review of Network Reliability Evaluation research," *Reliability Engineering*, pp. 5-6, 2018.
[11] Y. Pan, "Research on reliability index of communication network," *Electronic product reliability and environmental testing*, vol. 24, no. 1, pp. 1–4, 2020.
[12] Duan Yunfeng Song Junde, *Softswitch Architecture*, CTI World, no. 4, pp. 8–10, Lauderdale, FL USA, 2022.
[13] K Keith, *NGN Arcllitecture:Generic Principles,Functional Architecturc,and Lmplementation*, pp. 49–56, IEEE Communications Magazine, New York, USA, 2018.
[14] J. Du, "An implementation method for data migration between heterogeneous databases," *Journal of Qingdao Institute of Chemical Technology (Natural Science Edition)*, vol. 23, no. 3, pp. 78–81, 2019.
[15] T. Toth, "Applying mobile agent technology to intrusion detection," *In ICSE Workshop on Software Engineering and Mobility*, vol. 10, pp. 57–60, 2019.
[16] B. Bhargava, "Secure wireless network with movable base stations," *IEICE Transactions on Communications*, vol. 12, pp. 101–104, 2020.
[17] H. Yu, Z. Zhang, and J. Liu, "Research on scaling technology of bitcoin blockchain," *Journal of Computer Research and Development*, vol. 54, no. 10, pp. 2390–2403, 2022.
[18] H. Zhu and F. Xiang, "Advances in 5G network architecture design and standardization," *Telecommunications Science*, vol. 34, no. 4, pp. 126–132, 2018.
[19] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," *Computer Science*, vol. 14, no. 7, pp. 38-39, 2021.
[20] Ke Ma, "Service identification and management system and network traffic management," *Modern Telecommunications Technology*, vol. 43, no. 4, pp. 47–51, 2018.
[21] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: concept and applications," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 10, no. 2, pp. 1–19, 2019.
[22] D. K. Cohen, S. W. Raudenbush, and D. L. Ball, "Resources, instruction, and research," *Educational Evaluation and Policy Analysis*, vol. 25, no. 2, pp. 119–142, 2003.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9842050, 1 page
https://doi.org/10.1155/2023/9842050

Retraction

Retracted: Exploring the Entrepreneurship Training Mode of Medical Students in Beijing-Tianjin-Hebei Universities under the Strategic Background of "Healthy China" from the Perspective of Mixed Big Data

Mobile Information Systems

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] L. Gong, C. Wang, and J. Bai, "Exploring the Entrepreneurship Training Mode of Medical Students in Beijing-Tianjin-Hebei Universities under the Strategic Background of "Healthy China" from the Perspective of Mixed Big Data," *Mobile Information Systems*, vol. 2022, Article ID 8177760, 17 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 8177760, 17 pages
https://doi.org/10.1155/2022/8177760

Research Article

Exploring the Entrepreneurship Training Mode of Medical Students in Beijing-Tianjin-Hebei Universities under the Strategic Background of "Healthy China" from the Perspective of Mixed Big Data

Li Gong, Chao Wang, and Jiangtao Bai

Chengde Medical University, Chengde 067000, Hebei, China

Correspondence should be addressed to Jiangtao Bai; baijiangtao302@cdmc.edu.cn

Received 25 August 2022; Revised 17 September 2022; Accepted 22 September 2022; Published 6 October 2022

Academic Editor: Jiafu Su

In order to solve the problem that postgraduates are affected by the traditional concept of career choice of the native family and their entrepreneurial intention is obviously low, a method to explore the entrepreneurial spirit training model of medical students in Beijing-Tianjin-Hebei colleges and universities from the perspective of mixed big data is proposed. This article chooses to conduct research by conducting a survey on the entrepreneurial willingness of postgraduates in universities in the Beijing-Tianjin-Hebei and its influencing factors. From the perspective of master students from five universities in the Beijing-Tianjin-Hebei region, the selection of personal background information, entrepreneurial attitude, subjective norms, perceived behavior control, and entrepreneurship education index factors establish the index system of the influencing factors of graduate students' entrepreneurial intention and use the TPB theoretical model and SPSS data statistics and other tools to analyze the obtained data results. The results show that gender, major, place of origin, part-time experience, and entrepreneurial experience of relatives and friends correspond to 0.000, 0.000, 0.482, 0.172, 0.003, and 0.004, respectively. It can be seen that the independent variables, major and place of origin, have no statistical significance. In the statistics of college students' willingness to start a business, 23.33% of the students said "nothing at all," and 27.8% of the students said "thought, but would not start a business." It can be seen that entrepreneurial attitude, subjective norm, and perceived behavioral control all positively affect entrepreneurial willingness, and only learning ability, opportunity grasping ability, self-recognition ability, and decisiveness have a significant effect on entrepreneurial willingness.

1. Introduction

With the continuous development of medicine, medical education has been paid more and more attention. Especially under the influence of the epidemic situation, the promotion and development of medicine is particularly important. It is against this background that China has put forward the "healthy China" strategy. In the "healthy China" strategy, it is very important to do a good job in cultivating students' entrepreneurial spirit. In particular, the employment form of medical students is becoming more and more severe. Some foreign medical universities attach great importance to the research of entrepreneurial strategy and entrepreneurial environment, and have presented a mature mode in the cultivation of entrepreneurial ability [1]. At present, China's health industry only accounts for 4%~5% of GDP, which is far lower than the proportion of 15% in the United States and 10% in Japan and Canada. According to the characteristics of medical disciplines, it is a problem that many medical universities need to explore and face to carry out mass entrepreneurship and innovation education [2]. Since 2015, more than 23 policy documents have been issued for "mass entrepreneurship and innovation." However, compared with other majors, medical students' entrepreneurial endogenetic motivation and practical ability are slightly lacking in the practice of mass entrepreneurship and

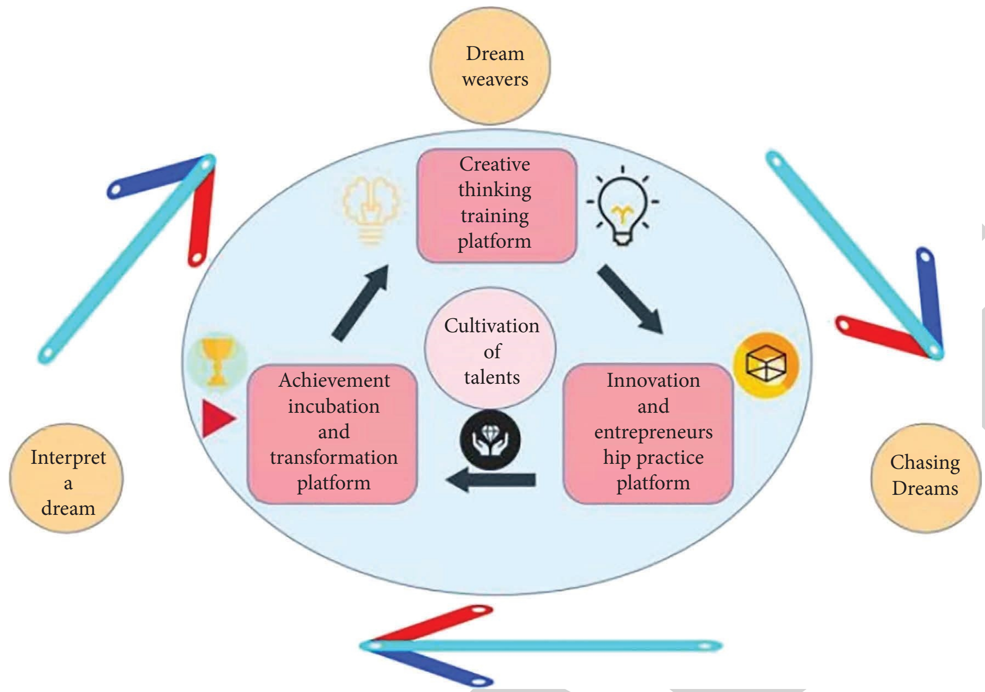


Figure 1: Talent cultivation turns to thinking.

innovation projects [3]. At present, there are many documents on the "healthy China" strategy in the field of higher education, but under its strategic background, there is less training on the entrepreneurial spirit of medical students. Most medical colleges and universities do not have in-depth entrepreneurial education on the integration of the "great health" industry. Therefore, through the innovative reform mode of "government, enterprise, university, and research," we will accelerate the development of the medical and health industry [4]. Taking the opportunity of vigorously developing medical and health undertakings and the goal of promoting students' entrepreneurship, major medical colleges and universities emphasize the cultivation of students' entrepreneurial awareness and knowledge application ability and, at the same time, link with the great health industry. Through improvement and innovation in the training mode, more entrepreneurial talents will be sent to the society [5].

2. Literature Review

Healthy China is a development strategy, and people's health is an important symbol of national prosperity and national prosperity. It is necessary to improve the national health policy and provide the people with all-round and full-cycle health services [6]. Since we want to promote the development of students' Entrepreneurship in the actual teaching environment and realize the implementation of the "healthy China" strategy, we need to understand the current situation of medical students' entrepreneurship training in the current actual teaching work. In this way, we can effectively realize the growth and promotion of students in the actual teaching activities. To cultivate talents, we must first determine a "dream" and carry out three stages of dream building, dream chasing and dream realization around the dream, and carry out training in stages and platforms, as shown in Figure 1.

Entrepreneurship education has also been well developed in foreign medical colleges. Foreign medical colleges and universities believe that interdisciplinary cooperation is crucial to the innovation of the healthcare industry, especially in transforming medical technology into effective clinical solutions. Therefore, almost all the courses in foreign medical colleges will be interdisciplinary. The courses include medical, engineering, science, business, or law majors. The concepts related to innovation and their applications in entrepreneurship, leadership, technology, medical care system, and pharmaceutical business will be taught [7]. A mixed method study of American idiopathic medical schools found that the number of entrepreneurial education programs in medical education was small but growing rapidly. These courses improve students' readiness for today's complex healthcare environment by covering novel educational topics with active and interdisciplinary pedagogy [8]. From 2007 to 2016, the entrepreneurship program began to be popular in the American College of allopathic medicine, forming an entrepreneurship education curriculum with seven education themes (innovation, entrepreneurship, technology, leadership, healthcare system, medical business, and enhanced adaptability) and two teaching methods themes (active learning and cross-disciplinary teaching) [9]. The University of Michigan (UM) School of Medicine designs and implements entrepreneurship programs to help students explore the business potential of medical innovation, support the establishment of entrepreneurial teams around the project, provide entrepreneurial mentors for team members, and provide tailor-made entrepreneurial education courses. Learning objectives are best provided in

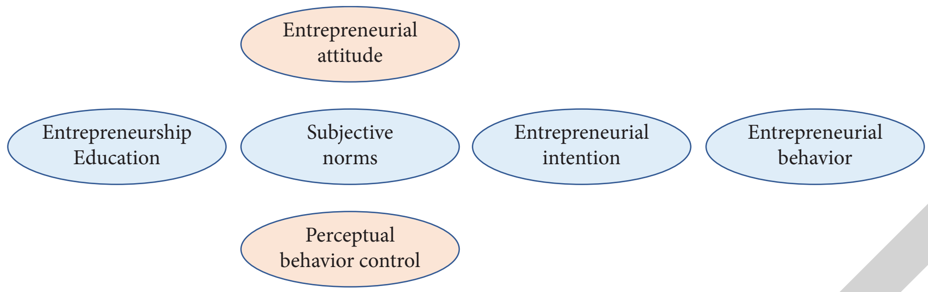


Figure 2: Framework of influencing factors of graduate students' entrepreneurial intention.

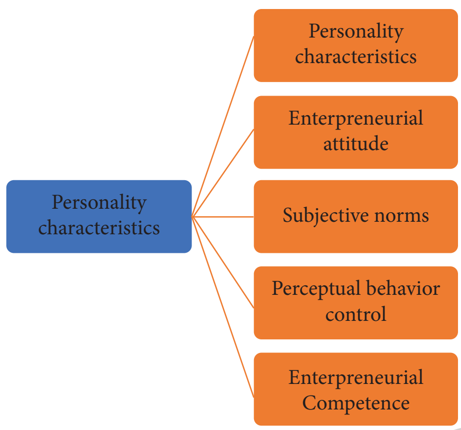


Figure 3: Model of influencing factors on entrepreneurial intention of medical college students.

the form of seminars/teaching. Teaching and seminars can provide more interactive teaching experience than lecture-based learning, increase communication between teachers and students, and promote students' active participation by holding entrepreneurial seminars, developing entrepreneurial ideas, forming team-led business plans and demonstrations, and developing skills including peer assessment, peer development, communication, critical assessment, creative thinking, problem-solving, and computing skills, to effectively promote and improve the innovative spirit and entrepreneurial skills of medical students.

Since our country put forward the entrepreneurship strategy, the major universities in the country have started to carry out the entrepreneurship education reform one after another, and more and more medical colleges are devoted to entrepreneurship education. There are certain differences and professional tendencies in the process of undertaking education in major universities. Different from comprehensive universities or other engineering classes, medical colleges lack resources for cross-learning due to their strong academic and professional characteristics. Moreover, most colleges and universities have misunderstandings in the entrepreneurship education of medical students, which is not significantly different from the innovation education of nonmedical students. The aim is to train students into big entrepreneurs, and there is insufficient correlation with the training objectives and employment characteristics of medical students. In addition, few medical colleges offer entrepreneurship courses for all students alone. Some medical colleges only set this course as an optional course of 2-3 class hours, with low credits, which is difficult to attract students' attention [10]. Although some colleges and universities have set up relevant courses, entrepreneurship education is only limited to theoretical courses and formalized entrepreneurship competitions, and there is no mature curriculum system. In addition, most of the teachers engaged in entrepreneurship education in medical colleges do not come from professional fields, with a single form of education, lack of entrepreneurial practice experience, and few innovation achievements [11].

3. The Influencing Factor Model of Entrepreneurial Intention Based on Plan Following

3.1. TPB's Entrepreneurial Education Factor Framework for College Students. Entrepreneurship education, as a medium for disseminating entrepreneurial knowledge and skills, cannot directly affect the entrepreneurial intention of students, but can exert influence by cultivating students' entrepreneurial attitude and individual characteristics. Entrepreneurship education can effectively cultivate and change the entrepreneurial attitude of college students, and social psychology research also shows that individual behavior and attitude can be effectively changed through practice and cognitive level, communication, and example encouragement. Therefore, based on the TPB theoretical framework, this article adds entrepreneurship education as an influencing factor. In the case of discussing how much entrepreneurship education occupies in the formation of postgraduate entrepreneurship intention, it also analyzes what kind of entrepreneurship education can maximize the driving effect in the formation of postgraduate entrepreneurship intention [12]. It is assumed that entrepreneurship education has an effect on entrepreneurial intention by changing individual entrepreneurial attitude, subjective norms, and perceived behavior control. The following theoretical factor framework is constructed as shown in Figure 2.

Based on the theory of planned behavior, this article introduces two variables: personality characteristics and entrepreneurial competence. Based on the above research hypotheses, it builds a perfect influencing factor model of Medical College Students' entrepreneurial willingness [13],

Table 1: Basic information of samples.

Category	Features	Quantity	Proportion (%)
Major	Medical science	291	29.82
	Pharmacy	206	21.11
	Nursing	142	14.55
	Psychology	116	11.89
	Management	160	16.39
	Engineering	61	6.24
Educational system	Four-year system	685	70.18
	Five-year system	291	29.82
Gender	Male	282	28.89
	Female sex	694	71.12
Place of origin	Town	461	47.38
	Countryside	512	52.62
Part-time experience	No	255	26.18
	Yes	719	73.82
Personal entrepreneurial experience	Never	886	91.34
	Once	76	7.84
	Starting a business	8	0.83
Entrepreneurial experience of relatives and friends	Never	263	27.11
	Once	482	49.69
	Starting a business	225	23.20
	Not at all	227	23.33
Entrepreneurial intention	Have thought about it, can try it if have a chance	467	48.01
	Have thought about it, but will not start a business	271	27.85
	Starting a business	8	0.83

which is in line with the current background environment of mass entrepreneurship. See Figure 3.

3.2. *Questionnaire Survey and Demonstration*

3.2.1. Sample Determination. The subjects of this study are junior undergraduates from five schools in Beijing, Tianjin, and Hebei. Their majors include medicine (traditional Chinese medicine and clinical medicine), pharmacy (traditional Chinese medicine and pharmacy), nursing, psychology, management (marketing and public utilities management), and engineering (medical information engineering).

The questionnaire survey was conducted by using the convenient sampling method and the combination of online and offline. A total of 1000 questionnaires were distributed, 988 were recovered, 12 invalid questionnaires were excluded, and 976 valid questionnaires were left, with an effective recovery rate of 97.60%. EpiData 3.1 was used for data entry, and SPSS 21.0 was used for data processing. Single choice questions are described and analyzed by frequency and percentage, and multichoice questions are described and analyzed by the percentage of respondents and the percentage of responses. See Table 1 for details.

3.2.2. Statistical Methods. EpiData 3.1 was used for data entry, and SPSS 21.0 was used for data processing. Cronbach's a was used to analyze the internal consistency of the questionnaire, factor analysis was used to test the structural validity of the questionnaire, and item domain correlation analysis was used to test the collective validity and discrimination validity of the questionnaire. According to the central limit theorem, when the sample size n gradually increases (usually requires $n > 30$), the probability distribution of the sample mean will approximately follow the normal distribution regardless of the overall distribution. The sample size of this study is about 1000, which has reached the sample requirements of the central limit theorem and can be regarded as approximately following the normal distribution. Therefore, the measurement data are statistically described by $(x + s)$, and the count data are described by component ratio or percentage (%). When the variance is the same, the two independent samples t-test is used for two groups of measurement data, and one-way ANOVA is used for multiple groups of data to analyze the entrepreneurial willingness and entrepreneurial competence of medical college students with different demographic characteristics. When the variance is not uniform, the rank sum test is used for analysis. Finally, we use multiple linear regression to analyze the influencing factors of entrepreneurial intention.

3.2.3. Descriptive Statistics. As shown in Tables 2 and 3, the overall score of entrepreneurial intention of the students of the two medical colleges is 2.54 ± 0.63, which is lower than the survey results of Ge Ruolan based on 290 medical undergraduates of Changsha Medical College in 2017 (2.94 ± 0.32). The scores of the four dimensions in the entrepreneurial willingness scale based on the TPB theory are 2.99 ± 0.78, 2.98 ± 0.75, 2.05 ± 0.76, and 2.30 ± 0.80, respectively, with an average score of less than 3. Among them, the lowest score is perceived behavior control, and the highest

Table 2: Scores of entrepreneurial intention of medical college students.

Dimension	Entry	n (%) Totally disagree	Relatively disagree	Commonly	Relatively agree	Totally agree	Score ($\overline{x} \pm s$)
Entrepreneurial attitude ($n = 974$)	AB1	112 (11.5)	149 (15.3)	497 (51.2)	182 (18.7)	33 (3.4)	2.87 ± 0.96
	AB2	88 (9.1)	143 (14.7)	456 (46.8)	253 (26.0)	34 (3.5)	3.00 ± 0.95
	AB3	58 (6.0)	149 (15.3)	525 (53.9)	211 (21.7)	31 (3.2)	3.01 ± 0.86
	AB4	44 (4.5)	96 (9.9)	333 (34.3)	406 (41.8)	92 (9.5)	3.42 ± 0.95
	AB5	107 (11.1)	285 (29.2)	453 (46.4)	105 (10.8)	23 (2.4)	2.64 ± 0.90
	Totality						2.99 ± 0.78
Subjective norms ($n = 973$)	SN1	77 (7.9)	217 (22.3)	506 (52.1)	140 (14.4)	32 (3.3)	2.83 ± 0.89
	SN2	63 (6.5)	155 (15.9)	515 (52.9)	206 (21.2)	34 (3.5)	2.99 ± 0.88
	SN3	62 (6.4)	165 (17.0)	534 (54.9)	177 (18.2)	35 (3.6)	2.96 ± 0.87
	SN4	69 (7.1)	128 (13.2)	427 (43.9)	289 (297)	60 (6.2)	3.15 ± 0.97
	Totality						2.98 ± 0.75
Perceptual behavior control ($n = 974$)	PBC1	278 (28.5)	415 (42.6)	222 (22.8)	54 (5.5)	5 (0.5)	2.07 ± 0.88
	PBC2	297 (306)	406 (41.8)	218 (22.4)	45 (4.6)	6 (0.6)	2.03 ± 0.88
	PBC3	263 (26.9)	423 (43.3)	230 (23.6)	49 (5.0)	8 (0.8)	2.09 ± 0.88
	PBC4	411 (42.3)	357 (36.8)	169 (17.4)	31 (3.2)	3 (0.3)	1.82 ± 0.85
	PBC5	283 (29.1)	360 (37.0)	285 (29.3)	42 (4.3)	3 (0.3)	2.10 ± 0.88
	PBC6	300 (30.8)	292 (30.0)	263 (27.0)	99 (10.2)	20 (2.1)	2.23 ± 1.06
	Totality						2.05 ± 0.76
Entrepreneurial intention (n-974)	E11	190 (19.5)	292 (30.0)	341 (35.0)	130 (13.3)	21 (2.2)	2.49 ± 1.02
	EI2	145 (14.9)	256 (26.3)	368 (37.9)	173 (17.8)	30 (3.1)	2.68 ± 1.03
	EI3	278 (28.5)	409 (42.1)	230 (23.6)	50 (5.1)	7 (0.7)	2.07 ± 0.89
	EI4	258 (265)	349 (35.8)	292 (30.0)	62 (6.4)	13 (1.3)	2.20 ± 0.95
	EI5	314 (32.2)	348 (35.7)	261 (26.8)	44 (4.5)	7 (0.7)	2.06 ± 0.91
	Totality						2.30 ± 0.80
Overall scale							2.54 ± 0.63

score is entrepreneurial attitude. The overall score of entrepreneurial competence of the students of the two medical colleges is 3.18 ± 0.52. Except for "learning ability" and "opportunity grasping ability," the scores of other variables are higher than 3. Among them, the highest score was "self-cognitive ability" (3.73 ± 0.66), and the lowest score was "opportunity grasping ability" (2.61 ± 0.82).

3.3. Reliability and Validity Test of the Scale

3.3.1. Reliability Test of the Scale. The Rasch model in the item response theory is applicable to the dichotomous response data. It is a special case of the single parameter model in the item response theory. Its probability is derived from the relationship between the subject's ability and the difficulty of the topic. Combining the subject's ability θ_i with the difficulty β_j of the item, a difference formula based on $(\theta_i - \beta_j)$ is proposed. The ratio of coding "correct/incorrect" or "yes/no" is

$$\ln\left[\frac{\Pr\left(y_{ij} = 1|\theta_i, \beta_j\right)}{1 - \Pr\left(y_{ij} = 1|\theta_i, \beta_j\right)}\right] = \theta_i - \beta_j, \tag{1}$$

where $\Pr\,(y_{ij} = 1|\theta_i, \beta_j)$ represents the probability that subject i evaluates item j as 1 or 0. θ_i represents the intrinsic ability of subject i; β_j represents the difficulty of item j, and there is a 50% probability that the answer is correct. Then, the formula of the two classification Rasch model is

$$P_r\left(y_{ij} = 1|\theta_i, \beta_j\right) = \log it^{-1}\left(\theta_i - \beta_j\right) = \frac{1}{1 + \exp\left(-\left(\theta_i - \beta_j\right)\right)}. \tag{2}$$

When the model contains the discrimination parameter a_j, which represents the discrimination of the item J, the formula can be rewritten as

$$P_r\left(y_{ij} = 1|\theta_i, \beta_j, a_j\right) = \frac{\exp\left[a_j\left(\theta_i - \beta_j\right)\right]}{1 + \exp\left(-\left(\theta_i - \beta_j\right)\right)}. \tag{3}$$

The partial credit model (PCM) constructed by the master is a special model when GPCM assumes that the discrimination parameter of each item is 1 ($a_j = 1$), which is applicable to multiclassification response data:

$$P_r\left(y_{ij} = y|\theta_i, \beta_{jh}\right) = \frac{\exp\sum_{h=0}^{k} a_j\left(\theta_i - \beta_{jh}\right)}{\sum_{k=0}^{m_j}\exp\sum_{h=0}^{k}\left(\theta_i - \beta_{jh}\right)}. \tag{4}$$

The PCM is extended to the generalized segment scoring model (GPCM). Since the project differentiation parameter is introduced into the model, GPCM brings more information than PCM:

$$P_r\left(y_{ij} = y|\theta_i, a_j, \beta_{jh}\right) = \frac{\exp\sum_{h=0}^{k} a_j\left(\theta_i - \beta_{jh}\right)}{\sum_{k=0}^{m_j}\exp\sum_{h=0}^{k} a_j\left(\theta_i - \beta_{jh}\right)}, \tag{5}$$

Table 3: Scores of entrepreneurial competence of medical college students.

Dimension	Entry	n (%)					Score ($\overline{x} \pm s$)
		Totally disagree	Relatively disagree	Commonly	Relatively agree	Totally agree	
Learning ability ($n = 973$)	A1	149 (15.3)	291 (29.9)	382 (39.3)	130 (13.4)	21 (2.2)	2.57 ± 0.97
	A2	122 (12.5)	208 (21.4)	420 (43.3)	201 (20.7)	22 (2.3)	2.79 ± 0.98
	A3	153 (15.7)	317 (32.6)	394 (40.5)	92 (9.5)	16 (1.6)	2.49 ± 0.92
	A4	97 (10.1)	164 (16.9)	375 (386)	293 (30.1)	43 (4.4)	3.02 ± 1.02
	Totality						2.72 ± 0.82
Ability to grasp opportunities ($n = 972$)	B1	117 (12.1)	269 (27.7)	443 (45.6)	127 (13.2)	15 (1.5)	2.64 ± 0.91
	B2	126 (13.2)	286 (29.5)	415 (42.8)	130 (13.4)	12 (1.2)	2.60 ± 0.92
	B3	129 (13.3)	289 (29.7)	413 (42.5)	131 (13.5)	10 (1.0)	2.59 ± 0.92
	B4	120 (12.4)	290 (29.9)	412 (42.5)	132 (13.6)	15 (1.5)	2.62 ± 0.92
	Totality						2.61 ± 0.82
Interpersonal skills ($n = 972$)	C1	36 (3.7)	108 (11.1)	408 (42.0)	358 (36.8)	62 (6.4)	3.31 ± 0.89
	C2	48 (4.9)	165 (17.0)	386 (39.7)	319 (32.8)	54 (5.6)	3.17 ± 0.94
	C3	38 (3.9)	167 (17.2)	416 (42.8)	309 (31.8)	42 (4.3)	3.15 ± 0.89
	C4	22 (2.3)	107 (11.0)	377 (38.8)	401 (41.3)	64 (6.6)	3.39 ± 0.85
	C5	25 (2.6)	109 (11.2)	359 (37.0)	411 (42.3)	67 (6.9)	3.40 ± 0.87
	Totality						3.28 ± 0.74
Planning ability (n-973)	D1	16 (1.6)	59 (6.0)	308 (31.7)	519 (53.4)	70 (7.2)	3.58 ± 0.78
	D2	20 (2.1)	63 (6.5)	378 (38.8)	450 (46.2)	62 (6.4)	3.48 ± 0.79
	D3	21 (2.2)	61 (6.3)	339 (34.8)	488 (50.2)	64 (6.6)	3.53 ± 0.80
	Totality						3.53 ± 0.72
Supply integration capability (n-973)	E1	74 (7.6)	236 (24.3)	443 (45.6)	193 (19.9)	26 (2.7)	2.86 ± 0.91
	E2	50 (5.1)	139 (14.3)	382 (39.3)	347 (357)	55 (5.7)	3.22 ± 0.94
	E3	37 (3.8)	140 (14.4)	408 (42.0)	324 (33.3)	63 (6.5)	324 ± 0.91
	E4	46 (4.7)	113 (11.6)	365 (37.6)	369 (38.0)	79 (8.1)	3.33 ± 0.95
	Totality						3.16 ± 0.77
	F1	14 (1.4)	34 (3.5)	274 (282)	538 (55.4)	111 (114)	3.72 ± 0.77
	F2	15 (1.5)	40 (4.1)	275 (28.3)	536 (551)	106 (10.9)	3.70 ± 0.78
	F3	10 (1.0)	35 (3.6)	272 (28.0)	548 (56.4)	107 (11.0)	3.73 ± 0.74
	F4	11 (1.1)	34 (3.5)	249 (256)	557 (57.3)	121 (12.4)	3.76 ± 0.75
	Totality						3.73 ± 0.66
Self-cognitive ability (n-972)	G1	22 (2.3)	82 (8.4)	388 (40.0)	408 (42.0)	71 (7.3)	$3.44 + 0.84$
	G2	22 (2.3)	100 (103)	356 (36.7)	408 (42.0)	85 (88)	3.45 ± 0.88
	G3	34 (3.5)	124 (12.8)	386 (39.9)	346 (357)	78 (81)	3.32 ± 0.92
	Totality						3.40 ± 0.73
Decisiveness (n-970)	H1	50 (5.2)	195 (20.1)	486 (50.2)	204 (21.0)	35 (3.6)	2.98 ± 0.87
	H2	36 (37)	173 (17.8)	451 (46.5)	269 (27.7)	41 (4.2)	3.11 ± 0.87
	H3	41 (4.2)	186 (19.2)	447 (46.2)	248 (25.6)	48 (4.9)	3.08 ± 0.90
	Totality						3.06 ± 0.75
Totality							3.18 ± 0.52

Table 4: Reliability analysis of the scale.

Gauge	Dimension	Cronbach's coefficient	Number of lowest scoring cases (%)	Number of highest scoring cases (%)
Entrepreneurial intention	Totality	0.941		
	Entrepreneurial attitude	0.902	32 (3.3)	10 (1.0)
	Subjective norms	0.849	42 (4.3)	12 (1.2)
	Perceptual behavior control	0.918	160 (16.4)	1 (0.1)
	Entrepreneurial intention	0.887	106 (10.9)	2 (0.2)
Entrepreneurial competence	Population	0.934		
	Learning ability	0.862	70 (7.2)	9 (0.9)
	Ability to grasp opportunities	0.923	90 (9.3)	4 (0.5)
	Interpersonal skills	0.886	15 (1.5)	23 (2.4)
	Planning ability	0.895	13 (1.3)	40 (4.1)
	Resource integration capability	0.850	27 (2.8)	15 (1.5)
	Self-cognitive ability	0.889	9 (0.9)	47 (4.8)
	Willpower	0.777	10 (1.0)	29 (3.0)
	Decisiveness	0.806	24 (2.5)	19 (1.9)

Table 5: Factor analysis results of the entrepreneurial willingness scale.

Entry	Factor 1	Factor 2	Factor 3
PBC4	0.881	0.079	0.068
PBC3	0.849	0.140	0.164
PBC2	0.848	0.118	0.192
PBC1	0.814	0.136	0.130
PBC5	0.784	0.202	0.188
PBC6	0.640	0.288	0.183
EI5	0.631	0.451	0.081
EI3	0.622	0.461	0.050
EI4	0.579	0.552	0.091
AB1	0.211	0.812	0.264
AB2	0.165	0.788	0.289
AB4	0.014	0.765	0.282
AB5	0.316	0.720	0.202
AB3	0.167	0.687	0.278
EI2	0.369	0.681	0.207
EI1	0.489	0.528	0.112
SN3	0.204	0.281	0.847
SN2	0.179	0.299	0.845
SN1	0.224	0.172	0.829
SN4	0.047	0.251	0.571

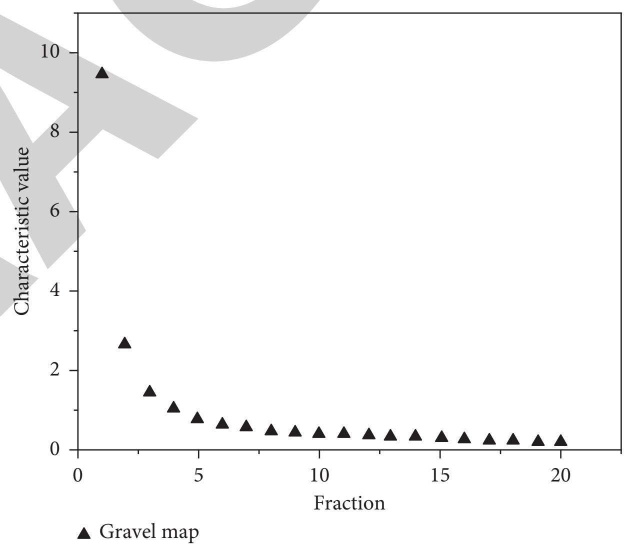


Figure 4: Factor analysis gravel chart of the entrepreneurial willingness scale.

wherein the parameter a_j is expressed as the potential health status or quality of life of the subject, the inherent health status or quality of life intensity measured by the response category threshold, and the discrimination of items under different potential health statuses.

3.3.2. Generalized Segment Scoring Model under Bayesian Theory. Definition $y = (y_{i1}, y_{i2}, \ldots, y_{ij})$ represents the observed item response vector, $\theta = (\theta_1, \theta_2, \ldots \theta_N)$ represents the subject's ability parameter, and the parameter vector of item j is $a = a_j$, $\beta = \beta_{jh}$. The likelihood function of the observed response vector is

$$\varsigma(y|\theta, a, \beta) = \prod_i \prod_j \prod_h \Pr\left(y_{ij} = y|\theta_i, a_j, \beta_{jh}\right). \tag{6}$$

N and $\log N$ represent normal distribution and log-normal distribution, respectively, and a is non-negative. The priori of parameter β is a priori with relatively no information. The distribution of latent trait θ was not set as standard normal. Instead, set the super priori of μ as standard normal and the super priori of μ as no information priori. The super priori of σ^2 obeys the inverse gamma distribution with an average value of 1.0 and a variance of 2 (when the shape rate is −0.5), which places most of the parameter values of β in the commonly observed [−3, +3] limits. Then, the joint posterior distribution is

Table 6: Factor analysis results of the entrepreneurial competence scale.

Entry	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
C2	0.852	0.088	0.111	0.167	0.055	0.165	0.050
C1	0.825	0.064	0.152	0.129	0.078	0.134	0.130
C3	0.814	0.126	0.107	0.092	0.060	0.140	0.050
C5	0.674	0.102	0.246	0.106	0.180	0.049	0.255
C4	0.667	0.130	0.203	0.094	0.150	0.020	0.262
B2	0.120	0.847	0.090	0.097	0.088	0.259	0.043
B3	0.092	0.835	0.062	0.133	0.104	0.242	0.048
B1	0.109	0.812	0.061	0.103	0.142	0.333	0.013
B4	0.155	0.777	0.025	0.140	0.136	0.295	0.114
F3	0.168	0.059	0.824	0.171	0.082	0.040	0.145
F2	0.165	0.062	0.793	0.146	0.151	0.095	0.208
F4	0.166	0.079	0.779	0.192	0.100	0.060	0.176
F1	0.226	0.065	0.716	0.118	0.196	0.042	0.258
H3	0.084	0.212	0.034	0.736	0.152	0.037	0.212
H2	0.268	0.188	0.109	0.676	0.111	0.098	0.175
G3	0.116	−0.081	0.364	0.668	0.135	0.188	−0.027
H1	0.092	0.414	0.103	0.654	0.167	0.002	0.167
G2	0.087	0.155	−0.091	0.430	0.569	0.220	0.159
G1	0.004	0.121	0.381	0.492	0.251	0.070	0.143
E2	0.084	0.059	0.153	0.126	0.847	0.060	0.097
E3	0.161	0.090	0.172	0.230	0.784	0.110	0.147
E1	0.092	0.269	−0.003	0.166	0.718	0.158	0.127
E4	0.114	0.089	0.260	0.150	0.689	0.108	0.153
A2	0.143	0.319	0.054	0.106	0.088	0.792	0.078
A1	0.131	0.268	0.058	0.145	0.074	0.775	0.007
A4	0.089	0.169	0.139	0.005	0.146	0.736	0.125
A3	0.131	0.380	0.013	0.117	0.104	0.734	0.033
D2	0.210	0.077	0.288	0.148	0.169	0.062	0.794
D3	0.213	0.089	0.268	0.183	0.202	0.129	0.777
D1	0.253	0.048	0.276	0.195	0.186	0.077	0.751

$$\begin{aligned} p(\theta,\alpha,\beta|y) &\propto p(\alpha,\beta)p(\theta|\alpha,\beta)p(y|\alpha,\beta) \\ &= p(a)p(\beta)p(\theta|\alpha,\beta)\iota(y|\theta,\alpha,\beta). \end{aligned} \tag{7}$$

The conditional posterior is proportional to the prior number likelihood ratio of the data. Then, the conditional probability distribution of θ, α, β is

$$\begin{aligned} &p(\theta|y,\alpha,\beta) \propto p(\theta)\iota(y|\theta,\alpha,\beta), \\ &p(a|y,\theta,\beta) \propto p(a)\iota(y|\theta,\alpha,\beta), \\ &p(\beta|y,\alpha,\beta) \propto p(\beta)\iota(y|\theta,\alpha,\beta). \end{aligned} \tag{8}$$

3.4. Consistency Test of Inspection Scale

3.4.1. Reliability Analysis of the Scale. GPCM was used to test the internal consistency reliability of the scale. The larger the coefficient, the higher the internal consistency of the variables, indicating that the measurement items between the variables have a good correlation. At present, the common criteria in academic circles are Cronbach's a is greater than 0.90, and then, the reliability of this test or scale is excellent; if Cronbach's a is between 0.80 and 0.90, the reliability is acceptable; if Cronbach's a is between 0.70 and 0.80, it needs to be revised, but it does not lose value; and if Cronbach's $a < 0.7$, the scale needs to be redesigned. The reliability analysis results of each part of the questionnaire are shown in Table 4. The

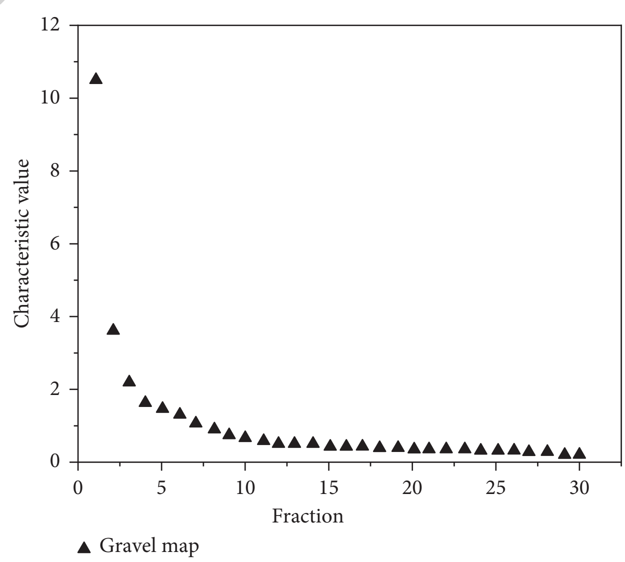


Figure 5: Factor analysis gravel diagram of the entrepreneurial competence scale.

reliability of each dimension of the TPB scale and the entrepreneurial competence scale is above 0.78. It can be seen that both scales have high reliability and high reliability.

Ceiling effect and floor effect refer to that most scores are concentrated at the very high or low end. These two

Table 7: Correlation coefficient matrix between demographic variables and entrepreneurial willingness.

	X_1 gender	X_2 major	X_3 place of origin	X_4 part-time experience	X_5 entrepreneurial experience of relatives and friends	Y_1 entrepreneurial intention
X_1	1					
X_2	0.012	1				
X_3	0.058*	−0.121*	1			
X_4	−0.136**	0.009	0.203**	1		
X_5	−0.015	−0.075**	−0.003	0.032	1	
Y_1	0.101***	−0.033	0.074*	0.094**	0.095**	1

Table 8: Regression analysis results of demographic variables on entrepreneurial willingness.

Variable	Nonstandard coefficient		Normalized regression coefficient	t	P
	Coefficient (B)	Standard error			
Constant	1.391	0.192		7.250	<0.000
X_1 gender	0.199	0.056	0.114	3.535	<0.000
X_2 major	−0.038	0.054	−0.023	−0.703	0.482
X_3 place of origin	0.071	0.053	0.045	1.368	0.173
X_4 part-time experience	0.176	0.059	0.098	2.978	0.003
X_5 entrepreneurial experience of relatives and friends	0.165	0.057	0.092	2.906	0.004

indicators reflect the important characteristics of score distribution. Calculate the proportion of people with the lowest and highest scores on the entrepreneurial willingness scale and entrepreneurial competence scale, respectively. If the scores of 1% of the subjects reach the limit (the highest score or the lowest score), there will be ceiling effect or floor effect, respectively. 15% of the floor or ceiling effect is considered acceptable. As shown in Table 4, only 16.4% of the cases with the lowest score in the perceived behavior control dimension in the entrepreneurial willingness scale have the floor effect. The proportion of the lowest score cases and the highest score cases in each dimension of the entrepreneurial competence scale are less than 15%, and there is no ceiling effect and floor effect.

3.4.2. Entrepreneurial Willingness Scale Based on TPB Theory. Table 5 and Figure 4 show the entrepreneurial willingness scale and the gravel chart. It can be seen that the KMO value is 0.942 greater than 0.7, Bartlett's sphericity test is approximately chi square 14372.737, the degree of freedom is 190, and the significance is 0.000, indicating that it is suitable for factor analysis. The principal component analysis is used for the orthogonal rotation of each item of the entrepreneurial willingness scale, and three common factors are extracted according to the feature root greater than 1 [14]. As shown in Table 5, PBC1~PBC6 have a high load in factor 1 and can be named perceptual behavior control factor; AB1~AB5 have a high load in factor 2 and can be named entrepreneurial attitude factor; SN1~SN4 have a higher load in factor 3 and can be named subjective criterion factor; EI1 and EI2 have higher loads in factor 2, and EI3~EI5 have higher loads in factor 1. The extracted common factors are basically consistent with the dimensions of the scale, and the cumulative contribution rate is 67.527%, indicating that the entrepreneurial intention scale has good structural validity.

3.4.3. Entrepreneurial Competence Table. Table 6 and Figure 5 are the entrepreneurial competence table and gravel chart. The KMO value of the entrepreneurial competence scale is 0.927, which is greater than 0.7. The sphericity test of Bartlett is approximately chi square 18390.291, the degree of freedom is 435, and the significance is 0.000, which indicates that it is suitable for factor analysis. The principal component analysis is used for the orthogonal rotation of each item of the entrepreneurial competence scale, and seven common factors are extracted according to the feature root greater than 1. As shown in Table 6, C1~C5 have a high load in factor 1 and can be named interpersonal relationship processing ability factor; B1~B4 have a higher load in factor 2, which can be named opportunity grasping ability factor; F1~F4 have a high load in factor 3 and can be named self-cognitive ability factor; H1~H3 has a high load in factor 4 and can be named decisive factor; E1~E4 have a higher load in factor 5 and can be named resource integration capability factor; A1~A4 have a higher load in factor 6, which can be named learning ability factor; D1~D3 have a higher load in factor 7, which can be named planning capacity factor; and G1~G3 also show a high load in factor 4 [15]. The extracted common factors are basically consistent with the dimensions of the scale, and the cumulative contribution rate is 71.743%, indicating that the entrepreneurial competence scale has good structural validity.

4. Analysis of Influencing Factors of Entrepreneurial Intention of Medical College Students

Multiple linear regression was used to analyze the influencing factors of Medical College Students' entrepreneurial intention. First, the ordered or disordered multiclassification variables are transformed into two classification variables, that is, majors: 1 = medical related (medicine, pharmacy,

Table 9: Regression analysis results of demographic variables on entrepreneurial willingness.

Variable	Nonstandard coefficient		Normalized regression coefficient	t	P
	Coefficient (B)	Standard error			
Constant	1.408	0.165		8.545	<0.000
X_1 gender	0.206	0.056	0.117	3.665	<0.000
X_4 part-time experience	0.192	0.058	0.107	3.336	0.001
X_5 entrepreneurial experience of relatives and friends	0.167	0.057	0.093	2.952	0.002

Table 10: Correlation coefficient matrix of entrepreneurial attitude, subjective norms, perceived behavior control, and entrepreneurial willingness.

	Entrepreneurial attitude	Subjective norms	Perceptual behavior control	Entrepreneurial intention
AB	1			
SN	0.572***	1		
PBC	0.454***	0.392***	1	
EI	0.646***	0.473***	0.685***	1

nursing) and 2 = nonmedical related (psychology, management, engineering); and entrepreneurial experience of relatives and friends: 1 = no entrepreneurial experience (never before) and 2 = entrepreneurial experience (ever and now) [16].

Secondly, the applicable conditions of multiple linear regression are analyzed, including the following:

(1) There is a linear relationship between the independent variable and the dependent variable. It is judged by drawing a scatter diagram.

(2) Normality of residuals. It is judged by drawing standardized residual histogram and normal probability diagram ($P-P$ diagram).

(3) Equivariance of residuals. By plotting the scatter plot of the predicted value of the standardized residual, if the standardized residual fluctuates below the zero level without obvious regularity, it can be judged that y satisfies the assumption of equal variance.

(4) Eliminate influential cases. It is generally believed that if the absolute values of standardized residuals and studentized residuals are less than 3, there is no strong influence point in the sample. However, the removal of strong influence points needs to be carefully selected in combination with professional knowledge.

(5) There should be no collinearity between independent variables. It is generally believed that if tolerance (TOL) < 0.1, variance inflation factor (VIF) > 5, and condition index (CI) > 30, there is severe collinearity.

4.1. Regression Analysis of Demographic Characteristics on Entrepreneurial Intention. Taking entrepreneurial willingness as the dependent variable and demographic characteristics as the independent variable, multiple linear regression analysis was conducted by using the forced entry method, and the regression equation is established as follows:

$$Y_1 = b_0 + b_{11}X_1 + B_{12}X_2 + b_{13}X_3 + b_{14}X_4 + b_{15}X_5, \qquad (9)$$

where Y_1 represents the entrepreneurial intention, X_1 represents the gender, X_2 represents the major, X_3 represents the place of origin, X_4 represents the part-time experience, X_5 represents the entrepreneurial experience of relatives and friends, and $b_{,i}$ represents the regression coefficients of various items.

There is a linear relationship between the independent variable and the dependent variable by plotting the scatter plot. The residual histogram and the normal probability diagram (P-P diagram) show that the residual is normal. The residual determined by the scatter plot of the normalized residual prediction value satisfies the isovariance [17]. As shown in Table 7, the minimum and maximum values of the standardized residuals for the correlation between demographic variables and entrepreneurial willingness are −2.011 and 3.779, respectively, and the minimum and maximum values of the studentized residuals are −2.016 and 3.797, respectively.

The test results of the goodness of fit of the model show that the complex correlation coefficient R is 0.183, the determination coefficient R^2 is 0.033, and the adjusted R^2 value is 0.028, indicating that personality characteristics can explain about 3.3% of entrepreneurial intention. The regression equation was statistically tested by ANOVA: regression regression, degree of freedom = 5, SS regression = 20.507, and Ms regression = 4.101; residual degree of freedom = 963, SS residual = 592.377, and Ms regression = 0.615. The F value is 6.667, and the P value is < 0.000, which indicates that the regression equation is meaningful. As shown in Table 8, gender, major, place of origin, part-time experience, and entrepreneurial experience of relatives and friends are 0.000, 0.000, 0.482, 0.172, 0.003, and 0.004, respectively. It can be seen that the independent variables, major and place of origin, have no statistical significance. Then, remove the two variables of major and place of origin, and do multiple regression of gender, part-time experience, and entrepreneurial experience of relatives and friends. The results are shown in Table 8.

Table 11: Regression analysis of entrepreneurial attitude, subjective norms, and perceived behavior control on entrepreneurial willingness.

Variable	Model 1					Model 2				
	Nonstandard coefficient		Normalized regression coefficient	t	P	Nonstandard coefficient		Normalized regression coefficient	t	P
	Coefficient (B)	Standard error				Coefficient (B)	Standard error			
Constant	1.389	0.192		7.243	0.000	−0.371	0.135		−2.743	<0.000
Gender	0.198	0.056	0.113	3.510	0.000	0.020	0.036	0.011	0.549	0.58
Major	−0.039	0.054	−0.023	−0.726	0.468	−0.015	0.034	−0.009	−0.456	0.64
Place of origin	0.073	0.052	0.046	1.400	0.162	0.081	0.033	0.051	2.450	0.01
Part-time experience	0.177	0.059	0.098	2.991	0.003	0.050	0.038	0.028	1.322	0.18
Entrepreneurial experience of relatives and friends	0.166	0.057	0.093	2.928	0.003	0.036	0.036	0.020	0.998	0.31
Entrepreneurial attitude						0.381	0.027	0.374	14.319	
Subjective norms						0.070	0.027	0.065	2.595	0.01
Perceptual behavior control						0.510	0.024	0.488	21.014	<0.000
R^2			0.034					0.617		
Adjust R^2			0.029					0.614		
F value			6.721***					193.275***		

Table 12: Correlation coefficient matrix between entrepreneurial competence variables and entrepreneurial willingness.

	X_1 learning ability	X_2 grasp the opportunity	X_3 interpersonal relationship	X_4 plan	X_5 resource integration	X_6 self-cognition	X_7 will	X_8 decisiveness	Y_1 entrepreneurial intention
X_1	1								
X_2	0.633***	1							
X_3	0.357***	0.330	1						
X_4	0.269***	0.268***	0.518***	1					
X_5	0.352***	0.357***	0.369***	0.479***	1				
X_6	0.257***	0.208***	0.476***	0.587***	0.432***	1			
X_7	0.287***	0.253***	0.393***	0.467***	0.467***	0.573***	1		
X_8	0.354***	0.437***	0.410**	0.449***	0.467***	0.431***	0.557***	1	
Y_1	0.615***	0.537***	0.230***	0.149***	0.240***	0.109***	0.191***	0.307***	1

Table 13: Regression analysis of entrepreneurial competence variables on entrepreneurial willingness.

Variable	Model 1					Model 2				
	Nonstandard coefficient		Normalized regression coefficient	t	P	Nonstandard coefficient		Normalized regression coefficient	t	P
	Coefficient (B)	Standard error				Coefficient (B)	Standard error			
Constant	1.384	0.193	0.111	7.173	0.000	0.286	0.187	1.526	0.127	
Gender	0.194	0.057	3.426	0.001	0.130	0.044	0.074	2.932	0.003	
Major	−0.041	0.054	−0.025	−0.769	0.442	−0.113	0.042	−0.067	−2.721	0.007
Main source	0.076	0.052	0.048	1.448	0.148	0.106	0.041	0.067	2.605	0.009
Part-time experience	0.178	0.059	0.099	3.008	0.003	0.068	0.046	0.038	1.479	0.140
Entrepreneurial experience of relatives and friends	0.168	0.057	0.094	2.934	0.003	0.059	0.044	0.033	1.349	0.178
Learning ability						0.444	0.032	0.457	14.007	<0.000
Grasp the opportunity						0.229	0.033	0.238	7.026	<0.000
Interpersonal relationship						0.009	0.033	0.009	0.286	0.775
Plan						−0.025	0.037	−0.023	−0.684	0.494
Resource integration						−0.008	0.032	−0.007	−0.242	0.809
Self-cognition						−0.106	0.041	−0.088	−2.580	0.010
Will						0.019	0.036	0.017	0.511	0.610
Decisiveness						0.073	0.035	0.068	2.072	0.039
R^2			0.034					0.440		
Adjust R^2			0.029					0.432		
F value			6.682***					57.093***		

Table 14: Hypothesis test results.

Hypothesis	Content	Support or not
H_{1a}	The gender of college students positively affects their entrepreneurial intention;	Support
H_{1b}	The majors of college students positively affect their entrepreneurial intention;	Do not support it
H_{1c}	The place of origin of college students has a positive impact on their entrepreneurial intention;	Do not support it
H_{1d}	The part-time experience of college students positively affects their entrepreneurial intention;	Support
H_{1e}	The entrepreneurial experience of relatives and friends positively affects the entrepreneurial intention of college students;	Support
H_{2a}	Entrepreneurial attitude positively affects entrepreneurial intention;	Support
H_{2b}	Subjective norms positively affect entrepreneurial intention;	Support
H_{2c}	Perceived behavior control positively affects entrepreneurial intention;	Support
H_{3a}	Learning ability positively affects entrepreneurial intention;	Support
H_{3b}	The ability to grasp opportunities has a positive impact on entrepreneurial intention;	Support
H_{3c}	Interpersonal relationship processing ability positively affects entrepreneurial intention;	Do not support it
H_{3d}	Planning ability positively affects entrepreneurial intention;	Do not support it
H_{3e}	The ability of resource integration positively affects entrepreneurial intention;	Do not support it
H_{3f}	Self-cognitive ability positively affects entrepreneurial intention;	Support
H_{3g}	Perseverance has a positive impact on entrepreneurial intention;	Do not support it
H_{3h}	Decisiveness positively affects entrepreneurial willingness.	Support

The test results of the goodness of fit of the model show that the complex correlation coefficient $r = 0.175$, the determination coefficient $R^2 = 0.031$, and the adjusted R^2 value is 0.028, indicating that the personality characteristics can explain about 3.1% of the entrepreneurial intention. The regression equation was statistically tested by ANOVA: regression regression, degree of freedom = 3, SS regression = 18.863, and Ms regression = 6.288; and residua residual freedom is 968, SS residual = 596.450, and Ms residual = 0.616. The F value is 10.205, and the P value is < 0.000, which indicates that the regression equation is meaningful. As shown in Table 9, the coefficients of the three variables are all positive, which indicates that gender, part-time experience, and entrepreneurial experience of relatives and friends have a positive impact on entrepreneurial willingness and have statistical significance ($P < 0.05$). Taking another look at the standardized regression coefficient, gender (0.117) > part-time experience (0.107) > entrepreneurial experience of relatives and friends (0.093), which indicates that the degree of influence of the three variables on entrepreneurial intention is in the order of gender > part-time experience > entrepreneurial experience of relatives and friends. To sum up, the regression equation is

$$Y_1 = 1.408 + 0.206X_1 + 0.19389X_4 + 0.167X_5. \quad (10)$$

4.2. Regression Analysis of Entrepreneurial Attitude, Subjective Norms, and Perceived Behavior Control on Entrepreneurial Willingness. First, the control variable (demographic variable) is used as an independent variable to conduct a regression analysis on entrepreneurial intention to obtain model 1, and then, the control variable and entrepreneurial attitude, subjective norms, and perceived behavior control are used as independent variables to conduct a regression analysis on entrepreneurial intention to obtain model 2.

First, the applicable conditions of multiple linear regression are tested. There is a linear relationship between the independent variable and the dependent variable by plotting the scatter plot. The residual histogram and the normal probability diagram ($P - P$ diagram) show that the residual is normal. The residual determined by the scatter plot of the normalized residual prediction value satisfies the isovariance [18]. As shown in Table 10, the correlation analysis of entrepreneurial attitude, subjective norms, perceived behavior control, and entrepreneurial intention is carried out. There is a strong correlation between Sn and PBC, and there is a moderate correlation between AB, Sn, and PBC. There may not be a multiple linear relationship. Further analysis showed that the TOL of the five variables in model 1 was between 0.935 and 0.993, all of which were >0.1, and the VIF was between 1.007 and 1.070, all of which were <5. The TOL of the eight variables in model 2 is between 0.584 and 0.977, all of which are >0.1, and the VIF is between 1.024 and 1.711, all of which are <5. And the Ci of the two models is less than 30, so it is judged that there is no collinearity between the independent variables. The minimum and maximum values of standardized residuals are −3.887 and 3.396, respectively, and the minimum and maximum values of student residuals are −3.914 and 3.421, respectively.

The test results of the model goodness of fit show that the complex correlation coefficient R of model 1 is 0.184, the determination coefficient R^2 is 0.034, and the adjusted

R^2 value is 0.029. The complex correlation coefficient R of model 2 is 0.786, the determination coefficient R^2 is 0.617, and the adjusted R^2 value is 0.614, which indicates that the TPB variable can explain 61.7% of entrepreneurial intention.

The regression equation was statistically tested by ANOVA: model 1 was regression, with a degree of freedom = 5, SS regression = 20.660, and Ms regression = 4.132; and residual residual freedom = 962, SS residual = 591.418, and Ms residual = 0.615; the F value is 6.721, and the P value is < 0.000, which indicates that the regression equation is meaningful. Model 2 was regression with degree of freedom = 8, SS regression = 377.772, and Ms regression = 47.221; residual residual, degree of freedom = 959, SS residual = 234.306, and Ms residual = 0.244; the F value is 193.275, and the P valueis < 0.000, which indicates that the regression equation is meaningful.

As shown in Table 11, the P values corresponding to entrepreneurial attitude, subjective norms, and perceived behavior control in model 2 are <0.000, 0.010, and <0.000, respectively, and the coefficients are all positive, which indicates that entrepreneurial attitude, subjective norms, and perceived behavior control have a significant positive impact on entrepreneurial willingness. According to the standardized regression coefficient, the effects of the three variables on entrepreneurial willingness are perceived behavior control, entrepreneurial attitude, and subjective norms from high to low.

4.3. Regression Analysis of Entrepreneurial Competence Variables on Entrepreneurial Willingness. First, the control variable (demographic variable) is used as an independent variable to conduct regression analysis on entrepreneurial willingness to obtain model 1, and then, the control variable and entrepreneurial competence-related variables such as learning ability, opportunity grasping ability, and interpersonal relationship processing ability are used as independent variables to conduct regression analysis on entrepreneurial willingness to obtain model 2 [19]. First, the applicable conditions of multiple linear regression are tested. There is a linear relationship between the independent variable and the dependent variable by plotting the scatter plot. The residual histogram and the normal probability diagram ($P-P$ diagram) show that the residual is normal. The residual determined by the scatter plot of the normalized residual prediction value satisfies the isovariance [20]. As shown in Table 12, the correlation analysis of entrepreneurial competence and entrepreneurial willingness shows that there is a moderate correlation between the variables, and there may not be a multiple linear relationship. Further analysis showed that the TOL of the five variables in model 1 was between 0.936 and 0.993, all of which were >0.1, and the VIF was between 1.007 and 1.068, all of which were <5. The TOL of thirteen variables in model 2 is between 0.5140.975, all of which are >0.1, and the VIF is between 1.025 and 1.944, all of which are <5. Therefore, it is judged that there is no collinearity between independent variables. The minimum and maximum values of standardized residuals are −3.956 and 3.589, respectively, and the minimum and maximum values of student residuals are −3.992 and 3.640, respectively [21].

The test results of model goodness of fit show that the complex correlation coefficient r of model 1 is 0.184, the determination coefficient R^2 is 0.034, and the adjusted R^2 value is 0.029. The complex correlation coefficient r of model 2 is 0.663, the determination coefficient R^2 is 0.440, and the adjusted R^2 value is 0.432, which indicates that the TPB variable can explain 44.0% of entrepreneurial intention [22].

The regression equation was statistically tested by ANOVA: model 1 was regression, with a degree of freedom = 5, SS regression = 20.546, and Ms regression = 4.109; and residual residual, degree of freedom = 953, SS residual = 586.034, and Ms residual = 0.615; the F value is 6.682, and the P value is < 0.000, which indicates that the regression equation is meaningful [23]. Model 2 was regression with a degree of freedom = 13, SS regression = 266.838, and Ms regression = 20.526; and residual residual, degree of freedom 945, SS residual = 339.742, and Ms residual = 0.360; the F value is 57.093, and the P value is < 0.000, which indicates that the regression equation is meaningful.

As shown in Table 13, the corresponding P values of learning ability, opportunity grasping ability, interpersonal relationship ability, planning ability, resource integration ability, self-cognition ability, perseverance, and decisiveness are P values 0.000, 0.000, 0.775, 0.494, 0.809, 0.010, 0.610, and 0.039, respectively [24]. Among them, the coefficient of learning ability, opportunity grasping, and decisiveness is positive, and the coefficient of self-cognitive ability is negative, which indicates that learning ability, opportunity grasping, and decisiveness have a significant positive impact on entrepreneurial intention, while self-cognitive ability has a negative impact (PG 0.05). Let's look at the standardized regression coefficient. The degree of effect of the four variables on entrepreneurial willingness from high to low is learning ability (0.457), opportunity grasping ability (0.238), self-awareness ability (0.088), and decisiveness (0.068).

The hypothesis test results can be obtained according to the above multiple linear regression analysis results, as shown in Table 14. In the dimension of personality characteristics, only gender, part-time experience, and entrepreneurial experience of relatives and friends affect entrepreneurial intention. Among the TPB variables, entrepreneurial attitude, subjective norms, and perceived behavior control have significant positive effects on entrepreneurial willingness, and there are also significant interactions among the three variables. Among the entrepreneurial competence variables, only learning ability, opportunity grasping ability, self-awareness ability, and decisiveness have a significant positive impact on entrepreneurial intention [25].

5. Conclusion

The results show that the overall score of entrepreneurial willingness of medical college students is 2.54 ± 0.63, 23.33% of the students said "no at all," and 27.8% of the students said "considered but would not start a business"; that is, half of the students clearly said that they would not start a business, and their entrepreneurial willingness is low, which is consistent with the research results of many scholars at home and abroad. However, with the rapid development of economy and the upsurge of national innovation and entrepreneurship, it is the general trend that medical college students' entrepreneurial willingness is rising. Improving the entrepreneurial willingness of medical college students is the root of the problem. Under the concept of frequency, the generalized partial scoring model uses the LTM package in *R* software to estimate the parameters of GPCM. However, it has a sample size limit when screening and evaluating the scale items, and generally requires a relatively large sample size. Then, in the face of relatively small samples, it is suggested to use the method of combining the Bayesian theory with the generalized segment scoring model, *R* software, and WinBUGS software for analysis. Based on the appearance of the Bayesian method and the applicability of item response theory, the Bayesian generalized segment scoring model enables us to still obtain more reliable parameter estimates when facing small samples and then make relevant statistical inference.

In short, the long-term students of traditional Chinese medicine have insufficient self-awareness and career awareness. Their awareness and ability of career planning need to be improved. The degree of innovation and entrepreneurship education and the innovation and entrepreneurship ability of students still need to be strengthened. Therefore, the school should attach importance to the career planning education of college students, improve the curriculum, establish a long-term learning mechanism for teaching teachers, improve the degree of specialization of guidance personnel, and pay attention to the psychological guidance of students. At the same time, we should follow the trend of the times, integrate innovation, and entrepreneurship education into the whole process of undergraduate and postgraduate training of long-term students of traditional Chinese medicine, provide all-round support for medical students' entrepreneurship projects, make the future career development of medical students more diversified, and promote the industrialization transformation of medical-related scientific research achievements in multiple dimensions, so as to form a good cycle of employment driven by entrepreneurship.

Data Availability

The labeled dataset used to support the findings of this study can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported by the Key Project of Humanities and Social Sciences Research of Hebei Provincial Department of Education in 2021 (project no. SD2021002).

References

[1] A. Y. Permana, R. D. Srihartati, J. Akbardin, A. Setiawan, D. Jatnika, and W. Rustandi, "Technopreneur training model in building an entrepreneurship," *Journal of Architectural Research and Education*, vol. 2, no. 2, pp. 190–197, 2020.

[2] E. M. Daniel and R. Owen, "Home-based self-employment: combining personal, household and employment influences," *Journal of Enterprising Culture*, vol. 30, no. 02, pp. 123–160, 2022.

[3] J. Dogra, S. Jain, A. Sharma, R. Kumar, and M. Sood, "Brain tumor detection from MR images employing fuzzy graph cut technique," *Recent Advances in Computer Science and Communications*, vol. 13, no. 3, pp. 362–369, 2020.

[4] M. Ubilla, C. I. Velez, and S. M. Afcha, "La educación para el emprendimiento como predictor de intención emprendedora de estudiantes universitarios," *Formacion Universitaria*, vol. 13, no. 2, pp. 63–72, 2020.

[5] J. Pardosi, R. A. Nugroho, J. R. Manullang, and A. N. Asikin, "Pengembangan kewirausahaan di universitas mulawarman untuk mencetak wirausaha muda yang inovatif," *Jurnal Pengabdian UntukMu NegeRI*, vol. 14, no. 2, pp. 2745–3782, 2020.

[6] E. O. Nwajiaku, A. C. Ananwude, and V. Obi-Nwosu, "Financial deepening and entrepreneurial growth in Nigeria: a time series analysis (1986 -2018)," *International Journal of Academic Research in Accounting, Finance and Management Sciences*, vol. 10, no. 2, pp. 257–263, 2020.

[7] K. Lamm, C. Mulvaney, A. Lamm, K. Davis, and F. Masambuka-Kanchewa, "Model of success: extension services helping ugandan youth find a career in agriculture," *Journal of International Agricultural and Extension Education*, vol. 28, no. 3, pp. 7–13, 2021.

[8] J. Jayakumar, B. Nagaraj, S. Chacko, and P. Ajay, "Conceptual implementation of artificial intelligent based E-mobility controller in smart city environment," *Wireless Communications and Mobile Computing*, vol. 2021, pp. 1–8, Article ID 5325116, 2021.

[9] M. Trihastuti, A. Abdulkarim, E. Danial, and A. Supriadi, "Learning strategy for students' entrepreneurial spirit through students' cooperation (kopma)," *Jurnal Iqra: Kajian Ilmu Pendidikan*, vol. 5, no. 2, pp. 83–94, 2020.

[10] A. A. Dolgaya and O. V. Kovbasyuk, "Transformative education: raising entrepreneurial spirit," *Russian Journal of Innovation Economics*, vol. 10, no. 2, pp. 663–672, 2020.

[11] P. Ajay, B. Nagaraj, R. A. Kumar, R. Huang, and P. Ananthi, "Unsupervised hyperspectral microscopic image segmentation using deep embedded clustering algorithm," *Scanning*, vol. 2022, pp. 1–9, Article ID 1200860, 2022.

[12] F. Zhou, M. K. Lim, Y. He, and S. Pratap, "What attracts vehicle consumers' buying," *Industrial Management & Data Systems*, vol. 120, no. 1, pp. 57–78, 2019.

[13] X. Luo, A. Cottam, Y. J. Wu, and Y. Jiang, "Hybrid-data approach for estimating trip purposes," *Transportation Research Record*, vol. 2675, no. 11, pp. 545–553, 2021.

[14] M. R. L. Bazargan and S. Taghipour, "A hybrid data-driven approach for forecasting the characteristics of production disruptions and interruptions," *International Journal of Information Technology and Decision Making*, vol. 21, no. 04, pp. 1127–1154, 2022.

[15] J. Qiu and S. Yang, "Material and interface engineering for high-performance perovskite solar cells: a personal journey and perspective," *The Chemical Record*, vol. 20, no. 3, pp. 209–229, 2020.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9813079, 1 page
https://doi.org/10.1155/2023/9813079

Retraction

Retracted: Interior Soft Decoration Product Design Based on 3D Modeling and Image Processing Technology

Mobile Information Systems

Received 22 August 2023; Accepted 22 August 2023; Published 23 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Jin, Y. Chen, and J. Qu, "Interior Soft Decoration Product Design Based on 3D Modeling and Image processing Technology," *Mobile Information Systems*, vol. 2022, Article ID 9095614, 7 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 9095614, 7 pages
https://doi.org/10.1155/2022/9095614

Research Article

Interior Soft Decoration Product Design Based on 3D Modeling and Image Processing Technology

Ye Jin, Yan Chen, and Jianing Qu

School of Design and Media, Wuxi Vocational Institute of Commerce, Wuxi, Jiangsu 214000, China

Correspondence should be addressed to Ye Jin; 2018212671@mail.chzu.edu.cn

Received 7 August 2022; Revised 2 September 2022; Accepted 24 September 2022; Published 4 October 2022

Academic Editor: Xuefeng Zhang

In order to solve the problems of the traditional 3D simulation system of indoor soft decoration, such as the lack of effective enrichment, rendering means, and poor authenticity when selecting the geometric modeling method to simulate the effect of indoor soft decoration, this paper proposes a 3D simulation system of indoor soft decoration based on OpenGL and Direct3D. The system uses 3DMAX software to establish the indoor model, imports the model into the OpenGL 3D graphics standard library module, draws the indoor soft decoration scene elements using the basic geometric elements, and constrains the scene elements through the wall, overlap, and size constraint functions, so as to ensure the reasonable layout of the indoor soft decoration agent area and enrich the indoor model. After enrichment, the indoor model is transmitted to the Direct3D rendering engine module transformation unit to implement the soft decoration layout, and the lighting unit and rasterization unit are used to render the indoor after layout so as to obtain the optimal indoor soft decoration effect and output the 3D simulation image of indoor soft decoration. The experimental results show that the favorability, operation sensitivity, design portability, design satisfaction, and soft fitting effect score of the system are higher than those of the Untiy3D system and Smart3D system. The average score of this system is about 95, while the average score of the Untiy3D system is lower than 89, and the average score of the Smart3D system is lower than 85. *Conclusion*. The three-dimensional simulation effect of indoor soft decoration of this system is good, the design is convenient, the operation is sensitive, and the user satisfaction is high, which provides a reliable analysis basis for indoor design.

1. Introduction

With rapid economic development, people's living standards have greatly improved, and the requirements for living quality are becoming higher and higher. Therefore, the interior soft decoration design is the key to interior design, which has become more and more important in the overall decoration in recent years [1]. However, when displaying the interior soft decoration design scheme, the designer needs to collect samples from different manufacturers to design the interior soft decoration that meets the needs of users, which consumes a lot of manpower and material resources. Therefore, it is of great significance to design a three-dimensional simulation system for interior soft decoration, which can bring great convenience to interior soft decoration designers and enable users to understand the designed soft decoration more intuitively. With the rapid economic development, people's living standards have improved, and they gradually pay attention to the quality of their home decoration. The so-called family decoration quality not only refers to the quality of furniture materials but also the color matching and living comfort. The decoration quality can directly reflect the cultural connotation and life taste of residents. Interior design is to create a comfortable and pleasant living environment by using design, technical techniques, and aesthetics according to the use of nature, corresponding standards, and the environment of each building [2]. This indoor space not only has practical value but also reflects the aesthetic needs of personal preferences, cultural atmosphere, and folk feelings. The traditional 3D simulation system for indoor soft decoration usually selects the geometric modeling method to obtain the indoor scene information. Although the cost is low, the performance of the development tools is poor, resulting in the lack of authenticity of the system and the inability to obtain real color matching and decoration materials. Therefore, when formulating the interior soft decoration scheme, designers usually need to communicate with customers and modify the design through software that can directly display the interior design [3]. Therefore, a three-dimensional simulation system capable of accurately simulating the interior soft decoration is of great significance to improve the quality of home decoration design.

2. Literature Review

The three-dimensional modeling technology is a three-dimensional virtual modeling technology. Its principle is to use the camera to take 360° ring shots so as to master the detailed status of the building's interior landscape and then use computer technology to restore the building's interior landscape in a real and three-dimensional way [4]. As an advanced virtual reality technology, the introduction of three-dimensional modeling technology can realize the three-dimensional display and virtual display design of the building interior landscape by creating a three-dimensional indoor scene [5]. Therefore, this paper proposes an interior landscape design system based on three-dimensional modeling that realizes the virtual display design of an interior landscape by generating image workflow, analysis result splicing, information digital analysis, panoramic information recording, etc. [6]. The interior landscape design system based on three-dimensional modeling can not only draw the plane interior landscape image but also generate the three-dimensional interior landscape image through three-dimensional modeling. Through three-dimensional modeling, it can provide customers with more clear and three-dimensional interior landscape details, which is convenient for customers to choose from. In the design communication stage, communication and negotiation are the most important links in the preparation stage. We will obtain the design-related intentions from customers and clarify the basic conditions, style intentions, functional requirements, modeling requirements, and other contents of the design. Make a general assessment of the physical and geographical environment around the building and the style characteristics of the building. Second, in the design conception stage, the overall design of the space is considered according to the functional requirements and design intention proposed by the customer. The design concept should be considered in relation to the indoor function layout, the style and genre of interior design, the use of decorative materials, soft decoration, and color matching. The third stage is the design and drawing stage. After communicating with the customer, the design scheme will be drawn up into detailed construction drawings using relevant software. Then, according to the construction drawing, draw the effect drawing with software and finalize the scheme. CAD software can be used in the industry to draw the design plan. The construction drawing belongs to the early-stage auxiliary software. After completing the plan according to the customer's requirements, 3DMAX is used to turn the plan into a three-dimensional space effect drawing so that the customer can more intuitively and objectively understand the design content, print it into a position drawing, and bind it into a volume [7]. The fourth is the construction stage, which is the implementation stage of the whole design. Each type of work completes the whole decoration process according to the early design drawings, construction drawings, and effect drawings. Fifth, in the stage of soft decoration, the designer arranges the indoor living space according to the principle of design style. "In the soft decoration design, no matter which type the design style belongs to, we should grasp the background color, main color, embellishment color, and their proportional relationship" [8].

On the basis of current research, this paper designs a 3D simulation system of indoor soft decoration based on OpenGL and Direct3D. The system combines OpenGL software with Direct3D software, which effectively improves the authenticity of 3D simulation of indoor soft decoration, makes customers truly feel the effect of indoor soft decoration design and layout, and improves the quality of indoor soft decoration design.

3. Research Methods

3.1. System R&D Platform

To determine the R&D platform of the system, a Windows7 system is adopted as the operating system, OpenGL and Direct3D are used as 3D program development tools, and 3dsMAX7.0 is used as a 3D model and animation tool [9]. OpenGL and Direct3D are both program interfaces. In the process of modeling, the OpenGL graphics library can provide simple point, line, and polygon drawing functions, as well as relatively complex curve and surface drawing functions. In the process of cross-platform application research and development, OpenGL can work on UNIX, Windows7, and MAC platforms, and its architecture can enable the desktop system to transfer the indoor software graphics processing to the server. Direct3D is a 3D drawing programming interface developed through the Microsoft Windows operating system, which is a part of DirectX. Direct3dapi abstracts hardware features in the same way, so that different 3D acceleration hardware features can be hidden.

In order to make the soft model more realistic, the most commonly used 3dsMAX7.0 software is used as the production tool for the 3D model. For more complex models, plug-ins related to 3DMAX may be used [10].

3.1.1. Direct3D Architecture. Figure 1 depicts the Direct3D architecture.

Direct3D is mainly composed of two types of drivers: the first type is implemented through the hardware abstraction layer Hal [11]. Hal is an abstraction layer directly related to hardware, and it is also a driver, which is generally provided by the manufacturer. When the hardware can directly support the required functions, the hardware abstraction

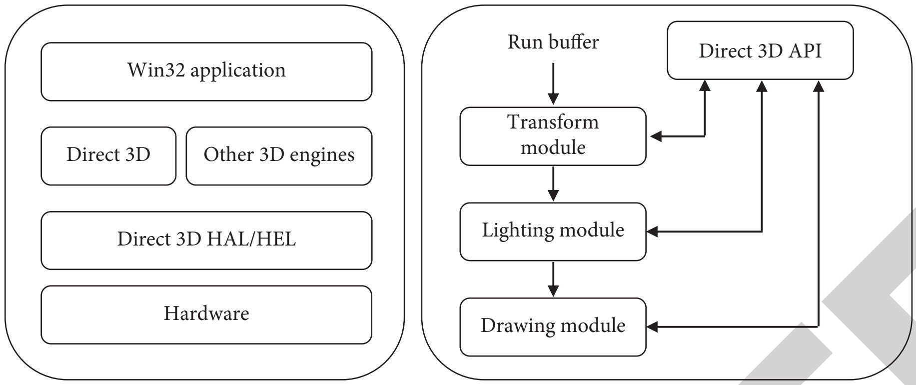


Figure 1: Direct3D architecture diagram.

layer will be applied so as to achieve acceleration; the other type of driver is realized through the hardware simulation layer HEL. When the hardware cannot support the required functions, the hardware simulation layer will simulate these functions through software to complete the specified tasks.

Direct3D includes a large number of c/c++ header files and interfaces for component object models [12]. Among them, there are many kinds and data structures in the header file, which transmit device information to the operating system and control the appearance of the 3D model according to this information. The 3D scene production of indoor soft decorations is realized through Direct3D. The whole process is divided into three levels, and each level is an independent dynamic loading software module. At the beginning of production, build corresponding graphics for each module through direct3dapi calls, and then introduce the running cache [13]. The first-floor module is the transformation module, which processes the geometric transformation of interior decoration. The lighting module is located on the second floor, which mainly calculates the indoor lighting and can process different types of light sources. The third layer is the rendering module, which constructs the final indoor soft decoration scene according to the output content of the upper two layers. For the above modules, the software can be used to convert them, so modules with strong functions can be used to replace the corresponding modules. Some modules can communicate with the hardware accelerator. At this time, any level in the whole process can be accelerated by hardware.

3.1.2. Design of User Login Registration Module. The user can enter the user's name and password in the login box in the system login interface to log in to the system. Unregistered users can log in to the system through registration. After successful login, they can perform various operations on the system. The user registration and login diagrams are shown in Figure 2.

3.2. Layout of Indoor Soft Decoration Agency Area. Soft decoration includes indoor furniture, household appliances, lamps, ornaments, and other objects, which are defined as scene elements, and on this basis, the agent area is defined, which refers to a group of objects that can realize a certain function together.

The essence of rational distribution of agent regions is to determine the optimal distribution position in the target space. Collect prior knowledge from interior decoration, combine it with functional requirements to determine constraint conditions, form a constraint cost function based on this, and solve this function so as to realize a soft decoration layout [14].

Use $A = (Q, S)$ to describe the layout of the proxy area in the target space, use Q to describe the proxy area set, $Q_i = \{Q_1, Q_2, \ldots, Q_k\}$, and use s to describe the feature set of the target space. Let the i-th proxy region $Q_i = (g, (a, b), \alpha, l, w)$, where G is the function type, (a, b) is the location, α is the direction, l is the length of the rectangle, and w is the width of the rectangle. The vertex coordinates at the lower left corner of the rectangle are the position of the proxy area [15].

The constraint function refers to the compliance with the layout constraints of the proxy area layout. For indoor soft decoration, the element layout must meet the requirements of function and vision [16]. This section gives three constraint functions to ensure the support and visual effect of indoor activities.

Constraint 1: through wall constraint [17]. Indoor soft objects need to be in the scope of the layout container. In the agent area layout problem, the phenomenon that the agent area penetrates outside the wall is punished by the above constraints, and the through wall constraint is obtained by the difference between the rectangle describing the agent area and the polygon describing the indoor contour:

$$M_1(A) = \frac{\sum_i B(Q_i - S)}{\sum_i B(Q_i)}, \tag{1}$$

where $B\ (*)\cdot$ is used to describe the polygon area calculation function.

Constraint 2: overlapping constraint [18]. When determining the optimal layout position, the objects shall be reasonably arranged; that is, the indoor soft fittings shall not overlap, and this shall be taken as a constraint. In the process of indoor soft decoration, for the layout of objects, the overlapping of agent areas will reduce the

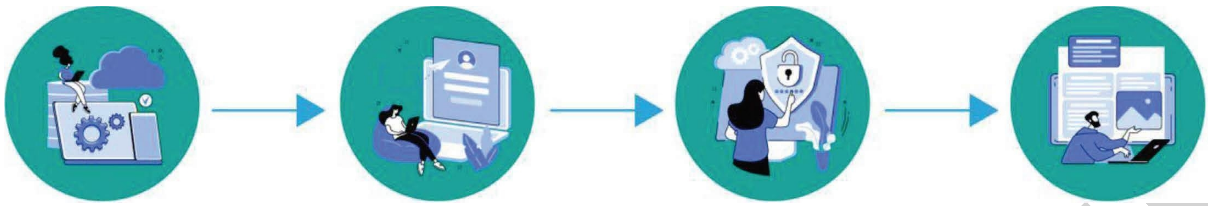


Figure 2: User registration and login diagram.

functions of objects in the area and affect the traffic, so it is necessary to prevent the overlapping of agent areas. The constraint function formed at this time is

$$M_1(A) = \frac{\sum_{1<i,j\in i<j} B\left(Q_i \cap Q_j\right)}{\sum_i B\left(Q_i\right)}. \tag{2}$$

Constraint 3: dimensional constraint [19]. From the aspect of comfort, when arranging indoor soft decorations, attention should be paid to the size of each object. There are certain requirements for its proportion in the interior. According to this constraint condition, the constraint function is determined as follows:

$$M_3(A) = \begin{cases} 1 - \left(\dfrac{A^*}{s_{\min}}\right)^2, & A^* < s_{\min}, \\ 0, & s_{\min} \leqslant A^* \leqslant s_{\max}, \\ 1 - \left(\dfrac{s_{\max}}{A^*}\right)^2, & A^* > s_{\max}, \end{cases} \tag{3}$$

$$A^* = \frac{\sum_i B\left(Q_i\right)}{B(S)},$$

where $s_{\max}$ is the maximum value of indoor area ratio and $s_{\min}$ is the minimum value of indoor area ratio. Here, $s_{\max} = 0.48$ and $s_{\min} = 0.36$ are taken.

According to the above three constraints, the total cost function is determined by a linear combination, $M(A) = \sum_{i=1}^{3} \omega_i M_i(A)$, and the indoor soft decoration layout can be realized by solving this function [20].

4. Result Analysis

In order to test the three-dimensional simulation of indoor soft decoration by the three-dimensional simulation system of indoor soft decoration based on OpenGL and Direct3D, this system is used to simulate the three-dimensional simulation of indoor soft decoration of a residential model room in a community. The sample room includes two bedrooms, a kitchen, and a living room. The system in this paper is compared with the Untiy3D system and the Smart3D system.

The system is mainly operated with the mouse and keyboard. It edits the three systems with different functions and counts the function operations and time consumption of the three systems. The test results of the operation of the three systems are shown in Table 1. By comparing the operation results of the three systems in Table 1, it can be seen that the three systems can achieve the basic operation of the system. The average running time of editing operations in this system is 327 ms, while the average running times of editing operations in the Untiy3D system and the Smart3D system are 755 ms and 637 ms, respectively. According to the statistical results, this system runs faster for editing operations and has better operation performance.

Make statistics on the indoor soft decoration of each room in the 3D simulation sample room of the system and compare the system with the Untiy3D system and the Smart3D system. The statistical results are shown in Figure 3. It can be seen from the statistical results in Figure 3 that the time used for indoor soft decoration 3D simulation of the system in this paper for 4 rooms is significantly lower than that of the Untiy3D system and the Smart3D system, and the time used for simulation of 4 rooms does not exceed 8000 ms. It shows that the system in this paper can complete the soft decoration 3D simulation of the target room in a short time and verifies the 3D simulation performance of the system in this paper.

Select 200 home decoration designers and divide them into 10 groups, so that the 10 groups can use the three systems for interior soft decoration design. After the completion of the design, 200 home decoration designers were counted, and the satisfaction of the three systems was verified by questionnaire. The survey's contents include favorable impression, operation sensitivity, design convenience, design satisfaction, and soft-fitting effect. The scoring results of 10 groups on the three systems are shown in Tables 2–4.

According to the survey results in Tables 2–4, the favorability, operation sensitivity, design portability, design satisfaction, and soft decoration effect of the system in this paper are higher than those of the Untiy3D system and the Smart3D system. The average score of the system in this paper is about 95, while the average score of the Untiy3D system is lower than 89, and the average score of the Smart3D system is lower than 85, indicating that the indoor soft decoration 3D simulation effect of this system is good, the design is convenient, the operation is sensitive, and the user satisfaction is high.

Table 1: Comparison of system operation.

Function	Text system		Untiy3D system		Smart3D system	
	Function operation key	Operation time (ms)	Function operation key	Operation time (ms)	Function operation key	Operation time (ms)
Shift left	4	251	4	725	4	574
Shift right	6	275	6	765	6	562
Forward	8	316	8	718	8	581
Back off	2	328	2	811	2	628
Narrow	q	298	q	765	q	647
Enlarge	W	364	W	724	W	685
Rotate	e	385	e	764	e	617
Pick operand	r	374	r	728	r	721
Distance measurement	t	396	t	734	t	768
Direction discrimination	u	374	y	807	a	592
Object tracking	o	285	u	826	S	648
Object search	p	274	i	698	d	624

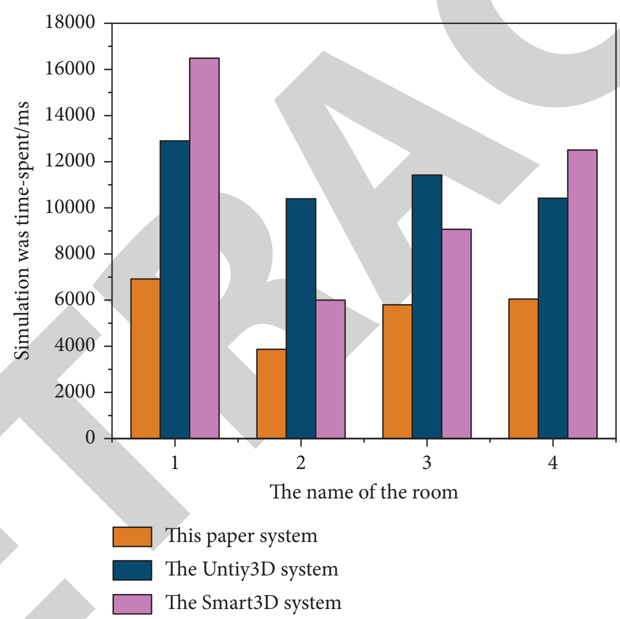


Figure 3: Time comparison among 3D simulation samples of three systems.

Table 2: Scoring results of the system.

Group number	Favorability	Operating sensitivity	Design convenience	Design satisfaction	Soft effect
1	95.85	95.74	94.85	95.27	94.28
2	96.85	94.28	95.27	94.38	93.58
3	93.54	96.85	96.37	96.38	96.58
4	94.71	95.47	95.74	92.68	95.37
5	96.27	93.28	93.27	93.58	94.57
6	95.17	93.47	94.85	94.69	95.69
7	94.85	95.74	96.37	95.74	94.37
8	95.68	94.85	94.28	94.68	92.58
9	96.74	95.27	93.68	96.28	96.57
10	93.87	96.04	94.58	95.47	94.58
Mean value	95.35	95.10	94.93	94.92	94.82

Table 3: Scoring results of Untiy3D system.

Group number	Favorability	Operating sensitivity	Design convenience	Design satisfaction	Soft effect
1	87.52	86.54	89.04	88.29	86.25
2	86.74	89.57	87.49	89.49	87.19
3	82.47	87.59	88.49	86.19	89.28
4	83.57	88.57	87.59	85.49	90.18
5	86.57	85.16	85.49	87.29	88.49
6	85.27	87.59	86.28	91.08	87.49
7	86.75	88.29	88.27	89.27	86.29
8	87.57	87.28	89.27	88.17	85.19
9	89.57	86.27	87.59	89.28	84.27
10	88.57	87.19	86.28	86.18	86.25
Mean value	86.46	87.41	87.58	88.07	87.09

Table 4: Scoring results of Smart3D system.

Group number	Favorability	Operating sensitivity	Design convenience	Design satisfaction	Soft effect
1	87.18	88.18	89.18	85.17	86.18
2	85.18	83.54	84.28	83.16	84.15
3	83.17	87.67	83.16	84.23	83.26
4	82.49	86.18	82.47	80.27	81.48
5	83.16	85.67	81.68	86.18	83.58
6	85.49	81.69	86.37	85.19	84.75
7	86.37	82.47	85.18	86.27	79.68
8	81.28	84.58	81.25	81.27	86.27
9	80.16	83.68	80.18	82.48	87.18
10	81.68	82.68	81.68	83.25	88.92
Mean value	83.62	84.63	83.54	83.75	84.54

5. Conclusion

In order to facilitate the interior designers in better designing the interior soft decoration and improve the authenticity and real-time performance of the 3D simulation effect of the interior soft decoration, this paper designs the 3D simulation system of the interior soft decoration based on OpenGL and Direct3D. The system combines OpenGL software with Direct3D software. Firstly, 3dMAX software is used to establish the indoor model, and then OpenGL software is used to add and constrain the indoor scene elements to enrich the indoor model. Finally, this paper carries out the final layout and rendering of furniture through Direct3D software to realize an effective three-dimensional simulation of indoor soft decoration and provide a reliable analysis basis for indoor design.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] J. Loy-Benitez, S. K. Heo, and C. K. Yoo, "Soft sensor validation for monitoring and resilient control of sequential subway indoor air quality through memory-gated recurrent neural networks-based autoencoders," *Control Engineering Practice*, vol. 97, pp. 104330–1043, 2020.

[2] A. Purnomo, "Interior design study of lecturer rooms at e12 building 2 nd floor unnes – learn from covid-19 pandemic," *IOP Conference Series: Earth and Environmental Science*, vol. 700, no. 1, Article ID 012051, 2021.

[3] J. Yang and S. Yang, "Design of urban landscape visual simulation system based on three-dimensional simulation technology," *International Journal of Industrial and Systems Engineering*, vol. 1, no. 1, p. 1, 2020.

[4] A. Manukovskiy, D. Kurdyukov, and V. Korotkov, "Experience of application of elements of information modeling technology," *Actual directions of scientific researches of the XXI century theory and practice*, vol. 8, no. 1, pp. 100–105, 2020.

[5] X. Man, F. Duan, Y. Piao et al., "Research on simulation ideas of relay protection device on load test based on virtual reality technology," *Journal of Physics: Conference Series*, vol. 2005, no. 1, Article ID 012196, 2021.

[6] S. Tkalenko and N. Sukurova, "The influence of fdi on sustainable economic development of Ukraine in terms of global digitization," *Journal of Digital Science*, vol. 2020, no. 1, pp. 40–53, 2020.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 5964054, 8 pages
https://doi.org/10.1155/2022/5964054

Research Article

Application of Informational Big Data in Case Study and Collection of Basic Educational Psychology

Juan Nie and Jianghong Mo

Guilin Tourism University, Guilin, Guangxi 541006, China

Correspondence should be addressed to Jianghong Mo; mojianghong0316@gltu.edu.cn

Received 20 July 2022; Revised 12 September 2022; Accepted 17 September 2022; Published 4 October 2022

Academic Editor: Jiafu Su

In the new situation of modernization, there is an influx of diverse social thinking. At the same time, coupled with the influence of COVID-19, which has swept the world, the ideological and psychological space of college students has been greatly impacted. In this context, the ideological and psychological health of college students is an important value for the education of college students. To be specific, as an important place for cultivating college students, colleges and universities should pay attention to students' thoughts and ideas from their hearts. In addition, colleges and universities should give full play to the role of educational psychology in colleges and universities and actively promote the synergistic development of all educational sectors in schools, so as to promote the realization of the goal of education in the new era. In recent years, a series of mental health problems such as anxiety, depression, low self-esteem, and interpersonal sensitivity have become frequent among college students, and some have even developed suicidal ideation. This has a very serious negative impact on individuals, families, and society. Therefore, if the mental health problems of college students can be detected early, the relevant school departments and counselors can provide timely and targeted help to such students. At the same time, these at-risk students can receive early treatment, thus reducing the harm. As a result, it is quite valuable to find an effective method to identify students with mental health problems. Traditionally, researchers have used questionnaires to survey students about their mental health. However, this approach has the disadvantage of being easily concealed and inefficient. In recent years, researchers have begun to use weblogs to identify students with mental health problems, but this approach still has shortcomings. First of all, they still use questionnaires to obtain labels. In addition, students' psychological activities may not only be reflected in their online behavior but also in their other daily behaviors. Big data in higher education plays a crucial role in analyzing and identifying students with psychological abnormalities. As a result, this research mainly extracts the behavioral characteristics of students by cleaning and transforming a large amount of disorganized student school data based on the educational data collected from school cards, academic affairs systems, access control systems, and related business systems. What is more, this study further analyzes the differences in behavioral characteristics between normal and abnormal students through hypothesis testing and finally establishes a model to identify abnormal students and evaluate the results.

1. Introduction

With the rapid development of information technology, artificial intelligence and big data are influencing various industries [1–3]. Nowadays, digitalization and informatization in colleges and universities have been promoted and improved, which has increased the amount of data in research results, students' daily behavior, academic performance, and network usage [4]. In this context, nowadays, there are relatively complete application systems among different departments of the university, such as academic application systems, access control systems, one-card data systems, and library circulation systems. Each of these systems stores a large amount of student data [5]. However, these data are only stored in the database and are not exploited. Therefore, one of the most important issues is that how to use a large amount of data effectively, analyze the information in the data through data mining technology, make the data into useful information, and accelerate the level of information construction in universities [6]. At the same time, with the arrival of big data, data mining technology has also flourished. Since the 20th century, data mining technology has been widely used in various industries [7]. In the field of education, many universities at home and abroad have started to use data mining technology to analyze educational data in order to understand students' various dynamics in time.

While China has achieved rapid economic development, the mental health of the population also needs attention [8]. To be specific, numerous psychologically related disorders are present in the population. These mental illnesses have many negative effects and are pervasive and harmful in nature [9]. As a result, mental health has become an important public health issue in our country. At the same time, with the rapid development of the knowledge economy and the increase in the popularity of higher education, the number of college students is increasing, and the number of students with psychological problems is also increasing [10, 11]. In recent years, there have been many cases of university students' psychological problems affecting their studies, dropping out of school, and even committing suicide. College students, as part of the high level of talent, have been considered to be outstanding members of society and are expected to have strong psychological qualities [12]. However, the reality is very disappointing. Specifically, the stresses of academics, social relationships, and employment have led to many tragedies that have left college students mentally exhausted [13]. As a result, the timely identification of students with psychological abnormalities has become one of the most essential and challenging issues for universities [14].

Mental health problems can have serious detrimental effects. For the individual, mental health problems can have a number of negative effects on the individual [15]. More seriously, some mental health problems may make individuals less adaptable to society and may even pose a serious threat to their physical health [16]. In addition, mental health problems can be a psychological burden to the family and an economic burden to the family. For society, mental health problems can disrupt public order, consume social resources, and increase the burden of disease on society [17]. As a result, mental health problems account for the majority of the negative effects of all illness-related hazards [18]. However, due to a lack of mental health knowledge and discriminatory attitudes toward the mentally ill, a large number of people with mental health problems are not motivated to seek professional help [19].

Therefore, in the context of the new era, how the ideological and political education work of college students can adapt to the policy, the call of the times, and the needs of the education subjects on the basis of maintaining their working principles has become an important issue of the times [20]. In recent years, relevant departments have proposed to attach importance to humanistic care and take it as an important method to strengthen and improve ideological and political work [21]. Therefore, psychological education has become one of the ten major education systems for ideological and political work in domestic universities. This can provide an important basis for the effective combination of ideological and political education and psychological guidance. At the same time, it also shows the future trend of ideological and political work of the party and the country [22]. Psychological guidance plays an indispensable role in daily ideological and political work. For example, during the outbreak of COVID-19 in 2020, psychological guidance played an active role in responding to sudden public events. This has inspired people to pay more attention to humanistic care in the current ideological and political work, always pay attention to students' physical and mental health, and constantly enter their hearts [23]. Only in this way can we better cultivate a healthy social mentality and a sound personality among college students and lead young students to help realize the Chinese dream of the great rejuvenation of the Chinese nation [24].

Currently, research on mental health issues is still dominated by questionnaires. However, there are many shortcomings in this approach. First and foremost, the scale of questionnaires is small. To be specific, both paper and electronic questionnaires require the participation of the subjects. As a result, there are a number of ways to attract subjects [25]. For example, each participant is paid a certain amount of money or a lottery is offered. However, many students are still reluctant to participate. What are more, questionnaires take a certain amount of time. It takes a few minutes to fill out each questionnaire, and some questionnaires can take up to half an hour [26]. Moreover, due to the lack of knowledge about mental health and the discriminatory attitude of college students towards mental health patients, a large number of people with mental health problems hide the truth when answering the questionnaires [27]. This can lead to a large discrepancy between the test results and the real situation. Last but not least, the mental health status of each student changes over time, usually lasting half a month or a month. As a result, if researchers want to know the mental health status afterward, they need to take the test again, which is both labor-intensive and costly.

The outbreak of COVID-19 has affected all walks of life, and the teaching format of schools has changed from offline to online. Therefore, colleges and universities have become important places for the prevention and control of the epidemic, which makes it more difficult for teachers to grasp the psychological state of mind of students [28]. This urgently requires the ideological and political education work in colleges and universities to insist on the combination of educating the mind and educating people. To be specific, colleges and universities should not only do a good job in ideological and political education but also pay attention to cultivating a sound personality and great psychological quality of college students [29]. As a result, under the background of a diversified society, college students face many problems such as academic pressure, employment pressure, interpersonal relationship trouble, and wrong thinking at home and abroad. In addition, many pressures make the desire to be valued and to have an equal communication status while having a strong sense of self [30].

Psychological counseling is a scientific psychological treatment method and also an important supplement to ideological and political education methods, which plays an important role in improving the educational effect and cultivating a positive and optimistic mindset of college students. In addition, ideological and political education in colleges and universities should meet the inherent demand of improving its effectiveness [31]. Modern ideological and political education should not only adhere to Marxism as the theoretical basis but also draw on the excellent achievements of different disciplines. The application of psychological guidance to ideological and political education in colleges and universities is not only an expansion and innovation of ideological and political education methods but also an effective reference for the new problems in ideological and political work [32]. In practice, ideological and political education should continuously improve its service function and highlight its social function [33].

With the advent of the Internet era, digital campuses have flourished. In this context, the campus life and learning data of college students are gradually recorded in the student management system. These massive data contain a huge amount of information and provide a better opportunity to understand students' behaviors. This study considers that the psychological status of college students is reflected by their behavioral status in school and is recorded in the relevant system. As a result, the purpose of this study is to identify students with mental health problems by analyzing the massive amount of data on campus and mining the behaviors related to mental health, so that machine learning algorithms can be effectively used to identify students with mental health problems. Furthermore, with the use of information-based big data, counselors can focus on these students based on the identified results and provide them with psychological guidance.

2. Technology Related to Big Data

2.1. Data Mining. With the rapid development of computer and network technologies, many human activities are no longer limited by time and space, such as telecommuting and information sharing. Increasingly sophisticated database technologies and pervasive data applications have led to the exponential growth of human data. In this context, people have started to explore the potential information from these massive data, resulting in a variety of data mining techniques. Among them, educational data mining is a new technology created to make full use of the huge data streams that emerge from the operation of digital campuses. To be specific, this technology integrates, classifies, and refines a large amount of data through the comprehensive application of various data mining technologies, thus making it very valuable in the ever-evolving education. As a result, educational data mining refers to a multidirectional and cross-cutting research area that encompasses three main disciplines: education, computer science, and statistics. At the same time, their combination has given rise to research areas similar to or related to educational data mining, such as computer-based education, machine learning, data mining, and learning analytics.

As shown in Figure 1, data mining can be generally divided into three phases: data collection and preprocessing, data analysis and mining, and evaluation and visualization of results. The first stage is data collection and preprocessing. This process mainly includes the collection of target data, the establishment of a data warehouse, data cleaning, data integration, statute and transformation, and other operations. To be specific, data collection and preprocessing are the foundation and key to the whole data mining process. In other words, the quality of data can determine the effect of data analysis and mining. By establishing a unified data warehouse, noise and missing data can be removed or filled with certain rules. Data integration is the unified management of data to facilitate subsequent use. Data preprocessing is a tedious and heavy workload part, but it also lays a solid foundation and support for the subsequent mining analysis. The second stage is the analysis and mining of data. This process is the core part of the whole data mining process. After using the preprocessed data, the data can be extracted, filtered, and modeled by using certain algorithms according to the data mining objectives. The third stage is the evaluation and visualization of the results. Finally, the mining model is evaluated to measure whether the model fulfills the expected tasks and objectives. The results of data mining are analyzed and filtered through certain visualization techniques, and useful results are selected and presented.

2.2. Random Forest Algorithm. The random forest algorithm mainly applies a decision tree as the base classifier, combined with the construction of Bagging, and incorporates the selection of random attributes. As a result, the detailed workflow of the random forest algorithm is illustrated in Figure 2. To be specific, the random forest algorithm can decide the class of samples by voting, and each base classifier has one vote. After that, the category of the sample can be determined based on the principle of majority rule. The randomness of the random forest is reflected in the randomness of attribute selection. Specifically, for each node of the base classifier, k attributes are first randomly selected from the set of attributes of the current node to form a subset. Then, the best splitting attribute is computed from the subset according to the attribute selection algorithm, instead of finding the best attribute from all attributes of the current node as in a traditional decision tree. As a result, the introduction of the parameter k reflects the randomness of the random forest algorithm. The random forest algorithm can effectively solve the problem of overfitting easily in decision trees. When dealing with high-dimensional data, it is not necessary to perform dimensionality reduction and feature selection before training because of the random property. In addition, since each tree is trained independently, parallel training can be realized, thus speeding up the training speed. At the same time, the error of the dataset can be balanced according to the error generated by each tree. However, when the noise of the dataset is high, overfitting may occur.

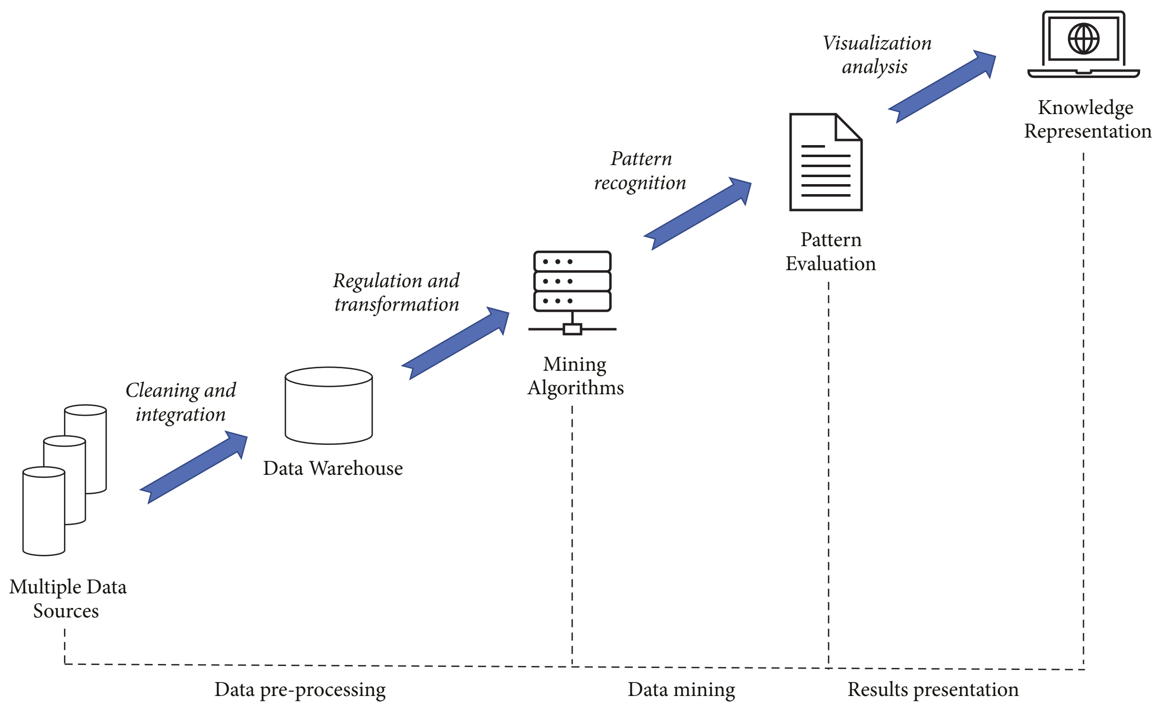


Figure 1: Process of data mining.

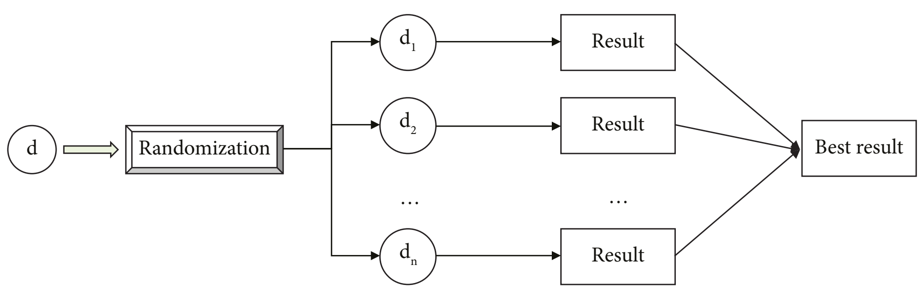


Figure 2: Workflow of the random forest algorithm.

2.3. Description of the Hypothesis Test. Hypothesis testing is a common method used in statistics to determine whether a random variable is consistent with a scientific hypothesis. Its main idea is the small-probability counterfactual method. To be specific, small probability generally refers to the probability of occurrence of 1% or 5% in a single experiment. The counterfactual method is to first formulate a hypothesis and then test it by a statistical method of sampling studies. In fact, the theory and method of determining the probability of the hypothesis are valid based on the principle of small probability, and the theory and method of determining the validity of the hypothesis make up the hypothesis testing. Figure 3 shows the confidence level and rejection domain of the hypothesis test.

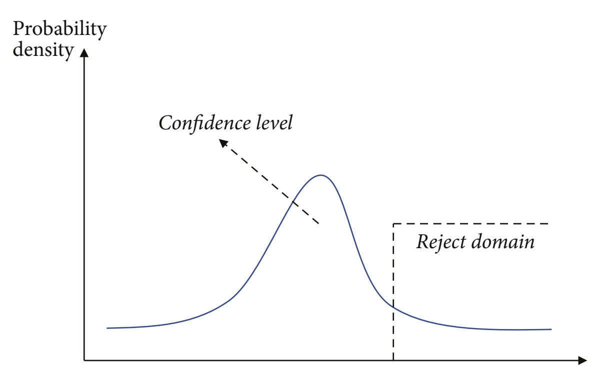


Figure 3: Confidence level and rejection domain of the hypothesis test.

2.4. Logistic Regression Model. Logistic regression is the transformation of a linear model into a nonlinear model to predict the actual test values and output as much as possible for a given data set. Although it is called regression, the actual solution is a classification model. Compared with other algorithms, logistic regression is simple, interpretable, and fast to train. As a result, this algorithm is mainly suitable for linear classification problems. The main idea is to fit the decision boundary as closely as possible and output the predicted values. The main idea of logistic regression is to transform the results of linear regression into the classification of predicted values by mapping them to functions. As a result, the sigmoid function as illustrated in Figure 4 is usually chosen as the mapping function, and its expression is

$$g(z) = \frac{1}{1 + e^{-z}}, \quad (1)$$

where z refers to the input value.

The model classification results can be optimized by the loss function of the model. When solving for the parameters in the model, the loss function is minimized by using the gradient descent method. The logistic regression classification model is simple to implement, computationally small, and easy to understand and implement. However, when the feature space is large, the classification performance is affected.

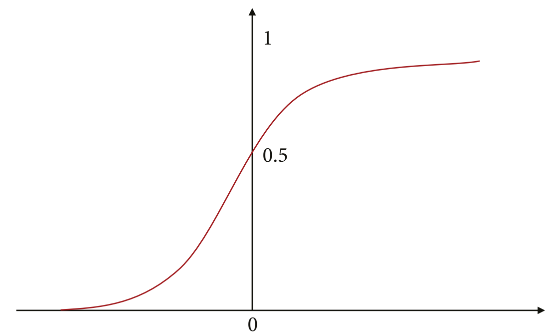


FIGURE 4: Sigmoid function.

3. Case Study of Basic Educational Psychology

3.1. Algorithm Framework. The overall framework of the educational psychological problem identification algorithm based on big data technology is shown in Figure 5. The whole algorithm process is divided into three parts, including data acquisition and preprocessing, feature extraction, and model training and recognition. In the data acquisition and preprocessing stage, four data sources are acquired, namely, weblogs, access control data, achievement data, and consumption data. In the feature extraction stage, relevant features such as students' online patterns and abnormal consumption scores can be extracted from the four data sources. In the model training and identification stage, the optimal classifier is selected among five common classification algorithms.

With the rise of digital campuses, more and more student behavior data are stored. These data have two characteristics, one is a large amount of data and the other is the complexity and diversity. So far, although various student management systems have been established in various universities in China and abroad, the data collected by these systems are still not well utilized. As a result, it is quite necessary to understand and analyze these data and establish relevant data models.

3.2. Student Behavior Characteristics. In the mining of educational data in colleges and universities, most scholars tend to focus only on the characteristics related to students' performance. This research focuses on extracting a comprehensive set of behavioral characteristics of students from the Big Five personality traits theory. Based on the original data, this study further extracts and quantifies the characteristics of students' behaviors, aiming at extracting the characteristics that can reflect the students' behaviors in school more intuitively. In addition, based on the basic data of students, the indexes can be established by referring to psychology and pedagogy in order to extract the characteristics of students. In addition, this research constructs a better behavioral profile to measure students' performance in school, so that we can detect abnormal students from their daily behavior.

Indeed, students with psychological abnormalities manifest in different ways in the relatively open and free environment of the university. Most students focus on the Big Five personalities of openness, responsibility, extroversion, agreeableness, and neuroticism (Figure 6). However, due to limitations in school data, this paper can only analyze certain traits to compare and contrast the differences between psychological abnormalities and normal students in order to better identify students with possible psychological problems from a large number of students promptly.

3.3. Social Relationship Mining. Traditional discovery of buddy networks is usually done by hypothesis testing or clustering. However, hypothesis testing requires the assumption that cooccurrence between students is independent, while clustering results are strongly influenced by the cluster centroids. Therefore, the main idea of this study is that the number of cooccurrences between two students is proportional to the probability of two students becoming friends. In other words, this study uses association rules to identify friends. On campus, it is assumed that if two students are friends, they will often be in the same place together, such as eating or spending money together, going to the library together, and so on. This is consistent with the pattern of friendships between college students. As shown in Figure 7, if students X and Y are friends, then students X and Y will be seen together many times.

As information technology and databases continue to evolve and mature, the size of data is exploding. There are more and more unbalanced data in various industries. However, since most classification algorithms are based on global balanced data, the prediction results of most unbalanced data are not promising. Therefore, it is necessary to equalize the data. Currently, there are two main types of unbalanced data processing. The first one is to resample the data from the data level to eliminate the majority samples or expand the minority samples. This can change the distribution of the original data samples and thus improve the imbalance of the data. The other is an improvement of the classification model algorithm. Specifically, the original

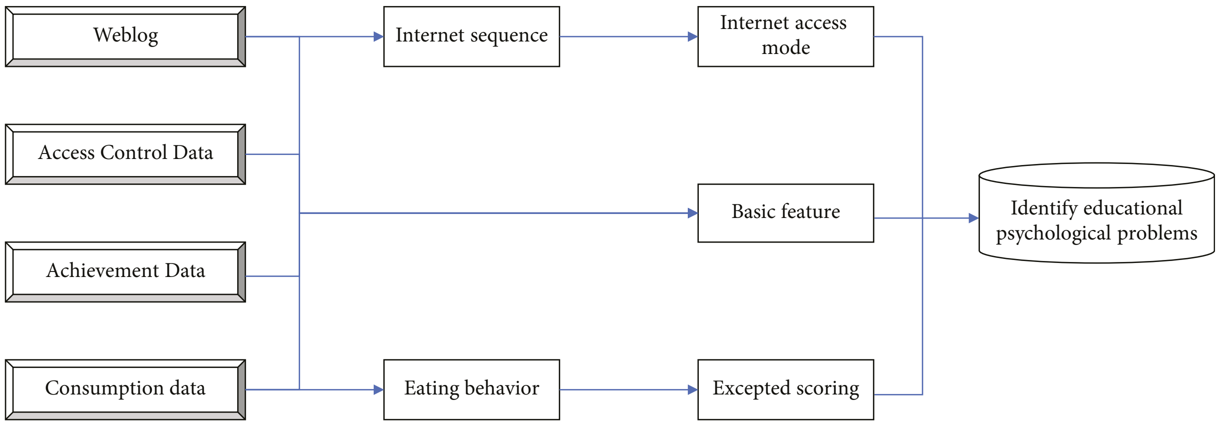


Figure 5: Framework of educational psychological problem identification algorithm.

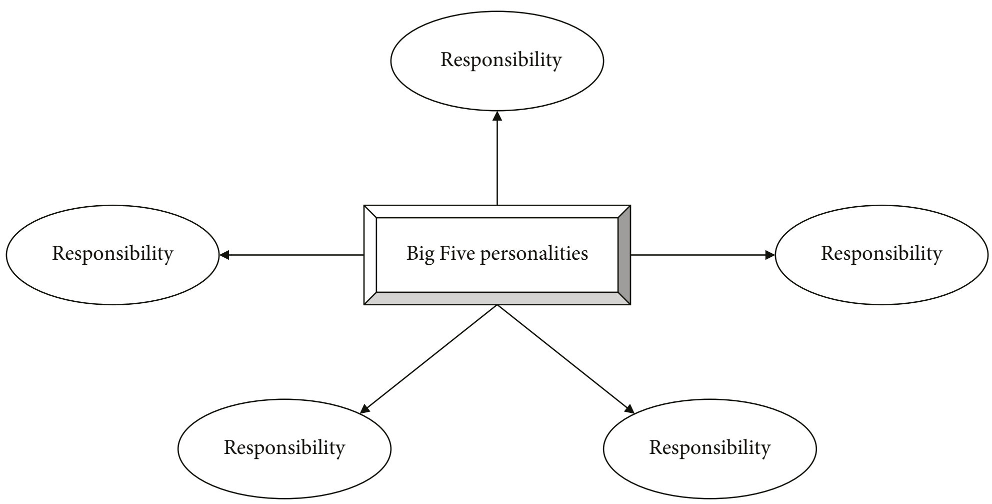


Figure 6: Big Five personalities.

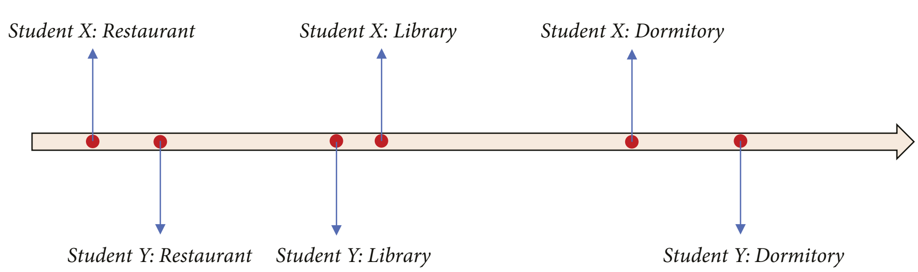


Figure 7: Student locations altogether indicate.

classification algorithm needs to be improved by introducing a penalty mechanism for misclassification costs so that the algorithm can be applied to unbalanced data. However, due to the diversity of data, unbalanced data usually have their unique characteristics. Therefore, it is difficult to improve the algorithm for unbalanced data, and the generalization ability and generalizability are not high. As a result, in general, most experts and scholars solve the unbalanced data problem at the data level.

4. Conclusion

The traditional way of a questionnaire survey on mental health problems has the problems of easy concealment and small scale. In recent years, a method of identifying mental health problems based on Internet logs has emerged. This method compensates for the shortcomings of questionnaires, but students' behaviors on campus are diverse. To be specific, Internet behavior is only a part of students' behaviors, and it is not enough to reflect all of their psychological activities. In addition, the method of identifying mental health problems based on Internet logs still uses questionnaires to obtain labels. Therefore, the labels are still unreliable. In this paper, we propose a model for mining students' social relationships in school by preprocessing a large amount of messy data and identifying students with psychological abnormalities. Based on this model, this research can extract other behavioral characteristics of students in school and analyze the difference between the performance of psychologically abnormal students and normal students in school. Furthermore, based on the preliminary data processing, a model for identifying students with psychological abnormalities was established to identify students with psychological abnormalities from a large amount of data.

However, the work done and the proposed method in this paper have the following areas for improvement in future research. In the model for identifying students with psychological abnormalities developed in this paper, due to the highly uneven data, the use of dichotomous classification requires data equalization to achieve better identification results. The one-category model is simple, but it does not contribute to the accuracy of the model. Therefore, further research is needed to establish a more convenient and accurate model for identifying students with psychological abnormalities based on big data from university education.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

The research was supported by 2021 National Social Science Fund General Project "Research on the Coupling Mechanism of Cultural Revitalization of Traditional Villages in Yunnan, Guizhou, and Guangxi and High-Quality Development of Rural Tourism" (21BMZ073); 2021 Guangxi University Young and Middle-Aged Teachers' Basic Research Ability Improvement Project "Research on the Influence Mechanism and Improvement Path of Farmers' Happiness in the Background of Rural Revitalization" (2021KY0825); 2018 Guangxi Zhuang Autonomous Region Philosophy and Social Science Planning Project "Guangxi Rural Culture in the Process of Tourism Urbanization' Research on the Coordinated Development of Ecology and Tourism" (18CGL001); and 2021 Research Project on the Theory and Practice of Ideological and Political Education for College Students in Guangxi "Research on the Influence Mechanism and Improvement Path of Guangxi College Counselors' Happiness in the New Era" (2021LSZ089).

References

[1] N. Singh, "Big data technology: developments in current research and emerging landscape," *Enterprise Information Systems*, vol. 13, no. 6, pp. 801–831, 2019.

[2] B. Cheng, J. Huang, K. Lu et al., "BIM-enabled life cycle assessment of concrete formwork waste reduction through prefabrication," *Sustainable Energy Technologies and Assessments*, vol. 53, (Part A), Article ID 102449, 2022.

[3] J. A. Lara, A. A. De Sojo, S. Aljawarneh, R. P. Schumaker, and B. Al-Shargabi, "Developing big data projects in open university engineering courses: lessons learned," *IEEE Access*, vol. 8, Article ID 23001, 2020.

[4] K. Vassakis, E. Petrakis, and I. Kopanakis, "Big data analytics: applications, prospects and challenges," *Mobile Big Data*, pp. 3–20, 2018.

[5] G. T. Reddy, M. P. K. Reddy, K. Lakshmanna et al., "Analysis of dimensionality reduction techniques on big data," *IEEE Access*, vol. 8, Article ID 54788, 2020.

[6] W. Villegas-Ch, X. Palacios-Pacheco, and S. Luján-Mora, "Application of a smart city model to a traditional university campus with a big data architecture: a sustainable smart campus," *Sustainability*, vol. 11, no. 10, p. 2857, 2019.

[7] L. W. Santoso, "Data warehouse with big data technology for higher education," *Procedia Computer Science*, vol. 124, pp. 93–99, 2017.

[8] C. K. Kaparounaki, M. E. Patsali, D. P. V. Mousa, E. V. Papadopoulou, K. K. Papadopoulou, and K. N. Fountoulakis, "University students' mental health amidst the COVID-19 quarantine in Greece," *Psychiatry Research*, vol. 290, Article ID 113111, 2020.

[9] M. J. Savage, R. James, D. Magistro et al., "Mental health and movement behaviour during the COVID-19 pandemic in UK university students: prospective cohort study," *Mental Health and Physical Activity*, vol. 19, Article ID 100357, 2020.

[10] R. Prowse, F. Sherratt, A. Abizaid et al., "Coping with the COVID-19 pandemic: examining gender differences in stress and mental health among university students," *Frontiers in Psychiatry*, vol. 12, Article ID 650759, 2021.

[11] T. M. Evans, L. Bira, J. B. Gastelum, L. T. Weiss, and N. L. Vanderford, "Evidence for a mental health crisis in graduate education," *Nature Biotechnology*, vol. 36, no. 3, pp. 282–284, 2018.

[12] J. J. Liu, Y. Bao, X. Huang, J. Shi, and L. Lu, "Mental health considerations for children quarantined because of COVID-19," *The Lancet Child & Adolescent Health*, vol. 4, no. 5, pp. 347–349, 2020.

[13] H. Fu, J. Niu, Z. Wu et al., "Influencing factors of stereotypes on wastewater treatment plants-Case study of 9 wastewater treatment plants in Xi'an, China," *Environmental Management*, vol. 38, pp. 1–10, 2022.

[14] A. Gulliver, L. Farrer, K. Bennett, and K. M. Griffiths, "University staff mental health literacy, stigma and their experience of students with mental health problems," *Journal of Further and Higher Education*, vol. 43, no. 3, pp. 434–442, 2019.

[15] S. P. Becker, M. A. Jarrett, A. M. Luebbe, A. A. Garner, G. L. Burns, and M. J. Kofler, "Sleep in a large, multi-university sample of college students: sleep problem prevalence, sex differences, and mental health correlates," *Sleep Health*, vol. 4, no. 2, pp. 174–181, 2018.

[16] G. Arslan, M. Yıldırım, Z. Karataş, Z. Kabasakal, and M. Kılınç, "Meaningful living to promote complete mental health among university students in the context of the

COVID-19 pandemic," *International Journal of Mental Health and Addiction*, vol. 20, no. 2, pp. 930–942, 2020.
[17] A. Kurtović, I. Vuković, and M. Gajić, "The effect of locus of control on University Students' Mental Health: possible mediation through self-esteem and coping," *Journal of Psychology*, vol. 152, no. 6, pp. 341–357, 2018.
[18] J. S. L. Brown, "Student mental health: some answers and more questions," *Journal of Mental Health*, vol. 27, no. 3, pp. 193–196, 2018.
[19] M. L. Zochil and E. B. Thorsteinsson, "Exploring poor sleep, mental health, and help-seeking intention in university students," *Australian Journal of Psychology*, vol. 70, no. 1, pp. 41–47, 2018.
[20] Y. Zolotov, A. Reznik, S. Bender, and R. Isralowitz, "COVID-19 fear, mental health, and substance use among Israeli university students," *International Journal of Mental Health and Addiction*, vol. 20, pp. 230–236, 2020.
[21] R. Lucas and A. I. James, "An evaluation of specialist mentoring for university students with autism spectrum disorders and mental health conditions," *Journal of Autism and Developmental Disorders*, vol. 48, no. 3, pp. 694–707, 2018.
[22] S. Cvetkovski, A. F. Jorm, and A. J. Mackinnon, "An analysis of the mental health trajectories of university students compared to their community peers using a national longitudinal survey," *Studies in Higher Education*, vol. 44, no. 1, pp. 185–200, 2019.
[23] I. Montagni, S. Qchiqach, E. Pereira, P. J. Tully, and C. Tzourio, "Sex-specific associations between sleep and mental health in university students: a large cross-sectional study," *Journal of American College Health*, vol. 68, no. 3, pp. 278–285, 2020.
[24] J. Huang, Y. T. Nigatu, R. Smail-Crevier, X. Zhang, and J. Wang, "Interventions for common mental health problems among university and college students: a systematic review and meta-analysis of randomized controlled trials," *Journal of Psychiatric Research*, vol. 107, pp. 1–10, 2018.
[25] A. Gulliver, L. Farrer, K. Bennett et al., "University staff experiences of students with mental health problems and their perceptions of staff training needs," *Journal of Mental Health*, vol. 27, no. 3, pp. 247–256, 2018.
[26] S. M. M. Sæther, M. Knapstad, K. G. Askeland, and J. C. Skogen, "Alcohol consumption, life satisfaction and mental health among Norwegian college and university students," *Addictive Behaviors Reports*, vol. 10, Article ID 100216, 2019.
[27] M. Visser and E. Law-van Wyk, "University students' mental health and emotional wellbeing during the COVID-19 pandemic and ensuing lockdown," *South African Journal of Psychology*, vol. 51, no. 2, pp. 229–243, 2021.
[28] M. Zohaib, "Dynamic difficulty adjustment (DDA) in computer games: a review," *Advances in Human-Computer Interaction*, vol. 2018, Article ID 5681652, 12 pages, 2018.
[29] K. J. Van Slingerland, N. Durand-Bush, and S. Rathwell, "Levels and prevalence of mental health functioning in Canadian university student-athletes," *Articles*, vol. 48, no. 2, pp. 149–168, 2019.
[30] S. Grégoire, L. Lachance, T. Bouffard, and F. Dionne, "The use of acceptance and commitment therapy to promote mental health and school engagement in university students: a multisite randomized controlled trial," *Behavior Therapy*, vol. 49, no. 3, pp. 360–372, 2018.
[31] B. Cheng, J. Huang, Z. Guo, J. Li, and H. Chen, "Towards sustainable construction through better construction and demolition waste management practices: a SWOT analysis of Suzhou, China," *International Journal of Construction Management*, vol. 41, pp. 1–11, 2022.
[32] J. V. Cleofas, "Student involvement, mental health and quality of life of college students in a selected university in Manila, Philippines," *International Journal of Adolescence and Youth*, vol. 25, no. 1, pp. 435–447, 2020.
[33] T. H. Duong, H. Q. Nguyen, and G. S. Jo, "Smart data: where the big data meets the semantics," *Computational Intelligence and Neuroscience*, vol. 2017, Article ID 6925138, 2 pages, 2017.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9754217, 1 page
https://doi.org/10.1155/2023/9754217

Retraction

Retracted: Scientific Research Management Helping the Development of Regional Cultural Industry from the Perspective of Artificial Intelligence

Mobile Information Systems

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Wang and X. Hu, "Scientific Research Management Helping the Development of Regional Cultural Industry from the Perspective of Artificial Intelligence," *Mobile Information Systems*, vol. 2022, Article ID 1032081, 14 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 1032081, 14 pages
https://doi.org/10.1155/2022/1032081

Research Article

Scientific Research Management Helping the Development of Regional Cultural Industry from the Perspective of Artificial Intelligence

Jun Wang and Xuying Hu

Quzhou College of Technology, Quzhou 324000, China

Correspondence should be addressed to Xuying Hu; 435689@qzct.edu.cn

Received 9 August 2022; Revised 5 September 2022; Accepted 13 September 2022; Published 3 October 2022

Academic Editor: Jiafu Su

In order to solve the problem that traditional scientific research management is restricted by management methods and tools, and it is difficult to realize the development of cultural industry, this paper proposes a research method that scientific research management helps the development of the regional cultural industry from the perspective of artificial intelligence. Based on a wavelet neural network based on artificial intelligence, this paper establishes the goodwill value evaluation of cultural industry development under scientific research management. Firstly, the paper uses the hierarchical structure theory to construct the theoretical analysis framework and evaluation index system of cultural value and divides the influencing factors of cultural industry value into three levels: input layer network structure, hidden layer network structure, and output layer network structure. Then the wavelet neural network method is used to analyze, and the three-layer network structure is superimposed to reflect the hierarchical relationship between the index systems. Through the combination of normative analysis and empirical analysis, it verifies that the method has unique advantages and feasibility and realizes the promotion of scientific research management to the development of the regional cultural industry.

1. Introduction

The cultural industry has become a new growth pole in the process of global economic development and has gradually become a pillar industry in the economy of some developed countries. How to promote the rapid development of the cultural industry is a problem that all countries in the world are thinking about and discussing at present. However, from the perspective of the general trend of the development of the cultural industry, scientific and technological innovation and management are becoming an important driving force for the rapid development of the cultural industry. National planning outlines have made a comprehensive deployment for the development of cultural industry and proposed to promote the integration of culture and science, and technology. In this macro context, strengthening research on science and technology management and the development of the cultural industry will not only help to broaden the original theoretical scope but also further summarize the practical experience in the development of the cultural industry. From a global perspective, the development of cultural industries and the trade of cultural products are constantly impacting and changing the traditional economic form, which has a sustained and significant impact on the sustainable development of the world economy and the market pattern of world products and has become an important factor in improving the potential and quality of economic development and enhancing the innovation ability of countries or regions. Scientific research management and the development of the cultural industry are of great significance to accelerate the transformation of economic mode, promote the adjustment of industrial structure and industrial transformation and upgrading, improve the national cultural soft power, and build an innovative country [1].

2. Literature Review

Western cultural industry theory is relatively mature in foreign countries. Scholars have carried out extensive research and discussion on the cultural industry from the perspectives of economics, sociology, history, philosophy, and other different disciplines [2]. From the perspective of scientific and technological innovation, foreign scholars' research and exploration of cultural industry mainly focus on three aspects: the relationship between scientific and technological development and cultural communication; the relationship between scientific and technological innovation and cultural industry institutional innovation and the specific form of cultural industry.

Around whether cultural system innovation should be based on technological logic, Zhi and others believe that although technological innovation is a huge driving force for the development of cultural industry, it may also lead to conflicts with the national system, or contradictions between media companies, personal freedom of expression and national system constraints [3]. Czako and others studied the evolutionary relationship between technological innovation and institutional change in the cultural industry and found that the application and promotion of printing technology led to the emergence of the printing media industry and gave birth to the corresponding legal system. Subsequently, the emergence of radio technology and film promoted the emergence of management systems related to content control and access control. Comparatively speaking, the emergence of the latter has made the authority of freedom of speech and expression more restricted [4]. Saeik and others' research on the cultural content industry found that the development of modern network technology not only promotes the enrichment of the forms and contents of the cultural industry, but also provides modern management equipment such as electronic information systems, communication facilities, monitoring systems, etc., which provides favorable external conditions for the reform of management modes and the innovation of management methods, and thus promotes the institutional innovation of many cultural industries. In general, most scholars agree that scientific and technological innovation has a significant impact on the institutional innovation of the cultural industry, but they have not reached a unanimous conclusion on whether to follow the technical logic and choose the same policy objectives, methods, and principles under the unified management framework [5]. Li and others studied the impact of digital technology on the publishing industry. He believed that the most fundamental change brought by digital technology innovation is to separate the published content from the physical carrier carrying it [6].

In the research on the countermeasures for the integrated development of cultural industry and scientific and technological innovation, it analyzes the similarities and differences between scientific and technological innovation and cultural innovation at the macro level, as well as the multi-level impact of science and technology on culture, and puts forward the goals and tasks of the deep integration of science and technology into culture. Some scholars have studied the historical evolution, key issues, and talent requirements of the integration and development of culture and science and technology and have proposed to deal with the mechanism of integration drive, the way of an integration and transformation, the cultivation of integration atmosphere, and the improvement of integration quality. They studied the relationship between scientific and technological development and the cultural industry management system and put forward strategic countermeasures to promote the integrated development of cultural industry and scientific and technological innovation from the perspective of management system innovation. From the five aspects of system innovation, concept innovation, technology innovation, capital innovation, and content innovation, this paper puts forward the relationship that must be handled well to realize the innovative development of the cultural industry. At the regional level, the development situation and mode of the cultural industry in Henan, Hebei, Guangxi, Beijing, Tianjin, Hangzhou, Dalian, Chengdu, Wuhan, Anyang, and other provinces and cities have been deeply studied, and the corresponding specific strategic countermeasures to promote the integration of cultural industry and science and technology and enhance industrial competitiveness that are suitable for the characteristics of each province and city have been put forward [7].

3. Value Analysis of Cultural Industry Based on Artificial Neural Network

3.1. Constituent Factors. Adopting the method of wavelet neural network to establish an index system and a model of cultural industry goodwill value evaluation will better reflect the hierarchy. The evaluation indicators are divided into three layers: input layer, the hidden layer, and the output layer. It can more clearly reflect the relationship between the indicator systems, better measure and determine the impact indicators of the value of cultural industry goodwill, make up for the shortcomings of the previous methods of evaluating the value of goodwill, and better measure the value of cultural industry goodwill. The elements of cultural industry goodwill value evaluation are mainly divided into two factors: soft environment and hard environment: soft environment refers to the sum of external reasons and conditions such as law, policy, system, culture, ideology, and so on in addition to material conditions, including personnel quality, management level, and public influence of cultural industry. The concept of hard environment relative to a soft environment is the sum of geographical conditions, resource conditions, infrastructure, basic conditions, and other external factors and conditions in economic development [8].

3.2. Analysis of Constituent Elements. First, the LM algorithm optimization equation is quoted. Assuming the number of iterations is k, the output layer function of cultural industry goodwill value is defined as $F(z)$, which is expressed as $z_k = [w_{ik}, r_k, a_k, b_k]$ in the matrix.

$$Z_{k+1} = z_k - A_{k-1} g_k. \tag{1}$$

Suppose $F(z) = \nu(z)^T \nu(z, \nu(z))$ is the error vector. Gradient formula of output layer function of goodwill value of cultural industry:

$$\nabla F(z) = 2J^T(z)\nu(z). \tag{2}$$

Hessaian matrix approximate estimation formula

$$\begin{aligned} \nabla^2 F(z) &\approx 2J^T(z)J(z), \\ z_{k+1} &= zk - \left[J\left(z^J J(z)\right)\right]^{-1} J^T(z)\nu(z). \end{aligned} \tag{3}$$

Converted into unit matrix, the output layer function gradient equation of cultural industry goodwill value can be expressed as

$$z_{k+1} = zk - \left[J(z)^J J(z) + u_k I\right]^{-1} J(z)^J \nu(z). \tag{4}$$

We introduce the global optimization algorithm GA to determine the initial state of the neural network so as to achieve the minimum fitting difference. The processing flow of the GA genetic algorithm is shown in Figure 1.

The goodwill value evaluation of the cultural industry based on wavelet neural networks in this paper can not only reflect the characteristics of wavelet neural network that effectively extract the local information of signals, have self-learning, adaptability and fault tolerance, high precision and fast convergence speed, and avoids the blindness of BP neural networks and other structural designs but also evaluate the goodwill value of the cultural industry from the perspective of combining qualitative and quantitative analysis from a unique perspective [9]. The hierarchical structure of the wavelet neural network is more suitable for the hierarchical stratification and classification of the factors affecting the goodwill value of the cultural industry, so as to more effectively and accurately build an evaluation index system and determine the value of the cultural industry in terms of goodwill according to the evaluation index.

3.3. Structural Analysis. The reputation of the cultural industry is the reputation of an enterprise. Its generation is not a gift, but only when the commodity economy develops to a certain stage. It is also gradually formed in the process of commodity producers' production and operators' circulation. The value of goodwill reflects not only the comprehensive evaluation of the enterprise but also the comprehensive evaluation of the enterprise by society, and it is the external overall image of the company. In essence, goodwill is the capitalized value owned by the enterprise, but it cannot be specifically identified. It can enable the enterprise to obtain additional income. In addition, goodwill not only reflects the current development of an enterprise, including its management level, performance, personnel quality, etc., but also reflects the future development of an enterprise. A wavelet neural network is a combination of wavelet analysis and neural network theory. Neural processing units can represent different objects. The processing units in the network can be divided into three types: input units, output units, and hidden layer units. The input unit is connected with external information and data, and the output unit refers to the output of system processing results, but the hidden layer unit is between the input unit and the output unit, which is a unit that cannot be observed outside the system [10].

This paper uses the wavelet neural network method to transform the constituent indicators of cultural industry goodwill value into neurons as evaluation indicators, which are divided into three levels of hierarchy, input layer, hidden layer, and output layer. We regard the factors that affect the goodwill value of the cultural industry as the input, the average value of goodwill value roughly calculated by experts in the industry as the output, and the neuron incentive function selected by the hidden layer as the Morlet wavelet.

$$h(t) = \cos(1.75t)\exp\left(-\frac{t^2}{2}\right). \tag{5}$$

As the input layer function of cultural industry goodwill value evaluation is.

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}}\psi\left(\frac{x-b}{a}\right)\left(a, b \in R^2\right). \tag{6}$$

The output layer function is

$$y_k = \sum_{j=1}^{n} r_j h\left[\frac{\sum_{i=0}^{m} w_{ij} x_k(i) - b_j}{a_j}\right]. \tag{7}$$

By transforming the indicators of various influencing factors into neurons in the wavelet neural network method, the cultural industry neurons are formed, the hierarchical network structure is formed, and the index system and model are formed, so as to more clearly show the various relationships of the internal composition of the goodwill value of the cultural industry.

4. Empirical Test of Cultural Industry Efficiency of Artificial Intelligence

In order to empirically test the above theoretical analysis of the impact mechanism of artificial intelligence on the efficiency of the cultural industry, this chapter will first use the DEA-Malmquist index method to calculate the efficiency of the cultural industry from 2005 to 2018. Secondly, the principal component analysis method will be used to measure the development level of artificial intelligence. Finally, OLS regression and 2SLS regression will be used to test the impact of artificial intelligence on the efficiency of the cultural industry. At the same time, because the endogenous explanatory variables that may exist in the model will affect the results, this chapter will continue to use the limited information maximum likelihood method LIML, two-step optimal GMM, and iterative GMM to test the stability of the results [11, 12].

4.1. Calculation of Cultural Industry Efficiency

4.1.1. Model Selection. Data envelopment analysis (DEA) is to determine the relatively effective frontier through mathematical programming and statistical data and then

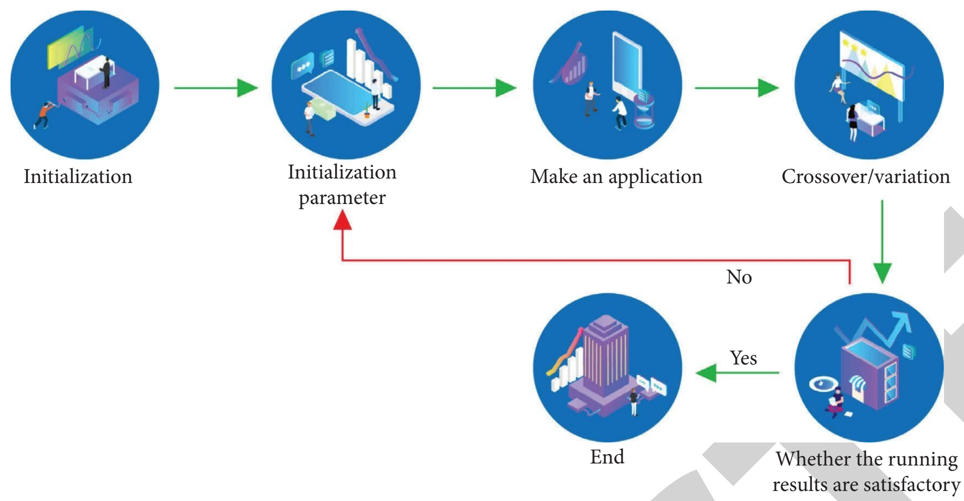


Figure 1: GA genetic algorithm processing flow chart.

analyze the deviation degree between the decision-making unit and the frontier of DEA, so as to calculate its relative efficiency. The DEA research method emphasizes the concept of relative efficiency, so as to evaluate the evaluation object, aiming to more effectively reflect the input-output efficiency of the analysis object, and select the best input-output scheme by adjusting the input-output of the analysis object. Because the DEA evaluation method takes relative efficiency as the evaluation index, it can analyze the efficiency of complex systems more simply and conveniently. From the perspective of output, the expressions of the previous $T-1$ period and the current t period of the Malmquist index are as follows:

$$\begin{aligned} M_0^{t-1} &= \frac{D_0^{t-1}\left(X^{t-1}\times Y^{t-1}\right)}{D_0^{t-1}\left(X^{t}\times Y^{t}\right)}, \\ M_0^{t-1} &= \frac{D_0^{t}\left(X^{t-1}\times Y^{t-1}\right)}{D_0^{t}\left(X^{t}\times Y^{t}\right)}. \end{aligned} \tag{8}$$

According to Fisher's ideal index, the comprehensive production index is the geometric average of the above formula:

$$\begin{aligned} M_0^{t-1,t} &= \sqrt{M_0^{t-1}\times M_0^{t}} \\ &= \sqrt{\frac{D_0^{t-1}\left(X^{t-1}\times Y^{t-1}\right)}{D_0^{t-1}\left(X^{t}\times Y^{t}\right)}\times\frac{D_0^{t}\left(X^{t-1}\times Y^{t-1}\right)}{D_0^{t}\left(X^{t}\times Y^{t}\right)}}, \end{aligned} \tag{9}$$

where X^{t-1} and Y^{t-1} represent the input and output vectors of the previous period, X^t and Y^t represent the input and output vectors of the current period, and D_0^{t-1} and D_0^t represent the technical level of the previous period and the current period.

Output oriented total factor productivity (TFP index) can be expressed as

$$\mathrm{TFP} = \sqrt{\frac{D_0^{t-1}\left(X^{t-1}\times Y^{t-1}\right)}{D_0^{t-1}\left(X^{t}\times Y^{t}\right)}\times\frac{D_0^{t}\left(X^{t-1}\times Y^{t-1}\right)}{D_0^{t}\left(X^{t}\times Y^{t}\right)}}. \tag{10}$$

TFP can be used to express the efficiency change level of the current period compared with the previous period. When $\mathrm{TFP}>1$, it means that the efficiency of the current period is on the rise; when $\mathrm{TFP}=1$, it means that the efficiency of the current period has not changed; when TFP $G<1$, it means that the efficiency of the current period is in a declining state [13].

Total factor productivity can be further decomposed into technical efficiency (EFFCH) and technological progress efficiency (CTECHCH), so the influence mechanism of total factor productivity of the cultural industry can be further analyzed. The expression is decomposed as follows:

$$\begin{aligned} \mathrm{TFP} = M_0^{t-1,t} &= \frac{D_0^{t-1}\left(X^{t-1}\times Y^{t-1}\right)}{D_0^{t-1}\left(X^{t}\times Y^{t}\right)} \\ &\cdot\sqrt{\frac{D_0^{t-1}\left(X^{t-1}\times Y^{t-1}\right)}{D_0^{t-1}\left(X^{t}\times Y^{t}\right)}\times\frac{D_0^{t}\left(X^{t-1}\times Y^{t-1}\right)}{D_0^{t}\left(X^{t}\times Y^{t}\right)}} \\ &= \mathrm{EFF}\times\mathrm{TECH}, \\ EFF &= \frac{D_0^{t-1}\left(X^{t-1}\times Y^{t-1}\right)}{D_0^{t-1}\left(X^{t}\times Y^{t}\right)} = \mathrm{PE}\times\mathrm{SE}. \end{aligned} \tag{11}$$

4.1.2. Construction of Index System. Considering the possibility and accuracy of data acquisition, combined with the above theoretical basis, this paper selects the cultural market's operating institutions as the representative variables of the cultural industry. The decision-making unit set of this paper selects the cost and revenue data of cultural market operating institutions in China's provinces, cities, municipalities directly under the central government and autonomous regions from 2004 to 2018. The data samples involve all provinces of the country. Due to the lack of data in Tibet, this paper will not consider Tibet for the time being. The data are mainly from the statistical yearbook of Chinese culture and tourism, the statistical yearbook of China, and the statistical database of China's economic network. The specific measurement indicators are shown in Table 1.

4.1.3. Calculation Results of Cultural Industry Efficiency

(1) Efficiency of Cultural Industry. The efficiency of the cultural industry in this paper is the total factor productivity of the cultural industry calculated according to the DEA-Malmquist index method. See Table 2 for the specific results of the efficiency of the cultural industry in various provinces and cities from 2004 to 2018.

As can be seen from Table 3, the average efficiency of Beijing's cultural industry from 2004 to 2018 was 1.122, indicating that the efficiency of its cultural industry increased by an average of 12.2%, ranking first. In addition, 20 provinces and cities have increased to varying degrees, and the remaining 10 provinces and cities are in a backward state. Specifically, Beijing and Chongqing are the only two regions where the efficiency of the cultural industry has increased by more than 10%; Fujian and Gansu increased by more than 5%; while Shaanxi and Jiangsu increased by only 0.2%; among the 10 provinces and cities that showed retrogression, the retrogression was not serious, with an average decrease of less than 0.5 in Hubei, Tianjin, Hainan, and Anhui; Qinghai has the largest degree of regression, with a decrease of about 5.3. Although the degree of regression is general, it has a large gap with Beijing, Chongqing, and other places ranking highly.

It can be seen from Table 4 that from 2004 to 2018, there were 6 years when the efficiency of the cultural industry was greater than 1, while there were 8 years when the efficiency was less than 1. Among them, the growth rate during 2006-2007 was the largest compared with other years, reaching 62.6%; during 2011-2012, it was followed by 10.1%; In the four years of 2004-2005, 2005-2006, 2013-2014, and 2016-2017, the growth rate of each year was 2% to 10%; In 2007-2008, 2008-2009, and 2010-2011, the recession was relatively serious, with more than 10% recession. The efficiency of the cultural industry in other years decreased by less than 10%.

(2) Technical Efficiency of Cultural Industry. It can be seen from Table 5 that the technological efficiency of the cultural industry has improved in 9 of the 14 observation periods from 2004 to 2018, and the technological efficiency of the remaining 5 periods has regressed. Among them, the technical efficiency during 2006-2007 has retreated the most. The technical efficiency index can be further divided into scale efficiency and pure technical efficiency, so the improvement of technical efficiency can be started from the aspects of enterprise scale, management and operation, technology, and other factors [14, 15]. The specific situation of the technical efficiency of the cultural industry in various provinces and cities in China from 2004 to 2018 is shown in Table 5.

As can be seen from Table 6, the average technical efficiency of Chongqing's cultural industry from 2004 to 2018 was 1.163, which means that the technical efficiency of its cultural industry increased by 16.3% on average, ranking first. In addition, only the average technical efficiency in Qinghai is lower than 1. Specifically, Chongqing is the only region where the technological efficiency of the cultural industry has increased by more than 10%; Guangdong, Fujian, Beijing, Ningxia, Sichuan, and Shanghai increased by more than 5%; the growth rate of other regions, except Qingdao overseas is between 1% and 5%. From the calculation results, it can be seen from the calculation results that there is a large gap between the technical efficiency of the cultural industry in Qinghai and other regions [16].

(3) Efficiency of Technological Progress in Cultural Industry. The average value of the technological progress efficiency of the cultural industry in each year shows that the technical efficiency of the cultural industry has exceeded 1 in 9 observation periods, while China's cultural technological progress efficiency has exceeded 1 in only 5 periods, such as 2005-2006, 2006-2007, 2011-2012, 2012-2013, and 2014-2015. This shows that the application of science and technology and the adoption of new technologies in the cultural industry still needs to be strengthened. The specific calculation results are shown in Table 7.

As can be seen from Table 8, the average technological progress efficiency of Beijing's cultural industry from 2004 to 2018 was 1.064, indicating that its technological progress efficiency of the cultural industry increased by an average of 6.5%, ranking first. In addition, technological progress in Fujian, Gansu, and Xinjiang increased by 2.8%, 1.9%, and 0.6% respectively. The efficiency of technological progress in other regions did not exceed 1.

4.2. The Impact of Artificial Intelligence on the Efficiency of Cultural Industry. This paper takes 30 provinces, municipalities, municipalities directly under the central government, and autonomous regions in China (excluding Tibet due to incomplete data acquisition) as the research sample, takes 2005–2018 as the research interval, and uses the Malmquist index to measure the efficiency of the cultural industry. On this basis, verify the impact of the AI level on the efficiency of the cultural industry. The measurement tool used is Statal 5.0.

4.2.1. Selection of Index Variables

(1) Core Explanatory Variables. The research object of this paper is the influence mechanism of artificial intelligence on the efficiency of the cultural industry. Therefore, this chapter selects the development level of artificial intelligence as the

Table 1: Efficiency measurement indicators of the cultural industry.

Input index	Number of employees in cultural market operation institutions Operating costs of cultural market operating institutions
Output index	Operating income of cultural market institutions

Table 2: Efficiency of cultural industry in various provinces and cities in China from 2004 to 2018.

Region	04-05	05-06	06-07	07-08	08-09	09-10	10-11	11-12	12-13	13-14	14-15	15-16	16-17	17-18
Beijing	0.963	1.254	1.716	0.341	0.816	0.845	0.812	1.442	1.576	1.161	0.972	0.394	1.001	0.730
Tianjin	0.263	1.261	1.612	0.664	1.032	0.861	1.033	0.932	0.947	0.919	1.414	2.251	0.997	1.013
Hebei	1.053	1.093	1.392	0.353	1.057	0.373	0.325	1.113	1.031	0.997	0.892	1.013	0.911	1.131
Shanxi	0.333	1.141	1.282	0.326	0.873	0.394	1.152	0.739	2.734	0.361	0.872	0.918	1.055	1.039
Inner Mongolia	1.061	1.044	1.861	0.311	1.142	0.982	0.332	0.995	0.952	0.906	0.973	0.967	0.941	0.891
Liaoning	0.639	1.723	3.084	1.351	0.212	0.892	0.959	1.102	1.019	0.931	0.879	0.726	1.139	1.191
Jilin	1.211	1.051	1.051	0.990	1.109	0.967	0.866	1.075	1.026	0.975	0.966	1.047	1.141	0.935
Heilongjiang	1.317	0.919	1.535	0.713	1.091	0.937	0.949	1.099	0.961	0.996	0.903	0.943	1.053	0.874
Shanghai	1.056	1.180	1.735	1.049	0.969	0.915	1.036	1.144	0.558	1.002	1.314	0.331	1.453	0.118
Jiangsu	0.999	1.026	1.391	0.901	0.996	0.911	0.336	1.131	0.919	1.013	1.027	0.922	0.951	0.970
Zhejiang	0.911	1.099	1.414	0.729	0.973	0.943	0.784	1.153	0.902	0.969	1.105	0.954	1.141	0.861
Anhui	1.043	0.979	5.784	0.253	0.946	1.165	0.637	1.057	0.348	1.013	1.136	0.923	0.391	0.996
Fujian	1.087	1.060	2.201	0.978	0.721	0.971	0.911	1.139	0.335	2.794	0.887	0.822	1.062	0.956
Jiangxi	1.102	1.011	1.436	0.354	1.098	0.937	0.937	0.868	1.201	0.997	0.917	0.841	1.071	0.901
Shandong	1.131	1.094	1.522	1.019	0.612	1.112	0.849	1.058	0.821	1.132	0.919	0.961	0.973	0.832
Henan	0.893	1.428	1.245	0.961	1.037	0.816	0.852	1.216	1.009	0.893	1.012	0.840	1.031	1.046
Hubei	1.229	0.946	1.413	0.935	0.943	0.965	0.818	1.127	0.938	1.187	0.811	0.737	1.001	0.962
Hunan	1.006	1.101	1.999	0.876	0.906	0.914	1.011	1.071	0.955	1.187	0.732	0.971	1.011	0.971
Guangdong	1.036	0.990	1.323	1.115	0.724	1.047	1.463	1.066	0.533	1.040	1.203	1.464	1.906	0.651
Guangxi	0.997	0.939	1.491	0.929	0.987	0.931	0.875	1.151	0.961	0.995	1.061	0.932	1.058	0.986
Hunan	1.084	0.992	1.907	0.725	0.876	1.036	0.921	1.025	1.597	0.614	0.381	1.146	1.038	0.810
Chongqing	2.321	1.462	1.741	0.893	0.825	0.957	0.901	1.049	0.972	1.041	1.706	1.186	0.437	1.031
Sichuan	1.582	1.131	1.899	0.941	0.751	0.836	1.012	0.991	0.784	1.025	0.944	0.119	0.763	1.432
Guizhou	1.013	1.129	1.245	0.977	0.904	1.011	0.682	1.431	0.915	1.064	0.968	0.871	1.115	0.663
Yunnan	1.017	0.778	2.421	0.641	1.121	0.919	0.737	1.251	1.026	1.036	0.917	1.001	1.025	1.022
Shaanxi	1.032	0.931	1.498	0.832	1.051	0.968	0.791	1.337	0.351	1.041	0.876	1.245	0.871	1.119
Gansu	1.101	1.131	1.214	1.003	1.142	0.393	0.356	1.186	0.899	0.901	0.645	0.532	1.586	1.053
Qinghai	1.021	1.087	1.458	0.582	0.879	0.871	0.791	1.069	0.918	0.829	0.884	1.033	1.110	0.919
Ningxia	1.191	1.093	1.116	0.910	1.089	1.010	0.874	1.223	1.002	1.491	0.664	0.858	1.009	0.991
Xinjiang	0.304	1.023	1.470	1.001	0.324	1.190	0.835	1.002	1.037	1.230	0.513	1.293	1.001	0.926

core explanatory variable of the empirical model. Based on the definition of the concept of artificial intelligence in Chapter 2, because artificial intelligence can affect the efficiency of cultural industry at the technical and institutional levels and it also has mutual influence at the technical and institutional levels, it cannot be completely analyzed independently. Therefore, a comprehensive, comprehensive, systematic, and representative index system is needed to reflect the development level of artificial intelligence. The indicators of the basic level of artificial intelligence mainly reflect the development level of the whole society's technical facilities, policy guidance, and other institutional levels; the indicators at the technical level mainly reflect the development level of talents and other technical levels; and the indicators at the application level comprehensively reflect the development level of artificial intelligence technology and system level.

Based on the initial formation of three primary indicators of the artificial intelligence foundation layer, technology layer and application layer, it is expanded into nine secondary indicators [17]. Including electricity consumption, total telecommunications business, the proportion of fixed asset investment in information, computer, and software industry in the total social fixed asset investment, the number of scientific and technological institutions of large and medium-sized industrial enterprises, the total number of R&D personnel at that time, the number of graduates of higher education institutions, the number of patent authorizations, the total output value of the high-tech industry, the turnover of technical contracts, etc. Finally, the principal component analysis method is used to measure the development level of artificial intelligence. The data in the above secondary indicators are mainly from the statistical database of the China economic network and the China Statistical Yearbook. The specific index system is shown in Table 9.

This paper uses principal component analysis to determine the weight of the above indicators, as shown in Table 10.

After comprehensive calculation, the comprehensive evaluation value of the development level of artificial intelligence is finally obtained. Figure 2 shows the overall trend of the average development level of AI in various regions of China from 2005 to 2018. It can be seen that the overall

Table 3: Average efficiency of cultural industries in various provinces and cities in China from 2004 to 2018.

Region	Numerical value	Ranking
Beijing	1.122	1
Chongqing	1.111	2
Fujian	1.083	3
Gansu	1.065	4
Xinjiang	1.047	5
Sichuan	1.043	6
Shanghai	1.041	7
Guangdong	1.038	8
Jilin	1.025	9
Hunan	1.023	10
Ningxia	1.018	11
Guangxi	1.017	12
Yunnan	1.011	13
Henan	1.011	14
Hebei	1.009	15
Jiangxi	1.004	16
Heilongjiang	1.003	17
Inner Mongolia	1.003	18
Shaanxi	1.002	19
Jiangsu	1.002	20
Hubei	0.996	21
Tianjin	0.996	22
Hainan	0.995	23
Anhui	0.994	24
Zhejiang	0.986	25
Guizhou	0.984	26
Liaoning	0.981	27
Shandong	0.974	28
Shanxi	0.966	29
Qinghai	0.947	30

Table 4: Efficiency of the cultural industry from 2004 to 2018.

	Technical efficiency (effch)	Efficiency of technological progress (techch)	Pure technical efficiency (pech)	Scale efficiency (sech)	Total factor productivity of cultural industry (tfpch)
2004-2005	1.634	0.633	1.423	1.141	1.031
2005-2006	0.891	1.228	0.884	1.008	1.093
2006-2007	0.502	3.223	0.575	0.884	1.625
2007-2008	1.271	0.663	1.276	1.004	0.847
2008-2009	1.793	0.496	1.636	1.100	0.895
2009-2010	0.975	0.964	0.961	1.015	0.944
2010-2011	1.077	0.828	1.051	1.025	0.886
2011-2012	1.050	1.049	1.046	1.005	1.101
2012-2013	0.530	1.859	0.793	0.666	0.986
2013-2014	1.478	0.730	1.131	1.307	1.076
2014-2015	0.626	1.565	0.627	0.999	0.984
2015-2016	1.289	0.751	1.236	1.043	0.968
2016-2017	1.407	0.734	1.372	1.025	1.037
2017-2018	1.020	0.932	1.030	0.990	0.952

development level of AI in China showed an upward trend from 2005 to 2018.

(2) Control Variables. For the control variables, four factors, including economic development level, human capital, cultural system factors, and market cultural demand, are taken as the control variables, as shown in Table 11.

First, the cultural industry is an industry that comes into being when the overall social and economic level of China has developed to a certain extent. Spiritual and cultural needs will appear only after meeting the general material needs of life. Therefore, the development of the cultural industry cannot be separated from the impact of the level of economic development. Based on this, this paper selects per

Table 5: Technical efficiency of cultural industry in various provinces and cities in China from 2004 to 2018.

Region	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Beijing	1.516	1.013	0.479	1.123	2.191	0.901	0.990	1.378	1.174	0.794	1.322	0.797	0.961	0.737
Tianjin	0.653	1.065	0.534	1.239	1.753	0.396	1.122	0.935	0.513	1.323	0.795	2.422	0.922	1.034
Hebei	1.543	0.371	0.412	1.412	2.220	0.913	1.009	1.097	0.524	1.357	0.661	1.399	1.326	1.131
Shanxi	1.533	0.917	0.385	1.053	2.147	0.912	1.444	0.766	1.324	0.614	0.631	1.243	1.397	1.057
Inner Mongolia	1.582	0.847	0.525	1.212	2.511	1.000	1.000	1.000	0.439	1.528	0.634	1.314	1.271	0.913
Liaoning	0.923	1.376	1.000	1.000	0.791	0.935	1.201	1.061	0.538	1.298	0.614	0.986	1.741	1.303
Jilin	1.331	0.837	0.289	1.711	2.125	0.982	1.122	0.994	0.531	1.336	0.676	1.423	1.511	1.000
Heilongjiang	1.871	0.727	0.471	1.123	1.031	0.969	1.202	1.065	0.511	1.241	0.704	1.282	1.326	0.956
Shanghai	1.035	1.000	0.847	1.113	1.000	1.000	1.000	1.000	0.435	1.531	0.731	1.051	1.933	0.931
Zhejiang	1.112	0.933	0.566	1.224	1.549	1.001	0.931	1.093	0.422	1.779	0.591	1.297	1.711	0.928
Anhui	1.475	0.831	1.637	0.492	1.721	1.13	0.817	0.999	0.411	1.635	0.763	1.255	1.271	1.062
Fujian	1.816	0.875	0.607	1.331	1.129	1.056	1.003	1.064	0.393	2.980	0.815	1.117	1.629	1.046
Jiangxi	1.772	0.822	0.412	1.526	1.056	0.971	1.156	0.864	0.623	1.374	0.671	1.142	1.586	0.961
Shandong	1.173	0.811	0.541	1.333	1.110	0.903	1.051	1.031	0.416	1.645	0.738	1.312	1.287	0.874
Henan	1.153	1.192	0.373	1.416	0.702	0.335	1.097	1.122	0.531	1.217	0.737	1.141	1.467	1.065
Hubei	1.843	0.755	0.720	1.479	1.356	0.971	1.063	1.053	0.476	1.963	0.504	1.137	1.462	1.043
Hunan	1.463	0.894	0.602	1.075	1.931	0.935	1.243	1.057	0.481	1.651	0.467	1.321	1.501	1.057
Guangdong	1.977	0.831	0.393	1.743	1.432	1.033	1491	1.000	0.438	1.241	0.633	1.313	1.603	0.836
Guangxi	1.530	0.787	0.433	1.474	1.973	0.941	1.099	1.105	0.503	1.411	0.726	1.335	1.469	1.032
Hainan	1.721	0.804	0.579	1.233	1.699	1.057	1.150	0.973	1.252	0.691	0.513	1.557	1.541	0.374
Chongqing	5.132	1.218	0.467	1.758	1.459	0.971	1.165	0.962	0.473	1.736	0.982	1.604	0.705	1.105
Sichuan	1.945	0.913	0.573	1.049	1.325	0.374	1.232	0.901	0.496	1.711	0.599	1.163	1.153	1.559
Guizhou	1.781	0.912	0.842	1.866	1.671	1.024	0.832	1.401	0.459	1.722	0.612	1.133	1.616	0.716
Yunnan	1.407	0.633	0.721	0.839	0.795	0.969	0.921	1.205	0.544	1.471	0.577	1.372	1.434	1.057
Shaanxi	1.699	0.764	0.396	1.534	1.391	1.006	0.993	1.216	0.436	1.519	0.576	1.359	1.235	1.193
Gansu	1.675	0.910	0.371	1.437	2.512	0.921	1.037	1.131	0.456	2.124	0.474	0.724	2.244	1.087
Qinghai	1.000	1.000	0.751	0.404	3.301	0.967	0.918	1.126	0.477	1.297	0.572	1.405	1.528	1.019
Ningxia	0.803	0.861	0.835	1.726	1.942	1.034	1.144	1.119	0.524	1.951	0.496	1.166	1.359	1.007
Xinjiang	1.314	0.821	0.866	1.525	1.792	1.237	1.146	0.913	0.523	0.408	0.878	1.757	1.504	1.000

Table 6: Average value of the technical efficiency of the cultural industry in various provinces and cities in China from 2004 to 2018.

Region	Numerical value	Ranking
Chongqing	1.162	1
Guangdong	1.060	2
Fujian	1.054	3
Beijing	1.053	4
Ningxia	1.052	5
Sichuan	1.051	6
Shanghai	1.050	7
Hainan	1.048	8
Zhejiang	1.048	9
Jilin	1.046	10
Gansu	1.044	11
Guangxi	1.043	12
Xinjiang	1.042	13
Jiangsu	1.038	14
Hunan	1.033	15
Yunnan	1.031	16
Inner Mongolia	1.031	17
Shaanxi	1.029	18
Henan	1.025	19
Hubei	1.025	20
Guizhou	1.024	21
Anhui	1.021	22
Heilongjiang	1.020	23
Shandong	1.018	24
Shanxi	1.013	25
Liaoning	1.011	26

Table 7: Technological progress of cultural industry in various provinces and cities in China from 2004 to 2018.

Region	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Beijing	0.638	1.238	3.584	0.749	0.372	0.936	0.824	1.046	1.342	1.462	2.248	0.894	1.042	0.928
Tianjin	0.402	1.184	3.021	0.536	0.589	0.963	0.966	0.998	1.829	0.748	1.779	0.929	1.081	0.935
Hebei	0.683	1.257	3.397	0.603	0.476	0.958	0.818	1.014	1.967	0.735	1.356	0.736	0.687	0.957
Shanxi	0.527	1.247	3.331	0.785	0.406	0.983	0.798	1.031	2.065	0.588	1.383	0.736	0.755	0.982
Inner Mongolia	0.671	1.233	3.546	0.672	0.455	0.988	0.831	0.995	1.954	0.593	1.534	0.736	0.741	0.975
Liaoning	0.689	1.256	3.084	1.352	0.268	0.954	0.798	1.038	1.893	0.716	1.432	0.736	0.657	0.914
Jilin	0.662	1.255	3.476	0.579	0.522	0.984	0.772	1.082	1.906	0.704	1.429	0.736	0.755	0.985
Heilongjiang	0.704	1.263	3.218	0.635	0.537	0.967	0.793	1.031	1.181	0.803	1.284	0.736	0.798	0.914
Shanghai	0.519	1.181	2.049	0.889	0.969	0.915	1.036	1.144	1.282	0.652	1.799	0.792	0.752	0.876
Jiangsu	0.569	1.191	3.012	0.574	0.631	0.939	0.859	1.015	2.099	0.568	1.788	0.736	0.658	0.907
Zhejiang	0.467	1.179	2.493	0.595	0.631	0.936	0.841	1.061	2.137	0.545	1.871	0.736	0.667	0.927
An Zheng	0.707	1.259	3.533	0.514	0.552	0.987	0.722	1.058	2.062	0.604	1.555	0.736	0.701	0.938
Fujian	0.598	1.211	3.626	0.535	0.638	0.921	0.893	1.071	2.097	0.937	2.139	0.736	0.652	0.923
Jiangxi	0.623	1.239	3.626	0.538	0.534	0.986	0.811	1.005	1.991	0.726	1.368	0.736	0.679	0.938
Shandong	0.637	1.242	2.805	0.586	0.551	0.974	0.808	1.027	1.974	0.688	1.321	0.736	0.755	0.952
Henan	0.775	1.198	3.337	0.678	0.471	0.977	0.777	1.065	1.906	0.728	1.361	0.736	1.524	0.982
Hubei	0.667	1.254	3.439	0.632	0.508	0.987	0.769	1.071	1.969	0.604	1.606	0.736	0.691	0.922
Hunan	0.685	1.232	3.318	0.815	0.469	0.978	0.813	1.013	2.011	0.719	1.567	0.736	0.673	0.917
Guangdong	0.556	1.193	3.381	0.641	0.505	0.967	0.977	1.066	1.149	0.648	1.921	0.808	1.189	0.733
Guangxi	0.652	1.256	3.406	0.631	0.491	0.987	0.796	1.012	1.454	1.644	1.241	1.234	1.257	1.045
Hainan	0.583	1.234	3.295	0.565	0.515	0.982	0.783	1.053	1.266	0.877	1.699	0.736	0.674	0.917
Chongqing	0.544	1.201	3.728	0.511	0.565	0.985	0.779	1.092	2.055	0.612	1.737	0.739	0.621	0.932
Sichuan	0.813	1.121	3.313	0.896	0.414	0.956	0.821	1.092	2.016	0.631	1.577	0.736	0.662	0.918
Guizhou	0.608	1.224	3.656	0.524	0.541	0.986	0.817	1.021	2.015	0.618	1.583	0.736	0.692	0.926
Yunnan	0.723	1.229	3.356	0.762	0.401	0.948	0.821	1.038	1.886	0.705	1.572	0.736	0.715	0.966
Shaanxi	0.637	1.217	3.789	0.525	0.565	0.962	0.796	1.099	1.955	0.655	1.534	0.736	0.704	0.934
Gansu	0.657	1.257	3.302	0.702	0.455	0.975	0.787	1.047	1.972	1.366	1.362	0.736	0.707	0.974
Qinghai	1.021	1.087	1.944	1.442	0.266	0.907	0.862	0.949	1.925	0.638	1.545	0.736	0.727	0.971
Ningxia	0.594	1.269	3.335	0.522	0.561	0.977	0.764	1.093	1.911	0.764	1.338	0.736	0.742	0.985
Xinjiang	0.612	1.246	3.411	0.659	0.46	0.962	0.772	1.077	1.982	1.424	1.354	0.736	0.666	0.926

Table 8: Average value of the technological progress of the cultural industry in various provinces and cities in China from 2004 to 2018.

Region	Numerical value	Ranking
Beijing	1.064	1
Fujian	1.027	2
Gansu	1.018	3
Xinjiang	1.005	4
Tianjin	0.996	5
Shanghai	0.993	6
Sichuan	0.992	7
Hunan	0.990	8
Henan	0.984	9
Heilongjiang	0.983	10
Yunnan	0.981	11
Jilin	0.980	12
Guangdong	0.976	13
Guangxi	0.975	14
Shaanxi	0.973	15
Inner Mongolia	0.972	16
Anhui	0.972	17
Hubei	0.971	18
Liaoning	0.970	19
Ningxia	0.968	20
Hebei	0.967	21
Jiangsu	0.964	22
Qinghai	0.964	23
Jiangxi	0.963	24
Guizhou	0.962	25
Chongqing	0.957	26
Shandong	0.957	27
Shanxi	0.953	28
Hainan	0.950	29
Zhejiang	0.940	30

Table 9: Selection of AI development level indicators.

Target	Primary index	Secondary indicators
Development level of artificial intelligence	Foundation layer	Power consumption
		Total telecommunication business
		Proportion of fixed asset investment in information, computer and software industry in the whole society
	Technical level	Number of scientific and technological institutions of large and medium-sized industrial enterprises
		R&d personnel at that time
		Number of graduates from higher education institutions
	Application layer	Number of patents authorized
		Total output value of high-tech industry
		Turnover of technology contract

Table 10: Index weight of artificial intelligence development level.

Index	Weight
Power consumption	0.04
Total telecommunication business	0.11
Proportion of fixed asset investment in information, computer and software industry in the whole society	0.08
Number of scientific and technological institutions of large and medium-sized industrial enterprises	0.14
R&d personnel at that time	0.15
Number of graduates from higher education institutions	0.01
Number of patents authorized	0.16
Total output value of high-tech industry	0.15
Turnover of technology contract	0.11

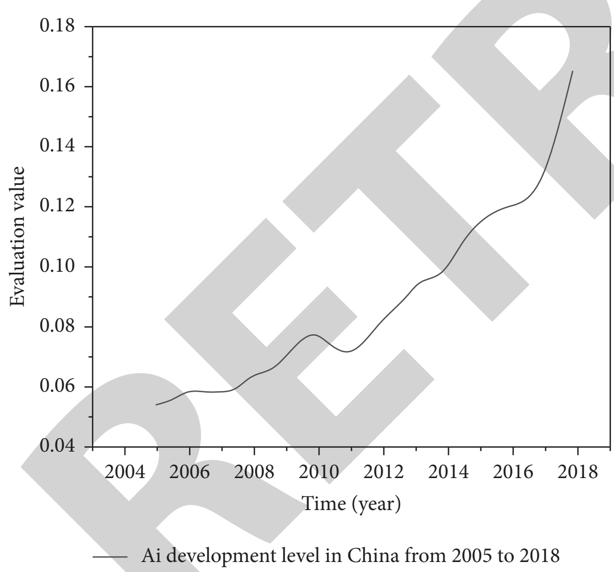


Figure 2: Development level of artificial intelligence in China from 2005 to 2018.

capita GDP as an indicator to measure the level of economic development. Second, human capital, as the human resource input of the efficiency of the cultural industry, is inseparable from the efficiency of the cultural industry. Theoretically, the higher the quality of human capital, the more obvious the role of human capital in improving the efficiency of the cultural industry. Therefore, this paper selects the proportion of high school and above in the total employment of society as the index to evaluate the quality of human resources. Third, national support can promote the development of the cultural industry, so this paper selects cultural broadcasting fees as the measurement index of national financial investment. Fourth, the market's cultural demand is the internal condition for the emergence and development of the cultural market. Under the influence of the market supply and demand mechanism, it can promote the innovation and development of enterprises in the cultural industry. Therefore, this paper selects the per capita consumption expenditure on education, culture, and entertainment as the index to measure the demand for cultural consumption [18].

4.2.2. Model Setting. In order to further verify the above analysis, we need to analyze the impact of artificial intelligence on the efficiency of the cultural industry from an empirical perspective and test the impact of artificial intelligence on the production efficiency of the cultural industry through the panel data model. The reasons for choosing panel data analysis include the following three aspects. First, the use of panel data models can increase the degree of freedom of samples and the amount of information, so as to reduce the gap between insufficient data supply and information demand in the efficiency model of the cultural industry. Second, this model can reflect the inter provincial gap in cultural industry and the regional differences in the impact of artificial intelligence on cultural industry in eastern, central, and western China. Thirdly, after Hausmann test and LM test of random utility model and fixed utility model, this paper believes that the OLS regression model is the most efficient;. At the same time, the

Table 11: Selection of control variables.

Other influencing factors	Specific measurement indicators	Control variable representation
Economic development level	Per capita GDP	Per capita GDP
Human capital	Proportion of high school and above in the number of employees	Human resource quality (human)
Cultural broadcasting expenses	Special financial expenditure for culture (ce)	Cultural system factors
Per capita consumption expenditure on education, culture and entertainment	Cultural consumption (yt)	Market cultural needs

Hausmann test results of ordinary least squares (OLS) and two-stage least squares estimation (2SLS) show that there may be endogenous explanatory variables in the model, and the instrumental variable method needs to be used. The specific construction model is as follows:

$$tfp_{it} = c + aai_{it} + \beta C + \mu_{it}, \tag{12}$$

where, i and t represent the region and year respectively; TFP is the explained variable, which indicates the efficiency of the cultural industry; a_i is the core explanatory variable, representing the development level of artificial intelligence; C are other factors that affect the total factor productivity of the cultural industry, and the per capita GDP (GDP), human resource quality (human), cultural consumption (YT), and special cultural fiscal expenditure (CE) are selected as control variables; α, β are parameters to be estimated; and μ is random disturbance term. In the model, a represents the influence coefficient of the development level of artificial intelligence on the total factor productivity of the cultural industry, which is the core result of this paper.

4.2.3. Empirical Results and Robustness Test. In order to explore the impact of artificial intelligence on the efficiency of the cultural industry, this paper will carry out OLS regression, 2SLS regression, limited information maximum likelihood (LIML), two-step optimal GMM, and iterative GMM, respectively. Firstly, OLS is used as the reference term for the model. As the Hausmann test results show that there may be endogenous problems in the model, the 2SLS model is used to add tool variables to solve the endogenous problems. However, there may be weak instrumental variables, and LIML is more insensitive to weak instrumental variables. For the sake of robustness, LIML is again used for regression analysis of the model [19]. At the same time, because the impression of some other unobservable factors will inevitably lead to heteroscedasticity and other problems, this paper uses the optimal GMM and iterative GMM models to control and reduce the heteroscedasticity and weak instrumental variables in the model.

(1) Regression Results. The empirical results of the impact of AI on the efficiency of the cultural industry are shown in Table 12. Column (1) shows the OLS regression results with control variables. The AI coefficient of the explanatory variable is significantly positive, indicating that the AI development level has a positive impact on the efficiency of the cultural industry. Column (2) is the 2SLS regression result. After solving some possible endogenous problems, the AI coefficient of the explanatory variable is still significantly positive. Column (3) is the regression result estimated by LIML, and the AI coefficient of the explanatory variable is significantly positive, indicating that this result has certain robustness. Column (4) is the regression result using two-step GMM estimation, and the AI coefficient of the explanatory variable is significantly positive. Column (5) is the regression result estimated by iterative GMM, and the AI coefficient of the explanatory variable is also significantly positive.

The empirical results of the impact of AI on the technical efficiency of the cultural industry are shown in Table 13. Column (1) shows the OLS regression results with control variables. The AI coefficient of the explanatory variable is significantly positive, but it does not pass the significance test. Column (2) is the 2SLS regression result. After solving some possible endogenous problems, the AI coefficient of the explanatory variable is positive and fails to pass the significance test. Column (3) is the regression result estimated by LIML. The AI coefficient of the explanatory variable is positive, and it still fails to pass the significance test. Column (4) is the regression result estimated by two-step GMM, and the AI coefficient of the explanatory variable is positive. Column (5) is the regression result estimated by iterative GMM, and the AI coefficient of the explanatory variable is also positive.

The empirical results of the impact of AI on the efficiency of technological progress in the cultural industry are shown in Table 14. Column (1) shows the OLS regression results with the control variable added. The AI coefficient of the explanatory variable is positive, indicating that the AI development level has a positive impact on the technological progress of the cultural industry. Column (2) is the 2SLS regression result. After solving some possible endogenous problems, the AI coefficient of the explanatory variable is significantly positive. Column (3) is the regression result estimated by LIML. The explanatory variable AI coefficient is significantly positive, indicating that this result is robust. Column (4) is the regression result using two-step GMM estimation, and the AI coefficient of the explanatory variable is significantly positive. Column (5) is the regression result estimated by iterative GMM, and the AI coefficient of the explanatory variable is also significantly positive [20].

(2) Result Analysis. From the above regression results, we can see that the coefficient of the impact of AI development level on the efficiency of cultural industry and the efficiency of technological progress is significantly positive. This means

Table 12: Impact of artificial intelligence on the efficiency of the cultural industry.

Variable	(1) OLS1	(2) 2SLS1	(3) LIML1	(4) GMM1	(5) IGMM1
Ai	0.7588**	0.7558*	0.7561*	0.8182**	0.8201**
	(1.9673)	(1.8724)	(1.8688)	(2.0339)	(2.0385)
Gdp	$-6.03E-06$**	$-5.68E-06$**	$-5.69E-06$**	$-5.62E-06$**	$-5.62E-06$**
	(−2.0373)	(−1.8665)	(−1.8655)	(−1.8412)	(−2.0750)
human	−0.0022	−0.0004	−0.0004	0.0004	0.0004
	(−0.5552)	(−0.1033)	(−0.1035)	(0.0855)	(0.0886)
yt	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.6421)	(0.2961)	(0.2960)	(0.0478)	(0.0434)
ce	0.0000	−0.0001	−0.0001	−0.0002	−0.0002
	(0.0441)	(−0.1173)	(−0.1172)	(−0.2380)	(−0.2397)
N	420	390	390	390	390
F inspection, *P* value	0.0558	—	—	—	—
Chi square test, *P* value	—	0.0961	0.0965	0.0677	0.0672

Table 13: Impact of artificial intelligence on the technical efficiency of the cultural industry.

Variable	(1) OLS2	(2) 2SLS2	(3) LIML2	(4) GMM2	(5) IGMM2
ai	0.9439*	0.2642	0.2711	0.1079	0.0971
	(1.8068)	(0.5156)	(0.5264)	(0.2123)	(0.1912)
gdp	$-4.45E-06$**	−0.00000157	−0.0000016	−0.0000035	−0.00000173
	(−1.3988)	(−0.4999)	(−0.5077)	(−0.1121)	(−0.0554)
human	0.0025	0.0065*	0.0065*	0.0074*	0.0076**
	(0.5814)	(1.7075)	(1.7056)	(1.9521)	(2.0030)
yt	−0.0002***	−0.0002***	−0.0002***	−0.0002***	−0.0002***
	(−4.6401)	(−5.2234)	(−5.2241)	(−5.2508)	(−5.2673)
ce	0.0005	0.0013	0.0013	0.0013	0.0013
	(0.5136)	(1.4521)	(1.4488)	(1.4545)	(1.4202)
N	420	390	390	390	390
F inspection, *P* value	0.0001	—	—	—	—
Chi square test, *P* value	—	0.0001	0.0001	0.0001	0.0001

Table 14: Impact of artificial intelligence on technological progress efficiency of cultural industry.

Variable	(1) OLS3	(2) 2SLS3	(3) LML3	(4) GMM3	(5) IGMM3
ai	0.9224	1.7330**	1.7330**	1.7571**	1.7571*
	(1.4341)	(2.2278)	(2.2277)	(2.2695)	(2.2694)
gdp	−0.00000459	$-7.76E-06$*	$-7.76E-06$*	$-7.92E-06$*	$-7.92E-06$*
	(−1.2075)	(−1.7854)	(−1.7854)	(−1.8349)	(−1.8350)
human	−0.0114**	−0.0134**	−0.0134**	−0.0135**	−0.0135**
	(−1.9871)	(−2.1809)	(−2.1809)	(−2.1956)	(−2.1956)
yt	0.0002***	0.0002***	0.0002***	0.0002***	0.0002***
	(5.2204)	(4.9067)	(4.9067)	(4.9035)	(4.9034)
ce	−0.0026**	−0.0036***	−0.0036***	−0.0036***	−0.0036***
	(−2.0554)	(−2.7237)	(−2.7237)	(−2.7141)	(−2.7140)
N	420	390	390	390	390
F inspection, *P* value	0.0001	—	—	—	—
Chi square test, *P* value	—	0.0001	0.0001	0.0001	0.0001

that the positive impact of AI on the cultural industry is mainly through promoting the technological progress of the cultural industry, so as to promote the efficiency of the cultural industry.

First, in terms of the efficiency of the cultural industry, the impact of artificial intelligence on the efficiency of the cultural industry is estimated to be 0.8200 under the iterative GMM model. Because the cultural industry is a technology intensive industry, the production of its products depends on technological creativity. The positive impact of artificial intelligence on the efficiency of the cultural industry shows that artificial intelligence can effectively promote the innovative development of the cultural industry. This result is consistent with the theoretical analysis above. Artificial intelligence can indeed break the current situation of path dependence in the cultural industry and promote the path creation of the cultural industry.

Second, from the perspective of technical efficiency, the regression result is not significant. From the perspective of objective conditions, the reason for this result may be that the government's public data has time constraints, so the length of data acquisition is limited, and it does not fully show the evolution process of artificial intelligence and cultural industry efficiency. Combined with the previous theoretical analysis, AI failed to promote the path creation of technology in the cultural industry. On the one hand, government policies have a great impact on the technical efficiency of the cultural industry. Although at the national level, China has issued many relevant policies on the integration of science, technology, and culture, there is a certain time lag in the cultural industry policies on digital technologies such as artificial intelligence at the local government level.

Third, in terms of the efficiency of technological progress, the impact of AI on the efficiency of technological progress of the cultural industry is estimated to be 1.7571 under the iterative GMM model, which is much larger than the impact coefficient of AI on the efficiency i the cultural industry. This shows that at this stage, the impact of artificial intelligence on the cultural industry is mainly through providing technical support to the cultural industry and providing new technology application in its production and operation process, so as to improve the efficiency of technological progress of the cultural industry and then promote the efficiency of the cultural industry. Combined with the above theoretical analysis, specifically speaking, at this stage, artificial intelligence can promote the cultural industry to realize the creation of a technological path, but the impact of artificial intelligence on the cultural industry is still relatively simple, mainly through technological breakthroughs and innovation, bring emerging technologies into the cultural industry, and realize the breakthrough technological path creation of the cultural industry, so as to promote the technological progress efficiency of the cultural industry. The quality of talents has a significant negative impact on the efficiency of technological progress in the cultural industry, that is, when the quality of talents continues to improve, it fails to promote the efficiency of technological progress in the cultural industry, which is consistent with the above analysis results of professional and technical talents and compound talents in the cultural market. At this stage, the impact of artificial intelligence on the efficiency of the cultural industry is restricted by the lack of compound talents. The situation of the gradual path creation formed by the integration of artificial intelligence and cultural industry technology has not been fully apparent.

Based on the analysis of the above results, the path creation process of artificial intelligence to promote the cultural industry first needs to experience the process of technology path creation such as the generation of new technologies and technology spillovers. With the continuous development of technological breakthroughs, product innovation, the formation of industrial formats, and other aspects, institutional path creation will occur synchronously. At this stage, the cultural industry has initially completed the process of technological path creation, but institutional path creation has not been fully realized.

5. Conclusion

When constructing the theoretical framework of cultural industry goodwill value evaluation, this paper divides the influencing factors of goodwill value evaluation into three levels: input layer network structure, hidden layer network structure, and output layer network structure according to the hierarchy theory and the characteristics of cultural industry goodwill value. Then the wavelet neural network method is used to analyze, and the superposition of the three-layer network structure is realized. The hierarchical network structure can more objectively reflect the hierarchical relationship between the index systems, which can be used as the theoretical basis for building the evaluation index system and model.

This paper uses the value chain theory to construct the cultural industry's goodwill value evaluation index system. When constructing the cultural industry goodwill value evaluation index system, according to the value chain theory, combined with the characteristics of the cultural industry, through the analysis of the constituent elements that affect the cultural industry goodwill value evaluation, the primary evaluation index system is determined; then, combining the questionnaire method and the wavelet neural network method, the preliminarily determined evaluation indicators are screened layer by layer, some special indicators are eliminated, and the index system of cultural industry goodwill value evaluation is constructed, which solves the disadvantage of subjectivity in the selection of indicators to a certain extent.

By constructing the efficiency index system of cultural industry, this paper uses the DEA-Malmquist index method to measure the efficiency of the cultural industry, the efficiency of technological progress of the cultural industry, and the technological efficiency of the cultural industry. By constructing the index system of the development level of artificial intelligence, the development level of artificial intelligence in China is measured by principal component analysis. At the same time, OLS regression and 2SLS regression are used to test the impact of artificial intelligence on the efficiency of the cultural industry, and LIML, two-step optimal GMM, and iterative GMM are used to test the stability of the results. The results show that first, AI can promote the efficiency of the cultural industry and promote the innovative development of the cultural industry. Second, limited by imperfect national policies, the shortage of compound talents and the incomplete market competition pattern, the impact of AI on the technical efficiency of the cultural industry is not significant. Third, at this stage, the role of artificial intelligence in promoting the efficiency of the cultural industry is mainly to provide technical support for the cultural industry and provide emerging technical support in its production and operation, so as to promote the technological progress of the cultural industry and

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9838372, 1 page
https://doi.org/10.1155/2023/9838372

Retraction

Retracted: Prediction of Long-term Variation of Ocean Circulation Based on Grey Correlation Clustering

Mobile Information Systems

Received 26 September 2023; Accepted 26 September 2023; Published 27 September 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] N. Ma and G. Yang, "Prediction of Long-term Variation of Ocean Circulation Based on Grey Correlation Clustering," *Mobile Information Systems*, vol. 2022, Article ID 4791054, 8 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 4791054, 8 pages
https://doi.org/10.1155/2022/4791054

Research Article

Prediction of Long-term Variation of Ocean Circulation Based on Grey Correlation Clustering

Na Ma [1] and Guoshu Yang [2]

[1] *Gaomi College, Qingdao University of Science and Technology, Qingdao, Shandong 261500, China*
[2] *School of Mathematics and Physics, Qingdao University of Science and Technology, Qingdao, Shandong 261500, China*

Correspondence should be addressed to Na Ma; 201811111121718@stu.hubu.edu.cn

Received 24 June 2022; Revised 11 September 2022; Accepted 17 September 2022; Published 30 September 2022

Academic Editor: Xuefeng Zhang

In order to study the change of circulation in the South China Sea, the author puts forward the research on the distribution and change of water masses in the South China Sea based on grey correlation clustering. Based on the wod21 temperature and salt observation data of the South China Sea from 1966 to 2013, this paper uses the grey correlation clustering algorithm to divide the water masses in the whole sea area and $5^{\circ} \times 5^{\circ}$ small areas and analyzes the distribution, temperature, salt properties, and seasonal changes of each water mass in the South China Sea combined with the systematic clustering method and T-S point clustering map. The experimental results show that in the vertical division, the South China Sea water mass is divided into five layers: surface water, subsurface water, sub-middle water, middle water, and deep water. The deep water in the South China Sea is mainly distributed in the basin with a depth of 900 m. The temperature value is lower than 5.5°C and the salinity range is 34.30–34.70. The properties and changes of water masses obtained in this paper are consistent with the existing conclusions, which shows that the grey correlation clustering algorithm is efficient and accurate in the division of water masses in the South China Sea.

1. Introduction

An accurate assessment of global ocean circulation change is a key in understanding the change of Earth's climate system and predict its future change trend [1]. For a long time, there is a basic scientific question that people urgently want to answer in terms of ocean circulation and climate "how does Earth's ocean circulation system change under the background of global warming?" However, there are great difficulties in answering this question. On the one hand, the changes in Earth's ocean circulation are very regional and complex. Under the forcing of greenhouse gas emissions, there are significant differences in the response of ocean circulation in different regions to climate change, as shown in Figure 1. For example, the subtropical western boundary current, the Pacific wind-driven circulation, and the Indonesian through current have a strengthening trend, which leads to rapid warming in relevant regions. The Angoras current has not accelerated since the 1990s, while the Atlantic meridional overturning circulation shows a slowing trend [2]. These regional differences are mainly caused by the dynamic process adjustment within the climate system. The dynamic adjustment of ocean circulation in different regions is very different, resulting in great differences in the estimation of its multidecadal trend. On the other hand, human beings still lack systematic and continuous direct observation of Earth's ocean circulation [3]. Historically, most of the observations of global ocean circulation have focused on some specific domains and time periods. The regional differences in ocean circulation make it difficult to monitor the overall changes of Earth's general circulation system with these specific observations.

2. Literature Review

Dias and others first used the equation of seawater motion to study the problem of ocean currents, trying to explain the cause of ocean currents with wind stress. However, they did

Figure 1: Variation of ocean circulation.

not understand the disaster state motion of seawater at that time and did not understand the significance of disaster in the momentum exchange of seawater, so they dealt with this problem with the molecular viscosity of seawater. Therefore, an unreasonable result is obtained: the wind stress needs to act continuously for thousands of years to form a stable ocean current. At the same time, the impact of Earth's rotation on ocean currents has not been understood in the study [4]. Scheen and others first recognized the important role of geostrophic deflection force (Coriolis force) in wind-driven ocean currents [5]; Liu and others derived a formula for calculating the current velocity using the slope of the isobaric surface. However, these two research results were not valued by oceanographers at that time [6]. Raj and others put forward the famous "wind ocean current theory". On the premise of considering the effect of Coriolis force on an ocean current, the velocity field of wind ocean current in the friction layer on the sea surface is successfully calculated for the first time according to the sea surface wind stress. The establishment of "circulation theory" and "wind ocean current theory" has made the research of ocean circulation have a solid mathematical theoretical foundation and began to use the motion equation to study the average motion of seawater [7]. Hao and others proposed the important role of inertial effect in the formation of current bending, but he applied the simplified disturbance equation, so it is still linear [8]. Sobisevich and others have played a direct role in promoting the study of nonlinear current theory [9]. On this basis, Zhao and others calculated the flow and velocity of the cross flow and the width of the Gulf Stream according to the principle of conservation of potential hazard degree, which is in good agreement with the measured results [10]. These nonlinear theories do not include any accidental parameters, which is also an advantage over the full flow theory. However, due to the complexity of mathematical processing of nonlinear problems, there are still some difficulties in the research of this theory in the 1990s when computer technology has not yet sprung up. With the rapid development of computer technology, the research of nonlinear theory has been greatly developed. Scholars can use the computer for numerical solutions so as to overcome the problem that it is difficult to obtain the analytical solution to nonlinear problems. Due to the significant regional differences in ocean circulation in different sea areas, the response of a specific sea area or a specific ocean circulation to climate warming will be significantly different. Therefore, Wang and others effectively overcome the regional differences by integrating the large-scale marine kinetic energy in the whole global sea depth and using the integrated marine kinetic energy as an index to judge the changes in marine circulation [11]. Based on the monthly average wod13 temperature and salt data of climate state, this paper divides the water masses in the South China Sea by using the grey correlation cluster analysis method and analyzes the temperature and salt distribution characteristics and seasonal variation law of each water mass by combining the systematic cluster tree method and temperature and Salt Point aggregation map.

3. Research Methods

3.1. Data Preparation. The data used for analysis in this paper is the world ocean database2021 (wod21). The temperature and salt data in this database mainly include observation data such as high-resolution CTD, water temperature detector (MBT, XBT), drifting buoy (DRB), profile buoy (PFL), and anchored buoy (MRB) [12]. The time range of n in this paper is from January 2021 to March 2021 in e-105° space. The stations in each season can be distributed throughout the South China Sea.

When processing the data, only the sections with more than three temperature and salt data are retained. Because this paper only uses the temperature and salt data to cluster the water masses in the South China Sea, only the elements of longitude, latitude, time, depth, temperature, and salinity are retained in the data. After deleting the unqualified sections through quality control, there are 29317 sections in total. The seasonal distribution of the number of sections is shown in Figure 2, including 7105 in winter, 7378 in spring, 7445 in summer, and 7389 in autumn. The seasonal distribution is relatively uniform [13].

Since wod21 data is measured data with wide sources and a large time span, and the quality of data is uneven, in order to ensure that the data used for analysis can get more reasonable results, some processing needs to be carried out on the original data before specific research, and finally, 12-month monthly average data of climate state need to be obtained. The data processing process is as follows:

(1) Delete the data whose salinity value is all NaN. Wod21 summarizes all the measured data. For some instruments (such as thermometers), the data obtained lack temperature or salinity data, and the salinity value is a very important element in water mass analysis, so these data are deleted.

(2) Extract the required elements and unify the format. The longitude and latitude, depth, temperature, and salinity elements in the original data are extracted.

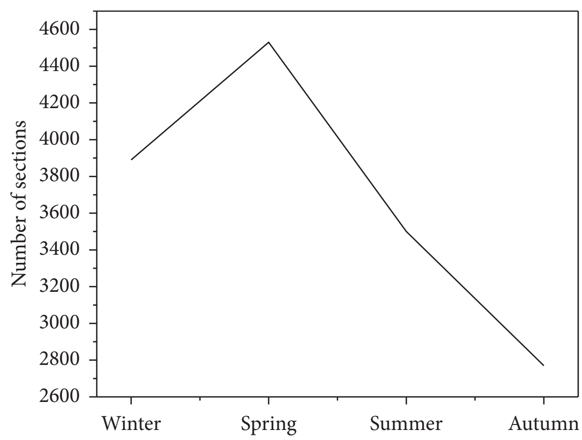


Figure 2: Distribution of section number.

The original format of time data is a string, which needs to be converted into an array of adult month day format. For data without time value, the time of adjacent data is taken [14].

(3) Extract the data of the South China Sea and delete the obviously wrong values. The temperature range is 0–40. For salinity, the coastal water is greatly affected by river water and the salinity is low, forming a large difference from the salinity values of other sea areas in the South China Sea. Therefore, the salinity value less than 31 is assigned as 31.

(4) Calculate the monthly average data of the climate state. Because the data are stored in years, and the site distribution is different every year. Therefore, when merging the data of the same month in different years, directly add the data with different site locations, and further compare the water depth with the data with the same site location. If the water depth is different, directly add the data, and if the water depth is the same, average the temperature and salt value.

(5) Delete sites with less than three vertical values. Site data with less than three values in the vertical direction.

(6) When used in clustering analysis, the data will be biased to the surface, resulting in the error of clustering results.

(7) Moving to average the profile. When drawing salinity depth and temperature-depth curves on the profile, it is found that the measured data are often prone to "burrs". Therefore, the sliding average of temperature and salinity is carried out, and the temperature and salt value corresponding to the original depth are taken.

(8) Interpolate data to grid points. The data processed according to the above steps can be used for cluster analysis. However, the uneven distribution of sites may lead to clustering results biased towards areas with more data distribution. Therefore, the data are interpolated, with the horizontal interval of 0.1° and the vertical interval of 5 m.

The data processed by quality control basically conforms to the water characteristics of the South China Sea, but there will still be a small number of unqualified points, which will not affect the cluster analysis of water masses. However, in the clustering process, the step of eliminating unqualified points is added, hoping to get more reasonable results [15]. In addition, the previous water mass classification results are mostly based on temperature and salt elements. Therefore, in this paper, only two elements of temperature and salinity are selected in cluster analysis, which is convenient to compare and analyze the water mass classification results with the previous conclusions.

3.2. Grey System Theory. In the field of information, people usually use the depth of color to express the clarity of information. "Black" means the information is unknown, "white" means the information is known, "gray" means some information is unknown and some information is known. For this part of the information system represented by "gray", we call it the "gray system". Grey system theory mainly studies those poor information uncertain systems with "some information being known and some information unknown". Through the analysis of the known part of the information, the possibility of the unknown part of the information is predicted, which is similar to semi-supervised learning. Grey correlation analysis and grey clustering are important branches of grey system theory. The grey correlation analysis method has no specific requirements for the sample size and regularity, and the amount of calculation is small, so there will be no inconsistency between the quantitative results and the qualitative analysis results. The grey relational clustering method simplifies the complexity of the system by merging similar factors [16].

In the traditional recommendation algorithm, the similarity of users or items is usually calculated according to the historical behavior data generated by users, then the similarity is sorted, and finally, recommended to users according to the sorting results. However, the historical behavior data of users are usually incomplete, heterogeneous, and loose. Such data will have a negative impact on the recommendation algorithm. The grey system theory has a good processing effect on these kind of data. Its advantages in processing data are very suitable for solving the problem of data sparsity in the recommendation system and can well alleviate the cold start problem.

3.2.1. Grey Relational Analysis Theory. Grey correlation mainly studies the uncertain correlation between various factors. Before grey correlation analysis, first find the mapping quantity that can reflect the behavior characteristics of the system, determine the effective factors affecting the system, and deal with them appropriately, then calculate the correlation coefficient and correlation degree between various factors, and finally analyze according to the calculation results [17]. In essence, this method judges the degree of correlation according to the similarity of the sequence curve of factors. The relevant definitions of grey correlation analysis theory are given below.

Define the mapping quantity of system behavior characteristics as the following formula:

$$X_i = (x_i(1), x_i(2), \cdots, x_i(n)), i = 0, 1, 2, \cdots, m. \tag{1}$$

Definition 2. Grey correlation degree: let x_0 be the mapping quantity of system characteristics and x_i be the mapping quantity of other relevant factors of the system.

$$\gamma(x_0(k), x_i(k)) = \frac{\min\limits_i \min\limits_k |x_0(k) - x_i(k)| + \zeta \max\limits_i \max\limits_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \zeta \max\limits_i \max\limits_k |x_0(k) - x_i(k)|}, \tag{2}$$

where, $\gamma(x_0(k), x_i(k))$ represents the correlation coefficient between x_i and x_0 at point k, and ζ represents the resolution coefficient and $\zeta \in (0, 1)$.

$$\gamma(X_0, X_i) = \frac{1}{n}\sum_{k=1}^{n} \gamma(x_0(k), x_i(k)), \tag{3}$$

where, $\gamma(X_0, X_i)$ represents the grey correlation degree of x_i to x_0, and formula (3) meets the following four characteristics:

(1) Normative, indicating that any two system behavior mappings cannot be strictly unrelated.

$$\begin{aligned} &0 < \gamma(X_0, X_i) \le 1 \\ &\gamma(X_0, X_i) = 1 \Leftrightarrow X_0 = X_i \end{aligned}. \tag{4}$$

(2) Integrity, indicating that the environment has an impact on the grey correlation degree. For b$X_i, X_j \in X = \{X_s \mid s = 0, 1, 2, \cdots, m; m \ge 2\}$, there

$$\gamma(X_i, X_j) \ne \gamma(X_j, X_i)\,(i \ne j). \tag{5}$$

(3) Even pair symmetry shows that when there are only two mapping quantities of system behavior characteristics, the pairwise comparison meets the pairwise symmetry. For $X_i, X_j \in X$, there is the following formula:

$$\gamma(X_i, X_j) = \gamma(X_j, X_i) \Leftrightarrow X = \{X_i, X_j\}. \tag{6}$$

(4) Proximity indicates that it restricts the quantification of grey correlation degree.

$$|x_0(k) - x_i(k)| \text{smaller}, \gamma(x_0(k), x_i(k)) \text{ bigger}. \tag{7}$$

According to the above definition, the steps of calculating the grey correlation degree are as follows:

(i) Step 1: find the initial value image of the behavior characteristic mapping quantity of each system, as shown in the following formula:

$$\begin{aligned} X_i' &= \frac{X_i}{x_i}(1) = (x_i'(1), x_i'(2), \cdots, x_i'(n)), \\ i &= 0, 1, 2, \cdots, m. \end{aligned} \tag{8}$$

(i) Step 2: calculate the difference mapping quantity, as shown in the following formula:

$$\begin{aligned} &\Delta_i(k) = |x_o'(k) - x_i'(k)|, \\ &\Delta_i = (\Delta_i(1), \Delta_i(2), \cdots, \Delta_i(n)), \\ &i = 0, 1, 2, \cdots, m. \end{aligned} \tag{9}$$

(i) Step 3: find the maximum difference and minimum difference between the two poles, as shown in the following formula:

$$\begin{aligned} M &= \max_i \max_k \Delta_i(k), \\ m &= \min_i \min_k \Delta_i(k). \end{aligned} \tag{10}$$

(i) Step 4: calculate the grey correlation coefficient, as shown in the following formula:

$$\begin{aligned} &\gamma_{0i}(k) = \frac{m + \zeta M}{\Delta_i(k) + \zeta M}, \zeta \in (0, 1), \\ &k = i = 1, 2, \cdots, n; i = 1, 2, \cdots, m. \end{aligned} \tag{11}$$

Step 5: calculate the grey correlation degree as follows:

$$\gamma_{0i} = \frac{1}{n}\sum_{k=1}^{n} \gamma_{0i}(k), i = 1, 2, \cdots, m. \tag{12}$$

3.2.2. Grey Relational Clustering Theory. Grey correlation clustering is a method that obtains grey correlation matrix according to a grey correlation analysis and integrates some observation indexes or observation objects into several definable categories [18]. The calculation steps of grey correlation clustering are given below.

(i) Step 1: determine the project characteristic data. There are n items, each item has m characteristic data, and all the data are as follows:

$$\begin{aligned} X_1 &= (x_1(1), x_1(2), \cdots, x_1(n)), \\ X_2 &= (x_2(1), x_2(2), \cdots, x_2(n)), \\ &\cdots\cdots\cdots \\ X_m &= (x_m(1), x_m(2), \cdots, x_m(n)). \end{aligned} \tag{13}$$

Step 2: calculate the grey absolute correlation degree to obtain the grey correlation matrix. For all $i \leq j, i, j = 1, 2, \cdots, m$, the grey absolute correlation degree E_{ij} of X_i and X_j is calculated in accordance with the following formula, and the grey correlation matrix A is obtained:

$$\varepsilon_{ij} = \frac{1 + |s_i| + |s_j|}{1 + |s_i| + |s_j| + |s_j - s_i|}, \tag{14}$$

where, $|s_i| = \int_1^n (X_i - x_i(1))dt, |s_j| = \int_1^n (X_j - x_j(1))dt$°

$$A = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \cdots & \varepsilon_{1m} \\ & \varepsilon_{22} & \cdots & \varepsilon_{2m} \\ & & \ddots & \vdots \\ & & & \varepsilon_{mm} \end{bmatrix}. \tag{15}$$

$\varepsilon_{ii} = 1; i = 1, 2, \ldots, m$in matrix.

Step 3: select appropriate parameters to cluster the characteristic variables.

The critical value $r \in [0, 1]$, generally $r > 0.5$. If $\varepsilon_{ij} \geq r (i \neq j)$, X_i and X_j are similar characteristics. In practical problems, the value of r can be changed according to needs. When $r \longrightarrow 1$, the finer the classification, the fewer are the variables in each group. On the contrary, $r \longrightarrow 1$, the coarser the classification, the more are the variables in each group.

3.3. T-S Point Aggregation Diagram. T-S point aggregation diagram has been used by many scholars for the analysis of water masses because of its simple method and strong intuition. In the existing analysis, water masses are mostly divided according to the density of scattered points on the T-S point aggregation diagram based on the relative uniformity of physical and chemical properties of water masses. However, the distribution of point sets on each map will change with the change in coordinate scale, resulting in different results due to the difference between people and the map [19]. In addition, the application of the T-S diagram to analyze shallow water masses has its limitations. It is only suitable for analyzing water masses with large differences in temperature and salinity.

Although the T-S point aggregation map cannot be directly used for water mass division, it can solve many other problems. For example, when the accuracy of the data is uncertain, the most reasonable results cannot be obtained by direct systematic clustering. The points with obvious observation errors can be eliminated according to the point clustering diagram before clustering. When the analysis results are not required to be very accurate, it can quickly give qualitative conclusions, which are essential basic materials for further quantitative calculation. In addition, for the mixing zone between water masses, the space occupied by it and the difference between the core properties of the two water masses mainly depends on whether there are points on the T-S point aggregation diagram, and each point represents a certain space. In this paper, by observing whether the scattered points of each type in the T-S point aggregation diagram gather on the point aggregation diagram, we can detect whether the clustering result of the grey correlation clustering analysis is reasonable. According to the distribution of water masses with different properties on the point aggregation diagram, we can roughly obtain the temperature and salt properties of each water mass, so as to analyze the degree of water mass degeneration and the law of seasonal change.

4. Result Analysis

4.1. Water Mass Properties in Winter. In winter (January), the vertical distribution of the South China Seawater layer is obtained by using the grey correlation clustering method. From the sea surface down, it is surface water, subsurface water, sub-middle water, middle water, and deep water, which has an obvious layered structure. In the T-S point aggregation diagram, the temperature salt range of surface and subsurface seawater is large. After further horizontal division, the surface water is coastal flushing water, near-shore mixed water, South China Sea surface water, and Kuroshio surface water, and the subsurface water is divided into the South China Sea subsurface water and Kuroshio subsurface water, as shown in Table 1 below.

The deep water in the South China Sea is mainly distributed in the sea area deeper than 900 m, which is the deepest water mass in the distribution range. The temperature is lower than 5.0°C, the salinity range is 34.35–34.62, and its temperature and salt properties are basically consistent with the deep water in the Western Pacific [20].

4.2. Nature of Water Mass in Spring. Spring is the transitional season from winter to summer, and the properties of water masses are often between the two seasons. Obvious stratification still exists in each water mass. It is not difficult to find through the corresponding T-S point aggregation diagram that the temperature of water in each layer increases slightly compared with that in winter, as shown in Table 2 below.

In order to study the regional change process of the nature of each water mass, cluster analysis was carried out according to the division method of water mass in winter. In spring (April), the change law of Pacific surface water with high temperature and high salt entering the South China Sea is similar to that in winter, which gradually decreases or even disappears. Deep water in the South China Sea exists in the sea area more than 800 m deep, the temperature value is lower than 5.5°C, and the salinity range is 34.30–34.70.

4.3. Water Mass Properties in Summer. Each water mass still has an obvious stratification phenomenon. From the sea surface down, it is surface water, sub-surface water, sub-middle water, middle water, and deep water. Under the influence of thermal factors in summer, the sea surface heats up, and the atmosphere transports heat to the ocean. This heat can affect the depth of the whole South China Sea, resulting in the temperature of each water mass being higher

Table 1: Water mass properties in winter.

Arrangement	Depth	Bit density	Temperature	Salinity
Surface	0–100	20.5–23.7	21.5–29	32.70–34.70
Subsurface layer	100–200	23.7–25.5	15–21.5	34.00–34.90
Secondary middle	200–350	25.5–26.6	9.8–15	34.40–34.70
Middle level	350–900	26.5–27.2	4.5–10	34.25–34.50
Deep level	>900	27.2–27.7	2–5	34.35–34.62

Table 2: Water mass properties in spring.

Arrangement	Depth	Bit density	Temperature	Salinity
Surface	0–100	20.5–24.5	21–30.5	33.00–34.80
Subsurface layer	100–200	24.5–25.5	15.5–21	34.10–34.90
Secondary middle	200–300	25.5–26.3	11–15.8	34.30–34.70
Middle level	300–800	26.3–27.2	5.5–11	34.25–34.50
Deep level	>800	27–27.8	1.5–5.5	34.30–34.70

Table 3: Water mass properties in summer.

Arrangement	Depth	Bit density	Temperature	Salinity
Surface	0–120	20–24	22–30.5	32.50–34.65
Subsurface layer	120–200	23.5–25.5	16–22	34.00–34.85
Secondary middle	200–350	25.25–26.4	11–16	34.35–34.65
Middle level	350–800	26.3–27.2	5.5–11	34.25–34.50
Deep level	>800	27–27.8	1.5–5.5	34.30–34.70

Table 4: Water mass properties in autumn.

Arrangement	Depth	Bit density	Temperature	Salinity
Surface	0–100	20.5–24	21.5–29	32.60–34.80
Subsurface layer	100–200	24–25.4	16–21.5	34.00–34.85
Secondary middle	200–350	25.4–26.4	11–16	34.30–34.70
Middle level	350–850	26.4–27.2	5.8–11	34.20–34.45
Deep level	>850	27–27.8	2–5.8	34.30–34.80

than that in winter. However, due to the increase of precipitation in summer and the increase of runoff of continental rivers, the seawater is diluted and the salinity of seawater is reduced. Therefore, the salinity value of each water mass is slightly lower than that of the winter water mass. Through the analysis of the T-S point aggregation map, it is found that the mixing degree of South China Seawater and external seawater in surface water and subsurface water is low, which is far less than that in winter, as shown in Table 3 below.

There are two kinds of water bodies in the surface water. Under the action of the southwest monsoon in summer, the low salt water from the Java Sea mixes with the water in the Gulf of Thailand and enters the south of the South China Sea to form the continental shelf water in the south of South China Sea. Because the temperature and salt properties of the water mass are similar to the surface water of the South China Sea in summer, the two are not divided when clustering by temperature and salt values. Considering the different sources of water, the water in the southern continental shelf of the South China Sea is classified as an independent water mass, and the water mass only exists for half a year of summer due to the influence of monsoon. The deep water in the South China Sea is distributed in the sea area with a depth of more than 800 m. The temperature value is lower than 5.5°C, and the salinity value is 34.30–34.70.

4.4. Nature of Water Mass in Autumn. Autumn is the intermediate period from summer to winter, and the temperature and salt properties of water masses are similar to those in spring. From the sea surface down, there are surface water, sub-surface water, sub-middle water, middle water, and deep water. In the T-S scatter diagram corresponding to each water layer, there are two kinds of water bodies with large differences in salinity sub-middle water, middle water, and deep water. The data of high salt water bodies are from 1985 and distributed in the central sea area of the South China Sea, and the specific causes of clustering are unknown, as shown in Table 4 below.

The clustering phenomenon of the lower water body of the South China Sea is often accompanied by the clustering of the upper water body, which indicates that this phenomenon is caused by different data years. In the surface and

Table 5: Properties of water masses in the South China Sea.

Name	Symbol	Source	Depth range/m	Temperature range/°C (early winter and late summer)	Salinity range (early winter and late summer)
Coastal freshwater mass	F	Continental runoff	<20		
Nearshore mixed water mass	M	The coastal fresh water is mixed with the external seawater to form	<75		
South China sea surface water mass	S	Western Pacific surface water	0–100 0–120		
Continental shelf water in the southern South China Sea (only in summer)	TS	Java Sea water and Gulf of Thailand water mixed to form	0–120		
Kuroshio surface water mass	KS	Kuroshio surface water	0–100		
Kuroshio subsurface water mass	KU	Kuroshio subsurface water	100–250		
South China Sea subsurface water mass	U	Western Pacific subsurface water	100–200 120–200	15.0–21.5 6.0–22.0	34.00–34.70 34.00–34.65
South China Sea sub-middle water mass	UI	Western Pacific subtropical thermocline water	200–350 200–350	9.8–15.0 11.0–16.0	34.40–34.70 34.35–34.65
Middle water mass in the South China Sea	I	Western Pacific middle water	350–900 350–800	4.5–10.0 5.5–11.0	34.40–34.50 34.39–34.55
Deep water mass in the South China Sea	D	Western Pacific deep water	>900 >800	<5.0 <5.5	34.35–34.62 34.30–34.70

subsurface water bodies, the mixed water between different water masses mainly comes from the sea area near the Luzon Strait. Due to the intense mixing of surface water and subsurface water, the surface water is further divided into coastal flushing fresh water, nearshore mixed water, South China Sea surface water, and Kuroshio surface water, and the subsurface water is further divided into Kuroshio subsurface water and South China Sea subsurface water. Deep water in the South China Sea is distributed in the sea area with a depth of 850 m, the temperature value is lower than 5.5 °C, and the salinity range is 34.30–34.80.

4.5. Seasonal Variation of Water Mass. In the vertical division, the South China Seawater mass is divided into five layers: surface water, subsurface water, sub-middle water, middle water, and deep water. Through the previous discussion, on the basis of vertical division, the surface water and subsurface water are further divided, and finally, the South China Sea is divided into 10 water masses: coastal alluvial fresh water, nearshore mixed water, Kuroshio surface water, Kuroshio subsurface water, South China Sea surface water, South China Sea shelf water (only in summer), South China Sea subsurface water, South China Sea sub-middle water, South China Sea middle water, and the South China Sea deep water. The deep water in the South China Sea is mainly distributed in the basin with a depth of 900m. The temperature value is lower than 5.5°C, and the salinity range is 34.30–34.70. The temperature and salt properties are basically consistent with the deep water in the Western Pacific. The seasonal and regional differences in temperature and salt properties of this water mass are not obvious. It is the most stable water mass in the South China Sea. The properties of each water mass are shown in Table 5 below.

By analyzing the properties of temperature and salt in each region, it is not difficult to find that the water masses in the South China Sea are formed after the Pacific water enters the South China Sea through the Luzon Strait. The surface and subsurface water are significantly affected and strongly denatured by the external seawater. The degree of denaturation of sub-middle, middle, and deep water masses is weak. After entering the South China Sea, the Kuroshio water is mainly distributed in the northwest of Luzon Strait. The invasion intensity is stronger in winter than in summer, and the sub-surface water of Kuroshio is stronger than the surface water of Kuroshio. The properties of the above water masses are basically consistent with the existing conclusions, but there are still a few differences, which need to be explained by the interannual variation law of water masses.

5. Conclusion

Based on the WOD13 temperature and salt observation data from 1966 to 2013, this paper uses the grey correlation clustering method to divide the water masses in the whole South China Sea and uses the system clustering method to support the division results. Combined with the T-S point clustering map of the whole region and sub-region of the South China sea, this paper analyzes the distribution, temperature and salt properties, and seasonal variation law of each water mass in the South China Sea. In the clustering process of water masses, this paper creatively uses the intraclass distance sum function and density value function to determine the “number of water masses” and “initial center” respectively, which improves the efficiency and accuracy of operation. Through the grey correlation clustering method, the water mass in the South China Sea is vertically divided into five layers, namely surface water, subsurface water, sub-middle water, middle water, and deep

Hindawi
Mobile Information Systems
Volume 2022, Article ID 7872897, 9 pages
https://doi.org/10.1155/2022/7872897

Research Article

Application of Edge Computing and Data Mining Processing System in Preschool Education Courses

Baoliang Xu

School of Education Sciences, Lingnan Normal University, Zhanjiang, Guangdong 524048, China

Correspondence should be addressed to Baoliang Xu; xubaoliang137@126.com

Received 29 July 2022; Revised 12 September 2022; Accepted 19 September 2022; Published 30 September 2022

Academic Editor: Jiafu Su

With the deepening of reform and opening up, my country's preschool education has entered a vigorous development. At the same time, higher requirements are put forward for the field of preschool education. Preschool education is an important part of education, and preschool education is the foundation of basic education. However, the informatization construction in the field of preschool education in my country has just started. In order to implement the planning requirements of the Ministry of Education on strengthening the informatization construction of preschool education, build a digital platform, improve children's quality, promote the harmonious development of children's body and mind, improve professional ethics and the professional level of preschool education practitioners with the help of modern new media, and meet the professional development and sense of achievement, acquisition, and satisfaction of preschool education practitioners. Taking multiple measures to ensure the diversified needs of preschool education practitioners is the important and only means for the stable, comprehensive, and sustainable development of preschool education. The data visualization platform of IOT architecture designed in this paper for edge computing ensures the application of massive data in preschool education courses, ensures the stability of system data transmission, and realizes the visualization of IOT terminal data. Therefore, the introduction of data mining technology in this paper can improve the utilization of resources, provide support for teachers, students, and ordinary users, and enhance the versatility of the preschool education resource platform.

1. Introduction

Today, with the reform of China's education system, preschool education has become more and more important [1]. Many scientific research institutions and educational institutions actively introduce information technology and develop and build a preschool education resource platform [2]. These platforms can provide teachers, students, and administrators [3]various resources, such as children's mental health, children's intellectual development, children's creativity improvement, and other knowledge content [4]. With the long-term operation of the platform, a large amount of data information has been accumulated, so it is very important to actively improve the utilization of preschool education resources [5].

As the foundation of basic education, the importance of preschool education is self-evident [6]. After the opening of the comprehensive two child policy, the number of preschool children has increased significantly [7]. These factors also make China's preschool education face greater pressure than before [8]. The 19th CPC National Congress proposed that preschool education should realize the deployment of early childhood education [9].

The opinions of the CPC Central Committee and the State Council on deepening the reform and standardized development of preschool education issued in 2018 clearly pointed out that the unbalanced and insufficient development of preschool education in China is still very prominent, and it is still a weak link in the whole education system [10]. One of the problems is that the construction of teaching staff lags behind, so the countermeasures of "Innovating the training mode and optimizing the training curriculum system" are put forward [11]. The quality of preschool education talent training depends on the curriculum of preschool education major in each school [12].

At this stage, the preschool education talents cultivated by secondary vocational schools are still an important part of kindergarten teachers, but their graduates' vocational adaptation level is low, and their comprehensive quality and ability have not yet reached the postobjectives [13]. The fundamental reason for this phenomenon is that at present, the preschool education curriculum of secondary vocational schools mainly adopts the traditional three-stage curriculum model, emphasizing theory over practice and ignoring the connection between knowledge and specific work tasks, and it cannot effectively cultivate students' comprehensive quality, key capabilities, and professional ability.

Based on this situation, it is necessary to comprehensively adjust the curriculum system of preschool education in secondary vocational colleges to truly meet the requirements of kindergarten teachers' professional ability [14].

From the development situation in recent years, the development level of preschool education has been significantly improved, and the demand for preschool education teachers is also rising [15]. It is undeniable that for modern society, higher education has obvious characteristics of popularization, but most kindergarten teachers are provided by the preschool education major of secondary vocational schools [16]. However, the phenomenon of low employment quality that does not match the high employment rate is very prominent [17]. Graduates generally do not have good job adaptability, and their comprehensive professional characteristics are poor, which is difficult to meet the requirements of kindergarten teachers' professional ability [18]· The low quality of talent training reflects the problems in the curriculum of preschool education in secondary vocational schools, which cannot achieve the goal of kindergarten vocational ability training [19]. From the current development situation, the preschool education major in secondary vocational schools in China is mostly divided into three curriculum stages, paying attention to theoretical research methods and ignoring the key impact of practical operation [20]. In the long-term development, the relationship between knowledge and specific work tasks is bound to be ignored, and students' key abilities and professional abilities cannot be effectively cultivated [21].

To sum up, with the continuous development of society, the importance of preschool education has increased significantly. These factors have led to the further improvement of the requirements for the training of preschool education professionals. In terms of curriculum design, it basically does not meet the requirements of the professional ability of nurses, and it is difficult to meet the actual needs. Since the concept of competency based vocational education was put forward, it has gradually attracted the attention and research of vocational education in China and has been widely used in the curriculum reform of vocational education.

As a front-line educator of secondary vocational preschool education, finally, the author has a real understanding of the current situation of the curriculum of this major and believes that we should jump out of the current three-stage curriculum system of secondary vocational schools and reform the curriculum model under the guidance of colleges and universities. Vocational education concept: this study will actively carry out case studies on the curriculum of preschool education in secondary vocational schools on the basis of the necessary ability-based vocational education, effectively analyze the current situation of the curriculum of case schools from the perspective of ability based, and put forward relevant suggestions. Edge learning is a data mining and analysis algorithm, which can recommend the required learning materials from a large number of preschool education resources according to the needs of users. It has an important role and significance.

2. Related Concepts and Theoretical Basis

2.1. Concept Definition. The full name of secondary vocational school is "secondary vocational and technical school." The secondary vocational and technical school plays an important role in China's vocational education system and is in the core development position. Secondary vocational and technical middle school mainly refers to the school that provides vocational and technical education for students in the whole high school stage. In secondary vocational and technical secondary schools, the academic system of general subjects is a three-year system, that is, the academic year is three years. The main types of secondary vocational and technical middle schools are ordinary middle schools, vocational high schools, and technical schools, which are referred to as "technical schools" in short.

Preschool education refers to the school education specially for preschool children aged 0–6. Due to the special situation of preschool children, the staff engaged in preschool education must have the corresponding professional knowledge. For secondary vocational and technical schools, a special "preschool education" major is set up to cultivate comprehensive talents who are competent for preschool education and have professional knowledge and skills of preschool education. For students, it is undoubtedly a very wise choice to work in nurseries and kindergartens, which is very consistent with the purposes and objectives of secondary vocational and technical schools.

As a professional term in the field of educational science, each researcher has different understanding and interpretation of the significance of curriculum based on different research contents and perspectives.

The education dictionary says that "the courses taught in schools are an organic whole with due structure and function, so it is also called *f* curriculum system" or "curriculum structure." It mainly refers to that the school usually determines the specific training objectives based on the overall school running objectives, the nature of the school itself, and the nature of disciplines and majors, and it defines the form of professional learning and the vocational skills and working skills that need to be developed. Therefore, the division of labor and cooperation between courses is reasonable and determined.

The concise encyclopedia of international education has different understandings on the curriculum. William Hubert, an American curriculum expert, believes that "curriculum design refers to the full range and characteristics of the curriculum arranged by schools or other institutions." Liang pointed out in the educational theory volume of the Encyclopedia of Chinese preschool education that from a macro perspective, it mainly refers to the whole learning plan. The course aims to achieve the training objectives of the school and selects educational content. The division of labor "reflects the purpose and nature of running a school," mainly the selection and arrangement of learning content.

Another domestic researcher pointed out that curriculum design is an important part of curriculum design, that is, under the guidance of certain educational values and educational development needs, talent training units will consider and formulate appropriate talent training goals according to the needs of social development, discipline research and development, as well as the special needs of students' personal career development and life pursuit. Standardize the development and setting of the curriculum system, organize the selection and arrangement of talents, determine teaching subjects and training objectives, and promote the establishment and improvement of the professional talent training system.

2.2. Theoretical Basis for the Curriculum Setting of Preschool Education in Secondary Vocational Schools. In the process of practice, we cannot do without the guidance of scientific theory. Only in the process of practice under the guidance of scientific theory can we achieve good results. The curriculum should be scientific and reasonable, with both professional theoretical basis and relevant policy basis. The concept of competency standard vocational education breaks the restrictions of China's traditional discipline standard talent training mode and is more conducive to improving the quality of vocational education talent training. Based on this, this study selects three secondary vocational schools in Qingdao as perfect kindergartens, which is the theoretical basis of educational curriculum. The theory of competency-based vocational education curriculum originated in the United States in the 1960s. It is believed that the traditional educational concept pays too much attention to the explanation of book knowledge points and ignores the cultivation and improvement of students' practical ability. Therefore, it greatly affects the effect of education. The ability-based vocational education curriculum is mainly to improve students' knowledge and professional skills on the basis of clarifying the established job requirements and vocational ability indicators, combined with students' own abilities and levels.

Compared with other previous educational theories, this theory pays more attention to students' practical ability in addition to the study of textbook theory. It is mainly under the basic premise of vocational ability training to ensure that students can be competent for corresponding employment positions after learning. It is a relatively advanced educational concept at this stage and has been widely recognized all over the world. Its basic characteristics include first, pay attention to the analysis of postvocational ability requirements, set corresponding professional courses and evaluation indicators on this basic premise, and take the vocational ability training objectives as the basic content requirements of the post. Curriculum education and reasonable training: second, pay attention to the improvement of students' comprehensive quality, carry out targeted training according to different students' comprehensive level and learning ability, and adopt a differentiated evaluation system to improve education efficiency; third, attach great importance to the improvement of students' autonomous learning ability. Teachers are the main guides of students, guiding students to make their own learning plans and arrangements, answering students' questions and doubts, giving full play to students' autonomy and initiative, and being responsible for students' personal growth and future. The teacher's first process here should not exceed 1000 predictions, and only the guidance and demonstration effect can be brought into play. Fourth, teaching methods should be more diversified and flexible, emphasizing the improvement and science of management mechanism.

In general, this theory plays a very important role in talent training. We need to truly meet the job needs and cultivate more excellent talents for the society. They not only have solid professional theoretical knowledge but also have good practical ability and can adapt to society and positions as soon as possible. Therefore, in the future curriculum system construction, we should take the vocational ability training as the basis of all work, pay attention to the cultivation of students' practical operation ability, meet the needs of diversified and personalized education, and establish a student-oriented professional ability. The scientific evaluation index system is used to comprehensively realize the construction of the education system. Secondary vocational preschool education is an important foundation and source to provide a large number of kindergarten teachers to the society and comprehensively improve the level of educational practice. Combined with the relevant concepts of competency-based vocational education curriculum, it can give full play to its guiding role in professional curriculum setting and improving students' comprehensive professional ability.

According to the vocational standards for kindergarten teachers (trial) issued in recent years, there are clear definitions and requirements for kindergarten teachers and preschool education practitioners. The document points out that kindergarten teachers should receive professional vocational education and must have noble morality, love education, and have sufficient professional knowledge and professional technical ability. This standard has become an important reference for kindergarten teacher training.

In addition, the vocational standards for this trial also put forward specific implementation suggestions: schools should comprehensively analyze the requirements of school employment, professional characteristics, and the current situation of existing teachers, pay attention to improving the level of relevant majors, ensure the rationality of curriculum, promote employment, pay full attention to pre-employment training and postemployment retraining and education, realize the key significance of professional ethics for their own development, guide them to actively participate in various social practices, and comprehensively improve the professional level and professional technical ability of kindergarten teachers and preschool education practitioners.

3. Related Technologies

3.1. Edge Computing. Edge computing was explained at the SIDCOM (Special Interest Group on Data Communication) conference in the field of communication networks. Edge computing originally originated from the media field. Its meaning is a platform that integrates key technologies such as data collection, network transmission, data storage, application calculation, and application on the side close to the data source or object, providing efficient services for applications under edge nodes. Edge computing is further divided into generalized edge computing and narrow edge computing. As shown in Figure 1, it presents a typical multilayer edge computing architecture.

Edge computing in a narrow sense means that the computing node is located on the data source, and the main application of edge computing in the narrow sense is the camera of video surveillance. Video surveillance cameras often contain CPUs and GPUs. Performing calculations in the camera equipment can realize face recognition in the monitored images, which can improve the speed at which the monitoring personnel receive alarm responses. As a result of using such a design, the cost of the terminal equipment is very high, the quality and volume of the terminal equipment are doubled compared to the traditional equipment, the power consumption is also very large, and the heat dissipation function is poor. In the industrial Internet of Things application environment, that is, data processing is performed on the sensor terminal, and most application scenarios require low-power products as terminal devices, and even some products are powered by batteries, which have certain requirements on the power consumption of terminal devices. Therefore, the edge computing described in this article refers to the generalized edge computing.

In a broad sense, edge computing refers to being far away from the cloud, being under the same network node as an object or data source that is close, and performing data processing. This kind of transmission process from the sensor of edge computing in the narrow sense to the cloud is called the computing node, and the calculation of the node after the data settlement in the same network environment is also called fog computing. At present, most of the edge computing technology research refers to generalized edge computing, and because the computing nodes of generalized edge computing are not limited by the conditions of IoT terminals, developers can configure servers with different computing power according to their needs so as to realize edge computing. The calculation of points can be applied in more and more complex scenarios.

To sum up, the three data visualization graphics libraries mentioned above will be compared in this article from the perspectives of compatibility, underlying implementation and open source, as shown in Table 1:

3.2. Data Mining. The extensive research and extensive application of data mining can be regarded as the result of the natural evolution of information technology. As a multidisciplinary field, data mining can be defined in many ways. Even if the term "data mining" itself is interpreted, it cannot fully cover the rich content it contains.

Both clustering and outlier analysis can be classified as unsupervised learning, as shown in Table 2. Clustering tasks analyze data objects without regard to class labels, usually due to the absence of labeled class data in the dataset. In general, outliers are discarded as noise or outliers in most data mining tasks. However, in some application scenarios such as fraud detection, outlier analysis can be interesting. Outlier analysis, also known as anomaly mining, is generally closely related to cluster analysis methods.

$$\Delta(x_1, x_2) = \Delta(x_2, x_1). \tag{1}$$

Then, (Ω, Δ) represents the metric space; (x, xj) represents the distance function between the elements xi and x. There are three functions that are often used to calculate distances: Manhattan (Equation 2), Euclidean (Equation 3), and P-norm (Equation 4).

$$\Delta(x_i, x_j) = \sum_{k=1}^{N} \left| f(x_i, a_k) - f(x_j, a_k) \right|, \tag{2}$$

$$\Delta(x_i, x_j) = \left(\sum_{k=1}^{N} \left(f(x_i, a_k) - f(x_j, a_k) \right)^2 \right)^{1/2}, \tag{3}$$

$$\Delta(x_i, x_j) = \left(\sum_{k=1}^{N} \left| f(x_i, a_k) - f(x_j, a_k) \right|^P \right)^{1/p}. \tag{4}$$

Let the nonempty finite set U = {x1, x2, . . . , xn} on the real space Ω, then the neighborhood δ-definition of ∀xi is

$$\delta(x_i) = \{x \mid x \in U, \Delta(x, x_i) \le \delta\}. \tag{5}$$

Let the nonempty finite set U = {x1, x2, . . . , xn} on the real number space Ω and the neighborhood relation N on it, that is, the two-tuple NS = (U, N), ∀X ⊆ U; then X is the upper and lower approximations in the neighborhood approximation space NS = (U, N) which are

$$N^{-} X = \{x_i \mid \delta(x_i) \cap X \ne \varnothing, x_i \in U\}. \tag{6}$$

Then, the approximate boundary of X (7)is obtained, the lower approximation N−X is the positive domain, and the area completely unrelated to X is the negative domain.

$$BN(X) = N^{-} X - N\_X. \tag{7}$$

Set the neighborhood decision system NDS = (U, A ∪ D), the decision attribute D divides the universe U into N equivalence classes (X1, X2, . . . ,XN), ∀B ⊆ A; then the decision attribute D is about the upper and lower approximations of set B which are expressed in the following equations Eq. 8 and Eq. 9, respectively.

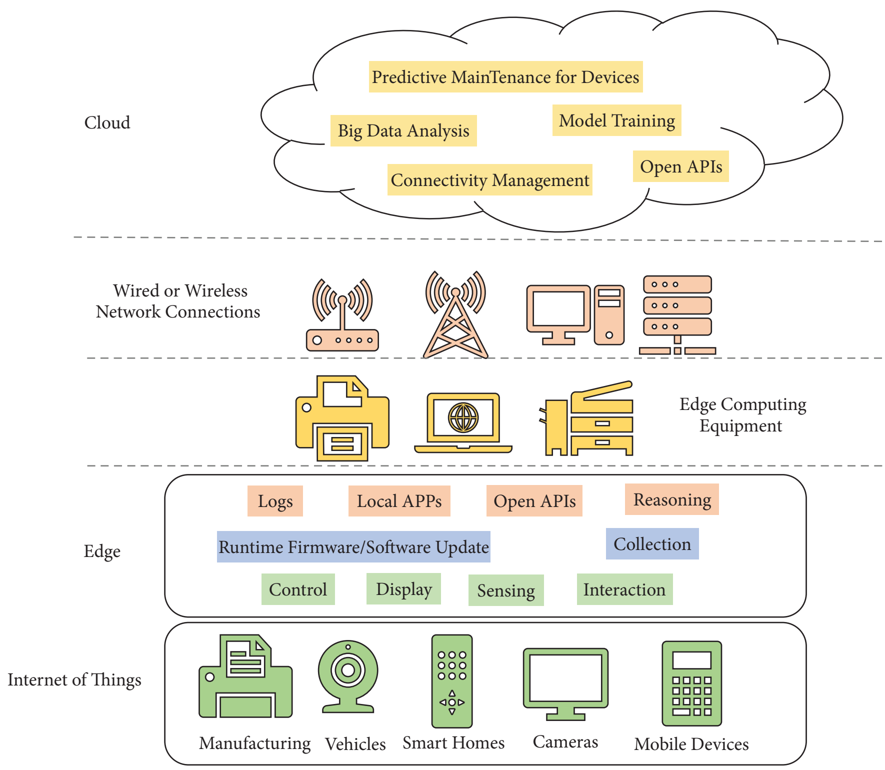


FIGURE 1: Edge computing architecture.

TABLE 1: Comparison of visual chart libraries.

Visual chart	Compatibility	The underlying implementation	Open source
ECharts	Compatible with all major browsers IE6 and above, also supports zooming and gestures on mobile	Canvas, based on vector graphics library ZRender, provides intuitive, interactive, and highly customized data visualization charts	Free Github 49.9k star
Highcharts	Internet Explorer 6 and above all major browsers, perfect support for mobile zoom, gesture operation	SVG, a simple and easy way to add interactive charts to web sites or web applications, has a limited chart library	Commercial payment Github 10.4k star
D3.js	Compatible with all major browsers IE9 and above	Canvas and SVG need custom development	Free Github 100k star

TABLE 2: Detailed introduction of data mining functions.

Subfunction	Meaning	Methods	Output	Example
Data characterization	Characterisation of data refers to a collection of general features or characteristics of the target class data, often using queries to collect data corresponding to user-specified classes.	Simple data summarization based on statistical measures and graphs, OLAP roll-up operation based on data cubes and attribute-oriented induction techniques are adopted.	Graph method (e.g., pie chart, bar chart, curve, and so on) and generalized relation and rule description method.	To investigate the characteristics of an aquatic product with a 5% increase in sales in the previous quarter, data on this product can be collected by executing SQL queries on the sales database.
Data to distinguish between	Compares the general characteristics of the target class data object with those of one or more contrast class data objects.	The methods used for data differentiation are similar to those used for data characterization.	The form generally includes comparison measures to help distinguish the target class from the comparison class.	Users want to compare the 5% increase in sales of aquatic products in the last quarter with the 20% decrease in sales of aquatic products in the same period.

Let $\Delta = R \times RN \longrightarrow R$ in N-dimensional real number space Ω; then, Δ is a metric (distance) on RN, if Δ satisfies the following conditions.

$$R = (r_1, \ldots r_{10}, ar_{11}, \ldots ar_{15}, br_{16}, \ldots br_{20}), \tag{8}$$

$$\text{SCDDF} = \left(\overrightarrow{OM}; r_1, \ldots r_M; dr_1, \ldots, dr_M; \theta_i\right). \tag{9}$$

Dependency of decision D on condition B

$$k_D = \gamma_B(D) = \frac{|Pos_B(D)|}{|U|}. \tag{10}$$

The algorithm is more effective in attribute reduction. This conclusion has been given. The verification data come from the UCI data set. The results after attribute reduction are shown in Figure 2.

4. Experimental Results and Analysis

4.1. Data Collection and Testing Environment. Qingdao vocational and technical middle school is the only public vocational and technical school in Qingdao Development Zone, and it is also a national key vocational middle school. Qingdao vocational and technical middle school has 89 classes, more than 5000 students, and 282 teaching staff, of which more than 70% of the teaching staff have intermediate and senior titles, and 117 double qualified teachers. There are 16 teaching majors, including preschool education, animation design, accounting, logistics management, business Korean, electromechanical technology application, computer, and so on. Qingdao vocational and technical middle school takes improving students' employability as the basic development direction, serves students and the society, pays attention to the cultivation of students' professional and technical ability, the teaching of professional knowledge, and the cultivation of professional and professional quality as the purpose, and adheres to the school running philosophy of "adult, talent, success."

The preschool education major of Qingdao vocational and technical middle school began to recruit students in 2001. Up to now, the number of graduates of the preschool education major of the school has exceeded 2000, mainly concentrated in kindergartens in small and medium-sized towns such as prefectures, cities, and towns. At present, the preschool education major of the school has 360 students and 25 full-time teachers. At the same time, the major is also equipped with piano room, art classroom, dance classroom, vocal music classroom, and other training classrooms.

The hardware configuration of the edge computing server of this system is shown in Table 3.

Chrome is currently the most widely used browser and provides a complete page testing function. This article will use the Chrome browser (Developer Tools) developer tools to analyze and test platform performance and metrics. In the Network control panel of the developer tool, the platform can initiate the request to the server to the whole process of receiving the returned data. The detailed information in the Network includes the URL path, HTTP request header, request parameters, HTTP status code, resource file, and server response information delay.

4.2. Analysis of the Current Situation of the Curriculum Setting of the Preschool Education Major in the Case School. According to the curriculum concept of this study, that is, according to the concept of the training objectives of the preschool education major of secondary vocational education, plan and design the curriculum, formulate a reasonable curriculum framework system and curriculum content structure, and provide targeted suggestions and opinions on class hour arrangement and curriculum arrangement according to the different teaching conditions of each academic year so as to create a complete preschool education curriculum system of secondary vocational schools. Through interviews, questionnaires, and data review, this study will analyze the current situation of the training objectives, curriculum structure, and discipline setting of preschool education major in the three case schools and investigate and understand the implementation of the curriculum.

According to the survey and statistics in Figure 3, only 8% of the students in school are satisfied with the training objectives of their major. These data fully show that students' satisfaction with the training objectives is very low. At the same time, nearly 20% of the students do not understand the training objectives of their major, and only 2.01% of the students think that they can achieve the training objectives by studying professional courses, but they are very pessimistic about the realization of the training objectives.

Through the results of interviews and questionnaires, it can be seen that the students and teachers in Qingdao secondary vocational and technical middle school generally pay less attention to and agree with the training objectives of preschool education. Many of them, not only students but also some teachers do not know how to achieve the training objectives.

It can be seen from the survey results in Table 4 that among the three schools, only Qingdao preschool normal school offers elective courses, accounting for only 5.19%. In the survey of "whether it is necessary to set up elective courses according to students' interests," 63.12% of students think it is necessary, 17.51% of them think it is necessary, and 11% of them think it is normal. Only 7.54% thought it was unnecessary. It can be concluded that students still generally believe that it is very necessary to set up elective courses.

Based on the optimization of Canvas instructions, the cache optimization method is used. After using the cache optimization method to execute the drawing instructions, the graphics are first saved in the off-screen Canvas, which can optimize the original many drawing instructions into one instruction. The optimized result can render the frame rate up to 60 fps.

As shown in Figure 4, the unoptimized page using the native ECharts graphics library and the page using 5 optimization methods are refreshed 20 times, and finally, the average rendering frame rate statistics are obtained. The frame rate of rendering using the rendering optimization method is significantly improved.

It can be seen from Table 5 that the arrangement of the school's professional skills classes and professional theory classes is not the same, but in general, the class hours of

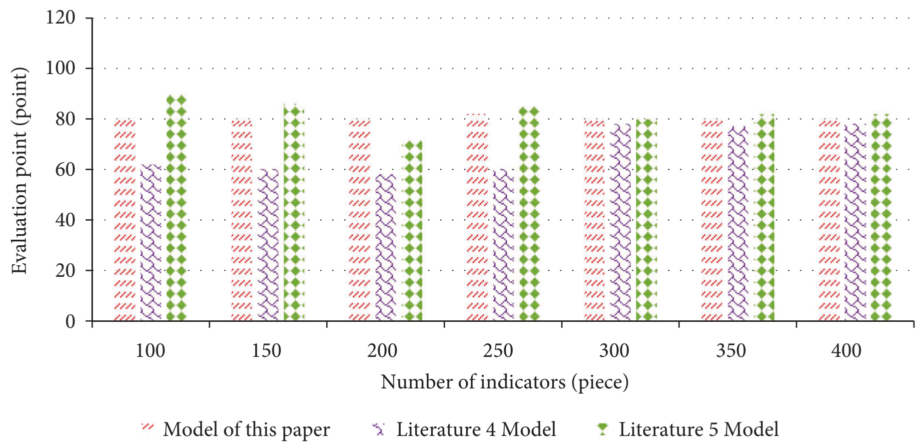


FIGURE 2: Comparison of the number of attributes and classification performance after reduction.

TABLE 3: Edge computing server configuration information table.

Browser	Rendering engine	Browser version	The operating system	System version
Edge	Trident	95.0.1020.40	Windows	Windows 11
Chrome	WebKit	95.0.4638.54	Windows	Windows 10
Firefox	Gecko	86.0	Windows	Windows 10
Safari	WebKit	11	MacOs	10.12.6
QQ browser	Chromium + Trident	10	Windows	Windows 10

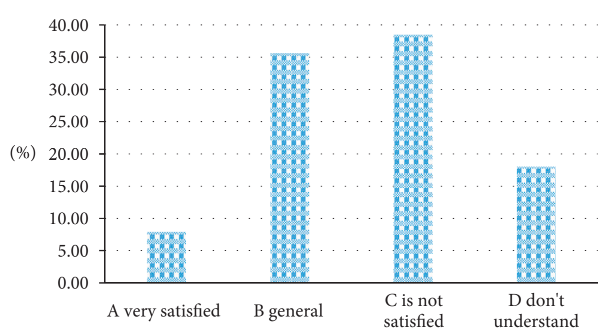


FIGURE 3: The survey of students' satisfaction with the training objectives.

TABLE 4: Comparison of school hours between elective courses and compulsory courses.

Optional course hour	The percentage	Compulsory course hours	The percentage
160	5.19	2090	94.81
0	0	3080	100
0	0	3080	100

professional skills classes are higher than the class hours of professional theory classes. In the questionnaire survey of "The Most Lack of Knowledge and Skills During the Internship," it can be seen from the survey results in Figure 4 that nearly 70% of the students believe that the basic theoretical knowledge of teaching and the teaching methods in the five major areas are relatively lacking, which are all majors. 13.68% of the students think that the content of professional skills in theoretical courses is relatively lacking, which shows that the proportion of teaching hours of teaching basic theories and teaching methods in the five major areas can be appropriately increased.

At present, there are still significant differences in quantity and quality of preschool education resources between urban and rural areas and between regions. Taking the student-teacher ratio as an example, since 2013, the student-teacher ratio in various regions has been changing in a better direction for five consecutive years, and the difference between regions has narrowed significantly as shown in Figure 5. However, although the difference has narrowed, the gap between the eastern region and the central and western regions is still relatively large. In 2017, the student-teacher ratio of kindergartens in the eastern region was 16.41 : 1, which was at a good level, while the student-teacher ratios of kindergartens in the central and western regions were 21.18 : 1 and 21.10 : 1, respectively, and the eastern region was significantly better than the central and western regions; the imbalance between regions has not been broken.

This chapter uses some statistical analysis methods to analyze some characteristics of preschool education. Relevant data were obtained in the form of questionnaires. Through descriptive analysis of the data, this paper introduces the basic information of the data to describe and analyze the students' network use and online learning behavior. This paper analyzes the differences and introduces the differences of different characteristics on E-learning behavior. Through correlation analysis, this paper

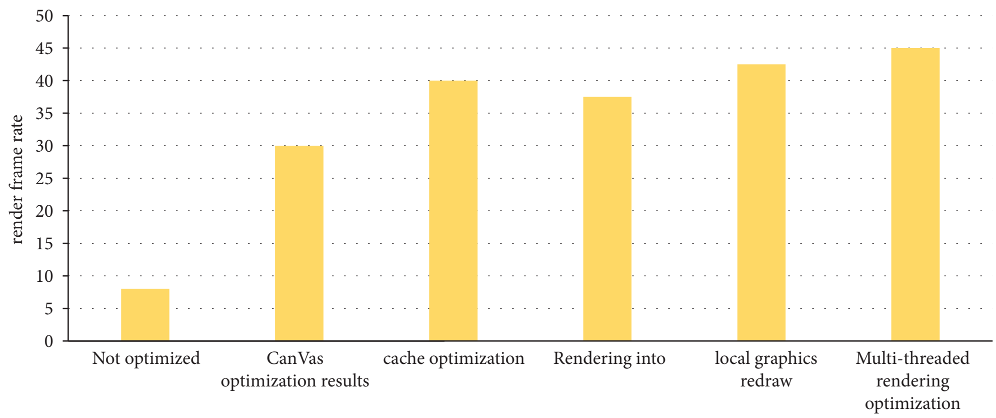


Figure 4: Rendering frame rate comparison results of five optimization methods.

Table 5: Basic information of professional theory courses and professional skill courses in the three schools.

Professional theory class hours	The proportion	In professional skill class	The proportion
850	27.60	986	32.01
444	14.45	890	28.90
470	15.23	842	27.34

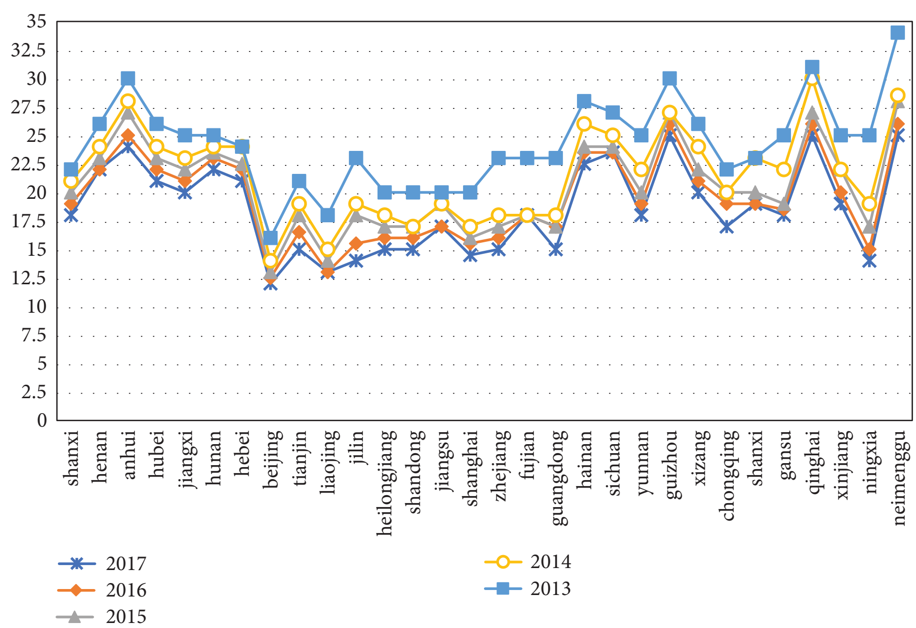


Figure 5: Proportion of kindergarten teachers and students nationwide (number of kindergarten children divided by number of full-time teachers) (2013–2017).

introduces the relationship between various factors in E-learning behavior.

5. Conclusion

The preschool education resource platform introduces edge learning and data mining technology, which can improve the utilization efficiency of preschool education resources. Edge learning and data mining technology are introduced into the early childhood education resource platform to improve the utilization efficiency of early childhood education resources. Edge learning can process video data, image data, and ordinary text data, further improve the data organization and discovery ability of the preschool education resource platform, and output the results to the display to realize children's visual data processing and operation service education. Edge learning is also one of the important technologies of the current computer, which can improve the service processing efficiency of big data platform and ensure the processing speed and automation level of the preschool education platform. Edge learning is one of the key technologies in data analysis of preschool education resources. At present, convolutional neural networks can be used to build big data analysis and recommendation models and dynamically update and process algorithms to ensure that the algorithms can accurately process preschool teaching resources. We need to improve the accuracy of the application of preschool education resources.

Data Availability

The labeled data set used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This work was supported by Lingnan Normal University.

References

[1] W. S. Barnett and K. Jung, "Effects of New Jersey's Abbott preschool program on children's achievement, grade retention, and special education through tenth grade," *Early Childhood Research Quarterly*, vol. 56, pp. 248–259, 2021.

[2] E. Smith, T. Sutarso, and G. L. Kaye, "Access with education improves fruit and vegetable intake in preschool children," *Journal of Nutrition Education and Behavior*, vol. 52, no. 2, pp. 145–151, 2020.

[3] N. Wieduwilt, S. Lehrl, and Y. Anders, "Preschool teachers' pedagogical beliefs in the field of language education," *Teaching and Teacher Education*, vol. 101, no. 10, Article ID 103296, 2021.

[4] M. Burchinal, K. Garber, T. Foster, M. Bratsch-Hines, X. Franco, and E. Peisner-Feinberg, "Relating early care and education quality to preschool outcomes: the same or different models for different outcomes? [J]," *Early Childhood Research Quarterly*, vol. 55, no. 19, pp. 35–51, 2021.

[5] O. Molchaniuk, O. Babakina, K. Dmytrenko, A. Kharkivska, and O Kapustina, "Partial programs in access to the social values: technology that changes preschool education," *Journal of Physics: Conference Series*, vol. 1691, no. 1, Article ID 012176, 2020.

[6] I. Shiyan, T. Le-Van, and O. Shiyan, "The key issues in implementing the federal state education standard for preschool education. Results of the national survey of preschool education quality in 2016-2017[J]," *Voprosy obrazovaniya/Educational Studies Moscow*, vol. 3, no. 6, pp. 854–864, 2021.

[7] N. Machynska, S. V. Lozynska, and Y. Y. Dekach, "Theoretical and practical aspects of valeology € the constituents of the system of preschool education[J]," *Revi sta romaneasca pentru educatie multidimensionala - Journal for Multidimensional Education*, vol. 13, no. 1, pp. 54–60, 2021.

[8] L. J. Liao and F. J. Gu, "5G and artificial intelligence interactive technology applied in preschool education courses," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 7629354, 11 pages, 2022.

[9] E. M. Melnick, K. Thomas, C. Farewell et al., "Impact of a nutrition education programme on preschool children's willingness to consume fruits and vegetables," *Public Health Nutrition*, vol. 23, no. 10, pp. 1846–1853, 2020.

[10] M. D. C. Ramírez-Rueda, R. Cózar-Gutiérrez, M. J. Roblizo Colmenero, and J. A Gonzalez-Calero, "Towards a coordinated vision of ICT in education: a comparative analysis of Preschool and Primary Education teachers' and parents' perceptions," *Teaching and Teacher Education*, vol. 100, no. 2, Article ID 103300, 2021.

[11] N. Megapanou, "Preschool teachers' views on the outcomes of preschool inclusive education for students with special educational needs and typically developing students[J]," *Open Access Library Journal*, vol. 9, no. 3, pp. 17–22, 2022.

[12] R. G. Fryer, S. D. Levitt, and J. A. List, "Introducing CogX: a new preschool education program combining parent and child interventions[J]," *NBER Working Papers*, vol. 7, no. 8, pp. 6722–6747, 2020.

[13] I. D. Emelyanova, S. V. Markova, and O. A. Podolskaya, "Professional competence of educators of the system of inclusive preschool education as a condition for sociolizing the physically challenged children[J]," *Revista Tempos e Espacos em Educacao*, vol. 54, no. 05, pp. 907–924, 2020.

[14] Y. Vyshpinska, "formation of creative personality of students majoring in preschool education in the process of studying the methods of musical education[J]," *Book chapters-LUMEN Proceedings*, vol. 16, no. 6, pp. 4672–4683, 2021.

[15] P. Spadaccino and F. Cuomo, "Intrusion detection systems for IoT: opportunities and challenges offered by edge computing [J]," *IEEE Communications Surveys & Tutorials*, vol. 32, no. 1, pp. 24–32, 2020.

[16] M. Garmaroodi, F. Farivar, and M. S. Haghighi, "Detection of anomalies and faults in industrial IoT systems by data mining: study of CHRIST osmotron water purification system," *The Journal*, vol. 56, no. 9, pp. 103–109, 2020.

[17] D. C. Nguyen, M. Ding, and Q. V. Pham, "Federated learning meets blockchain in edge computing: opportunities and challenges[J]," *arXive e-prints*, vol. 37, no. 05, pp. 241–247, 2021.

[18] J. Sharma, D. Kim, and A. Lee, "On differential privacy-based framework for enhancing user data privacy in mobile edge computing environment[J]," *IEEE Access*, vol. 12, no. 99, pp. 1–11, 2021.

[19] Y. Gordienko, Y. Kochura, and V. Taran, "Last mile" optimization of edge computing ecosystem with deep learning models and specialized tensor processing architectures[J]," *Advances in Computers*, vol. 33, no. 03, pp. 23–25, 2020.

[20] S. Weiwei, "Design of auxiliary teaching system for preschool education specialty courses based on artificial intelligence," *Mathematical Problems in Engineering*, vol. 2022, Article ID 4504707, 9 pages, 2022.

[21] V. Demirci, "Implementing quality management systems in preschool education[J]," *Book chapters-LUMEN Proceedings*, vol. 15, no. 4, pp. 73-74, 2021.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 4377286, 10 pages
https://doi.org/10.1155/2022/4377286

Research Article

Research on Development Strategy of Ethnic Sports Tourism Resources Based on Stochastic Forest Algorithm

Chunli Nie

College of Education, Guangxi Vocational Normal University, Nanning, Guangxi 530007, China

Correspondence should be addressed to Chunli Nie; ncl1985@126.com

Received 11 May 2022; Revised 15 July 2022; Accepted 28 August 2022; Published 26 September 2022

Academic Editor: Jiafu Su

The interactive development mode of combining sports industry with tourism industry has attracted more and more people's attention and gradually been accepted by people. The organic combination of the two can effectively promote the healthy growth of regional economy. The development and operation mode of minority STRs (sports tourism resources) has become an important topic in the research of sports tourism development. It is not only important for the successful integration of minority traditional sports into tourism industry but also of great significance for promoting the development of sports tourism industry. In this paper, the personalized sports tourism service recommendation system based on multi-objective optimization is studied, and a tourism service combination method based on multi-objective optimization is proposed. This method is based on multi-objective optimization, and the historical data of tourists and their current preferences are considered, respectively. The converted data are used to train the RF (random forest) model offline, and online recommendation only needs to be scored and predicted according to the rules of the RF model. The results show that the online recommendation time of the proposed algorithm is basically below 100 s, which is much lower than that of other recommendation algorithms. The experimental results show that setting the weights of user evaluation information and related tourism information can further improve the matching degree between recommendation results and users' needs.

1. Introduction

China is a multi-ethnic country, and each ethnic group has its own customs and habits. As a part of its excellent culture, ethnic traditional sports embodies the hard work and wisdom of the working people of all ethnic groups in the process of life and production, and it has survived through years of tempering. With the continuous improvement of people's living standard and the change of life concept, people have new requirements for spiritual life on the basis of satisfying material life. In recent years, the domestic tourism demand has been rising year by year. Combining traditional national sports with tourism, promoting the development of ethnic traditional sports, publicizing China's traditional culture, and creating economic benefits on the premise of ensuring sustainable development, how to coordinate the relationship between ethnic traditional sports and tourism has become an urgent problem to be solved.

Compared with foreign countries, the research on the development of sports tourism in China started late. In recent years, the research on sports tourism in China has developed rapidly, but there is still a certain gap compared with foreign countries. Generally speaking, the research on the basic theory of sports tourism in China is scattered and not thorough enough. Pu et al. believe that sports tourism is all forms of tourism activities related to sports activities [1], which are active or passive, accidental or spontaneous, and leaving home or workplace for non-commercial purposes. Sánchez-Martín et al. believe that sports tourism and tourism in general have both similarities and characteristics [2]. Huang et al. think that sports tourism, as a new field produced by the cross-penetration of tourism industry and sports industry, is a new form of tourism based on sports resources, which attracts people to participate in and feel sports activities and natural interests, a special leisure lifestyle combining sports and tourism, and an important part of sports industry [3]. Szromek et al. put forward the theory of internal and external driving forces of sports tourism marketing. It is believed that the marketing of sports tourism should have horizontal and vertical horizontal and vertical relations, and on the basis of horizontal and vertical relations and professional and correct goals, stable development institutions and positive and effective behaviors should be established [4]. Chen et al. proposed to formulate the intangible ethnic traditional sports cultural heritage protection plan [5], put forward innovative ideas to promote the intangible cultural heritage, that is, the specific measures for the industrialization development of ethnic traditional sports, and then strengthened the protection of intangible cultural heritage and the rational development of ethnic traditional sports.

As an important consumption mode of tourist citizens, it is an important symbol of the living standard of residents in a country or region. Developing tourism and promoting the development of tourism has become a pillar of social economy. The external conditions of sports activities actually refer to STRs (sports tourism resources), especially its natural resources, which is an important factor to attract people to participate in sports tourism activities and also a symbol different from ordinary daily sports activities and sports tourism activities. As one of many algorithms of machine learning, RF (random forest) algorithm can achieve high-precision classification and evaluation of large-scale unknown data by training and learning small-scale known samples. Based on this, this paper uses RF algorithm to study the development of ethnic STR. Through the research on the development and operation of ethnic STR, on the basis of previous studies on STR, we put forward our own opinions and ideas on the development of ethnic STR, so as to make some contributions to the development of China's sports industry and ethnic STR.

2. Related Work

2.1. STR Development Research. As sports tourism has an important influence on the economic development of a region or a country and the protection of local traditional social culture, sports tourism has gradually become the focus of academic and industry attention. To study the development strategy of sports tourism, we must first define the concept of sports tourism scientifically.

Sports tourism is an outbound activity of leaving the circle of life for noncommercial purposes to watch or participate in a sporting event. [6]. Hao et al. believe that sports tourism is essentially sports event tourism, which focuses on the influence and management of events [7]. Khandaker et al. believe that ethnic traditional sports refer to some national folk traditional sports and recreational activities as the predecessor of modern sports. China's ethnic traditional sports should include Han traditional sports and minority traditional sports [8]. Koval et al. started with the natural resources, human resources, and school education of ethnic sports and pointed out that the research on the development of ethnic sports should be combined with the integration of resources, so as to create a broader world for the development of ethnic sports resources in China [9].

Traditional ethnic minority sports is a field of traditional sports studies, focusing on the sports activities with ethnic characteristics that have been formed in the process of production, practice, and life of all ethnic groups for thousands of years. This kind of sports activity is presented to the world in a certain form of sports events. Livieris et al. believe that STR of ethnic minorities is a kind of exploitable resource form based on the basic points of ethnic minorities, which has obvious ethnic characteristics and the attributes of sports. It embodies the unique advantages of ethnic resources, natural resources, human resources, location resources, and sports resources [10]. Gao et al. discussed the research background of ethnic minority sports from the perspective of development and put forward their own opinions and feelings on the further research of ethnic minority sports. It is emphasized that ethnic minority sports under the background of the development of modern society is a new field worth studying [11].

2.2. Research Status of RF Algorithm. RF algorithm is an ensemble learning algorithm based on DT (decision tree), which has high prediction accuracy, good tolerance for outliers and noise, and is not prone to overfitting. It is used in a wide range of disciplines.

Mallinis et al. used NCL (neighborhood cleaning rule) technology to process the unbalanced data when using RF algorithm and then applied RF algorithm to improve the classification accuracy [12]. Ma et al. proposed a new oversampling method [13] to solve the problem of poor classification performance of RF algorithm on unbalanced datasets. Liu et al. sorted the DTs in the RF model in descending order according to AUC (area under curve) value, selected the DTs with high AUC value, then calculated the similarity of these DTs, and generated the similarity matrix. After clustering, the highest AUC value in each class was selected to form RF model [14]. Ghorbanian et al. applied RF algorithm to recommendation system and proposed a multi-dimensional context-aware recommendation method based on improved RF algorithm. Experimental results show that this method can reduce the average absolute error and root mean square error [15]. Although the application of this algorithm is more and more extensive and its function is more and more powerful, the random forest algorithm is not perfect, and there is still some room for improving the performance of the algorithm in dealing with different datasets and the classification accuracy of the algorithm. Therefore, this paper proposes a new algorithm to reduce the number of continuous variables, which can provide a concise dataset for C4.5 algorithm, thus improving the execution efficiency of C4.5 algorithm.

Zhang et al. proved the most direct simplified version of the original RF, that is, when selecting split features and split values, it randomly selects one of all features as split feature, and at the same time, it randomly selects one value from the selected feature value as split value [16]. Bao et al. put forward another version with consistency which is very close to the original RF. The difference lies in the selection of splitting values. They first randomly selected a subset of all the eigenvalues and then searched for the optimal splitting on this subset [17]. Zhang et al. further put forward the deep neural decision forest. They connected RF to the representation learning process of the deep neural network and turned the split function into a random decision function, so that it could update the parameters of the whole network through backpropagation [18]. He et al. also proposed a depth structure based entirely on RF [19]. They regarded RF as a node in the depth neural network. Select the feature with the best Gini value as the segmentation. In addition, they used the sliding window to scan the original features to generate new feature vectors, which were used as the input of the deep forest. It can be considered that the original features were transformed before being built. Sun et al. proposed a bi-objective elite differential evolution algorithm for multi-valued logic networks [20]. Faccioli et al. proposed a multi-objective evolutionary algorithm for protein structure prediction [21].

3. Research Method

3.1. Multi-Objective Optimization of Ethnic Sports Tourism Service Combination. STR development should not only consider the difficulty of resource development and the economic, social, and ecological benefits that resource development can receive but also consider the location and traffic conditions of resource location. According to the location and traffic conditions, STR development is generally divided into resource-based, customer-based, and resource-customer-based. If ethnic minority sports want to get rid of the previous traditional development model, they must integrate ethnic minority sports resources into the market and stand the test and choice of the market. Therefore, only by combining the ethnic minority sports cultural resources with the unique natural scenery and rich ethnic culture in minority areas can the sports resources in minority areas continue to develop and reflect the profound charm of traditional sports resources.

The resources of minority traditional sports tourism include natural resources, and much are humanistic resources. The development of traditional sports of ethnic minorities will certainly promote the development of various industries and realize the common development between tourist areas, local governments, and travel agencies. Therefore, in the process of developing the traditional STR of ethnic minorities, we must respect nature, love nature, and shape nature, so as to realize the organic unity of man and nature and achieve real harmonious coexistence. The strong appeal brought by sports tourism and entertainment activities is people's recognition and appreciation of their own national culture and history. Moreover, the joyful atmosphere, singing and dancing, and the warm atmosphere of unity in the festival enhance the consistency and coordination of the people within the clan.

In fact, the structure of multi-objective optimization problem is similar to that of single-objective optimization problem, which mainly includes objective function, constraints, and decision variables [20]. The difference is that the multi-objective optimization problem contains multiple objective functions and needs to be optimized simultaneously. A multi-objective optimization problem contains two or more objectives that need to be optimized at the same time, and the final result is a set of compromise solutions, the so-called Pareto optimal solution set. When designing algorithms, we should consider parallel computing, which mainly includes parallel computing when optimizing goals, parallel computing of fitness, and so on.

Tourism service composition is different from traditional web service composition, and tourism service composition has certain particularity. The final solution of the multi-objective service composition optimization model is not the only solution but a compromise of time, cost, quality of service, and other optimization objectives in the execution process. In this paper, the multi-objective problem is transformed into a single-objective problem. Firstly, by simplifying some constraints, a set of Pareto optimal solutions that can meet the constraints is obtained. Then, on this basis, these optimal solutions are sorted to get the most satisfactory combination for users.

The best results that users want are the shortest time, the lowest cost, and the best service quality. For convenience, this paper defines time d, cost p, and quality of service s as a function, as follows:

$$\begin{aligned} d &= f\left(x_1, x_2, \cdots, x_n\right), \\ p &= g\left(x_1, x_2, \cdots, x_n\right), \\ s &= h\left(x_1, x_2, \cdots, x_n\right), \end{aligned} \tag{1}$$

where $x_i\,(i = 1, 2, \cdots, n)$ is the parameter attribute of d, p, s.

0-1 variable x_{ij} is introduced. When $x_{ij} = 1$, it means that the jth alternative to realize the ith tourism service element is selected in the product design; otherwise, it is 0. Thus, the total cost function of product design can be obtained:

$$C = \sum_{i=1}^{6} \sum_{j=1}^{m} c_{ij} x_{ij} + c^*, \tag{2}$$

where c^* means the fixed cost of tourism product design, including staff salaries, research expenses, and so on. It is used to indicate the maximum cost that travel agencies can accept when designing travel products.

Tourism products are different from other service products. Optimizing the design of tourism products will inevitably involve the development of tourist attractions, which will inevitably have a certain impact on the ecological environment of tourist attractions. Let x be the eigenvector, and the final optimization problem can be expressed as

$$\begin{cases} \mathrm{Min}\left(\displaystyle\sum_{i=1}^{n} w_i x_i + c \sum_{i=1}^{n} w_i^2\right), \\ s.t. \displaystyle\sum_{i=1}^{n} \end{cases} w_i = 1. \tag{3}$$

Among them, the first term of the optimization objective is the above optimization equation, the second term is the regularization term of the parameters, and c is the balance parameter between them. Because we simplify the objective function to the linear combination of features, the constraint equation requires the sum of weights to be 1.

In the RF algorithm, C4.5 algorithm adopts the binary separation and dispersion method, which makes the efficiency of the algorithm have a great relationship with the value of continuous variables. Therefore, if we can effectively reduce the dataset and reduce the value of continuous variables, we can improve the execution efficiency of RF algorithm. In this paper, a new CVD (continuous variable disclusion) algorithm based on x^2correction is designed, and its calculation process is as follows:

(1) Calculate the number k of decision attributes in two adjacent intervals of a certain attribute value.

(2) Calculate the theoretical number E_{ij}, and the formula is

$$E_{ij} = \frac{R_i \cdot C_j}{N}, \tag{4}$$

where R_i is the number of samples in i interval; C_j is the number of class $j$samples; and N is the total number of samples in two adjacent intervals.

(3) The sequence of merging intervals is determined by the value of D, and its formula is as follows:

$$D = \frac{x_a^2 - x^2}{\sqrt{2v}}, \tag{5}$$

where v is the degree of freedom.

(4) Select the interval with the smallest D value to merge.

After all the continuous attribute variables in the dataset are subjected to the above steps, the reduction process of the dataset is completed, that is, the discretization of continuous variables is realized. The flowchart of CVD algorithm is as follows (Figure 1).

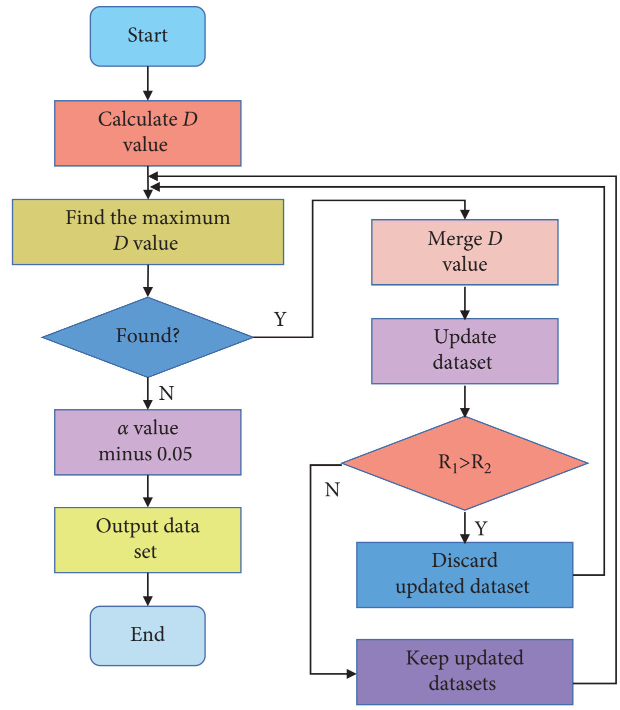


Figure 1: Flowchart of CVD algorithm program.

From the perspective of travel agencies or enterprises, we hope to design and optimize tourism products at the lowest cost and bring the greatest satisfaction to tourists. However, in real life, it is often difficult to achieve high-quality tourists' satisfaction and low cost at the same time, which requires enterprises or travel agencies to give weight to both according to the actual situation and make a scheme choice according to the final actual results.

3.2. STR Recommendation Based on RF. Every nation has its own unique traditional festivals. At present, these distinctive ethnic traditional festivals have become important tourism resources. Ethnic minorities will hold all kinds of traditional ethnic arts and sports activities when their traditional festivals come. However, the emergence of a new tourism model is bound to require the support of a large number of related manpower. Vigorously developing the research of ethnic traditional STR can enable more local people to return to their good jobs and show their local ethnic characteristic culture. As the market propagandist of traditional sports tourism of ethnic minorities, it is more important to promote themselves to the greatest extent. It is necessary to publicize the traditional sports tourism of ethnic minorities in various aspects and in all directions. Take various measures to encourage and mobilize citizens to participate in traditional sports of ethnic minorities, cultivate their habit of participating in traditional sports of ethnic minorities, and then cultivate the tourism market of traditional sports of ethnic minorities.

At present, the level of social technology and the changing trend have relatively little direct impact on the ethnic traditional sports tourism, but the indirect impact covers all aspects of the ethnic traditional sports tourism. Technological changes and breakthroughs have made the supporting industries of the ethnic traditional sports tourism more and more prosperous. In recent years, China's economy has been developing continuously, and residents' income level has been improved to a great extent. However, compared with developed countries, it is still at a lower level, and the proportion of income that can be used for traditional ethnic sports tourism is not very large, and it is still low compared with western developed countries. By developing and operating the ethnic traditional STR and forming the tourism industry with the ethnic traditional sports culture as the core, it can help more local minority people to engage in sports culture with ethnic characteristics and solve their employment difficulties.

With the popularization and development of network technology, the infinite potential of ethnic traditional sports tourism can be tapped to the maximum extent by the east wind of the network. The perfection of the construction of tourism platform is of great significance to the development of tourism. The recommendation system is used to recommend items and information that users may be interested in, which improves the information pushing efficiency of information providers and the advertising efficiency of commodity providers, so that specific information can be delivered to specific people, thus achieving a win-win situation for businesses and users.

The DT in the recommendation system can be used in the model-based method. The user can construct a DT after evaluating only two items, and the characteristics of each item are used to build a model to explain the user's rating. The DT can only be used to simulate a special part of the system.

The traditional CF algorithm calculates the similarity between this user and other users by using the paired evaluation scores of common scoring items [19] but ignores the scores that only one user evaluates. In this chapter, DCF (distributed collaborative filtering) algorithm is proposed. In addition, in order to significantly improve the efficiency, we provide the map reduce-based distributed version of this algorithm running on Hadoop. The framework of the algorithm Map_DCF is shown in Figure 2.

In the process of recommendation, the problem of personalized matching with users can be dealt with by means of knowledge reasoning. Therefore, it is necessary to establish rules that conform to reasoning according to data sources. Therefore, a general user data ontology needs to be established here.

The feedback factor N of ethnic sports tourism users is introduced, and the weight μ of the feedback factor is considered at the same time, and then the final total matching degree value M is calculated by

$$M = D(1-\mu) + N\mu. \tag{6}$$

CF recommendation algorithm based on proximity relation can be divided into UCF (user-based CF) algorithm and ICF (item-based CF) algorithm. The basic idea of the former is that similar users have common preferences, while the latter assumes that users will like projects similar to those they liked in the past, so ICF recommends projects similar to those they liked in the past.

In the UCF algorithm, the target user a scores the item i as follows:

$$\hat{r}_{ai} = \overline{r}_a + \frac{\sum_{u\in N(a)} (r_{ui} - \overline{r}_u) sim(a,u)}{\sum_{u\in N(a)} |sim(a,u)|}, \tag{7}$$

where $N(a)$ represents the k users who are most similar to the target users among the users who have scored item i; r_{ui}indicates the rating of user u on item i; $\overline{r}_a, \overline{r}_u$ represents the average value of user's a, u history score; and $sim(a,u)$ indicates the similarity between two users.

For the problem of dividing a sample into a certain category, that is, discrete variables, the DT in RF generally uses Gini value as the evaluation standard. It is defined as

$$Gini = 1 - \sum P(i)^2. \tag{8}$$

Among them, $P(i)$ is obtained according to the sample classification ratio, that is, the ratio of i classification to the total sample.

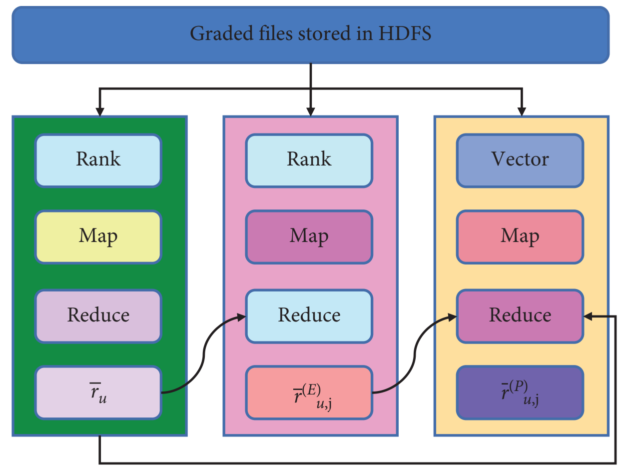


Figure 2: Map_DCF algorithm principle block diagram.

RF is a classifier that contains multiple DTs as the basis, and the final prediction result depends on the votes cast by the prediction results of the DTs, with the class with the most votes as the final prediction result. The advantage of the classifier is that its results are more accurate than those of the individual classifiers. The accuracy of RF is equal to AdaBoost, but it is more robust to errors and outliers.

With the increase of the number of trees in the forest, the generalization error of the forest gradually converges. Therefore, RF can avoid the problem of overfitting. For the conversations without purchasing behavior, further bootstrap sampling is carried out, and then the class is weighted by using the idea of weighted RF. The corresponding weight item is designed for one class, and a few classes are given more weight, so that it is valued.

Specific algorithm flow is as follows:

(1) Randomly extract the same amount of data to replace it, so that the samples in the majority class and the minority class form the sample data subset D_i^*, and ensure that the data of each class are included in the subset.

(2) Design a corresponding weight item for one class and assign a larger weight to a few classes, so that they are valued and assigned the corresponding weight value $w_1, \cdots, w_c$.

(3) The division of each node in DT is determined by the weighted majority vote, and the weighted vote of a class at the end node is multiplied by the number of data contained in the class. Get another prediction function sequence $h_1^*, \cdots, w_n^*$.

(4) Repeat the above steps k times, integrate the results of two random trees, and use the prediction function to vote on the classification problem to make the final prediction.

4. Result Analysis

For the sake of generality, Iris, Credit, and Blood datasets of UCI machine learning database are selected as experimental data. In the process of discretization, the algorithm has to obtain the chi-square quantiles. When it is implemented, the chi-square quantiles are all obtained by software R3.0.2, and the RF algorithm is implemented by using the language package random Forest4.6-6 of software R3.0.2.

The continuous variables in Iris are discretized by three algorithms, respectively, and the data reduction effects are shown in the following (Tables 1–3).

In the dataset, different reduction algorithms have good performance and achieve the purpose of discretization of continuous variables. From the reduction effect, the reduction rate of CVD algorithm is obviously higher than that of Ref [15] algorithm. Therefore, from the degree of information reduction, CVD algorithm is outstanding, which shows that CVD algorithm can reduce redundant information in datasets on a large scale.

Using the above datasets, the initial dataset and the datasets reduced by three different algorithms are used, respectively, and 200 experiments are carried out under the condition of keeping the parameters of RF algorithm consistent. The mean values of OOB (out-of-bag) estimation of RF on different datasets are compared, and the discretization performance of CVD algorithm is analyzed. The experimental data are as follows (Figure 3).

It can be seen that in the RF algorithm, the OOB estimation value of the algorithm is reduced after using CVD algorithm to reduce the dataset, which shows that the performance of the reduced dataset not only does not decrease but also improves the accuracy of RF. From the measurement standard of continuous variable discretization, CVD algorithm not only improves the dataset reduction rate but also keeps the information of the original dataset.

The construction of RF is the core of model building. The number of DT directly affects the operation speed and classification effect of RF classification algorithm, so the number of DT is very important for modeling [13]. As shown in Figure 4, with the increase of DT trees in RF, the accuracy of the model increases, but the recall rate does not change significantly.

It can be seen that, for this dataset, when the number of DT trees reaches 10, with the increase of DT trees, the classification accuracy rate of RF does not continuously increase, but the time cost will increase because of the increase of DT trees. It is ideal to select 10 DTs in the RF classification algorithm, which makes the recommendation system get accurate results quickly and meet the time limit of the recommendation system.

In this experiment, the Map_DCF algorithm proposed in this paper is compared with UCF, ICF, and Ref [13]. Figure 5 shows the change of online recommendation time of each algorithm with the number of clusters.

UCF and ICF algorithms have no clustering process, so the online recommendation duration is not affected by the change of cluster number. The online recommendation time of Map_DCF is basically below 100 s, which is much lower than that of other recommendation algorithms. As the number of clusters increases, the online recommendation time of Ref [13] decreases because the search range of the nearest users is narrowed. The online recommendation time of Map_DCF hardly changes with the number of clusters because there is no process of searching for the nearest user (or project) in online recommendation of Map_DCF, and it only needs to be recommended according to the rules obtained in advance by the RF model, so the online recommendation time is shorter than that of other models in the long run.

All features in an interval have a high correlation, and all features in the remaining intervals have a low correlation. When selecting features, on the one hand, the important features are selected according to Gini index; on the other hand, the features in two interval features are selected according to a certain proportion, which is partially random, so as to balance the strength and correlation of features. Repeat this process to build the whole forest. It can be found that, as shown in Figure 6, the effect of the model is improved.

The time of RF modeling and prediction is fast, and the time cost of feature evaluation by chi-square test is very small. Therefore, this method can ensure the running efficiency of the model. In addition, through the correlation analysis of features, the features are divided into intervals to improve the process of feature selection and get the effect of RF model optimization.

After resampling data, a lot of information will be lost. Therefore, on the basis of using the idea of balanced RF, we sample the conversations without purchasing behavior, and then use the idea of weighted RF to weight the classes, design the corresponding weight items for each class, and assign more weight to a few classes, so that they are valued. By weighting, the accuracy of the model can be guaranteed, and at the same time, the difference between DTs can be guaranteed. Combining the DT of the two to obtain the balanced weighted RF, the result is shown in Figure 7, and the result $F1$ is improved.

From the experimental results, it can be seen that after using balanced weighted RF, $F1$ index has been greatly improved, which reflects that threshold adjustment and class weighting method have obviously improved the unbalanced classification problem. Combined with the balance and weighting method of RF, the improvement of experimental results verifies the effectiveness of the improved method for unbalanced classification and also shows the importance of taking into account the retention and non-redundancy of data information and ensuring that a few classes are valued at the same time.

The difference between ethnic sports tourism and other tourism modes is that tourists can not only visit mountains and rivers, enjoy beautiful scenery, and experience traditional ethnic culture but also participate in them and feel the charm of sports. Scientifically set up ethnic sports activities that can enable tourists to participate in them, highlight the experience-oriented tourism mode, and form a cultural tourism scenic spot with "ethnic culture as its connotation and leisure experience as its feature." Only by balancing viewing and participation in this way can tourists feel happy physically and mentally when they play in the scenic spot and achieve the sublimation of their bodies and souls. This makes ethnic sports tourism a unique cultural and tourism approach with market competitiveness among many tourism resources.

Table 1: The algorithm reduces the effect table of Iris dataset.

Algorithm	Variable 1	Variable 2	Variable 3	Reduction rate (%)
Initial dataset	33	40	23	—
Ref [15]	30	40	23	4.86
Ref [17]	20	35	18	17.62
Algorithm in this paper	16	32	14	33.69

Table 2: The algorithm simplifies the effect table of Credit dataset.

Algorithm	Variable 1	Variable 2	Variable 3	Reduction rate (%)
Initial dataset	30	35	34	—
Ref [15]	28	32	34	3.36
Ref [17]	28	32	31	3.41
Algorithm in this paper	24	25	26	11.67

Table 3: The algorithm simplifies the effect table of Blood dataset.

Algorithm	Variable 1	Variable 2	Variable 3	Reduction rate (%)
Initial dataset	68	91	26	—
Ref [15]	17	13	7	91.33
Ref [17]	17	13	7	90.24
Algorithm in this paper	15	5	3	96.28

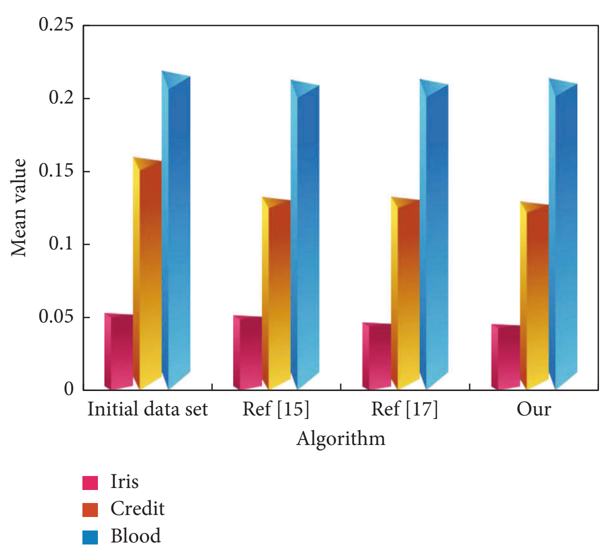


Figure 3: Use of the average value of OOB estimation after processing various datasets by RF.

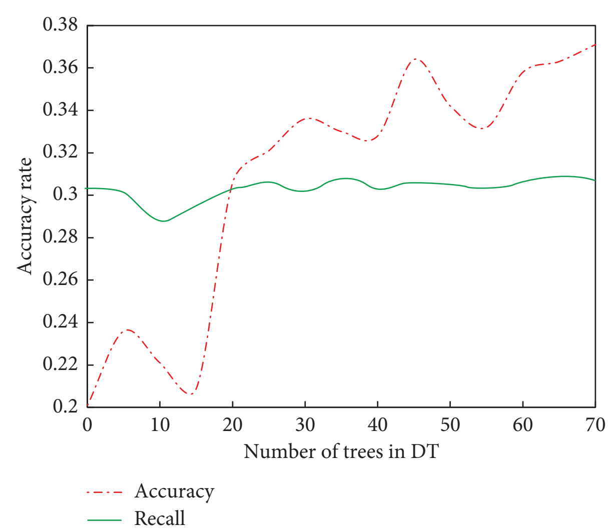


Figure 4: Influence of DT tree number on accuracy and recall rate.

Nowadays, the mainstream direction of the development of the world tourism market is from single sightseeing to diversified and thematic tourism. When developing ethnic traditional sports tourism projects, the relevant functional departments should keep the original ethnic style, dig deep into the connotation of ethnic traditional sports, gradually transform into modern people's tourism cultural concepts, aesthetic characteristics, and value orientation, develop towards the trend of international sports tourism integration, and speed up its socialization and industrialization development process in combination with the cultural characteristics of ethnic areas.

To meet the needs of tourism development and ethnic minority sports work, we should focus on building three teams: first, professional teams who know sports and are familiar with ethnic customs and culture, so as to prepare for the deep excavation of STR development of ethnic minorities. Second, it will manage a pioneering management team, promote the combination of ethnic minority sports resources and tourism, and plan excellent projects that can enter the market and manage related industries. Third, the all-round team of waiters, in the service place, reflects the quality level of service and leaves a good impression on tourists.

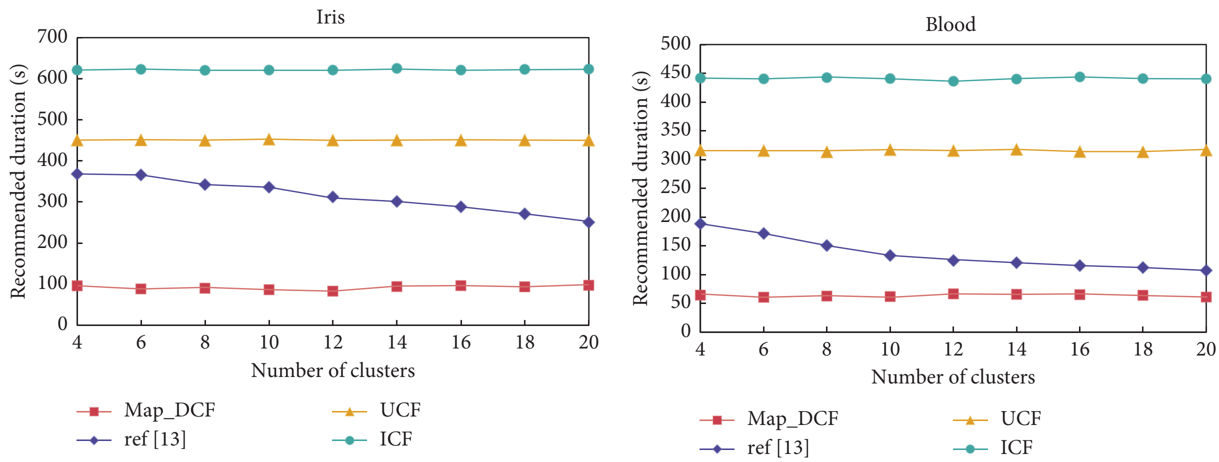


Figure 5: Comparison of recommended online time.

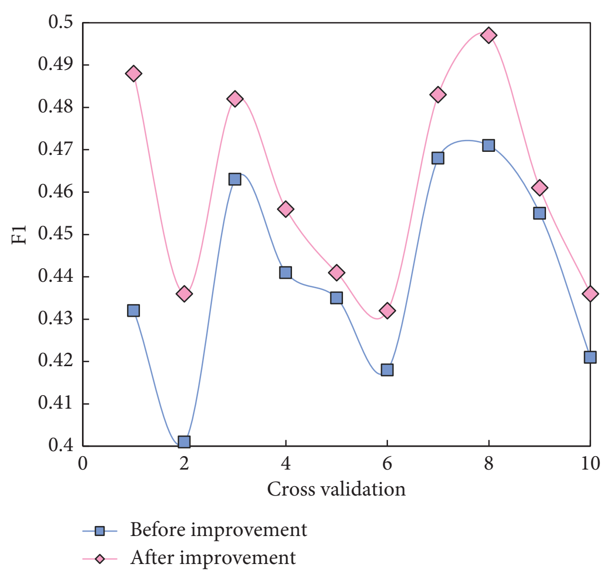


Figure 6: Comparison of results before and after feature selection and improvement.

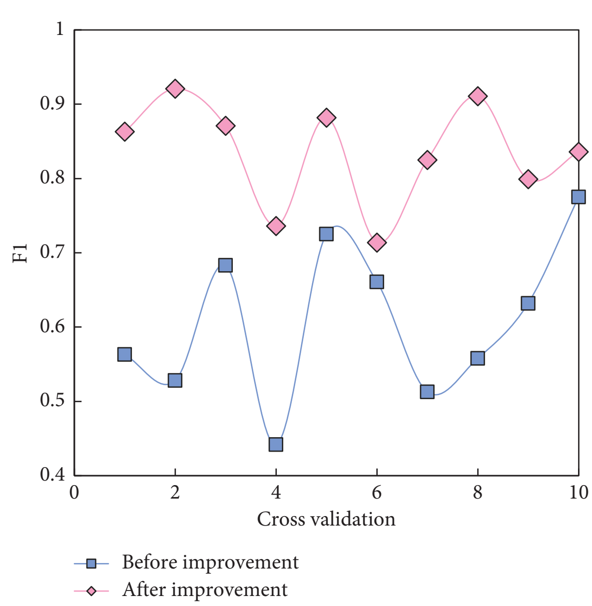


Figure 7: Comparison of effect before and after balanced RF.

5. Conclusion

The key to the development of ethnic traditional sports lies in the formation of a healthy cultural atmosphere, the establishment of sustainable development strategy, and the construction of internal development motivation. The integrated development of ethnic sports tourism not only enriches the tourism product system and promotes the transformation and upgrading of tourism but also revitalizes various resources and promotes the quality and efficiency of new industries [22]. Combining the characteristics of tourism services, this paper puts forward a multi-objective tourism service composition method. Finally, experiments show that our proposed method has high accuracy and efficiency in the multi-objective optimization of tourism service composition. Aiming at the defects of today's tourism recommendation system, an ethnic sports tourism information feedback recommendation algorithm based on RF is proposed. The algorithm reduces the dimensions of user and item vectors through clustering process, and the system only needs to make score prediction according to the RF model constructed when offline, without looking for the nearest user or item, thus greatly improving the recommendation efficiency. Through the set of balanced RF and weighted RF, the data information defect and redundancy caused by resampling can be solved, and all rare data can be selected. At the same time, the accuracy and difference of the model can be ensured by weighting.

Under the background of big data, it is a hot topic in the field of RF algorithm optimization in the future. How to realize distributed and concurrent computing of RF algorithm will be an important research direction.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This research was supported by Ministry of Education Humanities and Social Sciences Research Youth Fund Project (21XJC890002) and Guangxi Philosophical and Social Science Planning Research Topics (18CTY001).

References

[1] J. Pu and Y. Meng, "Application of visual design in the development of marine tourist resources," *Journal of Coastal Research*, vol. 98, no. sp1, p. 125, 2019.

[2] J. M. Sánchez-Martín, M. Sánchez-Rivero, and J. I. Rengifo-Gallego, "Water as a tourist resource in extremadura: assessment of its attraction capacity and approximation to the tourist profile," *Sustainability*, vol. 12, no. 4, p. 1659, 2020.

[3] W. Huang and C. S. Huang, "Discussion on engineering geological conditions and geological tourism advantage resources of coastal landscape highway in guangdong province," *Ground Water*, vol. 8, no. 10, p. 06, 2019.

[4] A. R. Szromek and K. Wybrańczyk, "Proposal of value for customer of spas: expectations of spa patients and tourist in polish spas," *Sustainability*, vol. 11, no. 13, p. 3598, 2019.

[5] H. Chen, S. Wang, H. Guo et al., "Study of marine debris around a tourist city in East China: implication for waste management," *Science of the Total Environment*, vol. 676, no. 1, pp. 278–289, 2019.

[6] Y. C. Lee and P. T. Liao, "The effect of tourism on teleconnected ecosystem services and urban sustainability: an emergy approach," *Ecological Modelling*, vol. 439, no. 1, Article ID 109343, 2021.

[7] J. Hao, H. Liu, J. Chen, and F. Sun, "Study on city-level optimization of tourism industry spatial organization nodes and organization mode for tourist destinations," *Complexity*, vol. 2020, no. 12, pp. 1–13, 2020.

[8] M. U. Khandaker, K. Asaduzzaman, A. F. B. Sulaiman, D. Bradley, and M. O. Isinkaye, "Elevated concentrations of naturally occurring radionuclides in heavy mineral-rich beach sands of Langkawi Island, Malaysia," *Marine Pollution Bulletin*, vol. 127, no. 7, pp. 654–663, 2018.

[9] G. M. Danilova-Volkovskaya, L. N. Koval, A. P. Babchenko, and N. Volkovskiy, "The tourist directions resources and sustainable development competitive advantage marketing model in the north caucasus," *Materials Science Forum*, vol. 931, no. 2, pp. 1182–1186, 2018.

[10] I. E. Livieris, E. G. Pintelas, T. Kotsilieris, S. Stavroyiannis, and P. Pintelas, "Weight-constrained neural networks in forecasting tourist volumes: a case study," *Electronics Times*, vol. 8, no. 9, p. 1005, 2019.

[11] W. Gao, Q. Zhang, Z. Lu, D. Wu, and X. Du, "Modelling and application of fuzzy adaptive minimum spanning tree in tourism agglomeration area division," *Knowledge-Based Systems*, vol. 143, no. 1, pp. 317–326, 2018.

[12] G. Mallinis, I. Chrysafis, G. Korakis, E. Pana, and A. P. Kyriazopoulos, "A random forest modelling procedure for a multi-sensor assessment of tree species diversity," *Remote Sensing*, vol. 12, no. 7, p. 1210, 2020.

[13] R. Ma, J. Ban, Q. Wang et al., "Random forest model based fine scale spatiotemporal O3 trends in the beijing-tianjin-hebei region in China, 2010 to 2017," *Environmental Pollution (Amsterdam, Netherlands)*, vol. 276, no. 10159, Article ID 116635.

[14] S. Liu, D. Mao, T. Xue et al., "A data-driven approach for online inter-area oscillatory stability assessment of power systems based on random bits forest considering feature redundancy," *Energies*, vol. 14, no. 6, p. 1641, 2021.

[15] A. Ghorbanian, S. Zaghian, R. M. Asiyabi, M. Amani, A. Mohammadzadeh, and S. Jamali, "Mangrove ecosystem mapping using sentinel-1 and sentinel-2 satellite images and random forest algorithm in google earth engine," *Remote Sensing*, vol. 13, no. 13, p. 2565, 2021.

[16] X. Zhang, J. Xu, Y. Chen, K. Xu, and D. Wang, "Coastal wetland classification with GF-3 polarimetric SAR imagery by using object-oriented random forest algorithm," *Sensors*, vol. 21, no. 10, p. 3395, 2021.

[17] S. Bao, H. Y. Pan, W. Zheng et al., "Multicenter analysis and a rapid screening model to predict early novel coronavirus pneumonia using a random forest algorithm," *Medicine*, vol. 100, no. 24, Article ID 26279, 2021.

[18] C. Zhang, J. Ren, F. Liu, X. Li, and S. Liu, "Three-way selection random forest algorithm based on decision boundary entropy," *Applied Intelligence*, no. 2, pp. 1–14, 2022.

[19] S. He, W. Chen, H. Liu et al., "Gene pathogenicity prediction of Mendelian diseases via the random forest algorithm," *Human Genetics*, vol. 138, no. 6, pp. 673–679, 2019.

[20] J. Sun, S. Gao, H. Dai, J. Cheng, M. Zhou, and J. Wang, "Bi-Objective elite differential evolution algorithm for multi-valued logic networks," *IEEE Transactions on Cybernetics*, vol. 50, no. 1, pp. 233–246, 2020.

[21] R. Faccioli, I. Silva, and L. O. Bortot, "A mono-objective evolutionary algorithm for Protein Structure Prediction in structural and energetic contexts," *IEEE Congress on Evolutionary Computation. IEEE*, vol. 13, no. 61, pp. 3–9, 2012.

[22] C. Nie, "Study on the strategy of developing ethnic sports tourism in Guangxi under the perspective of "one belt and one road"," *Journal of Guangxi College of Education*, no. 06, pp. 88–94, 2020.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 2497365, 11 pages
https://doi.org/10.1155/2022/2497365

Research Article

Explore on Online English Listening Evaluation System Using the Genetic Algorithm

Shiyu Yin, Jin Xu, and Liying Ge

School of Foreign Languages and Literature, Tianjin University, Tianjin 300350, China

Correspondence should be addressed to Jin Xu; amandaxujin@tju.edu.cn

Received 25 July 2022; Revised 20 August 2022; Accepted 3 September 2022; Published 22 September 2022

Academic Editor: Jiafu Su

Testing is a good teaching tool. The online testing system can realize automatic aggregation, scoring, statistics, and analysis of test questions, which saves a lot of time and money for the organizers and participants, and is an effective means to promote the modernization of education in China. The purpose of implementing English intelligence is to allow teachers and students to make better use of their strengths in the English learning process and to improve English teaching based on the results of quantitative analysis of intelligence, which will play a positive role in promoting quality language education in secondary schools, thus truly reducing the burden on teachers and students. The genetic algorithm is an intelligent algorithm based on natural selection and genetic variation in the world of bionic organisms. It has the remarkable advantages of simplicity and generality, robustness, parallel processing, efficiency, and practicality, and has been widely used in automation and other fields with good results. The method can effectively overcome the traditional problems of slow grouping speed, low success rate, and poor grouping quality. The system is scientific, reasonable, and practical and can meet the needs of users to the maximum extent.

1. Introduction

1.1. The Use of Genetic Algorithms in Computers. With the rapid development of computer technology and network technology, people have become more aware of this [1]. Most colleges and universities are connected to the Internet and have established networks on their campuses. The hardware equipment is getting better and better. Internet online examination has the characteristics of cross-territory, working anytime, working anywhere, and having a wide range of users, which is an important way to promote the modernization of education in China. It has become an important way to promote the modernization process of education in China [2]. It has become an important tool in practice, testing, evaluation, and analysis. In teaching, examination is a very important part. It is an important method to measure and evaluate the quality of education. It should also be integrated with modern information technology. Traditional essay-based quizzes, manual questioning, answering, manual reading, manual statistical results, and paper analysis can no longer meet the requirements of modern education. Online examination system can realize automatic question issuance, online examination, remote monitoring, automatic reading, and automatic statistical analysis, which optimizes the examination mode, greatly saves the time and cost of examination organizers and participants, and has important practical significance for improving the efficiency of school operation and promoting education modernization. These years, there are already some electronic testing systems that use computer and Internet technologies that have come into existence. However, all these methods have some defects as follows: on the basis of the C/S model, special clients need to be installed and it is difficult to update them [3]; only manual and randomized services are available, so the quality of the package is low and slow; the quality of group questions is poor, so the efficiency of the questions is low and the IQ is poor; only the automatic scoring function of objective questions is available, so the scoring of fill-in-the-blank and other subjective questions cannot be done automatically, and the security is low.

1.2. Concept of the Genetic Algorithm. As we know, genetic algorithms deal with binary chromosome encoding through encoding techniques that mimic the evolutionary process of binary string populations. The algorithm generates a new generation of binary strings, i.e., well-tuned segments and bits in each generation with the help of good genes from the previous generation [4]. Compared to traditional random methods, this method efficiently uses the available information to find characters with improved performance, a process similar to natural evolution, which in essence improves the chances of inheritance of its genes by screening for high quality chromosomes. In the solution space, the global optimal solution is found by iteration, as shown in Figure 1. The iteration of the population consists of three main operations: selection, hybridization, and mutation. The genetic algorithm is a kind of bionic optimization algorithm, which is a kind of knowledge obtained from human's own evolutionary process, that is, an intelligent algorithm that can imitate the continuous evolution of biological populations in nature. First, the optimization goal is approached gradually through steps such as selection, hybridization, and mutation.

2. State of the Art

2.1. Current Status of Domestic and International Research on Heritage Algorithms. Foreign countries have researched fast and efficient online examination systems, especially new distance learning software that not only achieves better interactivity and high efficiency of the system but also better functional extensions with the system by designing and implementing the latest system that enables real-time data transfer under the Internet for learners to achieve real-time on-demand and learning. As the global competition for talents has been fierce, in today's competitive society, most countries in the world are trying to develop online education to gain a dominant position in the education field so that Internet technology can be widely used in the education industry [5]. Since the birth of Internet technology, it has a series of advantages such as interactivity, sharing, wide distribution, and free and open access, which are more beneficial to the education industry and can save the cost of education; thus, today, online real-time exams are becoming more and more popular.

Most universities in China now have their own standalone or introduced online testing systems, as well as many other types of data support, talent assessment and recruitment, and education. The company has also developed its own sophisticated online testing system. Currently, in the past two years, the Ministry of Education's computer grade exams, Mandarin proficiency exams, and some professional certification exams from the Ministry of Human Resources and Social Security have been implemented online [6]. Currently, domestic online test systems are catching up with foreign development levels and conducting in-depth research in the direction of intelligence in terms of test question extraction, machine reading scope and test data mining. At present, many professional and technical examinations in the form of online web-based examinations exhibit the following characteristics, as shown in Table 1.

2.2. Application of the Genetic Algorithm. The genetic algorithm is an intelligent algorithm. In the previous section, the concept of genetic algorithm as well as the steps and key techniques of the algorithm are introduced in detail [7]. In the next paper, we use the genetic algorithm to implement automatic grouping of test libraries. There are many kinds of chromosome encoding, and binary encoding is the most common one. The biggest drawback of binary encoding is that it has a long encoding length and a large retrieval space. To address this problem, a packet-based natural number encoding method is proposed to reduce the total number of iterations and speed up the solving speed. Packetization implies that each packet represents one type of problem, and natural number encoding implies that the individuals in each problem are encoded in the form of natural numbers. For better understanding, an example is given in Table 2. This study assumes that there are three types of questions to be written: multiple-choice, fill-in-the-blank, and parsing questions, and Table 2 shows the chromosomes for the latter two topics grouped by group natural number coding.

2.3. General System Overview. The English intelligent diagnosis system is a computerized intelligent teaching, diagnosis, testing, and feedback system [8]. It is built on the basis of a sound knowledge structure system and a large question bank. After deep data mining of the question database, it uses four engines of test paper generation, assessment, diagnosis, and intelligent learning to complete a number of system functions such as thematic integrated practice, individual test and class test assessment, individual diagnosis, and class diagnosis, and provides them to users in web form. The basic block diagram of the system platform is shown in Figure 2.

The English online test management module is a key part of the English test [9]. The cloud-based English test system needs to be designed to address different test requirements and to process different test instructions in real time for different test sizes and requirements. The test system also needs to provide test resource sharing, support for cloud-based management, and other functions, and be managed by specialized technical staff to accommodate dynamic storage requirements. With the English online test management module, question writers can store all their test resources at any time and can call test resources at any time according to different test requirements. Candidates log into the system before the exam to complete the questions, and the answer time is automatically recorded and saved in real time; after the answers are completed, the relevant data is sent to the system [10]. The huge cloud storage space ensures that the stored data will not be damaged or lost.

2.4. System Structure and Functional Design. The ASP.NET online testing system is developed on a B/S basis [11]. Users can take online tests through the interaction between web browser and web server. The system is divided into three parts: system administrator, teacher, and student. The system administrator is responsible for the daily management and maintenance of the system. The main components are:

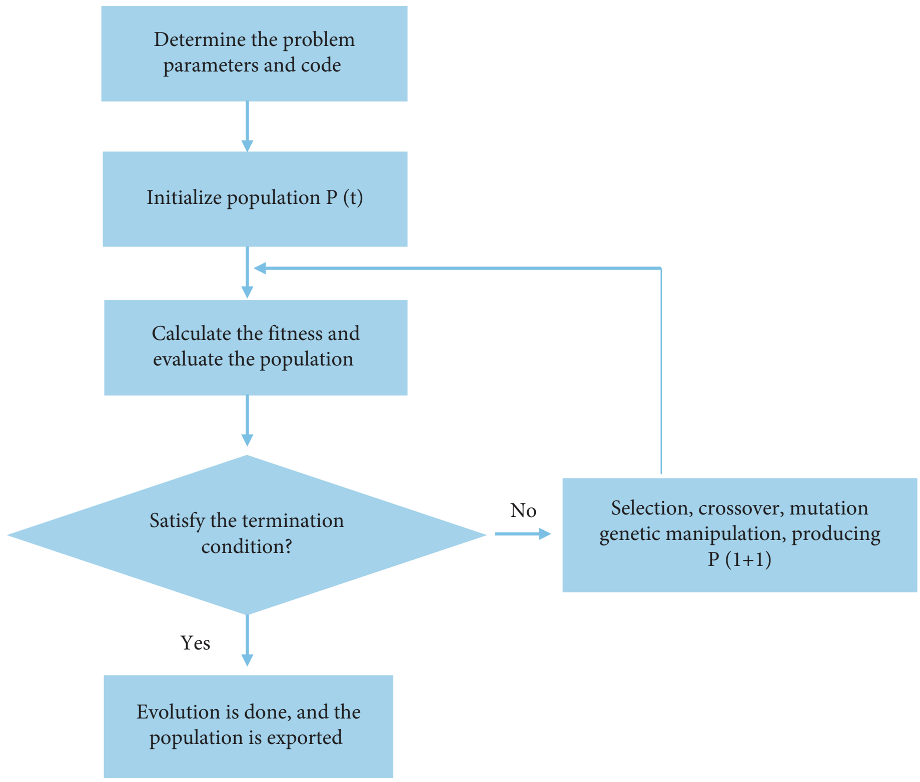


Figure 1: Schematic diagram of the operation process of genetic algorithm.

Table 1: The main types and characteristics of network testing.

Test type	Test features
Open	Online public test has unlimited location and limited time
Semiopen	Proctored exams, where candidates take online tests on time and on location

Table 2: Chromosome coding for both thesiss.

	Multiple choice questions				Fill-in-the-blank questions				Analytical questions			
Test thesis 1	20	31	1	64	24	2	68	7	5	91	37	68
Test thesis 2	4	90	67	45	56	99	35	57	74	8	10	27

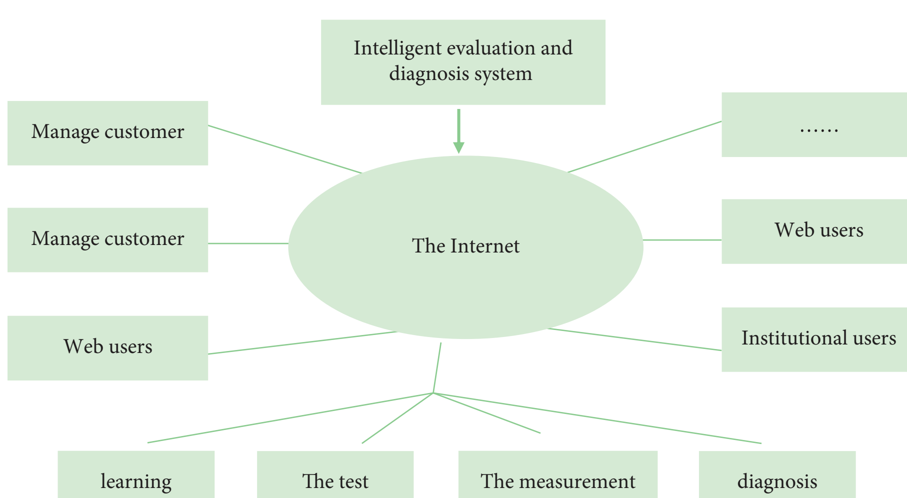


Figure 2: System basic framework diagram.

managing the database, maintaining the information of teachers and test takers, and managing the security of the system. The teachers are responsible for setting and maintaining the question bank, determining the parameters of the papers, distributing the test questions, scoring the subjective questions, and querying and analyzing the data. The students take the online quizzes and check their scores. The system consists of a user management sub-system, question bank management sub-system, essay writing sub-system, online examination sub-system, marking and scoring sub-system, score query sub-system, and statistical analysis sub-system.

The C/S model does not require intermediate links to connect the client directly to the server, so the response time is very fast; in addition, the C/S model has a friendly interface and strong transaction processing capability, and can complete a variety of complex business processes, see Figure 3. The biggest advantage of the B/S model is that it can work at any location without installing special software. As long as there is a computer with Internet access, the system can be easily expanded without any client maintenance [12]. As long as there is access to the Internet, the username and password are specified by the administrator, and it can be used. Therefore, the C/S approach is used in both test bank and score management subsystems, which places higher demands on the interactive capability and security, stability, and efficiency of the testing system. This system is a B/S system, using IE browser to log in to the system for examinations, self-tests and personal score inquiries; using IE browser, it can realize inquiries, analysis, and manual scoring of students' scores [13]. The system is based on test bank management and online testing and adopts a hybrid model architecture of B/S and C/S. The three levels of the testing system are shown in Figure 4.

3. Methodology

3.1. Intelligent Teaching Platform. Intelligent teaching platform is based on network practice technology, using computers as a teaching platform to assist teachers and students in teaching and making the teaching process more intelligent. In the teaching process, the intelligent teaching platform replaces the teaching work that teachers had to complete in the past and establishes an intellectual and collaborative partnership that can meet different course characteristics and different student groups [14]. Currently, in English teaching, the intelligent teaching platform breaks the traditional teaching method, changes the interaction between teachers and students, and realizes multimodal communication between teachers and students and between students, realizing real-time and non-real-time communication. In addition, it has many advantages [15]. First of all, it can build a speech recognition system that can help learners with speaking training. In traditional teaching, it is difficult for teachers to listen to each student's pronunciation, recitation, and meticulous evaluation of their spoken language one by one. However, the intelligent teaching platform is able to break this limitation by enabling the correction of errors and the ability to repeat the training according to the correct pronunciation. Secondly, through the smart teaching platform, targeted learning programs are designed to meet the needs and requirements of different students, thus achieving better results. It breaks through the limitations of time and energy of traditional classroom teaching, which prevented teachers from observing their classmates, much less each student in detail. With the development of society and the rapid development of science and technology, education must keep pace with the times, conform to the development of science and technology, use intelligent teaching platforms to guide students, and make the classroom more attractive, making it more efficient and nice.

In computerized online examinations, the design of the algorithm for drawing questions directly determines the efficiency and quality of the automatic set of papers. This paper focuses on a solution method that quickly and efficiently selects a set of optimal solutions, or a set of closest to optimal solutions, from a class of problems. Previous quiz systems have been implemented by random selection and inverse testing. The random selection method uses control indicators in the state space to randomly select a copy of the questions from the computer and place them in the test bank. This algorithm is simple in construction and fast in the extraction operation of a single question, but in the case of complex grouping conditions, it often leads to failure of the grouping due to local satisfaction of constraints. Therefore, this method is only applicable to the case of small question bank system with few constraints [16]. This system uses genetic algorithm as the main algorithm and has the advantages of adaptive optimization and intelligent search. It is widely used. The combined genetic algorithm with the genetic algorithmapplied the method to the neural network and obtained better results: the convergence of genetic algorithm under the condition of maintaining optimality was widely studied. After theoretical derivation and practical application, it is proved that the method has good optimization performance and convergence performance.

3.2. Ajax Technology. Ajax is not a language; it is simply a set of technologies that can be used to describe a set of techniques to improve the performance of the Internet. Ajax, whose full name is Asynchronous JavaScript, XML (XSTL), incorporates JavaScript scripting language, XHTML language, DOM documents, XML, CSS style sheets, XSTL, and several other technologies that play a role in collaboration. Ajax, Asynchronous JavaScript, XML, is a way of developing Web applications that uses client scripting to exchange data with a Web server. Ajax allows you to dynamically update pages without affecting the flow of interaction. The Ajax pattern introduces Ajax technology between the user and the server, which is like adding a layer of mechanism to the program to make it more responsive. When an action is performed on the interface, the web application pattern triggers a request for a connection to the server, which accesses the source, responds to the processing, and returns to the job execution interface. While the server is running, other client requests must wait, which wastes time and causes unreasonable resource allocation. Ajax technology uses

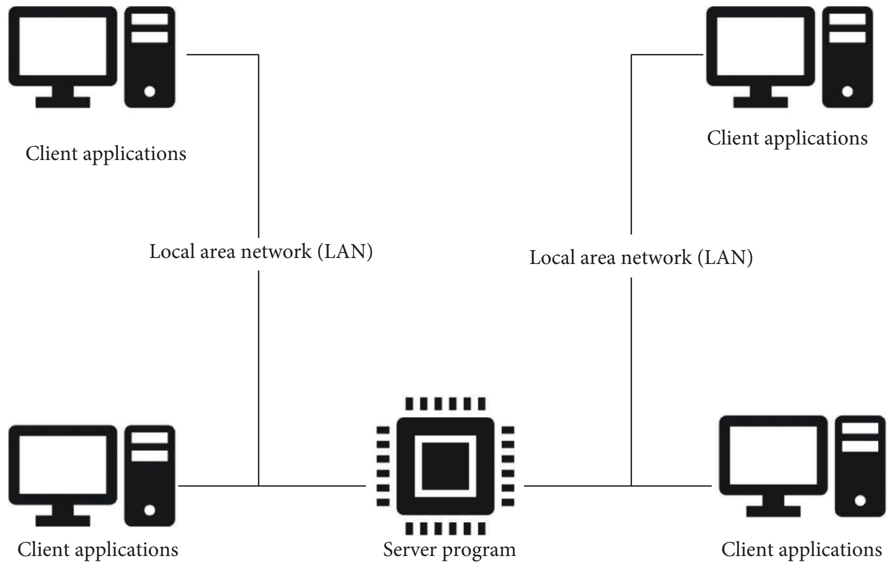


FIGURE 3: C/S mode of online examination system.

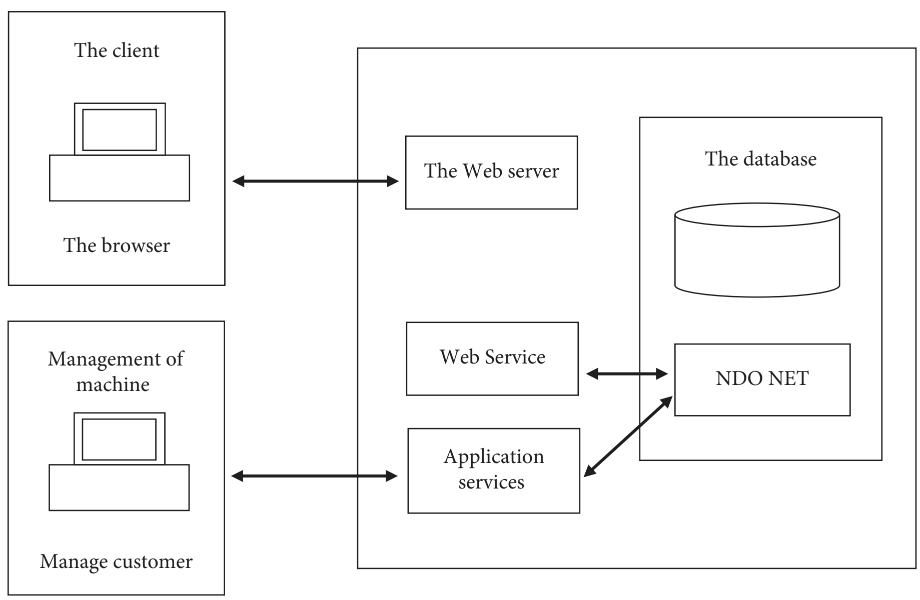


FIGURE 4: Examination system three—layer architecture.

asynchronous data exchange to enable user program interaction without waiting for server-side communication, helping users to save waiting time and improve the efficiency of program operation. Figure 5 shows a comparison diagram between the ajax model and the traditional Web workaround. Figure 5 shows the ajax model.

3.3. Theoretical Basis of Test Bank Construction. A test bank is simply a set of different exam questions. The test bank is set up with certain requirements for the test questions [17]. If all the questions can be entered into the test bank, then even if there is a question bank, this question bank is meaningless. In the development of educational econometrics, the construction of the test bank should be both theoretically based and index-constrained. The constraint of test questions in the test bank is the indicator constraint in the theory of educational measurement. These indicators are quantitative statistics that can be calculated by corresponding methods, and we call these indicators item statistics (or item parameters). The design of the test bank is based on modern examination theory, and there are currently two main approaches, namely, the classical measurement theory and the item response theory.

The test question index system contains a set of parameters that provide a qualitative or quantitative description of the intrinsic and extrinsic properties of the test questions and their role in the examination system. The automatic assembly of articles and the establishment of the

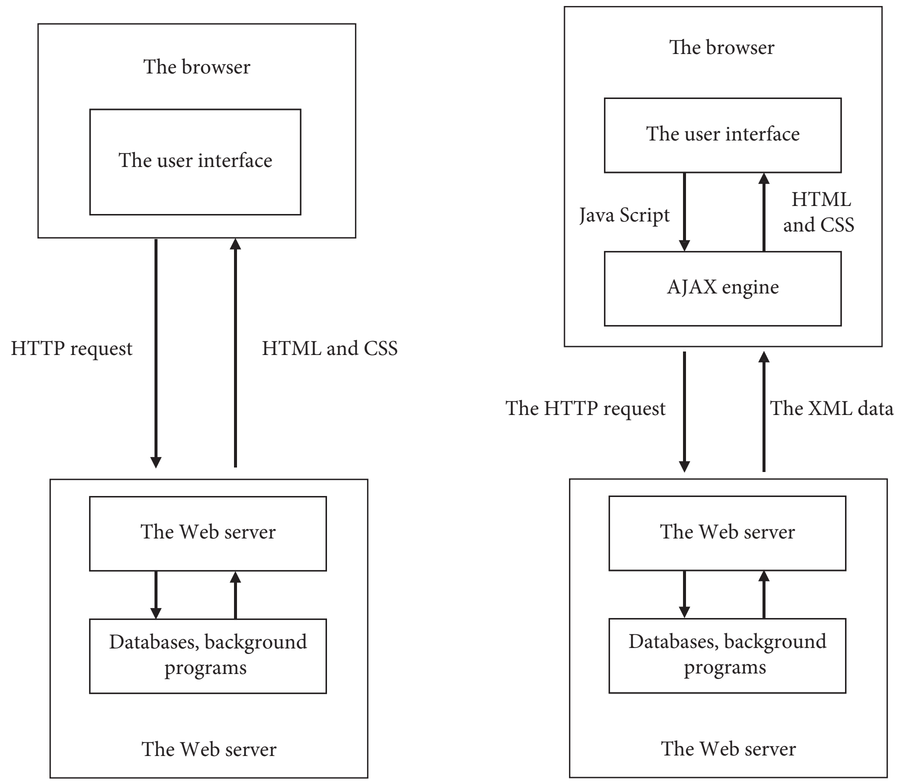


FIGURE 5: Diagram comparing the Ajax model to the traditional way the web works.

question bank management system are the prerequisites for the automatic assembly of test papers. The reasonableness of the parameters of the test indicator system directly affects the success or failure of group examinations. The test question indicator system defined in this article is: C = {knowledge, question type, chapter, difficulty, variance, ability requirement, score, and exposure}. There are several test question indicators reflecting subject matter information that can be provided directly to the user or computer for easy adjustment of the subject matter and computer grouping of papers. Difficulty refers to the difficulty procedure of the test questions. There are various ways to determine the difficulty of a test question, and the formula for calculating the difficulty of an objective question is as follows:

$$P = \frac{R}{N},$$
$$f = \sum_{i=1}^{t=5} f_i w_i, \tag{1}$$
$$f = k_1 \times W_N + k_2 \times W_J + k_3 \times W_C + k_4 \times W_Q.$$

$P = 0$ means that all the questions had no correct answers and the question was very difficult. $p = 1$ means that all the students got it right and the questions were very difficult. Obviously, the difficulty index refers to passing the test. The higher the rate of passing is, the easier the problem, and vice versa,. The formula for calculating the difficulty of subjective questions is as follows:

$$P_i = \frac{Xi}{K_i},$$
$$E = \frac{1}{\sum_{k=1}^{N} \left(T_k - Y_k\right)^2}, \tag{2}$$

P_i represents the difficulty level of the question, X_i represents the average score of the question, and K_i represents the full score of the question. The value of the difficulty index represents the degree of difficulty that is inversely proportional to the P value relative to the degree of difficulty of the objective questions.

The difficulty of a test paper is usually required to be normally distributed in order for the test paper to be of test value. For questions with too high or too low difficulty, test papers are written with low variability, which is not conducive to the evaluation of teaching effectiveness. Based on the difficulty and score of each question in the test paper, the difficulty of the entire paper can be found by the following equation:

$$P\left(b_i\right) = \frac{E_{b_i}}{E},$$
$$P = \frac{\sum_{i=1}^{n} P_i K_i}{\sum_{i=1}^{n} K_i}, \tag{3}$$

where P is the overall difficulty of the paper, n is the number of questions included in the paper, P_i is the difficulty of the ith question, and K_i is the score of the ith question.

3.4. Basic Overview of E-Rater Technology. E-rater, developed by Educational Testing Service (ETS), is a hybrid scoring system with a modular structure. E-rater combines the advantages of PEG and IEA to provide fast and accurate diagnostic results. Table 3 summarizes the analytical focus, core technologies, and advantages and disadvantages of these three systems.

3.5. ASP.NET Working Fundamentals. ASP.NET, called Active Server Web, is based on the foundation of Web application development, and is Microsoft's software for implementing Web applications, an updated form of desktop applications. Implementing dynamic web pages is also a mainstream development technology. In practice, ASP.NET uses a very convenient programming environment in the writing process, which is very conducive to implementing dynamic web pages, and the combination of framework-based development and scripting files greatly facilitates the implementation process for developers, using COM components and other integrations in the system, which allows programmers to give greater autonomy, and the nested HTML code can be better written [18]. The nested HTML code can be used to better write interfaces and functions related to the WEB application, as shown in Figure 6.

3.6. Implementation of the Online Examination System. The system is based on the original online examination system; the design of the test bank and the analysis process of the intelligent paper formation algorithm are selected, and the combination of the test bank formation method and the paper formation algorithm is finally determined, with the aim of improving the quality of intelligent paper generation and the efficiency of paper release, and realizing the separation of teaching and examination in the state of school building, which is finally realized by the system [19]. The functional requirements of the whole system were determined through a comprehensive analysis of the internal survey and the actual situation facing teaching. Based on the functional requirements, we conducted feasibility analysis and auxiliary prototype system, and finally used ASP.NET development technology by comparing the existing auxiliary technology with the design and implementation part of the original system, while SQL Server 2008 was used for the database. The whole architecture was developed using mainstream B/S web application, and the home page was designed and developed using the latest HTML5 technology, combined with Ajax technology and JavaScript scripting language to achieve asynchronous transmission and page interaction. The combination of Ajax and JavaScript technology can largely solve the server load due to the large number of students using the system and reduce the waiting time. According to the system requirements and the problems implemented, the system development environment is shown in Table 4.

4. Result Analysis and Discussion

4.1. System Requirements Analysis. The online testing system based on the genetic algorithm should maintain the original testing functions on one hand, and further improve them on the other. The basic functions of the traditional test, such as issuing questions, assembling papers, taking exams, correcting questions, and summarizing. Improvements to traditional exams, such as adding some functions, improving on the basis of the original functions to make it more scientific. The online examination system based on the genetic algorithm should have the characteristics of abundant questions in the test bank, scientific and reasonable information formation, more comprehensive and scientific implementation, effective prevention of cheating, more accurate and thorough information analysis, easy to operate, easy to use, and can save a lot of human and material resources [20].

After a thorough study of the system and the work that needs to be done, this paper proposes the following points:

(1) Multirole user management: the system is logged in as administrator, teacher, and student. The administrator has higher privileges and can add, delete, and modify different users.

(2) Management of the test bank: the establishment of the test bank is the core of the whole online testing system, and the quality of its construction directly affects the overall operation of the system. In the construction of the test bank, various types of data forms should be established to manage various types of questions and set the difficulty, variance, burstiness, score, knowledge point, and question type of the questions, so as to classify them and improve them. The administrator of test bank can add, modify and delete test questions.

(3) Formation of papers: formation of papers should include two functions of automatic formation of papers and manual formation of papers, of which automatic formation of papers is the main one and manual formation of papers is supplementary and complementary to each other. Automatic grouping must use scientific and reasonable calculation methods to achieve the maximum demand for test papers, speed up the processing speed of test papers, and improve the quality of test papers. Manual grouping mainly artificially adjusts the test questions that do not meet the requirements manually, which improves the quality and scientificity of the test questions.

(4) Online test: the design requirements of the online test system mainly focus on the online test part. The online test should be able to cope with various unexpected situations in order to ensure the successful passing of the candidates. After registration, candidates must complete the test within the specified time limit, and to ensure the fairness of the test,

Table 3: Table of comparison of the three systems.

System	Analysis of the key	The core technology	Advantages	Disadvantages
PEG	Language	1. Statistical techniques 2. Natural language processing	Language quality analysis	1. Not analyzing content 2. Only analyze the surface features of the text, which is easy to be recognized by examinees
IEA	Content	Information retrieval technology	Content quality analysis	1. Not analyzing the language quality of the composition 2. No analysis of discourse structure
E-rater	Language content Chapter structure	1. Statistical techniques 2. Natural language processing 3. Information retrieval technology	1. More consistent with artificial scoring elements 2. Discourse structure analysis is included	1. Poor analysis of content quality 2. The analysis of language quality is not comprehensive

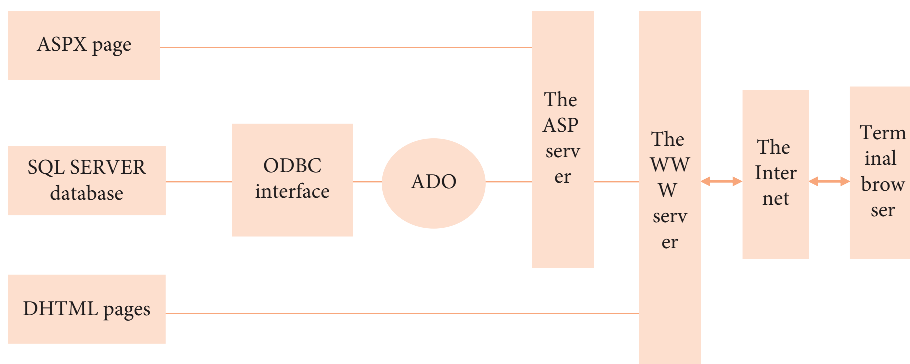


Figure 6: How ASP.NET works.

Table 4: System development environment requirements.

Projects	Detailed description
Operating system	Windows 7/8/10
Web server	IIS 7.0
Browser	Mainstream browsers are fine
Database	Microsoft SQL Server2008
ASP.NET	Microsoft NET framework 4.0
Unit testing tools	Unit

they must exit within the test time. In addition, to prevent cheating, the test should be supervised with the help of other auxiliary tools. The designed system data flow diagram is shown in Figure 7.

4.2. System Performance Testing. System performance testing aims to simulate the load capacity of the system under different load conditions in a given environment [21]. LoadRunner is a testing tool that can simulate a large number of users logging in and transmitting load at the same time to identify problems. Based on this, the relevant functions of the system were tested accordingly, and the test start clients were set to simulate the number of simultaneous logins of 700 and 2 seconds, respectively. The experimental results showed that the system had an average response time of 2.72 seconds and an average throughput of 48781.25 bytes, and are shown in Figure 8. During the test period, the data throughput of the system was basically stable except for fluctuations at the beginning and end of the system, and the average response time of the system was stable at other times. The tests showed that some performance metrics (e.g., system throughput and average response time) met the expected user requirements.

4.3. English Online Examination Management Module Design. The English online test management module is a key part of the English online test system. The design of the cloud computing English online test requires the ability to receive and process different test requirements in real time according to different test scales and requirements, and to make real-time adjustments to the cloud server. It also requires the sharing of test resources and support for the functions of the cloud management system, which is managed by professional technicians to accommodate dynamic storage requirements. Management module for English online exams. The English online test management module allows test takers to save their test resources at any time, and different test takers can store and retrieve them at any time. Test takers log in to the system through authentication to answer questions, and the process will be automatically timed and saved in real time; when the test is over, the relevant data is sent to the system. The huge space of cloud storage ensures that the stored data will not be damaged or lost. The structure of the English online test management module is shown in Figure 9.

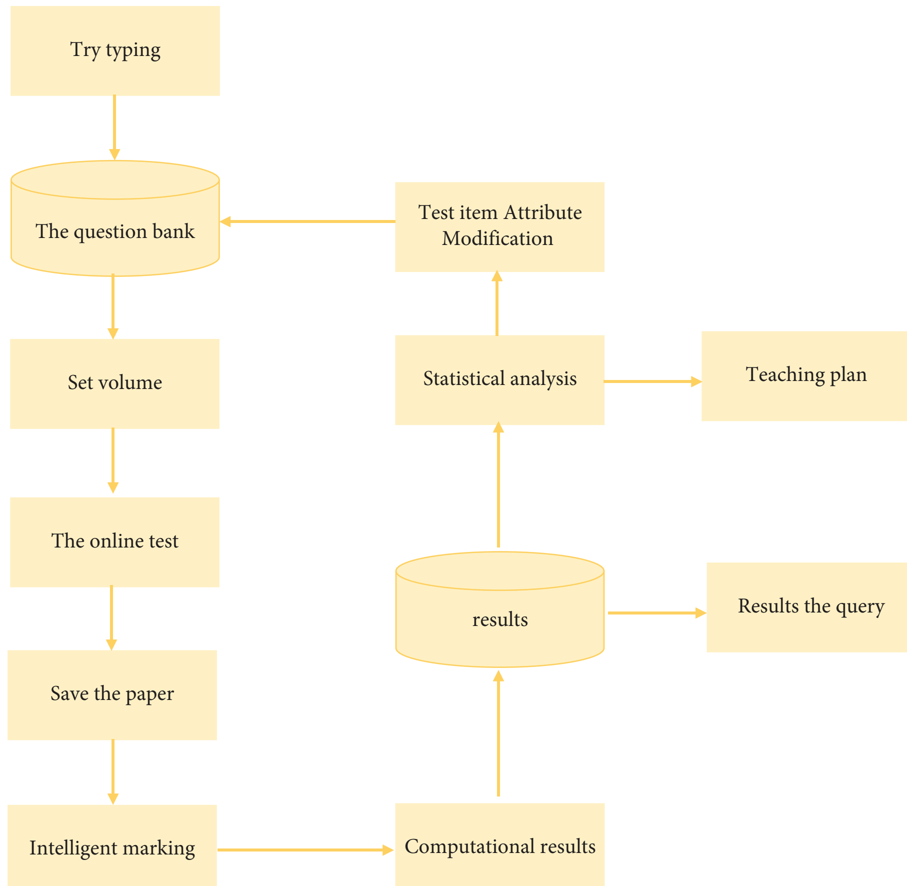


FIGURE 7: Data flow chart of online examination system based on the genetic algorithm.

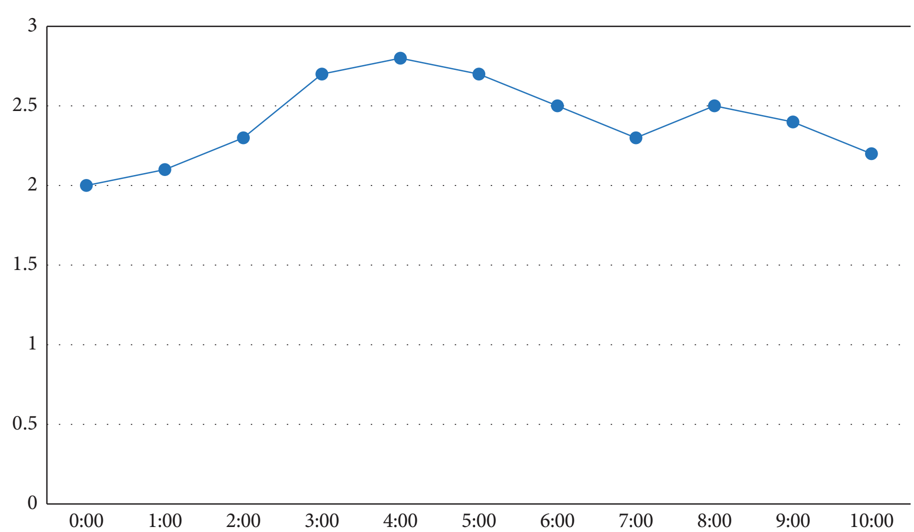


FIGURE 8: System response time diagram.

The English online test management module consists of paper grouping management, result management, test paper management, question bank management, the online test submodule, and the database value module. The paper management module can query, count, arrange, review, and print English online test questions; the web test subsystem can control the English online test, process monitoring, test paper storage, and other functions in real time, and use the database for data storage, backup, and recovery.

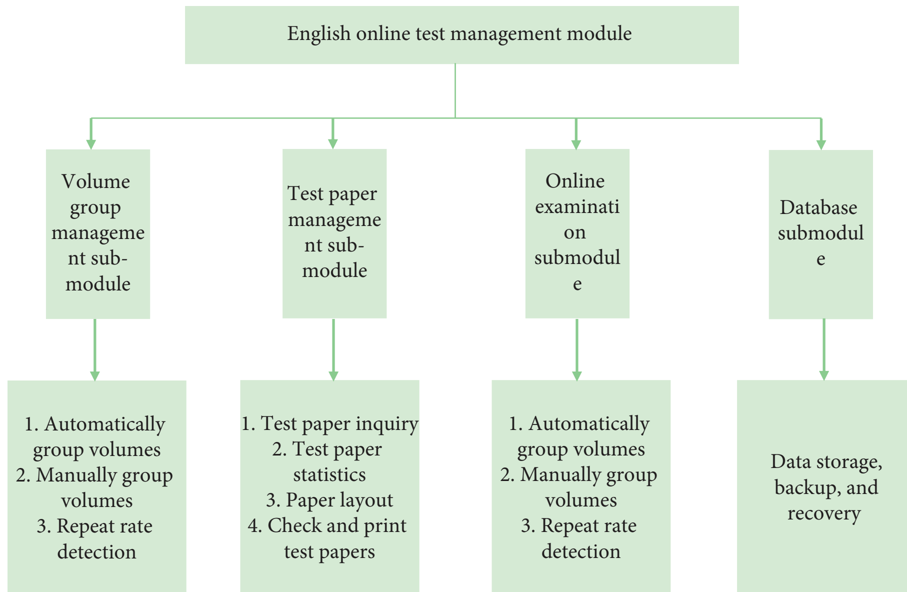


Figure 9: English online Examination management module structure diagram.

5. Conclusion

In this paper, firstly, we mainly analyze the current situation of online examination system, discuss the problem of insufficient intelligence of existing examination system, and design and implement a relatively intelligent and easy-to-implement online examination system based on the genetic algorithm. In the research and design of the whole system, the analysis focuses on the information grouping algorithm in the online examination system. Because of the good convergence of the genetic algorithm in the multi-constrained problem space, the system adopts a genetic algorithm based on the paper grouping algorithm. This paper first introduces the basic overview of the genetic algorithm, analyzes the constraint parameters to be used in the grouping model and the mathematical model of grouping, and then introduces the application of the genetic algorithm in the grouping algorithm strategy. Due to the limitation of the genetic algorithm itself, the basic genetic algorithm is integrated and designed with reference to the current relatively mature genetic algorithm and measures to reduce premature convergence during the design and implementation, and the finally the results are analyzed.

Development of English web-based teaching system. Through the establishment of this system, students will be able to conduct independent study and examinations, and obtain feedback information in time to make planned adjustments to the learning process and improve the teaching effectiveness. The next step is to conduct statistical analysis of the question bank and examination questions, and to analyze them to ensure the scientific and validity of the question bank, so as to improve the quality of the literature compilation and the effectiveness of online teaching.

Data Availability

The labeled data set used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This work is supported by the: The Philosophy and Social Science Project in Tianjin, TJZZ21-002.

References

[1] H. Wu, "Using modern tools to assist language testing," *Foreign Language E-Learning*, vol. 6, pp. 49–52, 2019.
[2] J. Yan, "The current situation and countermeasures for the development of distance education in China's universities," *Knowledge Economy*, vol. 96, no. l2, pp. 124–143, 2018.
[3] T. Luo and H. Z. Xu, "Web security issues," *Computer Applications*, vol. 20, no. 4, pp. 39–42, 2020.
[4] X. Chen, "Three-tier architecture of client/111E server," *Computer Applications*, vol. 20, no. 1, pp. 23–25, 2021.
[5] M. Hearst, "The debate on automated essay grading," *IEEE Intelligent Systems*, vol. 15, pp. 22–37, 2020.
[6] Apocrypha and S. Sun, *Principles and Applications of Genetic Algorithm Shirts*, pp. 6-7, National Defense Industry Press, Beijing, 2018.
[7] C. Wang, "Design of an online examination system based on JAVA technology," *Information/Communication*, vol. 93, no. 9, pp. 160-161, 2018.
[8] C. H. Chen, T. K. Liu, and J. H. Chou, "A novel crowding genetic algorithm and its applications to manufacturing robots," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 3, pp. 1705–1716, 2014.

[9] J. Francisco, V. Garcia, J. M. Moreside, and S. M. McGill, “MVC techniques to normalize trunk muscle EMG in healthy women,” *Journal of Electromyography and Kinesiology*, vol. 29, no. 7, pp. 574–780, 2022.

[10] G. Frosini, B. Lazzerini, and F. Marcelloni, “Performing automatic exams,” *Computers & Education*, vol. 31, no. 3, pp. 281–300, 1998.

[11] P. C. Gilmore and R. E. Geomory, “The theory and computation of knapsack functions,” *Operations Research*, vol. 14, pp. 1045–1074, 2021.

[12] Z. Li, “Research and implementation of online examination system,” *Henan Science and Technology*, vol. 21, no. 2, p. 12, 2020.

[13] Z. Yang, “Future classroom Future teacher Future education the integration of information technology and education change,” *China Modern Education Equipment*, vol. 3, no. 11, pp. 6–11, 2019.

[14] F. Tamburini, “A multimedia framework for second language teaching in self-access environments,” *Computers & Education*, vol. 32, no. 2, pp. 137–149, 1999.

[15] S. Z. Mirjalili, S. Mirjalili, S. Saremi, H. Faris, and I. Aljarah, “Grasshopper optimization algorithm for multi-objective optimization problems,” *Applied Intelligence*, vol. 48, no. 4, pp. 805–820, 2018.

[16] K. F. Curley, “Computer technology and knowledge workers: a pilot study of job impact,” *Interacting with Computers*, vol. 1, no. 2, pp. 171–182, 1989.

[17] A. D. Harris, “Online distanceeducation in the United Seats,” *IEEE Communications magzine*, vol. 37, no. 3, pp. 87–92, 2020.

[18] Y. Attali and J. Burstein, *Automated Essay Scoring With E-rater v.2.0*, Research & Development, vol. 11, pp. 1–7, 2022.

[19] L. Si, X. Hu, and B. Liu, “Image matching algorithm based on the pattern recognition genetic algorithm,” *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7760437, 9 pages, 2022.

[20] S. Roger, *Pressman Software Engineering A Practitioner’S Approach*, vol. 01, Qing Hua press, 5 edition, pp. 562–589, Beijing, China, 2020.

[21] L. M. Ellram, “Total cost of ownership: elements and implementation,” *International Journal of Purchasing and Materials Management*, vol. 29, no. 3, pp. 2–11, 1993.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 1898066, 11 pages
https://doi.org/10.1155/2022/1898066

Research Article

Research on the Application of Computer-Assisted Translation Technology in Translation Teaching

Yang Yang and Meng Su

North China Institute of Aerospace Engineering, Lang Fang 065000, China

Correspondence should be addressed to Yang Yang; wywl0110@nciae.edu.cn

Received 5 July 2022; Revised 22 August 2022; Accepted 26 August 2022; Published 15 September 2022

Academic Editor: Jiafu Su

In order to improve the quality of translation teaching, this study combines computer-aided translation technology to construct a translation teaching system, uses digital signal processing as a technical means to process waveform data online, and designs digital models for signal flow. Moreover, this study designs a digital filter to suppress waveform noise to improve the accuracy of amplitude measurement and uses a hardware algorithm for accumulating shift and averaging to correct the baseline stacking effect in real time. In addition, this study builds an intelligent translation software system with the support of algorithms in combination with translation requirements and conducts experimental verification of translation digital processing. Finally, the simulation experiment verifies the digital translation processing effect of the method in this study, and the application effect of computer-aided translation technology in translation teaching is verified through teaching evaluation research.

1. Introduction

The continuous progress of society and the increasing frequency of international exchanges have expanded the demand for compound translators in the international market. At present, many domestic colleges and universities have opened English majors, and even some colleges and universities have opened translation majors at the undergraduate level. However, although thousands of graduates pour into the market every year, they still cannot meet the needs of the society for translation talents. The main reason for this phenomenon is that the students' translation practice ability fails to meet the requirements of the translation market. The translation teaching mode of school teachers directly affects students' translation practice ability. The traditional "teacher-centered, text-based" translation teaching mode largely limits the cultivation of students' translation application ability. However, the continuous updating of computer information technology has promoted the innovation of translation methods and translation tools. Moreover, computer-aided translation (CAT) has become a necessary application technology for translation agencies.

Translation teaching has always followed the teaching of language skills from the bottom up. Traditional translation teaching takes teachers as the center, ignores the subject status of students, and cannot effectively cultivate students' translation practice ability and creativity. Translation teaching should change from "teacher-centered to student-centered" [1], in order to change the drawbacks brought by traditional translation teaching. The focus of the American interactive language teaching method is to take students as the center. According to the individual characteristics of students and the differences in teaching resources, teachers adopt diversified and flexible teaching methods according to their own teaching characteristics, so as to avoid students' passive learning state [2]]. The application of the interactive language learning method in translation teaching helps students to have a deep understanding of translation theories and techniques. In the mode of translation workshops, it is more conducive to creating a learning atmosphere of interaction between teachers and students. The learning exchange between students can cultivate students' autonomous learning ability, especially eliminate the tension of students in traditional classrooms, and then better participate in various teaching links designed by teachers [3]. In the computer-aided translation (CAT) environment, the process of simulating professional translation in the classroom can be realized, and the combination with the workshop style translation environment can truly build a professional translation ability training for students and collaborative ability development for the purpose. On the other hand, by supervising the translation process, teachers can monitor and summarize students' errors in the translation process, and provide a wide range of teaching data bases for teachers' translation teaching, which is conducive to teachers' design of personalized lesson plans, so as to better carry out translation teaching [4].

Compared with the traditional teaching concept of one-sided strengthening of social values, CAT translation teaching pays more attention to the individual development of students. In the era of "a hundred flowers blooming", it is undoubtedly better to cultivate talents with flamboyant personality, and education has achieved the purpose of "not only making people learn to do things (todo), but more importantly, making people learn to be people (tobe)" [5]. In this sense, the research on the training mode of translation talents under CAT translation teaching starts from the students, and the end point is also on the students. Only by fully understanding the needs of students, learning needs initiated by students to promote teaching reform, and changing top-level design to guide school-based design to school-based design to promote top-level design, this model is the focus of our research [6].

Translation teaching in colleges and universities is a special education category that integrates the characteristics of vocational education and higher education. On the one hand, translation teaching in colleges and universities should aim at cultivating practical translation talents [7]. However, translation teaching is only a subject teaching under the English major, and the teaching time is limited, only a short one year. Therefore, it is impossible to organize the teaching content according to the logical sequence of subject knowledge, emphasizing the comprehensiveness, system, and depth of the curriculum. Only if the curriculum structure is reasonable, the main body (students) of translation teaching can not only learn something in just one year but also lay the foundation for future personal development. The content of CAT translation teaching integrates computer skills, translation tools application, translation skills training, and other aspects of knowledge. The deeper content is that students can improve themselves and form their own values, professional ethics, and social responsibility through the self-construction of this knowledge and abilities [8]. Then, how to structure the CAT translation course is the core. Xu Bin proposed that in the CAT teaching system, the following teaching modules should be reasonably adjusted (but must be included): (1) the translation information technology module, including advanced word processing skills, digital text acquisition technology, input technology, search technology, corpus; (2) terminology work, that is, the establishment and management of termbase (TM) [9]; (3) CAT system application (such as SDL Trados, D6jaVu, Yaxin, etc.); and (4) translation project management, that is, familiarity with translation normal work of the company. These CAT teaching modules should be said to be very specific. These modules cover most of the content in CAT teaching and lay a solid foundation for students to move towards professional translation in the future. At the same time, the content of CAT practice also tends to be translated (for example, tenders, contracts, annual reports, etc., the content of the same field is highly overlapping), which fully shows that CAT teaching is an atypical professional teaching. But as mentioned above, translation teaching in colleges and universities is a special education category that combines the characteristics of vocational education and higher education [10]. "The teaching of translation can be divided into three levels: first, teaching translation as a purely foreign language teaching method; second, translation teaching as a foreign language professional course; third, translation professional teaching as a translation professional course." They further clearly pointed out: "First The third level belongs to the teaching translation, the purpose is to improve the foreign language level; the third level belongs to the translation teaching, the purpose is to cultivate professional translators; The primary course for translation talents [11]." In this sense, the teaching trend of CAT professionalization has weakened the meaning contained in the "advanced course of foreign language teaching" to a certain extent, that is, translation in the traditional "literary translation". Theoretical study and the cultivation of humanistic quality. CAT teaching weakens theoretical teaching (but also includes theoretical learning) and emphasizes applied translation practice (and occasionally literary translation) [12]. Therefore, how to resolve the contradiction between the two is also an important part of the curriculum. There are two solutions: one is to arrange a certain time (for example, 1/3 of one year's translation teaching time) to specially arrange to explain the theoretical part and do literary translation practice; the other is to use the MT (It consciously asks students to do bilingual parallel corpus (memory bank) of literary translation, and then through the query and comparison of corpus, analysis and explanation, the theoretical learning part and the practical content of literary translation are evenly distributed in the whole translation teaching activities. From the practical effect, the second method seems to be easier for students to accept [13]. Therefore, in addition to the four teaching modules mentioned by Xu Bin, the CAT teaching system should add a "translation appreciation and evaluation". This module focuses on the appreciation and evaluation of literary translation. In the appreciation and evaluation, it explains translation theory, experiences translation skills, and accepts humanistic influence [14].

Considering the class hours and task volume of translation courses, traditional translation courses are basically evaluated by a test study at the end of the semester. This summative evaluation method results in a single evaluation content, evaluation subject, and evaluation link, and it is difficult to achieve the evaluation effect [15]. The translation teaching of CAT should change this summative evaluation into a formative evaluation. The five major modules in the CAT teaching system are both independent and intersecting, especially the four modules mentioned by Xu Bin, if one is not learned, it will affect the learning effect of the entire CAT. For example, if you do not know "advanced word processing", you cannot import the CAT system (such as trados) normally when dealing with some manuscripts with messy formats such as Word documents. Even if you import people, many tags will be displayed in the CAT system, causing translation time trouble. For another example, if you do not create, maintain, and edit a memory library (MT) or term library (TM), you will not be able to fully and effectively utilize the functions of the CAT system, and will not achieve the purpose of saving time and efficiency. Therefore, CAT teaching is an interlocking teaching activity, which also automatically makes the traditional summative assessment evolve into a formative assessment [16].

This study combines computer-aided translation technology to construct a translation teaching system, improve the quality of translation teaching in colleges and universities, and promote the intelligent development of translation teaching.

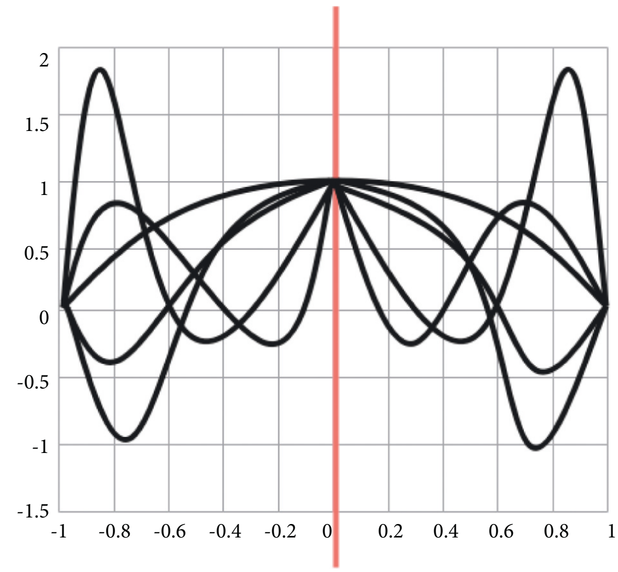


Figure 1: :Runge phenomenon.

2. Digital Processing of Translation Signals

2.1. DSP Digital Signal Processing. The main idea of digital Constant Fraction Timing (dCDT) is that after the nuclear pulse waveform is input, the ADC obtains accurate waveform amplitude information and collects enough waveform front-end points. By combining "coarse" counting and "fine" measurement, "coarse" counting is implemented by the FPGA high-order counter, which achieves a time resolution of 8 ns of ADC sampling time. Furthermore, "fine" measurements rely on DSP for waveform interpolation, which is reconstructed within one clock cycle. Its timing accuracy can be much smaller than the ADC clock period, reaching sub-nanosecond (100~10 ps) time resolution.

The AD9445 samples the entire waveform. The data enter the FPGA and undergoes FIR filtering and baseline recovery. After processing, the data is sent to the DSP, and the assembly function is used to send minidx (falling edge)/ maxidx (rising edge) to find the peak value in the data sequence. Although the more sampling points during the fitting, the more accurate the fitting accuracy is, at the same time the fitting time is also longer, resulting in an excessively long dead time window for DSP processing. Then, we select one sample point on the left side of the peak and two sample points on the right side, and we do a polynomial interpolation fit with a total of 4 points.

In scientific experiments or statistics, people often need to obtain an approximate expression $y=(x)$ for the independent variable x and the dependent variable y from a set of measured discrete data points (x, y). $i=0.1,\ldots, n$. This is to construct a mathematical approximation function $\varphi(x)$ from a given N points $(x.\ y)$, which requires the $\varphi(x)$ function curve to pass through all given points $(x.y)$ with equal function values, then $\varphi(x)$ is called an interpolation function.

The commonly used interpolation methods can be divided into two categories: algebraic polynomial interpolation and piecewise interpolation. For example, Lagrange interpolation and Newton interpolation belong to polynomial interpolation, and cubic spline interpolation belongs to piecewise interpolation. The advantage of algebraic polynomial interpolation is that the format is neat and standardized, but the disadvantage is that the Runge phenomenon will appear in high-order polynomials. As shown in Figure 1, the interpolation polynomial oscillates violently in the interpolation interval;

The Nth degree algebraic interpolation polynomial is:

$P_n(x) = a_0 + a_1x + \ldots + a_nx^n$ is an Nth algebraic interpolation polynomial, and the function is determined by $(n+1)$ coefficients $a_0, a_1, \ldots, a_n$:

When the curve passes through $(n+1)$ mutually different interpolation points (x_i, y_i) in a given plane, a linear system of equations can be obtained:

$$\begin{cases} a_0 + a_1x_0 + a_2x_0^2 + \ldots + a_nx_0^n = y_0, \\ a_0 + a_1x_1 + a_2x_1^2 + \ldots + a_nx_1^n = y_1, \\ \cdots \\ a_0 + a_1x_n + a_2x_n^2 + \ldots + a_nx_n^n = y_n. \end{cases} \quad (1)$$

The coefficient determinant of the system of (1) is the Vandermonde determinant:

$$V(x_0, x_1, \ldots, x_n) = \begin{vmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^n \\ 1 & x_1 & x_1^2 & \ldots & x_1^n \\ \ldots & & & & \\ 1 & x_n & x_n^2 & \ldots & x_n^n \end{vmatrix}. \quad (2)$$

When x_{i} are different from each other, there is $\prod_{0\le j\le i\le n}(x_i - x_j) \neq 0$, and the solution of the system of equations exists and is unique.

For $n+1$ different interpolation nodes (x, y), $i=0,1,\ldots$, n, based on the uniqueness of the Nth interpolation polynomial, a corresponding nth interpolation basis function 1

(x), $i-0,1, \ldots, n$ can be defined for each interpolation node x_i:

It is required that $x_0, x_1, \ldots, x_{i-1}, x_i, \ldots, x_n$ is the zero point of 1. (x), which can be set as:

$$l_i(x) = a_i(x-x_0)(x-x_1)\ldots(x-x_{i-1})(x-x_{i+1})(x-x_n), \tag{3}$$

$l_i(x_i) = 1, x = x_i$ is brought into the above formula to get:

$$l_i(x_i) = a_i(x_i-x_0)(x_i-x_1)\ldots(x_i-x_{i-1})(x_i-x_{i+1})\ldots(x_i-x_n) = 1$$

$$l_i(x) = \frac{(x-x_0)(x-x_1)\ldots(x-x_{i-1})(x-x_{i+1})\ldots(x-x_n)}{(x_i-x_0)(x_i-x_1)\ldots(x_i-x_{i-1})(x_i-x_{i+1})\ldots(x_i-x_n)} = \prod_{\substack{0\le j\le n \\ j\ne i}} \frac{x-x_j}{x_i-x_j} \tag{4}$$

The combined basis function is:

$$L_n(x) = \sum_{i=0}^{n} l_i(x) y_i. \tag{5}$$

That is, the Nth order Lagrangian interpolation polynomial.

Among them, linear interpolation is the simplest form of polynomial interpolation, and cubic Lagrangian interpolation polynomial is also called cubic interpolation.

$$L_n = \sum_{i=0}^{n} l_i(x) f(x_i) l_i(x) = \prod_{\substack{0\le j\le n \\ j\ne i}} \frac{x-x_j}{x_i-x_j} \quad n=3. \tag{6}$$

Piecewise interpolation is to divide the sequence (x_i, y_i) into segments. Only the discrete points in the small interval are fitted with the interpolation function, and each small interval is connected again to obtain the overall interpolation fitting. The advantage is that the local properties are good, and the disadvantage is that the interpolation nodes are not smooth.

The cubic spline interpolation polynomial is also called Spline interpolation: it is to construct a cubic polynomial in each sub-interval $[x_i, x_{i+1}]$, the function has equal function values at the nodes, and the first and second derivatives are required to be continuous. Each interval spline function can be uniquely determined by boundary conditions. The main advantage is that the degree of smoothness is high, and the parameters need to be solved by the undetermined coefficient method of the equation system, and the workload is larger than that of the polynomial interpolation method.

Cubic spline interpolation function $S(x)$: for $n+1$ nodes $[x_i, y_i], i = 0, 1, \ldots, n$ on a given interval $[a, b]$, the function $S(x)$ satisfies $S(x_i) = y_i$. $S(x)$ is at most a cubic polynomial on each cell interval $[x_i, x_{i+1}]$, and $S(x)$ has continuous second-order derivatives on $[a, b]$;

The existence and uniqueness of the constructed spline function can be determined by the method of undetermined coefficients.

Mark $M_i = S''(x_i), m_i = S'(x_i), i = 0, 1, \ldots, n$ is defined. Since $S''(x_i)$ is a linear function on the sub-interval $[x_i, x_{i+1}]$, a piecewise linear interpolation function of M is performed on $[x_i, x_{i+1}]$:

$$S''(x) = \frac{x-x_{i+1}}{x_i-x_{i+1}} M_i + \frac{x-x_i}{x_{i+1}-x_i} M_{i+1}, i = 0, 1, \ldots, n. \tag{7}$$

We set $h_i = x_{i+1} - x_i$ to get:

$$S''(x) = \frac{x_{i+1}-x_i}{h_i} M_i + \frac{x-x_i}{h_i} M_{i+1}, i = 0, 1, \ldots, n. \tag{8}$$

$S''(x)$ is integrated twice to get:

$$\begin{aligned} S(x) &= \frac{(x_{i+1}-x)^3}{6h_i} M_i + \frac{(x-x_i)^3}{6h_i} M_{i+1} + cx + d, \\ &= \frac{(x_{i+1}-x)^3}{6h_i} M_i + \frac{(x-x_i)^3}{6h_i} M_{i+1} \\ &\quad + C(x_{i+1}-x) + D(x-x_i) \end{aligned} \tag{9}$$

$S(x_i) = y_i, S(x_{i+1}) = y_{i+1}$ is substituted into the above formula, it can be solved:

$$C = \frac{y_i}{h_i} - \frac{h_i M_i}{6}, D = \frac{y_{i+1}}{h_i} - \frac{h_i M_{i+1}}{6}. \tag{10}$$

Therefore, on $[x_i, x_{i+1}]$, there is

$$S(x) = \frac{(x_{i+1}-x)^3 M_i + (x-x_i)^3 M_{i+1}}{6h_i} + \frac{(x_{i+1}-x) y_i + (x-x_i) y_{i+1}}{h_i} - \frac{h_i}{6}\left[(x_{i+1}-x) M_i + (x+x_i) M_{i+1}\right] x \in (x_i, x_{i+1}), \quad i = 0, 1, \ldots, n. \tag{11}$$

By $S_i'(x_i) = S_{i-1}'(x_i)$, the system of equations can be obtained:

$$\mu_i M_{i-1} + 2M_i + \lambda_i M_{i+1} = d_i, i = 0, 1, \ldots, n$$
$$\lambda_i = \frac{h_i}{h_i + h_{i+1}}. \tag{12}$$

Among them, there is $\mu_i = 1 - \lambda_i d_i = (6/(h_i + h_{i+1}))((y_{i+1} - y_i/h_i) - (y_i - y_{i-1}/h_{i-1})) = 6y(x_{i-1}, x_i, x_{i+1})$

When the values of M and M are given (when there is $M = 0$, $M. = 0$, it is called the natural boundary condition), at this time, the $n-1$ order equation system has n-1 unknowns, $i = 0, 1, \ldots, -1$, that is:

$$\begin{bmatrix} 2 & \lambda_1 & & & \\ \mu_2 & 2 & \lambda_2 & & \\ \cdots & \cdots & \cdots & & \\ & \mu_{n-2} & 2 & \lambda_{n-2} & \\ & & \mu_{n-1} & 2 & \end{bmatrix} \begin{bmatrix} M_1 \\ M_2 \\ \cdots \\ M_{n-2} \\ M_{n-1} \end{bmatrix} = \begin{bmatrix} d_1 - \mu_1 M_0 \\ d_2 \\ \cdots \\ d_{n-2} \\ d_{n-1} - \lambda_{n-1} M_n \end{bmatrix}. \tag{13}$$

The tridiagonal band matrix of the above (8) is solved, and the spline interpolation function of the discrete data system can be constructed. In the actual DSP time interpolation, linear interpolation, cubic interpolation, and spline interpolation are programmed through Matlab simulation. At the same time, it is tested in the DSP environment. In terms of interpolation accuracy, spline interpolation is the best, and in terms of execution speed, linear interpolation is the best. However, the DSP computing load and computing accuracy are considered comprehensively. From the real-time consideration, we use cubic interpolation to do our time interpolation function.

Through the cubic waveform interpolation expression obtained above, we construct the peak function curve from the 4-peak data given by the waveform peak finding function. Mathematically, to find the extreme value of the function, the numerical method — the chord tangent method is usually used, as shown in Figure 2.

If the first derivative of the peak function $L_z(x)$ is assumed to be $f(x)$, that is:

$$L_3'(x) = f(x). \tag{14}$$

Then, the root of $f(x)$ is the extremum of the function. The tangent method formula is defined as follows:

$$x_{k+1} = x_k = \frac{f(x_k)}{f'(x_k)} = x_k - \frac{f(x_k)(x_k - x_{k-1})}{f(x_k) - f(x_{k-1})}. \tag{15}$$

That is, the difference quotient is used to replace the derivative operation, and $f'(x_k) = \Delta f(x)/\Delta x = f(x_k) - f(x_{k-1})/x_k - x_{k-1}$.

The tangent method requires that two initial points x_0, x_1 are given at the beginning, and then $x_2, \ldots$ are obtained according to x_0, x_1. According to x_k、 x_k, x_{k+1} is obtained, and loop approximation is performed until the interpolated value of x obtained twice is $\leq \varepsilon$. We consider that approximate numerical roots are found.

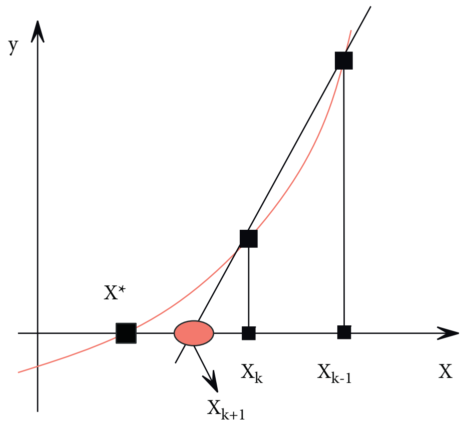


Figure 2: Chord tangent method to find function roots.

If it is assumed that the minimum value of the function is to be found, and x is known to be the minimum value of the sequence, then the extreme position of the function falls within the interval $[x_0, x_1]$ or $[x_1, x_2]$. First, the initial interval is given, and then the midpoint value of the interval is taken and substituted into the formula to calculate the midpoint function value y_m. Compared with the initial interval, the larger half interval is discarded, and the bisection approximation is recirculated until $y \leq \varepsilon$ is obtained twice, and the approximate value is considered to be obtained.

The extreme value is obtained by the peak interpolation curve, and the interpolation amplitude $S_{amplitude}$ of the waveform only needs to be multiplied by the constant ratio factor p. The reading value that fluctuates with the fluctuation of the waveform amplitude is obtained: $S_{threshold} = p * S_{amplitude}$

2.2. ADC Linearity Test.

2.2. ADC Linearity Test. ADC is the key performance factor of the timing measurement system. ADC has many test indicators. What we care about is the overall linearity test indicator of the analog part of the system including ADC acquisition and op amp filtering.

The ADC integral nonlinearity (INL) is the deviation of the actual conversion curve from the ideal conversion curve, as shown in Figure 3:

$$INL = \frac{V_D - V_{zero}}{V_{LSB\_ideal}} - D, 0 < D < 2^N - 1. \tag{16}$$

The principle of the integral nonlinear static test is as follows: a single channel is an input with a certain DC voltage Vs, and the corresponding amplitude value measured by the ADC is recorded. According to the test data, the ADC static conversion curve of the channel is fitted by Matlab software. The drive AD8139 used in the dCFT system is inverting amplification, and in the ^{22}Na positron experimental test platform, a gain of 2 input signals for the driver circuit is required to obtain the optimal input amplitude. The static integral nonlinearity of dCFT platform channel A is shown in Figure 4, and the static integral nonlinearity of dCFT platform channel A is shown in Figure 5.

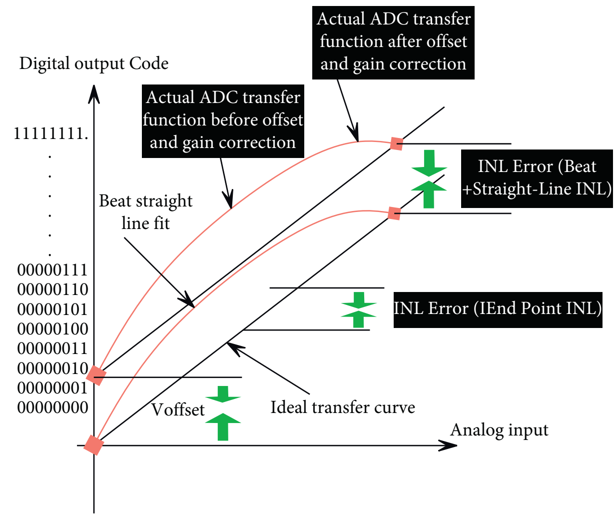


FIGURE 3: Schematic diagram of integral nonlinear.

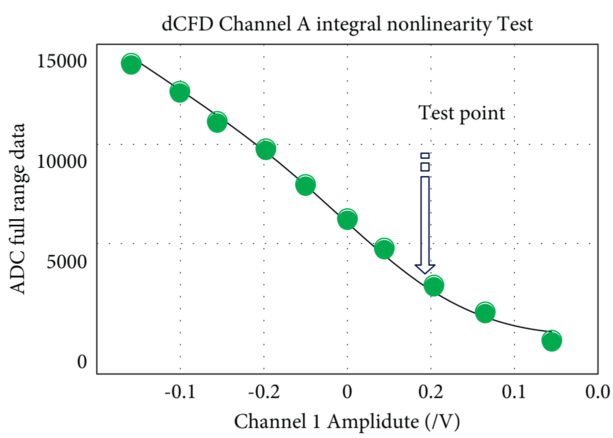


FIGURE 4: Static integral nonlinearity of dCFT channel A.

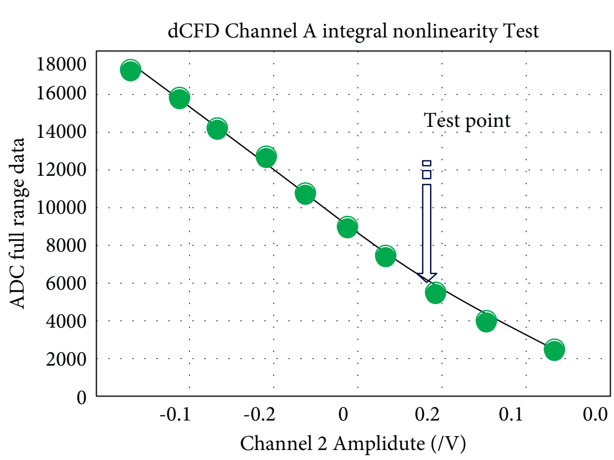


FIGURE 5: Static integral nonlinearity of dCFT channel B.

Differential nonlinearity (DNLA) is the relative deviation between the code width and the ideal code width in the actual conversion characteristics of the analog-to-digital conversion circuit, as shown in Figure 6.

$$DNL = \frac{V_{D+1} - V_D}{V_{LSB\_ideal}}, 0 < D < 2^N - 1. \tag{17}$$

The differential nonlinear dynamic test principle is as follows: a single channel is applied with a continuous analog signal, and the sampling output code is statistically analyzed to obtain the test circuit performance parameters, which is called the code density histogram test method. According to the difference in the added analog signal, it is divided into the ramp method (Ramp) and the sine method (Sine). The ADC samples a period of signal input, and the number of occurrences of different code outputs is called the code density. The ADC's output codes and their occurrences are plotted as coordinates, and the resulting graph is a code density histogram. In the histogram, each code is called a bin, and the number of occurrences of each code is called the bin width. The nonlinear parameters of the ADC can be calculated from the code density data.

The code density histogram test requirements are:

(1) Test data sampling data volume requirements

The code density histogram method is utilized to measure the dynamic performance of the ADC circuit, which is based on statistical theory. Therefore, in order to ensure the reliability and accuracy of the test, the data sample is required to be large enough. It can be obtained from the statistical theoretical analysis that the reliability (Z) and error (B) of the analysis method and the required number of samples and the number of ADC bits (N) have the following relationship:

$$N_{\text{Record}} = \frac{\pi \times 2^{N-1} \times (Z)^2}{\beta}. \tag{18}$$

For example, the number of ADC bits is 10Bit, the DNL and INL errors are required to be controlled within 0.1 LSB, and the analysis reliability is 95%. At the same time, the actual required sample data volume is $N_{\text{Resond}} = 617{,}920$, the analysis reliability is 99%, and the actual required sample data volume is $N_{\text{kscord}} = 1{,}070{,}678$. Moreover, the number of ADC bits is 14 Bit, the DNL and INL errors are required to be controlled within 0.1 LSB, and the analysis reliability is 95%. The actual required sample data volume is $N_{\text{Reced}} = 9{,}886{,}720$, the analysis reliability is 99%, and the actual required sample data volume is $N_{\text{Recowd}} = 17{,}130{,}848$;

(2) The sampling frequency and the sine wave input frequency cannot be correlated

The basis of the code density histogram measurement is to randomly sample the input signal, that is, to select an appropriate sampling frequency and signal frequency to ensure that the sampling circuit does not always repeatedly sample the same level. That is, the sampling frequency and the signal frequency are not allowed to have an integer multiple relationship, so as to ensure that the two are not related. In addition, the time jitter, reference drift,

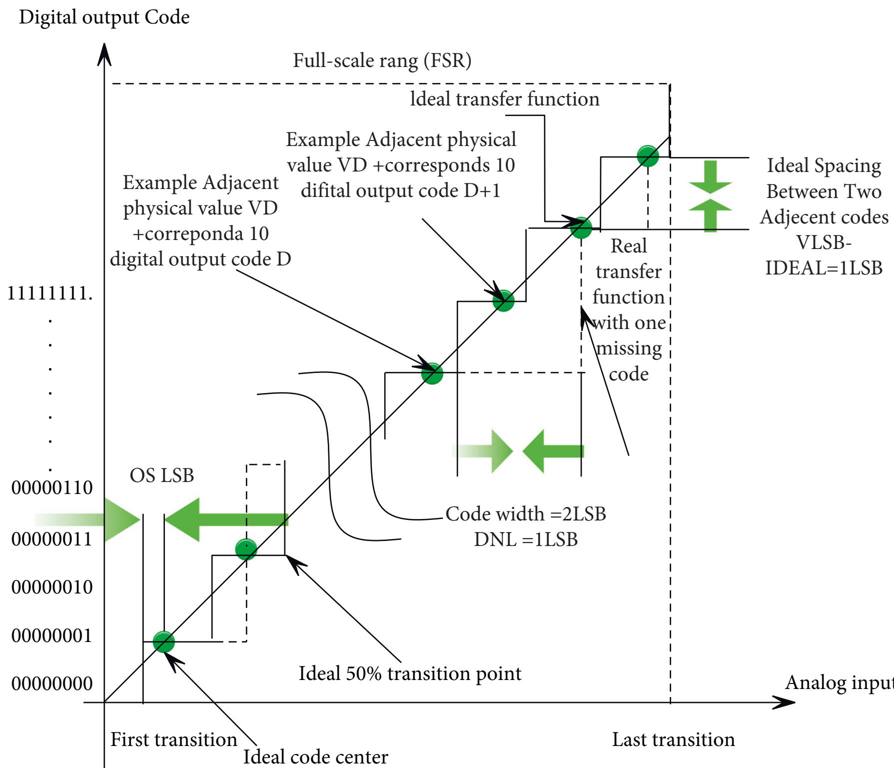


Figure 6: Schematic diagram of differential row linearity.

and frequency drift in the actual ADC circuit will make the sampling process more random.

(3) Selection of test signal

Usually, the code density test method often uses a ramp signal (triangular wave or sawtooth wave) or a sine signal as the input signal. The waveform code density histogram of the ramp method is ideal. Except for the two symbols at the edge, the number of output symbols is the same. However, the biggest problem is that when the frequency of the required input signal is high and the accuracy of the ADC circuit under test is high, it is difficult to obtain an input signal that meets the test requirements, and the nonlinearity and distortion of the input signal will have a greater impact on the test results.

For high-speed ADC circuit testing, a sine wave signal is usually used as the input signal. On the one hand, a sine wave can be obtained by filtering and other means to obtain a relatively ideal waveform. On the other hand, a sine wave can be precisely defined mathematically, and noise and dynamic distortion can be easily analyzed in the frequency domain. The ideal code density histogram of the sine wave is shown in Figure 7, and the actual discrete sampling code box histogram is shown in Figure 8.

The sine wave input signal is set to V_i, and there is $V_{IN} = A\sin(\omega t)$, we set $g(t) = A\sin(\omega t)$. Then, t can be regarded as a random variable, and within 0-T, t is uniformly distributed, that is:

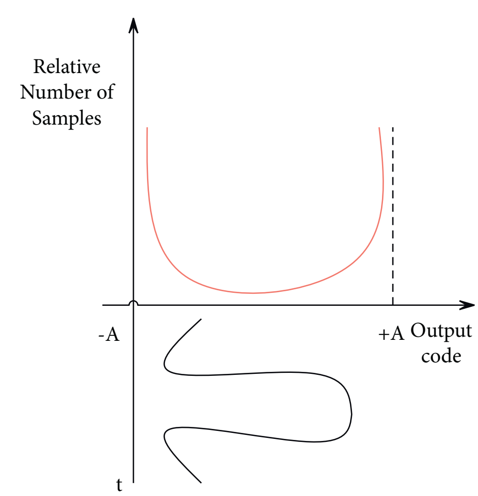


Figure 7: Histogram of ideal code density for sine wave input.

$$F_t(t) = \begin{cases} \frac{1}{T}, 0 \le t \ge T, \\ 0, otherwise, \end{cases} \tag{19}$$

where $F_t(t)$ is the probability density function of t, and V_{IN} is the function of the random variable t. From this, the probability density function of the sine wave V_{IN} can be obtained:

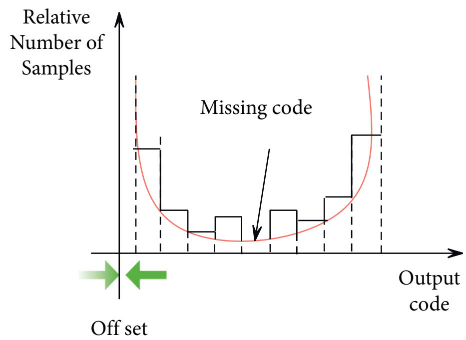


Figure 8: Density histogram of actual code of sine wave input.

$$P(V_{IN}) = \left|\frac{d}{d_{V_{IN}}} g^{-1}(V_{IN})\right| f_t(V_{IN}) = \frac{1}{\pi\sqrt{A^2 - V_{IN}^2}}. \quad (20)$$

Its distribution is shown in the figure. We integrate the probability density function of formula (14) and assume that the integration interval is $[V_a, V_b]$. Then, there is:

$$\begin{aligned} P(V_a, V_b) &= \int_{V_a}^{V_b} \frac{1}{\pi\sqrt{A^2 - V_{IN}^2}} dV \\ &= \frac{1}{\pi}\left\{\sin^{-1}\left[\frac{V_b}{A}\right] - \sin^{-1}\left[\frac{V_a}{A}\right]\right\}. \end{aligned} \quad (21)$$

$P(V_a, V_b)$ is the sampling probability in the $[V_a, V_b]$ voltage interval. We set $V_b - V_a = 1$ to get the discrete probability distribution function z:

$$P(i, A) = \frac{1}{\pi}\left\{\sin^{-1}\left[\left(\frac{2i - 2^n - 1}{2^n}\right)\frac{V_{ref}}{A}\right] - \sin^{-1}\left[\left(\frac{2i - 2^n - 3}{2^n}\right)\frac{V_{ref}}{A}\right]\right\}, \quad (22)$$

$P(i, A)$ is the probability that code i appears in the code box Bin $[i]$. V_{ref} is the full-scale voltage of the AD conversion circuit.

If the sine wave of the channel has a DC offset, which is set to V, there is $V_{IN} = V_0 + A\sin(\omega t)$, the probability density function is:

$$P(V) = \frac{1}{\pi\sqrt{A^2 - (V - V_0)^2}}. \quad (23)$$

Similarly, the probability density distribution function can be obtained:

$$\begin{aligned} P(V_a, V_b) &= \int_{V^a}^{V^b} \frac{1}{\pi\sqrt{A^2 - (V - V_0)^2}} dV \\ &= \frac{1}{\pi}\left\{\sin^{-1}\left[\frac{V_b - V_0}{A}\right] - \sin^{-1}\left[\frac{V_a - V_0}{A}\right]\right\}. \end{aligned} \quad (24)$$

We set $V_b - V_a = 1$, and in the same way, the discrete probability distribution function can be obtained:

$$P(i, A, V_0) = \frac{1}{\pi}\left\{\sin^{-1}\left[\left(\frac{2i - 2^n - 1 - 2V_0}{2^n}\right)\frac{V_{ref}}{A}\right] - \sin^{-1}\left[\left(\frac{2i - 2^n - 3 - 2V_0}{2^n}\right)\frac{V_{ref}}{A}\right]\right\}. \quad (25)$$

Derivation of calculation formula for ADC conventional dynamic conversion characteristic parameter test.

We set the code box corresponding to the output code i of the ADC to be $H[i]$, there are:

$$CH[i] = \sum_{j=0}^{i} H[i], \quad (26)$$

and it is an offset error.

If the ADC offset is zero, the sine wave histogram is symmetrical with respect to the midpoint, that is, the sum of the bin widths on both sides of the midpoint is equal. If the ADC has an offset V_d, the balance will be broken, and the sum of the widths of the code boxes on both sides of the middle point will be different. We set:

$$N_n = \sum_{i=1}^{2^{n-1}} H[i], N_p \sum_{i=2^{n-1}+1}^{2^n} H[i]. \quad (27)$$

We set P_d to be the probability that the randomly sampled voltage is positive, that is, the sampled voltage is within $(0, A + V_d)$. Pn is the probability that the randomly sampled voltage is negative, that is, the sampled voltage is within $(-A + V_d, 0)$. Then, there is formula (18) to get:

$$P_p = \frac{1}{\pi}\left\{\sin^{-1}(1) + \sin^{-1}\left(\frac{V_d}{A}\right)\right\} = \frac{1}{2} + \frac{1}{\pi}\sin^{-1}\left(\frac{V_d}{A}\right). \quad (28)$$

The entire sampling is an inevitable event, that is, $P_p + P_n = 1$. From this, we get:

$$P_n = \frac{1}{2} - \frac{1}{\pi}\sin^{-1}\left(\frac{V_d}{A}\right). \quad (29)$$

There are two formulas (22) and (23), we can solve:

$$V_d = A\frac{\pi}{2}\sin\left(P_p - P_n\right). \tag{30}$$

According to probability theory, P_p and P_n can be approximated by N_n/N_t and N_p/N_t, N_t is the total number of samples, and the estimated value of the offset V_d can be obtained:

$$V_d = A\frac{\pi}{2}\sin\frac{N_p - N_n}{N_p + N_n}. \tag{31}$$

In general, the offset V_d is small compared to the sine wave amplitude A, so the above formula can be further approximated as:

$$V_d = A\frac{\pi}{2}\frac{N_p - N_n}{N_p + N_n}. \tag{32}$$

The differential nonlinear DNL theoretical formula is the ratio of the actual probability of the code to the ideal probability minus 1, that is:

$$DNL = \frac{P_A(n)}{P_I(n)} - 1. \tag{33}$$

Among them, $P_A(n)$ is the actual probability of the occurrence of the nth code of the ADC, and $P_I(n)$ is the ideal probability of the occurrence of the nth code. In (18), $H[i]/N_t$ is used to approximate $P\,(V_a.\ V_b)$, and the estimated value $\hat{V}_b$ of V_b can be calculated. Because the offset voltage does not affect the integral and differential nonlinearity, it can be solved by formula (18):

It is obtained by the formula:

$$\begin{aligned}\hat{V}_b^2 = {}& V_a\cos\left(\frac{\pi H[i]}{N_t}\right)(2V_a)\hat{V}_b \\ & - A^2\left(1-\cos^2\left(\frac{\pi H[i]}{N_t}\right)\right) + V_a^2 = 0.\end{aligned} \tag{34}$$

$$\cos\left(\sin^{-1}\frac{V}{A}\right) = \frac{\sqrt{A^2 - V^2}}{A}. \tag{35}$$

We solve formula (18), and we can get:

$$\begin{aligned}&\hat{V}_b^2 - \left(2V_a\cos\left(\frac{\pi H[i]}{N_t}\right)\right)\hat{V}_b - A^2\left(1-\cos^2\left(\frac{\pi H[i]}{N_t}\right)\right) \\ &+ V_a^2 = 0.\end{aligned} \tag{36}$$

Only positive roots can be taken to solve the equation, and we can get:

$$\hat{V}_b^2 = V_a\cos\left(\frac{\pi H[i]}{N_t}\right) + \sin\left(\frac{\pi H[i]}{N_t}\right)\sqrt{A^2 - V_{i-1}^2}. \tag{37}$$

The discrete relation $V_b - V_a = 1$ is brought in, then we have:

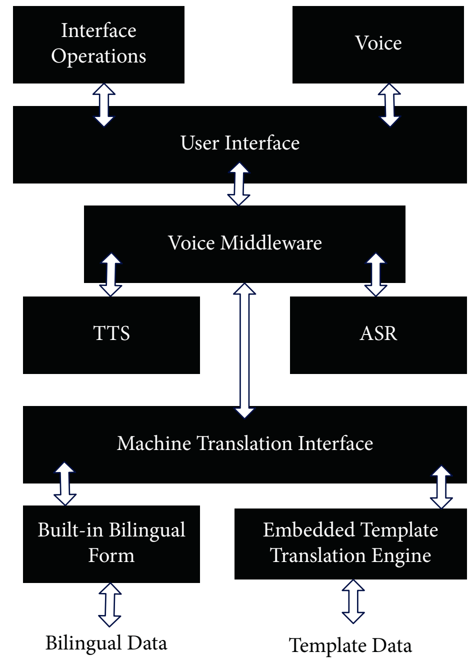


Figure 9: Intelligent computer-aided translation system.

$$\hat{V}_t = V_{i-1}\cos\left(\frac{\pi H[i]}{N_t}\right) + \sin\left(\frac{\pi H[i]}{N_t}\right)\sqrt{A^2 - V_{i-1}^2}. \tag{38}$$

The recursive formula does not produce a cumulative error. Therefore, $\hat{V}_i$ can be calculated from the boundary condition $V_1 = -A$ and the cumulative histogram CH $[i]$:

$$\hat{V}_i = -A\cos\left(\frac{\pi CH[i]}{N_t}\right). \tag{39}$$

In the equation, A is a linear unknown, which can be normalized, namely:

$$\overline{V}_i = -\cos\left(\frac{\pi CH[i]}{N_t}\right). \tag{40}$$

From this, the calculation formulas of differential nonlinearity (DNL) and integral nonlinearity (INL) can be obtained:

$$\begin{aligned}DHL &= \frac{\overline{V}[i+1] - \overline{V}[i]}{1LSB}, \\ INL &= \frac{\overline{V}[i] - \overline{V}[1]}{1LSB} - (i-1).\end{aligned} \tag{41}$$

3. System Construct and Test

Based on the algorithm of the second part, the intelligent computer-aided translation system constructed in this study is shown in Figure 9.

The dCFT translation system adopts the sine wave code density test, and the ratio of the ADC sampling frequency and the input sine wave frequency is irrelevant to meet the irrelevant condition. In order to achieve a 95% confidence

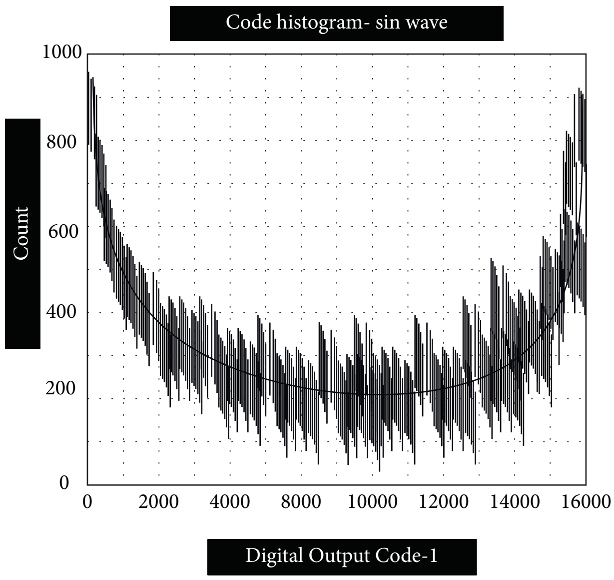


Figure 10: Histogram of ADC full-scale bin (bathtub curve).

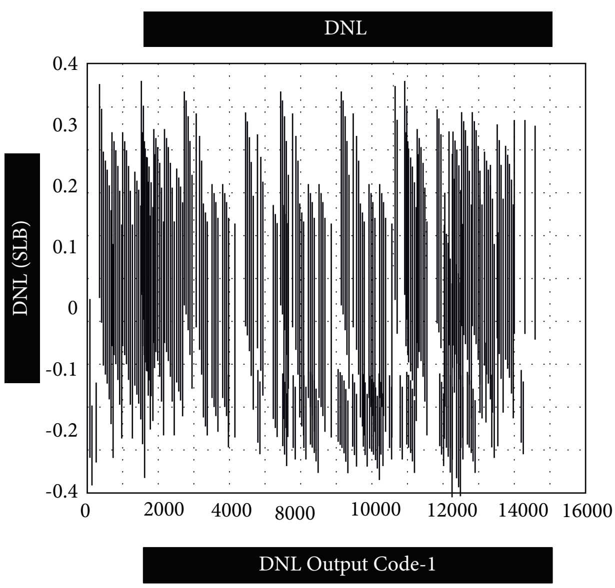


Figure 11: ADC DNL differential phenanthrene linear curve (95%). Therefore, at 95% confidence, $ADCDNL \approx -0.3 + 0.3\,(LSB)$.

rate, the 14-bit ADC needs to be tested more than 10,000,000 times. The sine wave data are drawn by Matlab, and the statistical histogram of each code box of the ADC is shown in Figure 10:

The ADC dynamic DNL curve can be calculated by bringing the code box statistical histogram into formula (18), as shown in Figure 11:

On the basis of the above research, the system proposed in this study is applied to translation teaching to explore its teaching effect. The results of the expert evaluation are shown in Table 1.

Table 1: The application effect of computer-assisted translation technology in translation teaching.

Num	Teaching effect	Num	Teaching effect
1	79.37	18	85.12
2	82.43	19	82.34
3	81.83	20	84.67
4	79.91	21	81.56
5	79.33	22	84.69
6	83.36	23	83.97
7	79.76	24	80.06
8	83.24	25	84.72
9	83.87	26	84.26
10	81.13	27	81.31
11	80.20	28	80.01
12	83.04	29	83.06
13	79.61	30	83.64
14	85.88	31	83.52
15	83.61	32	84.98
16	81.96	33	84.75
17	81.29	34	85.27

Through teaching evaluation research, it is verified that the application effect of computer-assisted translation technology in translation teaching is very good.

4. Conclusion

With the continuous development of computer technology, the degree of human dependence on computers is also rising rapidly. Therefore, how to solve the communication between people and computers has become the focus of the scientific community, which is also known as "intelligent human–computer interaction" technology. Among them, the most striking is undoubtedly the communication with the computer through language. The system proposed in this study is a multi-language intelligent translation terminal software system realized by using advanced technologies such as speech recognition, machine translation, speech synthesis, and an embedded multi-language service system framework. Moreover, this study combines computer-assisted translation technology to construct a translation teaching system to improve the quality of translation teaching in colleges and universities. In addition, this study verifies that the application of computer-assisted translation technology in translation teaching is very good through teaching evaluation research.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The research was supported by The Construction and Practice of Digital Resource of Moral Education in College English based on Blended-learning (Teaching Reform and Research Project of College English by Hebei Education Department).

References

[1] M. Rodríguez-Castro, "An integrated curricular design for computer-assisted translation tools: developing technical expertise," *The Interpreter and Translator Trainer*, vol. 12, no. 4, pp. 355–374, 2018.
[2] K. Taivalkoski-Shilov, "Ethical issues regarding machine (-assisted) translation of literary texts," *Perspectives*, vol. 27, no. 5, pp. 689–703, 2019.
[3] R. Shadiev, A. Sun, and Y. M. Huang, "A study of the facilitation of cross-cultural understanding and intercultural sensitivity using speech-enabled language translation technology," *British Journal of Educational Technology*, vol. 50, no. 3, pp. 1415–1433, 2019.
[4] R. Shadiev and Y. M. Huang, "Investigating student attention, meditation, cognitive load, and satisfaction during lectures in a foreign language supported by speech-enabled language translation," *Computer Assisted Language Learning*, vol. 33, no. 3, pp. 301–326, 2020.
[5] A. Tambusai, "Error analysis of instagram machine translation in Indonesia into English," *AFoSJ-LAS (All Fields of Science Journal Liaison Academia and Society)*, vol. 2, no. 2, pp. 261–271, 2022.
[6] F. Yasmin, "Challenges to computer assisted language teaching at university level," *INTERNATIONAL JOURNAL OF LANGUAGE & LITERATURE*, vol. 6, no. 2, pp. 86–93, 2018.
[7] S. M. Lee and N. Briggs, "Effects of using machine translation to mediate the revision process of Korean university students' academic writing," *ReCALL*, vol. 33, no. 1, pp. 18–33, 2021.
[8] R. Hussain, A. Lalande, C. Guigou, and A. Bozorg-Grayeli, "Contribution of augmented reality to minimally invasive computer-assisted cranial base surgery," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 7, pp. 2093–2106, 2020.
[9] D. Tafazoli, M. Gómez Parra, and H. Abril, "A cross-cultural study on the attitudes of English language students towards Computer-Assisted Language Learning," *Teaching English with Technology*, vol. 18, no. 2, pp. 34–68, 2018.
[10] P. Hendikawati, M. Z. Zahid, and R. Arifudin, "Android-based computer assisted instruction development as a learning resource for supporting self-regulated learning," *International Journal of Instruction*, vol. 12, no. 3, pp. 389–404, 2019.
[11] D. Zou, Y. Huang, H. Xie, D. Sa Zou, Y. Huang, and H. Xie, "Digital game-based vocabulary learning: where are we and where are we going?" *Computer Assisted Language Learning*, vol. 34, no. 5-6, pp. 751–777, 2021.
[12] R. Knowles, M. Sanchez-Torron, and P. Koehn, "A user study of neural interactive translation prediction," *Machine Translation*, vol. 33, no. 1-2, pp. 135–154, 2019.
[13] B. Raji, "Significance and challenges of computer assisted education programs in the UAE: a case study of higher learning and vocational education," *Education and Information Technologies*, vol. 24, no. 1, pp. 153–164, 2019.
[14] J. Kubicek, F. Tomanec, M. Cerny, D. Vilimek, M. Kalova, and D. Oczka, "Recent trends, technical concepts and components of computer-assisted orthopedic surgery systems: a comprehensive review," *Sensors*, vol. 19, no. 23, p. 5199, 2019.
[15] C. A. Lenkaitis, "Technology as a mediating tool: videoconferencing, L2 learning, and learner autonomy," *Computer Assisted Language Learning*, vol. 33, no. 5-6, pp. 483–509, 2020.
[16] A. D'Urso, V. Agnus, M. Barberio et al., "Computer-assisted quantification and visualization of bowel perfusion using fluorescence-based enhanced reality in left-sided colonic resections," *Surgical Endoscopy*, vol. 35, no. 8, pp. 4321–4331, 2021.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 8644276, 9 pages
https://doi.org/10.1155/2022/8644276

Research Article

Application of BP Neural Network in Teaching Quality Evaluation of Higher Vocational Education

Leimei Ai[1] and Tao Feng[2]

[1]Ganxi Vocational Institute of Science and Technology, Xinyu 338000, China
[2]School of Economics and Management, Beihang University, Beijing 100191, China

Correspondence should be addressed to Tao Feng; feng_tao@buaa.edu.cn

Received 15 June 2022; Revised 4 July 2022; Accepted 11 August 2022; Published 14 September 2022

Academic Editor: Jiafu Su

Higher vocational education has developed rapidly and attracted wide attention. Teaching quality evaluation is an important part of teaching management in classroom teaching and is an important part of teaching management. Based on the theory of teaching quality, 16 evaluation indicators are identified from the four aspects of teaching attitude, teaching content, teaching process, and teaching results and an HVE quality evaluation index system is constructed. The evaluation model is established by BP neural network algorithm. Finally, five higher vocational colleges in Henan Province are selected for empirical research. The results show that the evaluation model constructed in this paper has certain applicability and can provide a reference for HVE quality evaluation and improvement. Although the evaluation models herein have certain applicable values, further improvements are still needed in terms of indicator identification and correction.

1. Introduction

With the continuous development and progress of society, Higher Vocational Education (HVE) obviously has become an important part of the development of higher education. Improving the quality and development connotation has become a new focus, and higher education is increasingly becoming a concern. HVE is vocational education and technical education belonging to the level of education. Includes professional technical education before employment and postemployment-related continuing education. For example, some teaching programs of the American Institute of Technology and the Community College [1], Japan's short-term university partial teaching plan and specialized courses [2], French technical colleges and advanced technicians [3], Chinese early higher industry schools, specialized schools, specialists Schools [4], etc., as well as the education provided by some teaching plans such as adult colleges and universities. China began new development in the 1980s mainly cultivating liberal arts, science, engineering, agricultural and forestry, pharmaceutical, political and law, and finance and managerial professional assistant talents. Teaching quality assessment can provide a basis for improving the quality of running schools in higher vocational colleges, and establishing a foundation for the teaching quality assurance system.

Different subjects of HVE have different views on teaching quality. In the eyes of students, students are the main body of education, and all teaching behaviors are aimed at students' participation in teaching activities and are process-oriented. Teachers provide students with knowledge and skills through classroom teaching and develop students' values. For teachers, teaching ability and teaching level are the direct factors affecting the quality of teaching. Teaching behavior plays a guiding role in students, and it is the teacher's control to cultivate students' quality. The school level believes that the school's concept of talent training and the theme of running a school affect the quality of teaching, and the school's control of education quality is mostly reflected in the two levels of teachers' teaching and student participation. The social level believes that teaching quality is the satisfaction and adaptability of the public to the development of students. The teaching process should meet the development and needs of the society, and cultivate different talents in the society as the primary goal while ignoring the individual development of students. In summary, the quality of teaching should be a comprehensive consideration of the feelings of both students and teachers, and the quality of students directly reflects the level of teaching quality.

Teaching quality evaluation is to monitor teaching through a series of evaluation standards and control methods, obtain relevant data according to the evaluation standards, and build a reasonable evaluation index system. Guided by quality evaluation, teachers can form inertial thinking, optimize their own teaching behavior, and improve teaching effects. The evaluation of teaching quality is an activity that takes students and teachers as the main body and is guided by teaching goals. It is necessary to obtain correct evaluation results through technical analysis [5]. Therefore, it is very necessary to use relevant mathematical models to evaluate the quality of HVE.

HVE is in a period of economic transformation and development, due to the adjustment of industrial structure, it needs to be gradually perfected and improved. At present, HVE still has problems such as unclear training objectives, an imperfect education system, and differences in quality evaluation systems [6]. Lei and Zhong pointed out that in the teaching process of HVE, it is necessary to pay attention to the rational application of teaching evaluation and to use the multiple intelligence education quality evaluation systems to ensure the scientific and rational evaluation [7]. Li et al. and others analyzed the particularity of higher vocational education through research. The improvement of students' application ability is an important pursuit in the teaching process. At present, HVE needs to be improved in terms of teaching content, methods, teachers, and evaluation system. To meet the needs of higher vocational colleges [8]. Zhou explores students' ability development and its influencing factors in practical curriculum-oriented courses in higher vocational colleges through a fuzzy comprehensive evaluation method [9]. Miao started from the problems faced by the teaching reform of art courses in higher vocational colleges, established an evaluation index system from the evaluation of teaching reform ability, and used gray theory to build a teaching ability evaluation model [10].

In terms of research on constructing the evaluation model of HVE teaching quality, Liu and Lin constructed an evaluation index system from three aspects of teachers' technical ability, practical ability, and social ability under the background of the integration of production and education [11]. Kalkan and Cosguner have shown that according to the calculation results of the artificial neural network, academic self-efficacy is the most critical variable that affects students' achievements [12]. Based on improving BP neural network, the high vocational university performance innovation education model has been built by Zhang an evaluation index system from school supervision, teachers, and student education effects. Based on this, Fourier technology is used to improve the algorithm of the BP neural network to simplify the model [13]. In order to measure the quality of college physical education, Feng (2021) has built a teaching quality index system with big data technology and gives artificial intelligent quality data to construct a teaching quality assessment model [14]. The basic characteristics of teaching attitude, teaching content, teaching methods, and teachers, build a postgraduate teaching quality assessment index system and use BP neural network algorithm to build evaluation models, application sensitivity test identification key indicators [15]. Wei et al. has built a graduate nursing professional degree education quality assessment index system from the four aspects of input quality, process quality, output quality, and development quality, and empowering weights through Delphi law and level analysis [16]. Liu and Yin proposed a hybrid intelligence algorithm based on a genetic algorithm and reverse propagation neural network to assess the quality level of teaching, and play an intelligent algorithm in the evaluation model construction [17]. Considering the advantages of BP neural networks in evaluation [18], this study uses BP neural networks to analyze HVE evaluation indexes to establish a better evaluation index system and provide more accurate results.

This paper takes students and teachers as the main body of HVE, and constructs the quality evaluation index system of HVE from four aspects: teaching attitude, teaching content, teaching process, and teaching achievement. The HVE evaluation model is constructed based on the BP neural network algorithm, and empirical research is carried out with five universities in Henan province as the research object to verify the rationality and scientificity of the model constructed in this paper.

2. Evaluation Index System of HVE

The quality evaluation of HVE is guided by educational theory and teaching quality evaluation theory. Based on literature identification and brainstorming, this paper adopts the Delphi method to correct the evaluation index, and finally forms the quality evaluation index system of HVE. The construction of the evaluation index system is the core content of HVE teaching quality evaluation, and the scientific nature of the evaluation index system is the key to HVE teaching quality evaluation. In the process of teaching quality evaluation work, it is often necessary to pass the evaluation of peer teachers, students, and experts. In the evaluation process, teachers pay more attention to teachers' classroom performance, whether they can design teaching content based on the characteristics of the discipline, and whether they can combine theory with practice to achieve the purpose of higher vocational education. Expert evaluation mainly starts from the characteristics of the discipline, and examines whether teachers' teaching thinking and teaching content are closely related to the needs of social development. Expert evaluation is usually carried out in the form of lectures and courseware reviews. The evaluation from the student's point of view pays more attention to the content of the class and the conclusion after class. Students can understand the quality of teaching more intuitively by participating in classroom practice, and the teaching level of teachers is reflected in whether the knowledge imparted can be better absorbed and transformed by students. Teachers can also use evaluation indicators and evaluation standards as a reference to continuously adjust their teaching thinking, teaching content, and teaching methods to promote their own ability. It is the mission of higher vocational colleges to cultivate graduates with vocational skills. Therefore, in the construction of the quality evaluation index system of higher vocational education, it is necessary to focus on the interface between the teaching content and the needs of social talents.

The HVE evaluation index system is an evaluation of the teacher as a higher vocational education, from teaching preparation, teaching content, teaching process, teaching results, etc. As a multi-dimensional work, teaching quality evaluation should be a result of interacting with multiple factors. After preliminary identification and screening of the indicator, this article determines the first level indicator in the HVE evaluation index system as teaching attitude, teaching content, teaching processing, and teaching achievement, as shown in Figure 1. Teaching Attitude expressed the teacher's teaching work and attitude toward students. Teaching Content and Processing reflects the quality and ability of teachers to show in higher vocational education. Teaching achievement refers to the effects of students who have rendered after the education activities. The above four aspects can be more comprehensive to evaluate HVE.

Teaching attitude is an important factor in measuring HVE teaching value and teaching effect. The intuitive manifestation of teaching attitudes is whether the teacher can reach the teaching work on time, which reflects the self-discipline and integrity of teachers. The effect of the preparation determines the performance of teachers during the teaching process, and it is necessary to have a failure and low quality. In addition, teachers have determined students' attitude towards courses, and teachers should have strict requirements for students and maintain good communication with them. Based on the above discussion, the evaluation index of teaching attitude includes Punctuality in class A_{11}, Adequacy of preparation A_{12}, Attitude toward students A_{13}, Teacher-student relationship A_{14}.

The classroom teaching of HVE is different from undergraduate education, and its teaching content considers the cultivation of students' practical ability. When evaluating the quality of the teaching content, it is necessary to consider the cultivation of students' theoretical learning ability and hands-on ability. Based on the general teaching quality evaluation index, this paper considers two aspects of the combination of theory and practice in higher vocational education and theoretical teaching depth. Therefore, the evaluation indicators on the teaching content mainly include Well organized content A_{21}, Connection between theory and practice A_{22}, Clear content A_{23}, Cultivate comprehensive quality A_{24}, Moderate in depth A_{25}, etc.

The teaching process is the focus of HVE evaluation. In the process of higher vocational education, students should be student as the main body, considering the career development needs of students. HVE should jointly cultivate professional quality and overall quality, and the teaching method of practice classroom has a great difference in theory and theoretical learning. Therefore, in the teaching process, students should be highlighted in the cultivation of students, and teaching methods should be in both theory and practice. In terms of classroom organization design, teachers should pay attention to teaching logic and orderly, so that students can clearly understand the teaching intentions and teaching objectives, and complete their studies under the guidance of teachers to achieve learning goals. In addition, the proficiency of teaching content determines the authority of teaching, and inspirational artificiality reflected in the lecture process can effectively improve students' interest. In summary, the indicators in the quality assessment of the teaching process mainly include highlighting professional quality A_{31}, Applicability methods A_{32}, Logical and orderly A_{33}, Inspiring artistry A_{34}, etc.

Teaching achievements are the direct performance of HVE quality. The teaching achievements are mainly reflected in the quality of students' training, and the purpose of education is to cultivate people. HVE is the professional and technical personnel of society and workplace. Therefore, when evaluating the quality of the teaching results, teachers should consider the training and knowledge of students 'cultivation and students' learning and knowledge. Students' knowledge masters refer to whether students can fully master knowledge and apply the performance after they enter the workplace. In addition, the expansion of students' vision is also an important part of higher vocational education. Therefore, the evaluation index of teaching results mainly includes Students' level of knowledge A_{41}, Professional ability development A_{42}, Broadening of students' knowledge A_{43}, etc.

The evaluation index system can summarize the actual situation of HVE comprehensively. And these indexes can realize the evaluation of HVE as the input samples of the evaluation model, as shown in Table 1.

3. Evaluation Method of HVE

In this paper, a combination of fuzzy analytic hierarchy and BP neural network method is used to evaluate HVE. This can make full use of the advantages of both qualitative and quantitative methods. Fuzzy mathematics can better solve the fuzzy and difficult-to-quantify problems. The neural network has a strong non-linear mapping capability. In turn, the comprehensive method can effectively avoid the influence of too many subjective factors in the HVE assessment process and ensure the objectivity of the assessment results.

3.1. Fuzzy Analytic Hierarchy Process (FAHP). FAHP is a comprehensive evaluation method based on the theory of fuzzy mathematical affiliation, which has the characteristics of strong systematic and clear results, and is suitable for solving various nondeterministic problems. The core idea of this method is to convert qualitative evaluation problems into quantitative evaluation problems. It uses fuzzy mathematics to make an overall evaluation of the object which is subject to multiple factors. FAHP as a combined quantitative and qualitative evaluation method has the characteristics of strong systematic, clear thinking, and simple method. However, it still has certain shortcomings, for example, there is a certain bias in judging the consistency of the matrix, and

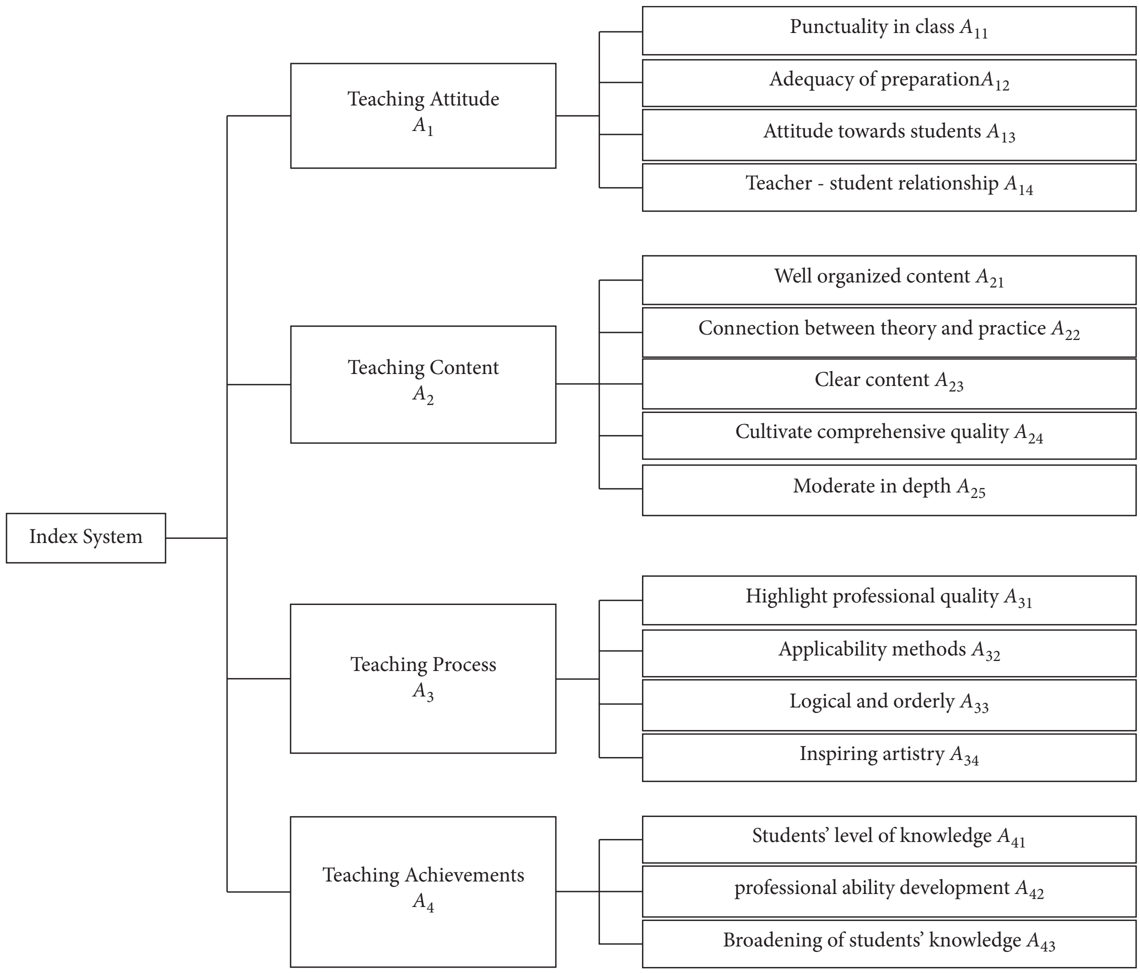


Figure 1: Evaluation index system of HVE.

Table 1: Explanation of Cantonese teaching evaluation index.

	Index	Explanation
A_{11}	Punctuality in class	Teacher arrives at the classroom on time, and completes the teaching content of the course in time
A_{12}	Adequacy of preparation	Teacher has prepared for teaching content before class
A_{13}	Attitude towards students	The attitude towards students should be strict and gentle
A_{14}	Teacher-student relationship	Teacher and students should maintain a great interaction relationship
A_{21}	Well organized content	The detailed level of teaching content should match important levels
A_{22}	Connection between theory and practice	Teaching content should meet the needs of students and career development needs
A_{23}	Clear content	The organization of teaching content should be clear and clear
A_{24}	Cultivate comprehensive quality	Based on the improvement of technical capacity, higher vocational teaching should pay attention to the cultivation of students' overall quality
A_{25}	Moderate in depth	Higher vocational teaching should pay more attention to the cultivation of hands-on operation capacity, rather than theoretical research
A_{31}	Highlight professional quality	Highlighting the characteristics of vocational education to cultivate high-end technical talents
A_{32}	Applicability methods	Whether teaching methods match professional skills
A_{33}	Logical and orderly	Classroom design should be both theoretical and practical, reaching a layer-in-depth effect
A_{34}	Inspiring artistry	Teachers should be very skilled in the teaching process, and can bring heuristic thinking for students
A_{41}	Students' level of knowledge	Students have the level of classroom knowledge and use to practice
A_{42}	Professional ability development	Students' ability to improve professional skills
A_{43}	Broadening of students' knowledge	Student's overall quality and knowledge growth

there is still no scientific basis for the consistency test criteria. Therefore, FAHP makes use of the advantages of fuzzy mathematics to make up for the shortcomings of the hierarchical analysis method. By introducing the fuzzy judgment matrix, the influence of subjective factors of experts is reduced.

Suppose the set of evaluation indicators is $X = \{x_1, x_2, \ldots, x_n\}$, where $x_i(i=1,\ 2,\ \ldots,\ n)$ is the evaluation factor. n is the number of individual factors at the same level. The set of evaluation results is assumed to be $V = \{v_1, v_2, \ldots, v_m\}$. where $v_j(j=1,\ 2,\ \ldots,\ m)$ is the evaluation result. m is the number of elements in the set, which is also the number of levels. This set specifies the range of results selection for a given evaluation element. The elements in the result set can be either qualitative or quantitative scores. Let the weight vector be $W = \{w_1, w_2, \ldots, w_n\}$, where $w_i(i=1,\ 2,\ \ldots,\ n)$ indicates the importance of factor x_i. The sum of the weights is 1, as shown in the following equation:

$$\sum_{i=1}^{n} w_i = 1, 0 \le w_i \le 1. \tag{1}$$

The relevant personnel in the education industry are organized to conduct fuzzy evaluation of each evaluation factor. According to the distribution of statistical evaluation levels of each indicator, assuming that the ratio of the number of evaluators who evaluate the indicator item x_i as level v_j to the number of all evaluators is r_{ij}, then the judgment matrix is $R = [r_{ij}]_{M\times N}$. Obviously, each row of this matrix is the evaluation result of every single factor, and the whole matrix contains all the information obtained from the evaluation of the set of evaluation factors X by the set of evaluation results V. The final evaluation result $\widetilde{B}$ is obtained by synthesizing the weight vector W with the judgment matrix R. That is, $\widetilde{B} = W \circ R = (b_1, b_2, \ldots, b_n)$, as shown in the following equation:

$$b_j = \sum_{i=1}^{n} w_{ij} \cdot a_{ij}, j = 1, 2, \ldots, m. \tag{2}$$

The evaluation result y of the first level evaluation factor can be obtained by normalizing $\widetilde{B}$. After that, the previous steps can be repeated to solve the evaluation results of the upper-level indicators in turn, and the overall fuzzy evaluation can be finally obtained.

3.2. Back-Propagation Neural Network (BPNN). BPNN is the most widely used neural network, which utilizes the error backpropagation algorithm to train multilayer feedforward neural networks. The learning memory capability and self-adaptive ability of BPNN are powerful, thanks to its nonlinear information processing system consisting of a large number of interconnected processing units. An important advantage of BPNN is that it eliminates the step of pretraining a fixed model. To obtain more desirable prediction results, it is only necessary to learn the training samples and then accumulate empirical knowledge.

As shown in Figure 2, the standard BP neural network model contains three layers, which are the input layer, the hidden layer, and the output layer. In the forward propagation process, the input samples are passed in from the input layer, processed through the hidden layer, and passed to the output layer. If the actual output of the output layer is not within the error range from the desired output, the backward propagation of the error is performed. After calculating the total error between the desired output and the actual value of all neurons in the output layer, the network weights and thresholds are adjusted. The error signals of the units in each layer are used to correct the weights of each unit so that the actual output of the neural network continuously approximates the desired output.

Where X_i is the input layer node, H_j is the hidden layer node, $Y_k$is the output layer node, and W_{ij} is the weight value between the layers. BPNN is a multilayer feedforward neural network with unidirectional propagation, and the standard BP network algorithm includes both forward propagation and backward transmission. For forward propagation, the data are corrected for weights through the input layer, and the resulting weighted sum is passed to the hidden layer, which is then activated by an activation function and passed to the output layer, as shown in the following equation:

$$H_j = \sum_{i=1}^{p} w_{ij} x_i \,(i = 1, 2, \ldots, p), \tag{3}$$

where p is the number of nodes in the input layer, and H_j is the weighted sum of the j-th node in the hidden layer. Through the action of the activation function and transmitted to the output layer, as shown in the following equation:

$$\Omega_k = f(H_j), \tag{4}$$

Ω_k is the result of the weighted sum of the inputs applied to the hidden layer by the activation function. For the evaluation of HVE, using the Sigmoid function as the activation function, the categories can be predicted and also approximate probability predictions can be obtained. The logarithmic odds function is a convex function of arbitrary order derivable, which has good mathematical properties and can be used directly by many numerical optimization algorithms to find the optimal solution.

For backpropagation, the purpose is to adjust the network weights and thresholds so that the error between the actual output value and the desired value is as close as possible to the predetermined value, and therefore a loss function needs to be designed. The mean square error (MSE) selected for this experiment, which is the mean value of the sum of squares of the errors at the corresponding points of the predicted and original data. It reflects the degree of difference between the estimated quantity and the estimated quantity, and the closer its value is to 0, the better the model selection and fitting, and the more accurate the data prediction, as shown in the following equation:

$$E = \frac{1}{2} \sum_{k=1}^{l} (y_k - y_k^*)^2, \tag{5}$$

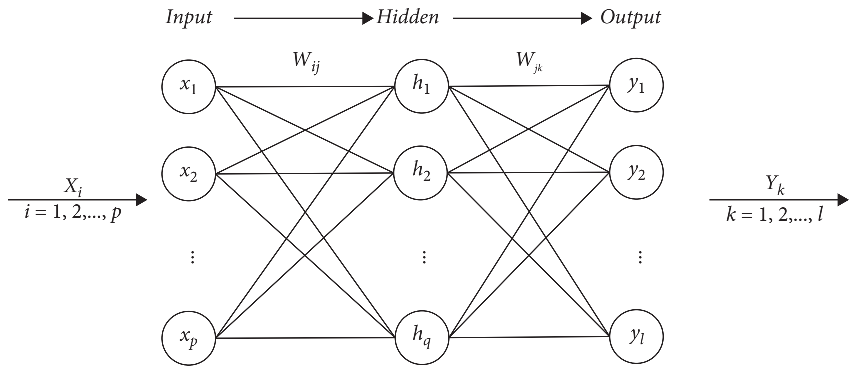


FIGURE 2: Classical BPNN model.

y_k is the expected value, y_k^* is the output value, and l is the number of nodes in the output layer. The connection weights are adjusted more accurately along the decreasing direction of the gradient of the mean square error E, as shown in the following equation:

$$\Delta w_{jk} = -\eta \frac{\partial E}{\partial w_{jk}}, \tag{6}$$

w_{jk} is the weight value between the hidden layer and the output layer, and η is the learning rate. The training is stopped when the actual output value of the output layer is almost indistinguishable from the desired value by continuously forward and backward propagation.

3.3. Comprehensive Evaluation Method. In this paper, an evaluation method based on the combination of FAHP and BPNN is proposed. First, we obtain a priori samples through FAHP and use these data to train BPNN. then, we further optimize the index weights through the advantage of the nonlinear calculation of BPNN. On the basis of combining quantitative and qualitative aspects, the final weights of each indicator are obtained, and then the purpose of evaluating HVE is achieved. The comprehensive evaluation process is shown in Figure 3.

4. Results and Discussion

In this paper, we take HVE assessment in Henan Province as an example. FAHP is used to realize the initial determination of index weights, construct the dataset, train BPNN, and obtain the final HVE assessment model. The experimental data in this paper are obtained from various higher vocational institutions in Henan Province. The big data are integrated and processed to obtain the HVE assessment index data.

4.1. Initial Determination of Index Weights. In the initial weighting period using FAHP, the relative importance of each index factor at each level is compared in the form of questionnaire. A total of 50 questionnaires are distributed to relevant experts in the field of education, and 46 valid questionnaires are recovered. The “0.1–0.9” scale shown in Table 2 is used to express their relative importance, so as to construct the fuzzy complementary judgment matrix of the corresponding indicators.

According to the results of the questionnaire, the scoring results of each index are averaged and five fuzzy complementary judgment matrices are calculated, which are R, R_1, R_2, R_3 and R_4. Based on the fuzzy complementary discriminant matrix $R = (r_{ij})_{n\times n}$, the corresponding fuzzy consistency matrix $R^* = (r_{ij}^*)_{n\times n}$ is calculated by using equation (7).

$$r_i^* = \sum_{k=1}^{n} r_{ik}\ (i = 1, 2 \ldots, n);\ r_{ij}^* = \frac{r_i^* - r_j^*}{2n} + 0.5. \tag{7}$$

Finally, based on the fuzzy consistency matrix, the initial weight vector of each layer of indicators is obtained by equation (8). The final calculation result is shown in equation (9).

$$w_i = \frac{1}{n} - \frac{1}{n-1} + \frac{2}{n^2 - n} \cdot \sum_{j=1}^{n} r_{ij}^*, \tag{8}$$

$$\begin{cases} W = (0.221, 0.256, 0.24, 0.283), \\ W_1 = (0.216, 0.225, 0.267, 0.292), \\ W_2 = (0.193, 0.25, 0.18, 0.213, 0.164), \\ W_3 = (0.304, 0.261, 0.239, 0.196), \\ W_4 = (0.312, 0.405, 0.283). \end{cases} \tag{9}$$

4.2. Indicator Weights Optimization. The linear weighted values of the sample indicators are calculated using the initial weights as the sample output labels, and the data set required for BPNN is constructed. The training set is divided into training and validation sets in the ratio of 8 : 2, where the training set contains 386 data and the validation set contains 97 data respectively. The evaluation grade of HVE is divided into four categories: A, B, C, and D according to the total score. Class A represents the best, i.e., the score range [4, 5], and class D represents the least, i.e. [1, 2].

A three-layer neural network structure is used with 16 nodes in the input layer, 3 nodes in the hidden layer, and 1 node in the output layer. The maximum number of training rounds is set to 1000 and the learning rate is 0.05. If the average error of the model on the validation set is less than 0.001, the training is ended early to prevent the overfitting

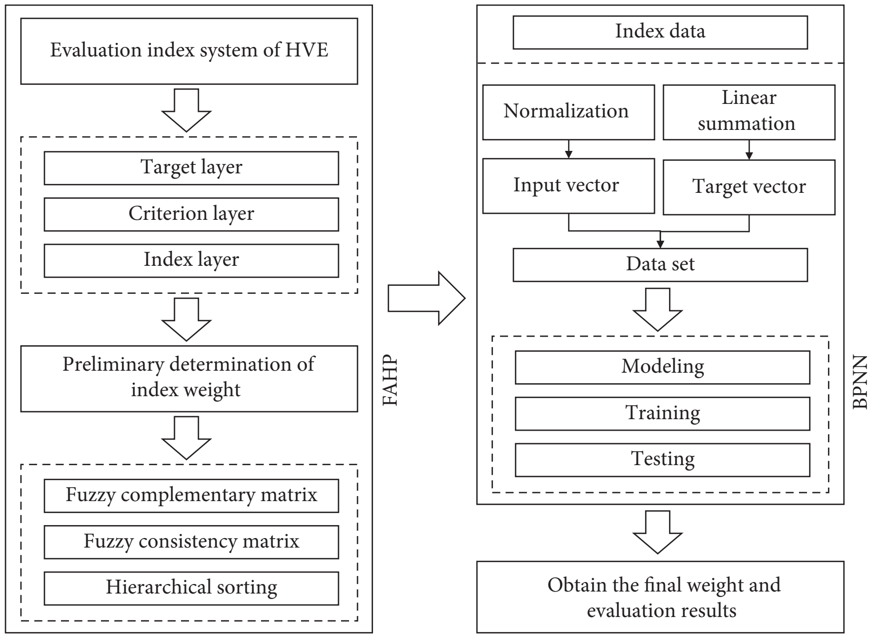


Figure 3: HVE evaluation model based on FAHP-BPNN.

Table 2: Comparison of importance weights of first-tier indicators.

Index	0.9 (extremely important)	0.8 (strongly important)	...	0.1 (extremely unimportant)	Index
A_1					A_4
A_2					A_3
A_3					A_2
A_4					A_1

Note. $r_{ij} = 0.5$ means i is as important as j.

phenomenon in further training. To verify the applicability of BPNN with HVE evaluation, the dataset was divided into 6 data sets and compared using the random forest (RF) and support vector machine (SVM) algorithms. The comparison results of the accuracy are shown in Figure 4.

The average accuracy of different machine learning methods is over 95%, which shows that the data in this paper is very reliable. And the applicability of BPNN with this paper is the strongest among these three methods. The role of BPNN in the combined model proposed in this paper is to reduce the influence of subjective factors of experts on the determination of weights, so the indicator weights of the neural network model after passing the training should be close to the initial weights, which will also increase the interpretability and reasonableness of BPNN. After several trials, the final model optimized indicator weights are shown in Table 3.

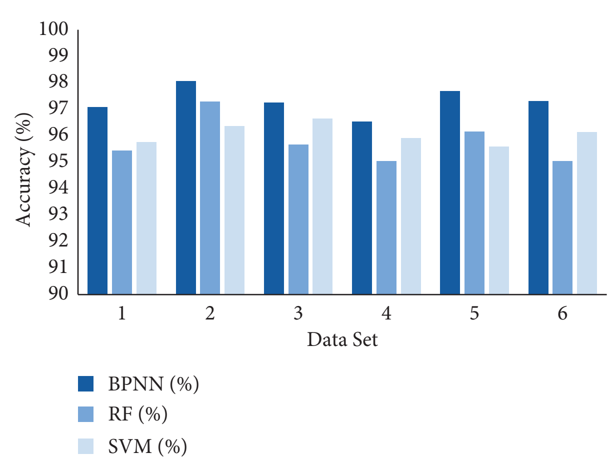


Figure 4: Accuracy comparison (BPNN, RF, SVM).

4.3. Comprehensive Evaluation Based on FAHP-BPNN. In order to prove the practical value of the model and make the model truly serve HVE, this paper invited relevant experts from higher vocational institutions in Henan Province to conduct interviews. The FAHP-BPNN model in this paper is used to evaluate the current educational status of five typical institutions so that the schools can have a better knowledge of today's teaching and the learning situation, and then make targeted improvements. The final score of HVE of each institution is shown in Table 4.

HZ pays attention to machine building and achieved the highest score of 4.0122, which belongs to grade A. While the comprehensive scores of TZ, KF, and SL are 3.5877, 3.3372, and 3.1264, respectively. Although all three are in the B grade, the LY has a much lower rating. SZ is the lowest score of 2.6773, a C grade.

Table 3: Evaluation indicators and weights.

Criterion layer	FAHP	BPNN	Index layer	FAHP	BPNN
A_1	0.221	0.228	A_{11}	0.048	0.052
			A_{12}	0.050	0.051
			A_{13}	0.059	0.057
			A_{14}	0.064	0.068
A_2	0.256	0.266	A_{21}	0.049	0.043
			A_{22}	0.064	0.074
			A_{23}	0.046	0.041
			A_{24}	0.055	0.061
			A_{25}	0.042	0.047
A_3	0.240	0.231	A_{31}	0.073	0.065
			A_{32}	0.063	0.059
			A_{33}	0.057	0.055
			A_{34}	0.047	0.052
A_4	0.283	0.275	A_{41}	0.088	0.085
			A_{42}	0.115	0.103
			A_{43}	0.080	0.087

Respondents generally give high ratings to the HVE effectiveness assessment index system. The effectiveness of teaching inside and outside the classroom is reflected by the indicators, and the indicators such as *Connection between theory and practice* not only reflect the technical but also the academic nature of higher vocational schools, which helps students to improve their overall quality. In addition, the weights in this paper are generally agreed upon. HVE values practical skills and therefore places the highest weight on teaching achievements. Teaching attitude, teaching content and teaching process complement each other and all obtain a weighting of 0.2 or more. And teaching content, as the core aspect of the classroom, achieves the second most important ranking.

In the final results, the indicators of *Teacher-student relationship*, *Connection between theory and practice*, and *Students' level of knowledge* receive the highest weights. This also points to the direction for the improvement of HVE. Testing and examinations are important initiatives that need to be used in vocational education in order to ensure the quality of education. Each program must establish an industry testing team to provide sound assessment and testing feedback on the results of vocational education. In order to improve the teacher-student relationship and to ensure the teaching of students according to their abilities, a new educational philosophy is needed. The ultimate goal of vocational education is to socialize and professionalize human development. To concretize and document the responsibilities and roles of vocational institutions, cooperative teams, and other relevant stakeholders to improve management and create a better environment for students' learning.

HZ received an A rating in the evaluation, and its teaching philosophy of serving development and promoting employment is worthy of consideration. HVE has the essential characteristics of applicability and technicality. Students should be instructed to use their strongest abilities in different environments to improve their adaptability and ability to solve problems effectively. While re-emphasizing skill cultivation, the education of the humanistic spirit should not be neglected. With economic globalization and social informatization, higher education institutions should also focus on students' personalized development and cultivate them into higher technical application talents with innovative abilities.

Table 4: Institutional HVE score.

Institution	HZ	TZ	SZ	KF	LY
Score	4.0122	3.5877	2.6773	3.3372	3.1264
Grade	A	B	C	B	B

5. Conclusion

HVE is the advanced stage of vocational education, and from the viewpoint of vocational education itself, employment orientation requires schools to cultivate technically skilled talents who can meet job requirements. HVE evaluation needs to strengthen the internal construction, which is also the way to improve the quality of education in higher vocational institutions. This paper follows the basic principles of combining theoretical analysis and empirical research, unifying qualitative and quantitative research, and integrates research methods such as literature research, questionnaire survey, interview, comparative study, and case study.

According to the actual situation of HVE development, 16 HVE evaluation indexes are proposed in four dimensions: teaching attitude, teaching content, teaching procedure, and teaching achievement. The optimal weights of each index were obtained based on the FAHP-BPNN method. the BPNN also achieved optimal results in the comparison of accuracy rates of different machine learning. Among the first-level indicators, the teaching achievement received the highest weight of 0.275. This fits with the practicality of higher education institutions. After the comprehensive evaluation, SZ school obtained a C grade, which indicates that its overall teaching quality is poor and needs to be enhanced according to the evaluation index system and the actual situation. the practical application of FAHP-BPNN comprehensive evaluation model in five typical higher vocational schools in Henan Province shows that the model can be widely used to guide the teaching quality improvement of higher vocational schools.

Due to the limitations of the research conditions, there are some aspects of this study that need to be improved. First, there is a need to improve the HVE evaluation indicators around the core of quality. HVE to change the type of education involving the enterprise society, there is a need to establish and improve the indicator dimensions related to the integration of industry-education and school-enterprise cooperation. Data collection and analysis can also be further improved to expand the range of respondents.

Data Availability

The labeled dataset used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] M. P. Ison, "The viability of tuition-free community college," *Educational Policy*, vol. 12, 2020.

[2] D. Criscuolo, K. Imamura, and I. Klingmann, "The specialist in medicines development (SMD) as a vocational program in pharmaceutical medicine: the Japanese and Italian experience [J]," *Frontiers in Pharmacology*, vol. 11, 2020.

[3] A. Colette and C. Marjolaine, "Perception by French students of the gendered nature of material artifacts studied in technology education," *International Journal of Technology and Design Education*, vol. 27, no. 1, pp. 1–18, 2017.

[4] A. G. Maria, "Prc education system development as an instrument for building a moderately prosperous society," in *Proceedings of the Annual Scientific Conference of the Centre-for-Political-Research-and-Prognosis of the Institute-of-Far-Eastern-Studies of the Russian Academy of Sciences (IFES RAS)*, Moscow, RUSSIA, March 2020.

[5] J. V. J. Chen, "Preparing students for class: a clinical trial testing the efficacy between multimedia pre-lectures and textbooks in an economics course [J]," *Journal of College Teaching & Learning*, vol. 13, no. 2, p. 37, 2016.

[6] Z. Hou, C. Yang, and Y. Lin, "Exploration of the cultivation mode of higher vocational education based on collaborative innovation," in *Proceedings of the 5th International Conference on Physical Education and Society Management (ICPESM 2017)*, Amman, JORDAN, March 2017.

[7] M. Lei and Y. Zhong, "Multiple intelligence evaluation system in the physical education of higher vocational colleges," in *Proceedings of the 6th International Conference on Applied Social Science (ICASS 2017)*, Singapore, May 2017.

[8] J. Li and I. Destech Publicat, "Research on strategy of career oriented higher vocational English teaching reform," in *Proceedings of the 3rd International Conference on Social Science, Management and Economics (SSME), Guangzhou*, PEOPLES R CHINA, May 2017.

[9] F. Zhou, "Reform and innovation of practice oriented higher vocational education curriculum [J]," *Agro Food Industry Hi-Tech*, vol. 28, no. 3, pp. 723–727, 2017.

[10] L. Miao, "Evaluation model and enhancement strategies for teaching reform capacity of art courses in higher vocational colleges," *International Journal of Emerging Technologies in Learning*, vol. 15, no. 18, pp. 124–138, 2020.

[11] X. Liu and H. Lin, "Research on innovation and entrepreneurship competency evaluation of teachers in industry-oriented higher vocational colleges [J]," *International Journal of Electrical Engineering Education*, vol. 1, 2021.

[12] O. K. Kalkan and T. Cosguner, "Evaluation of the academic achievement of vocational school of higher education students through artificial neural networks [J]," *Gazi University Journal of Science*, vol. 34, no. 3, pp. 851–862, 2021.

[13] H. Zhang, "Research on innovation education evaluation for vocational college based on improved BP neural network," in *Proceedings of the 2nd International Conference on Economics and Management, Education, Humanities and Social Sciences (EMEHSS)*, PEOPLES R CHIN, Wuhan, March 2018.

[14] B. Feng, "Dynamic analysis of college physical education teaching quality evaluation based on network under the big data," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 5949167, 13 pages, 2021.

[15] C. Liu, Y. F. Feng, and Y. L. Wang, "An innovative evaluation method for undergraduate education: an approach based on *BP* neural network and stress testing," *Studies in Higher Education*, vol. 47, no. 1, pp. 212–228, 2022.

[16] W. T. Wei, J. Y. Liu, Y. H. Liu, Y. Kang, R. Luo, and X. Zhang, "Evaluation index system of education quality for nursing professional degree postgraduate using the analytic hierarchy process," *Medicine*, vol. 100, no. 47, Article ID e27771, 2021.

[17] T. H. Liu and S. L. Yin, "An improved particle swarm optimization algorithm used for BP neural network and multimedia course-ware evaluation," *Multimedia Tools and Applications*, vol. 76, no. 9, pp. 11961–11974, 2017.

[18] Y. F. Chen, Z. B. Yu, and W. H. Zhao, "Based on optimization research on the evaluation system of English teaching quality based on GA-BPNN algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5949167, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9796124, 1 page
https://doi.org/10.1155/2023/9796124

Retraction

Retracted: Ideology in Sino-Foreign Cooperative Education: The Application of Big Data Mining Technology in the Work Conducted by Counselors

Mobile Information Systems

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Z. Shu, "Ideology in Sino-Foreign Cooperative Education: The Application of Big Data Mining Technology in the Work Conducted by Counselors," *Mobile Information Systems*, vol. 2022, Article ID 4062281, 12 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 4062281, 12 pages
https://doi.org/10.1155/2022/4062281

Research Article

Ideology in Sino-Foreign Cooperative Education: The Application of Big Data Mining Technology in the Work Conducted by Counselors

Ziqin Shu

Department of Geography and Tourism, Chongqing Normal University, Chongqing 401331, China

Correspondence should be addressed to Ziqin Shu; shuziqin@cqnu.edu.cn

Received 13 June 2022; Revised 2 August 2022; Accepted 9 August 2022; Published 7 September 2022

Academic Editor: Xuefeng Zhang

The globalization of today's world economy is developing in depth, and international education has gradually taken a new step in order to cultivate higher education talents who can adapt to the new context of development. Sino-foreign cooperative education is the main means for the country to carry out international education and foreign exchange, and it aims to cultivate complex talents who can adapt to the new international trade and exchange in the world. Due to the special nature of the student source and the complexity of the purpose of the school, the teaching concept and mode of Sino-foreign cooperative schools differ greatly from the traditional general higher education, which also puts forward strict requirements on the professional level of counselors. As an emerging technology, big data mining technology has a unique advantage in optimizing the processing and management of a large amount of data. The application of big data mining technology to the management of counselors in Sino-foreign cooperative education is important for improving the quality of education, establishing an efficient educational model of Sino-foreign cooperative education, and promoting the international modernization of national education.

1. Introduction

Ideological work has an important strategic position in the overall situation of national development, and firmly grasping the Party's dominant power and right to speak in the field of ideology is the top priority of the work of the whole Party. In recent years, the state and the government have issued many legal documents to strengthen the guidance in the field of ideological work, and the issuance of the Measures for the Implementation of the Party Committee's Responsibility System for Ideological Work has brought ideological work down to the grassroots organizations [1]. In the report of the 19th Party Congress, the Party made a complete elaboration on strengthening ideological work, indicating the direction of further consolidating the Party's leadership of ideological work in the new era. And the frontier of ideological work is higher education institutions, and colleges and universities, as an important stage to make people move from maturity to immaturity, have an important role in cultivating college students' core values and outlook on life [2].

As an important frontier zone of education for internationalization and modernization, Sino-foreign cooperative education has an extremely important position in the field of education. It follows the development trend of economic globalization and makes an important contribution to the construction of the community of human destiny. Compared with the traditional teaching mode, Sino-foreign cooperative education has cultivated a batch of comprehensive talents who can adapt to the new international trade in the world and have an international vision with the advantages of a unique faculty team, new specialties, and employment prospects. However, under its unique mode of schooling, some ideological problems have arisen. Since the training objectives of Sino-foreign cooperative education are different from those of ordinary higher education, the management level of counselors is more demanding. There are big differences between Chinese and Western cultures;

therefore, the daily behavioral habits and thinking directions of international students from different countries are often different. Besides, the values of college students also show diversified characteristics. It is worth being alert to the fact that Western culture is easy to infiltrate the values of Chinese students by using its hidden characteristics. During the critical period when students' worldviews, outlook on life, and values are not yet formally formed, the pluralistic values tend to "westernize" students' thinking patterns, and students may have deviations in their perceptions of social forms, cultural systems, and political directions. In addition, students are easily influenced by the "individual heroism" and "money worship" in Western culture and develop the concept of flattering foreigners and losing confidence in their own excellent traditional culture, resulting in "national nihilism" [3]. The ultimate consequence is that students have a weak sense of the collective concept, their ideals and beliefs are not firm enough, and their moral concepts are biased and slippery, which is contrary to the original purpose of the national cooperative education. Therefore, improving the effectiveness of counselor management is a key point in the work of improving the management level of cooperative schools.

At the meeting of the Political Bureau of the Central Committee of the Communist Party of China (CPC), in response to the implementation of the national big data collective learning strategy, President Xi Jinping pointed out that the development of big data is very fast today, and all people should carefully lay out and pay attention to the impact of big data on economic and social development; observe its development status and future direction through investigation; and build a digital China that contributes to national economic and social development and people's happiness and well-being [4]. According to the report of the 19th Party Congress, the primary position in the cause of people's livelihood is education, and accelerating the modernization of education is an urgent task for the country. Nowadays, the development of various social industries is being profoundly influenced and changed by big data mining technology, and school education is no exception. Higher education is an important stage to bring people from immaturity to maturity and plays a pivotal role in cultivating talents for national modernization. Therefore, promoting Internet + education and deepening the application of big data mining technology into the daily work of counselors is a key decision to solve the problem of insufficient and unbalanced development of education.

Big data mining technology not only has the commonly used functions of finding, comprehensively managing all kinds of information, collecting and analyzing data, and conducting statistics but also has the ability to process those lines of information that are hidden and useful in the database through data mining technology [5], providing valuable guidance for school management and teaching work and promoting the work of college counselors. At present, most colleges and universities only deal with data mining technology in low-level searching and simple analysis, so school leaders and the majority of research workers should focus on the key issue of how to mine the information data with high value and high guidance, which is also the common point in the application and research of data mining technology in various industries of the country. For example, at the end of each semester, the school organizes a survey on students' satisfaction with the work of counselors to consider the work of counselors. It is very meaningful for universities to make good use of the survey information so as to draw out valuable information for the management of school counselors and to make scientific analysis and judgment on each stage of student management and education through this information.

2. Research Background

2.1. Sino-Foreign Cooperative Education

2.1.1. Development Status. In 1986, the Sino-American Cultural Studies Center institution was established through the joint efforts of Nanjing University and Hopkins University, which is, in a strict sense, the first real Sino-foreign cooperative school institution in history. From 1995 to 2017, the number of Chinese-foreign cooperative institutions has grown from more than 70 to 2,572 under the approval of relevant state auditing departments, spreading across all types of teaching modes, covering 28 provinces across the country, and offering more than 200 majors. In 2015, the Strategic Alliance of Belt and Road Universities with the function of cultivating cross-border mobile talents was jointly constructed by 46 universities from East and West [6].

Sino-foreign cooperative education has been strongly supported by national policies, and the number of approved Sino-foreign cooperative education projects and institutions has been climbing in recent years, and Sino-foreign cooperative education projects are on a stable upward trend [7]. Except for 2018, when the number of school running projects decreased by 4 compared with 2017, all other years showed an increasing trend, and the overall level was on the rise. The number of schooling institutions also shows an upward trend in the general trend, especially in the first half of 2020 alone, there were five more schooling institutions than in 2019. Detailed data are shown in Figure 1.

By the end of 2019, among all the double first-class construction universities in China, 189 Chinese-foreign cooperative school running projects have been reached [8]. The top five provinces in the number of projects in the sub-provinces of Sino-foreign cooperative schooling institutions and projects in double first-class construction universities are Jiangsu, Beijing, Henan, Shanghai, and Heilongjiang (Figure 2).

The foreign cooperative institutions and projects of "double first-class" universities are mainly located in 16 countries and regions [9]. Among them, the cooperative objects of the undergraduate stage of Sino-foreign cooperative education projects in the "double first-class" construction universities are mainly concentrated in the United States, the United Kingdom, Australia, France, Germany, and other Western countries with developed higher education, and the number of cooperative education projects

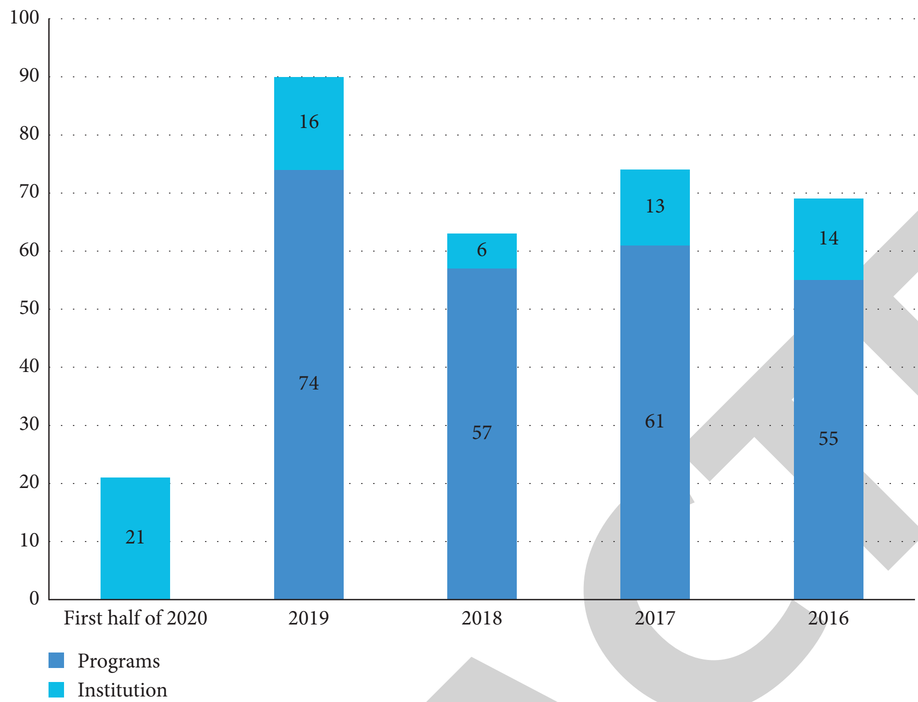


Figure 1: Number of approved Sino-foreign cooperative school projects and institutions in the past 5 years.

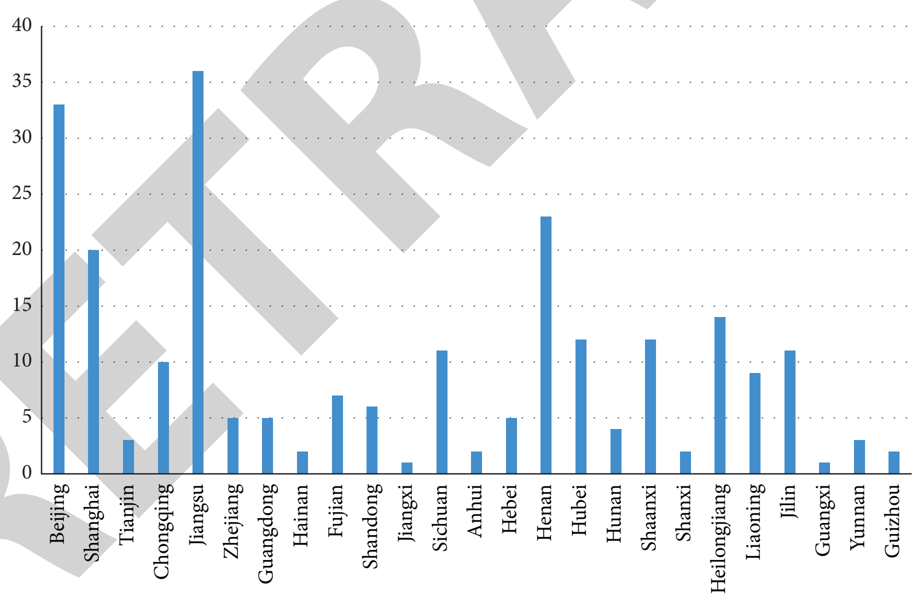


Figure 2: "Double first-class" construction universities of Chinese and foreign cooperation institutions and projects by province (as of 2019, excluding independent legal person institutions).

held with the universities of the above five countries accounts for about 10% of the "double first-class" construction universities. The number of cooperative education projects with universities in the above five countries accounts for about 80% of the total number of cooperative education projects at the undergraduate level of the "double first-class" construction universities. Other countries and regions with a high number of cooperative education projects are mainly concentrated in countries and regions with high level of higher education in Europe, America, and Asia, such as Ireland, Canada, Korea, etc. The detailed situation is shown in Figure 3.

In order to ensure the quality of education, the number of students enrolled in the first year of Sino-foreign cooperative education institutions is smaller than the approved enrollment scale but increases year by year as the education

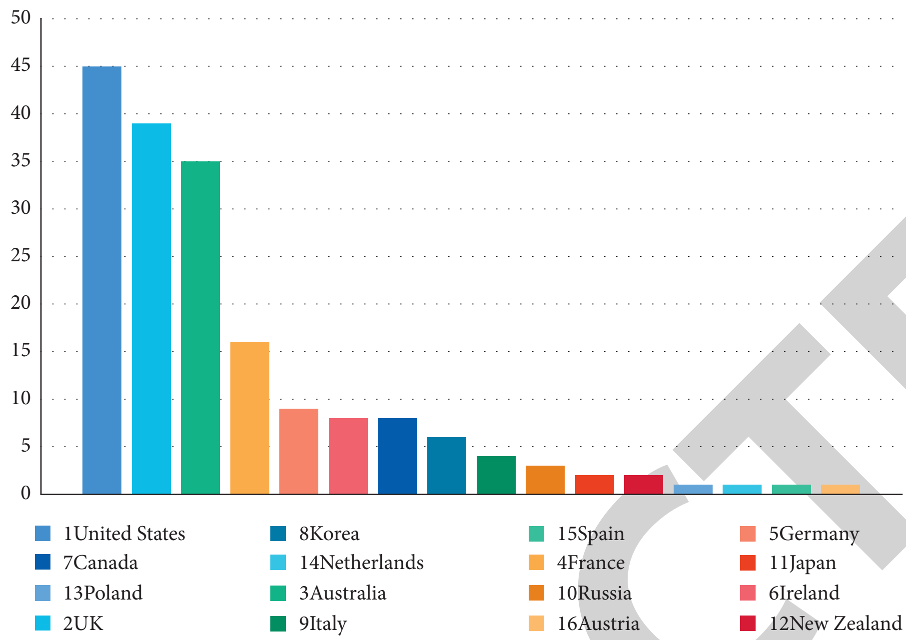


Figure 3: Distribution of countries and regions of cooperation between Chinese and foreign cooperative education programs and institutions at the undergraduate level of "double first-class" construction universities (excluding independent legal person institutions).

matures [10]. Among them, the enrollment of NYU Shanghai has increased from 151 in 2014 to 251 in 2020 (Figure 4).

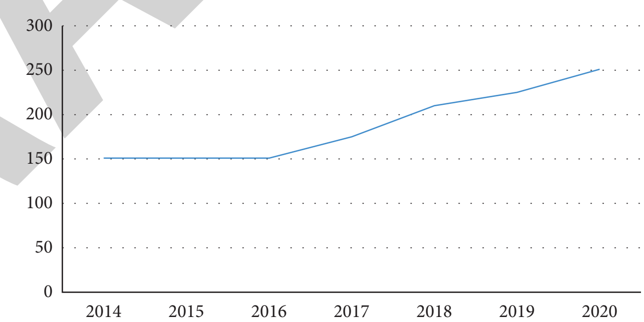


Figure 4: Trend of enrollment size of NYU Shanghai (2014–2020).

2.1.2. Problems and Challenges. Firstly, there are various challenges in terms of professional course exhaustion and student life adaptation. Under the national enrollment mechanism, the admission score line of Sino-foreign cooperative schools is usually below the undergraduate line, which means that the vast majority of students' academic level is weak, their learning ability needs to be improved, and their level is uneven, which prevents centralized and effective unified management. Since students have to face the pressure of language score assessment when they go abroad, some of them need to devote a lot of time and energy to study English, which makes them go off the beaten track and often use the leave of absence from school to improve their foreign language skills. In the long run, their enthusiasm for learning basic and specialized courses will be extinguished, and it is difficult for them to adapt to the diversified assessment mode of the school, and they are unable to effectively regulate themselves under the double pressure of specialized and foreign language courses. In addition, the tuition fees of Sino-foreign cooperative schools are often high, so students' families are generally well-off. Like the flowers in the greenhouse, they are highly cared for by their families since they are young, and it is difficult for them to adapt well to the high-pressure study life, so they tend to show the disadvantages of discipline breakdown and inability to take care of themselves. In addition, some students do not know enough about foreign folk customs and legal system and their seemingly insignificant actions often bring safety hazards to themselves, making them unable to integrate into a foreign environment better.

Second, the management of cross-border students is not sustainable and effective. Compared with other models, Sino-foreign cooperative education models are quite different, as they generally adopt the "2 + 2" and "3 + 1" models, in which students study in China for 2 to 3 years and then complete the rest of their studies abroad to obtain dual degrees from domestic and foreign schools [11]. During the school period, students' ideological and political education work is not fully prepared and in place, and the key period to effectively carry out patriotic collective education and ideal and belief education is exactly when students are studying in China, which makes students spend most of their time and energy on studying professional courses and foreign language courses, thus causing the lack of ideological and political education work. This brings a great challenge to the management of counselors, and how they can optimize the ideological and political education work of college students within the specified time becomes the most important work. Because of the spatial distance and global time difference, the counselors cannot manage the overseas students uniformly and effectively, which seriously restricts the cultivation and guidance to the overseas students.

Based on the above differences, the work in the field of ideology in colleges and universities with Sino-foreign cooperation is often more complicated, which puts forward more precise and special requirements for the management level of counselors, who need to formulate targeted and effective management measures according to the special characteristics of colleges and universities with cooperation. They should make correct analysis and judgment on the ideology of universities run by Sino-foreign cooperation, constantly strengthen the political function of the party organization of universities under the premise of following the orientation of socialism in running schools, precisely position the direction of running schools and talent cultivation mechanism in the reform of universities run by Sino-foreign cooperation, effectively carry out ideological work, and do it according to the time and advance with the trend.

3. Materials and Methods

3.1. Big Data Mining Technology

3.1.1. Meaning and Development Status. The researcher studies the object and content-related knowledge from the rhesus data, which are in various forms and can be rules, concepts, etc. This information is not obvious and predicted in advance, but it must be useful information. And data mining is the decision support process of finding the object of study in a collection composed of many data. Data mining is a process in which humans and machines are constantly exchanging with each other and there are numerous steps closely connected [12]. Some of the key steps are: asking a question, selecting data, organizing and determining data, constructing a model, judging, and interpreting (Figure 5).

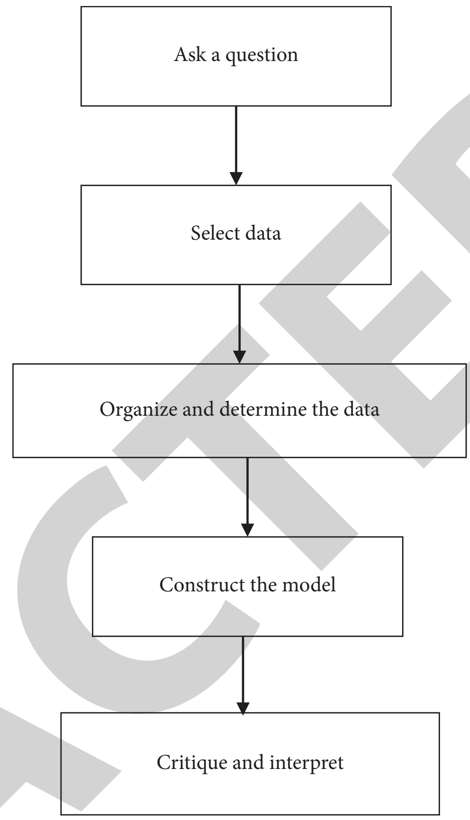


Figure 5: Data mining steps.

All these steps are subject to a continuous iterative process of proceeding. Data mining is not a single discipline, but it is a blend of many disciplines. Previously, the application of data was relatively superficial, but now, through data mining, researchers can fully exploit the empirical and theoretical knowledge in it to provide decision support, which is the task of data mining. The obtained knowledge can supplement and improve the existing knowledge system as well as provide support and assistance to decision makers in their decision-making process, and can also be stored as new knowledge in the relevant stored knowledge institutions. Data mining technology has not appeared in the limelight only in recent years but has been studied abroad for many years. With the rapid development of data mining technology and decision support systems, many industries represented by the retail industry have achieved high-profit growth, which has attracted the attention of many universities and research institutions, and they have invested a lot of money and energy in deep mining development and research. Data mining technology in China, on the other hand, started late and is still in its infancy. A large part of the work lacks comprehensive system integration design and is only limited to the top of local algorithm design. At present, only finance, banking, GIS, and other fields have initially applied data mining technology, which is inseparable from the lack of core technology [13].

3.1.2. Application. Data mining, as a deep method of data research and analysis, is a process of affirmation from the proposal of a program, and if this technology is put into the work assessment and examination system of counselors in cooperative schools, it can improve the management utility level of counselors and make the work achieve twice the result with half the effort. For example, the data on the "quantitative assessment form of counselors' work" can be used by data mining [14] to make a comprehensive performance evaluation of the management level of counselors in a university, listing which tasks counselors have done well and which tasks are lacking in management. In the past, colleges and universities usually used to look up information data in the form of information on the effectiveness of counselors' work. Now, the researcher can use data mining technology to process the information and data on counselors' work effectiveness and then make good use of the data and apply it to the management of counselors. The data mining process is shown in Figure 6.

Step 1: Determining the object and purpose of data mining. Very definitely clarifying the problem and discovering and finding the target of data mining is the

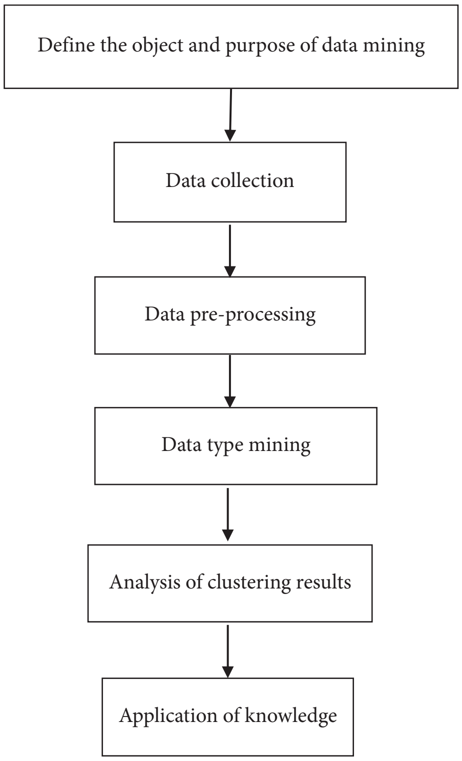


Figure 6: Clustering data mining process.

first and very important step in the process. The final result cannot be predicted, but the research problem can be anticipated in advance through data mining.

Step 2: Data collection. The data collection is more laborious and requires a lot of time and effort. This requires the collector to carefully collect all kinds of data and information in the usual educational activities, some of which can be directly obtained, while some data must be discovered through research and study.

Step 3: Data preprocessing. This step is to convert the collected data into a data model, which is analyzable and based on different algorithms, so the requirements of different kinds of algorithms for data models are very different.

Step 4: Data clustering mining. The data model is split into multiple groups by the data clustering mining technique. These words are split based on the magnitude of similarity; the greater the similarity, the greater the likelihood of being grouped together. This step is the selection and implementation process of the clustering algorithm and the input process of the data model.

Step 5: Analysis of clustering results. This process is the selection judgment and analysis of the information after clustering data mining and the judgment of the results of multiple group attributes.

Step 6: Application of knowledge. This is the final step, which is to apply the useful information obtained in the previous steps to the management activities of the counselors' education so that the counselors can draw useful conclusions to improve their management teaching and form a practical and effective management policy. This is also the purpose of the thesis research.

3.2. Design Scheme

3.2.1. Identifying Data Mining Objects and Targets. The researcher collected and compiled 120 "quantitative assessment forms of counselors' work" from a university with Sino-foreign cooperation and tried to answer questions like what is the overall management level of counselors in this university? What work is done well and what work is not done well in management? We hope that by processing these data through data mining technology, we can draw some important conclusions and bring practical guidance for teaching and management.

3.2.2. Data Collection. Through the data collection of 120 "quantitative assessment forms of counselors' work" of a Sino-foreign cooperative university, the researchers were well prepared for the information sources to be used in the data mining technique.

3.2.3. Data Preprocessing. The "quantitative assessment form for counselors" has several evaluation indicators for counselors who are required to be evaluated: "to reward and punish each student clearly and treat each student objectively and fairly," "to grasp and understand the situation of poor students and do a good job of grant allocation," and "to do a good job of grant allocation." The indicators are: "rewarding and punishing, treating every student objectively and fairly," "understanding the situation of poor students and doing a good job in allocating grants," "doing a good job in granting student loans," and many other indicators. Based on the performance level of the counselors, the counselors are divided into five grades, which are "excellent, good, average, poor, and very poor."

In the work assessment quantification table, there are a total of 15 items of assessment levels. How to synthesize these data into a template for cluster analysis through data mining techniques was a challenge for the researchers to transform the data. In this regard, the researchers chose to reorganize and integrate the information data in the quantitative job evaluation form in four areas: "management attitude," "management ability," "management style," and "management performance."

The "management attitude" corresponds to the table's "clear rewards and punishments, treating each student objectively and fairly," "appropriate behavior and good personal qualities," "friendly and harmonious relationship with students," and "acting objectively and fairly, and being honest."

The "management ability" corresponds to the table's "understanding of the situation of poor students and the distribution of grants," "using a special approach to special groups of students and meeting the specific needs of special students," "not to protect shortcomings, in accordance with the school regulations to deal with disciplinary students seriously," "in the evaluation of merit and awards to achieve a high level of transparency and adhere to the principles of fairness, impartiality, and openness," and "have strong organizational management skills, actively organize strong organizational and management skills, and actively organize and mobilize students to carry out various work activities."

The "management style" corresponds to the "care for students' study, work, and life, at least twice a week to check the dormitory," "insist on weekly inspection of students' dormitory hygiene," "talk with each student at least once a year," "actively participate in and supervise students' morning jogging activities," and "complete procedures for school scholarships, grants, and student loans and issue them." The procedures for granting scholarships, grants, and loans are complete and in place.

"Management performance" corresponds to the table's "conducts informative and effective class meetings," "understands students' learning situation in depth during class," and "can give reasonable suggestions to students about their learning and criticize and correct students' deficiencies."

The attribute values for these four attributes can be approached in several ways.

First, the researchers converted each individual assessment rating in the table into a more easily understood and calculated data type. The five levels of the appraisal scale, "excellent, good, fair, poor, and very poor," can be regarded as ordered data types in statistics, and they are arranged in an order that has a special meaning.

The values of the four attributes can be calculated using the arithmetic mean of the items they contain.

"Managerial attitude" = (objectivity and impartiality + rapport with students + decent behavior + integrity)/4.

"Management ability" = (mastery of understanding the situation of poor students + meeting the specific needs of special students + serious handling of disciplinary students in accordance with school regulations + transparency and openness in the evaluation of merits, prizes, and awards + strong organizational and management skills)/5.

"Management style" = (achieving at least two or more dormitory inspections per week + insisting on weekly inspections of student dormitory hygiene + talking to each student at least once per academic year + actively participating in and supervising students' morning jogging activities + complete procedures for scholarships, grants, and student loans and their disbursement)/5.

"Management performance" = (carrying out informative and effective class meetings + in-depth understanding of students' learning conditions in class + reasonable learning suggestions to students, criticism, and correction of their shortcomings)/3.

As a result of the above work, the 15 items in the quantitative work evaluation form as well as an after-class mastery and comments and suggestions section were all grouped into four attributes: "management attitude, management ability, management style, and management performance." The following is the process of the researcher's analysis and study of data from 120 quantitative job evaluation forms using big data mining technology methods.

3.2.4. Data Clustering Mining. Using the preprocessed data sample information statistics, the researchers collected the sample data to be used in the analysis process as shown in Table 1.

(1) Expected results obtained: the data sample information listed above was classified into three levels: better, moderate, and poor. The mining technique of cluster analysis is used to further derive the proportional distribution of the sample data information distribution in relation to the above three levels of ranking. The researcher tried to answer the question of "how good or bad is the overall management of counselors" and "what is working well? What is not working well enough?," etc.

(2) The selection of mining methods: From the clustering method, it can be seen that the algorithm based on density, grid, and model is not strong for these small and medium data types, so the classical algorithm k-means and k-central value algorithm in the clustering method can be used to study and analyze this type of data mining. According to the definition of the algorithm, k can represent the number of categories needed for the study. The most basic idea of these two classical algorithms is to divide n objects into k classes so that objects in the same class have high similarity and objects in different classes have low similarity to each other. The mean value in each class is the reference value for the k-means algorithm, while the k-centroid algorithm uses the point object at the center of the class as the reference value for calculating the dissimilarity.

The k-centroid algorithm has an advantage over the k-means algorithm when there are isolated or outlier points in the data object, which is related to the fact that centroids are less susceptible to isolated points, while means are susceptible to isolated points [15]. However, the k-centroid algorithm is more complicated compared to the k-means algorithm. Assuming that the complexity of the k-means algorithm is represented by X, it is calculated that $X = nkt$ (n represents the number of classes and t represents the number of iterations); assuming that the complexity of the k-centroid algorithm is represented by Y, when the k-centroid algorithm performs one iteration (an iteration is an activity that repeats the feedback process, usually to approximate the desired goal or result), the complexity of the algorithm afterwards has become $Y = k(n - k)^2$.

Table 1: Sample data for clustering.

Management attitude	Management ability	Management style	Management performance
0.65	0.6	0.56	0.58
0.65	0.6	0.56	0.58
0.35	0.35	0.31	0.33
0.76	0.76	0.68	0.75
0.8	0.8	0.81	0.83
...	...	...	...

$$X = nkt\,(3-1),$$
$$Y = k\,(n-k)^2\,(3-2), \tag{1}$$

where X represents the k-means algorithm complexity and Y represents the k-centroid algorithm complexity.

Based on the above sample data, four types of attributes, namely "management attitude, management ability, management style, and management performance," are obtained by data transformation. The data used are the arithmetic mean of items 3, 4, and 5 in the quantitative work appraisal table, so it is unlikely to generate isolated points or outlier points. Based on the above analysis, the researchers selected the k-means algorithm to perform data mining.

(3) Improvement of data samples: When executing the k-means algorithm, in order to get the global optimal solution as much as possible, it is necessary to try to reduce the skew of the sample data, which requires the researchers to improve the sample data. To this end, in the initial state, the top three data samples are classified as the central case of clustering, and after taking further steps to improve the k-means algorithm by combining the desired results with the integrated proportional distribution on the three rating levels described earlier, the samples that can represent the three levels are derived to replace the first three sample data that have been defined, thus serving as the center of clustering at the beginning, to the greatest possible extent possible to reduce the skewness of the data and the number of iterations performed [16]. The data as samples for the three classes are shown in Table 2.

(4) The flow of the algorithm is shown in Figure 7.

In the k-means algorithm, the function LoadPatterns (Char $*$ fname) is the process of loading the sample data information into the program. The main purpose of this function is to read out the relevant information from the database file km.dat and convert the data in this file into the sample Patten[i][j] array.

The function InitClusters() functions as the initialization process of the class centers, and the function starts with the first K data in the data sample as the starting class centers.

The function RunKMeans() functions as the main procedure of the algorithm, and the main procedure is to compare the distance of all objects with the center of each cluster. Then the objects are divided into the classes with the nearest class centers, and the cluster centers are calculated according to the redistribution, and if the class centers do not change, then the clustering process is terminated.

The ConvFlag is used as an identifier for whether the clustering is complete or not, the function ShowCenters() represents the cluster centers described by the algorithm, and the function ShowClusters() represents the identifier number of the samples described by the algorithm [17].

4. Results and Discussion

4.1. Results. In the analysis of the data obtained after data conversion using the k-means algorithm, 120 sample data of the work appraisal quantification table, which contains three standard samples representing the degree, namely better, moderate, and worse, and another 117 sample data were obtained by data conversion of the work appraisal quantification table. In this, all the sample data have four major aspects, namely management attitude, management ability, management style, and management performance. The data mining analysis was conducted on these four aspects, and the initial k value was set to 3, and the final mining results were obtained as shown in Table 3.

4.2. Analysis and Discussion. From the data in the table, it can be seen that the final proportional distribution of the sample data contained in each class is characterized as follows.

The first class, that is, the better group, has a total of 36 samples, and after removing a standard sample defined in advance, that is, "better," there are 35 samples left, accounting for $35/117 = 30\%$ of the 117 samples obtained after data transformation.

The second category, the moderate group, has 74 samples in total, and after removing a predefined standard sample, that is, "moderate," 73 samples remain, accounting for $73/117 = 62\%$ of the 117 samples obtained after data transformation.

The third category, that is, the poor group, has a total of 10 sample data, and after removing a standard sample defined in advance, that is, "poor," there are 9 sample data left, accounting for $9/117 = 8\%$ of the 117 sample data obtained after data transformation.

In order to estimate the final results of the data mining, the researcher took a whole group sampling method to randomly select a total of 248 students of a certain major from the university and then investigated the comprehensive quantitative scores of the ten activities organized by the students of the major. By setting the total score at 100, the

Table 2: Sample data representing the three ranks.

	Management attitude	Management ability	Management style	Management performance
Four categories of attribute values indicate a better situation	0.75	0.75	0.75	0.75
The four types of attribute values indicate the moderate situation	0.5	0.5	0.5	0.5
The four types of attribute values indicate the worse case	0.25	0.25	0.25	0.25

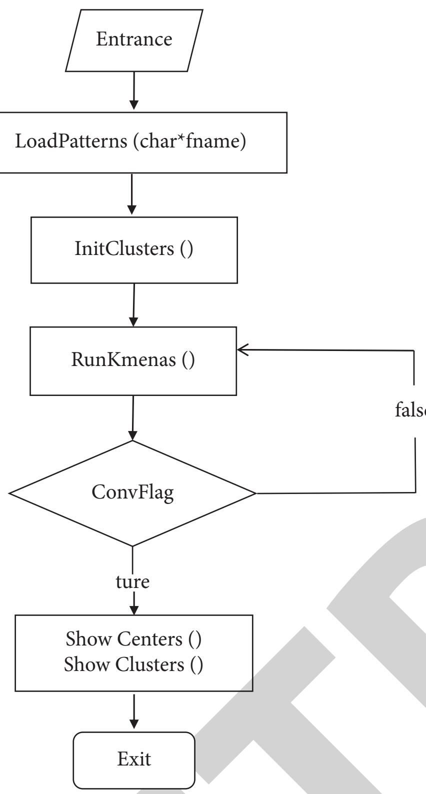


Figure 7: Flowchart of algorithm implementation.

three rating levels were below 60, 60 to 80, and above 80. The final classification results are shown in Figure 8.

When comparing the composite quantitative scores of the two categories (above 80, 80 to 60 (including 80 and 60), and below 60) with the levels used for clustering (0.75, 0.5, and 0.25), the percentage of branches in each category has increased, and the above analysis shows that the percentage of the three categories of scores is 21%, 68%, and 11% with the three categories of clustering. The sample proportions of the three categories of score bands, that is, 21%, 68%, and 11%, coincide with the sample proportions of the three categories of clusters, that is, 30%, 62%, and 8%.

The above results verified that such a data mining model adopted based on the quantitative table of counselors' work assessment is a practical and effective model and also showed that data mining technology can be applied to the management of counselors, which brings significant reference and guidance significance for improving the management level of Sino-foreign cooperative universities and for improving and developing the management and education of college counselors.

The results of the data mining were recombined, and the standardization degree of the three rating levels was carefully considered. After careful comparison, it was found that the attributes of the three rating levels directly improved compared with the well-defined standard sample, except for the attribute value of 0.74 for the first-class management method item, which was lower than that of 0.75 for the well-defined standard sample. This indicates that the overall scores of all classes have been improved. The overall score of each individual item was then used for further analysis and judgment. Due to the insufficient sample size, the researcher used a weighted approach for each item. The overall scores for each of the four individual items in the following categories were as follows:

"Management attitude" $= 0.77 * 30\% + 0.61 * 62\% + 0.31 * 8\% = 0.634$

"Management ability" $= 0.77 * 30\% + 0.57 * 62\% + 0.31 * 8\% = 0.6092$

"Management style" $= 0.74 * 30\% + 0.54 * 62\% + 0.28 * 8\% = 0.5792$

"Management performance" $= 0.79 * 30\% + 0.56 * 62\% + 0.30 * 8\% = 0.6082$.

The overall scores for the four items above were ranked from lowest to highest: management style (0.5792), management performance (0.6082), management competency (0.6092), and management attitude (0.634). These data all have one thing in common, that is, all scores are greater than 0.5, that is, they are all in the middle to upper level, which indicates that the overall management level is also in the middle to upper level, which is a relatively good result. Moreover, it is easy to find that the difference between the highest score of "management attitude" and the lowest score of "management method" is only 0.0548, which is equivalent to a difference of 5.48%. This means that while coordinating the work of each project, they should focus on strengthening the "management style" (caring for students' study, work, and life and checking the dormitory at least twice a week), "insist on weekly inspection of students' dormitory hygiene," "talk with each student at least once a school year," "actively participate in and supervise the students' morning run." "The requirements and management of the school scholarships, grants, and student loans are complete and the procedures for granting them are in place. Relatively speaking, the student work department has a "management attitude" (rewards and punishments are clearly defined, and each student is treated objectively and fairly), "good behavior and good personal qualities," and "friendly and harmonious relations with students." We do not need to

Table 3: Clustering results.

	Management attitude	Management ability	Management style	Management performance	Sample size number
Class 1 (better)	0.77	0.77	0.74	0.79	36
Class 2 (medium)	0.61	0.57	0.54	0.56	74
Class 3 (poorer)	0.31	0.31	0.28	0.30	10

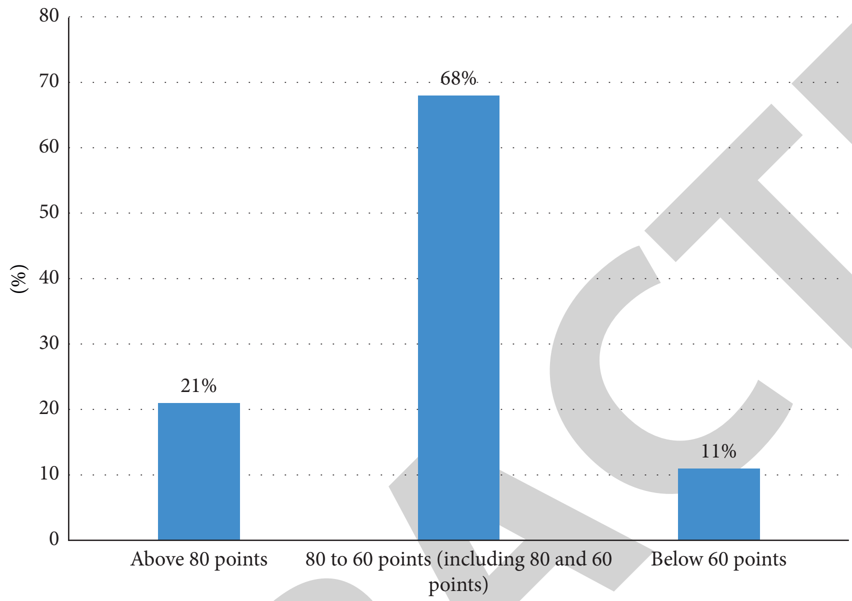


Figure 8: Comparison of analysis results of comprehensive quantitative score situation.

focus so much on "management attitude" (rewarding and punishing students objectively and fairly, treating each student objectively and fairly, "behaving well and having good personal qualities," "having a friendly and harmonious relationship with students," and "acting objectively and fairly, and being honest"). For "management performance" ("conducts informative and effective class meetings," "understands students' learning conditions in class," "gives reasonable suggestions to students about their learning," and "criticizes students' shortcomings," and criticizing and correcting students' deficiencies) and "management skills" ("understanding the situation of poor students and doing a good job of allocating grants," "adopting a special approach to special groups of students and meeting the specific needs of special students," "not to protect shortcomings, in accordance with the rules of the school to deal with disciplinary students seriously," "in the evaluation of merit and awards to achieve a high level of transparency and adhere to the principles of fairness, impartiality, and openness," and "have strong organizational management skills, actively organize strong organizational management skills, and actively organize and mobilize students to carry out various work activities") can be coordinated and balanced, rational and organised activities. The most important thing is that as counselors of universities with Sino-foreign cooperation, they should correctly grasp the relationship between these four aspects and adhere to the principle of unity of two points and focus, not only should they not look at these four aspects without prioritizing them, but also should not focus on one of the items and ignore the work of other items.

In addition, using big data mining technology to establish the evaluation system of student management in colleges and universities with Sino-foreign cooperation is a good choice for improving school management information [18] and reducing the work pressure of counselors. The system can provide counselors with data, materials, background, and various information needed for management; help clarify management goals; screen problems; establish various evaluation models; develop a series of alternative solutions and be able to evaluate and screen various solutions; use artificial intelligence to make analysis and comparison of information; and provide reasonable and effective management solutions for managers. As an emerging information technology industry, student evaluation management system has advantages that other industries cannot match. It can greatly reduce the work pressure of counselors engaged in various low-level information processing and analysis, allowing counselors to focus more on the work that requires the most management wisdom and experience, effectively solving the problem of working blindly but getting half the result with twice the effort.

The low efficiency of student management and evaluation, which is detached from reality and lacks scientific rationality, is a common phenomenon in today's Sino-foreign cooperative schools, and the lack of strong support for all aspects of student management and evaluation activities seriously limits the depth of cooperative school management reform. The data mining technology can now be applied to the management of counselors, which has to be said to be a forward-looking and far-sighted change activity for Sino-foreign cooperative universities. Applying data mining technology for student management evaluation system to the current student management in universities not only makes it easier for counselors to cope with various work tasks but, more importantly, counselors will get various unexpected gains and discoveries in this mining process.

Although big data mining technology has so many advantages in the application of counselor management in Sino-foreign cooperative universities, there are still many shortcomings that need to be improved. These include flaws in data management, insufficient attention to big data technology in schools, and also inadequate mathematical modeling techniques [19]. At the same time, the data among university systems should be reasonably standardized and unified, and standardized requirements should be formulated to prevent the occurrence of the problem of poor data circulation and inability to share among departments when the big data mining technology is used to analyze the data. In addition, for the processing of complex data types, it is unrealistic for universities to expect one system to mine all types of data due to the diversity of data types and different objectives of mining, which requires universities to build a targeted data mining system for different types of data. In addition, for data stored in databases, the data may reflect situations such as noise exceptions or incomplete data objects. These objects may mess up the analysis process and cause the knowledge model constructed from the data to be incompatible [20]. This requires universities to have cleanup methods for dealing with data noise and incomplete data as well as outsider mining methods for discovering and analyzing exceptions.

5. Conclusion

In summary, with the booming development of Chinese-foreign cooperative schools, the expectations of all parties in the society for the schools are getting higher and higher. And the various aspects of the ideological existence of Sino-foreign cooperative schools have challenged the management requirements of counselors in various aspects, which makes it difficult for universities to carry out reasonable and effective management of students according to traditional methods. At this time, an emerging technology can be used to discover the hidden patterns and information in the data through data mining to provide effective support and assistance to the management.

This article verifies through research that data mining technology is of great value to counselors in their management work. It is the result of human wisdom and practice, breaking the inherent thinking of traditional teaching and providing directions and methods for building models that are impossible to obtain in traditional teaching. It can provide scientific decision-making solutions for counselors of Sino-foreign cooperative schools and reference models for school education. Applying this technique to the management of counselors is of great significance both for the counselors themselves and for the school as well as for other types of universities to learn from and refer to.

In view of the limitation of research time and personal ability, this article has some incomplete points. For the application of big data mining technology in the field of Sino-foreign cooperative education, there are still many issues to be further explored.

Data Availability

The labeled dataset that supports the findings of this study is available from the author upon request.

Conflicts of Interest

The author declares no conflicts of interest.

Acknowledgments

The research was supported by the achievements of the Humanities and Social Sciences Research Project of the Ministry of Education in 2021 "Research on the Ideology-Related Work at College by Political Assistant in Chinese-Foreign Cooperation in Running Schools" (project no. 21JDSZ3023).

References

[1] L. I. Zhipeng, "The innovation strategy of Chinese-foreign cooperation in running schools in China's double first-class universities," *International Journal of Education and Teaching Research*, vol. 3, no. 2, 2022.

[2] Z. Chen, "Learning from and breaking through: the current situation, problems and countermeasures of sino-foreign cooperation in higher education," *International Journal of Social Science and Education Research*, vol. 5, no. 2, 2022.

[3] Y. Chen, "Students' Learning Experiences in the Foundation Year of a Sino-Foreign Joint University," in *Proceedings of the 2021 4th International Conference on Humanities Education and Social Sciences (ICHESS 2021)*, Chongqing, China, October 2022.

[4] Q. Wang, "Strategic Analysis of Sino-Foreign Higher Education Cooperative Universities Based on SWOT Model," in *Proceedings of the 2021 4th International Conference on Humanities Education and Social Sciences (ICHESS 2021)*, Chongqing, China, October 2022.

[5] W. Yang, "The present situation, problems and countermeasures concerning sino-foreign joint education programs in higher education in the new era," *Scientific Journal Of Humanities and Social Sciences*, vol. 3, no. 12, 2021.

[6] Z. Li, "A brief analysis of the ways to improve the quality of foreign teachers' specialized courses for junior college students in sino-foreign cooperation," *International Journal of Higher Education Teaching Theory*, vol. 2, no. 4, 2021.

[7] X. Zhu and A. Feng, "ICC nurturing strategies in the context of sino-foreign institutes," *Theory and Practice in Language Studies*, vol. 11, no. 9, pp. 1072–1079, 2021.

[8] Q. Wang and Y. Yin, "Analysis on the cultivation path of national security awareness of the students of sino-foreign cooperative education major in the new era," *International Journal of Higher Education Teaching Theory*, vol. 2, no. 2, 2021.

[9] X. Wang and Y. Yin, "A study on student management of sino-foreign cooperative education majors from the

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9858357, 1 page
https://doi.org/10.1155/2023/9858357

Retraction

Retracted: Psychological Analysis of Athletes during Basketball Games from the Perspective of Deep Learning

Mobile Information Systems

Received 22 August 2023; Accepted 22 August 2023; Published 23 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Q. Meng, "Psychological Analysis of Athletes during Basketball Games from the Perspective of Deep Learning," *Mobile Information Systems*, vol. 2022, Article ID 4319437, 9 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 4319437, 9 pages
https://doi.org/10.1155/2022/4319437

Research Article

Psychological Analysis of Athletes during Basketball Games from the Perspective of Deep Learning

Qinghui Meng

Shanxi University, Taiyuan, Shanxi 030006, China

Correspondence should be addressed to Qinghui Meng; 202023705033@email.sxu.edu.cn

Received 5 July 2022; Revised 15 August 2022; Accepted 20 August 2022; Published 6 September 2022

Academic Editor: Jiafu Su

Due to the influence of psychological factors during sports, basketball players often change the shooting rhythm in sports, which will reduce the shooting rate and directly affect the training effect. Therefore, we need to analyze the basketball shooting process of some athletes, which is not a complete negation of traditional training methods, but a psychological analysis of some athletes' shooting in the course of competition, so as to complement the feedback and training information after the game. With the great improvement of computer computing power, deep learning natural language processing can help people analyze and solve previously unsolvable problems in production and life. The psychological emotion of adolescence has a great influence on the study and life of middle school students. At present, the unified monitoring and analysis of the daily life of middle school students need not only a lot of manpower, but also slow speed. If the psychological problems are not found in time and feedback is given, it will cause a series of adverse effects on individual athletes. The neural network model based on deep learning can process students' daily mass text information quickly and accurately, and then give comprehensive judgment, which is a good solution. This paper applies the neural network algorithm of the Bi-LSTM model and CNN model to study the text data, and finally has 95.55% and 90.03% accuracy in the psychological analysis experiment, which provides a feasible solution to solve the batch rapid analysis of psychological changes reflected in the daily text of athletes during basketball. Some suggestions are put forward on how to strengthen and improve the psychological quality of college basketball players and their ability to bear pressure and difficulties.

1. Introduction

With the development of market economy in modern society and the increasing improvement of material living standards, modern people begin to yearn for a fitness lifestyle, and sports gradually become a fashion [1]. Among many sports, basketball has become the most popular sport in China's sports industry. According to the white paper on China's Basketball Industry 2018, both the number of basketball courts and the number of basketball fans are far ahead of other ball sports, and basketball has become a fashion [2]. As a team project, basketball can improve people's physical quality; especially in the youth group, basketball can help them control their weight, cultivate teamwork ability, optimize their physical quality, correct sports attitude, and spread the competitive sportsmanship of basketball [3].

American expert Gruby believes that "for junior and intermediate players, 80% of the success of modern basketball is due to the biological power, 20% to the soul, and for high-level players, 20% to the biological power and 80% to the soul." Therefore, people began to fully understand the impact of psychology on the shooting rate. With the increasing improvement of basketball technology, psychology has become more critical. Therefore, this chapter gives some opinions on the psychological analysis and methods of basketball players' shooting skills for colleagues to explore.

With the increasing influence of NBA in the world, basketball is attracting more and more people's attention and love [4]. With the rapid development of basketball, basketball tactics have become more complex [5]. In today's information age, basketball is developing toward high-speed, high-skill, high-altitude competition and fierce competition. In order to adapt to the rapidly changing situation in the arena, the current basketball technology will be rapidly changed to the trend of high-speed evolution. The cross-penetration and interaction of basketball attack and defense and basketball confrontation technology have different psychological effects on athletes. On the basketball court, the game becomes more intense and exciting. Through several years of development and exploration, the current basketball technology has not only been limited to physical quality, but also developed into confrontation in psychology, will, combat skills, and so on. The strong physical quality is the cornerstone, and the advanced fighting technology is the guide. However, the impact of psychological quality on basketball players in basketball matches cannot be underestimated [6]. If the psychology is not perfect, it will be difficult to show your ability in the competition; athletes can often outperform if they are mentally tough. In the world, many countries attach great importance to psychological training of basketball players.

Basketball is a comprehensive and direct competition, strong objective dependence on the collective project [7]. The process of the game is complex and changeable, and fierce confrontation and the psychological changes of basketball players are also very complex. The outcome is unpredictable and can change quickly. Therefore, basketball players need to have a strong psychological quality to cope with any changes on the court, so that it is not affected [8]. In a large number of practice matches, basketball players not only consume a lot of physical energy, but also consume a lot of psychological ability in their competitive activities [9]. With the increasing improvement of the technical level of modern basketball in China, the overall technical level and comprehensive strength of Chinese basketball players have significantly improved, but the differences between basketball players are gradually decreasing [10]. In this case, the psychology of tennis players is very important. In foreign high-level basketball matches, the main assessment is not only the physical fitness and combat skills of tennis players, but also a psychological competition [11]. At these times, especially at the end of the competition, the decisive role is usually the psychological quality of the players. In the basketball game, it is unpredictable on the court, which makes the psychological activities of basketball players in the process of participation constantly change [12]. Through the research on the game and relevant data, basketball players generally have two mental states: one is conducive to the players' technology display and application in the game, and the other is not conducive to the players' technology application and use in the game [13]. This paper provides a feasible solution to solve the problem of batch rapid analysis of the psychological changes reflected in the psychological daily texts of athletes during basketball. Put forward relevant suggestions on how to strengthen and improve the psychological quality and ability to withstand pressure and difficulties of college basketball players. Therefore, "in order to improve the shooting rate of basketball players to enhance the psychological stability of shooting ability and increase the degree of exciting matches, the study of psychological stability factors affecting the shooting rate is very key."

2. State of the Art

2.1. Analysis of Psychological Factors Affecting Shooting in Basketball Match

2.1.1. Lack of Preparation for the Competition and Rush to Participate. The athletes are not well prepared for the competition and do not analyze the strength of the opponent [14]. Mental preparation does not form accurate consciousness, so attention cannot be focused. Central nervous system disorder is caused by excessive excitement or inhibition, sudden or untimely handling of events, tight muscles, busy movements, and fear of losing.

2.1.2. Lack of Understanding of the Other Party before the Competition. Without understanding the other side, you may make unnecessary mistakes in the game [15]. For example, psychological fear that the other party is better than themselves, or contempt for the other party, thinking that the other party is poor and vulnerable, helps to ignite their fighting spirit [16]. Athletes are too eager to win the game, and if there is a score behind, they cannot control their emotions, which needs to be avoided. This has resulted in the distortion of the movement, the previously formed movement habits, the rhythm changes, and out of control.

2.1.3. Lack of Training and Psychological Pressure. The training at ordinary times is lax, often unable to reach the normal training quality, and the playing is often interfered by external influences, which is easy to lead to rigidity of action, lack of actual combat awareness, and low shooting hit rate [17].

2.1.4. Lack of Competition Experience and High Psychological Tension. Players do not participate in the competition for a long time and have no experience in the competition. Every time after a major competition, their mentality is easy to be too nervous and thus reduce the scoring and shooting rate [18].

2.1.5. Score Versus Shooting Percentage. Shooting score is the yardstick for judging the success or failure of a basketball game. Whether a game is won or not mainly depends on the number of shots and the percentage of hits [19]. Performance is the standard to evaluate the technical and tactical ability of team players, and also the comprehensive embodiment of their performance on the court [20]. Therefore, the score also plays a certain role in shooting. When the score is leading, the skills will be more normal, the shooting action will be more simple and easy, and the hit rate will be higher. When the score falls behind, there will be impatience and the shooting action will be more rigid, thus reducing the hit rate. When the score falls behind, some players will lack self-confidence and even abandon themselves, which is not conducive to giving full play to their shooting skills. When the score is nearly uniform, this is the best time to test the psychological quality. People with good physical and psychological quality generally have a higher shooting rate, while people with good physical but poor psychological quality have a limited shooting rate.

2.2. The Importance of Mood, Confidence, and Concentration in Shooting. Emotion, belief, and attention are also closely related. Emotion contributes to confidence and attention. When athletes are excited, they are confident, focused, and confident in shooting. Otherwise, people will be distracted, lack confidence, hesitate, and delay the flight. Therefore, mentality, belief, and attention are interrelated and mutually restricted.

Stable state of mind: a stable state of mind can make the physiological process and psychological process of shooting consistent, so as to better mobilize the regulatory function of the nerve, so that the muscle can more smoothly and directly carry out the shooting action. The state of mind is stable, and he can control himself under all conditions. Especially under the most difficult conditions, he should be full of confidence, guide himself to fulfill the choices he has made, eliminate any obstacles with strong perseverance, and consciously adjust his best state of mind to maintain stable shooting effects. When the players have strong desire to overwhelm the other side or strong desire to shoot and will to shoot in the game, they will carry a huge psychological burden on their backs, form a bad ideological pressure, cause failure in the game, and even lose confidence in shooting. It can be seen that the balance of mind in the shooting process is very important. A positive and excited attitude will reduce the physical and mental pressure of the players and make it easy to complete high-quality events. A positive, excited, and relaxed attitude often makes the players run actively in the attack and makes full and complete actions when shooting, thus improving the shooting hit rate. Table 1 lists the analysis and investigation of the causes of athletes' psychological and emotional states during ball shooting.

Enduring faith: confidence refers to the ability of athletes to prove that they can complete the specified shooting activities during practice and competition. Lack of self-confidence will restrict the player's thinking and make him hesitant, which will affect the next good shot. Even at the beginning of a good shot, he will be unable to hit the ball because of lack of confidence. When players are confident, shooting skills can be used freely and easily. Confidence is important to the level of the players' practice. During the practice, we should learn to use various teaching methods and other languages to exercise the confidence of the players. Only after hard practice and mastering good shooting skills can we have sufficient material foundation for confidence.

High attention: attention refers to people's mental movement, pointing and focusing on specific objects. Attention can be attracted under certain conditions and time and focused on a certain focus. The high concentration of attention is the main psychological factor when shooting, which can keep the psychological activities in a positive state. Modern psychologists have confirmed that the range of attention is positively related to the tension of attention. The higher the tension of attention, the smaller the range of attention. In the whole shooting process, the attention of the players is pointed to according to the situation of different stages. For example, when getting a basket, players will focus on the distance of the shot, whether the position is blocked by the opponent, what kind of shot to use, and so on.

2.3. Strengthen Psychological Quality and Improve Shooting Accuracy

2.3.1. Pay Attention to the Psychological Preparation before the Game. The competition practice has also proved that the mentality preparation before the competition is important to the victory of the competition. The mentality of players and their preparations for the game usually depend on the understanding of the coach and players on the significance of the game. Before the competition, the psychological countermeasures shall be put forward according to the abilities of the players of both sides and the mental state of both sides in the preparation activities, and transmitted to the players. Let the players have full confidence in the event, so that no matter what happens before and during the event can be in a good psychological state.

2.3.2. Strengthen Self-Psychological Adjustment Training. The situation in the competition is complex and changeable. The failure and achievement in the competition will lead to the change of the players' mentality. Be realistic, people cannot let a successful setback affect their life mentality. Train the players' strong will and keep a clear mind. During the competition, the players realized self-regulation by changing their skills and tactics on the field. They can use self-suggestion, self-imagination, and self-regulation to maintain a good state of mind.

2.3.3. Cultivate the Quality of Decisive Will. Train the team members to create positive shots. Of course, comprehensive psychological analysis should be carried out by moving, getting rid of, grabbing the ball, and breaking through. If positive shots are formed, they should be decisive and do not worry. To do this well, the players need to have a bold style and strong attention, and at the same time, they also need to have strong shooting ability and determination.

2.3.4. Strengthen the Difficulty of Simulating Confrontation in Training. In the ordinary practice, it can be set in different training venues, different technical conditions, tactical conditions, and different confrontation conditions to improve the antagonism of offensive shooting, increase the stability and hit rate of shooting, and also use more practice means such as hypothetical luck and time pressure, such as hypothetical time is the last 3 minutes, 1 minute, 30 seconds, 5 seconds. Carry out special cooperative shooting training under the condition of backwardness or draw at all times, so that the team members will not be nervous and soft at all times and in adversity, and have the courage to fight and dare to stop and win, so as to improve and consolidate the shooting percentage. In the ordinary practice, in addition to

Table 1: Questionnaire on the causes of psychological and emotional states of athletes when shooting three-point ball.

name	Delegation	Position	Emotional state	Cause
Xie XX	Northeast Normal University	Point guard	Relax	Have a correct understanding of the game
Zhu XX	Xi'an Jiaotong University	Forward	Absorption	Strong confidence in the match
Yan XX	Tianjin Polytechnic University	Forward	Relax	Treat the game normally
Liu XX	Liaoning University	Forward	Nervous	The score fell behind in the game
Chen XX	Guangdong University of Technology	Forward	Nervous	The score fell behind in the game
Luo XX	Hunan University	Point guard	Decisive	Highly trained and confident
Wang XX	Huazhong University of Science and Technology	Forward	Panic	Inadequate preparation for the game
Ma X	Wuhan University of Technology	Forward	Nervous	Lack of self-confidence

training the players to learn more points and various shooting training methods, the quality of shooting training must also be increased. Especially after the training or a large number of training, the high-quality shooting training cannot be ignored, nor can the players form arbitrary and irresponsible shooting behavior in the ordinary practice, such as the methods in some competitions, such as penalty kicks or fixed-point shooting competitions. All of this requires the training team members to be more careful and serious, and in view of the characteristics of the competition, so as to enhance their confidence in the game and their shooting percentage when their physical quality is significantly reduced.

2.4. Promote the Basketball Players Smooth Play of the Psychological State of Tactics and Skills. The psychological conditions that promote the smooth development of basketball players' skills and tactics are divided into the following four aspects:

First is emotional stability. A stable mood is the prerequisite and cornerstone for the normal development of basketball players' skills and tactics. This kind of stable mood is to make the basketball players in the competitive activities, not affected by personal feelings, or sad or happy, to control the results of the game. Whether a player has a stable mood in the competition and normal practice plays an important role in whether a player can play a stable role in the unpredictable competition. Emotional stability is the key reason to maintain good psychological quality. There are many factors that cause emotional changes. Master the various factors that affect the emotional stability of basketball players, and point out the development direction for routine practice.

Second is strong will. It is generally believed that the willpower refers to the basketball player's ability to control his own actions and technical behaviors spontaneously in order to win the final game in competitive activities. Its main feature is the players' positive attitude in the competition. In competitive activities, players must have strong willpower to overcome their own or external obstacles. This will reflects the shaping ability and perseverance during the competition and is not afraid of difficulties. Basketball, as a team activity, also contains a huge antagonistic component, which leads to the variability and complexity of the will of basketball players. Therefore, this requires the strong will of the team members to overcome many difficulties in the competition.

Third, we should have the ability to adapt to change. Basketball is not only fierce, but also complex and changeable. The game is characterized by uncertainty. There is no way to predict what will happen on the court. It is impossible for the opponent to play according to the skills you have prepared.

3. Methodology

3.1. Feedforward Neural Network. Neural network is a machine learning algorithm that simulates human brain. It has strong expression ability and good data fitting ability. Neural networks are generally composed of injection layer, hidden layer, and input-output layer. By setting multiple hidden layers, it can automatically learn features from the bottom to the top of the training data. Expression of neural network is

$$y = NN(x). \tag{1}$$

X is the input vector, and y is the task-dependent output.

The simplest neural network model is feedforward neural network, also known as multilayer forward neural network. As shown in Figure 1, each item has a value even for money, and there are two hidden layers.

The basic working unit of neural network is neuron. The information of the upper level can be summed by weighted matrix to obtain neurons and then activated by activation method. The output is calculated to be consistent with the input value of the next stage, so it has a good fitting function.

Among θ, I represents the weight of Xi input, and the machine can learn independently. F is the activation parameter. The machine can realize nonlinear fitting of data, but usually uses other activation parameters such as sigmoid, tanh, and ReLU. The mathematical expression of the multilayer feedforward neural network of a group of hidden layers n is

$$\begin{aligned} NN(x) &= y, \\ h' &= x, \\ h' &= g^i\left(h^{i-1}W^i + b^i\right), \\ y &= h^{n-1}W^n. \end{aligned} \tag{2}$$

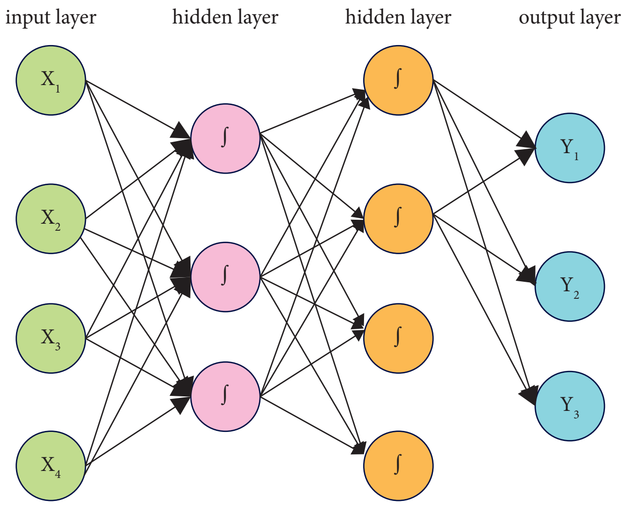


Figure 1: Feedforward neural network.

3.2. CNN Model. Convolutional neural network (CNN) model is a neural network model for obtaining local information. CNN initially used image recognition technology. Most pictures on social networks have simple language structure and compact structure, and can also be described using embedded word vector technology. Through the application of CNN model, it also achieved good results in text analysis. The CNN model first arranges the embedded vectors of words from the text and then performs the following three processes after input:

Step 1. Convolution layer: Each neuron was treated as a filter. The filter type is $f * \mathrm{F}$. It scans the $k * x$ matrix of a sentence filled with s steps, where k is the width of the sentence and X is the word embedding vector of a word. In the inner product of a scan matrix, the obtained (rounded down) matrix is finally generated. Convolution algorithm is a step-by-step feature extraction method. Designing multiple filters can extract high-level features.

Step 2. Pool layer: The pool is carried out in the process of continuous convolution. By obtaining relevant data in the text, and by using cell average technology in the models of Max and LSTM, the parameters of the data source are reduced to reduce the amount of overfitting and improve the feature extraction technology.

Step 3. Full connectivity layer: The calculation result of convolution layer is ReLU activation function or nonlinear mapping. Finally, the softmax function is used to calculate the tag probability of more types of emotion. ReLU function is

$$f(x) = \begin{cases} x\,(x > 0), \\ \lambda x\,(x \leq 0), \end{cases} \tag{3}$$

where λ is a variable that can be learned through the back propagation algorithm.

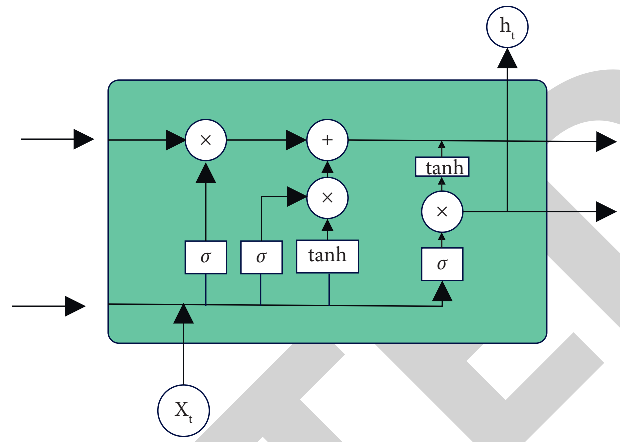


Figure 2: A cell in the LSTM model.

3.3. LSTM Model. LSTM was first proposed by Hochreiter et al. to deal with the phenomenon of ladder disappearance and ladder destruction caused by the short memory time of recurrent neural network (RNN). It is also an RNN. LSTM can fit the sequence signal well. LSTM is mainly composed of neurons, which is equivalent to the neural cells of memory. On the basis of the entrance and exit of RNN mode, a special forgetting gate is also added, which has the ability of long-term memory. The entry gate, output gate, and forgetting gate are adjusted by sigmoid trigger parameters. Because the sigmoid function can take the value from zero to one, it can limit the opening rate of the forgetting gate at the current node. Both the input gate value and the cell state can be converted by the tanh trigger function. Figure 2 shows a cell in the LSTM mode.

Forgotten door expression is

$$f_t = sigmoid\left(W_f x_t + U_f h_{t-1} + b_f\right). \tag{4}$$

Enter the door expression as

$$\begin{aligned} i_t &= \mathrm{sigmoid}\left(W_i x_t + U_i h_{t-1} + b_i\right), \\ \overline{C}_t &= \mathrm{s\,tanh}\ \left(W_C x_t + U_C h_{t-1} + b_C\right). \end{aligned} \tag{5}$$

4. Result Analysis and Discussion

4.1. Experimental Data and Environment. The model in this paper will be trained by the flowchart above to obtain the final neural network parameters suitable for the needs of this paper. (1) Use Python's Jieba library to convert the text into one-hot code and then transfer it to the input layer; (2) the word vector of one-hot code pretrained by word2vec is transformed into the word embedding vector as the input of neural network in the embedding layer; (3) various outputs are obtained after the model is trained by Bi-LSTM network or CNN model; (4) after full connection layer multiplication, the affective tendency results are obtained by softmax function multi-classification. Based on this process, this paper will compare and summarize Bi-LSTM and CNN models.

Table 2: Bi-LSTM parameters.

Parameters	Value
Batch size	64
Unit-num	32
Bi-LSTM layer	32
Epoch	6
Activate function	LeakyReLU
Learning rate	0.001
Loss function	Cross-entropy
Optimization function	Random gradient descent
Word vector dimension	256

The main data collection of this paper is from the public data set of IMDB, with a total of 50000 data sets, and the ratio of positive and negative samples is 1 : 1. After pretreatment, it can be divided into 20000 training sets, 5000 development sets, and 25000 test sets. The training set is used to train the neural network modeling, and then the data of the development set are used to select the debugging parameters of the model. Finally, the generalization ability of the model is evaluated by the test set. The test parameters directly affect the test results of the model. After the pretest, we compared the hidden layer function, optimization parameters, CNN loss function, and filter of different Bi-LSTM, and obtained the best initial parameters of the two modes. Table 2 shows the parameters of all Bi-LSTM, and Table 3 shows all parameters of CNN.

Table 3: CNN parameters.

Parameters	Value
Filter size	3∗3
Filter quantity	100
Activate function	ReLU
Pooling method	Max
Dropout rate	0.5
Epoch	6
L2	3
Word vector dimension	256

4.2. Experimental Results and Analysis. This chapter not only compares the accuracy of Bi-LSTM and CNN mode, but also practices the classical RNN mode and LSTM mode to help analyze the advantages, disadvantages, and effects of this mode. The experimental results are shown in Table 4. It can be seen that Bi-LSTM has achieved quite good results compared with LSTM in context analysis, and its accuracy is 5.52% higher than that of CNN model. The prediction should be because the different sentence lengths of the data sets in this paper lead to CNN's lack of high accuracy in analyzing long paragraphs. However, the accuracy of RNN is significantly lower than that of the other three due to gradient disappearance and gradient explosion. Although LSTM is superior to RNN, it is still impossible to make accurate estimation because the following contents cannot be merged. The conclusion of this paper also indirectly shows that the emotional expression of middle school students for daily texts is scattered in language segments. If the analysis is performed manually, an error may occur. The accuracy range of Bi-LSTM is 95.55%, so it can effectively realize large-scale text processing.

Table 4: Comparison of experimental results.

Model	Acc
Bi-LSTM	95.55
CNN	90.03
LSTM	85.07
RNN	81.33

As can be seen from Figure 3, the solid line represents the measured sample data, and the dotted line represents the predicted data of the network. In most cases, it can be said that the trend of the two curves is basically the same, but it does not rule out that there is a large difference between the actual measured data and the predicted data at individual prediction points. Compare the prediction results of the network with the data obtained from the actual measurement, and the relative error between the prediction results and the actual measurement results is within the acceptable range, as shown in Figure 4. Therefore, the prediction results can meet the requirements of measurement, but it also shows that the prediction model can be further improved to improve the accuracy and accuracy of network prediction results.

As can be seen from Table 5, the null hypothesis is rejected. Therefore, the performance of each algorithm is different. Based on the above analysis results, it can be determined that the proposed method performs better than K-means, k-meanspo, k-meansfa, and fuzzy *k*-means algorithms in statistics.

In order to further verify the performance of the SECNN model, different amounts of training data were selected to conduct a comparative experiment on the topic recognition task. The number of training data selected under each topic increased from 1000 to 10000, and 1000 pieces of data were added each time. Ten comparisons were made. And the representative *f*-value is used as the main judgment index. The test results are shown in Figure 5. It can be seen from the experimental results that the improved neural network model has shown its advantages when the amount of data is

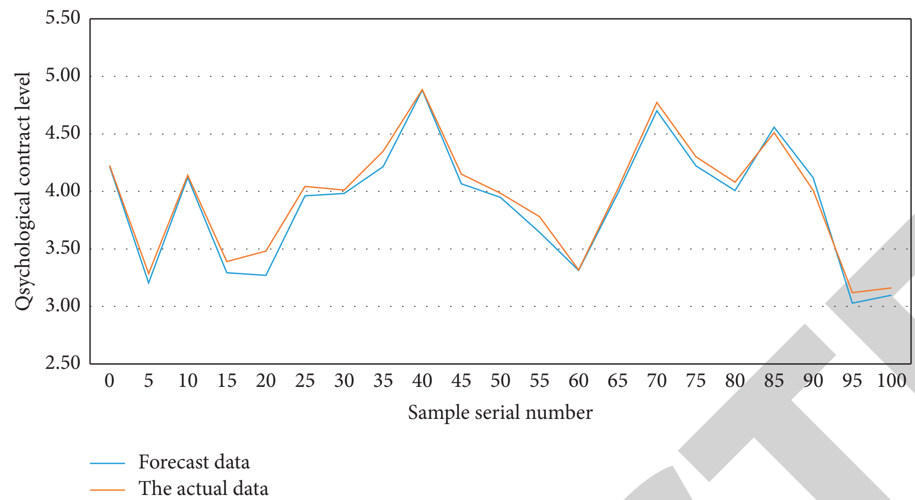


FIGURE 3: Neural network prediction results.

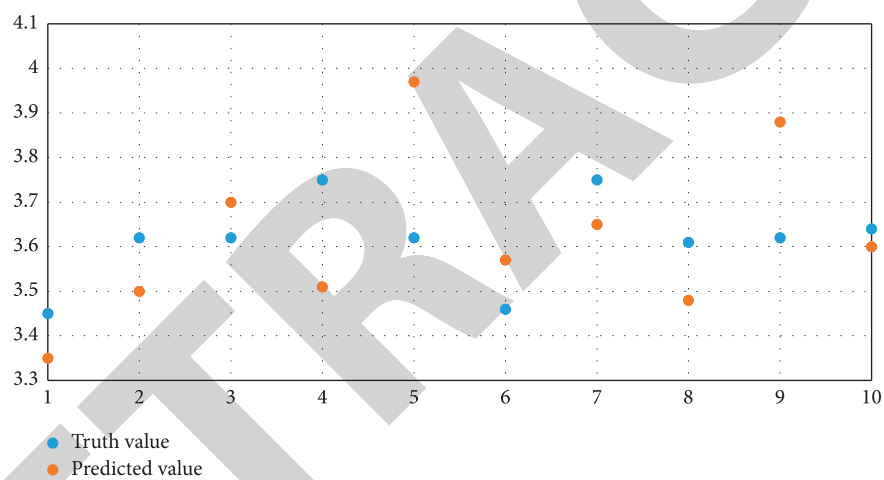


FIGURE 4: Corresponding error diagram between predicted data and actual data.

TABLE 5: Statistical results of Friedman test.

Algorithm	z value	p value	α/(k-1)	Result
K-means	−2.8571	0.00219	0.02	Refuse
K-means PSO	−2	0.02275	0.025	Refuse
K-means FA	−2	0.02275	0.0333	Refuse
Fuzzy K-means	−2.8571	0.02219	0.05	Refuse
Method in the text	−3.1428	0.00084	0.10	Refuse

relatively small; especially with the increase of the amount of data, this advantage gradually increases. This is because there are many learning parameters in the neural network model, which requires a large number of data to achieve better learning results. However, the learning ability of the traditional machine learning model does not increase with the increase of the data scale. The $F1$ value of the three models only reaches about 86.7%, which is also a deficiency of the traditional machine learning model. In addition, compared with CNN model, the $F1$ value of SECNN model reaches about 94.9% when 7000 pieces of data are available, while the $F1$ value of CNN model is about 95.0% when 11000 pieces of data are available, which indicates that SECNN requires less training data when there is only a difference of about 0.1% $F1$ value, proving that SECNN model has stronger feature extraction ability and learning ability.

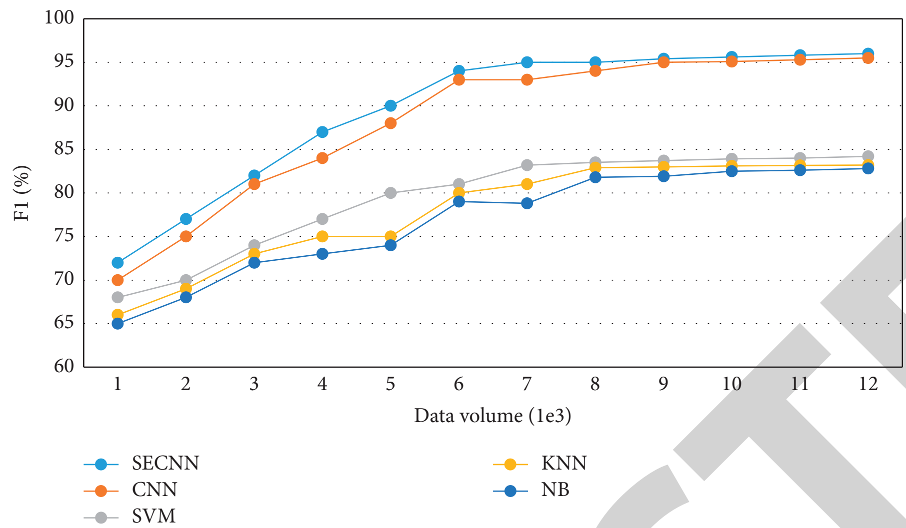


FIGURE 5: Performance comparison of different data sizes.

5. Conclusion

Since the birth of basketball, the theoretical research on basketball psychological training has gradually deepened. In today's international basketball arena, the psychological adjustment and psychological training of the players have become the main part of the whole training process and play a vital role in the athletes' competition. In the paper, the batch analysis and the solving methods of athletes' psychological problems are compared and analyzed, and some conjectures are made in view of the different properties of Bi-LSTM model, CNN model, and traditional neural network model. Experimental results show that the *F*1 value of SECNN model reaches about 94.9% when it has 7000 data, while the *F*1 value of CNN model is about 95.0% when the data volume is 11000, indicating that SECNN requires smaller training data when it differs only by about 0.1% *F*1 value, which proves that SECNN model has stronger ability to extract features and learning ability. This paper puts forward a feasible scheme for batch analysis of athletes' psychological text.

Data Availability

The labeled data set used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This work was supported by Shanxi University.

References

[1] S. Oehme, P. Moewis, H. Boeth et al., "PCL insufficient patients with increased translational and rotational passive knee joint laxity have no increased range of anterior–posterior and rotational tibiofemoral motion during level walking," *Scientific Reports*, vol. 12, no. 1, p. 13232, 2022.

[2] Y. Yang, "An analysis of hazel motes in wise blood from the perspective of posthumanism," *Journal of Higher Education Research*, vol. 3, no. 2, pp. 188–190, 2022.

[3] B. Arla, D. Tfuac, and E. Lga, "Impact of quality, type and volume of data used by deep learning models in the analysis of medical images," vol. 14, no. 2, p. 9, 2022.

[4] T. Fechter, I. Sachpazidis, and D. Baltas, "The use of deep learning in interventional radiotherapy (brachytherapy): a review with a focus on open source and open data," vol. 13, no. 5, pp. 7–13, 2022.

[5] P. A. Karwande and V. K. Chechi, "Effect of yogic activities on selected physiological and psychological parameters among STEM undergraduates," *IOP Publishing Ltd*, vol. 10, no. 5, pp. 8–18, 2022.

[6] K. Sjblom, S. Juutinen, and A. Mkikangas, "The Importance of self-leadership strategies and psychological safety for well-being in the context of enforced remote work," *Challenges*, vol. 13, no. 1, Article ID 012001, 2022.

[7] C. E. Gonalves and H. M. Carvalho, "Revisiting the relative age effect from a multidisciplinary perspective in youth basketball: a bayesian analysis," *Frontiers in Sports and Active Living*, vol. 2, no. 1, pp. 83–88, 2021.

[8] Z. Xiang, W. Gao, and T. Tao, "Psychological test of patients with mental disorders based on eye movement data fusion algorithm," vol. 3, no. 1, pp. 47–58, 2021.

[9] D. Sprouts, Y. Gao, and C. Wang, "The development of a deep reinforcement learning network for dose-volume-constrained treatment planning in prostate cancer intensity modulated radiotherapy," *Biomedical Physics & Engineering Express*, vol. 9, no. 23, 2022.

[10] J. Su, "Scene matching method for children's psychological distress based on deep learning algorithm," *Complexity*, vol. 2021, no. 3, Article ID 6638522, 11 pages, 2021.

[11] Q. Dong, H. Wang, and Z. Hu, "Commentary: using goal-driven deep learning models to understand sensory cortex," *Frontiers in Computational Neuroscience*, vol. 12, p. 4, 2018.

[12] A. Whiten, "The psychological reach of culture in animals' lives," *Current Directions in Psychological Science*, vol. 30, no. 3, pp. 211–217, 2021.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9756026, 1 page
https://doi.org/10.1155/2023/9756026

Retraction

Retracted: Design of Intelligent Speech Translation System Based on Deep Learning

Mobile Information Systems

Received 22 August 2023; Accepted 22 August 2023; Published 23 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Tan, "Design of Intelligent Speech Translation System Based on Deep Learning," *Mobile Information Systems*, vol. 2022, Article ID 2463812, 7 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 2463812, 7 pages
https://doi.org/10.1155/2022/2463812

Research Article

Design of Intelligent Speech Translation System Based on Deep Learning

Ying Tan

Jiangsu Maritime Institute School of International Education, Nanjing, Jiangsu 211170, China

Correspondence should be addressed to Ying Tan; 20148441@stu.sicau.edu.cn

Received 21 June 2022; Revised 29 July 2022; Accepted 17 August 2022; Published 6 September 2022

Academic Editor: Jiafu Su

In order to solve the problem of low translation accuracy caused by complex sentence parameters in traditional machine translation systems, a method based on deep learning was proposed. First, MCU SPCE061A is used to study the problem of complex digital signal. The training data in the synchronous translation server support the translation services of a large number of users, and the translation results were displayed through the session interface of the user terminal. The PMDL model is used to detect the keyword signal, record the PCM audio data, and slice the collected pulse code modulation signal, so as to wake up the artificial intelligence voice service. Then, this study establishes a speech recognition process that accurately outputs the speech-related semantics. In this paper, a manual interactive synchronous translation program is designed with the input text as the search criterion, and the set is trimmed to obtain the best translation effect. The experimental results show that the sentence translation accuracy of the system is 0.9 ~ 1.0. It is proved that the method based on deep learning solves the problem of low accuracy of the traditional translation system.

1. Introduction

Governments, businesses, academic organizations, humanitarian organizations, and other organizations have recently faced unprecedented internationalization and globalization. The effectiveness, market size and scope of competition of security, and trade and commerce all depend on global information awareness and global interaction and communication ability. Strengthening world integration requires natural and effective international cross language communication, and the language gap is a huge obstacle to world integration [1]. With the increasing popularity of real-time translation, it is urgent to develop a system supporting synchronous translation. Speech to text translation (s2t) refers to the process that allows the machine to automatically translate the text of the target language from the speech signal of the source language.

A traditional speech translation system usually consists of two parts: speech recognition module and machine translation module. Cascading the two modules can form a speech translation system [2]. This cascading approach can improve the overall performance from the improvement of a single component. For example, in recent years, neural machine translation has generally replaced statistical machine translation, which not only improves the quality of text translation but also improves the quality of speech translation. For this model which separates speech recognition and translation models, two models with good performance may be generated, respectively. However, when the two models are cascaded together, the inherent error propagation problem in this way will affect the performance of the whole system.

Cascading speech recognition and machine translation is still the mainstream method of speech translation [3]. However, the result of speech recognition is spoken sentences, which contain a large number of nonstandard languages, such as repetition, ellipsis, inversion, unclear semantic logic, broken sentences, and so on. In addition, influenced by the accent of the speaker, environmental noise, homonyms, and confusing words in the language, there will be errors in the results of speech recognition. These problems bring great challenges to the follow-up machine translation. Therefore, a speech recognition postprocessing module is usually added between speech recognition and machine translation to reduce the impact of verbalization and recognition errors as much as possible by regularizing, breaking sentences, smoothing, punctuation prediction, and even error correction on the recognition results. Due to the complexity of oral English, these problems have not been completely solved. Combined with artificial intelligence technology, a machine synchronous translation system with a speech recognition device is designed. The translator configuration system helps to improve the synchronous translation service, to improve the user experience effect, and to promote the healthy development of the online translation market.

2. Literature Review

Liu and others proposed the famous dynamic time warping (DTW) algorithm, which can effectively compare two time series with different lengths and calculate the similarity through dynamic programming. Because the algorithm is easy to implement, it once dominated the research of speech recognition technology. Subsequently, the statistical model-based method gradually stepped onto the historical stage of speech recognition, among which the well-known algorithm is the hidden Markov model (HMM). After the statistical model, artificial neural network (ANN) has opened a new door for research in the field of speech recognition and has gradually become the mainstream method of speech recognition [4]. Luo and others proposed a new time-frequency convolution neural network (TFCNN) framework, which convolutes the feature space in both time and frequency scales. Under all test conditions, for all feature sets, the performance of the framework is always better than the convolution neural network, which can significantly reduce the word error rate (WER) [5]. Zhang and others proposed an end-to-end automatic speech recognition model for mono-multispeaker on ICASSP (International Conference on Acoustics, Speech and Signal Processing). Compared with other researches on mono-multispeaker speech recognition, this model can improve the performance of end-to-end model in separating overlapping speech and recognizing separated streams and bring about a relative performance improvement of about 10.0% in character error rate (CER) and WER, respectively [6]. Das and others proposed a new acoustic model for far-field speech recognition tasks. The long short term memory recurrent neural network (LSTM) based on attention mechanism and multitask learning framework reduced the absolute word error rate by 1.5% [7]. Xue and others proposed a digital Chinese continuous recognition system. Aiming at the characteristics of Chinese speech recognition, they adjusted the acoustic model of Sphinx speech recognition system so that the recognition rate of the model for speech digital string can reach 98% [8]. HORII and others proposed a speech recognition framework—deep full convolutional neural network (DFCNN). This framework has achieved 15% improvement over the two-way RNN-CTC, a widely used speech recognition model in the industry, and provides the recognition of Cantonese, Henan, Sichuan, and other dialects in the dictation test of Chinese voice messages within iFLYTEK [9].

The translated sentences contain complex sentence parameters, which are not eliminated immediately, resulting in poor translation results. Therefore, combined with artificial intelligence technology, a machine synchronous translation system with a speech recognition device is designed. The translator configuration system helps to improve the synchronous translation service, improve the user experience effect, and promote the healthy development of the online translation market.

3. Research Method

3.1. System Hardware Structure Design. In order to realize the reasonable connection between translation services and the needs of users at all levels and also effectively allocate translator resources, this paper designs a machine synchronous translation system based on artificial intelligence technology and speech recognition [10]. The system is mainly composed of user terminal, server, and translator terminal. The feedback module of the user terminal is responsible for communication between internal modules or between the interpreter end and the outside, displaying the input and output content. The sound input and output part is synchronously translated by the translation module, and the feedback module finally gives feedback to the input interpreter user. The modules of the server are, respectively, used for the communication between the internal modules of the server or between the server and the external modules and to store the generated or received external information. The translation scoring module scores the information received by the computer after machine translation, and the order allocation module assigns orders to specific interpreters. The corpus learning module mainly transforms the translated content into corpus through machine learning technology to improve the level of machine translation. The translator terminal includes a communication module, which is responsible for the communication between internal modules or between the interpreter terminal and the outside. The display module is responsible for displaying the input and output contents; voice module inputs and outputs voice to the module; and translation module is responsible for synchronous translation.

3.1.1. SPCE061A Single-Chip Microcomputer. 32 KB flash is embedded in the SPCE061A chip. After high-speed processing, it can easily and quickly process complex digital signals. The structure of SPCE061A single-chip microcomputer is shown in Figure 1.

An intelligent speech recognition module based on SPCE061A is designed, and the corresponding control program was written. After when different voice command signals are received, the MCU pin sends the predetermined high and low levels [11, 12]. After when the microprocessor is connected to the power level, it will immediately and intelligently identify the sentences to be translated, send the

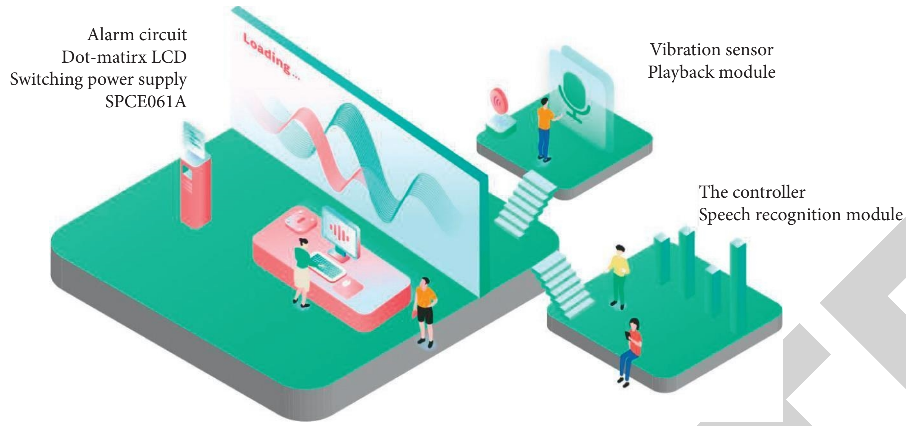


Figure 1: SPCE061A single-chip microcomputer structure.

letter composition instructions to the system in combination with the actual language conditions, and then combine the letters with idioms in combination with the artificial intelligence technology. The single-chip microcomputer with its own loudspeaker is used to play the intelligent recognition results synchronously in real time.

3.1.2. Machine Synchronous Translation Server. One or more translation servers and network servers running the decoder constitute a machine synchronous translation system, and the translation services between different languages are usually handled by different translation servers. Therefore, a distributed server system was built to support the translation services of a large number of users. The system was equipped with a decoding server and a network server. The online translation system runs on the decoding server. Users can query online synchronously through the HTTP network server interface.

The establishment of the machine translation server architecture is divided into two stages: first, learning the step size from the large-scale parallel corpus to obtain the maximum probability, and then second, the maximum probability solution is obtained by training. This method obtains the training data by counting the sentence pairs in the parallel corpus, then prunes them to remove the redundant data, and finds the most likely translation results according to the training data.

3.1.3. User Terminal. The user terminal provides a user interface, i.e., a session interface, which includes a session module and a translation module. This is the user interface. When the session button on the main interface is pressed, the program will jump to the session interface. After clicking the dialog button, the program will enter a dialog list page. When you click an entry in the dictionary, the program jumps to the dictionary page.

The user terminal session interface frame displays the chat list, the language information of each message, the time when the message is translated, etc. There is a language input button at the bottom of the dialog box to input Chinese and English buttons, respectively [13, 14].

3.1.4. Microphone Array. Due to various noise interference, the speech signal is inaccurate, and even the speech is submerged. Therefore, installing a microphone array in the system can convert sound signals into digital signals. Using this array can not only improve the resolution of sound but also extract accurate pure speech from noisy speech signals. This method extracts the time and space information of the sound source through the microphone array and suppresses the noise so as to accurately recognize the voice instructions in the noisy environment.

3.2. System Software Design

3.2.1. Design of Artificial Intelligence Voice Wake-Up Function. In order to realize the voice wake-up function, it is necessary to set up a process in the background regularly and real time monitoring of the surrounding environment of the device and detect whether there is a signal of the keyword required by the device in the signal. The specific process is shown in Figure 2.

Figure 2 shows that after the keyword signal detection was completed, the detection callback or static detection was performed and the action is recorded. During the static detection of the recorded action, the process returns the recorded PCM audio data to the main process, and then continues to the next step. On this basis, the PMDL model was used to slice the collected pulse code modulated signal [15, 16].

3.2.2. Speech Recognition Function Design. With the continuous development of the Internet of things and artificial intelligence, the demand for free interaction between people and computers is more and more frequent, and people pay more and more attention to it. Now, it is generally believed that speech recognition is a part of language technology, which also includes speech synthesis and natural language processing. The main purpose of speech recognition technology is to complete the conversion from speech to text, so speech recognition is also the basic condition of natural language processing. The speech signal is generated by the vibration of the vocal organ, so the lip radiation and glottic

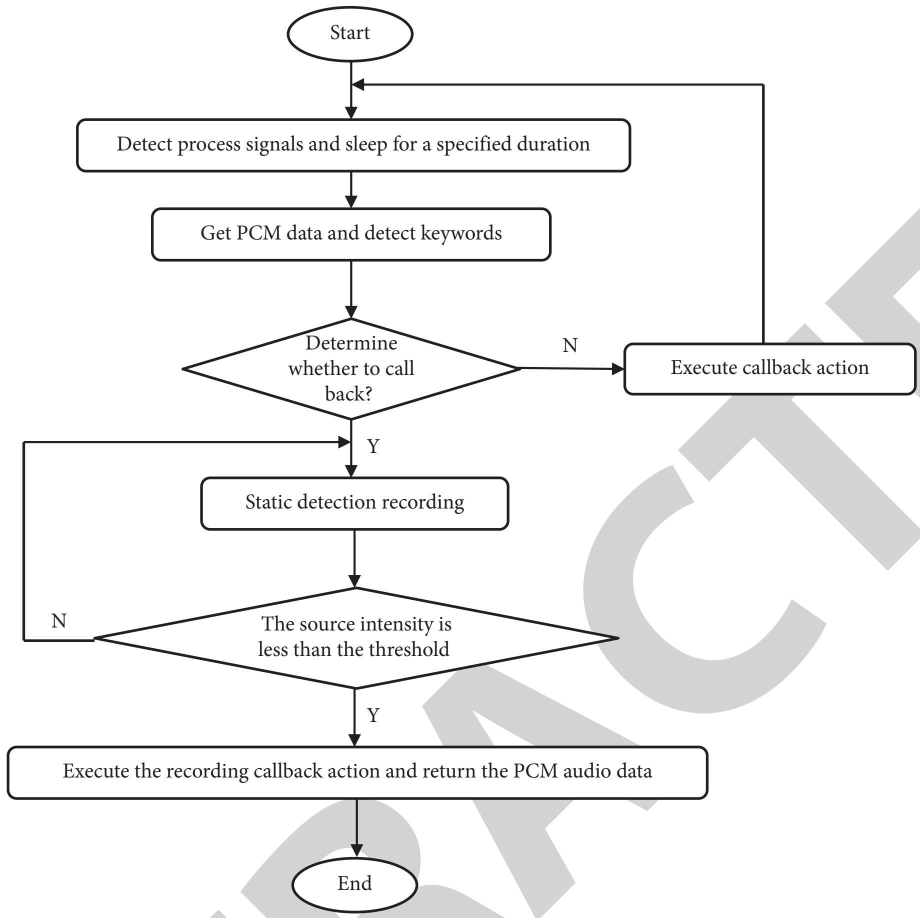


Figure 2: Artificial intelligence voice wake-up process.

excitation will have a certain impact on it, resulting in the reduction of the high-frequency band of the speech by 6 db/octave band above 800 Hz. Therefore, in the process of speech transmission, the high-frequency part is easy to be lost. Therefore, the preemphasis step should be introduced before the speech signal processing. The main function of speech signal preemphasis is to reduce the influence of lip radiation and improve the resolution of high-frequency part. Speech signal is a typical nonstationary time-varying signal; that is, its distribution law will change with time. Since most signal processing systems can only process stable signals at present, it is necessary to process speech signals first. After research, the speech signal has the characteristic of "short-term stability" because the occurrence of speech is closely related to the movement of the vocal organ. When speaking, the vocal organ will produce a certain inertial movement. Within a period of time (generally 10 ms ~ 30 ms), the speech signal can be regarded as approximately unchanged, so this speech signal can be treated as a stable signal [17, 18]. The speech recognition process is presented in Figure 3.

Figure 3 demonstrates that the speech recognition system can complete speech recognition based on the following four working principles. First, the speech recognition library and detection terminal are selected directionally and then combined with antialiasing bandpass filtering technology, which can effectively eliminate individual speech differences,

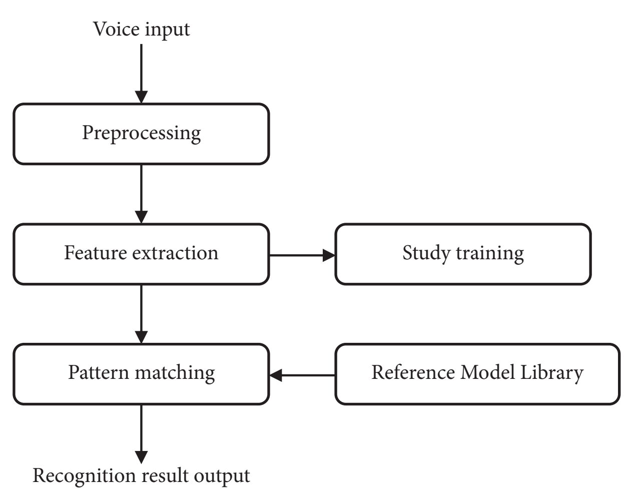


Figure 3: Speech recognition process.

sampling equipment, and sampling environment noise. Then the speech acoustic parameters such as average energy, vibration peak, and average zero crossing rate are extracted, which can quickly and accurately reflect the sound quality characteristics of speech; After that, a speech pattern database is established. The main link of language repetition training is to let the speaker repeat his pronunciation, directly delete the redundant speech information from the

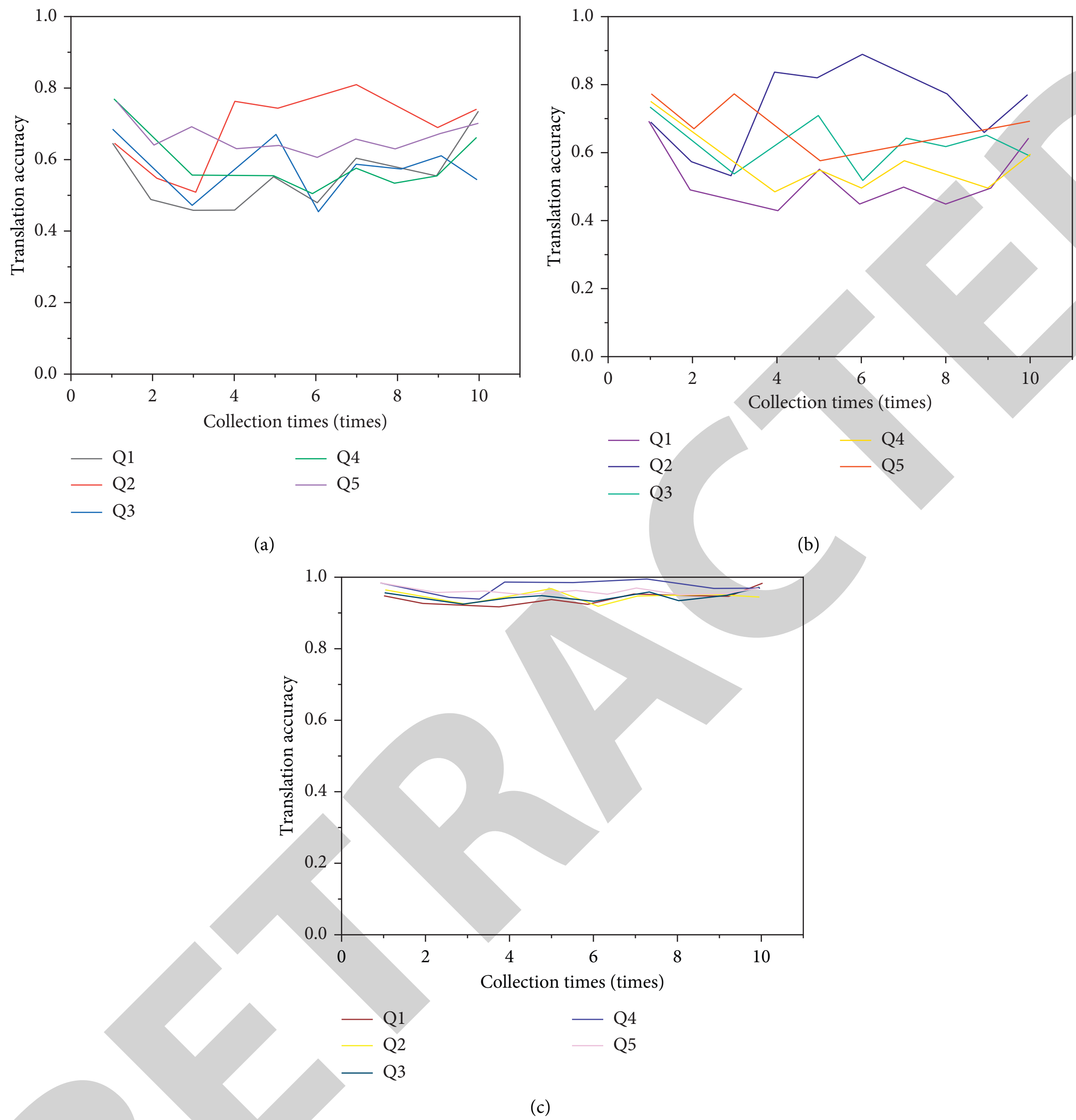


FIGURE 4: Comparative analysis of translation accuracy of the two systems. (a) Machine synchronous translation system based on the Moses statistical machine translation framework. (b) Corpus based machine translation system. (c) Machine synchronous translation system based on artificial intelligence and speech recognition.

original speech samples one by one, only retain some key speech data, and scientifically classify the key speech data according to the relevant schemes. Finally, the relevant semantics of speech are accurately output according to the speech similarity [19, 20].

3.2.3. Manual Interactive Synchronous Translation. The machine translation method of human-computer interaction function includes the following steps:

Step 1. Read the machine translation model and select the corresponding prestored domain according to the user of the translation domain.

Step 2. After reading the text, divide the text into a series of sentences to facilitate subsequent modules.

Step 3. By inputting text as search criteria, after receiving the user's input text, search the matching translation on the search network, and different users can get different translation results corresponding to the input [21]. Since the number of states in the W_n set increases exponentially with the increase of n, it will take a lot of time if the set is not pruned. Therefore, pruning is required. The pruning process is as follows: for the determined source language sentences $R_1, R_2, \ldots, R_n$, there is a phrase model (f, g, s, d), and the four elements in the model represent phrase library, grammar model, distortion limit, and distortion parameter, respectively.

Suppose W_{n} translates n words into m sets. If an element in the M statement in W_2 consists of two words, it means that only two words are translated into idioms. For W_{i}, each state has a transition state, all possible states will be added to the corresponding set, and finally the state with the highest score will be returned [22].

Let β be the search parameter and p be the transfer parameter, and the resulting syntax model is as follows (1):

$$g = \mathrm{next}(p, d). \tag{1}$$

After determining the search parameters, remove all the parameters in the set that do not meet (2) to achieve the purpose of pruning:

$$a(g) > a(d) - \beta. \tag{2}$$

Step 4. When translating sentences in the source language, first read the translation options of the sentences in the source language, and then expand the translation assumptions from a small container to a large container. At each transition stage, if the difference between a score and the highest score in the container is greater than the threshold, then the state will decline. If the state remains unchanged, all available transition options expand. If the old and new assumptions are the same, the score increases. The best translation result is to find the translation statement with the highest score in the largest container.

4. The Experimental Results

The purpose of the experiment was to test the functions of the machine synchronous translation system based on artificial intelligence technology and speech recognition so as to ensure that the system can meet the purpose of synchronous translation.

4.1. Experimental Data Acquisition. 123425 English sentences were selected from the translation database, out of which, 1000 were derived from translation materials in the field of news. Five of the 1000 sentences were randomly selected and translated by five teachers. The translated results are as follows: Q1 is Japan is interested in China's new round of strategic technology; Q2 is that the United States exerts pressure on South Korea; Q3 is that the United States and China hold negotiations on the issue of border deployment; Q4 is thatthe US side attaches great importance to China US relations; and Q5 is that China will not pose a threat to other countries. In the experimental results, the Moses statistical machine translation framework translation system, corpus translation system, artificial intelligence technology, and speech recognition translation system are used to translate the above five sentences. The accuracy of the translation results is shown in Figure 4.

It can be seen from Figure 4(a)) that using the machine synchronous translation system based on Moses statistical machine translation framework, the translation accuracy of five sentences was between 0.4 and 0.8. It can be seen from Figure 4(b)) that using the corpus based machine synchronous translation system, the translation accuracy of the five sentences was between 0.4 and 0.9. It can be seen from Figure 4(c)) that the machine synchronous translation system using artificial intelligence technology and speech recognition had a translation accuracy of 0.9 ~ 1.0 for five sentences [23]. According to the above analysis results, the machine synchronous translation system using artificial intelligence technology and speech recognition has high translation accuracy.

5. Conclusion

A single-chip microcomputer has the characteristics of convenient and rapid processing of complex digital signals in speech recognition hardware, which has become the main advantage of machine synchronous translation system to realize artificial intelligence technology and speech recognition. The system combines the artificial intelligence technology and speech recognition technology. It has the advantages of high recognition efficiency and high accuracy and thus meets the needs of people for different speech communication. However, the system still has many shortcomings, which need to be further improved in the design process: in the next development stage, a more efficient decoding algorithm can be used to improve the decoding speed. A more robust caching mechanism and common word search mechanism can be adopted to speed up offline translation and meet the requirements of efficient translation. The application of deep learning technology can ensure the organic unity of high accuracy and high efficiency of scenic spot text translation and speech recognition. It provides a convenient and simple way for tourists to obtain scenic spot information. At the same time, it provides a good help for tourists who are unfamiliar with the local language of the tourist destination or have certain visual impairment so that they can effectively understand the scenic spot information and improve their tourism experience.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Z. Z. Gafarova, M. A. Bozorova, S. S. Jumayeva, L. I. Idiyeva, and L. U. Radjabova, "Synchronous translation – a complex set of cognitive processes," *International Journal of Psychosocial Rehabilitation*, vol. 24, no. 1, pp. 403–407, 2020.

[2] S. M. Shavkatovna, N. N. Soliyevna, S. M. Ahmadovna, and K. S. Shaydulloevna, "Strategies for mastering synchronous translation art," *International Journal of Psychosocial Rehabilitation*, vol. 24, no. 03, pp. 404–415, 2020.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9765805, 1 page
https://doi.org/10.1155/2023/9765805

Retraction

Retracted: The Biomechanical Analysis of Jumping Difficulty Movement in National Traditional Sports

Mobile Information Systems

Received 22 August 2023; Accepted 22 August 2023; Published 23 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] L. Zhang and W. Yang, "The Biomechanical Analysis of Jumping Difficulty Movement in National Traditional Sports," *Mobile Information Systems*, vol. 2022, Article ID 7679116, 8 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 7679116, 8 pages
https://doi.org/10.1155/2022/7679116

Hindawi

Research Article

The Biomechanical Analysis of Jumping Difficulty Movement in National Traditional Sports

Linbao Zhang [1,2] and Wei Yang [3]

[1]Physical Education College, Shangrao Normal University, Shangrao, Jiangxi 334001, China
[2]Cavite State University, Cavite 99900, Philippines
[3]Research Center Wujia School Suzhou University, Anhui, Suzhou 234000, China

Correspondence should be addressed to Wei Yang; 1700210465@stu.sqxy.edu.cn

Received 30 May 2022; Revised 18 July 2022; Accepted 28 July 2022; Published 5 September 2022

Academic Editor: Jiafu Su

In order to explore the relevant factors and the completion of difficult movements, this paper is an important theoretical basis for improving and constructing the evaluation of athletes' movement quality. It presents a kinematic analysis of the difficult movements of competitive Tai Chi, taking the difficult movements of the swirl group of competitive aerobics as the research object. The data is collected by different methods. When completing difficult movements, due to different personal exercise habits, this paper takes the right leg as the swinging leg and the left leg as the take-off leg when completing the difficult movements. The test object in this paper adopts the traditional three-step up-step method. The starting time of the take-off stage is consistent with the illustration. By consulting experts for the stage division of the difficult movement, the two movement stages are equally divided into three stages according to the needs of the research: take-off, soaring, and landing. The results show that, from the moment the left foot hits the ground to the moment of maximum cushioning, the angles of the right knee are 176.48° and 165.06°, respectively, and the flexion of the right knee is not large; especially when the left foot touches the ground, the right leg is close to a straight line, indicating that the athlete's body is in a straight line. The posture is relatively good, the left foot is used as the grounding leg, the angle of the knee joint is reduced from 160.82° to 124.41°, and the left knee buffer angle is 36.40°. As the right leg swings to the ground, the left knee and left hip are slightly stretched, and the right knee and right hip are slightly flexed. This change is also reflected in the fact that the center of gravity changes from the direction of the horizontal speed at this stage. Move the left side of the body slightly to the right. Difficulty movements are different in the landing stage, and their joint parameters are also different. When the landing leg touches the ground, there is no significant difference in the maximum cushioning time, the degree of torsion of the shoulder and the hip is different, the relative rotation position of the shoulder and the hip is opposite, and the posture of the shoulders at the time of landing is different. In teaching, the landing leg and the end direction of the rotation are not the same; it is necessary to strengthen the strength training of the legs, to experience the force feeling of the different degrees of rotation of the shoulder joint and the hip joint, as well as the spatial perception ability, and take the initiative to actively land. The magnitude of the rear cushion and the coordination of the trunk and the limbs are used to avoid unnecessary damage during the landing phase.

1. Introduction

Wushu has rich cultural connotations. It is a traditional sports event that has been continuously created and gradually formed by countless sages of the Chinese nation in the course of thousands of years of historical evolution. It is an important part of China's traditional excellent historical culture [1]. As early as 1960s, China included martial arts as a course in the physical education syllabus of primary and secondary schools. Social popularization can achieve the purpose of promoting the spirit of China's national traditional culture while promoting the effect of fitness [2]. With the development of Wushu, it now occupies a place in physical education in colleges and universities. Physical education majors in colleges and universities all have martial arts specialties. Students majoring in physical

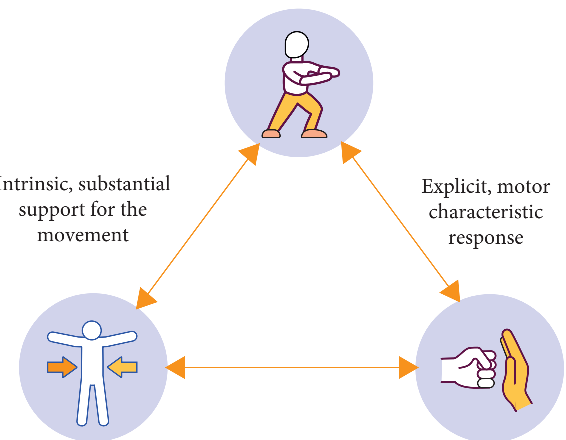


Figure 1: Research progress of sports biomechanics energetics and actions, such as simple jumping and flying.

education with higher skills and accomplishments choose one of the many sports (including martial arts) they have learned as their major. Students who choose Wushu as their major mainly learn simple Wushu routines, equipment, basic jumping movements, and Wushu theory and will mainly focus on Wushu teaching in primary and secondary schools in the future. At present, the analysis and research of high-difficulty movements in Wushu movements are the main direction of this research, thus ignoring the research on the basic movements [3]. For ordinary martial arts learners and enthusiasts with average physical fitness, most of them face these difficult martial arts' moves with an attitude of a dispensable attitude, and they are most exposed to the basics of martial arts moves in their learning and practice. Actions include simple jumping and flying. Figure 1 shows the progress in motion, biomechanics, and energetics. However, studies on such relatively simple and highly ornamental movements are rarely seen in the current research. The movement of "running and swaying the lotus 360° to land on one foot" is one of the basic jumping movements of martial arts, and it is also the movement learned by the martial arts students specializing in physical education [4]. This action is relatively difficult to master than other basic jumping action technical routes, but it is highly ornamental. There are a large number of studies at different levels in the research on Wushu technical movements, such as the research on the flying classes B- and C-level movements in Wushu routines, the research and analysis of a classic Taijiquan movement, and the analysis of a certain movement in Sanda.. Summarizing these studies, it is found that most of the technical movements studied and analyzed are highly difficult movements in competitive martial arts competitions, such as "flying the lotus 360° to lift the knees independently, swinging the lotus 540°, cyclone feet 720°, and spinner rotation of the body 360°;" there are few studies on the "run-up, swaying, and 360° single-leg landing" action.

2. Literature Review

Aiming at this research problem, Koryagina et al. obtained the characteristics of each joint angle, joint speed, force sequence, and other characteristics of the experimental subjects in difficult movements through three-dimensional kinematics analysis of the selected difficulty. Combined with sports anatomy, sports physiology, and according to the characteristics of sports mechanics, the dynamic parameters required in the process of Shoumin's superaction can be obtained, so as to analyze and study the results and provide a more regular and reasonable reference for the improvement and training of the difficult movements of the competition [5]. Rusdiana et al. analyzed Chinese outstanding competitive aerobics athletes through biomechanical research, obtained various kinematic parameters and data, obtained the parameters of athletes' movement techniques, obtained the laws of movement activities, and provided reference for actual training, judging the athletes' movement problems and the reasons for the problems, and put forward specific methods to solve the problems [6]. Keisuke et al. believe that difficult movements play a very important role in the complete set of competitive Taijiquan movements. The rules stipulate that a difficult movement should be selected, which determines not only the completeness of the set, but also the score and difficulty of the difficult movements which is an important factor in the score [7]. Rajkumar et al. discussed the characteristics of human body changes after aerobics exercise and the aspects of aerobics competition and training from the aspects of human physiology and anatomy [8]. da Silva Sores et al. also used biomechanics to dissect the perspective of body shaping and aerobic exercise [9]. A large number of documents reviewed by Biscarini et al. found that the current biomechanical research is mainly used in the analysis of human body structure and sports function, movement technology, sports injuries, planning and improvement of sports equipment, general fitness, and fitness guidance for special groups, etc. The proportion of research in this area is also increasing [10]. Perry et al. studied the mechanical characteristics of human activities and the general characteristics of movement changes and the science of sports equipment development [11]. Powell et al. pointed out that the current kinematic research methods of biomechanics mainly include camera and image analysis, infrared light point motion capture and analysis, electromagnetic induction motion capture, and analysis, among which three-dimensional camera and analysis methods have been widely used [12]. Hsiao et al. used the observation video method and the image analysis method to analyze the technical movements of the two research subjects. The width provides a reference in [13]. Lempke et al. conducted a biomechanical comparative analysis of five different levels of badminton players' backcourt smashing techniques by using three-dimensional camera measurement method. By capturing these reflective points, the obtained moving images are more vivid, which reduces the workload for subsequent moving image analysis [14]. The kinematics of aerobics movement decomposition based on the video tracking algorithm is discussed and analyzed. The feasibility of the aerobics motion decomposition based on multi-motion target video tracking algorithm is summarized. The experiment concluded that at the beginning the right hip angle was 167.45. As can be seen from the angles of both knees, the legs are relatively straight and remain in good position. When the jump contacts the ground, the difference between the two moments of the right hip angle is 29.8. At this time, the upper body starts to bend, stretching the left muscle, so that the hip joint and shoulder joint form a larger rotation angle [15].

Based on the current research, this paper proposes a kinematic analysis of the difficult movements of competitive Tai Chi. Taking the difficult movements of the swirl group of competitive aerobics as the research object, this article takes the right leg as the swing leg and the left leg as the take-off leg when completing the difficult movement, as an example. The results show that, from the moment the left foot hits the ground to the moment of maximum cushioning, the angles of the right knee are 176.48° and 165.06°, respectively, and the flexion of the right knee is not large; especially, when the left foot touches the ground, the right leg is close to a straight line, indicating that the athlete's body is in a straight line. The posture is relatively good, the left foot is used as the grounding leg, the angle of the knee joint is reduced from 160.82° to 124.41°, and the left knee buffer angle is 36.40°. As the right leg swings to the ground, the left knee and left hip are slightly stretched, and the right knee and right hip are slightly flexed. This change is also reflected in the fact that the center of gravity changes from the direction of the horizontal speed at this stage. Move the left side of the body slightly to the right. Difficulty movements are different in the landing stage, and their joint parameters are also different. When the landing leg touches the ground, there is no significant difference in the maximum cushioning time, the degree of torsion of the shoulder and the hip is different, and the relative rotation position of the shoulder and the hip is opposite, and the posture of the shoulders at the time of landing is different. In the teaching, the landing leg and the end direction of the rotation are not the same; it is necessary to strengthen the strength training of the legs, to experience the force feeling of the different degrees of rotation of the shoulder joint and the hip joint, as well as the spatial perception ability, and take the initiative to actively land. The magnitude of the rear cushion and the coordination of the trunk and the limbs are used to avoid unnecessary damage during the landing phase.

3. Methods

3.1. Research Objects. This paper takes the difficult movements of the swirl group of competitive aerobics as the research object.

3.2. Research Methods

3.2.1. Documentation Law. According to the research purpose and significance of this paper, through the CNKI retrieval platform, with the keywords of "competitive Taijiquan," "difficulty movement," "kinematics," and "sports biomechanics," the literature related to the research since 2018 was searched. 26 representative works, consulting the relevant books in the library and literature database of Beijing Sports University, understand the current research status of the current related competitive aerobics projects, especially the kinematic analysis of difficult movement techniques, carefully study the relevant materials, and classify the collected materials. The classification, summary, and analysis provide a theoretical basis for the writing of the research content of this paper.

3.2.2. Interview Method. According to the research needs of this article, consult the coaches in the field of competitive Tai Chi on the basic movement structure, technical key points, matters that need attention, and main training methods of jumping. At the same time, experts in sports biomechanics were consulted about the shooting design, stage division of difficult movements, and kinematic parameter requirements in this study [16].

3.2.3. Experimental Method. Competitive Taijiquan athletes were selected as test subjects, males, 173 cm tall, 63 kg in weight at the time of the test, and had a training period of 15 years.

3.2.4. Data Collection. Test equipment: the experimental data collection was completed in the Taijiquan gymnasium of a comprehensive training hall of a sports university. An 8-lens infrared light spot high-speed motion capture system (QUALISYS-MCU500) was used to collect the kinematic data of the test object completing difficult movements, and the sampling frequency was 200 Hz.

Site preparation: before data collection begins, draw all the curtains on the test site to block unnecessary light spots, calibrate the test area, and place the 8 lenses of the infrared light spot high-speed motion capture system in the front, left, and rear of the test area. For the positions of the right and left front, left rear, right front, and right rear, adjust the angle and height of the lens, connect to the computer for multiple corrections, and eliminate the interference of excess light spots [17, 18].

Athlete preparation: athletes arrive at the test site in advance to prepare for warm-up activities and change into shooting clothes. Paste marker points according to the Hele-Hayes mannequin, completed by students majoring in sports biomechanics.

Infrared spot calibration: the athlete enters the test area and calibrates the infrared spot on his body.

Test process: after the athletes are ready, the testers issue the start password and collect data on the athletes completing the specified difficulty movements, each of which is completed 6 times; 3 referees on the scene will score the completion of each action, according to the referee's score. In each case, each action with the highest score is selected for later data processing [19].

Table 1: Lower limb angle parameters in the take-off stage.

Feature screen	Time	Hip joint		Knee joint		Ankle joint	
		Right	Left	Right	Left	Right	Left
Start time	0	163.84	128.85	168.88	178.1	109.3	137.5
Take-off leg landing moment	0.346	137.24	143.51	169.96	155.03	96.22	117.74
The minimum moment of the knee angle of the take-off leg	0.374	67.96	48.24	144.77	121.21	89.08	86.7
The lowest moment of the head	0.424	76.33	32.68	153.13	127.54	88.41	87.52
Moment of swinging leg off the ground	0.48	116.24	43.94	168.63	141.35	118.74	87.67
The moment when the take-off leg leaves the ground	0.664	153.88	124.72	165.76	157.95	145.65	125.25

3.2.5. Division of Action Stages. Due to different personal exercise habits when completing difficult movements, this paper takes the right leg as the swing leg and the left leg as the take-off leg when completing the difficult movements. The test object in this paper adopts the traditional three-step up-step method. The starting time of the take-off phase is consistent with the diagram. By consulting experts for the division of the difficult movement, and, according to the needs of the research, the two movement stages are equal. It is divided into three stages: take-off, air-to-air, and landing.

The specific division is as follows.

Take-off stage: it is divided into single-foot support, double-foot support, and single-leg extension stage. In this study, when taking three steps up, the moment when the athlete swings the leg to the highest center of gravity in the third step was taken as the starting moment, and the moment when the take-off leg kicked off the ground was taken as the end point, and the process was defined as the take-off stage. The single-leg support stage is from the starting time to the time when the take-off leg touches the ground, the double-foot support stage is from the time when the take-off leg touches the ground to the time when the swing leg leaves the ground, and the single-leg stretch stage is from the time when the swing leg leaves the ground to the time when the take-off leg leaves the ground [20].

Flying stage: it is divided into rising and falling stages. The flying stage refers to the moment when the take-off leg leaves the ground to the moment when one foot touches the ground. The highest moment in the sky is the cutting point, and it is divided into the rising stage and the falling stage. Landing stage refers to the process in which the athlete is swinging his legs to the ground with both feet when one foot hits the ground.

3.2.6. Data Processing. The research data is processed and analyzed through the software in the Qualisys system to obtain the kinematic parameters required for the research, and the Cortex software is used to derive the body center of gravity parameters and the Euler angle of the torso relative to the pelvis [21].

3.3. Mathematical Statistics. According to the actual needs of the research, the data obtained are summarized, sorted, classified, calculated, and charted using Excel software.

3.4. Comparative Analysis Method. By analyzing the kinematic parameters of the athlete's swirl and swivel 360° difficult movements and comparing the technical characteristics of the two movements at the same stage, the similarities and differences in each link of the two difficult movements were found out [22].

4. Results and Analysis

4.1. Kinematics Analysis of Difficult Movements in the Take-Off Stage

4.1.1. Kinematic Analysis of Joint Parameters of Lower Limbs. It can be seen from Table 1 that the right hip is fully extended, the right hip joint angle is 163.84°, and the left hip joint angle is 128.85°. It can also be seen from the angle of the two knee joints that the legs are relatively straight and maintain a good body posture. When the take-off leg touches the ground, the difference between the right hip angles is 26.5°, and the left hip angle difference is −14.65°; that is, the right hip angle decreases and the left hip angle increases [23, 24]. The right hip angle decreases, indicating that the upper body begins to bend down while the upper body turns to the right, the left hip continues to abduct, and its value increases. When the left external oblique muscle is elongated, the shoulder-hip torsion angle also increases accordingly.

In the take-off stage, the upper body is also bent down and folded, and the hip joint angle first decreases and then increases. The left hip joint is flexed to a minimum value of 32.68° at the lowest moment of the head. In the subsequent hip extension process, the hip angle changes greatly; especially when the take-off leg leaves the ground, the right hip angle is 153.88°, and the left hip angle is 124.72°. Judging from the knee angle data, the minimum flexion degree of the left knee is 121.21°, while the hip flexion is more obvious at this moment, and the corresponding ankle joint angle is close to 90°, indicating that the knee flexion and squatting range is small during the take-off stage, mainly the upper body trunk. Movement is dominant. Combined with the angle of the hip and knee joints when the take-off leg is off the ground, the range of flexion of the hip and knee is small, which is conducive to the rotation of the body on the same longitudinal axis during the flight.

4.1.2. Kinematics Analysis of the Trunk Action Link. According to the definition of the positive and negative values of the shoulder-hip torsion angle in this paper, the positive value is the left rotation state of the shoulder relative to the hip, and the negative value is the right rotation state.

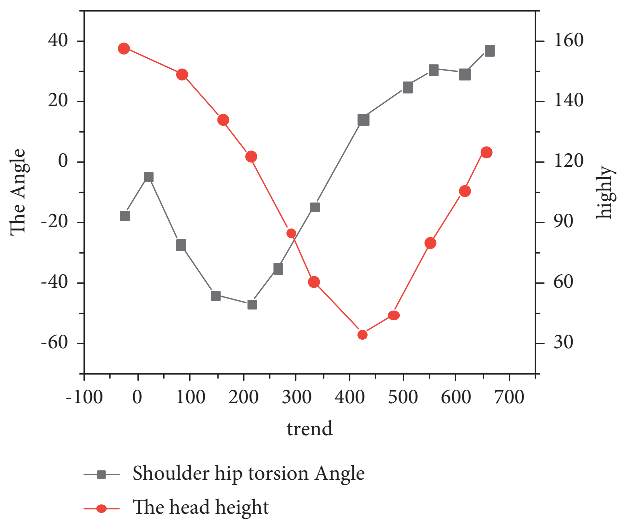


Figure 2: Shoulder-hip torsion angle and head height-time curve in take-off stage. It also means that the rotator rotates 360° at the lowest moment of the head to turn to the front and left of the body.

In the take-off stage, the amplitude of the shoulder and hip rotation also increases first, then decreases, and then increases, and the shoulder joint turns from right to left relatively to the hip joint.

It can be seen from Figure 2 that the shoulders are twisted to the right of the hips before the take-off and the reverse rotation of the hips and shoulders reaches the maximum value of −47.32°, which corresponds to the moment when the take-off leg touches the ground. With the movement of the upper body during the take-off phase, the shoulders are rotated to the left, reaching the maximum value of 37.75°; that is, when the shoulders and hips rotate at the same time, the shoulders are rotated 85.06° relatively to the hips. Combined with the change in the height of the head, the head movement curve first falls and then rises. The lowest moment of the head is after the rotation angle of the shoulder and hip is 0°, that is, the position in front of the body to the left. It can be seen from the data of the intersection of the vertical dotted line and the two curves in the figure that he height of the head at the lowest moment of the head is 25.97 cm and the corresponding torsion angle of the shoulder and hip is 15.32°. It also means that the rotator rotates 360° at the lowest moment of the head to turn to the front and left of the body.

Through the parameters of the torso relative to the pelvis torsion, the relative position of the torso space can be reflected more three-dimensionally. As shown in Table 2, when the take-off leg hits the ground, the trunk is flexed forward to −160.38°, and the lateral flexion is flexed from the original left to the right to −7.31. At the same time, the trunk turns right to −20.85°, which elongates the left external oblique muscle. When the buffer reaches the minimum knee angle of the take-off leg, the trunk has been bent forward to the minimum value of −123.34°, but in fact the forward flexion is greater, indicating that the pelvis is in a forward tilted state at this time. At the lowest moment of the head, the shoulders and hips are almost parallel, and only 1.28° left. At the moment of taking off and leaving the ground, the trunk turned to the left to a maximum value of 21.76°, and the left area continued to increase to 30.79°. The trunk moved from flexion to extension, and the angle was 179.00°, which almost coincided with the vertical axis of the body, indicating that the rotator rotates. The body 360° has a small upward lift when taking off, and it is mainly for the rotation during the flight phase to keep close to its vertical axis.

4.1.3. Kinematics Analysis of the Single-Leg Extension Stage. In this paper, the shoulder-foot torsion angle is defined as the angle of the line between the shoulders and the left foot in the horizontal plane, and the hip-foot torsion angle is the angle of the line between the two hips and the longitudinal axis of the left foot in the horizontal plane. In the single-leg extension stage, the shoulder-foot torsion angle and the hip-foot torsion angle continued to increase, and the trends were similar; that is, the shoulder and hip joints were twisting synchronously. When the rotation reaches the maximum value and then falls back, the shoulder-foot torsion angle is the same as the maximum hip-foot torsion angle. At this moment, the athlete's upper torso has rotated to the maximum in the horizontal plane. The shoulder-foot torsion angle is the line connecting the shoulders and the left foot in the horizontal plane. The maximum value of the included angle is 136.98°, and the maximum value of the hip-foot torsion angle, that is, the angle of the line connecting the two hips and the longitudinal axis of the left foot in the horizontal plane, is 105.03°. Due to the limitation of the physiological structure of the hip joint, the torsion amplitude is smaller than that of the shoulder joint.

After swinging the leg off the ground in the single-leg extension stage (Figure 3), the projection angle of the right hip on the horizontal plane has a maximum value and is close to 180°, indicating that the projections of the right shoulder, hip, and knee on the knee angular velocity show an upward trend after reaching the minimum value of 127.7°/s; that is to say, the minimum value corresponding to the left knee angular velocity is the timing of single-leg take-off when the rotor rotates 360°. In addition, the right hip angle at the time of single-leg extension was 139.88°. Although the hip was not fully extended, it can be seen from the curve of the right hip angle that, during the single-leg extension stage, the right hip angle continued to increase and the maximum value was about 170°; it means that the upper body torso and the right leg are close to the same-vertical axis, which is conducive to the rotation of the body around its own vertical axis in the subsequent flight stage.

4.2. Kinematics Analysis of the Flight Stage. In the flight stage, the rotation is mainly driven by the shoulder joint. According to the angle between the shoulders and the hips and the horizontal plane in the original data, combined with the height of the shoulder and hip joints, the rotation of the hip joint at the characteristic moment can better reflect the flight stage. Therefore, in order to facilitate the study, this paper selects the angle of the connection line between the two hip joints and the ground, that is, the hip inclination

Table 2: Torsion angle parameters of the trunk relative to the pelvis during the take-off stage.

Feature screen	Time (s)	Sagittal plane (°)	Frontal plane (°)	Level (°)
Start time	0	148.97	10.80	21.29
Take-off leg landing moment	0.346	−160.38	−7.31	−20.85
The minimum moment of the knee angle of the take-off leg	0.374	−123.34	6.70	−4.03
The lowest moment of the head	0.424	−126.51	10.29	1.28
Moment of swinging leg off the ground	0.48	−137.14	19.84	2.84
The moment when the take-off leg leaves the ground	0.664	179.00	30.79	21.76

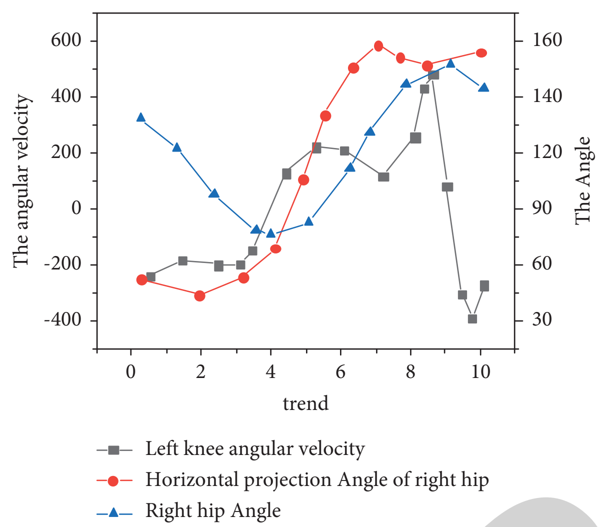


Figure 3: Left knee angular velocity, right hip horizontal plane projection angle, and right hip angle-time curve during single-leg extension. It means that the upper body torso and the right leg are close to the same-vertical axis, which is conducive to the rotation of the body around its own vertical axis in the subsequent flight stage.

angle as the parameter. In addition, due to the rotation process of the vertical plane of the hip joint, there will be a certain deviation, but the data deviation is relatively and highly tiny, so the maximum value in each revolution is defined as a special moment. In this paper, after the jump, the first time when the line between the two hips is parallel to the ground is defined as the moment when the hip joint rotates 0°, and there are four characteristic moments when the hip joint rotates 0°, 90°, 180°, and 270° relatively to the ground after taking off to analyze the state of the body in the air.

According to the data in Table 3, the duration of each 90° rotation of the athlete is 0.11s, 0.09s, and 0.125s, respectively. It can be seen that the rotation speed in the air is extremely fast, and their body posture is also changing all the time. Spread your legs apart, keep your body parallel to the ground, and turn your body 360°. The following is an analysis of the flight stage based on the rules and standards, combined with the side and top views of the four characteristic moments. Combined with the data in Table 3, at the beginning of the take-off and after the take-off, since the shoulder joint has begun to rotate during the take-off and extension stage, the shoulder and hip twist is the largest at 4 moments, which is 60.13°. Through the early rotation of the shoulder joint, effectively drive the body to rotate, increase the angular speed of rotation, and reduce exercise time. It can be seen from the table that the upper body and the right leg move towards each other and are obviously higher than the horizontal plane and the angle of the legs is 73.93°.

4.3. Kinematic Analysis of Landing Stage. During the flight stage, the shoulders and hips are mainly driven by the rotation. When preparing to land, the right leg and the torso should be rotated as much as possible while maintaining a horizontal posture. At the same time, the left leg will actively move to the ground to prepare for landing, and the arms will gradually expand. In the landing stage, after the left foot hits the ground, at the moment of landing, the body is almost facing the ground. During the transition from landing on one foot to landing on both feet, since the left foot hits the ground first, and the overall speed of the body in the direction of the horizontal plane is on the right side of the body position, the rotation speed in the air is relatively large. The right leg also continues to swing in the direction of motion while moving downward. The right leg continues to swing and arc down in the direction of the horizontal plane speed. It is shorter and has a higher movement speed, which increases the load of the landing leg accordingly. The upper body should not be lifted too early to ensure the stability of the center of gravity.

From the data in Table 4, at the moment when the left foot touches the ground, the torsion angle of the shoulder and hip is 18.2°, and the relative angle of the line connecting the shoulders and the hips is much smaller than that in the flying stage. The hip joint gradually returns to the same plane. At the moment when the left foot touches the ground, the left hip joint is 103.98°, slightly larger than 90°. This is the angle between the left leg and the horizontal plane at this moment, which is the upward direction of the landing. The degree of flexion and extension of the body trunk is related, and, from the video screen, it has a greater correlation with the direction when landing. From the moment when the left foot hits the ground to the moment of maximum cushioning, the angles of the right knee joint are 176.48° and 165.06°, respectively, and the flexion of the right knee is not large; especially when the left foot touches the ground, the right leg is close to a straight line, indicating that the athlete's body posture is relatively ok, the left foot is used as the grounding leg, the angle of the knee joint is reduced from 160.82° to 124.41°, and the left knee buffer angle is 36.40°. As the right leg swings to the ground, the left knee and left hip are slightly stretched, and the right knee and right hip are slightly flexed. This change is also reflected in the fact that the center of gravity changes with the direction of the

Table 3: Rotation parameters of the hip joint at four moments in the flight stage.

Feature screen	Duration	Split angle	The angle between the thigh and the horizontal plane		Shoulder torsion angle
			Right	Left	
Hip rotation 0° (rise in the air)	0	73.93	−28.61	42.05	60.13
Rotate the hip joint 90° (in the air)	0.10	40.23	1.44	6.41	56.67
Hip rotation 180° (highest in the air)	0.08	49.36	33.51	−14.14	46.43
Hip rotation 270° (flying down)	0.124	45.91	10.10	32.92	25.28

Table 4: Joint parameters in landing stage.

Feature screen	Hip angle		Knee angle		Shoulder torsion angle
	Right	Left	Right	Left	
The moment when the left foot touches the ground	155.36	103.98	176.48	160.82	18.2
Maximum buffer time	111.91	80.14	165.06	124.41	21.0
The moment your feet touch the ground	97.92	89.52	155.94	137.83	10.7

horizontal speed at this stage. Move the left side of the body slightly to the right.

In general, the difficult movements have different landing legs in the landing stage, and their joint parameters are also different. When the landing leg touches the ground, there is no significant difference in the maximum cushioning time and the degree of torsion of the shoulder, and the hip is different, and the relative rotation position of the shoulder and the hip is opposite, and the posture of the shoulders at the time of landing is different. In teaching, the rotation direction of the landing leg is different from the rotation direction of the upper body. It is necessary to strengthen the leg strength training to experience the force sense of the shoulder joint and hip joint at different degrees of rotation, as well as the spatial perception ability. The magnitude of the rear cushion and the coordination of the trunk and the limbs are considered to avoid unnecessary damage during the landing phase.

5. Conclusion

This paper proposes a sports biomechanical analysis of jumping difficult movements in Tai Chi. Using the literature method, interview method, experimental method, etc., with the help of infrared light spot high-speed motion capture system, the kinematic data of athletes completing difficult movements is collected, the technical characteristics of difficult movements are analyzed, and the key technical characteristics of this group of difficult movements are explored to promote Tai Chi diversified development of boxing. In the take-off stage, the hip is also mainly flexed. When the single foot is extended, the projection angle of the horizontal plane of the right hip tends to be a straight line, and the right hip angle is 140°. It is beneficial to the rotation around its own vertical axis in the subsequent flight stage. At the beginning of the flight, the angular velocity is increased by the prerotation of the shoulder and hip joints, and the rotation radius is reduced and the kinetic energy loss is reduced with the approach of the limbs to the vertical axis of the body; the torsion of the shoulder and hip is relatively large in the stage of flight and ascent, and the rotation is dominated by the shoulder landing. During the stage, the load on the grounding leg is relatively large, and the upper body should be avoided to be lifted prematurely. In the future, in the practice of strengthening students' basic skills, solid basic skills are the basis for improving the quality of technical movements. In view of the unreasonable athlete's approach route and take-off foot abduction angle, it is explained in a targeted manner, and repeated training can strengthen the athlete's movement awareness and improve the movement quality.

Data Availability

The data used to support the findings of this study can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This paper was supported by Jiangxi Provincial Social Science Fund Project of the 14th Five-Year Plan (2021) (project no. 21TY01).

References

[1] A. V. Shpakov, A. V. Voronov, A. A. Artamonov, D. O. Orlov, and A. A. Puchkova, "Biomechanical characteristics of walking and running during unloading of the musculoskeletal system by vertical hanging," *Human Physiology*, vol. 47, no. 4, pp. 419–428, 2021.

[2] D. Häske, L. Schier, J. O. Weerts et al., "An explorative, biomechanical analysis of spine motion during out-of-hospital extrication procedures," *Injury*, vol. 51, no. 2, pp. 185–192, 2020.

[3] M. Pabón-Carrasco, A. Castro-Méndez, S. Vilar-Palomo, A. M. Jiménez-Cebrián, I. García-Paya, and I. C. Palomo-

Hindawi
Mobile Information Systems
Volume 2022, Article ID 9554359, 11 pages
https://doi.org/10.1155/2022/9554359

Research Article

Ecological Integrity Evaluation of Organically Evolved Cultural Landscape

Wenting Xu

Hainan University, Haikou 570228, China

Correspondence should be addressed to Wenting Xu; xuwt@hainanu.edu.cn

Received 10 June 2022; Revised 27 June 2022; Accepted 18 August 2022; Published 5 September 2022

Academic Editor: Jiafu Su

When the traditional method is used to evaluate the cultural landscape as a whole, due to the lack of historical and cultural factors, the impact of the ecological environment on the landscape change is not considered. Therefore, it cannot objectively reflect the law of landscape change. Therefore, from the perspective of the combination of cultural landscape composition characteristics and cultural ecological characteristics, this paper puts forward the concept of cultural ecological integrity based on the existing concept of cultural heritage integrity, aiming at the organic evolution of cultural landscape in cultural landscape heritage, and expounded the dynamic sustainability and effectiveness of natural environment and social environment on the process of cultural landscape. Considering the relevant characteristics and different forms of cultural landscape and cultural ecological integrity, according to the concept of cultural ecological integrity, this paper gave the principles and standards of cultural landscape ecological integrity evaluation and put forward the evaluation method of cultural ecological integrity. At the same time, each index in the ecological integrity evaluation model was decomposed and explained. Finally, taking the cultural heritage of Long-men Grottoes as an example, the paper used the cultural and ecological integrity evaluation model proposed in this paper to evaluate the material cultural landscape, intangible cultural landscape, and local traditional dwellings of Long-men Grottoes and made a comparative analysis between the control group and the observation group. The results showed that the cultural landscape ecological integrity evaluation model proposed in this paper can not only objectively evaluate the cultural landscape but also provide a certain theoretical and practical reference for the effective classification, protection and sustainable development, and utilization of the subsequent cultural landscape.

1. Introduction

Cultural landscape usually has the characteristics that nature and culture complement each other, and material and spirit promote each other. It is known from the existing research that the organically evolved cultural landscape comes from the basic needs of society, economy, administration and religion [1]. The research shows that due to the mutual integration and adaptation of landscape and natural environment, various existing cultural landscapes have been formed. For the dynamic cultural landscape existing in the region or widely distributed, for example, the type of cultural landscape with organic evolution, its value lies in the sustainability of the landscape [2]. It has always maintained a positive social role in today's natural and social environment related to tradition, and is the material evidence of its historical evolution and development.

The cultural ecological integrity of cultural landscape is a reference system and classification reference standard for the evaluation and protection of cultural landscape based on the registration standard of cultural landscape heritage and the evaluation of world cultural heritage issued by the World Heritage Committee (WHC) from the basic standpoint of cultural ecology [3]. In view of the complexity of cultural landscape, cultural ecological integrity effectively considers the natural environment, cultural environment and cultural landscape architecture, which can provide a more comprehensive and dynamic unique perspective for the protection and evaluation of world cultural landscape.

The 2005 edition of the operational guidelines for the implementation of the convention for the protection of the world cultural and natural heritage (hereinafter referred to as the operational guidelines) clearly points out that the authenticity and integrity of the landscape has become an organic part of the outstanding universal value of the world heritage [4, 5]. The analysis of the integrity assessment of relevant cultural landscape heritage shows that people's classification and analysis of the integrity of general cultural heritage can be applied to cultural landscape [6]. However, for the dynamic cultural landscape existing in the region or widely distributed, for example, for the type of cultural landscape with organic evolution, its integrity is more important. Because the existing landscape integrity evaluation is not comprehensive or targeted, this paper proposed to establish a cultural ecological integrity evaluation system of cultural landscape on the basis of ensuring the authenticity of cultural landscape from the perspective of cultural ecology.

2. Related Concepts of Cultural Ecological Integrity of Cultural Landscape

2.1. Characteristics of Cultural Ecological Integrity. The research shows that the cultural ecological integrity of cultural landscape is a landscape attribute formed through continuous evolution and continuation on the basis of the existing operation guidelines on integrity [7]. It reflects the dynamic sustainability and effectiveness of the natural and social environment on the evolution of cultural landscape. Since 1992, experts in cultural landscape and architecture related fields have begun to pay attention to the integrity, emphasizing the integrity, paying attention to the integrity of function, structure and vision, and paying attention to the current situation of static cultural landscape [8, 9]. Therefore, the cultural ecological integrity of cultural landscape should have the following characteristics:

(1) Cultural ecological integrity belongs to the cultural landscape of organic evolution: the organic evolution process of cultural landscape is generally closely related to traditional living habits, and is constantly changing with the changes of modern society. Among many cultural landscapes, some cultural landscapes with relevance are called composite landscapes. Their cultural meaning is determined by the relationship between natural factors and human religion, art, history and culture. Most of these complex landscapes are natural scenic spots protected by human beings, such as scenic spots and religious shrines, which have certain characteristics of ecological integrity.

(2) The cultural ecological integrity of cultural landscape has a certain linear change law of cultural landscape: International Society for Landscape Ecology (ISLE) and American National Geographic Society (ANGS) have conducted in-depth research on cultural landscape on the basis of landscape natural ecology, emphasizing the necessity of integrating cultural landscape and cultural ecology [10]. Some scholars put forward the theory of cultural continuity from the perspective of anthropology and sociology, holding that different cultures have certain continuity in their own development process. Landscape culture itself has more social culture. It not only involves social ethics, religion, customs and other concepts, but also contains a lot of artistic and cultural content [11, 12]. Therefore, with the continuous change of cultural landscape and environmental conditions, the cultural ecological integrity of cultural landscape can make the landscape culture show a certain linear change law.

(3) Cultural landscape has certain communication, and the landscape culture of the same culture has regional differences: From the perspective of the development of cultural ecology, cultural ecology mainly studies the interaction and correlation between natural geographical environment and human culture. Cultural ecosystem is a complex of nature, culture and economy [13]. Due to the interaction between culture and ecological environment, within a certain regional scope, the complex natural and social environment usually makes the development of different landscape cultures present a certain diversity, and thus forms regional differences.

As shown in Figure 1, it is the schematic diagram of the discipline relationship of cultural landscape.

2.2. Main Manifestations of Cultural Ecological Integrity. From the perspective of integrity operation guide, the ecological integrity of organic evolutionary cultural landscape is mainly reflected in the following four points:

(1) Cultural landscape can adapt to the existence of integrity and constantly evolve with the changes of natural or social environment.

(2) Cultural landscape can provide a certain use function and social value for related fields, can coexist harmoniously with today's social environment and natural environment, and has a certain cultural value.

(3) In the historical process of the coexistence of single or multiple cultural landscapes of the same type, the changes and development of cultural landscapes can be effectively observed and identified. For example, various forms of relevant historical data can be used for recording and archiving.

(4) For the cultural heritage to which the cultural landscape belongs, various forms and styles of cultural landscapes that meet the above three characteristics and have the same subordinate cultural types can coexist, thus forming a relatively clear geographical space and reflecting a representative cultural transmission path.

As shown in Figure 2, it is the basic composition of cultural landscape.

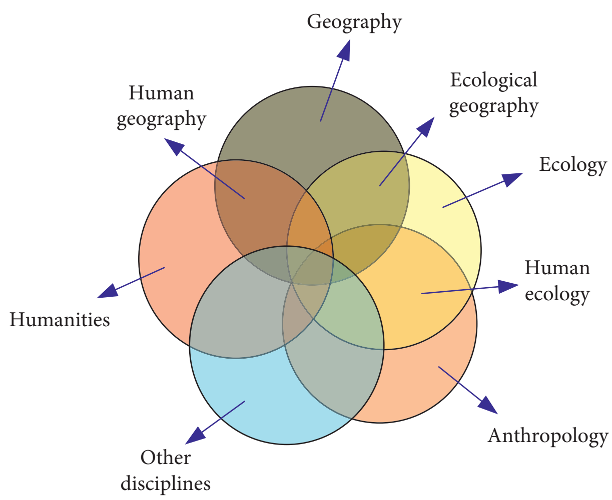


Figure 1: Schematic diagram of discipline relationship of cultural landscape.

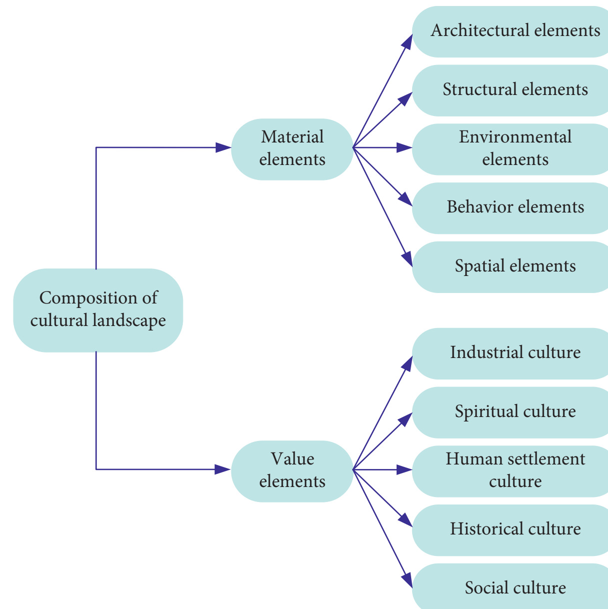


Figure 2: Basic composition of cultural landscape.

2.3. Research Value of Cultural Ecological Integrity. According to the landscape integrity theory put forward by predecessors, through the improvement of the concept of cultural landscape, it is known that the current cultural landscape is still related to history in terms of function and survival characteristics. At present, the integrity of cultural landscape seen from the visual point of view is only a result and fragment of the historical process. Therefore, in order to accurately understand the changes and authenticity of the heritage, it is crucial to analyze the authenticity of its function and historical structure [14].

Cultural landscape cultural ecological integrity has the significance of single and overall layout in terms of landscape scale and spatial characteristics. The cultural ecological integrity of cultural landscape mainly emphasizes the dynamic change process of the form, field environment and function, land use, life and spiritual belief of cultural landscape. The significance of studying cultural landscape ecological integrity from the perspective of cultural ecology is to fully reflect the social and historical value of cultural landscape integrity [15]. The cultural landscape with good cultural ecological integrity can provide more complete social exchange information for people in the historical process. From the utilization of natural resources to the progress of architectural technology, from the migration of cultural communicators and the spread of culture to the changes of society and the persistence of cultural beliefs are important historical materials.

3. Study on Cultural Ecological Integrity Evaluation of Cultural Landscape

3.1. Standards to be Followed for Cultural Ecological Integrity Evaluation. Cultural ecological integrity based on cultural landscape is a concept put forward from the perspective of cultural ecology, which can reflect the ecological characteristics of cultural landscape. Therefore, according to the integrity of traditional cultural landscape heritage and modern cultural ecology, the evaluation criteria of modern cultural ecological integrity can be obtained, as shown in Table 1.

According to the evaluation requirements in Table 1, after a certain historical period, the traditional cultural landscape heritage evaluation mainly focuses on the evaluation criteria of relative static and existing status. On this basis, a new evaluation standard can be formed from the perspective of modern cultural ecology, that is, pay more attention to the relationship between modern culture, landscape and architecture in the evaluation process. For example, landscape culture plays an important role in modern cultural ecology [16]. Therefore, from the perspective of traditional cultural landscape and modern cultural ecology, we can put forward the ecological evaluation standards of cultural landscape.

(1) Reference of evaluation criteria: due to the lack of research on the existing evaluation methods or systems for the ecological integrity of cultural landscapes at home and abroad, on the basis of the existing concept of cultural heritage integrity, we can refer to the existing evaluation systems for the integrity and authenticity of cultural landscapes, which mainly include empirical evaluation, single factor evaluation and comprehensive factor grading and quantification from the theoretical basis and technical means. Then, combined with the existing views and research methods of cultural ecology, further improve the evaluation criteria of cultural landscape ecological integrity.

(2) The scientificity of the evaluation criteria: in the process of evaluating the ecological integrity of cultural landscape, the established evaluation index system should not only fully reflect the internal mechanism of the ecological integrity of

Table 1: Integration of traditional cultural landscape integrity and modern cultural ecology.

Integrity of traditional cultural landscape heritage	Evaluation concept of modern cultural ecology
Landscape elements and their composition	Integration of modern culture and ecosystem
Landscape material and technology	Dynamic regulation of constituent elements
Traditional cultural landscape and its heritage	Tolerance of modern culture
Complementarity of traditional cultural heritage	Stratification of modern cultural system

the research object, but also systematically describe the connotation and target expectation of the evaluation object and its resources on the basis of relevant cultural landscape system theories.

(3) The dominance of evaluation indicators: the selected evaluation indicators should be representative and typical. When designing the cultural landscape integrity evaluation system, we should not only carefully screen and demonstrate the relevant indicators, but also reflect the importance of different indicators through weighted processing. In addition, the structure of the evaluation index system should be clear and concise.

(4) Guidance of evaluation criteria: the purpose of evaluating the ecological integrity of cultural landscape is to better protect the existing cultural landscape and achieve its sustainable development. Therefore, the establishment of corresponding index points in the evaluation index system can not only provide guidance for the relevant government departments to manage cultural landscape, but also provide scientific reference for the protection, development and utilization of cultural landscape resources.

3.2. Main Methods of Cultural Ecological Integrity Evaluation. The main methods of cultural ecological integrity evaluation include the comprehensive analysis method of qualitative and quantitative treatment for multifactor, multilevel and multi-index problems [17]. In order to make the cultural landscape ecological integrity evaluation model accurately and objectively reflect the evolution law of landscape objects, this paper fully draws on the experience of domestic and foreign scholars in the integrity and authenticity evaluation of cultural landscape heritage, and the proposed landscape evaluation method can not only describe the evolution characteristics of cultural landscape, but also reflect the value of cultural landscape. At the same time, based on the existing evaluation methods of predecessors, the evaluation system of cultural landscape and cultural ecological integrity is established by using sampling survey method, fuzzy evaluation method, principal component analysis method and analytic hierarchy process [18]. As shown in Figure 3, it is a schematic diagram of the research process of cultural landscape ecological integrity.

3.3. Construction of Cultural Ecological Integrity Evaluation Index System. When constructing the evaluation index system of cultural ecological integrity, this paper mainly analyzes the structure and constituent factors of cultural landscape according to the existing cultural ecological integrity standards of cultural landscape and the requirements of future protection and sustainable development, so as to make it related to the cultural ecology of cultural landscape [19]. In this paper, analytic hierarchy process (AHP) is mainly used to construct the evaluation index system of cultural ecological integrity, as shown in Figure 4.

In the cultural ecological integrity evaluation index system, the determination of each index weight is mainly based on analytic hierarchy process, including the construction of hierarchical structure model and the establishment of pairwise comparison matrix [20]. Using the paired comparison method, the paired comparison matrix is established from 1 to 9 comparison scales to the lowest level. The meaning of pairwise comparison matrix is shown in Table 2.

For the calculation of the weight vector, if the maximum eigenvalue of matrix A is $\lambda_{\max}$, the corresponding eigenvector is $w = \begin{pmatrix} w_1 & w_2 & \cdots & w_n \end{pmatrix}^T$, then $a_{ij} = (w_i/w_j)$, where $i, j = 1, 2, \cdots, n$, as follows

$$A = \begin{bmatrix} \dfrac{w_1}{w_1} & \dfrac{w_1}{w_2} & \cdots & \dfrac{w_1}{w_n} \\ \vdots & \vdots & \vdots & \vdots \\ \dfrac{w_n}{w_1} & \dfrac{w_n}{w_2} & \cdots & \dfrac{w_n}{w_n} \end{bmatrix}. \tag{1}$$

For consistency inspection, calculate the consistency index CI first, shown as follows:

$$CI = \frac{\lambda_{\max} - n}{n - 1}. \tag{2}$$

Then, the average random consistency index RI is searched correspondingly. For $n = 1, 2, \cdots, 9$, the consistency ratio CR can be calculated by the ratio of CI to RI. When $CR < 0.1$, it is considered that the consistency of the judgment matrix is acceptable, otherwise the judgment matrix should be properly modified.As shown in Table 3, it is the average random consistency index RI value.

Finally, calculate the combination weight vector and check the consistency of the combination results, as shown in Table 4.

The combination consistency test is similar to the above method, and the calculation formula is

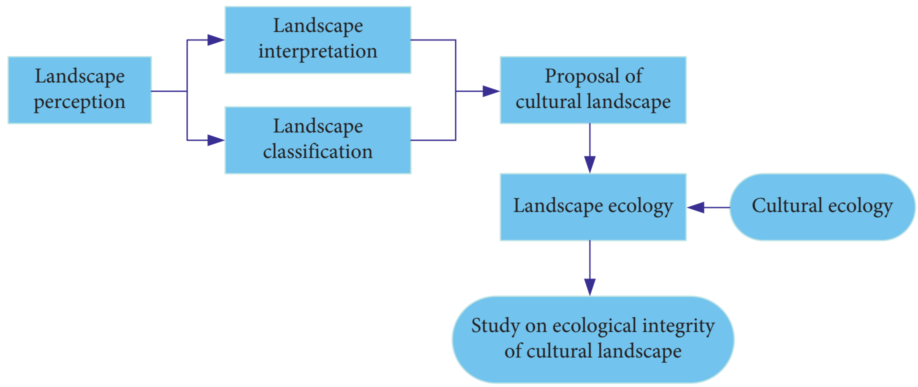


Figure 3: Schematic diagram of cultural landscape ecological integrity research process.

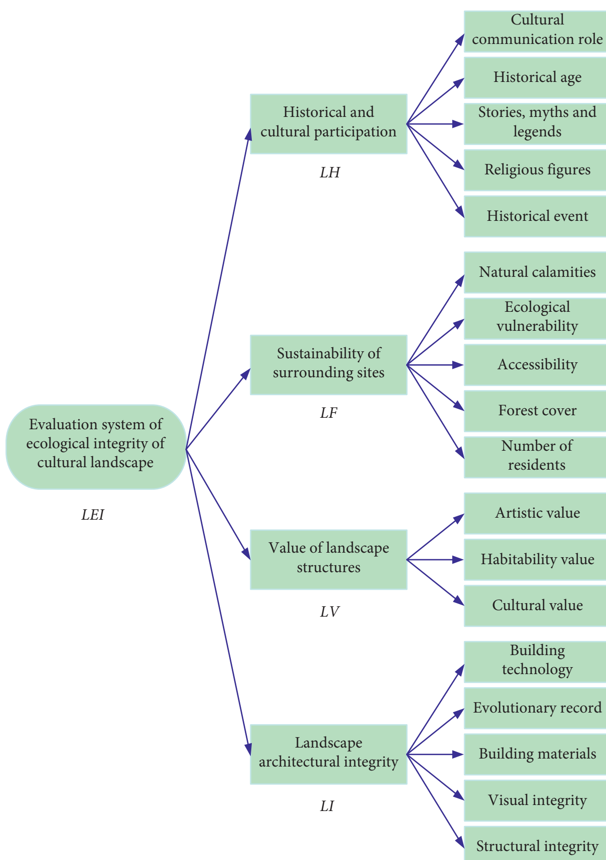


Figure 4: Evaluation index system of ecological integrity of cultural landscape.

$$b_i = \sum_{j=1}^{m} b_{ij} a_j, \quad i = 1, 2, \cdots, n,$$
$$CR = \frac{\sum_{j=1}^{m} CI(j) a_j}{\sum_{j=1}^{m} RI(j) a_j}. \tag{3}$$

When $CR < 0.1$, it is considered that the consistency of the judgment matrix is acceptable; otherwise, the judgment matrix should be properly modified.

4. Cultural Ecological Integrity Evaluation Model and Index Description

In the target layer of the evaluation index system, it mainly reflects the comprehensive index of cultural landscape ecological integrity evaluation by integrating different evaluation factors, which is expressed by the comprehensive index A of cultural landscape ecological integrity, and its evaluation index is expressed by LEI. It consists of four indicators, among which, the landscape history and cultural participation are represented by B_1, the indicator is named LH, the sustainability of surrounding sites is represented by B_2, the indicator is named LF, the landscape structure value is represented by B_3, the indicator is named LV, and the landscape architectural integrity is represented by B_4, the indicator is named LI. For the cultural ecological integrity evaluation of cultural landscape, descriptive indicators and evaluation indicators are selected to reflect the ecological integrity of the landscape in time and space respectively.

The ratio of standard layer and factor layer represents the evaluation weight of each index, which is marked as p_i^h, p_i^f, p_i^v and p_i^i respectively. The standard value can be comprehensively determined and recorded as f according to the evaluation and scoring of experts in relevant research and cultural landscape related fields [21, 22]. Each index evaluation model of the ecological integrity evaluation standard layer can be expressed as follows:

$$\begin{aligned} LH &= \sum_{i=1}^{5} p_i^h f_i^h \sum_{i=1}^{5} p_i^h = 1, \\ LF &= \sum_{i=1}^{5} p_i^f f_i^f \sum_{i=1}^{5} p_i^f = 1, \\ LV &= \sum_{i=1}^{3} p_i^v f_i^v \sum_{i=1}^{3} p_i^v = 1, \\ LI &= \sum_{i=1}^{5} p_i^i f_i^i \sum_{i=1}^{5} p_i^i = 1. \end{aligned} \tag{4}$$

Using the participation of landscape culture, regional sustainability, and the integrity of landscape architecture structure and value, we can build a comprehensive evaluation system of cultural ecological integrity of cultural

Table 2: Significance of pairwise comparison matrix.

Scale	Meaning
1	When the two factors are compared, they are of the same importance
3	When two factors are compared, the former is slightly more important than the latter
5	When two factors are compared, the former is obviously more important than the latter
7	When two factors are compared, the former is more important than the latter
9	When two factors are compared, the former is much more important than the latter
2, 4, 6, 8	Represents the intermediate value of the above judgment

Table 3: Average random consistency index value.

n	1	2	act	4	5	6	7	8	9	10	11	12
RI	0	0	0.51	0.87	1.14	1.26	1.33	1.42	1.47	1.48	1.53	1.56

Table 4: Calculation of combined weight vector.

A	A_1	A_2	⋯	A_m	Total sorting weight of layer$B$
B	B_1	B_2	⋯	B_m	
B_1	b_{11}	b_{12}	⋯	b_{1m}	$\sum_{j=1}^{m} b_{1j}a_j$
B_2	b_{21}	b_{22}	⋯	b_{2m}	$\sum_{j=1}^{m} b_{2j}a_j$
⋯	⋯	⋯	⋯	⋯	⋯
B_n	b_{n1}	b_{n2}	⋯	b_{nm}	$\sum_{j=1}^{m} b_{nj}a_j$

Table 5: Detailed rules for the evaluation of cultural ecological integrity from the perspective of historical and cultural participation.

Evaluation criteria	Evaluation factor	Evaluation weight (p_i^h)	Description of evaluation indicators (f_i^h)
Historical and cultural participation (B_1)	Cultural communication role	0.30	Assign values according to the level of cultural transmission, and refer to the size of cultural radiation area
	Historical age	0.18	The buildings of different ages are divided into five categories, including contemporary buildings, modern buildings, near ancient buildings, middle ancient buildings and ancient buildings
	Stories, myths and legends	0.13	Determined according to the size and amount of landscape influence
	Religious figures	0.18	Determined according to the size and amount of landscape influence
	Historical event	0.21	Determined according to the size and amount of landscape influence

landscape. Through the organic integration of the above four landscape cultural attributes, the comprehensive evaluation of cultural ecological integrity can be realized. Because these cultural attributes are independent of each other and have the same status for the comprehensive evaluation of cultural ecological integrity, the weights of these four factors are equal [23]. The relationship between the comprehensive evaluation index of cultural landscape ecological integrity and each evaluation index can be expressed as follows:

$$LEI = \sqrt[4]{LH * LF * LV * LI}. \tag{5}$$

In each index classification of the evaluation index system, landscape architecture, ecological environment, and cultural environment are the three subsystems of cultural landscape, which integrate the dynamic and open ideas of cultural landscape. In the cultural landscape ecological integrity evaluation system, the standard layer B_1 means to evaluate the historical and cultural participation, so as to dynamically describe the integrity of the cultural environment from the time dimension, as shown in Table 5.

After experts in the field of urban and rural planning and landscape ecology assess the ecological integrity of cultural landscape, the average value of relevant indicators of urban and rural planning and landscape design is determined [24]. The research goal of this paper is to evaluate the cultural ecological integrity on the basis of analyzing the organic evolutionary characteristics of cultural ecological integrity. Therefore, this paper mainly establishes the cultural ecological

Table 6: Evaluation rules of cultural ecological integrity from the perspective of space opening.

Evaluation criteria	Evaluation factor	Evaluation weight (p_i^f)	Description of evaluation indicators (f_i^f)
Sustainability of surrounding sites (B_2)	Natural calamities	0.18	Taking the occurrence frequency of earthquakes with magnitude greater than 4 as the characteristic of dominant disasters, the lower the frequency, the higher the corresponding index value. Similar evaluation methods shall be adopted for other related natural disasters
	Ecological vulnerability	0.15	The ecological stability of cultural landscape is evaluated from the five ecosystems of cold and drought, grassland, agriculture and animal husbandry, farming and forest, so as to ensure the ecological integrity of the whole cultural landscape
	Accessibility	0.38	The evaluation standard of accessibility index is mainly based on the sustainable impact of traffic on cultural landscape, taking it as a reference, taking advanced modern traffic as the highest level, simple modern traffic as the second level, human and animal power buffer as the third level, and human and animal power as the fourth level directly
	Forest cover	0.1	Based on forest coverage
	Number of residents	0.19	Determined according to the size and amount of landscape influence

Table 7: Evaluation rules of cultural ecological integrity from the perspective of landscape structure and cultural value.

Evaluation criteria	Evaluation factor	Evaluation weight (p_i^v)	Description of evaluation indicators (f_i^v)
Value of landscape structures (B_3)	Artistic value	0.3	Evaluation and assignment by relevant experts from the qualitative aspect
	Habitability value	0.24	Evaluate the residential function from the perspectives of heating, heat preservation, storage, and sanitation
	Cultural value	0.46	Evaluation and assignment by relevant experts from the qualitative aspect

Table 8: Evaluation rules of cultural ecological integrity from the perspective of landscape structure itself.

Evaluation criteria	Evaluation factor	Evaluation weight (p_i^i)	Description of evaluation indicators (f_i^i)
Landscape architectural integrity (B_4)	Building technology	0.16	Qualitative rating shall be carried out according to the content of traditional building technology used in cultural landscape from high to low
	Evolutionary record	0.24	The grading evaluation shall be conducted according to whether there are site selection literature records, landscape construction process records, repair records, destruction records and the number of recorded items
	Building materials	0.16	The landscape structure materials are evaluated according to the sensitivity. The stone has the lowest sensitivity, the highest relative integrity, the soil is in the middle, the wood is the lowest, and the stone soil mixture and stone wood mixture take the middle value in turn
	Visual integrity	0.24	Rating based on visual effects of existing landscape structures
	Structural integrity	0.2	The evaluation shall be conducted according to the following conditions: no major renovation, multiple improvements, complete reconstruction, remains of the site and the site of the site

integrity evaluation system from the perspective of cultural ecology, without involving the specific index weight value.

B_2 is mainly used to evaluate the ecological environment integrity of spatial open regulation, as shown in Table 6. It can describe the spatial open regulation of ecological environment integrity from five aspects, namely accessibility, number of residents, ecological vulnerability, forest coverage, and natural disasters.

B_3 is mainly used to evaluate the integrity of landscape structure and cultural value, as shown in Table 7. Among them, functionality refers to the functional value of human settlements, which mainly evaluates the ecological integrity

Table 9: Index classification and comprehensive evaluation standard of cultural ecological integrity.

Integrity grading assignment HL		Incomplete	Inferior	Secondary	Superior	Advanced
Assignment		1	3	5	7	9
Grading standard		[1, 2]	(2, 4]	(4, 6]	(6, 8]	(8, 9]
LH	f_1^h	Contemporary	Modern	Paleozoic	Medieval	Ancient
	f_2^h	No	Few	Some	Multi	Several important
	f_3^h	Pure	Individual	General	Important	Important nodes
	f_4^h	No	Small	Many	Wide	Wide and numerous
	f_5^h	No	Few	Some	Important	Many and influential
LF	f_1^f	Inaccessible	Direct entry	Buffer entry	Simple	Advanced
	f_2^f	<10	10~100	100~1000	1000~5000	>5000
	f_3^f	Cold arid	Grassland	Husbandry	Farming	Forest
	f_4^f	<0.1	0.1-0.2	0.2–0.4	0.4–0.6	>0.6
	f_5^f	Frequently	Multiple	Episodic	Less	No record
LV	f_1^v	No	Unsuitable	Poor suitability	Basically suitable	Suitable
	f_2^v	—	—	—	—	—
	f_3^v	—	—	—	—	—
LI	f_1^i	No	Any record	Any two	Any three	Four or more
	f_2^i	Timber	Earth	Stone wood mixture	Stone soil mixture	Stone
	f_3^i	Fully modern	Modern and traditional	Traditional and modern	Traditional technology	Traditional architectural
	f_4^i	Relic land	Remains of the site	Complete reconstruction	Repaired	No major renovation
	f_5^i	Dilapidated	Damage	Relatively complete	Basically complete	Complete

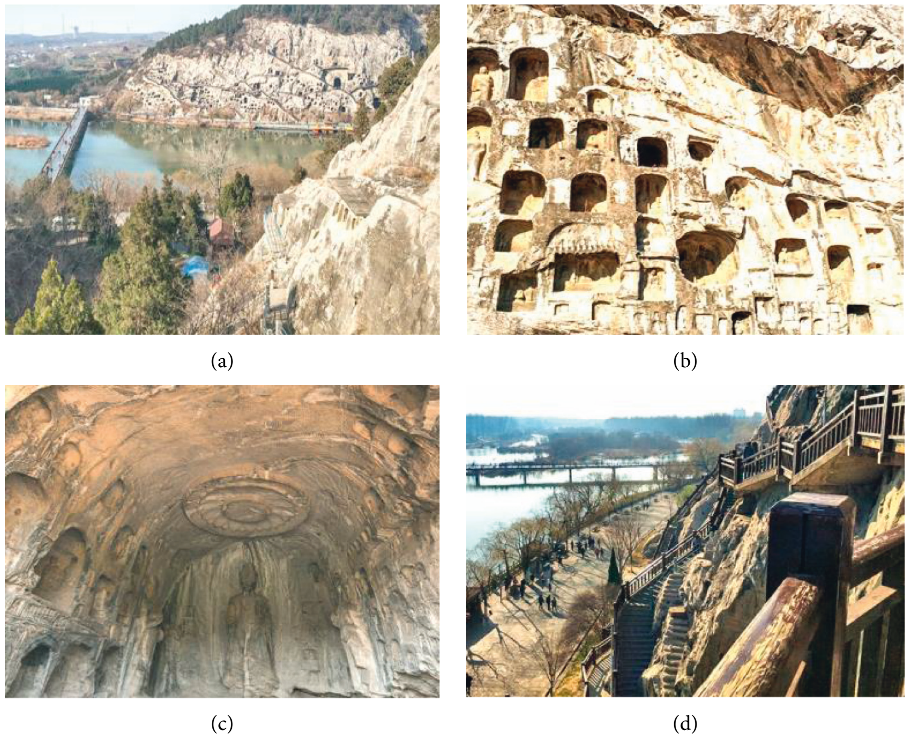


Figure 5: Some cultural landscapes of Long-men Grottoes. (a) Remote landscape, (b) exterior view of Grottoes, (c) Giant Buddha, and (d) Plank road.

Table 10: Evaluation results of physical cultural heritage integrity.

Site cultural landscape integrity (100%)						Total
Quantitative assessment (60%)			Qualitative assessment (40%)			
Historical time (10%)	Degree of scarcity (10%)	Landscape scale (30%)	Landscape richness (10%)	Artistic value (25%)	Scientific value (15%)	43%
5%	7%	8%	5%	12%	6%	

Table 11: Evaluation results of intangible cultural heritage integrity.

Integrity of intangible cultural landscape (100%)						Total
Quantitative assessment (60%)			Qualitative assessment (40%)			
Historical time (10%)	Degree of scarcity (10%)	Landscape scale (30%)	Landscape richness (10%)	Artistic value (25%)	Scientific value (15%)	26%
3%	2%	5%	4%	8%	4%	

from the aspects of heating, insulation, storage, and sanitation. The other two items need to be qualitatively assessed according to relevant experts. B_4 is mainly used to evaluate the integrity of the landscape structure itself, as shown in Table 8.

To evaluate the cultural ecological integrity of cultural landscape from the perspective of cultural ecology, because it involves the natural field, cultural ecological characteristics and cultural factors of cultural landscape, it is necessary to consider the comprehensive evaluation of time, space and perceptual activities. Referring to the cultural landscape evaluation index system, the index adopts a five-level standard combining qualitative and quantitative.

Since the evaluation grading values of LH, LF, LV, and LI are all between 1 and 9, expressed as $LH \in [1, 9]$, $LF \in [1, 9]$, $LV \in [1, 9]$, and $LI \in [1, 9]$, we can know $LEI \in [1, 9]$ from formula (5). Therefore, LEI can be divided into [1, 2], (2, 4], (4, 6], (6, 8], and (8, 9] five ecological integrity, which respectively correspond to the incomplete, low integrity, medium integrity, excellent integrity, and high integrity of cultural ecological integrity of cultural landscape. Different types of cultural landscape can be refined into different evaluation data for targeted classification research. As shown in Table 9, the evaluation index classification and evaluation criteria of cultural landscape ecological integrity are described.

Table 12: Damage assessment results of traditional dwellings around Long-men Grottoes.

Extent of damage (%)	Number	Proportion (%)
≥85%	18	41
75%–85%	11	25
60%–75%	8	18
45%–60%	5	11
≤45%	2	5

5. Application and Analysis of Evaluation Model

In order to verify the scientificity and effectiveness of the cultural landscape and cultural ecological integrity evaluation system proposed in this paper, this paper takes the Long-men Grottoes landscape in Luo-yang City, Henan Province as an example for investigation and analysis. Long-men Grottoes are located in the southern suburbs of Luo-yang, Henan Province, China. Here, the two mountains stand against each other, the Yi-river flows in the middle, and the scenery is beautiful. As shown in Figure 5, it is the landscape of some Long-men Grottoes.

In order to evaluate the application effect of the cultural landscape ecological integrity evaluation model proposed in this paper, taking the material cultural landscape, intangible cultural landscape and local traditional dwellings of Long-men Grottoes as samples, 10 places were randomly selected as the control group and the observation group. Among them, the control group used the traditional empirical evaluation method, and the observation group used the cultural landscape ecological integrity evaluation model proposed in this paper.

For the physical cultural heritage and intangible cultural heritage, the evaluation is mainly conducted from the quantitative and qualitative perspectives. Quantitative indicators account for 60%, mainly including historical time (10%), scarcity (10%), landscape scale (30%), and landscape richness (10%). Qualitative indicators account for 40%, mainly including artistic value (25%) and scientific value (15%). As shown in Table 10, the evaluation results of the ecological integrity of the physical cultural heritage are shown, and Table 11 is the evaluation results of the ecological integrity of the intangible cultural heritage.

In addition, this paper uses the cultural landscape ecological integrity evaluation model to evaluate the damage of traditional dwellings around Long-men Grottoes, as shown in Table 12.

In order to analyze the evaluation effect of the cultural landscape and cultural ecological integrity evaluation system proposed in this paper on the Long-men Grottoes site, the two groups of experimental objects were compared from three aspects: the material cultural landscape of Long-men Grottoes: the intangible cultural landscape and the local traditional dwellings. As shown in Figure 6, from the score results of the two groups, the total score of the observation group is higher than that of the control group, and the scores of various indicators of the observation group are also higher than that of the control group. Through comparative experiments, we know that the cultural landscape ecological integrity evaluation model proposed in this paper can objectively reflect the evolution process of cultural heritage.

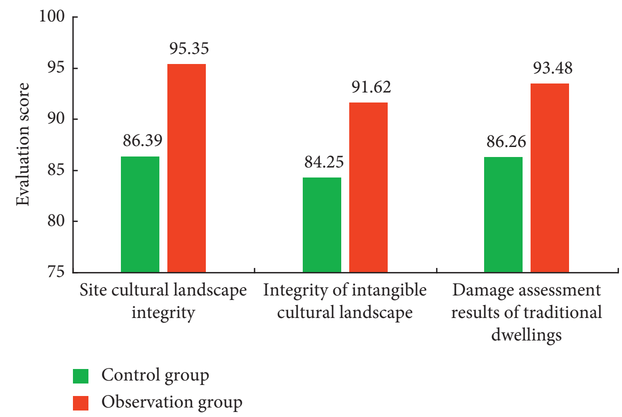


Figure 6: Comparison of actual operation effect between the two groups.

6. Conclusion

In view of the fact that the existing cultural landscape integrity evaluation system was not perfect enough to objectively reflect the evolution law of cultural landscape, this paper proposed a cultural ecological integrity evaluation model of cultural landscape based on the organic evolution characteristics of cultural landscape in cultural landscape heritage. Through the elaboration of ecological integrity and ecological integrity evaluation methods, the evaluation system and methods of cultural landscape ecological integrity were given theoretically. The cultural ecological integrity evaluation system not only fully considered the characteristics of the organic evolutionary cultural landscape but also objectively reflected the particularity and complexity of the cultural landscape. Through the application and analysis of examples, the results showed that the cultural landscape ecological integrity evaluation system proposed in this paper had important research significance for the protection, sustainable utilization, and development of organic evolutionary cultural landscape. In addition, this study can not only provide some theoretical reference and practical guidance for the effective classification of different organic evolutionary cultural landscapes but also provide a basis for the retrospective study of organic evolutionary cultural landscapes.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] M. Agnoletti, “Rural landscape, nature conservation and culture: some notes on research trends and management approaches from a (southern) European perspective,” *Landscape and Urban Planning*, vol. 126, pp. 66–73, 2014.

[2] M. Lazdinis, P. Angelstam, and H. Pülzl, “Towards sustainable forest management in the European Union through polycentric forest governance and an integrated landscape approach,” *Landscape Ecology*, vol. 34, no. 7, pp. 1737–1749, 2019.

[3] M. Antrop, J. Brandt, I. Loupa-Ramos et al., “How landscape ecology can promote the development of sustainable landscapes in Europe: the role of the European Association for Landscape Ecology (IALE-Europe) in the twenty-first century,” *Landscape Ecology*, vol. 28, no. 9, pp. 1641–1647, 2013.

[4] M. Cocks, S. Vetter, and K. F. Wiersum, “From universal to local: perspectives on cultural landscape heritage in South Africa,” *International Journal of Heritage Studies*, vol. 24, no. 1, pp. 35–52, 2017.

[5] C. H. Wang and C. C. Fu, “The conservation of disappearing sugar industry cultural landscapes in Taiwan,” *Journal of Asian Architecture and Building Engineering*, vol. 10, pp. 1–6, 2011.

[6] F. M. Y. Roxas, J. P. R. Rivera, and E. L. M. Gutierrez, “Framework for creating sustainable tourism using systems thinking,” *Current Issues in Tourism*, vol. 23, no. 3, pp. 280–296, 2020.

[7] K. Taylor, “Cultural landscapes and asia: reconciling international and southeast asian regional values,” *Landscape Research*, vol. 34, no. 1, pp. 7–31, 2009.

[8] T. Winter, “Beyond eurocentrism? Heritage conservation and the politics of difference,” *International Journal of Heritage Studies*, vol. 20, no. 2, pp. 123–137, 2014.

[9] B. Cooke and K. Moon, “Aligning “public good” environmental stewardship with the landscape-scale: adapting MBIs for private land conservation policy,” *Ecological Economics*, vol. 114, pp. 152–158, 2015.

[10] L. W. C. Lai, “Sustainable development of heritage conservation and tourism: a Hong Kong case study on colonial heritage,” *Sustainable Development*, vol. 28, no. 5, pp. 1181–1188, 2020.

[11] E. Fuentes-Montemayor, D. Goulson, and K. J. Park, “The effectiveness of agri-environment schemes for the conservation of farmland moths: assessing the importance of a landscape-scale management approach,” *Journal of Applied Ecology*, vol. 48, no. 3, pp. 532–542, 2011.

[12] L. E. Hannon and T. D. Sisk, “Hedgerows in an agri-natural landscape: potential habitat value for native bees,” *Biological Conservation*, vol. 142, no. 10, pp. 2140–2154, 2009.

[13] C. M. Raymond, C. Bieling, N. Fagerholm, B. Martin-Lopez, and T. Plieninger, “The farmer as a landscape steward: comparing local understandings of landscape stewardship, landscape values, and land management actions,” *Ambio*, vol. 45, no. 2, pp. 173–184, 2016.

[14] K. F. Tieskens, C. J. E. Schulp, C. Levers et al., “Characterizing European cultural landscapes: accounting for structure, management intensity and value of agricultural and forest landscapes,” *Land Use Policy*, vol. 62, pp. 29–39, 2017.

[15] H. Zhong, L. Wang, and H. Zhang, “The application of virtual reality technology in the digital preservation of cultural heritage,” *Computer Science and Information Systems*, vol. 18, no. 2, pp. 535–551, 2021.

[16] A. Campiani, A. Lingle, and N. Lercari, “Spatial analysis and heritage conservation: leveraging 3-D data and GIS for monitoring earthen architecture,” *Journal of Cultural Heritage*, vol. 39, pp. 166–176, 2019.

[17] J. Burger, M. Gochfeld, D. S. Kosson et al., “The costs of delaying remediation on human, ecological, and eco-cultural resources: considerations for the Department of Energy: a methodological framework,” *Science of the Total Environment*, vol. 649, pp. 1054–1064, 2019.

[18] F. Han, “World heritage cultural landscapes and new international trends,” *Chinese Landscape Architecture*, vol. 11, pp. 18–21, 2007.

[19] N. Claidière and J. B. André, “The transmission of genes and culture: a questionable analogy,” *Evolutionary Biology*, vol. 39, pp. 12–24, 2012.

[20] B. S. Frey and A. Briviba, “A policy proposal to deal with excessive cultural tourism,” *European Planning Studies*, vol. 29, no. 4, pp. 601–618, 2021.

[21] R. Camprubi and L. Coromina, “Content analysis in tourism research,” *Tourism Management Perspectives*, vol. 18, pp. 134–140, 2016.

[22] M. Agnoletti, “Rural landscape, nature conservation and culture: some notes on research trends and management approaches from a (southern) European perspective,” *Landscape and Urban Planning*, vol. 126, pp. 66–73, 2014.

[23] B. G. Jonsson, J. Svensson, G. Mikusiński, M. Manton, and P. Angelstam, “European union’s last intact forest landscapes are at A value chain crossroad between multiple use and intensified wood production,” *Forests*, vol. 10, no. 7, pp. 564–585, 2019.

[24] S. Mansourian, J. Parrotta, P. Balaji et al., “Putting the pieces together: integration for forest landscape restoration implementation,” *Land Degradation & Development*, vol. 31, no. 4, pp. 419–429, 2020.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 9322949, 9 pages
https://doi.org/10.1155/2022/9322949

Research Article

Validation of a Research Instrument to Measure Generation Y Parents' Perception of Service Quality and Effect on Satisfaction and Word of Mouth in International Schools

Howard Stribbell and Somsit Duangekanong

Graduate School of Business and Advanced Technology Management, Assumption University of Thailand, Bangkok, Thailand

Correspondence should be addressed to Howard Stribbell; howard.stribbell@tis.edu.mo

Received 14 June 2022; Revised 26 July 2022; Accepted 4 August 2022; Published 30 August 2022

Academic Editor: Jiafu Su

This quantitative research aimed to (1) validate an adapted instrument used for measuring service quality in private hospitals to measure the perceived service quality of K-12 international schools and (2) implement the instrument to assess service quality's impact on word of mouth and satisfaction. It was necessary to contextualize the scale items through item objective congruence test using industry experts. Scale items were adjusted to reflect the service provided by teachers, staff, and leadership of the school. The resulting 27 scale items for service quality were shown to be contextually valid and internally reliable. The instrument was then implemented to measure parental service quality's effect satisfaction and word of mouth. The survey was piloted by 33 parents and verified for internal consistency before being administered to 422 Generation Y parents. The results showed that the modified instrument was reliable and valid. The results showed that service quality had a direct and positive effect on both satisfaction and word of mouth but it had a greater effect on satisfaction.

1. Introduction

Schools are good at assessing the academic achievement of their students. Evaluating academic achievement is essential for schools. However, as a service industry, schools must also ensure that the quality of their service meets the expectations of parents and students. Schools will routinely monitor the satisfaction level of staff, parents, and students. Higher education institutes often measure the quality of the service they deliver by asking their students [1–5]. However, K-12 schools face the challenge that their students are not their direct customers; parents are the ones paying the tuition. As such, there is often a gap in measuring the service quality as perceived by their parents. Yet, satisfaction and perceived service quality are essential for service industries to be successful in an ever-increasing competitive market [6, 7]. As a result, schools must seek to understand the needs of their parents and students allowing them to adapt and improve so that they can not only meet but also even exceed their expectations. Therefore, there exists a need for K-12 schools to develop an instrument for measuring the perceived service quality of their parents.

This study sought to develop and validate an existing service quality instrument to make it contextually relevant and statistically reliable for use with parents of K-12 international schools.

Objectives for this research were, first, to develop and validate the service quality instrument according to the specific context of Generation Y parents in The International School of Macao, and second, to use a reliable modified instrument to measure Generation Y parents' perception of service quality and the effect on parent satisfaction and word of mouth.

1.1. Research Question. Can parent perception of service quality in K-12 international schools be reliably measured and what effect does service quality have on parent satisfaction and on parent word of mouth?

1.2. Research Objectives. The objective of the current study was to develop a reliable instrument for measuring service quality by adapting and contextualizing one used in private healthcare. The second objective was to implement and validate the instrument by determining the effect that service quality has on parent satisfaction and parent word of mouth.

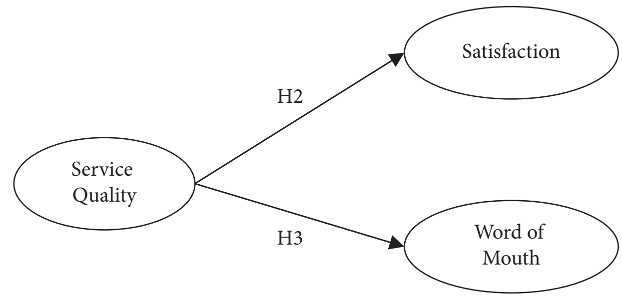


Figure 1: Conceptual framework for the current study.

2. Literature Review

2.1. Service Quality. Service quality is most often based on the expectancy disconfirmation theory [8–10]. A review of the literature suggests that two service quality models are typically accepted. Parasuraman [10] and colleagues proposed the SERVQUAL model for measuring service quality that consisted of 5 dimensions: tangibles, reliability, responsiveness, assurance, and empathy. This model has been used widely by many researchers [11–15] and specifically to measure service quality in higher education [2, 16–18]. Gronroos [9] developed a service quality model specifying the technical and functional aspects of service quality [9, 19].

Research into service quality has been used in many industries but its use in education has been limited to higher education [16, 20] and the researcher was unable to find any studies where it was used in K-12 contexts. Unfortunately, the use in HE is not suitable for K-12 contexts. In a university setting, the student receiving the service also decides whether it fits their expectations. In K-12 education, parents determine if the service meets their expectations though they may have only received some of the services directly and some of the services indirectly. As such, there is a need for a different measurement instrument that addresses this gap.

2.2. Expectation Disconfirmation Theory. Service quality and satisfaction is a measure between what is expected and what is experienced and is commonly referred to as expectation disconfirmation theory [21] and previously known as expectation confirmation theory [22, 23]. Expectation disconfirmation theory can be applied to a product or a service. A person with low expectations may be satisfied with the same level of experience that a person with high expectations would not. As such, expectation disconfirmation theory informs perceived service quality. Customers are satisfied with a service experience if it meets or exceeds their expectations [8]. In the same way, parents will be satisfied with their service experience from the school if it meets or exceeds their expectations.

2.3. Satisfaction. The satisfaction of both students and parents is often closely monitored by schools using both internal and external procedures. School-related nonacademic and academic aspects are frequently included in the multifaceted concept of satisfaction. Parents who are not happy will complain and spread negative messages, and if things do not change, they will withdraw their children from the school and look for another one.

In the research, satisfaction is informed by the expectancy disconfirmation theory [22–24] as the "consequence of the difference between the expected and perceived performance" [25]. Consumer satisfaction is a unique type of customer attitude used in the service sector that takes into account how much a customer likes or dislikes a service after using it [26].

Research is also interested in how satisfaction affects customer loyalty. Customer loyalty can be exhibited in repurchasing, continued use, or positive word of mouth. Jain et al. were not able to establish a link between satisfaction and WOM or brand loyalty. While satisfaction was not sufficient by itself in generating WOM, building customer-brand relationships on social media did have a significant and positive impact on brand trust, brand loyalty, and WOM for the brand [27].

2.4. Word of Mouth. Word of mouth (WOM) occurs when a consumer expresses their individual experiences with the company to other consumers. Previous research has shown that WOM has a direct effect on a consumer's expectation and on a consumer's perceived benefit which leads to a decision to purchase. WOM is generated by the consumer after purchasing and thereby influences other potential customers. In this way, the consumer is the producer of the WOM [27]. WOM is often a behaviour associated with customer loyalty. Like customer loyalty, parent loyalty can be defined as parents who give positive word of mouth, recommends the school to others, and encourages others to use the school service [28].

A special kind of WOM includes the liking, commenting, and sharing of posts on social media [29]. The effect of WOM is directly related to the strength of the tie between the author and receiver [30]. Prospective parents put a lot of faith in the WOM of their friends who are already enrolled in the institution.

2.5. Generation Y. Parents of international school students span multiple generations. Generation Y parents, also called Millennials, were born between 1981 and 2004 inclusively. Generation Y parents place greater emphasis on caring for

Table 1: Results of the item objective congruence test.

Item	1st expert	2nd expert	3rd expert	Total scores	IOC scores	Result
SQ1	1	1	1	3	1	Accepted
SQ2	1	1	1	3	1	Accepted
SQ3	1	−1	1	1	0.33	Revised and resubmitted
SQ3 revised	1	1	1	3	1	Accepted
SQ44	1	−1	1	1	0.33	Revised and resubmitted
SQ4 revised	1	1	1	3	1	Accepted
SQ5	1	1	1	3	1	Accepted
SQ6	1	1	0	2	0.67	Accepted
SQ7	1	1	1	3	1	Accepted
SQ8	1	1	0	2	0.67	Accepted
SQ9	1	1	1	3	1	Accepted
SQ10	1	1	0	2	0.67	Accepted
SQ11	1	0	1	2	0.67	Accepted
SQ12	1	1	1	3	1	Accepted
SQ13	1	0	1	2	0.67	Accepted
SQ14	1	0	1	2	0.67	Accepted
SQ15	1	1	1	3	1	Accepted
SQ16	1	1	1	3	1	Accepted
SQ17	1	0	1	2	0.67	Accepted
SQ18	1	1	1	3	1	Accepted
SQ19	1	1	1	3	1	Accepted
SQ20	1	1	1	3	1	Accepted
SQ21	1	1	1	3	1	Accepted
SQ22	1	0	1	2	0.67	Accepted
SQ23	1	0	1	2	0.67	Accepted
SQ24	1	1	1	3	1	Accepted
SQ25	1	1	0	2	0.67	Accepted
SQ26	1	1	1	3	1	Accepted
SQ27	1	1	1	3	1	Accepted

and protecting their children while also encouraging their children to be involved in community service. Academic achievement of the children of Generation Y parents is increasing and parents are likely to have higher expectations of their children's teachers and schools. [31].

3. Methods

3.1. Research Framework and Hypotheses. In addition to developing a reliable instrument to measure service quality, the current study developed the following research model to investigate the effect of service quality on satisfaction and word of mouth where service quality is considered an independent variable and satisfaction and word of mouth are considered dependent variables. The following hypotheses are proposed:

> *H1*. A reliable and contextually valid instrument can be developed to measure service quality of parents in international schools.
>
> *H2*. Service quality has significant and positive impact on satisfaction.
>
> *H3*. Service quality has significant and positive impact on parent word of mouth.

The following conceptual framework provided a model to examine the relationships between service quality and satisfaction and service quality and word of mouth as seen in Figure 1.

3.2. The Contextualized Scale Items. Previously developed for private healthcare by Lam [32] and operationalized by Cham [33], the 23 scale items were selected based on the reliability evidenced through its Cronbach Alpha value of 0.839 [33].

The researcher converted the scale items for the K-12 private international school context. The conversion was verified through an IOC test with 3 experts in the K-12 education sector. The experts were asked to determine the suitability of the construct to measure the given variable. Items that did not achieve a majority approval (>0.6) were revised based on the expert feedback and resubmitted during the second round. Results of the IOC are shown in Table 1.

The 23 items of Cham (2016) were converted to 27 items as shown in Table 2 [33]. Industry experts specifically separated the identified subject in the original scale items from "staff" to "teachers" and "office and support staff" in the revised scale items to more closely reflect the different roles in a school.

3.3. Sample Size and Method. Given the framework of 3 latent variables with 37 observed variables, an anticipated effect size 0.2, and a probability level of 0.05, it was calculated

Table 2: Conversion of Cham (2016) original scale items for K-12 context.

Item	Original scale item	Revised scale tem
SQ1	This hospital has up-to-date equipment	This school has up-to-date equipment.
SQ2	The physical facilities of this hospital are visually appealing	The physical facilities of this school are visually appealing.
SQ3	The staffs of this hospital appearance are neat	The teachers at this school present and conduct themselves in a professional manner.
SQ4		The office and support staff at this school present and conduct themselves in a professional manner.
SQ5	The materials associated with this hospital are visually appealing	The educational materials associated with this school are visually appealing.
SQ6	The staffs of this hospital perform the medical service right on the first time	The school performs the educational service well.
SQ7	The staffs of this hospital provide dependable services as promised	The school provides dependable services as promised.
SQ8	The staffs of this hospital are sincere to solve my problems	The school is sincere in solving my problems.
SQ9	The staffs of this hospital provide services at the appointed time	The school provides meetings and events as scheduled.
SQ10	This hospital keeps accurate medical records	This school keeps accurate educational records.
SQ11	The staffs of this hospital are never too busy to respond to my requests	The school responds to my requests promptly.
SQ12	The staffs of this hospital tell me when the services will be performed	The school tells me when the school events and activities will occur.
SQ13	The staffs of this hospital are always willing to help me	The teachers of this school are always willing to help me or my child.
SQ14		The staff of this school are always willing to help me or my child.
SQ15	I received prompt service from the staffs of this hospital	I received prompt service from the school.
SQ16	The staffs of this hospital are trustworthy	This school is trustworthy.
SQ17	I feel safe in receiving services from the staffs of this hospital	I feel positive about receiving education and support services from the school.
SQ18	The staffs of this hospital are consistently courteous to me	The teachers are consistently courteous to me.
SQ19		The office and support staff are consistently courteous to me.
SQ20	The staffs of this hospital have the knowledge to answer my questions	The teachers of this school have the knowledge to answer my questions.
SQ21		The staff of this school have the knowledge to answer my questions.
SQ22	The staffs of this hospital give individual attention to me	The office and support staff of this school give individual attention to me or my child.
SQ23		The teachers of this school give individual attention to me or my child.
SQ24	This hospital has convenient operating hours for my needs	This school schedules meetings that are convenient for my needs.
SQ25	This hospital has my best interests at heart	This school has the best interests of my child at heart.
SQ26	The staffs of this hospital understand my specific needs	The teachers of this school understand my child's specific needs.
SQ27		The staff of this school understand my child's specific needs.

that a minimum sample size of 296 would be required to determine effect. As such, a goal of 400 responses was set and exceeded. In this study, nonprobability sampling was used. The school granted access to 1937 parent emails and contact information. The researcher was also able to collect contact details of alumni's parents who would fit the Y population. The researcher was able to use judgement sampling by emailing all parents whose children were currently enrolled in the school. Upon completing the questionnaire, respondents were asked to forward the questionnaire to another potential participant including alumni parents whose contacts were not available at the school. The use of snowball sampling further increased response success.

A pilot test using the modified scale items and a 5-point Likert scale was administered by emailing 99 current parents and resulted in 33 valid responses. The collected data was used to test the reliability of the modified scale items. The service quality construct achieved a Cronbach Alpha value of 0.956 which indicates that the items are internally consistent and reliably indicate the service quality construct.

The resulting prototype instrument was then distributed to the parents at The International School of Macao by email and they were asked to complete the survey in English or Chinese. There were 475 total responses by Generation Y parents. Of the 475 valid responses, 74.3% (353) were female, 24.0% (114) were male, and 1.7% (8) preferred not to state their gender.

4. Results Analysis

4.1. Demographics. Of the 475 responses, 422 were complete and valid and 74.9% (316) were female, 23.2% (98) were male, and 1.9% (8) preferred not to state their gender. Most respondents were married or living with a partner (90.8% (383)), 5.7% (24) were single or divorced, and 3.5% (15) preferred not to state their marital status.

Table 3: Mean, standard deviation, and assessment of normality.

Construct	Items	Mean	Std dev	Skewness	Std error	Kurtosis	Std error
Service quality	SQ1	4.08	0.678	−0.415	0.119	0.281	0.237
	SQ2	4.24	0.713	−0.789	0.119	0.918	0.237
	SQ3	4.23	0.608	−0.413	0.119	1.134	0.237
	SQ4	4.24	0.585	−0.169	0.119	−0.078	0.237
	SQ5	4.13	0.666	−0.399	0.119	0.172	0.237
	SQ6	4.13	0.597	−0.186	0.119	0.293	0.237
	SQ7	4.14	0.601	−0.460	0.119	1.812	0.237
	SQ8	4.12	0.622	−0.266	0.119	0.257	0.237
	SQ9	4.26	0.554	0.020	0.119	−0.420	0.237
	SQ10	4.08	0.587	−0.157	0.119	0.432	0.237
	SQ11	4.18	0.605	−0.364	0.119	1.173	0.237
	SQ12	4.36	0.554	−0.188	0.119	−0.272	0.237
	SQ13	4.39	0.581	−0.316	0.119	−0.725	0.237
	SQ14	4.36	0.575	−0.228	0.119	−0.715	0.237
	SQ15	4.24	0.598	−0.139	0.119	−0.491	0.237
	SQ16	4.28	0.559	−0.023	0.119	−0.513	0.237
	SQ17	4.20	0.590	−0.222	0.119	0.290	0.237
	SQ18	4.39	0.556	−0.167	0.119	−0.856	0.237
	SQ19	4.37	0.586	−0.373	0.119	−0.271	0.237
	SQ20	4.29	0.578	−0.201	0.119	−0.148	0.237
	SQ21	4.27	0.543	0.077	0.119	−0.425	0.237
	SQ22	3.85	0.741	−0.146	0.119	−0.400	0.237
	SQ23	4.01	0.737	−0.297	0.119	−0.363	0.237
	SQ24	3.87	0.781	−0.576	0.119	0.844	0.237
	SQ25	4.12	0.606	−0.063	0.119	−0.346	0.237
	SQ26	4.03	0.690	−0.295	0.119	−0.091	0.237
	SQ27	3.87	0.721	−0.299	0.119	0.397	0.237
Satisfaction	SAT1	4.33	0.609	−0.328	0.119	−0.654	0.237
	SAT2	4.27	0.685	−0.409	0.119	−0.850	0.237
	SAT3	4.24	0.640	−0.311	0.119	−0.419	0.237
	SAT4	4.28	0.682	−0.969	0.119	2.485	0.237
Word of mouth	WOM1	4.42	0.595	−0.554	0.119	−0.228	0.237
	WOM2	4.38	0.635	−0.584	0.119	−0.328	0.237
	WOM3	4.27	0.733	−0.771	0.119	0.428	0.237

The following descriptive statistics and assessment of normality are presented in Table 3. While the values for skewness and kurtosis between -2 and +2 are considered acceptable to prove normal univariate distribution [34], only SAT4 exceeds the -2, +2 limit. However, Kline suggests that when using a large sample population procedure, such as SEM, one could reject the null hypothesis (of consistency with the normal distribution) and adopt a more descriptive approach to assessing normality. As such, the results are considered acceptable [35].

4.2. Confirmatory Factor Analysis. To evaluate the convergent and discriminant validity of the constructs and to determine the model fit, confirmatory factor analysis (CFA) was utilized. To test convergent validity, the following four criteria should be met as suggested by Hair et al. [36]; namely, the construct reliability (Cronbach Alpha value) should be greater than 0.7; the explained variance (AVE) within each construct should be larger than 0.5; the standardized factor loading of each observed variable to the latent construct should be at least 0.60; and the composite reliability (CR) should be at least 0.70. As seen in Table 4, all four criteria were met. All the observed variables had a factor loading greater than the recommended minimum of 0.6. The Cronbach Alpha values are well above the minimum of 0.7. The AVE values for all constructs exceed the minimum of 0.5 and the composite reliability (CR) values meet the minimum of 0.7. These results demonstrate that all the constructs in this study achieved the acceptable level of convergent validity. Two observed variables from service quality, namely, SQ1 and SQ2, were removed to get model fit.

Discriminant validity of this study was assessed using the Heterotrait-Monotrait Ratio of Correlations (HTMT) method proposed by Hensler et al. [37]. The statistical variances between the constructs are measured using discriminant validity. By comparing the correlations of indicators across constructs to the correlations of indicators within a concept, discriminant validity may be evaluated. The model has discriminant validity if the correlation ratio is below 0.9 [37]. The HTMT plugin by Gaskin was used in AMOS to determine the HTMT values [38]. As seen in Table 5, the correlation ratios between all variables are below 0.9; thus discriminant validity for this study was achieved.

Table 4: Convergent validity assessment.

Variables	Factor loading	*t*-value	CR	AVE	Cronbach Alpha
Service quality (SQ)			0.964	0.517	0.963
SQ3	0.715	—			
SQ4	0.638	15.534			
SQ5	0.608	12.240			
SQ6	0.753	15.224			
SQ7	0.734	14.838			
SQ8	0.771	15.588			
SQ9	0.724	14.640			
SQ10	0.709	14.321			
SQ11	0.753	15.209			
SQ12	0.652	13.148			
SQ13	0.749	16.717			
SQ14	0.736	14.886			
SQ15	0.810	15.433			
SQ16	0.835	16.911			
SQ17	0.771	15.580			
SQ18	0.742	14.991			
SQ19	0.656	12.320			
SQ20	0.765	15.469			
SQ21	0.783	15.840			
SQ22	0.602	12.139			
SQ23	0.641	12.945			
SQ24	0.618	12.459			
SQ25	0.781	15.810			
SQ26	0.733	14.817			
SQ27	0.629	12.683			
Satisfaction (SAT)			0.900	0.694	0.893
SAT1	0.853	—			
SAT2	0.845	21.101			
SAT3	0.870	22.241			
SAT4	0.759	18.645			
Word of mouth (WOM)			0.886	0.721	0.876
WOM1	0.836	—			
WOM2	0.908	22.290			
WOM3	0.800	19.055			

Table 5: HTMT analysis of the current study.

	WOM	SAT	SQ
WOM			
SAT	0.792		
SQ	0.586	0.742	

Table 6: Model fit criteria and values for the current study.

Index	Acceptable values	Statistical values
CMIN/DF	<5.00 [38]	2.584
GFI	≥0.85 [39]	0.854
AGFI	≥0.80 [39]	0.817
NFI	≥0.80 [40]	0.903
CFI	≥0.80 [41]	0.938
TLI	≥0.80 [42]	0.927
RMSEA	<0.08 [43]	0.061
Model summary		Acceptable model fit

Since convergent and discriminant validity were proved, construct validity was established.

Based on the suggestion by Hair et al. [36] the following criteria were used to determine model fit including chi-square fit statistics over degrees of freedom (CMIN/DF), goodness of fit index (GFI), adjusted goodness of fit index (AGFI), normed fit index (NFI), comparative fit index (CFI), Tucker-Lewis index (TLI), and root mean square error of approximation (RMSEA). A research model with a good fit meets the following criteria: CMIN/DF greater than 5; GFI greater than 0.85; AGFI, NFI, CFI, and TLI greater than 0.8; and REMSEA less than 0.08. As seen in Table 6, the current research model exceeded every criterion and achieved acceptable model fit.

4.3. Structural Equation Model. A structural equation model (Figure 2) was created to determine the standardized path coefficients of the constructs, given construct validity and an acceptable model fit. The initial model required SAT and WOM to be covaried in order to achieve model fit as

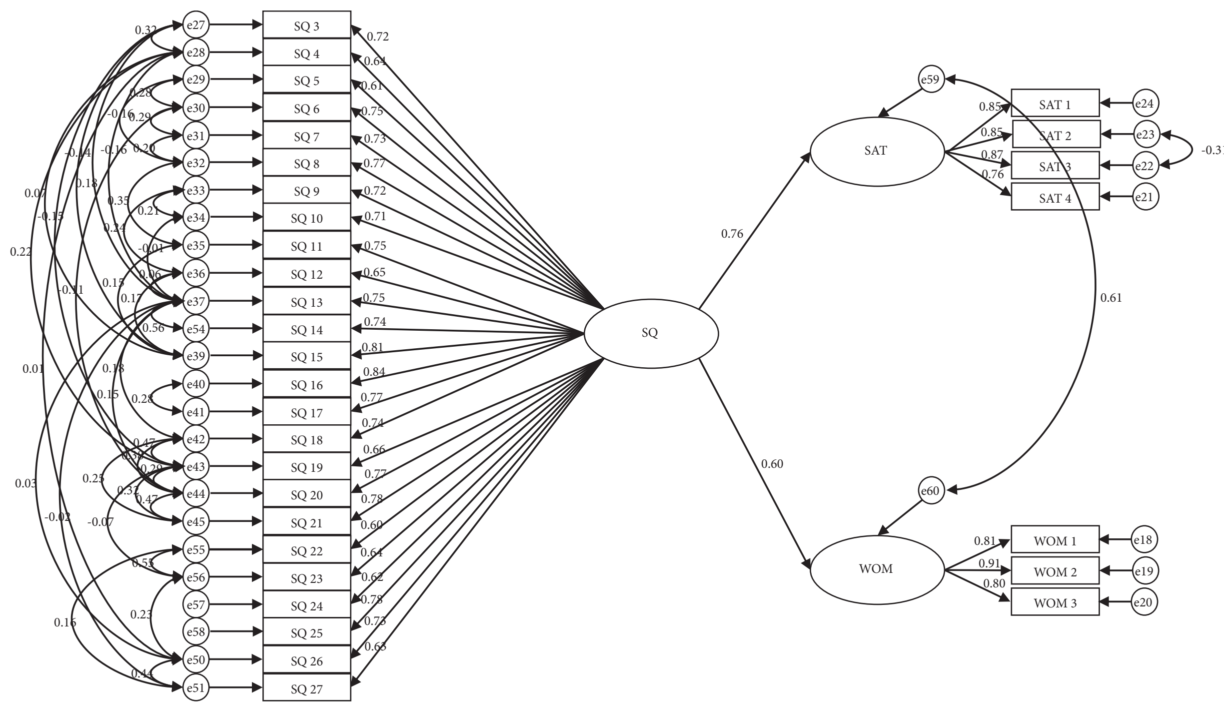


Figure 2: Structural Equation Model of the current study.

Table 7: Model fit of SEM.

Index	Acceptable values	Before adjustment	After adjustment
CMIN/DF	<5.00 [38]	2.864	2.584
GFI	≥0.85 [39]	0.840	0.854
AGFI	≥0.80 [39]	0.800	0.817
NFI	≥0.80 [40]	0.893	0.903
CFI	≥0.80 [41]	0.927	0.938
TLI	≥0.80 [42]	0.914	0.927
RMSEA	<0.08 [43]	0.067	0.061
Model summary		No model fit	Model fit

reported in Table 7. Thus, we can conclude that the model is valid [36].

As can be seen in Figure 2, the standardized direct effects reported by the AMOS software indicate that service quality had 0.76 factor on satisfaction and a 0.60 factor on word of mouth. The results are discussed in the following section.

5. Discussion

This study's primary goal was to develop a contextually reliable and valid measure instrument for assessing the perceived service quality of parents in international schools. The findings demonstrate that, through IOC and CFA, a reliable instrument can be developed. This study's secondary objective was to use the new instrument to assess the impact service quality had on satisfaction and word of mouth. The impact of these variables and the proposed hypotheses are discussed below and shown in Table 8.

Table 8: Hypotheses results.

Hypothesis	Goal	Result
H1	Service quality instrument	Supported
H1	SQ ⟶ SAT	Supported (.756, ***)
H2	SQ ⟶ WOM	Supported (.596, ***)

*** = $p < 0.001$; ** = $p < 0.01$; * = $p < 0.05$; ns = "not significant".

5.1. Influence of Service Quality on Satisfaction. The current study revealed that service quality had a significant and direct impact on parent satisfaction. As such, H2 was supported. Like satisfaction, service quality is a measure of the gap between expected results and actual results. When a parent receives lower than expected service, they perceive it as lower quality. Responding to a parent complaint with good service quality can lead to parent satisfaction. When a parent receives higher than expected service, they perceive it as higher quality. This expectancy confirmation or expectancy disconfirmation applies to satisfaction as well [8]. When a parent's needs are exceeded, they are more likely to be perceived as being satisfied.

5.2. Influence of Service Quality on Word of Mouth. The results of the current study revealed that there is a significant and direct relationship between service quality and parent word of mouth. As such, H3 was supported. When a parent receives greater than expected service at the school, they are more likely to recommend the school to others. It is understandable that when more parents feel that they get higher levels of service at the school, their satisfaction of the school and the likelihood of them recommending the school are increased.

6. Conclusion

This study reports the development of a contextualized service quality instrument for K-12 parents from one sample of Generation Y parents. The results demonstrate that modified scale items are reliable and may be used as an appropriate instrument for assessing parents' perception of service quality. The objectives were all answered satisfactorily. The adaptation of the service quality model used in healthcare required industry expert consultation for contextualization and measuring for internal consistency to ensure that the scale items would measure the intended variables. The use of the Item Objective Congruence provided a meaningful context of the scale items for a K-12 international school. Service quality scale items originally referring to staff were split to identify teachers and office and support staff. The development of reliable and contextualized scale items can be used in future research. As suggested by Chatfield and Collins [44] modified scale items can be used reliably as part of a measurement model and further used in the structural equation models (SEM) to identify causal relationships that use service quality as one of the constructs.

School administrators and leaders need a reliable method for measuring the quality of the service that staff were providing in their schools. School administrators and leaders can also benefit by understanding the role that the service that a parent experiences has a direct impact on their satisfaction and on the likelihood to recommend the school. By improving the experience associated with service, parents will have a greater perception of the school quality. The instrument can be used by schools to better evaluate the quality of the service that it provides to parents. Furthermore, the use of the new instrument can be used to survey all the parents in the school and does not need to be limited to a specific subpopulation such as Generation Y.

One limitation is that the developed instrument was only tried out in one international school and specifically limited to one specific population within the larger parent population. Future studies should expand on the population and be implemented within a greater range of international schools. Such development would yield further refining of the instrument.

Data Availability

The data used to support the findings of this study are available from the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

References

[1] R. S. Bellamkonda, "The effects of service quality on student loyalty: the mediating role of student satisfaction," *Journal of Modelling in Management*, vol. 11, no. 2, pp. 446–462, May 2016.

[2] B. R. Barnes, "Analysing service quality: the case of post-graduate Chinese students," *Total Quality Management and Business Excellence*, vol. 18, no. 3, pp. 313–331, May 2007.

[3] S. Pokharel, "Stakeholder's perception of service quality: a case in Qatar," *Quality Assurance in Education*, vol. 27, no. 4, pp. 493–510, Jan. 2019.

[4] G. Singh and M. Kumar, "Exploratory factor Analysis of service quality dimensions for higher educational institutes: a students perspective," *Glob. J. Manag. Bus. Res.*, vol. 14, no. 8, 2014.

[5] P. Yin Wong and H. Yin Wong, "An integrated-process model of service quality, institutional brand and behavioural intentions," *Managing Service Quality*, vol. 24, no. 5, pp. 487–521, Sep. 2014.

[6] J. J. Cronin Jr., M. K. Brady, and G. M. Hult, "Assessing the effects of quality, value, and customer satisfaction on consumer behavioral intentions in service environments," *Journal of Retailing*, vol. 76, no. 2, pp. 193–218, 2000.

[7] V. A. Zeithaml, L. L. Berry, and A. Parasuraman, "The behavioral consequences of service quality," *Journal of Marketing*, vol. 60, no. 2, pp. 31–46, 1996.

[8] K. E. Clow, D. L. Kurtz, J. Ozment, and B. Soo Ong, "The antecedents of consumer expectations of services: an empirical study across four industries," *Journal of Services Marketing*, vol. 11, no. 4, pp. 230–248, Jan. 1997.

[9] C. Grönroos, "A service quality model and its marketing implications," *European Journal of Marketing*, vol. 18, no. 4, pp. 36–44, 1984.

[10] A. Parasuraman, V. A. Zeithaml, and L. L. Berry, "A conceptual model of service quality and its implications for future research," *Journal of Marketing*, vol. 49, no. 4, p. 41, 1985.

[11] H. Arasli, S. Mehtap-Smadi, and S. Turan Katircioglu, "Customer service quality in the Greek Cypriot banking industry," *Managing Service Quality: International Journal*, vol. 15, no. 1, pp. 41–56, Jan. 2005.

[12] K. Choudhury, "Service quality and word of mouth: a study of the banking sector," *International Journal of Bank Marketing*, vol. 32, no. 7, pp. 612–627, Jan. 2014.

[13] N. Asiah Abdullah and N. Asiah Abdullah, "The effect of perceived service quality dimensions on customer satisfaction, trust, and loyalty in e-commerce settings," *Asia Pacific Journal of Marketing & Logistics*, vol. 22, no. 3, pp. 351–371, Jan. 2010.

[14] S. Marković, D. Lončarić, and D. Lončarić, "Service quality and customer satisfaction in the health care industry - towards health tourism market," *Tourism and Hospitality Management*, vol. 20, no. 2, pp. 155–170, 2014.

[15] T. Riani and A. L. Riani, "The effect of health service quality and brand image on patients loyalty, with patients satisfaction as mediating variable (A study in vip ward of prof. Dr R soeharso ortopedics hospital in surakarta)," *Sebelas Maret Business Review*, vol. 2, no. 1, Sep. 2017.

[16] Y. Nguyen Thi Hoang, "Measuring service quality in the context of higher education in vietnam," *Journal of Economic Development*, pp. 77–90, Dec. 2013.

[17] S. E. McDaniel and J. R. McDaniel, "Assessing service quality in schools of business," *International Journal of Quality & Reliability Management*, vol. 14, no. 3, pp. 204–218, Jan. 1997.

[18] C. A. Ruby, "Assessing satisfaction with selected student services using SERVQUAL, a market-driven model of service quality," *NASPA Journal*, vol. 35, no. 4, pp. 331–341, 1998.

[19] G. N. Soutar, "Word of mouth antecedents in an educational context: a Singaporean study," *International Journal of Educational Management*, vol. 26, no. 7, pp. 678–695, Jan. 2012.

[20] A. Quinn, G. Lemay, P. Larsen, and D. M. Johnson, "Service quality in higher education," *Total Quality Management and Business Excellence*, vol. 20, no. 2, pp. 139–152, Feb. 2009.

[21] P. M. Swamidass, "Expectation theory," in *Encyclopedia of Production and Manufacturing Management*p. 197, Springer, Boston, MA, USA, 2000.

[22] R. L. Oliver, "Effect of expectation and disconfirmation on postexposure product evaluations: an alternative interpretation," *Journal of Applied Psychology*, vol. 62, no. 4, pp. 480–486, 1977.

[23] R. L. Oliver, "A cognitive model of the antecedents and consequences of satisfaction decisions," *Journal of Marketing Research*, vol. 17, no. 4, pp. 460–469, Nov. 1980.

[24] S. McQuitty and A. Finn, "Systematically varying consumer satisfaction and its," *Implications for Product Choice*, vol. 19, 2000.

[25] S. Kataria and V. Saini, "The mediating impact of customer satisfaction in relation of brand equity and brand loyalty," *South Asian Journal of Business Studies*, vol. 9, no. 1, pp. 62–87, 2019.

[26] Wu, "The impact of hospital brand image on service quality, patient satisfaction and loyalty," *African Journal of Business Management*, vol. 5, no. 12, pp. 4873–4882, 2011.

[27] N. K. Jain, S. Kamboj, V. Kumar, and Z. Rahman, "Examining consumer-brand relationships on social media platforms," *Marketing Intelligence & Planning*, vol. 36, no. 1, pp. 63–78, 2018.

[28] C. Li and C. Hung, "Marketing tactics and parents' loyalty: the mediating role of school image," *Journal of Educational Administration*, vol. 47, no. 4, pp. 477–489, 2009.

[29] J. H. Liu, M. North, and C. Li, "Relationship building through reputation and tribalism on companies' Facebook pages," *Internet Research*, vol. 27, no. 5, pp. 1149–1169, 2017.

[30] A. M. Abubakar, M. Ilkan, and P. Sahin, "EWOM, eReferral and gender in the virtual community," *Marketing Intelligence & Planning*, vol. 34, no. 5, pp. 692–710, 2016.

[31] N. Howe, *The Lineup of Generations*, AASA | American Association of School Administrators, 2010, https://www.aasa.org/SchoolAdministratorArticle.aspx?id=11024accessed.

[32] S. S. K. Lam, "SERVQUAL: a tool for measuring patients' opinions of hospital service quality in Hong Kong," *Total Quality Management*, vol. 8, no. 4, pp. 145–152, Aug. 1997.

[33] T. H. Cham, Y. M. Lim, N. C. Aik, and A. G. M. Tay, "Antecedents of hospital brand image and the relationships with medical tourists' behavioral intention," *International Journal of Pharmaceutical and Healthcare Marketing*, vol. 10, no. 4, pp. 412–431, Jan. 2016.

[34] D. George and M. Mallery, *SPSS for Windows Step by Step: A Simple Guide and Reference*, Pearson, 10a ed edition, 2010.

[35] R. B. Kline, "Principles and practice of structural equation modeling," *Principles and Practice of Structural Equation Modeling*, Guilford Press, New York, USA, 4th ed edition, 2016.

[36] J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson, *Multivariate Data Analysis: A Global Perspective*, Prentice-Hall , 7th ed edition, 2010.

[37] J. Henseler, C. M. Ringle, and M. Sarstedt, "A new criterion for assessing discriminant validity in variance-based structural equation modeling," *Journal of the Academy of Marketing Science*, vol. 43, no. 1, pp. 115–135, 2015.

[38] Y. H. Al-Mamary and A. Shamsuddin, "Testing of the technology acceptance model in context of Yemen," *Mediterranean Journal of Social Sciences*, vol. 6, no. 4, 2015.

[39] C. Sica and M. Ghisi, "The Italian versions of the beck anxiety inventory and the beck depression inventory-II: psychometric properties and discriminant power," in *Leading—Edge Psychological Tests and Testing Research*, M. A. Lange, Ed., pp. 27–50, Nova Science Publishers, USA, 2007.

[40] J.-H. Wu and Y.-M. Wang, J. Gaskin, Measuring KMS success: a respecification of the DeLone and McLean's model," *Information & Management*, vol. 43, no. 6, pp. 728–739, 2016.

[41] P. M. Bentler, "Comparative fit indexes in structural models," *Psychological Bulletin*, vol. 107, no. 2, pp. 238–246, 1990.

[42] G. P. Sharma, R. C. Verma, and P. Pathare, "Mathematical modeling of infrared radiation thin layer drying of onion slices," *Journal of Food Engineering*, vol. 71, no. 3, pp. 282–286, 2005.

[43] R. Pedroso, L. Zanetello, L. Guimarães et al., "Confirmatory factor analysis (CFA) of the crack use relapse scale (CURS)," *Archives of Clinical Psychiatry (São Paulo)*, vol. 43, no. 3, pp. 37–40, 2016.

[44] A. Collins, "Introduction to multivariate analysis," *Introduction to multivariate analysis*, p. 248, 2018.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 6026875, 10 pages
https://doi.org/10.1155/2022/6026875

Research Article

Markov Model-Based Learning Aid for Students' Civics Course

Xiaohui Zang [1] and Chengming Bai [2]

[1] *Students' Affairs Office, Shaanxi University of Science & Technology, Xi'an 710021, Shaanxi, China*
[2] *School of Marxism, Shaanxi University Science & Technology, Xi'an 710021, Shaanxi, China*

Correspondence should be addressed to Xiaohui Zang; zangxiaohui2021@163.com

Received 27 May 2022; Revised 11 July 2022; Accepted 18 July 2022; Published 29 August 2022

Academic Editor: Jiafu Su

Since the 19th National Congress, students' ideological education has become more and more one of the national priorities, so the Civics course has become one of the essential compulsory courses for students at all stages of school and university, and the learning methods of Civics course have also become a hot issue of concern to students, which makes the learning process of supplementary learning methods very important. In this paper, a Markov model was developed to calculate the probability transfer matrix and predict the supplementary learning methods used by students. This paper also establishes a Markov model to predict the frequency of students' online classroom learning at different stages, and it is found that, in the future, more and more students will use the Internet for their Civics course assisted learning; therefore, it is very important to establish a perfect Civics course online assisted learning platform, and this paper also puts forward some suggestions for establishing a Civics course online assisted learning system, which provides some methods for subsequent students' Civics course learning. This paper also proposes some suggestions for establishing a web-assisted learning system for Civics courses and provides some methods for subsequent student learning in Civics courses.

1. Introduction

Since the 19th National Congress, students' ideological education work has become more and more important to the Party Central Committee, and higher requirements have been put forward for this purpose. Therefore, learning the Civics course well is very important for implementing and enforcing the Party's educational philosophy and policy, cultivating high-level and high-quality talents who can develop comprehensively for the country, developing socialism with Chinese characteristics, and guiding the ideology of China and other fields. Under such conditions, the study of Civics course becomes an essential task for students at all stages in the learning process. The study of Civics course is not the same as other subjects, so how to learn Civics well also becomes a concern for students, which makes the auxiliary methods in the learning process very important [1].

Markov model (Markov Model) is a statistical model that is widely used in the fields of recognition and identification, such as the labeling of different lexical properties, sound and text conversion, probabilistic grammar, etc. It can also be applied to make predictions about engineering, weather, etc. In the definition of Markovianity, "now" refers to a certain definite moment, but in practical problems, the concept of "now" in Markovianity is usually generalized, i.e., the stopping moment [2]. For example, if the Brownian motion is in the plane from the center of the circle, if we want to study the conditional independence of the event at the moment t_0 when the circle is first reached and the subsequent events, we set t_0 as the stopping moment and consider t_0 as the "now." If the "now" is extended to the "now" in the case of stopping time, under the condition that the "now" is known. This property is called strong Markovianity. The Markov process with this property is called strong Markov process. For quite some time, many people thought that Markov processes are necessarily strong Markov processes. It was J. L. Dubb who first suggested that a rigorous proof was needed for strong Markovianity. It was not until 1956 that an example of a Markov process that was not a strong Markov process was found.

The use of Markovian models to study learning assistance methods for students' Civics courses is a better research method. A learning aid method is a type of learning method [3]. Broadly speaking, learning methods refer to the motivation students have and the approach they take in learning, which determines the effectiveness and quality of their learning; a large aspect of learning methods is classroom instruction, or course experience. In Western academia, research on students' experiences and approaches to learning is part of the Tradition of the Students' Approaches to Learning (SAL), which has been developing for more than 40 years. In 1976, Swedish scholars used phenomenological descriptions to identify two types of approaches to learning, namely, deep and shallow learning approaches, thus laying the foundation for the birth of the SAL research tradition [4]. The deep learning approach points to a focus on the meaning of the text and is a learning approach that aims at deep understanding. The shallow learning approach points to the memorization of text in order to be able to answer questions rather than to comprehend it. A large number of subsequent studies have supported the idea that whether students adopt a deep learning approach is influenced by a combination of individual student characteristics (e.g., prior learning experiences, existing conceptions), teacher teaching methods, instructional evaluations, and students' perceptions and experiences of the learning environment, but of the many influencing factors, the actual lesson experience students have in the classroom is the one with the greatest intensity [5].

However, with the rapid development of modern technology, students' learning forms have gradually become diversified, and students are exposed to more and more supplementary learning tools in addition to traditional classroom learning. At the same time, China's higher education is developing from elite education to mass education, universities and schools are expanding their enrollment, and the examination and test modes of Civics and Political Science classes are becoming diversified [6]. This has the consequence that teachers leave soon after class and cannot communicate much with students, and students have to rely on rote memorization when they review what they have learned. And the consequence of this approach for students is a superficial understanding of the content and meaning of educational courses. It creates a "test-taking" mentality in students' learning [7]. Therefore, how to make efficient use of extracurricular time, strengthen teachers' perception of ideological and political course itself, deepen students' knowledge of ideological and political course and understanding of the significance of ideological and political course, and study students' auxiliary learning methods in ideological and political course learning have become a new focus.

Based on Markov model, this paper analyzes and studies the auxiliary methods for students' ideological and political course learning in view of the current situation and problems encountered in students' learning, which is conducive to higher efficiency and more fruitful learning of ideological and political course in the future. According to the analysis results, the frequency of assisted learning methods used by college and primary and secondary school students in the future is predicted, which provides some reference suggestions for the assisted learning of ideological and political courses in the future.

2. Literature Review

This paper mainly studies the auxiliary methods of students' ideological and political course learning through the establishment of Markov model, so it first introduces some common research problems based on Markov model [8].

Technology of asphalt pavement based on Markov model to predict the process is as follows: to Xinjiang Ili region G3016 line on the highway engineering, use grey prediction model to analyze the change of the pavement performance and then according to the Markov models get road state transition probability matrix and the state transition probability in the future, again carry on with the forecast, and predict the technical condition of road surface. The prediction method of grey system can achieve ideal prediction effect when the historical data is less, and compared with other prediction methods, this method is closer to the actual situation [9]. In the process of investigation and analysis of road conditions, some relevant data of road use and performance are sometimes difficult to investigate or lack investigation, so the analysis method based on Markov model is very effective. According to the above method, the road performance of this section can be predicted and relatively accurate data can be obtained. The measured value of road performance in 2013 is 96.86 and predicted value is 98.14; the measured value of road performance in 2015 is 88.61 and predicted value is 87.15; the measured value of road performance in 2017 is 82.24 and the predicted value is 83.01; the measured value is 79.61 in 2018 and the predicted value is 78.84. It can be found that the error between the predicted value and the actual value is about 0.99, which meets the expectation of the predicted value [10]. It can be concluded that the establishment of Markov model to predict the road performance is a good method [11].

The monitoring of air quality based on Markov model is also an important work. Air quality is an important indicator of environmental monitoring. The quality of air has far-reaching significance and influence on human survival and development, food, clothing, shelter and transportation, and harmonious coexistence of nature. If we can accurately predict the level of air quality in an area, we can do a good job of protection and relevant countermeasures in advance and minimize the harm and adverse impact of air quality deterioration on life [12]. 730 air quality grade data pieces of Taiyuan city from April 1, 2019, to March 31, 2021, are selected and modeled by Markov model. The transition probability matrix and its corresponding probability matrix are calculated and its marginal probability is obtained. After data collection and model calculation and analysis, the air quality level on April 15, 2021, is used as the initial state to predict the air conditions for the next five days. The predicted states of air quality in the five days after April 15 are 4, 2, 2, 2, and 2, respectively, which means that the probability of air quality grade being moderate pollution, good, good,

good, and good is the highest, while the actual air quality grade is good, good, good, moderate pollution, and good, and the prediction accuracy is 80% [13]. By forecasting the weather of the whole year in 2020 and comparing with the actual situation, it can be obtained that the forecast days of the air quality of the year 2020 in six grades (excellent, good, mild pollution, moderate pollution, severe pollution, and severe pollution) are, respectively, 56 days, 218 days, 65 days, 18 days, 5 days, and 3 days. The actual days are 54 days, 223 days, 65 days, 15 days, 8 days, and 0 days, respectively, and the relative error is 4.38%, indicating that the prediction effect of Markov model is relatively accurate [14].

Markov model can also be used to analyze the industrial structure of energy. With the passage of time and the development of science and technology, the industrial structure of energy is also constantly changing; the analysis of the rule of the change of energy industrial structure found that this change can be described by Markov chain. Therefore, the establishment of Markov model to analyze the development and change trend of energy industry structure in several decades can better understand the law of energy industry change and then carry out analysis and suggestions. In the U.S., for example, in the National Energy Data System (SEDS) agenda, datasets from four states - California, Arizona, New Mexico, and Texas - from 1960 to 2009 were collected and organized, the data were preprocessed, useless data were removed, and valid data were extracted for analysis [15]. Firstly, the energy consumption of the four states was analyzed and plotted, and the following conclusions were drawn: the proportion of fossil fuels used in the four states was basically 50%~60%, but the use of fossil fuels in Arizona showed a downward trend in the past 50 years, while the other three showed an upward trend. The share of clean renewable energy used in all four states has declined over the past 50 years, with Arizona experiencing a slower decline in the share of clean renewable energy and the remaining three states experiencing a faster decline in the share of clean renewable energy. The use of thermal power, motor gasoline, and other energy sources in these four states has largely maintained a steady upward trend over a 50-year period [16]. The reasons for the above situation were thus analyzed. Arizona has few fossil fuel resources but is rich in solar resources due to the semiarid climate of most of the state, with long hours of sunlight, and is in a highland region with abundant wind resources and abundant geothermal resources. In addition, the type of industry in the state is a small industrial sector with low per capita energy consumption. In contrast, the other three states have abundant fossil fuel resources, such as natural gas and crude oil. However, because of California's mild climate and less frequent use of heating equipment in the winter, per capita energy consumption is lower. While renewable clean energy has been declining in the other three states over the past 50 years, they all have abundant renewable energy resources that can still be developed and used vigorously today [17].

In addition, Markov models allow for design-theoretic analysis of intelligent systems, such as an intelligent Chinese character to Braille conversion system designed for the blind. The principle of conversion from Chinese characters to Braille is "Chinese characters-Pinyin-Braille." By building a Markov model, the steps of this system design can be simply divided into changing Chinese characters into pinyin and then pinyin into Braille. The Markov model can recognize random Chinese text statements and change Chinese characters into pinyin [18]. However, it is not enough to use this method alone. It is necessary to combine the inverse maximum matching word division method for utterance division of Chinese text and to use the phrase dictionary to match phrases or single characters with the Character Set of Chinese Character Codes for Information Exchange. Next, a phonetic code dictionary is used to solve the problem of pinyin to Braille conversion. As we all know, Chinese characters are unique, so Chinese is very prone to ambiguity when translated into other languages. The process of reading Chinese text from pinyin, recognizing words and phrases through pinyin, and converting pinyin or outputting Braille may result in translation errors due to misunderstanding [19]. Therefore, it is necessary to establish a database and a statistical state database to count the correct corresponding word meanings, register the errors in the reading process and conversion process, and correct the errors in the translation process. The application of Markov process in the Braille conversion system can be elaborated as follows: In the input process of random Chinese text, if the Chinese character to be converted is after a previously recognized Chinese character and is related only to that Chinese character, but not to the Chinese character before it, it is represented by a state space M_n. Given that a large number of utterances of Chinese characters have a coherent expressive meaning, the probability of occurrence of individual linguistic symbols in each sentence or utterance is not independent of each other, and the current Chinese character of each random Chinese character text must be related to the previous Chinese character, which can be processed by using Markov chains for utterances.

In addition to making predictions, Markov models can also analyze the performance of safety systems, for example, to measure the working process of safety instruments [20]. By using Markov models in chemical plants, the safety integrity of the whole system can be tested and the structure and processing of the plant can be optimized according to the test structure, thus improving the safety factor and fault tolerance of the whole plant and reducing the risk. For safety instrumentation, there are many factors that affect its safety performance, such as the time of device operation, cycle function testing, etc. Therefore, based on Markov model, it is possible to establish a test method for safety instrument function [21]. The more factors considered, the more channels for different state generation and transfer in Markov model naturally. The calculated results of the safety assessment show that the proportion of dangerous failures in the actuator is 78.6%, in the sensor 18.2%, and in the logic controller 3.2%. The abovementioned percentage of dangerous failures leads to the conclusion that the actuator of the device has the greatest influence in the protection of the investigated device, followed by the sensor. If we want to improve the safety of the whole device and reduce the risk, the first thing to consider is the optimization of the processing of the actuators of the device.

From the above examples, it can be seen that Markov model has a wide application space in practical species, especially in probability statistics and prediction, with more accurate experimental results. This establishes a solid theoretical foundation for the research in this paper [22].

3. The Establishment of Markov Model and the Analysis of Civics Course-Assisted Learning Method

3.1. Establishment of the Markov Model. Markov model is a kind of stochastic time series analysis method, which predicts the future state of things by studying the initial probability of different states and the transfer probability between states. The most important feature of Markov model is that it has no posteriority; i.e., it is considered that the conditional distribution of the state of the process or system at the moment $t > t_0$ is independent of the state of the process before the moment t_0, if the state at the moment t_0 is known [23]. That is, the future state does not depend on the past but is only related to the current state. This property is very suitable for analyzing data with high volatility and no obvious time-varying characteristics.

3.1.1. Markov Process. Let $X = (X_1, X_2, X_3, \ldots, X_t)$ be a sequence of random variables, where each random variable takes values in a finite set $S = \{s_1, s_2, s_3, \ldots, s_n\}$, called the state space. Markov is characterized by the following.

(1) Finite history assumption:

$$P(X_{t+1} = s_k | X_1, \ldots, X_t) = P(X_{t+1} = s_k | X_t). \tag{1}$$

(2) Time invariance assumption:

$$\forall i \in \{1, 2, 3, \ldots, T\} \forall x, y \in S, P(X_i = y | X_{i-1} = x) = p(y|x). \tag{2}$$

If X has these characteristics above, then this sequence of X is called a Markov process (chain).

If this sequence is a Markov chain, it has the $n \gg 0$ following $i, j \in I$, properties for all $p_{ij}^{(n)}$ integers, and n-step transfer probabilities.

$$\begin{gathered} p_{ij}^{(n)} = \sum_{k \varepsilon I} p_{ik}^{1} p_{kj}^{(n-1)}, \\ P\{X(n_1) = i_1, X(n_2) = i_2, \ldots X(n_m) = i_m\} \\ = \sum_i p_i^{(0)} p_{ii_1}(n_i) p_{i_1} p_{i_2}(n_2 - n_1) \ldots p_{i_{m-1}} p_{i_m}(n_m - n_{m-1}), \\ \mathbf{P}^{(n)} = \mathbf{P}^{(n-1)} \mathbf{P}. \end{gathered} \tag{3}$$

3.1.2. Markov Analysis Method. The basic model of the Markov analysis method is

$$\mathbf{X}(K+1) = \mathbf{X}(K) \times \mathbf{P}, \tag{4}$$

.

Let $p_{ij} = P(X_i = j | X_0 = i), i, j \in I$; then, p_{ij} is called transfer probability, which denotes the transfer probability from state i to state j. The transfer p_{ij} probability matrix thought of as an element is shown as follows:

$$\mathbf{P} = \begin{pmatrix} p_{11} & \cdots & p_{1m} \\ \vdots & \ddots & \vdots \\ p_{m1} & \cdots & p_{mm} \end{pmatrix} = (p_{ij})_{m \times n}. \tag{5}$$

3.1.3. Test of the Markov Model. If a Markov model is to be used to make predictions about something, it is a very important prerequisite that such a thing must have Markovianity. Therefore, it is necessary to test the thing first and use χ^2 to judge the test results. Assume that the index series is divided into m states, the frequency of transformation from state i to state j is recorded as f_{ij}, the transfer probability is p_{ij}, and the conditional probability of the state transfer frequency matrix is taken as the marginal probability [24]. The calculation formula is as follows:

$$P_j = \frac{\sum_{i=1}^{m} f_{ij}}{\sum_{j=1}^{m} \sum_{i=1}^{m} f_{ij}}. \tag{6}$$

The formula for calculating the χ^2 statistic is as follows:

$$\chi^2 = 2 \sum_{j=1}^{m} \sum_{i=1}^{m} f_{ij} \left| \log \frac{p_{ij}}{P_j} \right|. \tag{7}$$

It should be noted here that the log in equation is because $\ln x$ is often written as $\log x$ in programming.

The degrees of freedom are obtained by checking the table given the significant $\chi_\alpha^2((m-1)^2)$ level α. So if $\chi^2 > \chi_\alpha^2[(m-1)^2]$, then the test is passed.

3.1.4. Constructing the Multistep Transfer Probability Matrix. The calculation of the multistep transfer probability matrix requires the use of the C-K equation (Chapman–Kolmogorov equation), which is calculated as follows:

$$P(u+v) = P(u)P(v). \tag{8}$$

In the equation, if we make $u = 1$ and $v = n - 1$, then we can get the following recurrence $P(n) = P(1)P(n-1) = PP(n-1) = \ldots = P^n$ relation, so for the chi-square Markov chain, the n-step transfer probability matrix is equal to the nth power of the one-step transfer probability.

3.1.5. Calculating the Invariant Probability Measure. It can be seen that the core problem of Markov model is to determine the transfer probability matrix **P.** In this paper, we take the number of people (ten people) as the scale, take the predicted 50,000 people as the initial value of Markov probability prediction, calculate the value of the frequency that various auxiliary methods of learning will appear, respectively, and then take the natural year as the scale, establish Markov model for the change process of the frequency of students using a certain auxiliary method at different stages in several years, and predict the change trend of the frequency of using this auxiliary method in the next few years [25].

3.2. Analysis of Auxiliary Learning Methods in Civics Courses

3.2.1. The Necessity of Aided Learning Methods in Civics Courses. First of all, in modern teaching, teachers are not exactly the "people who teach knowledge" in the traditional sense anymore. The teacher is transformed from "the person who teaches knowledge" to "the person who guides students to learn." The main body of learning is not the teacher but the students, and the process of learning can be seen as the process of improving knowledge in the minds of the learners themselves [26]. Therefore, teachers should not only teach literature knowledge, but also guide students to learn and help them to learn and improve themselves and try to build their own learning system [27]. In traditional teaching, teachers and textbooks are the only sources for students to acquire knowledge, but with the progress of information technology, the development of the Internet, and the diversification of values, students have far more ability and opportunities to acquire knowledge than before, but on the contrary, in the vast information network, students who have not yet entered the world may also lose their way and lose their judgment. Under such conditions, teachers should, on the one hand, rely on their own rich experience and grasp of subject knowledge to connect what students have learned in various places into a complete knowledge system; on the other hand, teachers should guide students to correctly distinguish between the best and the worst and consciously resist the erosion of bad culture. In the traditional teaching, students are taught by teachers in the classroom according to the knowledge points in the textbook. Nowadays, with the development of multichannel information technology, students' knowledge acquisition not only is concentrated in the classroom, but also has many other supplementary learning methods [28].

Secondly, students in China not only are active in their thinking and personality, but also like to pursue new things and are enthusiastic about the Internet and offline activities. Most of the students learn computer-related knowledge by looking up information, reading news, or chatting with online friends through the Internet. Einstein once said that science is a powerful tool, and it depends on mankind itself, not on the tool, how to use it to bring happiness or disaster to mankind. Ideological and political theory is not an obscure theoretical derivation; on the contrary, it is very easy to understand; in the traditional teaching, teachers only need to prepare lessons carefully before class, study the textbook carefully, integrate the knowledge in the textbook, pass the knowledge to students in the classroom, and correct homework carefully after class to complete the basic work, but ideology and politics are closely related to the real society and everyone, so only the classroom learning is not enough [29]. In addition to the knowledge imparted by teachers in the classroom, it is more important to practice, which requires the use of, for example, television, the Internet, lectures, newspapers, extracurricular books, etc., to change the traditional static teaching materials into dynamic teaching materials composed of sound, images, real people, etc. to attract students' interest and, on this basis, with the help of multimedia networks to constantly update the supplementary methods. The teaching contents are constantly updated by multimedia network [30].

3.2.2. Analysis of Students' Learning Aids in Civics Courses. First of all, we investigate the auxiliary learning methods of students in the Civic Science course, and we can get the following more common auxiliary learning methods: news reading, online classroom, lectures and conferences, and visits to red bases. The distribution of the number of people who use the above four methods for the auxiliary learning of Civics and Political Science courses is shown in Figure 1.

The steps of using Markov model for data analysis of students' study assistance methods in Civics courses are shown in Figure 2. Firstly, the Markov transfer matrix needs to be established, secondly, the Markov transfer matrix is initially solved by using statistical method, quadratic programming method, and other methods of mathematical calculation, then the results are normalized to obtain the transfer probability matrix, and finally the probability matrix is analyzed and summarized [31].

Randomly selecting 1000 college and university students to conduct a survey of Civics course learning, to investigate their main auxiliary learning methods used in Civics course learning, where the numbers from 1 to 5 represent news reading, online classes, lecture meetings, visits and exchanges and other learning methods, and data processing and construction of transfer matrix, auxiliary learning methods statistical results are shown in Figure 3.

Based on the preliminary understanding of the above students' situation and the analysis of the pattern, the following hypotheses can be made about the Markov model of students' learning auxiliary learning methods for Civics and Political Science courses:

(1) The number of students using learning methods that are transferred from one state to adjacent states is 10 (data are counted by number).
(2) The transfer order is from state 1 to state 2, state 2 to state 3, state 3 to state 4, state 4 to state 5, and state 5 to state 1, respectively.

For a sample of 1000 school and university students, the number of transfers using different learning methods for the number of students can be calculated, and the transfer probability matrix can be calculated by the formula. Using this transfer matrix, the auxiliary learning methods used by 50,000 students in their Civics courses can be predicted, and the results obtained are shown in Figure 4.

4. Results and Discussion

According to the predicted results of the auxiliary learning methods used by 50,000 students in the study of Civics and

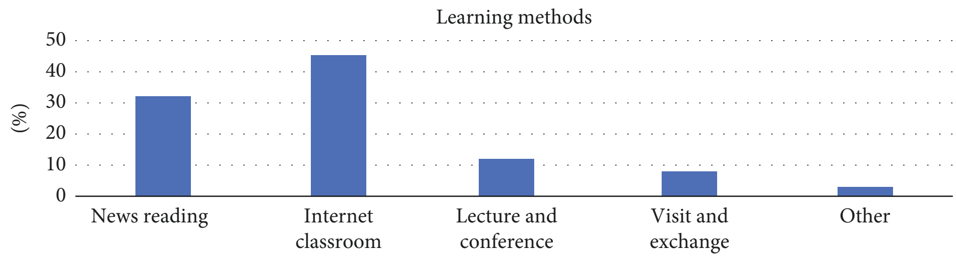


Figure 1: Distribution of the number of people using different auxiliary learning methods for Civics courses.

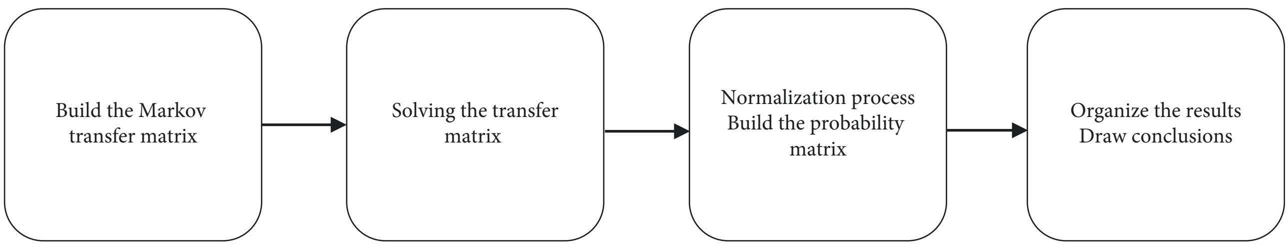


Figure 2: Steps of analyzing students' auxiliary methods of learning Civics courses using Markov model.

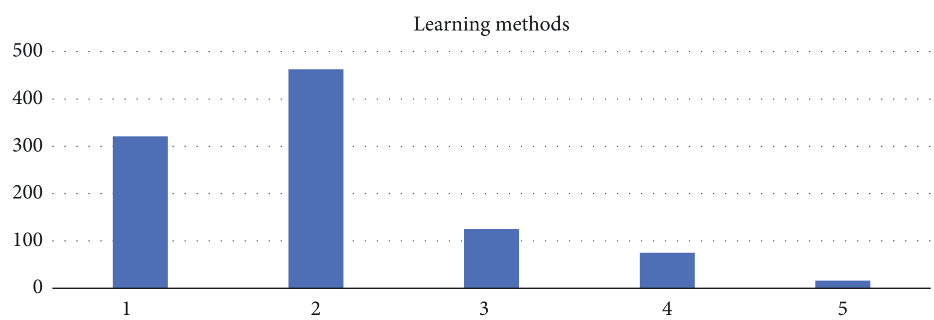


Figure 3: Statistics of 1000 university, school, and college students' auxiliary learning methods of Civics and Political Science courses.

Political Science courses, the fan chart shown in Figure 5 can be obtained. The largest number of students, 47.2%, used online classes for learning, followed by those who learned by reading news, about 33.7%, followed by those who learned by attending lectures and conferences, 11.8%, and finally those who learned by visiting red bases and other attractions, 6.8%. This is a small difference from the actual results of the statistics. This shows that the Markov model can be used to predict the auxiliary methods used by students in Civics courses in the future and to better utilize the results of this study for course improvement and study guidance.

From the above results, it can be seen that the online classroom is the most widely used learning support method among students. Combined with the survey on the frequency of Internet use by college and university students in recent years, the above results can be used to predict the frequency of Internet use by college and university students in the next three years.

After the questionnaire survey, the curves of the change of frequency of using online classroom for learning by junior and senior high school students in recent years are shown in Figure 6, and the Markov model is established according to the method outlined above, and the Markov transfer matrix

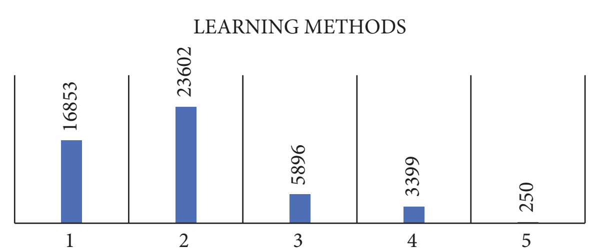


Figure 4: Prediction of the number of different supplementary methods used by 50,000 students in their Civics course study.

is set, then the mathematical calculation methods such as statistical method and quadratic programming method are used to solve the Markov transfer matrix initially, then the results are normalized to obtain the transfer probability matrix. The analysis and summary were conducted to obtain the frequency change curve of university and school students using online classroom for the three years from 2022 to 2024, as shown in Figure 7.

From the above results, we can see that online classroom has played an increasingly important role in students' learning life, and setting up online classroom for students' Civic Science course assisted learning has become urgent.

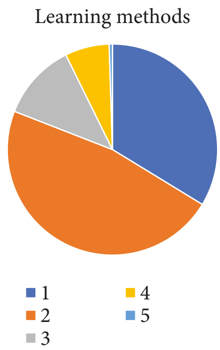


Figure 5: Sectoral graph of the percentage of students using different assistive methods in their Civics courses.

Among them, constructing Civic Science network-assisted teaching system is a good method, which can combine some of the abovementioned methods of assisted learning and share relevant videos and electronic versions of book materials about Civic Science courses for all students including primary and secondary school students and also provide a platform for teachers and students to exchange and learn. Here, university, high school, and elementary school students can exchange relevant knowledge. It can also provide a platform for teachers and students to exchange knowledge and expand their knowledge and make up for the shortcomings of classroom teaching, which is of great importance and far-reaching significance to students' development. The working module diagram of the constructed Civic Science Course Learning Support System is shown in Figure 8. The main modules of the system can be divided into five types, which are online courses, assignments and quizzes, lecture meetings, discussion and exchange, and system management.

The main purpose of the "online course" learning module is to provide students at different stages with different information about classroom knowledge, hot news, and real-time national events. Because students at different stages learn different specific knowledge, this module can show the textbook knowledge and extracurricular knowledge that need to be read and understood at different stages in the form of videos and pictures, so that students with different needs can study selectively according to their actual situation and improve the learning efficiency.

The "homework assessment" is an area open to both teachers and students. Teachers mainly assign and correct online homework for their own classes and answer students' questions when they are asked online. Students submit their assignments on time and provide feedback on their completion according to the teacher's schedule. This is a very efficient way to learn, and in the case of objective questions, the system can automatically correct them, which to a certain extent also improves efficiency. Classroom quizzes, as the name implies, are tests on what is learned in class. The main purpose of this module is to give students the opportunity to consolidate and test themselves on relevant knowledge points after systematic learning of Civics in the classroom. Therefore, the module can include special exercises and quizzes. The special exercises are the quizzes after each chapter to consolidate what they have learned and try to do some extracurricular extensions, while the thematic quizzes are in the form of online exams for students to answer within a certain period of time, so as to test the degree of students' integration of what they have learned.

The module "Lectures and Conferences" can replace students' participation in lectures and seminars offline, with more flexible time online, and can include such current affairs conferences as the Fourth Plenary Session of the 19th Central Committee, which are of national concern and significance, for students to study and research. It is also possible to hold activities such as lectures and symposiums online, which can avoid offline meetings where the number of people cannot get together and can reduce contact, which is also helpful for epidemic control and can be attended by students at all levels of school and university. Teachers at all levels can also make presentations, providing a good platform and opportunity for cross-learning and disciplinary integration.

The "Discussion and Communication" module is primarily open to students. Students of all grades can ask questions and communicate with each other on this platform, and they can reply to each other and speak freely.

Finally, the module of "System Management" is mainly to coordinate and organize, including user registration management, information release, question bank management, video upload, and other functions. Registration management is to verify the identity of users who register for the learning platform; information release is mainly for system administrators to release and modify information about courses and issue meeting notices, etc.; question bank management is mainly for teachers to examine and test their students according to the actual situation; and video upload is mainly for teachers to upload Civic Science course videos and current affairs content to the website for students to watch and study. The video uploads allow teachers to upload the videos of Civics courses and current affairs content to the website for students to watch and learn.

This kind of ideological and political auxiliary teaching system provides a new way for students to learn ideological and political courses. Through this system, students can not be restricted by time, place, and other situations, according to their own actual situation to learn but can communicate and discuss; the development of students learning is of great significance.

5. Summary and Prospect

5.1. Summary. This paper analyzes the auxiliary learning methods used by students in ideological and political courses through the establishment of Markov model. The main work is as follows:

(1) This paper introduces the significance and importance of studying the auxiliary methods of students' ideological and political courses.

(2) This paper introduces the concept, significance, and establishment method of Markov model and investigates the application scenarios of Markov model in practice through literature.

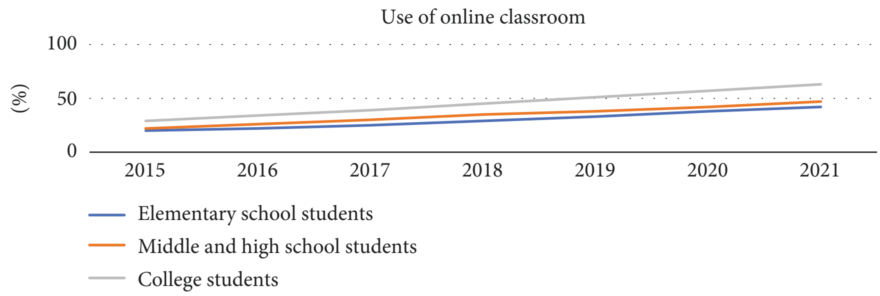


Figure 6: Curve of the change of the frequency of using online classroom for learning by school and university students in recent years.

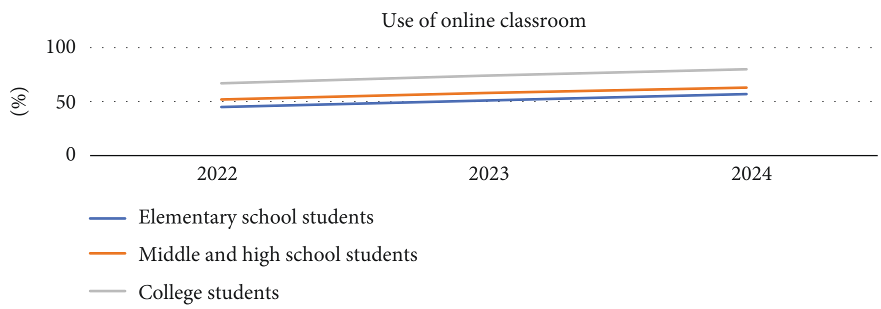


Figure 7: Curves of the change of frequency of using online classroom for learning by university and junior high school students in the latter three years using Markov model.

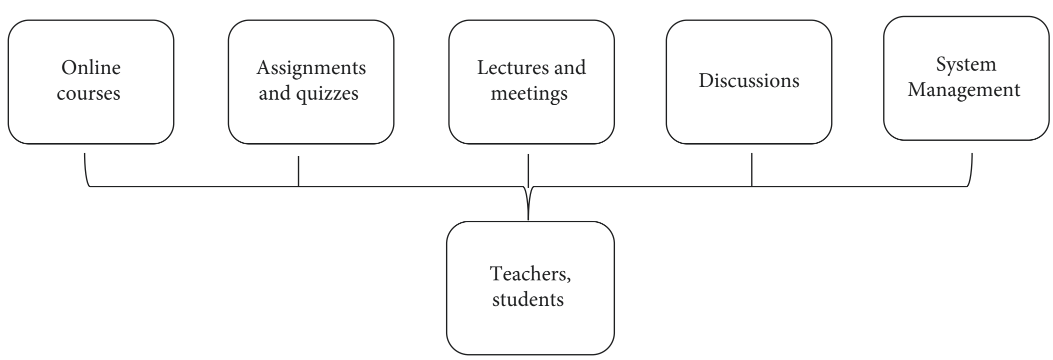


Figure 8: Working module of the simulated civic studies learning support system.

(3) A Markov model, a supplemental learning method used by students in political science courses, was developed and results were obtained by creating transfer matrices and calculating transition probabilities.

(4) In view of the research results, suggestions are put forward for further improvement of ideological and political curriculum, especially for the use of network classroom assisted teaching method.

5.2. Prospect. Markov model is a statistical probability model, which has great value and significance. The establishment of Markov model can be used for statistics, prediction, and evaluation of many different fields of research. It can be imagined that, with the rapid development of various technologies, Markov model can be combined with a variety of other models and algorithms, providing more help to many deeper fields of research.

In the future, the network classroom learning will become one of the main auxiliary methods for the student to study the course, because the network class has many advantages; it is not restricted by time and place, and through the network platform, the students of all ages can have education curriculum content of what they learn to share and ask questions and to expand the scope of study. It is of great significance to the growth of students.

The prediction and statistics of the assisted learning method of students' ideological and political course based on Markov model proposed in this paper still have many shortcomings, which need to be further improved.

First of all, establish a model of one thousand media and students in the process of the statistics; only do the statistics of the single option, but in fact many students in the education course learning will use a lot of kinds of methods; the alternative to the follow-up to establish state transition matrix and probability matrix has a great influence; matrix becomes more complex and needs further consideration.

Second, organize students to use Markov models to predict future learning, because the progress of science and technology, the international situation, the evolution of the epidemic, etc. all have a big impact on students' learning manner, but these factors are not calculated in the model, which will reduce the accuracy of the model. Later, multiple variables should be set and other statistical models combined to make a more accurate prediction of students' use of online classroom learning.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This research was supported by Xi'an Social Science Planning Fund Management Office (no. ZX24).

References

[1] G. Gruner, F. Feinberg, J. Venables et al., "An undergraduate medical education framework for refugee and migrant health: curriculum development and conceptual approaches," *BMC Medical Education*, vol. 22, no. 1, p. 374, 2022.

[2] J. Kendall, "The utility of game-based approaches within the PE curriculum design and implementation process to develop more knowledgeable others," *Strategies*, vol. 35, no. 3, 2022.

[3] Y. Tang, J. Wei, Y. Xin, and X. Mou, "Curriculum design for vocational industrial robotics: the quality function deployment (QFD) approach used in China for vocational course design," *SN Social Sciences*, vol. 2, no. 5, p. 64, 2022.

[4] R. Mishra, B. Shraddha, and R. L. Raina, "Towards transforming higher education institutions (HEIS) of India: an approach at JKLU, Jaipur on "curriculum" & "pedagogy"," *GYANODAYA - The Journal of Progressive Education*, vol. 13, 2022.

[5] M. Harrison and H. Alberti, "How does the introduction of a new year three GP curriculum affect future commitment to teach? An evaluation using a realist approach," *Education for Primary Care*, vol. 33, no. 2, pp. 92–101, 2022.

[6] Z. Jianfang and Z. Shihua, "Research on the development and practice approach of home-school Co-education curriculum," *International Journal of Educational Management*, vol. 7, no. 1, 2022.

[7] L. Estrada, M. A. Williams, and C. S. Williams, "A competency-guided approach to optimizing a physician-scientist curriculum," *Medical Science Educator*, vol. 32, no. 2, pp. 523–528, 2022.

[8] A. P. Silberman, R. A. Rozenfeld, and D. O. Kessler, "Core components of a pediatric critical care transport communication curriculum: a modified delphi approach," *Air Medical Journal*, vol. 41, no. 2, pp. 217–221, 2022.

[9] Juttafrst, F. Sinéad, and C. Victoria, "Ro1e deve1opment in psychodrama training findings and challenges," *Prospects*, 2022.

[10] J. Bron, C. Bovill, and W. Veugelers, "Students experiencing and developing democratic citizenship through curriculum negotiation: the relevance of Garth Boomer's approach," *Curriculum Perspectives*, vol. 42, 2022.

[11] H. J. Wong, M. Attaar, M. Campbell et al., "A modular simulation curriculum to teach endoscopic stenting to practicing surgeons: an "Into the fire" approach," *Surgical Endoscopy*, pp. 1–9, 2022.

[12] M. Gaotlhobogwe, S. Trivedi, J. Kasozi, and T. Kebalepile, "Effectiveness of reception class teachers' pedagogical approaches in delivering pre-primary curriculum - evidence from practice," *South African Journal of Chemical Engineering*, vol. 12, no. 1, 2022.

[13] B. Michael, H. Michael, and R. M. Katie, "An integrated point-of-care ultrasound curriculum: an evidence-based approach," *Journal of Physician Assistant Education*, vol. 33, 2022.

[14] T. Sun and X. Zhao, "Evaluation method of ideological and political curriculum reform effect in colleges and universities based on fine-grained learning," in *Proceedings of the 2021 4th International Conference on Humanities Education and Social Sciences (ICHESS 2021)*, p. 615, Xishuangbanna, China, October 2022.

[15] A. Barbara, M. Lavinia, and B. Cole, "Diversifying the social policy curriculum: a collaborative approach," *Social Policy and Society*, vol. 21, no. 1, 2022.

[16] J. R. Brown, T. Devick, and C. Zielinski, "A culturally relevant approach to food customs in the German curriculum," *Die Unterrichtspraxis/Teaching German*, vol. 54, no. 2, pp. 181–199, 2021.

[17] J. Kennedy, A. Percy, L. Thomas, C. Moyle, and J. Delahunty, "Holding space for an aboriginal approach towards curriculum reconciliation in an Australian university," *Teaching in Higher Education*, vol. 26, no. 7-8, pp. 1060–1076, 2021.

[18] H. ElGeed, M. S. El Hajj, R. Ali, and A. Awaisu, "The utilization of simulated patients for teaching and learning in the pharmacy curriculum: exploring pharmacy students' and recent alumni's perceptions using mixed-methods approach," *BMC Medical Education*, vol. 21, no. 1, 2021.

[19] J. A. Adams and B. M. Dewsbury, "Student preference for course approach to pedagogically different methodologies in anatomy and physiology," *Advances in Physiology Education*, vol. 46, no. 1, pp. 45–55, 2022.

[20] D. W. Fleenor, H. G. Atkinson, R. Karani, S. Lerner, S. Leisman, and D. Marin, "An innovative approach for integrating mandatory, longitudinal spirituality training into the medical school curriculum," *Academic Medicine*, vol. 97, no. 2, 2021.

[21] H. J. Wong, M. Attaar, H. M. Hedberg et al., "Implementation of a medical student surgical skills simulation curriculum: a learner-centered "into the fire" approach," *Journal of the American College of Surgeons*, vol. 233, no. 5, 2021.

[22] B. N. Guest, C. Chandrakanthan, K. Bascombe, and J. Watkins, "Pharmacology for physician associate programmes: a collaborative, flexible and responsive approach to curriculum design," *Future Healthcare Journal*, vol. 8, no. 3, pp. e580–e584, 2021.

[23] V. Emily, S. Lueders Kerry, C. Killen, and E. Kenny, "Pre-employment transition services and the expanded core curriculum: a low vision rehabilitation approach," *Journal of Visual Impairment & Blindness*, vol. 115, no. 6, 2021.

[24] A. J. Mills and J. F. Hair, "Consumer insights: a turning point for marketing research education," *Journal of Marketing Education*, vol. 43, no. 3, pp. 279–284, 2021.

[25] D. Chadha, J. Campbell, M. Maraj et al., "Engaging students to shape their own learning: driving curriculum re-design using a theory of change approach," *Education for Chemical Engineers*, vol. 38, 2022.

[26] E. Jennifer, "Critically reading the canon: culturally sustaining approaches to a prescribed literature curriculum," *Journal of Adolescent & Adult Literacy*, vol. 65, no. 4, 2021.

[27] J. L. Piñeiro, E. Castro-Rodríguez, and E. Castro, "What problem-solving knowledge is required in mathematical teaching? A curricular approach," *Curriculum Perspectives*, vol. 42, 2021.

[28] V. P. Goby, "A point of view: can a narrative imagination approach to the business curriculum contribute to global citizenship and sustainability?" *Globalisation, Societies and Education*, vol. 19, no. 5, pp. 605–609, 2021.

[29] C. James and P. Camilleri, "Teachers' learning in extraordinary times: shifting to a digitally facilitated approach to lesson study," *International Journal for Lesson & Learning Studies*, vol. 10, no. 2, 2021.

[30] A. Y. Supriadi and E. Syaiful, "The role of principal in implementing the 2013 curriculum-based learning at SMP negeri 6 unggul sekayu musi banyuasin regency," in *Proceedings of the International Conference on Education Universitas PGRI Palembang (INCoEPP 2021)*, p. 565, Parsippany-Troy Hills, NJ, USA, July 2021.

[31] N. Nefy Puteri and A. Lathifah, "Experiential learning method and hands-on approach for optimizing learning on signal and systems course," in *Proceedings of the 3rd International Conference on Educational Development and Quality Assurance (ICED-QA 2020)*, p. 506, Padang, Indonesia, February 2021.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 2871481, 9 pages
https://doi.org/10.1155/2022/2871481

Research Article

Analysis and Evaluation of Sports Effect Based on Random Forest Algorithm under Big Data

Kai Liang [1] and Dongdong Zang [2]

[1]*Department of Physical Education, Jiangsu Vocational College of Agriculture and Forestry, Zhenjiang, Jiangsu 212400, China*
[2]*Economics and Humanities College, Jiangsu Vocational College of Agriculture and Forestry, Zhenjiang, Jiangsu 212400, China*

Correspondence should be addressed to Kai Liang; liangkai@jsafc.edu.cn

Received 11 May 2022; Revised 29 June 2022; Accepted 21 July 2022; Published 29 August 2022

Academic Editor: Jiafu Su

Under the background of big data, all walks of life have carried out in-depth informatization construction. As an important part of national education, the informatization construction of universities cannot be ignored. In recent years, the state has promulgated various policies and regulations, which provide a guarantee for the normal development of school physical education and can improve the current situation of college students' declining physical health to a certain extent. This study attempts to set some specific indicators to promote the better implementation of college sports so that it can supervise and inspect them from all aspects during the actual development of college sports and provide a standard for measuring the implementation effect of college sports. Based on the RF (Random Forest) algorithm, this paper puts forward an evaluation algorithm of students' sports achievements, which can be used to predict students' sports achievements and at the same time, find out the factors that affect students' learning and rank them in importance. The results show that the confidence level of sports effect evaluation by this method is high, and the average confidence level is above 0.96. *Conclusion*. This method has improved the effect of sports effect evaluation, thus effectively guiding sports skills training and improving sports skills.

1. Introduction

Vigorously developing college students' physical education is a major event to promote the construction of a healthy China and realize the Chinese dream. It is necessary to make college students love sports and participate in sports as a living habit so that college students can have a healthy physique, good interpersonal relationships, and strong character. It is necessary to mobilize all social forces to protect the health of college students. The state has made constant regulations on the quality and quantity of sunshine sports, which is enough to show that the state attaches importance to college students' participation in sunshine sports and cares about students' physical health. The research shows that there are many declining trends of college students' physical health in China, and one of the most important reasons is that they do not cultivate good sports habits and cannot do extracurricular exercises by scientific methods or means.

Sports achievement is mainly a measure of whether there is a gap between students' physical fitness and expected goals and the means to measure this gap. It can only show the results of students' physical exercise but cannot reflect the causes of such results. Yang and Li introduced the classic Apriori algorithm and the famous DT (Decision Tree) ID3 algorithm of association rules [1] and used Apriori association rules algorithm to mine the influence degree of a course's excellence on other courses. Attar et al. put forward a design idea of multistrategies, combining DM technology with statistical analysis and using DT-based classification mining method to analyze the data in the student's sports achievement database, and generate the student's sports achievement DT, which can intuitively show the position of a certain achievement in different grade calculation methods and provide evaluation information for teaching departments [2]. Wu et al. used the classical C4.5 algorithm to construct the prediction model of students' professional achievement, including data processing, pruning optimization of DT, and evaluation of the performance of the model by cross validation, to find out the potential students' behavior rules [3]. Yuan et al., based on the theory of performance evaluation, designed a performance evaluation index system to measure students' growth according to the process performance, taking innovative and designed experiments in the teaching stage of students' professional courses as the process [4].

Big data has the characteristics of large amount of data, various types, low-value density, high speed, and high timeliness. The era of big data poses new challenges to people's existing data control ability and also provides unprecedented space and potential for people to gain deeper and comprehensive insight. With the continuous accumulation and expansion of data, data mining is particularly important. Therefore, how to scientifically and reasonably mine students' sports achievement data and find the potential knowledge and information in the achievement data so as to guide students and teaching staff in the next stage of study and work has a positive guiding significance for the improvement of teaching quality. Therefore, this paper proposes to use RF (Random Forest) algorithm to accurately show many aspects of college students' sports performance analysis from a quantitative point of view so as to help teachers, students, and physical education departments formulate corresponding measures, which is conducive to the cultivation of students' physical exercise habits, the promotion of healthy physique, and the improvement of physical education teaching quality.

2. Related Work

2.1. Analysis and Research on Sports Effect. When human beings are engaged in sports activities, they cannot live without the environment, so they must exchange material, energy, and information with the natural and social environment. The factors that make up the environment are bound to be connected with sports activities as a whole or with each other.

Yi and Fan think that the school sports environment refers to the sports environment within a school, and the school sports environment, school moral education, learning, and other environments together constitute the school education environment [5]. Ouyang and Liu divided the campus sports environment into five subenvironments from the content, namely, sports consciousness environment, training environment, teaching environment, hygiene environment, and competition environment, and discussed the influence of subenvironments on students' physical and mental development [6]. Bideau et al., under the guidance of the concept of lifelong physical education, analyzed the result evaluation and the process evaluation so that the university physical education teaching evaluation can be evaluated from two aspects: process and result, making it more scientific and reasonable [7]. Li et al. calculated the weights of each index and each subsystem by AHP (analytic hierarchy process) and made single ranking and total ranking. The weighted linear method is used to synthesize the indexes of the system to evaluate the development level of each subsystem and the system [8].

Yin and Cui studied the application of PCA (principal component analysis) and Bayesian KNN (K-nearest neighbor) algorithm in DM (data mining). PCA was used to reduce the dimension of characteristic data, and Bayesian K-nearest neighbor algorithm was used to sell and classify, that is, to predict career direction [9]. Sarlis and Tjortjis put forward an improved algorithm after researching and analyzing ID3 algorithm, which can mine and analyze the data stored in the educational administration management system, so as to find out the relationship between curriculum settings and provide some data basis for the university's achievement statistics to assist with decision-making [10]. Harvey et al. used model long-distance running to establish fuzzy evaluation model for data analysis. They solved the problem of conversion between national physical health standards and students' sports scores [11].

2.2. Research on RF Algorithm. RF is a multi-DT model that combines multiple DTs to make predictions. It has the advantages of ranking the importance of variables, high prediction accuracy, strong antinoise ability, few adjustment parameters, strong adaptability to data sets, and effectively avoiding "overfitting" phenomenon.

Poona et al. used RF to select the characteristic genes of five groups of gene expression profile data, such as lung cancer, stomach cancer, and colon cancer, and combined the selected characteristic genes with support vector machine to classify the original data set [12]. Provost et al. improved the node partition method of each tree in RF algorithm and proposed an improved RF algorithm. This algorithm takes the customer's life value as the dividing node and solves the problem of unbalanced data distribution [13]. Lee et al. used RF algorithm to classify different land types, and the results showed that RF algorithm could accurately classify land cover, which was consistent with expectations [14]. Naghibi et al. adjusted the weight of each sample according to whether the samples were correctly classified and proposed a weighted RF combining algorithm based on adaptive lifting [15]. Ming et al. applied greedy random adaptive search program with annealing randomness to RF algorithm, and its feature selection can improve the accuracy of classification [16].

Ghorbanian et al. designed and implemented an improved RF algorithm by using Spark distributed storage computing technology and proposed an improved RF algorithm based on Spark [17]. Bei et al. put forward the RF model of fusion factor analysis. The main innovation lies in constructing feature groups by factor analysis and then randomly extracting features according to the number ratio of features to form candidate subsets of each split node. Improve the accuracy and convergence speed of the model [18]. Raphael et al. used RF algorithm to study the nonlinear relationship between biomass pyrolysis kinetic parameters, biomass species, and various heating conditions and put forward an RF model to predict biomass pyrolysis kinetic parameters. The training results show that RF model can predict biomass pyrolysis kinetic parameters under different heating conditions well [19]. Lin et al. used the improved RF algorithm to build an evaluation model based on the historical data of wheat breeding and put forward the application of RF algorithm in the auxiliary evaluation of wheat breeding [20]. The experimental results show that the proposed algorithm has achieved good results in wheat germplasm evaluation and can assist breeders in variety selection.

3. Research Method

3.1. Analysis Method of Sports Effect. Nowadays, the computer network has become very popular in universities. Using DM technology to establish a university sports achievement management system should be able to provide administrators, teachers, and students with sufficient information and quick query means, complete the work of teachers' scoring, and make statistics, analysis, and processing of data. It is conducive to scientifically set and arrange one's own fitness goals and exercise plans. Physical exercise can increase the opportunities of contact and communication between people. By taking part in sports activities, you can forget your troubles and pains, eliminate your loneliness, and gradually form the awareness and habit of interacting with others. It also helps students to change their bad behaviors and living habits that affect their health.

Cooperation ability is not only a necessary quality for participants in sports activities, but also a kind of ability that needs to be developed through sports activities. College students' participation in sports activities, especially in collective sports activities, requires the full cooperation between themselves and others, which not only enables the collective goal to be achieved, but also gives full play to the role of individuals.

DM is the process of extracting hidden information and knowledge from a large number of incomplete, noisy, fuzzy, and random data, which people do not know in advance but are potentially useful. It is of great significance to detect these anomalies and deviations, which can better optimize the model. There are many different types of deviations, such as special cases that do not conform to the basic rules, abnormal cases in clustering, values that change with time, and so on. From the perspective of DM, the application of DM technology in scientific management of physical examination results can be divided into three stages: data preparation, pattern search, and visual representation. Figure 1 shows the general framework model of scientific management of physical examination results.

The scientific management framework of physical examination results described in Figure 1 can be realized by four steps: the establishment of data objects, data preprocessing, data mining, and the representation and evaluation of pattern knowledge. In the scientific management of physical examination results, there are many data sets involved, such as teacher information, student information, examination results, and so on. Therefore, it is the first condition to establish a suitable DM object and make clear the goal of DM.

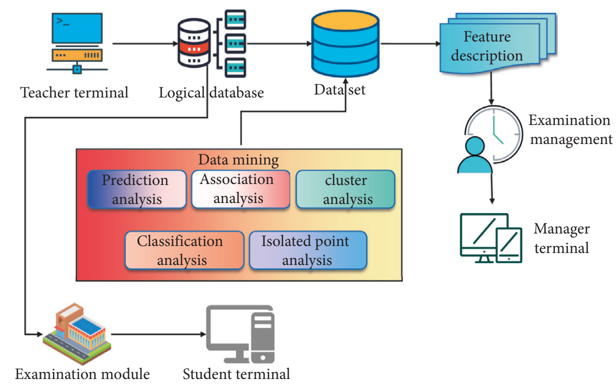


Figure 1: General framework of scientific management of general physical examination results.

Data preprocessing is an important link in the process of DM, especially when the data containing noise, incompleteness, or even inconsistency are DM; it is even more necessary to improve the quality of DM objects and finally achieve the purpose of improving the quality of pattern knowledge acquired by DM.

Information gain metrics tend to test with many outputs, that is, tend to select attributes with a large number of values [15]. Information gain rate is the expansion of information gain, which can overcome this problem. It normalizes the information gain by using the "split information" value. Split information is defined as follows:

$$\operatorname{splitinfo}(A) = -\sum_{j=1}^{v} \frac{|D_j|}{|D|} * \log_2\left(\frac{|D_j|}{|D|}\right). \quad (1)$$

This value represents the information generated by dividing the training data set D into v partitions corresponding to the attribute A test.

The submodel of RF algorithm is CART (Classification and Regression Tree) model. CART algorithm uses binary tree construction in binary recursion mode, and every time the whole sample set is divided into two subsets, thus generating two branches of subtrees. Gini index is mainly used to indicate the impurity of data set. Gini index of sample set D is defined as

$$\operatorname{Gini}(D) = 1 - \sum_{i=1}^{m} p_i^2, \quad (2)$$

where p_i is expressed as the probability that the data in sample set D belongs to category C_j. If the binary partition of sample set D based on attribute A will be divided into two subsets D_1, D_2, then the Gini index based on this partition can be calculated as the weighted sum of the impurity of each partition:

$$\operatorname{Gini}_A(D) = \frac{|D_1|}{|D|}\operatorname{Gini}(D_1) + \frac{|D_2|}{|D|}\operatorname{Gini}(D_2). \quad (3)$$

Applying DM technology to improve the level of systematic analysis, through in-depth and scientific analysis of examination results, students can know the relative position of individuals in the group from horizontal and vertical aspects, make clear the gains and losses of individual learning and the differences with others, and promote individual learning. At the same time, teachers and school education decision-makers can understand the problems existing in teaching and reflect on the quality of education and teaching.

Introducing the observation set of sports effect evaluation, comprehensively analyzing and scheduling the sports effect evaluation model, and setting the value of the reliability weight ω of sports effect evaluation are represented as

$$\omega = \omega_{\max} - t\frac{\omega_{\max} - \omega_{\min}}{T_{\max}}. \tag{4}$$

$\omega_{\max}, \omega_{\min}$ represents the regulation coefficient of sports effect evaluation, $T_{\max}$ is the time delay, and t is the correlation factor of sports effect evaluation. Combining with the scheduling method of association rules, the cloth fusion model of sports effect evaluation is obtained [6], and the calculation formula is

$$SL_i = \begin{cases} L_i, & i = 1, \\ \text{New}_i, & \text{otherwise}. \end{cases} \tag{5}$$

New_i represents the distributed scheduling set of sports effect evaluation, thus constructing the fuzzy association rule set and prior distribution set of sports effect evaluation. By using the big data fusion analysis method [7], the characteristic analysis model of sports effect evaluation is established to improve the effectiveness of sports effect evaluation.

3.2. Construction of Evaluation Model of Students' Sports Performance. The construction of university sports environment indicators must follow the scientific principle, and the selection and construction of evaluation indicators should be based on scientific nature so as to avoid mistakes caused by empiricism and subjective judgment as far as possible so as to ensure the reliability of research results. Therefore, we should select the indicators from multiple dimensions, comprehensively and systematically, and try our best to make the selected indicators more comprehensive and systematic. Unless there are related scientific methods for conversion and measurement, it will affect the credibility of the evaluation. Ensure the operability, availability, and practicability of the evaluation process of college students' campus sports environment, which are the basis of accurate and feasible evaluation results and popularization and application.

Weight is a kind of quantitative value that compares and balances the evaluation object with a quantitative relationship to show its relative importance. According to different evaluation purposes, the weight set according to the importance of indicators is called the weight coefficient [17]. The general subjective weighting methods are illegal, analytic hierarchy process and weighting method. Subjective weighting requires evaluators to score the importance of each index and make an evaluation. Its characteristic is that it can humanize the opinions of experts and evaluate the importance of each index [19].

Although the accuracy and efficiency of the traditional RF algorithm are better than other classification algorithms in dealing with high-dimensional data, there are some defects in some data: if the data dealing with classification problems are unbalanced data, the classification results predicted by RF algorithm are not ideal; if the voting weights are the same, there will always be errors in the final results, especially on the unbalanced data sets, and the classification results tend to be biased towards most categories. The default parameters are not necessarily the best choice, and it is not efficient to use the traditional grid search method to find the best parameters.

In view of the above problems existing in the evaluation of students' sports performance, this paper puts forward an evaluation algorithm of students' sports performance based on RF algorithm, which can be used to evaluate students' sports performance, predict the students' related sports test scores of senior grades through the students' sports test scores of junior grades, sort the variables in importance, and find out the key factors that affect students' academic performance.

On the basis of the previous research, this paper proposes a new comprehensive optimization algorithm, which mainly integrates simulated annealing algorithm into the execution process of RF algorithm and uses binary coding, $\text{OOB}_{(\text{OutofBag})}$ error minimization, and simulated annealing operation to obtain the best combination of feature selection O, DT scale K, feature subset scale N, and DT weight W_t.

We set the objective function of the improved RF algorithm as follows:

$$\begin{aligned} f\left(\begin{matrix} K^*, O^*, \{\text{Attribute}_i | i = 1, 2, \ldots, M\}, \\ \{w_j | j = 1, 2, \ldots, K\} \end{matrix} \right) \\ = \arg\min(\text{avg OOB error}), \end{aligned} \tag{6}$$

where K, O is a real number, $K \in [0, 500]$, and $w_j \in [0, 15]$. The value of Attribute_i is 0 or 1, where 0 indicates that the feature has not been selected, and 1 indicates that the feature has been selected.

If the sample set is F, a few classes in the data set are $F_{\min}$, most classes are $F_{\max}$, and the imbalance coefficient E is shown in the following formula:

$$E = \frac{F_{\min}}{F_{\max}}. \tag{7}$$

The classification ability of each DT in RF is different. In the traditional RF algorithm, the votes of each DT are the same, which is unreasonable. In order to improve the classification ability of RF algorithm on unbalanced data, avoid the situation where the classification results of RF algorithm are seriously biased towards the majority class and make it perform well on the minority class of unbalanced data sets. The weight of DT voting is shown in the following formula:

$$W_i = E_i + \frac{\min_i}{L}. \tag{8}$$

E_i is the unbalance coefficient of the data sampled from the training set of the ith DT, $\min_i$ is the number of minority samples with correct prediction of this DT, and L is the length of data samples to construct this DT.

According to the Gini index, the importance of these factors is ranked, in order to help students make targeted tutoring according to the experimental results, guide the teaching work of teaching staff, provide decision-making support for education management departments, and improve teaching quality. Sample collection of students' sports achievements:

$$SC = \{(x_i, y_i), \quad i = 1, 2, \ldots, n\}. \tag{9}$$

$X = X(X_1, \ldots, X_n)$ is the explanatory variable, that is, the course name related to the course to be predicted, and Y is the target variable, that is, the student's sports achievement to be predicted.

The optimized prediction value of sports effect evaluation is as follows:

$$x(t_{n+1}) = X_{m+1}(m). \tag{10}$$

The structure of student sports performance evaluation model based on RF algorithm is shown in Figure 2.

The basic idea of RF algorithm is as follows:

(1) From the original data set of students' sports scores after the above data preprocessing, the Bootstrap sampling method is adopted to randomly extract N scores training sets.

(2) DT is established for each performance training set, and according to the principle of minimum Gini index, the best attribute is selected for internal node splitting.

(3) Collect the prediction results of N DTs, and decide the grade category of the new sample according to the voting method.

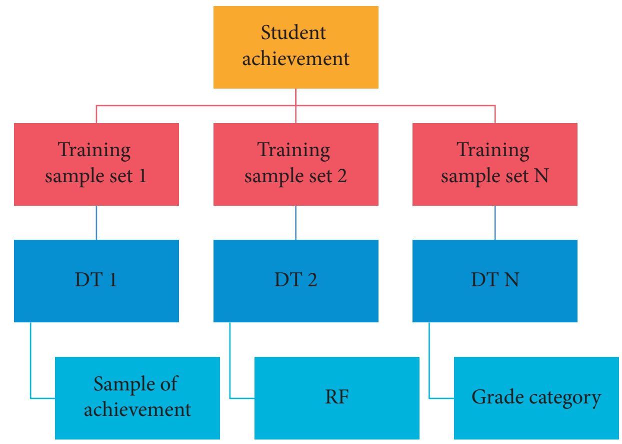


Figure 2: Evaluation model of students' sports performance based on RF algorithm.

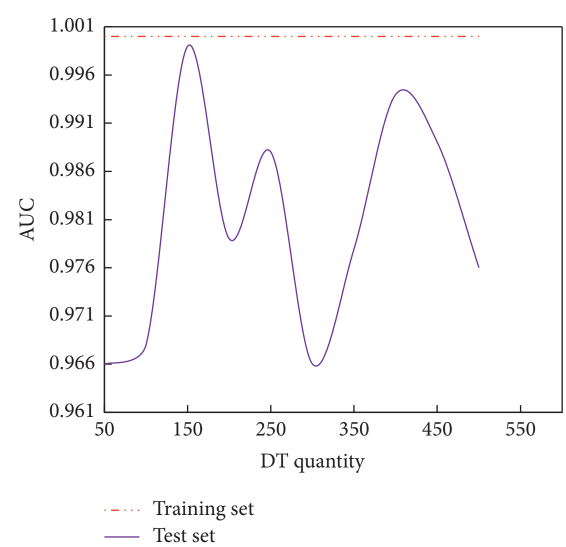


Figure 3: Influence of DT number on data set.

4. Result Analysis

In order to verify the effectiveness of the design model in sports effect evaluation, MATLAB is used for simulation analysis, and SPSS statistical analysis software is used for big data analysis of sports effect evaluation. The distribution sample length of statistical data of sports effect evaluation is 1026, the frequency of feature sampling is 1.44 kHz, and the training set size is 300. The results are shown in Figures 3 and 4.

It can be seen that the classification ability of this algorithm is improved compared with that of RF algorithm. The same AUC (Area Under Curve) value under the default parameters of this algorithm and the AUC value after the optimization of RF algorithm parameters indicate the effectiveness of the improved algorithm. From the comparison of parameters, it can be seen that parameter optimization on data sets is helpful to improve the classification ability of the algorithm. The smaller the RF scale, the lower the depth of the tree, and the higher the efficiency, the better the algorithm, under the condition that the Kappa coefficient and AUC value of the algorithm are guaranteed.

The quality of data determines the result of classification, so the data is preprocessed and then classified by RF. It can effectively improve the sensitivity of RF to unbalanced data. It is an important stage of college students' personal development, and all physical skills are developing vigorously. Excessive learning pressure may lead to students' physical and mental exhaustion. Therefore, we encourage students to participate in leisure sports. Moderate leisure sports can increase their vital capacity, promote students' quick thinking and clear mind, and help improve their learning efficiency. Through in-depth analysis of the existing achievements of students of the same major in universities, this paper explores the potential rules of achievement data, samples a large number of achievement data to form different training sets, and then forms a number of DTs to predict the students' achievements, respectively. Finally, the

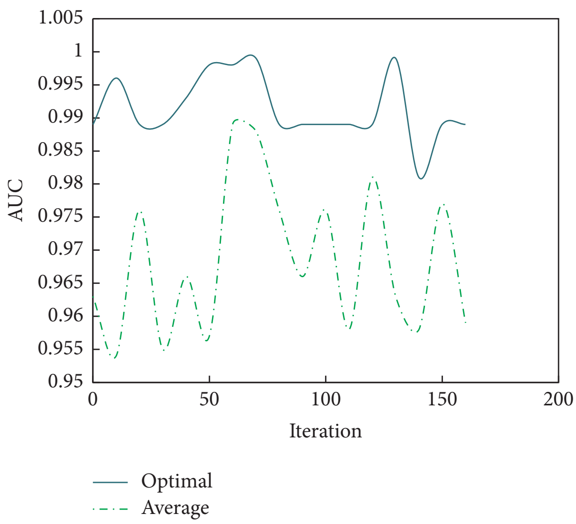


Figure 4: RF model iteration diagram.

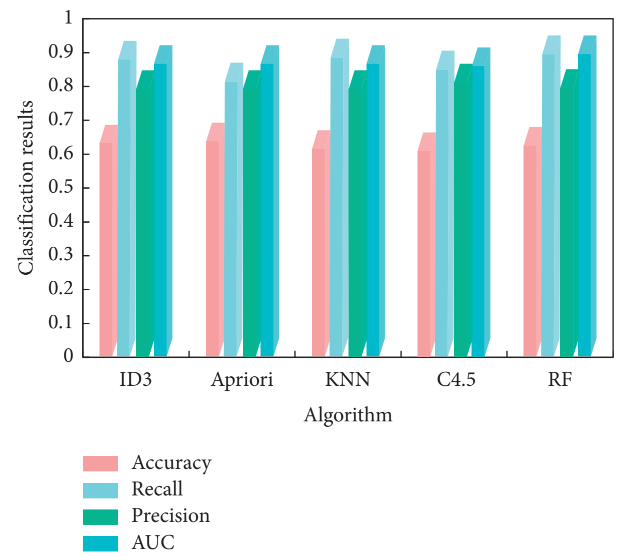


Figure 5: Comparison of classification results of students' sports performance data set on various basic algorithms.

risk level of students' sports achievements is obtained by integrating all DTs' prediction results.

Generally speaking, in independent tasks, competition has advantages, because in such tasks, the requirements for mutual cooperation among members are not very high, and the goal of individual activities is not to defeat others, but to accomplish the task. With the increasingly fierce competition in modern society, it is helpful for college students to adapt to the society well after they step out of school and enter the society by cultivating their competitive consciousness and ability. In order to ensure the fairness of the evaluation of college sports, it is convenient and accurate to use unified evaluation methods and means in the evaluation process, to make the subjective evaluation content tangible and digital, and to use computer data processing software for evaluation. In addition, it is necessary to combine statistical evaluation with expert evaluation to establish a fair and just evaluation system.

For the data set of students' sports achievements, the feature variables that need feature planning include students' activity in class, students' absenteeism, the number of times that students visit teaching resources after class, the number of times that students participate in course discussions, and students' satisfaction with the course. Other feature variables are calculated by using original values, such as birthplace with coded values and gender with Boolean values of 0 or 1.

In the process of realization, firstly, the discrete data is normalized to [0, 1] based on the log function standardization method, and the corresponding weight coefficient is assigned to each factor so that the students' classroom activity data can be weighted. Figure 5 shows the comparison of the classification results of students' sports performance data sets on various basic algorithms.

Figure 5 shows that we use a variety of basic DM algorithms for comparison and classification, trying to analyze the differences of various DM algorithms. From the experimental results, it can be seen that the RF algorithm has a slight advantage over other data algorithms in the student characteristic data set, but the overall difference is not big.

Table 1: Training data.

Project	Passed number	Number of failed students
Long jump	80	77
Long-distance run	55	102
Basketball	30	122
Short-distance run	28	126
Hurdle race	39	115
Volleyball	97	79
Shot	36	120
High jump	99	59

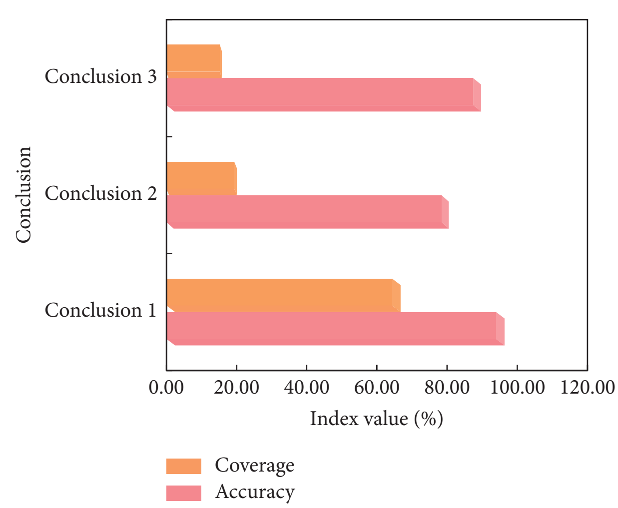


Figure 6: Schematic diagram of conclusion.

Table 2: Statistical analysis of sports effect evaluation.

Group	Sample number	Standard deviation	Maximum	Minimum
Contrast group	286	0.533	33.68	12.96
Training group	632	0.306	32.17	15.62

What attributes can determine whether students can pass the sprint by using DT technology? Take the whole sampling method to select some students' sports achievements as training sets, and copy these achievements records to the training example worksheet. Using the function, find out the number of people who passed and failed in a single subject in the training set, as shown in data Table 1.

The key to construct a good DT lies in how to choose a good logical judgment or attribute. It has been found that, in general or with a high probability, the smaller the tree, the stronger the prediction ability of the tree. To construct DT as small as possible, the key lies in choosing appropriate logical judgment or attribute. Here, the information gain is used for attribute selection.

According to the information gain of each attribute, the attribute of long-distance running should be selected as the root node of DT. Other results are used as the test set and used to test the accuracy of the generated DT. The following rules can be obtained from this DT, as shown in Figure 6.

The learning level of students' long-distance running will directly affect their learning effect of sprint events. The learning of hurdle events also has a certain influence on the learning of sprint events. Therefore, teachers should consider students' long-distance running foundation when teaching sprint. Students with good long-distance running and average sprint should pay more attention to the study of hurdle events.

Compared with the culture course, physical education class is not valued by many schools and parents, and because of the pressure of entering a higher school, the duration of students' physical exercise is shortened. Parents cannot fully recognize the benefits of physical activities for students and attach too much importance to the achievements of culture course, which leads to their neglect of the cultivation of students' physical quality. Due to the negative effects on the body, leisure and psychology after taking part in physical exercise, such as physical fatigue after long exercise time or high intensity, learning collapse will be affected. In addition, under the pressure of traditional ideas and crazy social environment, physical exercise and academic performance are negatively correlated in form.

By comparing the factors that affect students' physical exercise at each level, we can find that the experience of individual participation in physical education class or physical exercise constitutes the common factors that affect students' participation in physical exercise. In behavioral learning theory, operant conditioning holds that learning is based on the response results. Poor students in physical education class: some students are often criticized in class because they do not obey discipline or behave badly. In this case, even if they have their own favorite sports, their performance in physical education class is restricted, and

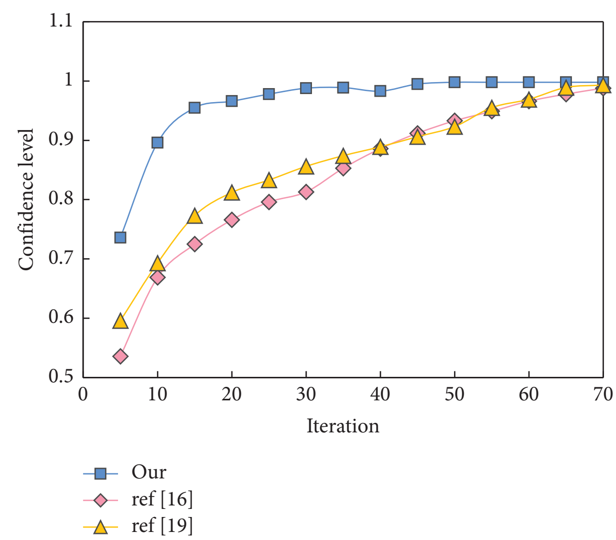


Figure 7: Comparison of confidence degree of sports effect evaluation.

their sports performance is not very good, which affects their understanding of sports to a certain extent.

Some students, though not doing well in the cultural course, have been active in physical education class, received positive responses, and had a positive experience in physical education class. According to the above simulation parameters, the sports effect is evaluated, and the statistical distribution sample set is shown in Table 2.

According to the results in Table 2, the big data analysis of sports effect evaluation is carried out, and the big DM results of sports effect evaluation are shown in Figure 7.

According to the analysis of Figure 7, the confidence level of sports effect evaluation by this method is high, and the average confidence level is above 0.96. The results of these professional practice courses are mainly graded by teachers according to the experimental results made by students. After observing the data, it is found that the scores of these courses are concentrated compared with those of professional theoretical courses, and the difference of scores is not obvious, which has little effect on analyzing students' later professional courses. Therefore, the experimental results are consistent with the actual situation. Among many factors that affect students' poor physical performance, "exercise situation" has the greatest influence on students' performance. For a student who never exercises, if his attendance rate in class is low and his foundation is average, then his exam results must be poor.

College students play an important role among the members of the society and undertake a strong social mission. Whether one can fulfill one's own responsibilities and gain good trust from others in sports not only plays an important role in promoting one's role in sports, but also makes certain adaptation for college students to enter the society in advance. Strengthen the education of college students' safety knowledge, improve their self-management ability, formulate a reasonable and planned extracurricular physical exercise schedule, and select sports with low sports risks characterized by fun, safety, health, and positivity according to the value of the risk factors studied so as to reduce the occurrence of risks.

Physical education teachers should constantly optimize the educational methods. Teachers should learn and progress from each other, probe into each other, sum up experiences, form an educational method suitable for college students, and guide college students to love sports, cultivate their interest in sports, and participate in sports consciously. Set up the leading organization of extracurricular sports together, formulate corresponding rules and regulations, and accomplish all tasks in organizing extracurricular sports activities in schools. In organizing extracurricular sports activities, more college students have the opportunity to participate in sports, feel the charm of sports, and experience the fun of sports so as to participate in sports exercises and improve their physical quality.

5. Conclusion

Physical education is an important part of school teaching. How to make computer and network technology serve the daily physical education teaching in schools is an inevitable requirement under today's big data. Based on RF algorithm, this paper puts forward a scientific evaluation algorithm for students' sports achievements. Through this research, combining pattern recognition and big data analysis methods, the optimization of sports effect evaluation is realized. The research shows that this method has a high confidence level in sports effect evaluation, and the evaluation results are accurate and reliable. It can correct and intervene students' learning in advance so that students can achieve better results in professional courses and play a very good scientific guiding role in assisting students' learning behaviors and optimizing educational decisions.

This study only considers the relationship between participation in physical exercise and academic performance from the perspective of a school. As the author's ability is limited, this study does not take into account factors such as school differences, which is also something that needs to be explored in future studies.

Data Availability

The labeled data set used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

References

[1] F. Yang and F. W. Li, "Study on student performance estimation, student progress analysis, and student potential prediction based on data mining," *Computers & Education*, vol. 123, pp. 97–108, 2018.

[2] W. S. A. A. Attar, N. Soomro, P. J. Sinclair, E. Pappas, and R. H. Sanders, "P-7 Effect of injury prevention programs that include the nordic hamstring exercise on hamstring injury rates in soccer players: a systematic review with meta-analysis," *British Journal of Sports Medicine*, vol. 50, no. Suppl 1, pp. A35.2–A35, 2016.

[3] Z. Wu and C. Zhou, "Construction of an ielligent processing platform for equestrian event information based on data fusion and data mining," *Journal of Sensors*, vol. 2021, no. 6, 9 pages, Article ID 1869281, 2021.

[4] C. Yuan and Y. L. YYang, "Sports decision-making model based on data mining and neural network," *Neural Computing & Applications*, no. 2, pp. 1–14, 2020.

[5] D. Yi and G. Fan, "Sports video mining via multichannel segmental hidden markov models," *IEEE Transactions on Multimedia*, vol. 11, no. 7, pp. 1301–1309, 2009.

[6] J. Q. Ouyang and R. Liu, "Ontology reasoning scheme for constructing meaningful sports video summarisation," *IET Image Processing*, vol. 7, no. 4, pp. 324–334, 2013.

[7] B. Bideau, F. Multon, R. Kulpa et al., "Virtual reality applied to sports: do handball goalkeepers react realistically to simulated synthetic opponents?" *Journal of Multivalued Logic and Soft Computing*, vol. 18, no. 5-6, pp. 228–233, 2012.

[8] Y. Li, J. Ge, and W. Hao, "Construction and simulation of a multiattribute training data mining model for basketball players based on big data," *Wireless Communications and Mobile Computing*, vol. 2021, no. 7, 14 pages, Article ID 6399266, 2021.

[9] Z. Yin and W. Cui, "Outlier data mining model for sports data analysis," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, pp. 1–10, 2020.

[10] V. Sarlis and C. Tjortjis, *Information Systems*, vol. 93, Article ID 101562, 2020.

[11] S. Harvey, O. Atkinson, and B. P. Hyndman, "Chapter 7: an investigation into sports coaches' twitter use," *Journal of Teaching in Physical Education*, vol. 39, no. 4, pp. 481–490, 2020.

[12] N. K. Poona, A. Van Niekerk, R. L. Nadel, and R Ismail, "Random forest (RF) wrappers for waveband selection and classification of hyperspectral data," *Applied Spectroscopy*, vol. 70, no. 2, pp. 322–333, 2016.

[13] F. Provost, C. Hibert, and J. P Malet, "Automatic classification of endogenous landslide seismicity using the Random Forest supervised classifier," *Geophysical Research Letters*, vol. 44, no. 1, pp. 113–120, 2017.

[14] K. Lee, M. Lee, and D. Kim, "Utilizing random Forest QSAR models with optimized parameters for target identification and its application to target-fishing server," *BMC Bioinformatics*, vol. 18, no. S16, p. 567, 2017.

[15] S. A. Naghibi, K. Ahmadi, and A. Daneshi, "Application of support vector machine, random forest, and genetic algorithm optimized random forest models in groundwater potential mapping," *Water Resources Management*, vol. 31, no. 9, pp. 2761–2775, 2017.

[16] D. Ming, T. Zhou, M. Wang, and T. Tan, "Land cover classification using random forest with genetic algorithm-based parameter optimization," *Journal of Applied Remote Sensing*, vol. 10, no. 3, Article ID 035021, 2016.

[17] A. Ghorbanian, S. Zaghian, R. M. Asiyabi, M. Amani, A. Mohammadzadeh, and S. Jamali, “Mangrove ecosystem mapping using sentinel-1 and sentinel-2 satellite images and random forest algorithm in google earth engine,” *Remote Sensing*, vol. 13, no. 13, p. 2565, 2021.
[18] Z. Bei, Z. Yu, N. Luo, C. Jiang, C. Xu, and S. Feng, “Configuring in-memory cluster computing using random forest,” *Future Generation Computer Systems*, vol. 79, pp. 1–15, 2018.
[19] C. Raphael, P. Philipp, and B. Anne-Laure, “Random forest versus logistic regression: a large-scale benchmark experiment,” *BMC Bioinformatics*, vol. 19, no. 1, p. 270, 2018.
[20] W. Lin, W. Fan, H. Liu, Y. Xu, and J. Wu, “Classification of handheld laser scanning tree point cloud based on different KNN algorithms and random forest algorithm,” *Forests*, vol. 12, no. 3, p. 292, 2021.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 2966888, 10 pages
https://doi.org/10.1155/2022/2966888

Research Article

Intelligent Decision Making for Constructing Students' Entrepreneurship System in Colleges and Universities

Cui Liu

Xinyang Vocational and Technical College, Xinyang 464000, China

Correspondence should be addressed to Cui Liu; liucui7060@xyvtc.edu.cn

Received 20 June 2022; Revised 19 July 2022; Accepted 28 July 2022; Published 27 August 2022

Academic Editor: Xuefeng Zhang

In order to test the efficiency of colleges and universities to carry out innovation and entrepreneurship education and cultivate the innovative talent system required by economic and social development and change, this paper proposes to combine the fuzzy comprehensive evaluation method based on the analytic hierarchy process. First, gather the students' ideas, then create a set of measures for each measure, count the weight and significance of each measure, and create a set of criteria for the assessment of the educational process of innovation and entrepreneurship. Census of colleges and universities' innovation and business measurement model based on student perceptions and determining the weight of each measurement scale will help students understand what is more important in evaluating college and universities' innovation and business education. Other metrics include innovation and training market, it needs approach, design industry analysis examines business, opportunities for school-to-business engagement, innovation, business education, training, entrepreneurship, and platform information. Therefore, in order to improve the innovation and business education of colleges and universities, it is necessary to carefully establish school management, faculty, classes, activities, innovation, business platform, business education environment, support, advocacy mechanism, and so on. We will improve the teaching of innovation and entrepreneurship and focus on training faculty of innovation and business education and focus on strengthening the innovation and entrepreneurship information platform, innovation, and entrepreneurship incubation base.

1. Introduction

Business education, a unique field of study in human life, aims to develop new memories, new minds, and new abilities for the objects of learning. Its origin and development are worthy of purpose and constant. Looking back at the growth of higher education in China in recent years, people are increasingly paying more attention to business education in the framework of "public utilization and new construction." The new skills have become new practices in vocational training in many colleges and universities [1]. The centralized release of relevant policy documents from the central to local governments is being manifested by vigorous development of entrepreneurial activities, deepening of the accumulation of basic theories of entrepreneurship education, continuous expansion of research methods, and obvious improvement of practical results. It can be said that discipline comprehensive research plays an important role in the theoretical and practical innovation and development of entrepreneurship education. In recent years, paying attention to ecology, including relevant thinking and methods of educational ecology and social ecology, is a frontier issue in the research of entrepreneurship education. Figure 1 shows the analysis of the Internet entrepreneurship environment.

2. Literature Review

In response to this research problem, Pratama et al. pointed out in the "construction of entrepreneurship education ecosystem for college students in local colleges and universities" that the lack of systematic construction of a perfect entrepreneurship ecosystem is one of the fundamental reasons for the lack of remarkable results in China. The educational and teaching resources and systems of colleges and universities are not perfect, and there is a lack of communication between colleges and universities and society [2]. At present, the important work of entrepreneurship education is to establish an entrepreneurship education ecosystem in colleges and universities and analyze the problems existing in teaching staff, curriculum system construction, and entrepreneurship education support mechanism one by one [3]. Santosoet al. analyzed the construction of college entrepreneurship education ecosystem from internal and external factors. They pointed out that the external factors of the education ecosystem mainly refer to the policy contents proposed by the government and the resources and services provided by social enterprises, while the internal factors mainly refer to the entrepreneurship education micro system established by the school, including the introduction of the strategic policy of national entrepreneurship education, the implementation of entrepreneurship courses, the organization of entrepreneurship activities, the creation of a campus cultural atmosphere and incentive system, the incubation of entrepreneurship projects, and the docking and matching of investment resources inside and outside the school [4]. Ribeiro et al. pointed out in the research on “Introduction and risk analysis of entrepreneurship education ecosystem in Chinese colleges and universities,” although some colleges and universities have begun to change to entrepreneurial universities, most of them are still in the stage of theoretical exploration and lack of practice. In the case of doubts, if we rashly introduce the entrepreneurship education ecosystem, it is difficult to achieve the expected results, which will affect the normal development of teaching and scientific research in colleges and universities to a certain extent [5]. Munaiseche et al. believed that the construction logic of the entrepreneurship education ecosystem is to realize the integration of “entrepreneurship” and “education” and must reflect the two key characteristics of open interconnection and endogenous growth. Entrepreneurship education ecosystem must follow the path of spontaneous evolution, constantly adjust its own boundary, function, and structure and construct a closed-loop evolution of the university entrepreneurship education ecosystem, so as to realize the complete value chain of knowledge production, knowledge diffusion, and value creation [6]. Teller et al. proposed that for 2050, China’s higher entrepreneurship education should focus on building a horizontal cooperation system and a vertical school system. The strategic planning of higher entrepreneurship education should plan the overall and long-term development, innovate the concept of development, break the obstacles of system and mechanism, build an entrepreneurship ecological chain for coordinated development with regions, and form a global entrepreneurship education community with competition and sharing [7]. Based on the investigation and analysis of a university, Al et al. proposed that the internal functions of the ecosystem should be continuously improved, the external ecological environment of the system should be optimized, and the ecosystem of entrepreneurship education in colleges and universities should be constructed [8]. From the perspective of international vision, Tessrasit et al. analyzed three driving modes of college entrepreneurship education ecosystem construction: professional

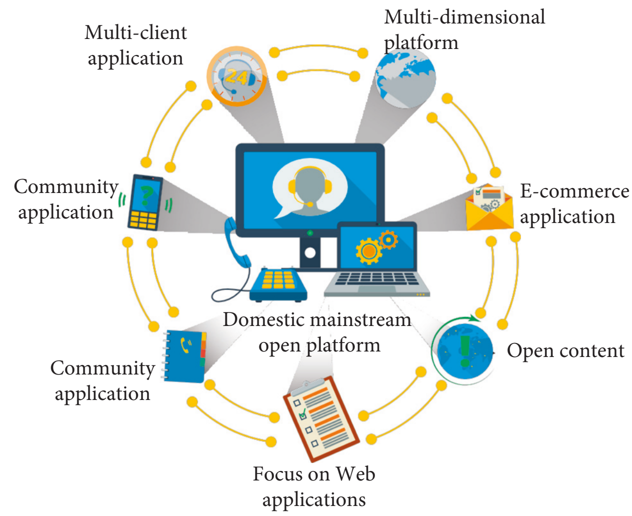


Figure 1: Analysis of the Internet entrepreneurship environment.

college driving mode, special institution driving mode, and whole school system driving mode. They pointed out the main problems existing in China’s entrepreneurship education and proposed that China should make a choice based on China’s national conditions and make a strategic choice of internal cooperation and external connection around factor development, system operation, and organizational innovation, build a college entrepreneurship education ecosystem with Chinese characteristics, and provide a Chinese solution to the problems related to the construction of college entrepreneurship education ecosystem [9]. Permana et al. believed that the environment plays an extremely important role in the construction of an ecosystem. Only a suitable ecological environment of entrepreneurship education can create the most effective environment of entrepreneurship education [10]. Pbedn et al. believed that the construction of entrepreneurship education ecosystem in Chinese colleges and universities is affected by six key factors, such as the vision and promotion of senior leaders of colleges and universities, the investment of teachers and resources, entrepreneurship courses, entrepreneurial activities and practice, organizational structure, and environmental factors [11].

Based on the current research, it is proposed to collect the students’ opinions according to the hierarchical analytical procedures and ambiguous measurement procedures, establish a standardized measurement system at each level and count the weights and values of each measure. It is necessary to develop a set of criteria for measuring research on innovation and market education based on students' perspectives, including models of innovation and market development.

3. Method

3.1. Internet. Due to the special technical characteristics of the Internet and the special operation and management mode of Internet enterprises, Internet entrepreneurship has some differences from entrepreneurship in traditional industries. The new wave of entrepreneurship led by the new wave of the Internet also has its new characteristics. "Internet +" is mainly user-oriented and combined with the latest technology, not only forms an "iron triangle" with entrepreneurship, innovation, and venture capital but also provides technical support and thinking innovation for the renewal of traditional industries [12]. Internet entrepreneurship has the characteristics of strong derivation and has more extensive cooperation space with traditional industries. For example, Vipshop, Taobao, Dangdang, Alipay, and Apple pay are new entrepreneurial innovation models formed by effectively combining traditional department stores, traditional bank counters, traditional cash payment methods, etc. With the "Internet" big data platform, this diversified combination mode of Internet entrepreneurship and traditional business not only shortens the distance between enterprises and users but also speeds up the pace of innovation of enterprises and also enables entrepreneurs to have a direct contact with users and meet various needs of users. This "special function" brought by the "Internet" has greatly promoted the cultivation of entrepreneurs' innovative spirit and the development of China's market economy. Therefore, "Internet +" means to rely on Internet information technology to promote the deep integration of the Internet and traditional industries, to give full play to the advantages of Internet big data, to upgrade the industrial product structure, to improve economic productivity, and finally to increase social wealth.

3.2. Related Concepts

3.2.1. Innovation and Entrepreneurship Education. Innovation and business development are designed to develop students' business skills and old skills, such as market knowledge, have a desire to innovate, and the ability to innovate. At the same time, every organization is faced with the task of new contracts and one-step entrepreneurship skills. The concept of innovation and business education is provided by the concept of business education. Innovation-oriented and business education, as the goal of business education, is the essence of education [13, 14]. In this study, the concept of innovation and business education in colleges and universities is as important as in colleges and universities.

3.2.2. Evaluation Index System. Indicators refer to the units used to measure objectives and the standards expected to be achieved. The index in educational evaluation refers to a standard that is easy to observe and measure, which is used to determine the evaluation content in the process of educational evaluation. At the same time, the index is also a standard for the evaluation content. The index in educational evaluation is a standard to evaluate the specific parts of teaching content in the field of education. Index system refers to the reasonable collection of many single index systems. The index system of educational evaluation is a complete set of several specific evaluation indexes under the premise of educational objectives and according to a specific evaluation dimension.

3.3. Selection Principles of Evaluation Indicators of Innovation and Entrepreneurship Education in Colleges and Universities. Once we have decided on the terms and standards that guide the selection of colleges and universities, new developments, and market assessment standards, we need to define the terms for college and university selection, innovation, and market assessment. In accordance with the principles of educational evaluation, studies on the principles of business evaluation in colleges and universities have been compiled and the principle of selecting the tests for this study has been determined.

Educational evaluation is a work with strong truth and value. In its operation, it should not only comply with the occurrence and development law of the evaluation object but also consider the interest requirements of the evaluated object and give appropriate direction guidance [15]. Therefore, a series of scientific principles must be followed in the evaluation to make the evaluation true and credible, which can not only objectively reveal the problems but also motivate the evaluated to make progress and promote the development of various types of work.

3.3.1. Principle of Objectivity. When carrying out educational evaluation, we should seek truth from facts, not make subjective assumptions, and not judge the length by personal likes and dislikes. Evaluation must be carried out according to scientific evaluation standards. Once the standards are determined, they should not be changed at will. If personal feelings are mixed into the evaluation, it is bound to frustrate the enthusiasm of the judged.

3.3.2. Principle of Comprehensiveness. In the evaluation, we should comprehensively and appropriately evaluate various projects or specific indicators in the educational work. We should not exaggerate some projects or indicators and ignore others, which is easy to make the work lose balance and one-sided. One-sidedness is incompatible with the policy of cultivating people with all-round development. It is necessary to give different degrees of weight to projects with different properties and functions. It is wrong to completely ignore some aspects or pay insufficient attention to the project with more attention.

3.3.3. The Principle of Combining Consistency and Difference. The principle of consistency refers to the adoption of consistent standards for all faculty and students in educational evaluation, rather than two or more sets of standards [16, 17]. Only by following the principle of consistency and respect can we distinguish good from bad and good in the group and make the evaluated know their actual situation and position in the group. Only in this way can we find out the gap, find the deficiency, and motivate people to make progress.

Due to different school conditions, different leadership levels, and different strengths and weaknesses of teachers and students, in the evaluation, we should start from the actual situation of the evaluated person, look for the differences within the individual and judge whether they have made progress or regressed compared with the past. This can enable the evaluated object to find out how much their subjective efforts are and what aspects they need to redouble their efforts compared with their past. Paying attention to the different evaluation objects can spur and motivate people.

3.3.4. The Principle of Combining Quantitative Analysis with Qualitative Analysis. Qualitative analysis is to judge the status and nature of the evaluated object. This analysis method usually first describes the performance of the evaluated object according to certain standards and then analyzes its characteristics, problems, essence, development trend, and so on. This analysis method is characterized by strong comprehensiveness and fuzziness and pays attention to the focus and essence of things [12, 18]. Compared with qualitative analysis methods, quantitative analysis is bright, simple, and easy to compare. Since the 1970s, it has been applied in some quantifiable fields of educational evaluation. However, the quantitative method itself has some shortcomings worthy of attention: firstly, when quantifying the evaluation elements, we must first accurately distinguish them and make them independent in order to correctly measure them. In this way, when we evaluate something, we must subjectively separate its connection with other things in the observation stage, which often leads to a simplified and superficial analysis of the problem and cannot fundamentally look at the problem; secondly, when quantifying the evaluation elements, we should make them static and regard them as a certain point at a certain time to measure accurately. Thirdly, many social factors and human psychological factors cannot or are not easy to be quantified. According to these characteristics, the two should be organically combined in the evaluation, and the lack of any aspect is not enough to fully grasp the actual situation of the evaluated object.

3.3.5. The Principle of Combining Others' Evaluation with Self-Evaluation. This principle has two meanings: first, the evaluation should not only be attended by leaders, peers, classmates, or experts but also by the evaluated himself. The task of modern education evaluation is not only to give the evaluation a conclusion but also to improve their work through evaluation. Without the enthusiasm of the evaluated, the evaluation will lose its significance. At the same time, the self-awareness and self-evaluation of the evaluated are an indispensable factor to ensure a comprehensive and fair evaluation. Without the information and views they provide on their work and performance, the evaluation will lose one aspect of the basis, and its objectivity and impartiality are difficult to guarantee [19].

It should be noted that since the evaluated person participates in the evaluation activities as a subject, he should be a subject with full rights. He should not only make self-evaluation but also have the right to defend the final conclusion of the evaluation. This is not only to respect the evaluated person but also to further carry out the communication and guidance between the manager and the managed person through the reply and defense, so as to better improve the work.

Second, the evaluation of educational work should not only be carried out by the education system itself but also by people or institutions outside the education system. As the evaluation subject, education system personnel have the advantages of understanding the law of education, but they may also have the disadvantages of self-protection. Moreover, people outside the education system can receive the effect of "target free mode" and broaden their thinking.

The above principles are only the elaboration of the general principles of educational evaluation, taking into account the general situation in educational evaluation [20]. When the evaluated object and its activities have particularity, the principles of educational evaluation should be appropriately supplemented or refined according to the specific situation.

Based on the research purpose of this paper, in the general principles of education evaluation, combined with the characteristics of the personality principles of innovation and entrepreneurship education evaluation in Colleges and universities, this paper establishes the following principles for the selection of indicators:

First is the principle of comprehensiveness. Previously, there was little research perspective from the perspective of students in literature in this field. Therefore, in the selection of indicators, we should comprehensively consider all dimensions and try to cover the indicators from all angles related to students, so as to ensure the integrity of the selection of evaluation indicators.

Second is the principle of scientificity and hierarchy. Students' evaluation of education is closely related to the school, teaching process, curriculum, teachers, hardware and software environment construction, and other aspects. As mentioned in the previous literature review, the sub-indicators of teaching, teachers and curriculum permeate each other [21, 22].

Third is the principles of development and dynamics. Innovation and entrepreneurship education in colleges and universities itself is a dynamic and developing process. Students and education implementers interact and influence each other in the education implementation environment. Therefore, the selection of student evaluation indicators should not only be fixed, rigid, and one-way but should also be two-way flow and circular [23].

3.4. AHP Weight Calculation Method Combined with Fuzzy Comprehensive Evaluation

3.4.1. Introduction to Analytic Hierarchy Process. Analytical hierarchical process (AHP) is a method of determining the weight of hierarchical processes based on network considerations and various types of measurement methods. This process breaks down the concepts that affect the entire decision-making process into several stages, such as goals and procedures and are interrelated decision-making processes. Well combined and multiple reviews: Analytical hierarchical procedures are necessary for determining with hierarchical and hierarchical evaluation indices, and it is difficult to quantify the target value. The main steps are as follows [24].

To build a hierarchical structure, first list all index items requiring weight, determine the overall goal, criterion layer, and index layer of the index system and then establish a multi-level independent and orderly hierarchical model. Create a conclusion matrix, divide the data obtained by experts according to 9 points of measurement at the same level and write it into the end matrix. Calculate the weight of each measure, open the product of each element of the line of the filter matrix by n degrees (n is the number of indices in a layer) and then normalize the result and finally get the weight vector. Identify the consistency of the filter matrix. An appropriate test is one in which the relative matrix is similar. If the value of parameter a is higher than parameter B and the value of parameter B is greater than parameter C, then the parameter must be greater than parameter C. If the test result is yes, it has logical consistency; otherwise, the internal consistency of the matrix is inconsistent.

3.4.2. Combined with the Calculation Flow of Fuzzy Comprehensive Evaluation Method. In general, the analytical hierarchical process (AHP) is a way in which the decision maker (usually an expert in this field) receives the benefits of the decision from the competition, paired according to a hierarchical model.

This paper examines the measurement process of measuring innovation and the business education system generally based on the students' perspectives and measures the standards for measuring innovation and business ventures in colleges and universities from a student perspective.

4. Results and Analysis

4.1. Questionnaire Design and Data Recovery

4.1.1. Design and Distribution of Evaluation Index Questionnaire. In the hierarchical model of the College and University Innovation and Entrepreneurship Education Evaluation Index system, the importance of one measure is measured at the next level by student thinking using a nine-step approach. Suppose there are n parameters at a level.

$$\{X_1, X_2, X_3, X_4, \ldots, X_n\}. \tag{1}$$

In the questionnaire design, we need to compare X_i and X_j in pairs, and then refer to a satty's 1 ~ 9 main scale method. As shown in Table 1, the ratio and meaning of importance between each index are shown. For example, if indicator X_i is considered more important than indicator X_j, the importance scale is 7. If indicator X_j is considered more important than indicator X_i, the scale is 1/7.

After excluding all invalid questionnaires with inconsistent answers and incomplete options, 274 valid questionnaires were recovered, and the effective recovery rate was 83.6%. The regional distribution characteristics of the survey samples are shown in Table 2.

4.1.2. Analysis of Importance Indicators of Questionnaire. An important principle is to ensure the validity of the questionnaire by examining the accuracy of each question included in the questionnaire. In general, the reliability of the questionnaire is very good if the Cronbach's alpha coefficient is greater than 0.9. Coefficient of 0.8 ~ 0.9 indicates reliability. If it is less than 0.8, the reliability is not very good. The examiner should revise the questionnaire and repeat the assessment. The results of the Cronbach's alpha coefficient analysis of this questionnaire are shown in Table 3.

As can be seen from Table 3 above, the Cronbachsa coefficient of reliability measure is greater than 0.96 or 0.8, indicating that the reliability of the data in this questionnaire is very high. The fact that the reliability of the damaged material has not been improved indicates that everything can be stored, and the data from this research questionnaire show that it is reliable in the future. Next, Although the importance ratio represented by the scale of 1 and 1/9 is very large, the difference in value is 8/9, less than 1. Therefore, it is inaccurate to use the general standard deviation to represent the dispersion of importance scale data. Therefore, in this paper,, the 1 ~ 9 scale method is not used to calculate the dispersion degree but the equal difference of 1/9, 1/8, 1/7, 1/6,... 1/2, 1, 2, 3, ..., 8 and 9 is assigned as 1 ~ 17, as shown in comparison Table 4.

To sum up, the reliability and dispersion of the survey and research data are good. There is no need to delete or adjust the index items, and the index weight can be calculated in the next step.

4.2. Comprehensive Group Opinions. This study uses the fuzzy comprehensive evaluation method to synthesize the opinions of the student group. The steps are as follows: Taking the comparison items between the two secondary indicators under the primary index "innovation and entrepreneurship practice teaching" as an example, this paper constructs the student group scale vector under the index of "innovation and entrepreneurship practice teaching".

The membership degree ofeachquestionnaire item scale is shown in Tables 5–7.

The final group decision scale is calculated.

By multiplying the membership matrix D and the scale matrix R, the final decision scale vector of three pairwise contrast terms can be obtained as

$$W\colon\ (10.9088, 6.9715, 6.2673). \tag{2}$$

Then, according to the comparison Table 4, adjust its scale to the 9-degree scale method, then the 9-degree scale of the relative importance of d1 and d2 is 3; that is, the student group believes that d1 is slightly more important than d2; the scale of the 9-degree method for the relative importance of d1 and d3 is 1/3, which means that the students think d3 index is slightly more important than d1 index; the 9-degree scale of the relative importance of d2 and d3 is 1/4, which means that the student group believes that d3 index is significantly more important than d2 index.

Table 1: Main scale method.

Relative importance	Meaning	Explain
1	Equally important	Both contribute equally to the goal
3	Slightly important	According to experience, i is slightly more powerful than j
5	More important obviously important	According to experience, i is more powerful than j
7	Very important	i is more powerful than j, and its advantages have been proved in practice
8	Absolutely important	The degree of importance can be asserted as the highest
2, 4, 6, 8	Intermediate value of two adjacent degrees	Use when compromise is required

Table 2: Characteristics of student sample.

Region	Number of people	Proportion
Province A	78	23.79
Province B	75	22.87
Province C	63	19.21
Province D	58	17.69
Province E	26	7.94
Province F	19	5.78
Province G	9	2.75

Table 3: Total reliability statistics.

Cronbach's alpha	Number of items
0.951	77

Table 4: Comparison Table of re equal difference assignment.

9-Degree scale	1/9	1/8	1/7	...	1/3	1/2	1	2	3	...	7	8	9
Reassign	1	2	3		7	8	9	10	11		15	16	17

4.3. Weight Determination

4.3.1. Build Judgment Matrix and Calculate Weight Vector. Construction of judgment matrix: after the summary of students' opinions, the judgment matrix can be constructed. Still take the secondary index under the primary index "innovation and entrepreneurship practice teaching" as an example. There are three secondary indicators under innovation and entrepreneurship practice teaching D: d1 innovation and entrepreneurship competition, d2 innovation and entrepreneurship lecture; and d3 innovation and entrepreneurship training practice. According to the scale data obtained from the questionnaire, the judgment matrix D is obtained as follows:

$$\begin{aligned} D &= 1, 3, \frac{1}{3}, \\ D &= \frac{1}{3}, 1, \frac{1}{4}, \\ D &= 3, 4, 1. \end{aligned} \tag{3}$$

Calculate eigenvector: calculate the mean value of each line - eigenvector W is calculated:

$$w_i = \frac{1}{n}\sum_{j=1}^{n}\overline{a}_{ij}\,(1, 2, 3, \ldots, n). \tag{4}$$

4.3.2. Consistency Test. Consistency test room judges whether the matrix has high or low consistency. Check the consistency by calculating the consistency ratio CR (consistency ratio). It is generally believed that when the required consistency ratio CR is less than 0.1, it meets the consistency inspection standard. It shows that the judgment matrix constructed by it has high consistency. It is meaningful when used in subsequent research, otherwise the judgment matrix constructed is not ideal and meaningless. The steps to calculate the consistency ratio CR are as follows:

Calculate the maximum characteristic root as

$$\lambda_{\max} = \sum_{i=1}^{n}(AW)_t/nW_i. \tag{5}$$

Find consistency index CI (Consistency Index) as

$$CI = \frac{(\lambda_{\max} - n)}{(n-1)}. \tag{6}$$

The average random consistency index RI is found in Table 8.

Consistency ratio CR(Consistency Ratio):

$$\begin{aligned} CR &= \frac{CI}{RI}, \\ CR &= \frac{(\lambda_{\max} - n)/(n-1)}{RI}. \end{aligned} \tag{7}$$

YAAHP software is the auxiliary software of analytic hierarchy process, which provides the functions of model construction, data entry, calculation and analysis. Due to the large amount of weight calculation of the judgment matrix, this study uses YAAHP software to input and calculate the judgment matrix. The consistency CR of the judgment matrix D of the secondary index is $0.0707 < 0.1$. It is proved that the consistency of D matrix is good. Similarly, the judgment matrix data of the remaining indicators at the same level are entered, and all judgment matrices pass the consistency test after calculation.

4.3.3. Determination of Index Weight. Through the calculation of matrix D by YAAHP software, the weight vector W_i

Table 5: Membership degree of each questionnaire item scale (1).

	Isochronous important scale								
	1	2	3	4	5	6	7	8	9
Comparison between “d1 innovation and entrepreneurship competition” and “d2 innovation and entrepreneurship lecture”	0	0.37	0.37	0.73	1.105	0.73	2.92	4.007	5.08
	10	11	12	13	14	15	16	17	
	19.35	28.47	20.805	8.39	4.465	1.825	1.09	0.37	

Table 6: Membership degree of each questionnaire item scale (2).

	Isochronous important scale								
	1	2	3	4	5	6	7	8	9
Comparison between “d1 innovation and entrepreneurship competition” and “d3 innovation and entrepreneurship lecture”	0.37	0.37	2.56	7.28	12.077	20.45	24.46	13.135	8.76
	10	11	12	13	14	15	16	17	
	4.02	2.56	1.47	1.465	0.37	0.37	0	0.37	

Table 7: Membership degree of each questionnaire item scale (3).

	Isochronous important scale								
	1	2	3	4	5	6	7	8	9
Comparison between “d2 innovation and entrepreneurship competition” and “d3 innovation and entrepreneurship lecture”	0.73	3.29	3.648	11.68	19.4	22.27	14.58	10.23	6.3
	10	11	12	13	14	15	16	17	
	1.47	2.56	0.37	1.08	1.825	0.37	0	0	

Table 8: Average random consistency index RI.

n	1	2	3	4	5	6	7	8	9
RI	0	0	0.53	0.88	1.13	1.27	1.37	1.42	1.47

Table 9: Weight of secondary indicators under Innovation and entrepreneurship practice teaching.

Practical teaching of innovation and entrepreneurship	Innovation and entrepreneurship competition	Lecture on innovation and entrepreneurship	Innovation and entrepreneurship training practice	W_i
Innovation and entrepreneurship competition	1	3	1/3	0.2685
Lecture on innovation and entrepreneurship	1/3	1	1/4	0.1173
Innovation and entrepreneurship training practice	3	4	1	0.6145

Table 10: Weight of secondary indicators at the school level.

School level	School running idea	Management system	Policy implementation	Financial support	Scientific research	W_i
School running idea	1	1	1/3	1/3	1/4	0.748
Management system	1	1	1/3	1/3	1/4	0.0748
Policy implementation	3	3	1	1/3	1/4	0.1508
Financial support	3	3	3	1	1/3	0.247
Scientific research	4	4	4	3	1	0.4535

of the index of this layer can be obtained. The data are shown in Table 9 below.

Similarly, the weights between theremaining indicators ofthe same layercan be obtained. See Table 10–13.

4.4. Weight Synthesis, Sorting and Determination. After calculating each index at the same level of the index judgment matrix, it needs to repeat the above steps to calculate the weight coefficient of the secondary index corresponding

Table 11: Weight of secondary indicators under teachers of innovation and entrepreneurship 2ducation.

Innovation and entrepreneurship education teachers	Teacher qualification	Team structure	Teaching ability	Scientific research ability	Teaching enthusiasm	Training	W_i
Teacher qualification	1	1/4	1/3	1/3	1/3	1/2	0.0628
Team structure	4	1	1/3	2	1	1/3	0.1447
Teaching ability	3	3	1	1	1	1	0.215
Scientific research ability	3	1/2	1	1	1/3	1/3	0.1167
Teaching enthusiasm	3	1	1	2	1	1/4	0.1529
Training	2	3	1	3	4	1	0.3093

Table 12: Weight of secondary indicators under incentive mechanism.

Excitation mechanism	Credit setting	Innovation and entrepreneurship award	School enterprise cooperation job opportunities	W_i
Credit setting	1	1	1/3	0.1918
Innovation and entrepreneurship award	1	1	1/4	0.1745
School enterprise cooperation job opportunities	3	4	1	0.6338

Table 13: Weight of secondary indicators under feedback mechanism.

Feedback mechanism	Teaching situation tracking	Feedback demand channel	Evaluation mechanism	W_i
Teaching situation tracking	1	2	3	0.5498
Feedback demand channel	1/2	1	1	0.2403
Evaluation mechanism	1/3	1	1	0.2099

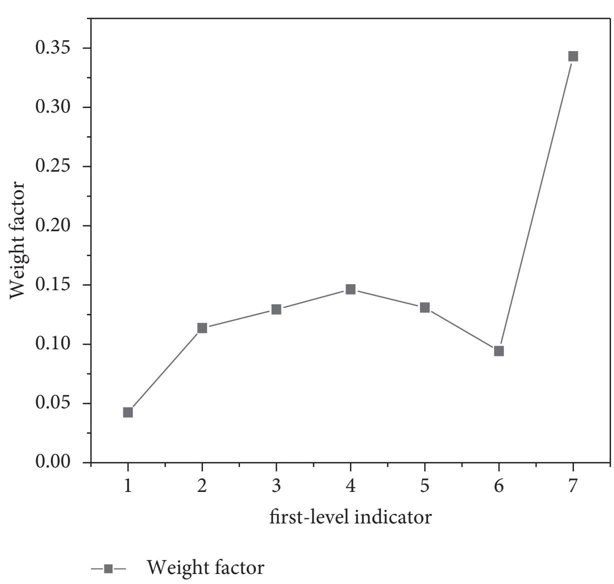


Figure 2: Weight coefficient of primary indicators of innovation and entrepreneurship education evaluation in colleges and universities from the perspective of students.

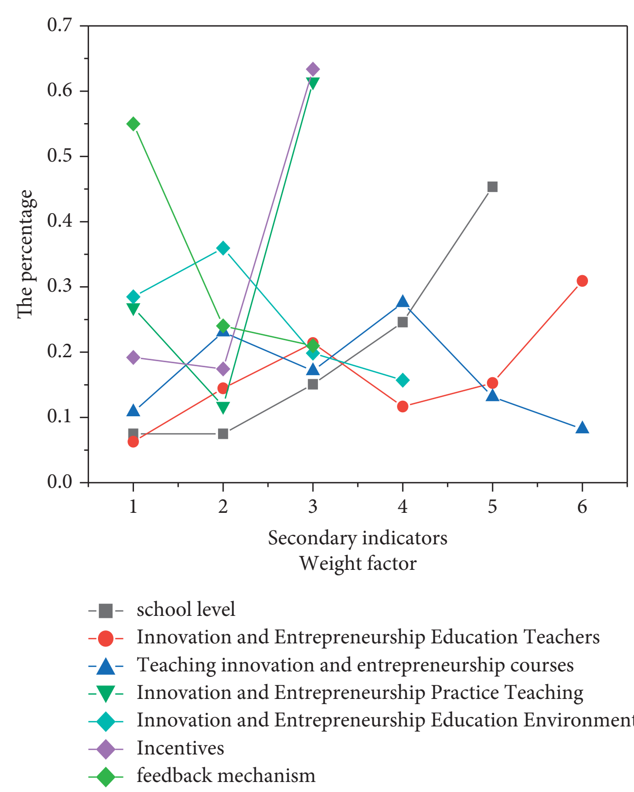


Figure 3: Weight coefficient of secondary indicators of innovation and entrepreneurship education evaluation in colleges and universities from the perspective of students.

to each level index. For example, there are m elements in layer A of primary index, including A1, $A2,\ldots, Am$, and there are n factors in layer B of secondary index. Let $B_{1i}, B_{2j},\ldots, B_{nj}$ be the weight of layer B to layer A, then the weight

$B_i(i = 1,2,3, ..., n)$ of each factor of layer B to layer is shown in formula (8). The synthetic weight can be obtained according to the formula, that is, the final weight for the target layer index. Finally, the synthesis weight of the index system is obtained, as shown in Figures 2 and 3.

$$b_i = \sum_{j=1}^{m} b_{ij} a_j. \quad (8)$$

5. Conclusion

This article discusses how to create a business partnership for college and university students and how to create a benchmark for students. The system includes innovative and business assessment courses, learning from them, integrating with student assessment features and developing new standards and assessment practices. Marketed by students' perspectives, college and university students are the main objects of research, using assessment models designed to question, research, and evaluate the reliability of metrics.. This section examines the importance of innovation, innovation in Business intelligence research, and implications for business education using data research and analysis. It is unfortunate that the measurement system developed in this document is not used in the actual assessment due to the time dependence after the definition of the index system of the analysis of innovation and market education. It is hoped that the benchmark study will continue into practice and further improvement in future research.

Data Availability

The dataset used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest in relation to this article.

Acknowledgments

This study did not receive any funding in any form.

References

[1] X. Na, Z. Weng, Z. Zheng, J. Wang, D. Han, and Y. Wei, "Exploration and practice of student-centered "ternary" talent cultivation mode," *Creative Education*, vol. 12, no. 9, p. 8, 2021.

[2] G. Pratama, C. Nagara, and Amrullah, "The effectiveness of entrepreneurship building development in Wates campus, Yogyakarta state university," *IOP Conference Series: Earth and Environmental Science*, vol. 832, no. 1, Article ID 012029, 2021.

[3] S. J. Beetson, S. Pradhan, G. Gordon, and J. Ford, "Building a digital entrepreneurial platform through local community activity and digital skills with ngemba first nation, Australia," *International Indigenous Policy Journal*, vol. 11, no. 1, pp. 1–19, 2020.

[4] N. Towers, A. S. Santoso, N. Sulkowski, and J. Jameson, "Entrepreneurial capacity-building in HEIs for embedding entrepreneurship and enterprise creation – a tripartite approach," *International Journal of Retail & Distribution Management*, vol. 48, no. 8, pp. 881–899, 2020.

[5] A. Ribeiro, C. A. Ferragi, F. T. Strixino, and A. Cardoso, "Entrepreneurship education going remote: a response to COVID-19 restrictions," *Journal of Entrepreneurship Education*, vol. 23, p. 1, 2020.

[6] M. Munaiseche, T. Saroinsong, H. Rumambi, M. Kondoj, J. Rangkang, and A. Kumaat, "The concept of digital entrepreneurial ecosystem for smes in north sulawesi, Indonesia," *American Journal of Industrial and Business Management*, vol. 12, no. 03, pp. 405–419, 2022.

[7] T. Cruz Teller, "Collective resilience: building wellness in an entrepreneurship programme," *AI Practitioner*, vol. 23, no. 1, pp. 16–20, 2021.

[8] H. H. Etal, "Building entrepreneurship behavioural intention: an empirical examination of the areas upon which to focus," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 3, pp. 2526–2536, 2021.

[9] K. Tesprasit, P. Aksharanandana, and A. Kanchanavibhu, "Building green entrepreneurship: a journey of environmental awareness to green entrepreneurs in Thailand," *Journal of Information Technology Applications and Management*, vol. 27, no. 5, pp. 35–47, 2020.

[10] A. Y. Permana, R. D. Srihartati, J. Akbardin, A. Setiawan, D. Jatnika, and W. Rustandi, "Technopreneur training model in building an entrepreneurship," *Journal of Architectural Research and Education*, vol. 2, no. 2, pp. 190–197, 2020.

[11] P. Bednář and L. Danko, "Coworking spaces as a driver of the post-fordist city: a tool for building a creative ecosystem," *European Spatial Research and Policy*, vol. 27, no. 1, pp. 105–125, 2020.

[12] A. Pratama, S. Mukaromah, S. A. Ithriah, and E. M. Safitri, "Entrepreneurship information system design with iconix process for a student business unit marketplace," *Journal of Physics: Conference Series*, vol. 1569, no. 2, Article ID 022041, 2020.

[13] H. Sutton, "Building a student success network helps keep adult learners on track," *Recruiting & Retaining Adult Learners*, vol. 22, no. 8, p. 12, 2020.

[14] M. Zheng, N. M. Ruzgar, N. R. Angoff, and D. I. Rosenthal, "Tips for building a community for incoming medical students virtually in the COVID-19 era," *Medical Science Educator*, vol. 31, no. 6, pp. 2033–2040, 2021.

[15] N. Djam'An, Bernard, and Sahid, "Developing students' creativity in building city mathematics through project based learning," *Journal of Physics: Conference Series*, vol. 1899, no. 1, Article ID 012147, 2021.

[16] Febrianty, H. Hadiwijaya, and D. T. Octafian, "System development management e-school as a students information media," *Journal of Physics: Conference Series*, vol. 1783, no. 1, Article ID 012026, 2021.

[17] M. M. Saeed and A. Fahim, "Designing a smart system for building student's schedule-an application at Neelain university," *International Journal of ADVANCED AND APPLIED SCIENCES*, vol. 8, no. 1, pp. 82–94, 2021.

[18] Q. Zhou, "Construction of industry-university-research-application system under the background of mass entrepreneurship and innovation education," *OALib*, vol. 07, no. 10, pp. 1–11, 2020.

[19] L. Zhong, "Research on the practice information system of university students' innovation and entrepreneurship based

on the new computer technology support," *Journal of Physics: Conference Series*, vol. 1744, no. 3, Article ID 032151, 2021.

[20] J. L. Ding and B. Shi, "Analysis and modeling of enterprise competitive intelligence based on social media user comments," *Entrepreneurship Research Journal*, vol. 11, no. 2, pp. 47–69, 2021.

[21] B. Gavurova, J. Belas, M. Cepel, and I. Kmecova, "Perception of the quality of educational system for entrepreneurship -comparative analysis," *Acta Polytechnica Hungarica*, vol. 18, no. 3, pp. 65–86, 2021.

[22] D. Dubin and W. Xiaoxia, "Human-computer system design of entrepreneurship education based on artificial intelligence and image feature retrieval," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 4, pp. 5927–5939, 2020.

[23] X. Hu, S. Marlow, A. Zimmermann, L. Martin, and R. Frank, "Understanding opportunities in social entrepreneurship: a critical realist abstraction," *Entrepreneurship: Theory and Practice*, vol. 44, no. 5, pp. 1032–1056, 2020.

[24] L. Mao and W. Zhang, "Analysis of entrepreneurship education in colleges and based on improved decision tree algorithm and fuzzy mathematics," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, pp. 2095–2107, 2021.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 3090115, 10 pages
https://doi.org/10.1155/2022/3090115

Research Article

English Long Sentence Segmentation and Translation Optimization of Professional Literature Based on Hierarchical Network of Concepts

Di Shao and Rui Ma

College of Foreign Languages, North China Institute of Aerospace Engineering, Langfang, Hebei 065000, China

Correspondence should be addressed to Di Shao; shaodi6666@126.com

Received 7 June 2022; Revised 4 July 2022; Accepted 17 July 2022; Published 25 August 2022

Academic Editor: Jiafu Su

The structure of English long sentences is generally complicated, with complicated logical levels, many parallel elements, many modifiers and conjunctions, and long post attributives. Moreover, pronouns in English long sentences often need to be judged according to the context. Complex English long sentences are very common in English and even run through the whole English article. The analysis results of these complex long sentences will seriously affect the quality and readability of machine translation. In this paper, HNC (hierarchical network of concepts) method is improved to realize the segmentation of long sentences, so as to simplify sentence patterns. According to the characteristics of professional literature, this paper puts forward a translation optimization method of professional literature combining HNC with statistics, which greatly improves the recognition efficiency of unknown words by extracting professional terms. The results show that the word segmentation method based on HNC and statistics proposed in this paper has achieved a good word segmentation effect in the open test environment, with an accuracy rate of 93.38% and a recall rate of 94.51%. The conclusion shows that our method can make full use of the knowledge of the source tree database, thereby improving the accuracy of the syntactic model on the target tree database.

1. Introduction

In recent years, the domestic English translation business has been developing constantly, from the translation of English works in the past to the diversified English translation of English movies, TV dramas, news, etc. As a whole, great progress has been made. Machine translation has become an effective way to solve English translation. However, due to the complexity of language knowledge and the limited cognition of language rules, there is a big deviation between the accuracy of machine translation and language artistic conception [1, 2]. At present, the machine translation algorithms of English complex long sentences are based on syntactic analysis, multistrategy analysis, and corpus translation, which mainly focus on word meaning checking and semantic feature processing [3, 4], but these machine translation algorithms have low translation accuracy, high recovery rate, and poor reliability of translation results. Therefore, this paper studies the segmentation of English long sentences based on HNC (hierarchical network of concepts) and the optimization of professional literature translation, so as to realize the automatic proofreading of English translation and improve the efficiency and quality of English translation.

Professional literature translation involves multidisciplinary fields. Theoretically, it is appropriate that professional literature should be translated by professionals. However, the reality is that most professional people may not be proficient in foreign languages or they know foreign languages but lack basic translation skills, so it is difficult for them to be competent translators. In English translation, how to solve the problem of long sentence translation is a key point worthy of attention. From the perspective of English language structure, a long sentence means that the components in a sentence are complex; for example, it may consist of several clauses or cover many other modifiers. Duan et al. put forward the concept of document processing automation based on word frequency statistics. It establishes a theoretical basis for selecting important words and abstract sentences [5]. Giannella et al. provide information of different levels of knowledge according to users' personal interests and try to shield this information that users do not like. HNC reasoning becomes the key technology of knowledge organization and knowledge query in intelligent information retrieval [6]. Matsuuchi et al. modeled the problem as finding the dependency tree with the highest probability (score) from a directed multiple graph (complete dependency graph) [7]. Chen et al. proposed a dependency analysis method based on integer linear programming [8]. They used the first-order graph model and added some strong constraints inspired by linguistics. Their model needs the constraint of exponential number, because every possible ring needs to introduce a constraint. Corpus from different fields plays a decisive role in statistical models. Therefore, even if a system has excellent effect and precision of word segmentation in its own domain, it often fails to achieve satisfactory results for word segmentation in other domains.

Complex English long sentences are very common in EST, some of which are even as long as dozens of lines, including hundreds of words, and contain many clauses and nonpredicate verbs. These clauses and phrases are interdependent and have very distinct semantic hierarchical relationships. In HNC, the phrase corresponds to the concept of semantic block. As the next semantic unit of a sentence, semantic block is the function of sentence type, and it plays a very important role in sentence type. The effective segmentation of long sentences can be used in the actual machine translation system, which can simplify the sentence structure and improve the overall performance of the machine translation system. In addition, it can also be used in the fields of natural language processing such as information retrieval and text classification. At present, under the premise that there is little research on word segmentation technology of professional literature and the word segmentation technology of professional literature is immature, the research on translation optimization of professional literature has great practical significance.

2. Related Work

2.1. Research on Machine Translation Technology. With the development of machine learning methods [9], artificial intelligence, computer technology, and linguistic theory, people have a deeper understanding of the background, objectives, and application prospects of machine translation, so machine translation has made great progress again.

Luo et al. believe that the biggest barrier to machine translation is semantic ambiguity. The diversity of language semantics requires that machine translation must have a set of grammar and syntax systems that can fully analyze language structure and semantic problems [10]. However, the conversion rules and syntactic analysis models applied at that time obviously could not solve all kinds of complicated language problems. Motoyuki and Kazuhiro discussed traditional machine translation methods based on dictionaries and conversion rules and example machine translation methods based on parallel corpora [11]. Condon et al. believe that any language problems faced by translation can be attributed to word meaning problems at the lexical level and structure problems at the grammatical level [12]. Wen et al. put forward a diagonal alignment model and an alignment model based on hidden Markov model. The diagonal alignment model can better solve the influence of small-scale corpus on translation results, and the alignment model based on hidden Markov model can make alignment smoother in general [13].

Wei et al. put forward a vocabulary domain labeling method based on the annotation information in dictionaries, which uses the annotation information of words in general dictionaries to label the domain of words, expanding the scale of existing HNC [14]. Wu et al. put forward the principle of machine translation method based on semantic unit theory, which regards the translation between natural languages as the conversion between different expressions of the same semantics in two natural languages. Firstly, semantic analysis is carried out in the source language to obtain sentence meaning expressions, which are then substituted into semantic units of the target language to generate sentences in the target language [15]. Liu et al. have studied the classification system of machine translation specialty, the mapping of specialty dictionaries to specialty classification systems, and the mapping of International Standard Classification ICS standards to specialty classification systems [16].

2.2. Professional Literature Translation Research. Over the years, many scholars at home and abroad have done a lot of research work in the field of term extraction. The methods used are rule-based methods, statistics-based methods, and the combination of statistics and rules. Jiang et al. introduced the concept of mutual information to measure the combination ability of two words and extracted various combinations of words [17]; Wang et al. adopted the method based on pure statistics to automatically extract terms from open corpus and jointly applied two statistical parameters, mutual information and log-likelihood ratio, to the automatic term extraction algorithm, extracting the seeds of two-word words to be expanded from corpus [18]. Geng adopts word-based maximum entropy model and word-based conditional random field model to automatically extract terms and transforms the problem of term extraction into the problem of term labeling, thus realizing an automatic term extraction system based on the combination of word-based conditional random field model and rules. The accuracy of the system is 89.9%, and the f value is 88.4% [19].

Kedar proposed a global training method using linear model for transition-based model, and the training algorithms of graph-based and transition-based methods tend to be consistent [20]. Tubau found through detailed comparison that the error distribution of the two methods is different. Compared with the graph-based method, the transition-based method has high accuracy on the dependency arc with short distance (the distance between the core word and the dependency word) and on the dependency arc with far distance from the root node (from this arc to the root node in reverse) [21]. Pereira proposed a method to reorganize the results of multiple models [22]. The main idea is to establish a new dependency graph according to the analysis results of several models. The weight of each arc in the dependency graph is obtained by voting multiple model results. Ge proposed a fine-grained fusion method [23]. The idea of this method is to fuse the features used in transition-based and graph-based models and train the weights together.

3. Research Method

3.1. Segmentation of English Long Sentences by HNC Algorithm. Due to the extensive use of relative words, conjunctions, prepositions, etc., English long sentences usually contain a large number of coordinate elements, and the relationship between the main clause and the clause, as well as the clause and clause, makes the sentences complicated and difficult to grasp. Therefore, when translating English long sentences, translators often need to adopt the split translation method; that is, the middle participle, phrase, and even clause of the original text are split from the original text and then translated to facilitate the overall arrangement of the sentence. English is a language that emphasizes hypotaxis. It is customary to use conjunctions to show the relationship between sentences and the components in sentences by explicit means, while Chinese pays attention to parataxis, which mainly depends on the logical connection between sentences and components in sentences. No matter what kind of English long sentence it is, these methods and steps are involved in the translator's use of split translation.

Compared with short English sentences, long English sentences involve many subclauses, and the relationships among these subclauses are complicated. However, if we ignore the correlation, it will directly lead to readers' deviation in understanding the whole sentence. Actually, in order to express the meaning more clearly, people will make more use of the mixture of complex sentences and compound sentence. Such sentences are called mixed long sentences. In this sentence structure, sentences will contain more parallel structures, and these parallel structures will also have their own subordinate structures. Therefore, English long sentences should have the following characteristics: long sentences, many words, many sentence components, and many main or clause components. Moreover, the sentence structures are complex and changeable, and some of them have nested structures, which contain more grammatical and semantic information.

The HNC statistical algorithm used in this paper is mainly used to improve the feature selection algorithm, and the core metric of feature selection is the feature weight. If a concept is repeatedly mentioned in the article, it is often closely related to the main object of the article, and it is also of great value for category judgment. Judge whether the analysis results meet the design requirements of symbol system, how much they meet, and the next improvement direction. In addition, the consistency of each knowledge base is tested quantitatively. In this way, the process of machine translation will involve the ordering of English sentences. There are also adverbial clauses with independent grammatical structure, which are mainly located at the beginning or end of a sentence.

In HNC method, this paper mainly analyzes sentences in three steps, and the process is shown in Figure 1.

The specific process is as follows:

(1) The part of speech of sentences is marked, and then the parts of speech are merged according to the regular expression matching rules, so as to achieve the purpose of simplifying sentence components. In this way, the number of "words" in the whole sentence is reduced, and then merging these components will bring convenience to the subsequent processing of sentences.

(2) Analyze compound sentence. In English sentences, a common problem is the word order before and after compound sentence; that is, after English is translated into Chinese, the translation must conform to the logical order of Chinese. For example, in common adverbial clauses of cause, the result in an English sentence is usually at the beginning of the sentence, while the reason is at the end of the sentence.

(3) If there are clauses in the segmented sentence, further segmentation is needed. These clauses can be subject clauses, object clauses or attributive clauses, etc.

Through the above three steps, try to make the sentences more concise and shorter, which will bring more convenience to the subsequent machine translation work.

The higher the level of a concept is, the more abstract it is and the wider the scope it covers. The concept covers too wide a range, and its significance to classification is relatively small. Use a generalization formula to measure:

$$F_R(c) = 1 - \frac{\max_{i=1}^{n}\left(F_t\left(S_i\right)\right)}{\sum_{i=1}^{n} F_t\left(S_i\right)}, \tag{1}$$

where $F_t(S_i)$ represents the number of times that the subconcept itself and all its subordinate subconcepts appear in the article and n represents the total number of subconcepts of this concept c.

In order to be like an artificial topic analyst, the text topic can be determined only according to several concepts that have strong predictive effects on the topic, and the hierarchical order of the text topic can be ensured. Semantic network is used to represent the hierarchical concepts of topics, and the concepts that have strong decisive effect on topics are selected as the central nodes in conceptual reasoning network.

Information gain is a criterion that is often used in machine learning to measure the importance of a certain predictor. First of all, the evaluation function is used to get a

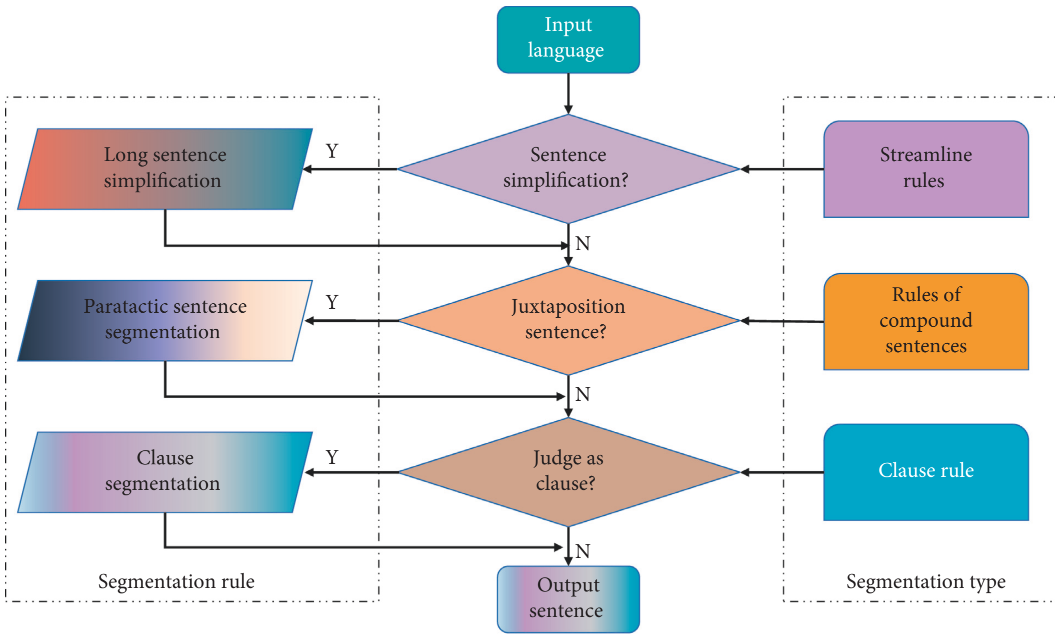


Figure 1: Streamline process based on HNC long sentences.

score for all the predictive words, and then they are sorted according to their size. Finally, according to the size of the threshold, it is decided to retain the predictive words that better reflect the characteristics of the theme. The evaluation function is as follows:

$$\mathrm{InfGrain}(F) = P(F)\sum_i p(w_i|F)\log\frac{p(w_i|F)}{p(w_i)} + P(\overline{F})\sum_i p(\phi_i|\overline{F})\log\frac{p(\phi_i|\overline{F})}{p(\phi_i)}. \quad (2)$$

Here, F is a word, $P(F)$ indicates the probability of the occurrence of the word F, $\overline{F}$ indicates that the word F does not appear, $p(\phi_i)$ indicates the probability of the occurrence of class i, $p(\phi_i|\overline{F})$ indicates the probability of the occurrence of the word F, and ϕ_i indicates the probability of its occurrence.

$\mathrm{InfGrain}(F) > w$ words are the core words, and InfGrain is defined as the subordinate degree of the subject words. The number of core words is limited by obtaining a larger w.

The algorithm is based on the semantic knowledge base of HNC. Each word in the knowledge base gives the knowledge of HNC concept type, cluster code, etc. The type of the word is judged from the concept category, and the cluster code is the logical combination unit of the word matching.

Cosine similarity model is an important model commonly used to measure the word meaning difference between two phrases. It is based on multidimensional space, and the difference between two vectors is represented by the cosine of the angle between two vectors. If the cosine between two phrases is larger, the smaller the angle between two semantic vectors is, the closer the meaning of two phrases is. On the contrary, if the cosine value between two phrases is smaller, the semantic difference between the two phrases will be greater.

Let two phrases in the same corpus be multidimensional semantic vectors u, v, and let $u = [a_1, a_2, \ldots, a_n]$, $v = [b_1, b_2, \ldots, b_n]$, n be the dimensions of the vectors; then the English translation similarity $\mathrm{Sim}(u, v)$ between the two phrases is calculated by the following formula:

$$\mathrm{Sim}(u, v) = \frac{u \cdot v}{\|u\| \times \|v\|} = \frac{\sum_{i=1}^{n} a_i \times b_i}{\sqrt{\sum_{i=1}^{n} a_i^2 \times \sum_{i=1}^{n} b_i^2}}. \quad (3)$$

In order to ensure the efficiency of global search algorithm based on dynamic programming, the graph-based model needs to make strong independent assumptions. It is assumed that, in the dependency tree corresponding to a sentence, only the dependency arcs existing in some special structures are interrelated and influence each other, while other dependency arcs are independent of each other. Under this assumption, the score of a dependent tree is decomposed into the sum of the scores of some subtrees.

$$\mathrm{Score}(x, \mathrm{d}) = w \cdot f(x, \mathrm{d}) = \sum_{p \subseteq \mathrm{d}} \mathrm{Score}_{\mathrm{subtree}}(x, p), \quad (4)$$

where p represents a subtree allowed by an independent hypothesis. p contains one or more dependent arcs in d. $\mathrm{Score}_{\mathrm{subtree}}(x, p)$ represents the score contributed by p. $f(x, \mathrm{d})$ is the aggregated syntactic feature vector corresponding to (x, d), and w is the weight vector of syntactic features.

Through explicit means such as morphological changes, relative words, and conjunctions, the relations between words, sentences, and clauses in the language are closely connected, so that the sentence structure is complete and is standardized, thus forming the syntactic features of English hypotaxis, as well as the features of English long sentences with complex structure, many logical levels, many modifiers, and many parallel components.

3.2. Optimization of Professional Literature Translation. Professional words can be divided into two categories: common words and technical terms. Compared with the general corpus, the boundaries between these two types of words are clearer. Among them, common words appear frequently, which reflects the writing norms and sentence characteristics of patent corpus, but their vocabulary is not large. First of all, a term is a fixed or semifixed word or phrase that is closely combined. It should be a linguistically established word. However, it is different from common words. The main difference is that it is a language unit with strong domain characteristics and is used in a specific professional field. Therefore, terms have three characteristics: tight combination, language completeness, and domain.

The part of speech words can also be determined by disambiguation according to the context information, but there is no new word that has never appeared in the training corpus, and naturally there is no context information that can be associated with it. Therefore, the processing of new words in part of speech tagging faces greater challenges. The training data corresponding to each affix is sparse and the part of speech of Chinese affixes has great ambiguity. Because Chinese words are shorter than English, some existing metagrammatical features may not be applicable to Chinese. Statistical methods are used to predict the part of speech of unknown words, so there is no need to add artificial customized grammatical rules. Therefore, when the training corpus is large enough, statistical methods can be applied to any language and have strong adaptability and practicability.

HNC divides cognitive structure into local and global associations. Local association refers to the association at the lexical level, which can be simply summarized as follows: dividing concepts into abstract concepts and concrete concepts, expressing abstract concepts with quintuple and semantic network, and expressing concrete concepts with anchored expansion approximation.

Specifically, it is to translate English long sentences directly according to the fixed order, without considering too much linguistic and artistic effects, mainly to ensure that the translation is clear and accurate or even to change jumbled long sentences into multiple short sentences in the translation process. According to the limited objects of auxiliary semantic blocks, the semantic blocks and even sentence types are further analyzed to reduce the difficulty of semantic block analysis and anatomy.

Technical terms are the linguistic expressions of concepts in various disciplines, the achievements of scientific research, and the crystallization of knowledge language in the course of human progress. In order to clearly explain the functions and technical characteristics of patents, it is necessary to use a large number of technical terms, and even some terms are self-coined. The domain difference between professional literature and general corpus determines that the word segmentation method in general domain is not suitable for professional literature. According to the inherent characteristics of professional literature, this paper adopts the word segmentation method based on HNC and statistics. The extraction process of technical terms in this paper is shown in Figure 2.

There are two steps to extract professional terms:

(1) Firstly, according to the rules of word formation, the rules of term extraction are summarized, and the candidate terms are evaluated and selected by using algorithms and forbidden word lists, so as to obtain preliminary professional terms.

(2) Secondly, aiming at the problem that low-frequency technical terms are difficult to identify, the preliminary technical terms are used as the template training text, and the conditional random field model is used to construct the term extraction template, so as to finally extract meaningful low-frequency technical terms and improve the extraction accuracy of technical terms.

The basic idea is to inspect one training example at a time. According to the current weight vector $w^{(k)}$, the prediction result of the current instance is obtained, and, according to the error in the prediction result, the current feature vector is updated to obtain a new feature vector $w^{(k+1)}$. The algorithm determines the final weight vector through multiple iterations. Every iteration needs to traverse all training examples.

According to the current weight vector $w^{(k)}$, the approximate optimal dependency tree of top $-\,K$ is obtained by using the decoding algorithm based on column search and stored in $\widehat{\mathrm{d}}^{K}$. The update criteria are

$$\begin{aligned}&\min\left\|w^{(k+1)}-w^{(k)}\right\|,\\&s.t.\forall \mathrm{d}\in\widehat{\mathrm{d}}^{K} w^{(k+1)}\cdot f\left(\left(x^{(j)},\mathrm{d}^{(j)}\right)-f\left(x^{(j)},\mathrm{d}\right)\right)\geq \operatorname{loss}\left(\mathrm{d}^{(j)},\mathrm{d}\right).\end{aligned}\tag{5}$$

That is, the current feature vector $w^{(k)}$ is adjusted minimally, so that, for any dependency tree d in top $-\,K$, the difference between the score of the correct dependency tree $\mathrm{d}^{(j)}$ and the score of d is not less than the error number of d.

When the encountered new words have both prefixes and suffixes, how to select more effective affix information to improve the accuracy of new word labeling becomes the key. You can make a trade-off in the following ways:

(1) Choose the longer of prefix and suffix.

(2) Select the one with less parts of speech in the prefix and suffix distribution. If the prefix has a possible part of speech and the suffix has a possible part of speech, the prefix distribution is selected to estimate the lexical probability.

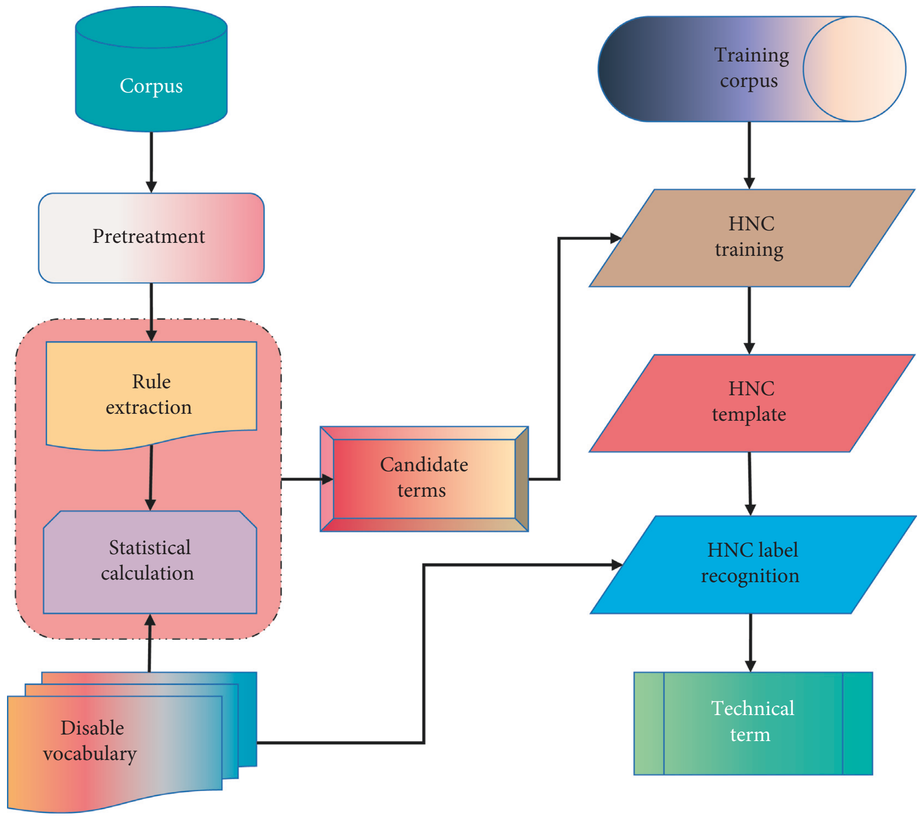


Figure 2: Technical term extraction process.

It is more useful to calculate the entropy of prefix and suffix distribution and then determine which one is more useful. The method of calculating entropy is used in part of speech tagging to measure how much information is necessary for independent data. For example, the formula for calculating the entropy of affixes is

$$S_a = -\sum_i \frac{n_{ai}}{N_i}\log_2\left(\frac{n_{ai}}{N_i}\right), \quad (6)$$

where n_{ai} is the number of times that part of speech i appears in all words with a as affixes and N_i is the total number of times that affix a appears. Through calculation, the affix distribution with the smallest gun is selected to estimate $P(w_i|t_i)$.

By transforming the problem of part of speech tagging of new words into the probability of launching survival words, the unification of statistical methods can be well realized, and the problems of acquiring rule knowledge and building dictionaries can be avoided. Assume that the transition probability is only related to the previous state; that is, the part of speech c_j of x_j is determined by the part of speech of w_{j-1}. Namely,

$$P\left(c_j|x_j\right) \approx \sum_{m=1}^{M} P\left(c_m|w_{j-1}\right)P\left(c_j|c_m\right), \quad (7)$$

where M represents the total number of parts of speech categories. According to the Bayesian formula, the probability of lexical emission is

$$P\left(x_j|c_j\right) = \frac{P\left(x_j\right)}{P\left(c_j\right)}P\left(c_j|x_j\right). \quad (8)$$

In the formula, the number of times in which marked symbol c_j appears in the training expectation is represented by $C(c_j)$. The number of times that the symbol string $c_m c_j$ appears together in the training corpus is indicated by $C(c_m c_j)$.

4. Result Analysis

In some cases, the translation of English long sentences is limited by time, which will increase the difficulty of translation. For example, for simultaneous interpretation and consecutive interpretation in English, it is necessary for the translator to respond in time and translate the information during or after speaking. Accordingly, Chinese expression pays more attention to the overall semantic expression in language, ensuring harmony, conforming to artistic conception, and reflecting the characteristics of parataxis. However, English expression pays attention to the structure and form of sentences in language, showing the characteristics of hypotaxis. For the translation of long English sentences, it is necessary to translate English sentences that focus on hypotaxis from parataxis level.

In order to investigate the role of HNC statistical model in automatic text classification system, the system uses the same training text and test text to conduct classification experiments on HNC statistical model and morphology-based model, respectively. The difference between the two systems is limited to feature weighting and selection. Take different numbers of training corpus and classify the same test set. Both HNC-based and word-form-based methods are used for comparative experiments. The experimental results are shown in Figure 3.

It can be seen that the classification accuracy of the system increases with the increase of training corpus, but, with the continuous increase of training corpus, the growth of the classification accuracy of the system slows down. No matter how much the training corpus is, the method based on HNC has higher accuracy than the method based on morphology. The highest difference is about 10% percentage points.

By the time of 80 articles, the accuracy rate of HNC-based method reached 92.35%, while that of morphology-based method was only 84.55%, with a difference of 7.8 percentage points. This shows that the classification method based on HNC has obvious advantages over the method based on morphology.

Table 1 is the comparison table of segmentation accuracy, recovery rate, and average cross-connection number between traditional algorithm and this algorithm. The accuracy rate and recovery rate of segmentation, respectively, indicate the accuracy rate of dividing complex English long sentences and the usage rate in specific translation, which are the basis for ensuring the accuracy of English long sentence translation and also important indicators. For the translation of complex English sentences, the higher the segmentation accuracy, the lower the average recovery rate, and the higher the translation accuracy.

It can be seen that the number of cross-connections of this algorithm is 17.46, which is 12.68 less than that of the traditional translation algorithm, indicating that the performance of this translation algorithm is better. Compared with traditional algorithms, this translation algorithm has higher translation accuracy and recovery rate, so it has higher practicability.

Figure 4 lists some experimental results. The experimental results mainly consider the recall rate, accuracy rate, and F value.

It can be seen that the recall rate of each step has achieved good results and the application of comprehensive method has improved the correct rate of sentence segmentation. However, the correct rate of comprehensive method is not very high, so the method needs further improvement.

English pays attention to the continuity of logical thinking and is used to connect the components of sentences by explicit means, so that the relationships among the components of sentences can be revealed, while Chinese pays more attention to parataxis. There will be some mistakes in word segmentation and labeling on the training corpus of this patent, and there will be some mistakes in the corresponding term recognition results. It is easy to divide the whole professional term or part of the term into a string with other words, and this phenomenon of word adhesion often occurs in the front and back boundaries of terms. Particularly, for long sentences, the analysis time of high-order models increases rapidly. This

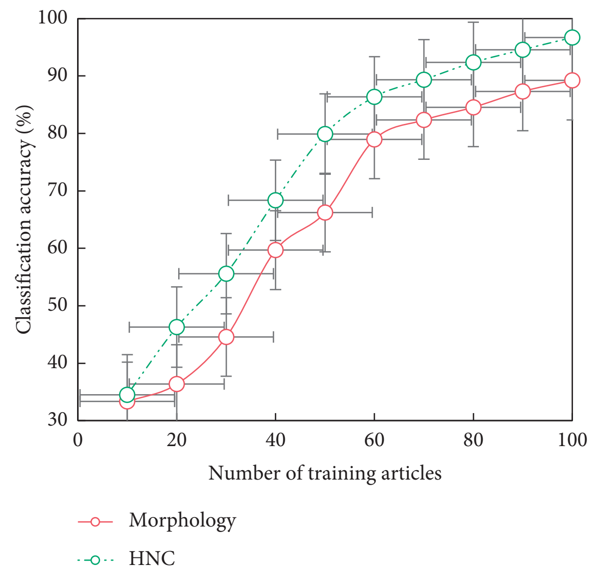


Figure 3: Classification result.

Table 1: Comparison results of various indexes between traditional algorithm and this algorithm.

	Traditional algorithm	Algorithm in this paper
Number of clauses after segmentation	1369	1683
Correct cut fraction	1224	1579
Cross-connection number	30.14	17.46
Average accuracy	89.41%	93.82%
Average recovery rate	7.16%	15.38%

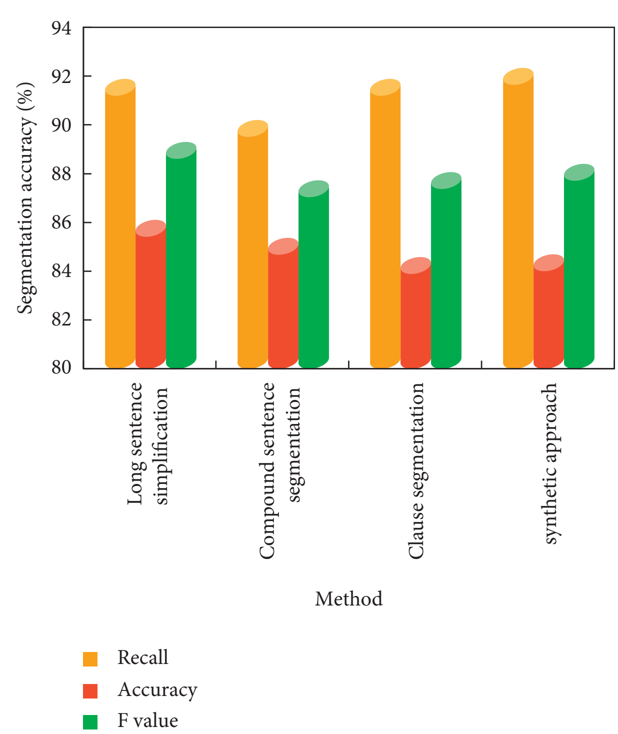


Figure 4: Experimental results of long sentence segmentation.

restricts the application of high-order dependency parsing model in high-level natural language processing systems, such as machine translation and information extraction. Therefore, according to the characteristics of Chinese, we propose a fast dependency parsing method based on punctuation.

The translation results are evaluated by the international general evaluation indicators BLEU and mteval-v13a. Table 2 shows the evaluation results of the three systems.

Under a standard answer, the BLEU values of this system and [10] translation system are 0.193, and 0.142, respectively. The BLEU value of the former is much higher than that of the latter. It can be seen that the translation quality of this system is obviously better than that of the comparative translation system.

The corpus used in this experiment is the Chinese patent data provided by the patent MT task of NTCIR-9 conference. In this paper, HNC and CRF++ tools are used to train the training corpus, so as to extract patent terms. The sampling statistical results of HNC and CRF algorithm terminology recognition are shown in Figure 5.

It can be seen that although the training accuracy of CRF model is lower than that of HNC, more technical terms are identified. In this experiment, 1863 preliminary terms were identified by HNC, and 2014 terms were identified by CRF model training. Compared with HNC, the vocabulary difference (number of changes) was 56893.

In this paper, the patent HNC is constructed by rule extraction and statistical learning, and there are 440, 247 entries in the constructed HNC. Figure 6 is the experimental result of segmenting all words in the test corpus in the open test environment.

It can be seen that, through word segmentation experiments on professional literatures in many fields, the general word segmentation tool in [17] has only 70.66% accuracy rate for patents and the effect is not satisfactory. The word segmentation method based on HNC and statistics proposed in this paper has achieved good results in the open test environment, with 93.38% accuracy and 94.51% recall in the open test.

The imperfection of the corpus makes it impossible to contain all linguistic phenomena, which leads to the deviation of the acquired linguistic information and the inability of correctly reflecting the real linguistic phenomena. It also causes serious data sparseness of the calculated model parameters, which affects the accuracy of labeling. Because the training corpus cannot contain all possible words and parts of speech, the language is constantly developing, and new words will be added. So, when there are a large number of new words, because the relevant information cannot be obtained from the training corpus, the probability of wrong tagging will be greatly increased, which will seriously affect the accuracy of tagging. Therefore, a better solution to the problem of new words is also a big problem that puzzles part of speech tagging.

For this reason, we extract the dependency structures corresponding to all segments from the training corpus. Figure 7 compares the efficiency of different methods in different sentence lengths.

Table 2: Evaluation of translation evaluation results.

System type	BLEU-4
Ref. [10]	0.142
Ref. [13]	0.176
This paper system	0.193

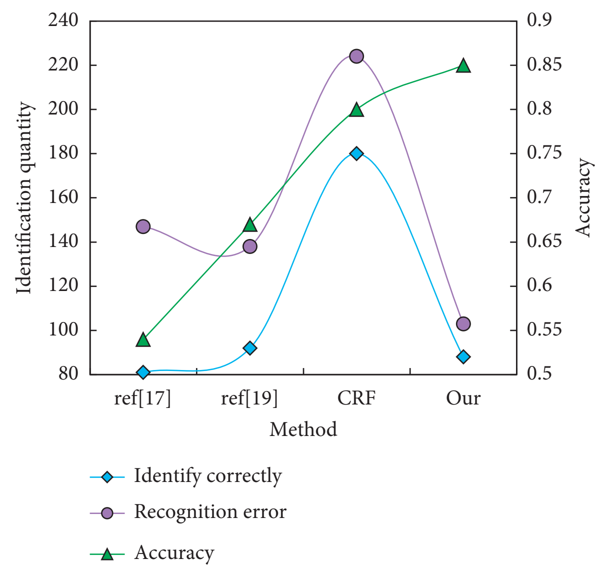


Figure 5: Sample results of term extraction experiment.

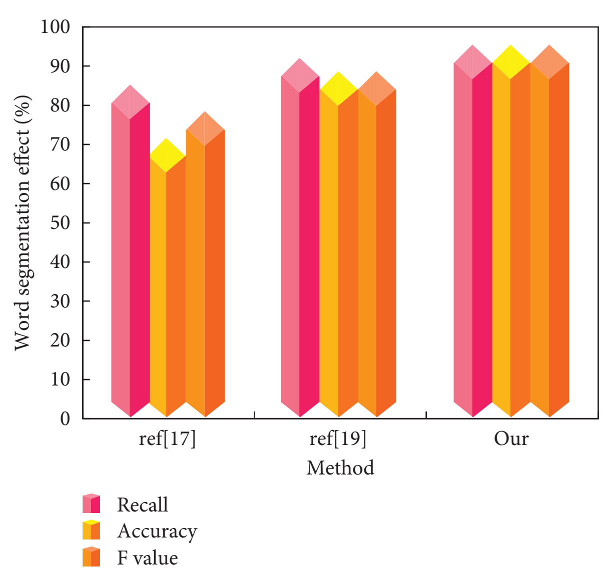


Figure 6: Experimental results under open test.

For the sake of clarity, we have omitted sentences longer than 110. Time analysis includes all operations from the input of sentences to the output of analysis results, such as feature extraction and decoding. It can be seen that the two-stage method greatly improves the efficiency of analyzing long sentences and the influence of sentence length on the analysis time is very small.

Because the average sentence length of the corpus is 25, we used three intervals: sentences less than 29, sentences

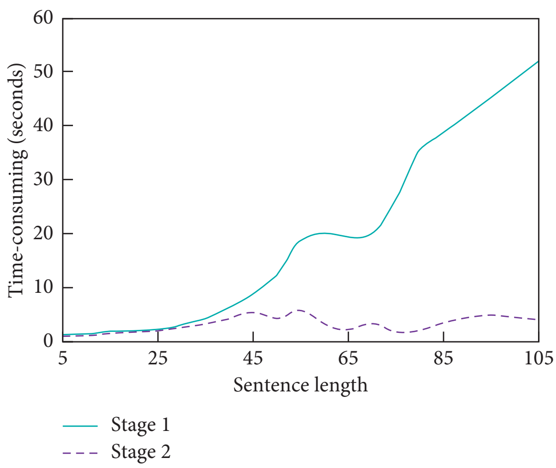


Figure 7: Efficiency comparison of different sentence lengths.

greater than 50, and sentences in the middle. It can be seen that, compared with the one-stage method, the two-stage method increases the syntactic accuracy with the increase of sentence length. For sentences longer than 28, the accuracy of the two-stage method is greatly improved.

Merging the converted source tree database with the target tree database to form a larger training data, finally, the syntactic analysis model is trained on the merged tree database. As the training data increases, the accuracy of the syntactic analysis model will naturally increase. However, it is a very difficult problem to deal with the inconsistency of annotation structures in different tree repositories. Through a clever strategy to reduce the noise contained in the transformed tree library, the performance of syntactic analysis can be improved. In the experimental part, we indirectly compare our method with theirs, and the results show that our method is more effective.

5. Conclusion

There are many long sentences in English, which are easy to be translated into blunt and dull Chinese due to the different syntactic features between English and Chinese, and sentence translation is the cornerstone of text translation, which makes English long sentence translation a major focus and difficult in translation. This paper applies HNC algorithm to English long sentence segmentation. A regular expression is used to describe or match a series of strings that conform to a certain syntactic rule; that is, it provides a way to match strings. The word segmentation method combining HNC with statistics can effectively improve the recognition rate of unknown words by extracting technical terms. The word segmentation method based on HNC and statistics proposed in this paper has achieved good results in the open test environment, with 93.38% accuracy and 94.51% recall in the open test. Careful error analysis shows that the joint model can better resolve syntactic sensitive part of speech ambiguities and correct resolution of these ambiguities can further help syntactic analysis.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This research was supported by "Model Course of Moral Education-College English," the Education Department of Hebei Province (KCSZ202104S).

References

[1] K. Shaalan, A. Hendam, and A. Rafea, "Rapid development and deployment of bi-directional expert systems using machine translation technology," *Expert Systems with Applications*, vol. 39, no. 1, pp. 1375–1380, 2012.

[2] Y. Zhang, "Research on English machine translation system based on the internet," *International Journal of Speech Technology*, vol. 20, no. 4, pp. 1017–1022, 2017.

[3] K. N. Dew, A. M. Turner, Y. K. Choi, A. Bosold, and K Kirchhoff, "Development of machine translation technology for assisting health communication: a systematic review," *Journal of Biomedical Informatics*, vol. 85, pp. 56–67, 2018.

[4] H. Wu, "Multimedia interaction-based computer-aided translation technology in applied English teaching," *Mobile Information Systems*, vol. 2021, no. 5, pp. 1–10, 2021.

[5] G. Duan, H. Yang, and K. Qin, "Improving neural machine translation model with deep encoding information[J]," *Cognitive Computation*, vol. 13, no. 3, pp. 1–9, 2021.

[6] C. R. Giannella, R. K. Winder, and J. P. Jubinski, "Annotation projection for temporal information extraction," *Natural Language Engineering*, vol. 25, no. 3, pp. 385–403, 2019.

[7] K. Matsuuchi, T. Fukami, N. Naoe, R. Hanaoka, S. Takata, and T Miyamoto, "Performance prediction of a hybrid-excitation synchronous machine with axially arranged excitation poles and permanent-magnet poles," *Electrical Engineering in Japan*, vol. 150, no. 2, pp. 43–49, 2005.

[8] W. Chen, D. Kawahara, K. Uchimoto, Y. Zhang, and H Isahara, "Using short dependency relations from auto-parsed data for Chinese dependency parsing," *ACM Transactions on Asian Language Information Processing*, vol. 8, no. 3, pp. 1–20, 2009.

[9] X Chun and M.-Y. Zhang, "Multiple linear regression for extracting phrase translation pairs[J]," *Journal of Computers*, vol. 6, no. 5, pp. 905–912, 2011.

[10] Y. Luo and Y. Xiang, "Application of Data Mining Methods in Internet of Things Technology for the Translation Systems in Traditional Ethnic Books[J]," *Data Mining for Internet of Things*, vol. 8, pp. 93398–93407, 2020.

[11] H. Motoyuki and H. Kazuhiro, "Media processing technology for achieving hospitality while on the go.[J]," *NTT Technical Review*, vol. 13, no. 4, pp. 1–7, 2015.

[12] S. Condon, M. Arehart, D. Parvaz, G. Sanders, C. Doran, and J. Aberdeen, "Evaluation of 2-way Iraqi Arabic–English speech translation systems using automated metrics," *Machine Translation*, vol. 26, no. 1-2, pp. 159–176, 2012.

[13] X. Wen, "Hierarchical phrase machine translation decoding method based on tree-to-string model enhancement[J]," *Acta*

Technica CSAV (Ceskoslovensk Akademie Ved), vol. 62, no. 1, pp. 531–539, 2017.

[14] Y. Wei and Y. Guo, "Research and analysis of improved extraction based on information processing technology[J]," *Journal of Digital Information Management*, vol. 10, no. 2, pp. 142–146, 2012.

[15] X. Wu, Y. Xia, J. Zhu, L. Wu, S. Xie, and T Qin, "A study of BERT for context-aware neural machine translation," *Machine Learning*, vol. 111, no. 3, pp. 917–935, 2022.

[16] X. Liu, W. Wang, W. Liang, and Y Li, "Speed up the training of neural machine translation," *Neural Processing Letters*, vol. 51, no. 1, pp. 231–249, 2020.

[17] Y. Jiang and H. Liu, "Research on the construction of parallel corpus for the specific field of machine translation[J]," *Boletin Tecnico/Technical Bulletin*, vol. 55, no. 19, pp. 77–82, 2017.

[18] D. Wang, J. Su, and H. Yu, "Feature extraction and analysis of natural language processing for deep learning English language," *IEEE Access*, vol. 8, pp. 46335–46345, 2020.

[19] B. Geng, "Text segmentation for patent claim simplification via bidirectional long-short term memory and conditional random field," *Computational Intelligence*, vol. 38, no. 1, pp. 205–215, 2022.

[20] Y. Kedar, "Interaction between two determiner systems: the acquisition of English articles by a Hebrew-speaking child," *First Language*, vol. 39, no. 1, pp. 111–136, 2019.

[21] S. Tubau, "The asymmetric behavior of English negative quantifiers in negative sentences," *Journal of Linguistics*, vol. 56, no. 4, pp. 775–806, 2020.

[22] R. C. Pereira, "Epenthesis and deletion as strategies to acquire complex syllabic structures: strategies in the Interlanguage of Brazilian learners of German as a foreign language," *System*, vol. 98, no. 2, Article ID 102479, 2021.

[23] L. Ge, "Study on the translation techniques in English Chinese translation of passive sentences in educated[J]," *Linguistics*, vol. 3, no. 1, pp. 407–420, 2021.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9783613, 1 page
https://doi.org/10.1155/2023/9783613

Retraction

Retracted: Hierarchical Strategies for Building Multiethnic Communities from the Perspective of Data Mining Analysis

Mobile Information Systems

Received 22 August 2023; Accepted 22 August 2023; Published 23 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] S. Qin, "Hierarchical Strategies for Building Multiethnic Communities from the Perspective of Data Mining Analysis," *Mobile Information Systems*, vol. 2022, Article ID 7506323, 10 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 7506323, 10 pages
https://doi.org/10.1155/2022/7506323

Research Article

Hierarchical Strategies for Building Multiethnic Communities from the Perspective of Data Mining Analysis

Suhan Qin

Guangdong Polytechnic Normal University, Guangzhou 510665, China

Correspondence should be addressed to Suhan Qin; qinsuhan@gpnu.edu.cn

Received 28 June 2022; Revised 27 July 2022; Accepted 4 August 2022; Published 23 August 2022

Academic Editor: Xuefeng Zhang

The formation of a multi-ethnic community is the need of the times for ethnic relations. Its construction must first establish a homogeneous space that can support interethnic coexistence, including national public living space and national public cultural space, and, on this basis, form a heterogeneous space that is compatible with the existence of national individuality, so as to appreciate in national understanding. The expression, communication, and broad recognition of national individuality determine its important position in the construction of national community. In the process of building the Chinese nation community, the main purpose is to strengthen the Chinese cultural identity, build a spiritual home shared by all ethnic groups, and consolidate the community consciousness of the Chinese nation. Based on this, this paper mainly studies the engineering application based on data mining algorithm and data mining algorithm in the construction of multi-ethnic community. By summarizing the experience in the process of the construction of the Chinese nation community, some solutions are put forward to the problems in the construction of the national community and strive to focus the value orientation of all members of the nation on the identification of the common values of the Chinese nation.

1. Introduction

The construction of ethnic communities in multi-ethnic countries is a systematic project, which is different from the natural evolution of indigenous people [1]. As a constructive cultural group, the establishment and improvement of the ethnic minority community (state) requires external intervention. Through cultural integration, a common ideology is formed, and it is established and regarded as the common spiritual home of all citizens, so as to improve the conscious ability of the ethnic minority community, to lay a solid foundation for the establishment of ethnic minority communities [2]. In the process of establishing a national community, the French nation's formation and integration paradigm is distinctive and has achieved good results, providing reference and reference for other multi-ethnic countries to establish ethnic communities [3].

My country has been a unified multi-ethnic country for a long time in human history, and it has created an excellent external environment for us [4]. In China's long-term feudal society, due to the fact that the exchanges and unions within China's various ethnic groups were at a relatively spontaneous stage, and due to the ethnic contradictions caused by the mutual oppression of ethnic groups, there was no ethnic community consciousness (national consciousness) among the various ethnic groups. Under the external threat of genocide brought about by China's imminent national subjugation and imperialist invasion, all ethnic minorities in China are united unprecedentedly under the unified leadership of the CCP, and their national consciousness has awakened [5]. Table 1 shows the regional distribution of some ethnic minorities in my country. Through the bloody battles of the people of all ethnic groups and the success of the new democratic revolution, the People's Republic of China has established a democratic regime, and the people of all ethnic groups have entered the era of democracy together. During the founding stage of the People's Republic of China [6], the Communist Party of China established a system of regional ethnic autonomy with Chinese characteristics, combined Marxist theory with the reality of China's ethnic

Table 1: Distribution of some ethnic minorities in my country.

Ethnic name	Main distribution area
Mongolian	Inner Mongolia Autonomous Region, Liaoning Province, Jilin Province
Hui nationality	Ningxia Hui Autonomous Region, Gansu Province, Henan Province
Tibetan	Xinjiang Autonomous Region, Sichuan Province, Qinghai Province
Uighur	Xinjiang Uygur Autonomous Region, Hunan Province
Yi people	Guizhou Province, Yunnan Province, Hunan Province
Zhuang	Sichuan Province, Yunnan Province, Guizhou Province
Buyi	Guangxi Zhuang Autonomous Region, Yunnan Province, Guangdong Province
Korean	Guizhou Province
Manchu	Jilin Province, Heilongjiang Province, Liaoning Province
Dong nationality	Liaoning Province, Jilin Province, Hebei Province
Yao nationality	Guizhou Province, Hunan Province
Bai people	Guangxi Zhuang Autonomous Region, Hunan Province, Yunnan Province
Tujia	Yunnan Province, Guizhou Province
Hani	Hunan Province, Hubei Province, Sichuan Province
Kazakh	Yunnan Province
Ethnic name	Xinjiang Uygur Autonomous Region, Gansu Province

minorities, actively dealt with ethnic relations and strived to achieve ethnic equality, and the reunification of the motherland and the community of the Chinese nation provide the basic and necessary conditions for the establishment of the country [7].

As of December 2021, the author has retrieved a total of 1,968 related academic papers under the title “Community of the Chinese Nation” in CNKI, of which there are 977 papers in 2021 (see Figure 1) [8]. After an overall analysis of it, it is found that the most frequently occurring keywords are “Chinese nation,” “cultural identity,” “five identities,” and “national unity” [9]. It can be seen that in previous studies, most scholars have focused on the theoretical interpretation at the macro level, the development of the context, the mining of value implications, and the design of the cultivation path [10]. The research focuses on the single-ethnic research at the meso-level and the research on colleges and universities [11]. As far as the relevant research results have been found so far, few scholars have discussed the relevant issues of “forging the consciousness of the Chinese nation’s community” from the perspective of the daily production and life practice of villagers (residents) within the “interembedded community” [12].

In today’s era of big data and artificial intelligence, data analysis, algorithms, and computing capabilities are particularly important [13]. It is the core of information mining research to establish a method in the analysis of a large amount of data and discover the knowledge points after “information association,” so as to visualize its meaning. In the data mining algorithm, neural network is generally a systematic and effective data analysis and processing method for dividing and processing digital information such as words, pictures, and sounds [14]. According to the principle of the big neural network, the data mining method is used to realize the national community big data analysis system. At the same time, according to the software engineering concept, the requirements analysis, summary and detailed analysis, program implementation, and debugging are completed. Using the analysis of data from the visualization platform, important information on issues such as “politics,” “culture,” and “society” was discovered.

So far, although various basic theoretical research results have laid a good foundation for the consciousness of the Chinese nation, there is still a large research gap, and there are still many new theoretical research contents that need to be supplemented [15]. The progress of scientific research can promote the research and development of ethnic minorities in various disciplines [16]. Therefore, this study will pay more attention to the theoretical system and practical application value of the construction of the Chinese nation’s community, put forward operational strategies and measures to better protect the Chinese nation’s civilization and the cohesion of the Chinese nation, and provide the necessary measures for the realization of the great rejuvenation of the Chinese nation. The basic theoretical guarantees a useful framework for evaluating and managing environmental health risks across different settings [17].

2. Related Concepts and Theoretical Basis

2.1. The Essence of the National Community. The 5,000-year history of civilization of the Chinese nation is the history of the harmonious inheritance of civilizations of all ethnic groups and the history of the process of mutual recognition and peaceful development of civilizations of various ethnic groups [18]. As an open and inclusive family, the Chinese nation gathers ethnic minority compatriots of different life backgrounds, languages, appearances, and skin colors to ensure harmonious coexistence and common development with people of all ethnic groups. The free flow and open dissemination of various civilizations have always been the most important law affecting the development of human civilization. The Chinese nation should retain the diversity of civilization and rich and colorful spiritual life style, rather than undermining the civilizational advantages of all ethnic groups. The following situations can be considered.

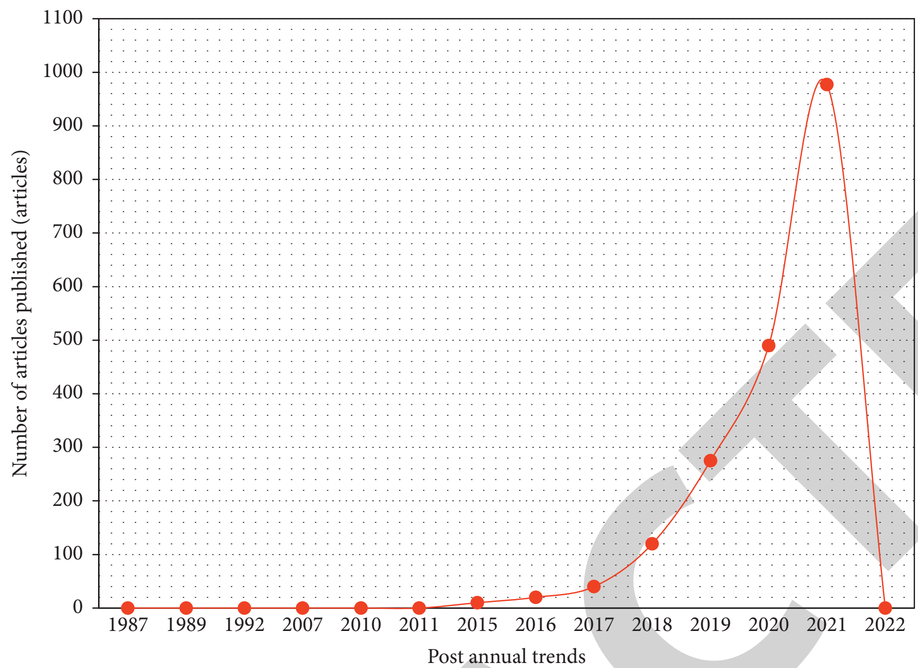


FIGURE 1: Search results with the title "Community of the Chinese Nation".

2.1.1. Maintain the Coexistence of National Cultural Integration and Independence. History is produced under certain conditions and is the result of historical progress and development. The cultural form will change with the development and change of the economy, and various places in a city will also produce different cultural forms [4].

In order to protect our own fresh blood, we must accept foreign cultures. Teacher Chen Yinke also talked about the combination of his own cultural theory and foreign culture: "On the one hand, we must absorb foreign theories; on the other hand, we must not forget the status of primitive peoples." We learn to use the power of family and have a deeper understanding of foreign cultures to realize "one's own words."

The development of multiculturalism broke the "quiet" loneliness of the local culture. Out of the awareness of self-protection, when a foreign culture enters, the culture will first show its own rejection and try to protect itself, but after the test of time, it is gradually found that a single cultural form is not enough to meet the needs of current social development. Through continuous running-in and practical testing, Chinese culture is gradually liberated in national independence, and more diversified cultural existence is accepted. The uniqueness of culture is manifested in a certain natural environment. It can neither be seen as a narrow self-knowledge nor can it be defined separately from the national environment. The uniqueness and integration of culture are formed together, and there is no uniqueness. People's values will lose their personal development, and if they lack integration, national communities cannot be built. The two coexist and coexist.

2.1.2. Multicultural Conflict and Inclusive Coexistence. In the reality of affirming and accepting the existence of multiculturalism, people should recognize the opposition and contradiction within civilization. Contradiction is the root and foundation of everything. People should dialectically deal with the conflict within ethnic pluralism developmentalism. "The imbalance of things is also the love of things." If all things maintained the same principle, law, consistency, and order, the development of the world and the progress of the world would stop. Cultural differences can also effectively reflect the huge space for the progress of civilization. With the development of information network and the deepening of economic internationalization, understanding the gap, strengthening communication, and identification are the only way for the development of national culture.

As Marxism said, people must be clearly aware of the objective existence of the problem. "Since ancient times, every country has been superior to other countries in some aspects." Seeking common ground while shelving differences is the gentlest way to deal with cultural conflicts. Seeking common ground "Cunqi" conforms to an important category of traditional Chinese social culture "reconciliation." "Harmony" means harmony, unity, public world, and inclusiveness. Harmony means symbiosis and coexistence. "Reserving differences" means leaving space for both parties to be more tolerant and valuing of each other. Various forms of cultural exchanges can develop together in a period of time, inclusive and mutually beneficial, exchanges and mutual learning, mutual promotion, and common development, which further reflects the inclusive characteristics of Chinese culture.

2.1.3. The Dynamic and Innovative Coexistence of Multiculturalism. The spread of Chinese culture is a process of not forgetting the origin of Chinese civilization and the

Table 2: Differential pattern of ethnic communities in my country.

Differential structure	Chinese nation of all ethnic groups	Chinese nation of the Chinese people	Chinese nation of the world Chinese
Properties of home	Home of people of all nationalities	Chinese people's home	The home of Chinese all over the world
National attribute	Multi-ethnic republic	Sovereign state of the Chinese people	The "Motherland" of Chinese cultural identity in the world
National attribute	The Chinese nation formed by the Han nationality and various ethnic minorities	The Chinese nation on the basis of the Chinese nation	Chinese nation
People attribute	Chinese nation of all ethnic groups	Chinese nation	Chinese nation of the world Chinese
Diversity of the Chinese nation	Multi-ethnic cohesion of ethnic and cultural diversity	Chinese culture of the Chinese people	Integration of Chinese culture
Cultural identity	The cultural identity of the Chinese national community united by all ethnic groups	The cultural identity of the Chinese nation community	The cultural identity of the Chinese nation community
Cultural awareness	Cultural consciousness of various nationalities	The connection between Chinese culture and world culture	Cultural consciousness of Chinese civilization
Cultural connection	The connection between the cultures of various Chinese nationalities and the cultures of the world's nationalities	Narrow nationalism	The connection between Chinese civilization and the human community
Cultural transcendence	Narrow self-nationalism	United the Chinese people	Narrow Chinese nationalism
Chinese culture	Bringing together all nationalities	Chinese nation of the Chinese people	Bringing together Chinese at home and abroad

spirit of the main civilization. This is a process of re-practice, re-innovation, and sublimation of mainstream civilization. Cultural creation can also enhance the ability of social subjects to absorb, digest, and apply scientific knowledge, and it is also a reflection of the cultural enthusiasm of social subjects. Cultural creation in each era is a node of the cultural heritage of the entire Chinese nation. By absorbing overseas advanced civilizations and blending multiple civilizations at home and abroad, it reflects the general needs of human society for spiritual and cultural development. The meaning of "new" is to emphasize the contemporary value of culture to human free will.

In order to adapt to the changes of the times, conform to the trend of economic and social development, and maintain vitality, all ethnic groups need to constantly innovate and update to ensure the renewal of the blood of civilization. All members of the ethnic community should not only maintain a humble attitude and jointly absorb the essence of Chinese culture but also actively resolve internal conflicts, promote healthy development, constantly explore new ideas, and constantly create new content, so as to contribute beneficial cultural nutrition to the construction of ethnic communities.

2.2. Problems in the Process of Building a National Community. Theoretical knowledge comes from concrete reality. In the process of forming a human community, each human being must go through a stage of integration and growth, from isolation to mutual communication and understanding, from passive "freedom" to active "self-action." My country's development and construction will inevitably require the introduction of a series of national policies and regulations to meet the needs of modern development. The reform and opening up not only provided opportunities for the establishment and improvement of our society but also caused the dilemma of national identity. Among them, Table 2 shows the differential order pattern of my country's national community.

Various political ideological trends emerged while different ethnic groups interacted. In addition, my country's frontier management system is still not perfect, which makes some ethnic minorities feel a sense of separation; some localities and individual classes have accepted the influence of separatism and religious beliefs in Western regions. Because of the impact of extremism and the indifference of community awareness, the understanding and practice of racial independence or ethnic separation in some places have had a certain impact on the development of China's community.

2.3. Overview of Data Mining Perspective. As the name suggests, a frequent pattern is a process in which data are continuously generated. Mining frequent data helps to find interesting problems hidden in a large amount of information. The mining of frequent items is the cornerstone of frequent pattern mining, and the concepts and methods are elaborated in Chapter 4.

Another important data mining task for predictive data analysis includes classification and regression. In it, the class pattern is deduced by the class of the training data set (statistical objects with class labels), and the class labels used to analyze the uncertain data objects in the class labels. Chapters 3 and 4 will elaborate on the research and application of neural network algorithms in analytical tasks. In addition, regression analysis, as one of the most common big

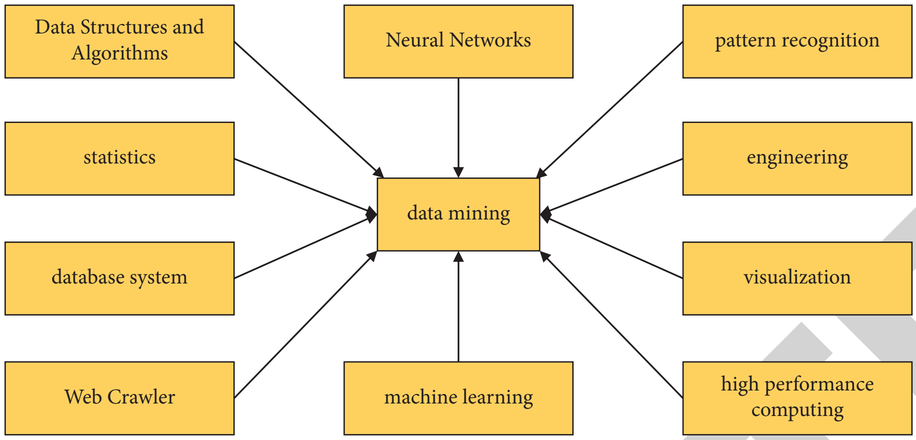


Figure 2: Data mining draws techniques from multiple domains.

Table 3: Functions and applications of data mining algorithms.

Category	Function	Algorithm	Typical application
Prediction model segmentation	Classification	Decision trees, neural networks, and difference analysis	Target marketing, quality control, and risk analysis
	Numerical prediction	Linear regression, nonlinear regression	Profitability analysis
Link analysis	Cluster analysis	k-means and neural network	Market segmentation and customer segmentation
Predict category prediction model	Association discovery	Statistics and set theory	Cross-marketing
	Sequence association discovery	Statistics and set theory	Time-series basket analysis
	Similar time series discovery	Statistics and set theory	Sales flow and stock price volatility
Segmentation	Time series forecasting	Statistical time series models	Sales forecast, interest rate forecast, and inventory control

data prediction and calculation methods, can also be applied to the identification of data distribution trends, as well as image fitting and trend prediction. It can be applied in the fields of stock research and biological population derivation.

After decades of vigorous development, the data resource mining profession can be said to be changing with each passing day and never out of date. The types of its applications are becoming more and more diverse, the application methods are also changing rapidly, and the use environment will also expand day by day. It can be summarized as the following characteristics: multi-disciplinary integration, oriented to specific needs and applications, large-scale and rich data, and interesting models. Figure 2 shows examples of disciplines that have had an important impact on the development of data mining.

3. Related Technologies

This chapter describes algorithms for three classic data mining tasks. There is no best method for data mining, only the best method for a specific goal. Table 3 shows the different functions and typical uses of different algorithms in data mining techniques. The specific reasons are as follows:

TextCNN is used in text analysis tasks because CNN is the most commonly used module in neural networks, and because TextCNN is also a technical framework that is easy

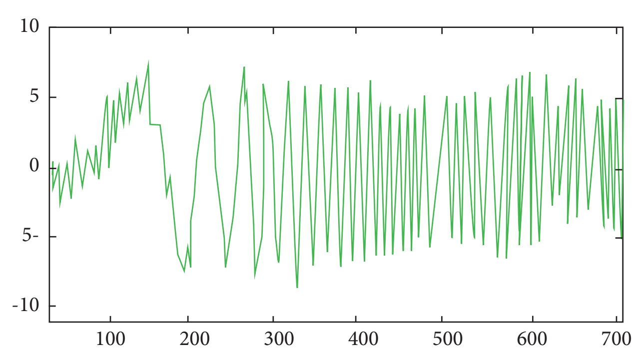


Figure 3: Noischir signal.

to master and apply, it is increasingly mature in engineering applications. The output works fine; in the task of association rules, FP-Growth is selected because the improvement of its algorithm greatly improves the computational efficiency compared with the traditional similarity calculation. Its essence is to convert the operation space (memory) into operation time and perform the operation on the laboratory server. In this way, sufficient network resources are used to effectively deal with the time consumption problem.

LDA is chosen in the task of unsupervised clustering, so it can process text-like data more efficiently and is a rare algorithm suitable for the implementation on subsequent platforms. Figure 3 summarizes the actual use of the three data mining algorithms described in this paper and introduces the application scenarios, advantages, and disadvantages of several algorithms in detail. We will describe it in detail in Chapter 5. The successful application of the big data analysis platform can test the feasibility, scope, and effect of the three methods in the design.

Here, x represents the input of the j th node $(j = 1, \ldots, M)$ of the input layer; wj represents the weight between the i th node of the hidden layer and the j th node of the input layer; θ represents the i th node of the hidden layer; $\phi(x)$ represents the activation function of the hidden layer; wi represents the weight between the k th node of the output layer to the $i(i = 1, \ldots, q)$ node of the hidden layer; and a represents the k th node of the output layer (k).

The four models of its data mining are as follows:

(1) Node output model is as follows:

Hidden layer node is as follows:

$$O_i = f\left(\sum w_{ij} \times x_i - q_j\right). \tag{1}$$

Output node is as follows:

$$Y_k = f\left(\sum T_{jk} \times O_j - q_k\right). \tag{2}$$

f is nonlinear action function and q is neural unit threshold.

(2) Action function (Sigmoid excitation function) model is as follows:

$$f(x) = \frac{1}{1_e^{-x}}. \tag{3}$$

(3) Error calculation model is as follows:

$$E_p = \frac{1}{2} \times \sum \left(T_{pi} - O_{pi}\right)^2. \tag{4}$$

T_{pi} and O_{pi} are the expected and calculated values of i, respectively.

(4) Self-learning model is as follows:

$$\Delta w_{ij}(n+1) = h \times \Phi_i \times O_j + a \times \Delta w_{ij}(n). \tag{5}$$

h and a are learning factor and momentum factor; ϕ and O are the calculated outputs of i and j.

According to the flow of the BP algorithm, the formula is derived as follows:

(1) The forward propagation process of the signal, equations (6) to (9), is as follows:

$$\text{net}_i = \sum_{j=1}^{M} w_{ij} x_j + \theta_i, \tag{6}$$

$$y_i = \Phi(\text{net}_i) = \Phi\left(\sum_{j=1}^{M} w_{ij} x_j + \theta_i\right), \tag{7}$$

$$\text{net}_k = \sum_{i=1}^{q} w_{ki} y_i + a_k = \sum_{i=1}^{q} w_{ki} \Phi\left(\sum_{j=1}^{M} w_{ij} x_j + \theta_i\right) + a_k, \tag{8}$$

$$\begin{aligned} a_k &= \Psi(\text{net}_k) = \Psi\left(\sum_{i=1}^{q} w_{ki} y_i + a_k\right) \\ &= \Psi\left(\sum_{i=1}^{q} w_{ki} \Phi\left(\sum_{j=1}^{M} w_{ij} x_j + \theta_i\right) + a_k\right). \end{aligned} \tag{9}$$

Note: net is the input of the i th node of the hidden layer; y is the output of the ith node of the hidden layer.

Net is the input of the k th node of the output layer.

o is the output of the k th node of the output layer.

(2) The back-propagation process of the error is as follows:

$$\begin{aligned} \Delta w_{ki} &= -\eta \frac{\partial E}{\partial w_{ki}}, \\ \Delta a_k &= -\eta \frac{\partial E}{\partial a_k}, \\ \Delta w_{ij} &= -\eta \frac{\partial E}{\partial w_{ij}}, \\ \Delta \theta_i &= -\eta \frac{\partial E}{\partial \theta_i}, \\ \Delta w_{ki} &= -\eta \frac{\partial E}{\partial w_{ki}}, \\ &= -\eta \frac{\partial E}{\partial \text{net}_k} \frac{\partial \text{net}_k}{\partial w_{ki}}, \\ &= -\eta \frac{\partial E}{\partial O_k} \frac{\partial O_k}{\partial \text{net}_k} \frac{\partial \text{net}_k}{\partial w_{ki}}. \end{aligned} \tag{10}$$

Note: Ep is the quadratic error criterion function of sample P; E is the total error of the training sample P.

Δwi is the correction amount of the output layer weight.

Δa is the correction amount of the output layer threshold.

Δwj is the correction amount of the hidden layer weight.

$\Delta \theta$ is the correction amount of the hidden layer threshold.

$$\Delta w_{ki} = \eta \sum_{p-1}^{p} \sum_{k=1}^{L} \left(T_k^p - O_k^p\right) \psi'(\text{net}_k) \cdot y_i\,. \tag{11}$$

Additional momentum method is as follows:

$$\Delta w(i+1) = \eta\left(\Delta w(i) + u\frac{\partial E}{\partial w}\right), \quad (0 < \eta < 1), \tag{12}$$

where η represents the momentum factor.

The purpose is to increase the weight value on the basis of error back-propagation to achieve the effect of accelerating convergence.

Adaptive learning rate is as follows:

$$\mu(i+1) = \begin{cases} 1.05\mu(i), & E(i+1) < E(i), \\ 0.7\mu(i), & E(i+1) > 1.04E(i),\ (0 < \mu < 1), \\ \mu, & \text{otherwise.} \end{cases} \tag{13}$$

In the basic BP algorithm, its learning rate is fixed; therefore, when the algorithm reaches the extreme point, it may swing left and right, causing it to fail to converge. By adding formula (14) to the BP algorithm, the purpose of automatically adjusting the learning rate can be achieved, so that the algorithm can converge to the minimum value as much as possible.

The weight w adjustment method is as follows:

$$\Delta w = \left(J^T j + ml\right)^{-1} J^T E, \tag{14}$$

where E is the error vector; J is the Jacobian matrix of E to Δw; and M is the scalar.

The L-M method can change in two extreme cases $m \longrightarrow 0$ and $m \longrightarrow \infty$, so its convergence speed is faster.

4. Experimental Results and Analysis

4.1. Experimental Conditions. Matlab7.10.0 (R2010a) implements wavelet analysis and wavelet packet analysis noise reduction algorithms, respectively.

4.2. Experimental Environment. HP Presario CQ36 laptop, Inter(R) Core(TM) i3 CPU M330 @ 2.13 GHz, installed memory (RAM) 2.00 GB (1.87 GB available), hard disk 320 GB; running on Windows 8.1 Professional Edition performed on the operating system.

4.3. Experimental Data. Select the Noischir signal that comes with MATLAB, and the low-frequency and high-frequency parts of the signal contain noise at the same time, as shown in Figure 3.

National culture witnesses the history, is the spiritual wealth that has been passed down from generation to generation, and is the symbol of characteristic culture. In the ecology of national culture dissemination, national culture must keep pace with the times while retaining its own uniqueness. Therefore, how to get rid of the shackles of content homogeneity in the current social environment has become a problem that we need to think about. Rooting in national characteristics, insisting that content is king, and mining innovative content output methods are important issues to improve the communication power of national culture at present.

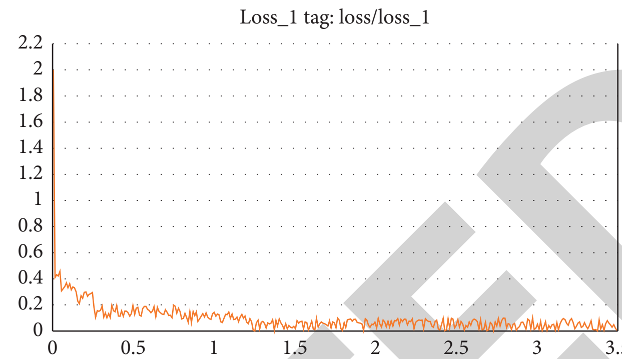


Figure 4: Text CNN loss rate curve.

In the research, a data set equivalent to the content of the previous section is selected for research; that is, the content of the subset of THUCNews is trained and tested. These data sets are the same as those used by fastText and NNF, and the results are similar. These eight categories are provided, including sports, finance, real estate, education, science, fashion, news, and entertainment. These categories have a total of 6,500 short text data, as shown in Figure 4.

The data set includes 5000 ∗ 8 classification training sets, a total of 40000 pieces of data information; 500 ∗ 8 classification test sets, a total of 4000 data information; and 1000 ∗ 8 levels of test sets, a total of 8000 pieces of data. The number of rounds of practice is set after the first twenty rounds, and convergence has been basically achieved in the fifteenth round of practice.

4.4. Selection of Training Data and Calculation Parameters. In this paper, the following four functions are taken as 5,000 sets of values, of which the first 4,900 sets of values are used for training and the last 100 sets of values are used for prediction. The BP network structure is 9–14–1, the maximum learning efficiency is set to 0.1, and the maximum training accuracy is 10–6. Therefore, the maximum training frequency of the BP neural network is 20,000 times, as shown in Figure 5.

From the core level, the construction of ethnic communities is the construction of national cultural soft power. The core of the Chinese nation's sense of community and the general recognition of the status of all members of the Chinese nation, as well as the consistent, broad and legally binding social values, and the special strength of all members. This is an inclusive value based on the common interests of all countries, addressing the common problems of all countries, and realizing the common development goals of all countries in the world. With the development of economy and society and the vital interests of the majority of citizens, the choice of life in the context of human cultural diversity should be regulated.

Practical activities refer to people's basic social activities, and reality itself is a concrete reflection of human subjective initiative. Regularity and purpose are the common

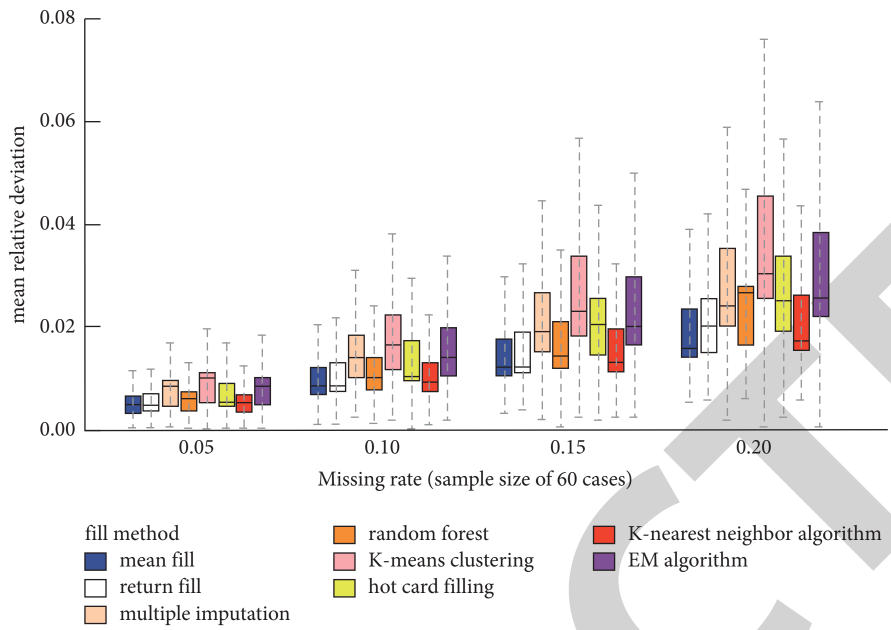


FIGURE 5: Relative error curves of the prediction results of the four algorithms.

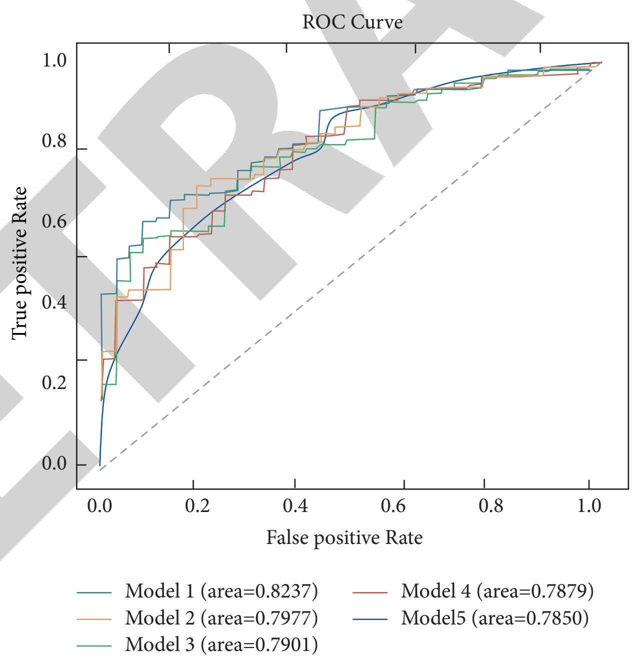


FIGURE 6: Consistency of the purpose and laws of practice and labor.

characteristics of people's activities, and the unity of the two is the essence of "real people." Because human practice is a purposeful and conscious activity, only through practice can we reveal the development process of the organic unity of purpose and meaning.

He also emphasizes, "How an individual can express his life (Figure 6), he is who he is. Therefore, what they have is the same as their product—both as they make and as how they make it consistent." He puts the concepts of history and social development into the corresponding concepts of productive forces and production relations, and verified them. What he means is that the form of production activities embodied by the labor force follows the basic purpose and principle of production, and its behavior itself reflects the personality and basic ability of human beings.

The full improvement of human nature and the advancement of social civilization are inseparable from the unremitting exploration and questioning of truth, goodness,

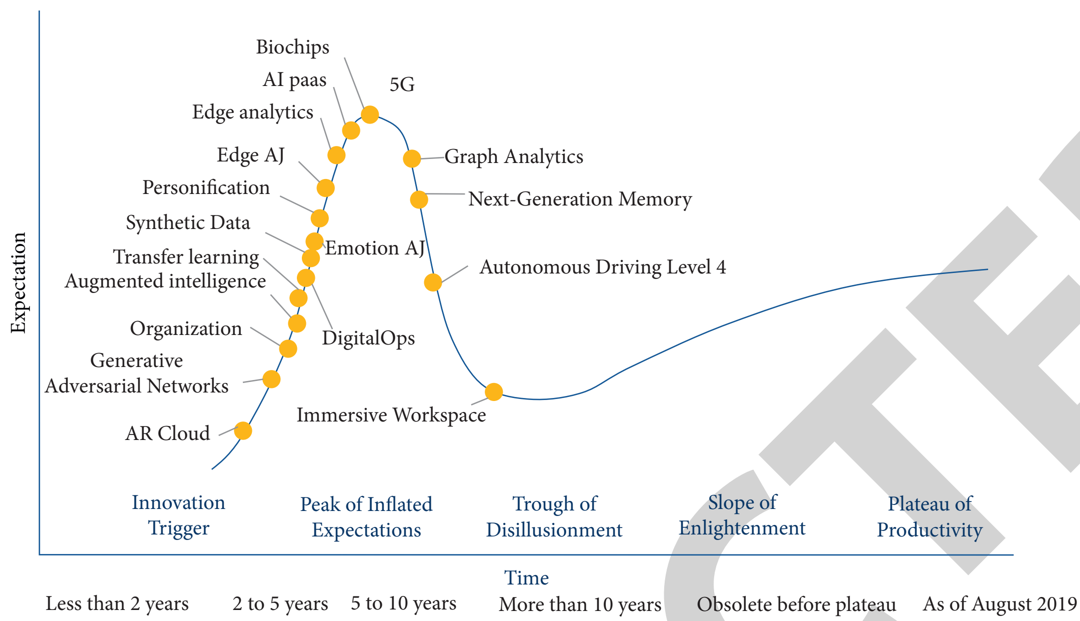


FIGURE 7: The application and development of emerging technologies such as data mining in the construction of national communities.

and beauty. As can be seen from Figure 7, seeking truth to understand objective things and to explore the law of changes in things is the spiritual activity of human beings to pursue science and adhere to truth. Seeking well can purify people's soul and ideological realm, which is the most fundamental spirit of human beings. To cultivate kindness is to cultivate virtue first. The yearning for beauty is the highest requirement for value based on "truth" and "goodness."

5. Conclusion

From the core level, the construction of the national community in the new era is the construction of the country's cultural soft power. It is a special spiritual force that brings together the inner strength, cohesion, and charisma of a nation. This paper presents the method of data preprocessing and analyzes its principle, which provides support for its application in the construction of multi-ethnic communities [19]. Thirdly, the performance of the current data mining is analyzed, and then, the genetic algorithm is used for analysis and improvement according to the defect that the initial weights and thresholds are not processed. Community-built applications provide support. In a word, in the process of building the Chinese nation community in China, by studying the development and integration experience of the French nation, it has brought useful reference and enlightenment to the Chinese people. Building a community of the Chinese nation is inseparable from the leadership of our country and the joint participation of people of all ethnic groups in the country. This is inseparable from improving the sense of identity of the Chinese nation, improving the sense of identity of the Chinese nation and society, and improving the spiritual home shared by the whole Chinese nation.

Data Availability

The labeled data set used to support the findings of this study is available from the author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This work was supported by the Research on Cultural Identity of Ethnic Minority College Students from the Perspective of Ethnic Community (18YJA710039).

References

[1] D. Dockter, "Quantifying the representation of plant communities in the protected areas of the U.S.: an analysis based on the U.S. National vegetation classification groups," *Journal of Forests*, vol. 12, no. 05, pp. 110–112, 2021.
[2] M. Zhao, H. Xiao, and S. Zhou, "Data analysis algorithms for mining online communities from microblogs," *International Journal of Web Based Communities*, vol. 16, no. 2, pp. 1–17, 2020.
[3] M. A. Gualhano and A. Vasconcelos, "Analysis of teaching certification programs in federal schools using data mining," *Educação e Pesquisa*, vol. 46, no. 2, pp. 89–101, 2020.
[4] C. Bratsas, E. Chondrokostas, K. Koupidis, and I. Antoniou, "The use of national strategic reference framework data in knowledge graphs and data mining to identify red flags," *Data*, vol. 6, no. 1, pp. 2–18, 2021.
[5] C. Bratsas, E. Chondrokostas, and K. Koupidis, "The use of national strategic reference framework data in knowledge graphs and data mining to identify red flags," *Data*, vol. 7, no. 05, pp. 110–112, 2021.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 7919963, 8 pages
https://doi.org/10.1155/2022/7919963

Review Article

The Impact of Public Service Motivation on Job Satisfaction in Public Sector Employees: The Mediating Roles of Work Engagement and Organizational Commitment

Dan Lu [1,2] **and Chih Hung Chen** [1]

[1] *International College, National Institute of Development Administration, Bangkok 10240, Thailand*
[2] *School of Finance and Public Administration, Yunnan University of Finance and Economics, Kunming 650221, China*

Correspondence should be addressed to Dan Lu; danlu@ynufe.edu.cn

Received 12 June 2022; Accepted 1 August 2022; Published 22 August 2022

Academic Editor: Xuefeng Zhang

In the public service sector, staff's motivation is of great significance for the developments of both the service team and the governmental service departments. This research develops a model to study the effects of public service motivation on job satisfaction and further assesses the mediation effects of work engagement and organizational commitment. Five-hundred and fifty employees from Chinese public service sectors were engaged in the questionnaire-based study. The results show that public service motivation plays important roles in work engagement, organizational commitment, and job satisfaction. Furthermore, work engagement and organizational commitment are the factors to mediate public service motivation and job satisfaction. Therefore, our findings suggest that staff in the public sector with enhanced public service motivation generally appear more involved in their work, are more loyal to their organizations, and demonstrate more positive work attitudes to improve productivity.

1. Introduction

The public sector's primary duty is to support citizens and organizations with necessary public products and services, in which service constitutes this sector's function and is also one of the most critical obligations in servicing governments. Employees in the public sectors play significant roles in realizing the principal duty, so there is great necessity to study the employees' work attitudes and behaviors in the public sector. The literature shows that work attitudes, behavior, and performance are not related solely to individuals or organizations but also may be attributable to the relationship between individuals and organizations.

There currently exist two theories in this area, which are public service motivation (PSM) theory and the person-organization fit (P-O fit) theory [1–3]. The PSM theory argues that PSM is positively correlated with work attitudes and behaviors and employees with higher PSM have elevated levels of job satisfaction and organizational commitment. Therefore, PSM is considered as a significant individual predisposition to explain the public employees' work attitudes [1, 4–6]. The P-O fit theory considers relationship between person and organization as the key factor that results in the individual's attitudes, behaviors, and others [7, 8]. It assumes that P-O fit completely mediates PSM and work attitudes, while PSM has no straightforward correlation with neither work attitudes nor performance when considering P-O fit [1, 9].

For both PSM and P-O fit studies, work attitudes, which include job satisfaction and organizational commitment, are the mostly frequently examined criteria [7, 10]. There are two distinct views: the first one is that PSM automatically increases public employees' job satisfaction and organizational commitment, while the second holds the standpoint that PSM imposes indirect and positive effect on work attitudes with P-O fit as the mediating factor [9, 11]. Now, there is the trend to accept the first view, but there still exists the research gap and further studies are warranted in the research area.

The PSM theory seeks to stress the sheer weight of public service motivation. Its theoretical assumption is that a stronger public service motivation tends to suggest improved job satisfaction, work performance, and organizational commitment [4]. Given the significant relations among PSM, work attitudes, and work performance, PSM has gradually become an essential element used to explain job satisfaction, work performance, and turnover intention [12, 13], which are associated effectively with individuals' attitudes and organizational behaviors. Additionally, previous studies not only have emphasized the association between PSM and work attitudes, but also have focused on evidence-based research of the correlation between PSM and social behaviors, participation in public serving activities, civil organizing, citizen's organizational behaviors, etc. [14].

In China, the central government has gradually raised its requirements to the civil servants with the purpose of developing a servicing government. In this circumstance, studies on the relationship between the civil servants' PSM and work attitudes would be of great meanings. The 17th and 18th National Congress of the Chinese Communists pointed out clearly that the overall goals for comprehensive reformation in China are to accelerate the reformation of the administration, to establish a servicing government, and to develop and comprehend the socialism with Chinese characteristics to promote the system of national governing and modernism of governing measurements. The staff throughout the public sector are the very executors and operators of this implementation. China is undergoing a new era of "quintuple ages," the age of reforming and opening, the age of cultural renaissance, the age of China's rise, the age of digitization, and the age of globalization. The 19th National Congress clearly illustrated that, by entering a new era, the most critical objective of the administration reformation is to recruit a highly competent and professional team to enable the public sector staff to achieve results and accomplishments that satisfy the citizenry. However, to build a servicing government and to improve the service's quality and capacity, the public service staff's work attitudes must be developed and shaped and their consciousness of altruism needs to be cultivated as well, as these can enable them to perform more satisfying and desirable jobs. In this way, the ultimate goal of reforming the national administration can be achieved eventually.

With most of the theories and findings obtained in the context of Western culture and the special social, politic, and economic conditions of China, there is need to answer the following question: does public service motivation also affect work attitudes in the Chinese cultural context, which Confucianism influences significantly? Therefore, it is essential to investigate public sector staff's work attitudes and their influencing factors in this context based upon the PSM theory. The findings will also contribute to a more rational and comprehensive understanding of PSM's relation to work attitudes. According to the above, given that this work has been conducted with reference to the Western theory of PSM and its measures, public sector staff are used as the subjects to investigate in more depth the influencing mechanism of PSM on work attitudes (job satisfaction) by introducing work engagement and organizational commitment as the intervening variable with the goal to reveal the functional system of PSM that affects work attitudes. By doing so, it may serve as a reference for the human resource departments in the Chinese public sector.

2. Literature Review and Hypothesis Development

2.1. PSM, Job Satisfaction, and Organizational Commitment. Public service motivation, initiated by Perry and Wise [4], is defined as individual's inclining reaction to the main or overall public service departments and organizations. Specifically, an improved level of PSM purports to suggest that the individuals thereof arguably tend to serve in the public sector due to their strong community spirit. Following this concept, more extensive studies have been conducted to establish a more comprehensive and accurate definition of PSM. Vandenabeele [15] contends that PSM needs to be viewed as a belief, value, and attitude that transcends people and organizations' self-interests, focuses instead on political entities' interests, and encourages individuals to engage in appropriate actions at the proper time. Although Rainey and Steinbauer [16] extended the definition of PSM from the public to private sector, they believe that PSM is a general motivation of altruism serving the communities' interests. Currently, PSM theory has become a critical criterion and benchmark to investigate public sector staff's work behaviors and performance, and both domestic and oversea researchers have also tested and applied it widely. Accordingly, it has gained practical value in the field of human resources management in the public sector.

Many practical studies have demonstrated that, with higher PSM levels, individuals' commitment to public organizations leads further to better job satisfaction [17, 18]. Public departments are governmental organizations with the goal to serve the social public interests. Thus, this working environment can fulfill individuals' altruism-related demands and motivation to serve the public's interests, which facilitates the level of job satisfaction of those with a higher PSM [12, 19]. Further, other studies have introduced mediating variables, such as the person-organization fit, organizational characteristics, and department variances, and the results have shown the indirect effects of public sector staff's PSM on their job satisfaction, which can be strengthened or weakened by the variables above [20]. Domestic studies have demonstrated that public sector staff's PSM affects their job satisfaction positively [21, 22] and the effect can be more obvious, especially when it comes to public spirits and self-sacrifice. In addition, it also functions significantly as a mediator between the value of work and job satisfaction [23].

Generally, organizational commitment concerns a mental state of the relationship between the staff and the organization or a psychological attachment to certain organizations, which reflects the consistency of staff and organizations' value directly [24]. From the perspective of person-organization fit, the compatibility of people and working environment is reached largely through the consistency fit and supplementary fit. While PSM is a general altruism motivation to serve the public's interests, it promotes the consistency of people and organizations' value in serving the public [16]. Further, public organizations also fulfill the internal demands, expectations, and preferences of the individuals who serve society [25]. Kim [1] states that higher level of PSM tends to reflect higher level of loyalty and emotional attachment to their departments, such that PSM can facilitate the staff's organizational commitment. Plenty of studies have verified the positive associations between PSM and organizational commitment [16]. In the same vein, this paper proposes the fact that altruism motivation can also be a function of a higher PSM in the public sector, which would then lead to stronger organizational attachment and greater commitment.

Based upon the studies above, PSM obviously serves to predict job satisfaction and organizational commitment. That is, on the one hand, a higher level of PSM can improve job satisfaction and organizational commitment, and, on the other hand, it may reduce or eliminate (if possible) turnover intentions. With this end in view, this paper formulates the following two hypotheses:

Hypothesis 1. Job satisfaction is significantly and positively conditioned by PSM.

Hypothesis 2. Organizational commitment is significantly and positively conditioned by PSM.

2.2. Work Engagement and Job Satisfaction. Work engagement, which has enormously attracted academic attention, was first proposed by in 1990 and excited great many positive psychological studies in work-related areas. However, many other researchers also believe that there are both similarities and differences among work engagement and involvement, job satisfaction, and organizational commitment [26], exploring the relation between engagement and work attitudes. In an empirical study that engaged 343 nurses, Mahiro et al. [27] discovered that nurses with higher work engagement worked more efficiently and were more willing to engage in their work and provide people-oriented services constantly. Zeffane and Melhem's study [28] suggested that work engagement can not only serve as a valuable indicator of people's job satisfaction and turnover intentions, but also serve as a mediating factor among organizational commitment, performance, and job satisfaction. Mahiro et al. [27] and other researchers have claimed that work-related organizational commitment and job satisfaction are positively and significantly conditioned by work engagement, whereas turnover intentions are negatively conditioned by work engagement. Further, team leaders' work engagement can improve employees' job satisfaction considerably and eliminate turnover intention effectively [29]. According to this, work engagement serves the function of predicting staff's job satisfaction, organizational commitment, and other work-related behaviors. Given the above, this study proposes the following hypotheses:

Hypothesis 3. Job satisfaction is significantly and positively conditioned by work engagement.

Hypothesis 4. Job satisfaction is significantly and positively conditioned by organizational commitment.

2.3. Work Engagement and Organizational Commitment as Mediators. In recent years, more and more researchers have focused on the critical theoretical value and practical potential of work engagement in the scope of organizational behavior studies and have been trying to demonstrate what Kahn [30] concluded: individuals with greater work engagement will dedicate themselves to their behavioral roles and try to express and present themselves physically, consciously, emotionally, and spiritually. Additionally, empirical analysis of survey data collected from central government agencies in the Republic of Korea found that organizational commitment acts as a mediator between PSM and job satisfaction [31]. This study engaged participants from the Chinese public sector to investigate whether work engagement and organizational commitment mediate PSM and job satisfaction. On the one hand, PSM has an "altruistic" nature and helps improve public sector staff's job satisfaction and eliminate turnover intentions. On the other hand, work engagement and organizational commitment are a positive, comprehensive, and work-related status that enhances staff's enthusiasm to work and therefore improves their work performance overall [32]. As a result, it can be said that PSM affects not only public sector staff's work attitudes directly, but also work attitudes indirectly via work engagement and organizational commitment. Hence, this study proposes the following hypotheses:

Hypothesis 5. Work engagement is significantly and positively conditioned by PSM.

Hypothesis 6. Work engagement acts as a mediator of PSM and job satisfaction.

Hypothesis 7. Organizational commitment acts as a mediator of PSM and job satisfaction.

3. Methods

Public sector staff were the main participants in this study, and the samples included civil servants and staff from public institutions and organizations. Questionnaires were distributed to MPA students at 4 universities in the Yunnan and Guizhou Provinces. We chose subjects from these two provinces owing to the approachability of the universities there. The universities' education centers coordinated the study, and the questionnaires were distributed and collected during the MPA classes. During the process, 600 questionnaires were distributed and 580 collected. Five-hundred and fifty questionnaires were retained after excluding 30 invalid ones (response rate 91.67%).

The percentage of males and females in all samples were 52% and 48%, respectively; with respect to age, 6.7% participants were under 25 years old, 35.5% were 25–30, 39.5% were 31–40, and 18.4% were 41 and above. 12.9% of the participants had college degrees and lower, 57.1% held bachelor's degrees, and 30% held master's degree and higher. With respect to work experience, 31.3% had 5 years or less, 24% had 6–10 years, 15.6% had 11–20 years, and 29.1% had 21 years or more. 59.5% were clerical staff, 26.9% were at the office-level, and 13.6% were at the county level and above.

Perry's PSM scale [33] was used. Examples of its 5 items include "meaningful public serving matters to me" and "I am prepared to better serve the society." Kanungo's work engagement scale [34] was employed, which includes "I have been 100% dedicated to my job" and 3 other items. Boateng and Hsieh's scale [35], which includes four items, was used to assess the public employee's job satisfaction. Sample items comprise "overall, I am satisfied with my current job" and "no matter what, I will not leave my current job." Organizational commitment was evaluated with a four-item scale that Meyer et al. [36] developed and validated. Sample items include such statements as "I feel emotionally attached to my organization."

The five-point Likert scale: 'strongly disagree' (1), 'disagree' (2), 'neither disagree nor agree' (3), 'agree' (4), 'strongly agree' (5) was employed for the quantifications. Cronbach's alphas were to 0.875, 0.793, 0.838, and 0.844, respectively.

4. Results

4.1. Descriptive Analysis. Table 1 tabulates the mean, standard errors, correlations, Cronbach's alpha values, and others of the study variables. The correlation matrix shows that the correlation coefficients (r) between PSM and job satisfaction, organizational commitment, and work engagement are, respectively, 0.356, 0.439, and 0.424, with all p values smaller than 0.01. These r values indicate positive relations, providing strong evidence for Hypotheses 1, 2, and 5. The positive correlation of organizational commitment ($r = 0.698$, $p < 0.01$) and work engagement ($r = 0.608$, $p < 0.01$) with job satisfaction strongly supports Hypotheses 3 and 4. Additionally, Cronbach's alpha values range from 0.793 to 0.875, which outweigh the minimum threshold value (MTV) of 0.70, suggesting the study scale's reliability.

4.2. Confirmatory Factor Analysis. The study employs the confirmatory factor analysis (CFA) to test the four-factor model's goodness of fit. Three fit indices, which are comparative fit index (CFI), Tucker-Lewis Index (TLI), and root mean square error of approximation (RMSEA), were used to examine the model fit. The present study attempts to propose five alternative model conceptualizations (Table 2). The hypothesized four-factor model fits good to the data (χ^2 (96) = 212.845, $p < 0.001$; CFI = 0.975, IFI = 0.975, TLI = 0.969, and RMSEA = 0.047) when comparing to the five alternative models. When it comes to the one-factor mode, the fit is less adequate (χ^2 (104) = 23.266, $p < 0.001$; CFI = 0.690, IFI = 0.691, TLI = 0.642, and RMSEA = 0.160). The CFA results demonstrated the proposed four-factor model's construct distinctiveness.

4.3. Test of Common Bias Methods. As this research was conducted using self-reported data, it may include common bias methods [37, 38]. The common bias methods, such as anonymity, forward, and backward integral, are used. Further, CFA was used to test the participants' self-evaluations, and model fit is unsatisfactory (χ^2 (104) = 23.266, $p < 0.001$; CFI = 0.690, IFI = 0.691, TLI = 0.642, and RMSEA = 0.160). Therefore, the presence of severe common bias methods is unlikely. Then, using common method factors to test the common method variances in comparison to the hypothesized four-factor model, $\blacktriangle X^2/df = 0.025$, $\blacktriangle$CFI = −0.006, $\blacktriangle$IFI = −0.006, $\blacktriangle$TLI = −0.002, and $\blacktriangle$RMSEA = 0.002, in which the variance of all fit indices was smaller than 0.03, indicated that the modeling had not been optimized significantly after common method factors were added and no noticeable common bias methods were found in the testing [39].

4.4. Hypothesis Tests. The data's fit was evaluated by Structural Equation Modeling (SEM) as there is necessity to verify the model proposed to test the hypothesis (Figure 1). With χ^2 (99) = 372.634, $p < 0.001$, CFI = 0.942, TLI = 0.929, and RMSEA = 0.071, there are excellent fit indices for the model. The CFI and TLI were both higher than 0.9, and the RMSEA was equal to or less than 0.08. As shown in Tables 3 and 4, PSM affected JS positively ($p < 0.001$, standardized coefficient (SC) = 0.424), which lent strong support to Hypothesis 1; PSM affected OC positively ($p < 0.001$, SC = 0.420), which lent strong support to Hypothesis 2. The direct relation between WE and job satisfaction ($p < 0.001$, nonstandardized coefficient (NSC) = 0.511) and between organizational commitment and job satisfaction ($p < 0.001$, NSC = 0.402) was both significant and positive, which supported Hypotheses 3 and 4. Further, PSM affected WE positively ($p < 0.001$, NSC = 0.439).

A bootstrapping method was adopted to test the indirect and mediating effects in the path structural equation model. Table 4 illustrates that PMS imposes statistically noteworthy indirect effect on job satisfaction (SE = 0.034; CI = 0.243 and 0.350; the CI did not include zero), which is through work engagement (PS ⟶ WE ⟶ JS). In this sense, work engagement serves as a mediator between PSM and job satisfaction fully, which provided strong evidence for Hypothesis 6. Further, PSM's indirect effect on job satisfaction through organizational commitment (PSM ⟶ OC ⟶ JS) was statistically robust (SE = 0.033; CI = 0.225 and 0.335), justifying the role of organizational commitment as a mediator between PSM and JS. Accordingly, Hypothesis 7 was supported as well.

Table 1: Means, standard errors, correlations, and Cronbach's alpha values of the study variables.

Var	Mean	SE	Gender	Age	Edu	Years	PSM	WE	OC	JS
Gender	0.520	0.500	1							
Age	2.856	1.129	0.094*	1						
Edu	3.118	0.758	0.035	0.073	1					
Years	2.784	1.637	0.113**	0.689**	0.297**	1				
PSM	3.700	0.680	0.072	0.014	0.025	−0.049	(0.875)			
WE	3.350	0.813	−0.076	−0.251**	−0.149**	−0.400**	0.356**	(0.793)		
OC	3.721	0.657	−0.049	−0.144**	−0.068	−0.283**	0.439**	0.534**	(0.844)	
JS	3.502	0.717	−0.065	−0.129**	−0.086*	−0.278**	0.424**	0.608**	0.698**	(0.838)

Note. $N = 550$, $^{*}p < 0.05$, and $^{**}p < 0.01$. SE means standard error; Var means variable; Edu means education; PSM means public service management; WE means work engagement; OC = meansorganization commitment; JS means job satisfaction.

Table 2: Comparison of measurement models for variables studied.

Model	X^2	df	X^2/df	CFI	IFI	TLI	RMSEA
Hypothesized four-factor model	212.845	96	2.217	0.975	0.975	0.969	0.047
Three-factor model							
PSM combined with WE	23.266	18	1.239	0.997	0.997	0.996	0.023
PSM combined with JS	56.171	25	2.247	0.987	0.987	0.981	0.048
PSM combined with OC	54.089	24	2.254	0.988	0.988	0.982	0.480
One-factor model (PSM, JS, and OC combined)	159.575	61	2.616	0.975	0.975	0.967	0.054
One-factor model (all constructs combined)	1561.730	104	15.017	0.690	0.691	0.642	0.160

Note. $N = 550$.

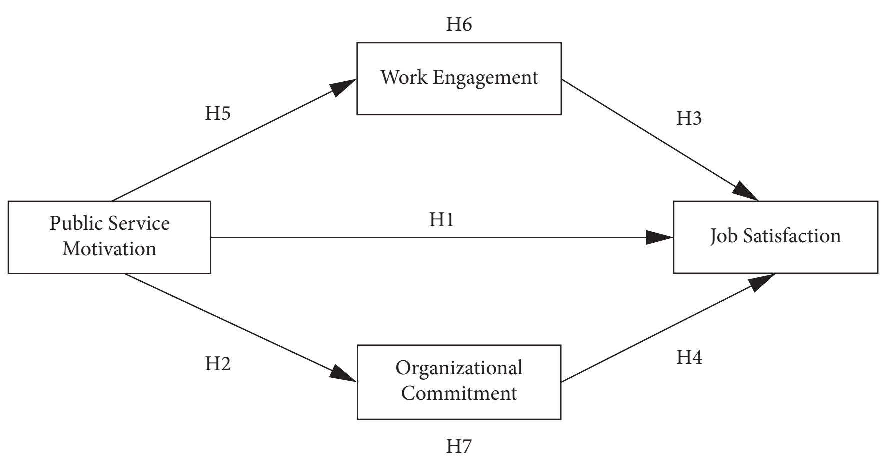


Figure 1: Research structure and hypotheses.

Table 3: Path test.

Path	Standardized coefficient (SC)	Nonstandardized coefficient (NSC)	SE	CR	p
PSM ⟶ OC	0.420	0.464	0.043	10.855	***
PSM ⟶ WE	0.432	0.403	0.036	11.216	***
PSM ⟶ JS	0.426	0.432	0.033	11.649	***
WE ⟶ JS	0.490	0.511	0.032	15.877	***
OC ⟶ JS	0.455	0.402	0.027	14.857	***

Table 4: Standardized bootstrap mediation effect test.

Path	Effect value	SE	Bias-corrected 95% CI			Percentile 95% CI		
			Lower	Upper	p	Lower	Upper	p
PSM ⟶ WE ⟶ JS	0.294	0.034	0.243	0.350	0.002	0.243	0.350	0.002
PSM ⟶ OC ⟶ JS	0.278	0.033	0.225	0.335	0.002	0.222	0.333	0.002

5. Conclusions

This study included 550 employees in the public sector as participants, proposed the mechanism which specifies the influence of PSM on job satisfaction, and presented the roles of work engagement and organizational commitment as mediators. Accordingly, the research concluded that the following: (1) job satisfaction is significantly and positively conditioned by PSM; (2) job satisfaction is significantly and positively conditioned by work engagement and organizational commitment, and (3) work engagement and organizational commitment mediate the relations between PSM and job satisfaction.

5.1. Theoretical and Practical Implications. Based upon the conclusions above, management should be aware of the following: (1) when the public sector realizes the practical value of PSM theory, it can improve its staff's level of PSM in various ways. Firstly, HR departments in the public sector should consider including PSM in their recruiting process and measure and test the potential candidates' PSM [40]. Currently, the main recruitment approach in China's public sector continues to be tests of writing documentation and administrative ability, which cannot detect and reveal the candidates' actual motivation. Under this circumstance, it would be difficult to shortlist candidates with the "be the first to show concern and the last to enjoy oneself" spirit, which would affect their work attitudes and professional ethics even further thereafter. In addition, establishing PSM should also be emphasized in the new staff training procedure. This study demonstrated that PSM plays a critical and direct role in job satisfaction, which can help the public sector's HR departments acknowledge PSM theory's value and provide solid evidence of the necessity to establish PSM in public service departments' staff; (2) work engagement and organizational commitment in the public sector can be promoted by cultivating employees' desire to serve the public. This research demonstrated that work engagement and organizational commitment not only affect job satisfaction directly, but also serve to mediate the relations between PSM and job satisfaction in part or entirely. Firstly, the public service spirit should be encouraged actively. Service is the most fundamental function of a servicing government. As suggested by Wilson [41], the question we are facing and merits government officers' attention, is whether the public service staff dedicate the most mind and efforts, and serve with their consciousness, not only to their superiors but also to the society. From the perspective of building a servicing government, the concept of "public servant" to which the Chinese government has always referred is consistent and is associated with the PSM theory that Perry developed [33]. They share the same goal to establish the public serving spirit in the public sector, promote altruism in service, and facilitate work engagement further to accomplish tasks more effectively. Secondly, public workers' professional ability and service quality should be improved via professional training. Currently, most system reformation has led to large scale government organizations' adjustment and redistribution of staff, which has affected both their work efficiency and work content. The public sector should also adapt to the trend of reforming to address the challenges of governmental systems' reform positively, optimize the training of occupational skills and quality, and improve the staff's adaptability [41]. Ultimately, this will help them adapt to their new posts and enhance their work engagement and organizational commitment further; (3) with the development of reasonable and diverse incentive schemes, public workers' job satisfaction can be improved. With the background of the comprehensive system reform, major shifts, including the public sector's wage standard, pension integration, streamlining duties, and titles, have been promoted gradually, and these reforms have had a noticeable positive effect on public workers' work attitudes and organizational citizenship behaviors. The public sector considers the involvement of both internal and external incentives to build more diverse remuneration schemes against the background of administration system reform. On the one hand, they should use effective external motivating methods to adjust the entire remuneration structure and attempt to establish the structure based upon performance and targets [34]. With the distinct proportion of basic and bonus incentives, a stage of "different efforts for different pay" can be reached, which will promote younger employees' motivation and involvement effectively. On the other hand, in addition to emphasizing external material incentives, the public sector should concentrate more on internal incentives to establish an administration culture motivated by the value of service [42]. Because of the traditional incentive theory's overemphasis on external material incentives, it monitors its own internalizing process, which contributes largely to the "isolating effects" that compromise the function of internal incentives. Therefore, the public sector can promote employees' PSM via recognition prizes, promotion, motivation of reputation, sense of competency, self-authority, etc. These methods would increase job satisfaction and organizational commitment and also balance the internal and external incentives, thereby avoiding the negative isolating effects from external to internal incentives [43].

5.2. Limitations and Future Research. Admittedly, further research is warranted to address some of the limitations associated with the current study. Firstly, the participants in this paper were limited to public service workers in only two provinces (Yunnan and Guizhou), and a more extensive regional effort is required for future studies that could even be extended to cross-regions and cross-industries research in the private sector. Secondly, as PSM uses Western scales, the compatibility with Chinese scenarios may make the effects in the results unclear. Therefore, a fully localized measurement scale that encompasses the Chinese cultural and political background is required. Thirdly, the measures used in this study primarily address research variables on the individual level, while there are a different organizational culture and workload among different public departments; thus, variables on the organizational level need to be introduced to assess the way in which PMS affects work attitudes via cross-level testing methods.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] S. Kim, "Does person-organization fit matter in the public -sector? Testing the mediating effect of person-organization fit in the relationship between public service motivation and work attitudes," *Public Administration Review*, vol. 72, no. 6, pp. 830–840, 2012.

[2] Q. Miao, N. Eva, A. Newman, and G. Schwarz, "Public service motivation and performance: the role of organizational identification," *Public Money & Management*, vol. 39, no. 2, pp. 77–85, 2019.

[3] W. Vandenabeele, A. Ritz, and O. Neumann, "Public service motivation: state of the art and conceptual cleanup," in *Proceedings of the palgrave handbook of public administration and management in europe*, E. Ongaro and S. Van Thiel, Eds., Palgrave Macmillan UK, London, pp. 261–278, November 2018.

[4] J. L. Perry and L. R. Wise, "The motivational bases of public service," *Public Administration Review*, vol. 50, no. 3, pp. 367–373, 1990.

[5] S. Castaing, "The effects of psychological contract fulfilment and public service motivation on organizational commitment in the French civil service," *Public Policy and Administration*, vol. 21, no. 1, pp. 84–98, 2006.

[6] S. T. T. Teo, D. Pick, M. Xerri, and C. Newton, "Person-organization fit and public service motivation in the context of change," *Public Management Review*, vol. 18, no. 5, pp. 740–762, 2016.

[7] A. L. Kristof-Brown, R. D. Zimmerman, and E. C. Johnson, "Consequences of individuals' fit at work: a meta-analysis of person–job, person–organization, person–group, and person–supervisor fit," *Personnel Psychology*, vol. 58, no. 2, pp. 281–342, 2005.

[8] S. Hassan, M. Fiaz, and R. Asif, "Public service motivation and quality of working life in public institutions: analyzing the moderating role of po fit," *Journal of Management Research*, vol. 8, no. 2, pp. 159–181, 2021.

[9] L. Bright, "Does person-organization fit mediate the relationship between public service motivation and the job performance of public employees?" *Review of Public Personnel Administration*, vol. 27, no. 4, pp. 361–379, 2007.

[10] S. K. Pandey, B. E. Wright, and D. P. Moynihan, "Public service motivation and interpersonal citizenship behavior in public organizations: testing a preliminary model," *International Public Management Journal*, vol. 11, no. 1, pp. 89–108, 2008.

[11] W. Vandenabeele, "The mediating effect of job satisfaction and organizational commitment on self-reported performance: more robust evidence of the PSM—performance relationship," *International Review of Administrative Sciences*, vol. 75, no. 1, pp. 11–34, 2009.

[12] D. Coursey, K. Yang, and S. K. Pandey, "Public service motivation (psm) and support for citizen participation: a test of perry and vandenabeele's reformulation of psm theory," *Public Administration Review*, vol. 72, no. 4, pp. 572–582, 2012.

[13] J. Taylor, "Public service motivation, civic attitudes and actions of public, nonprofit and private sector employees," *Public Administration*, vol. 88, no. 4, pp. 1083–1098, 2010.

[14] D. J. Houston, "Walking the walk" of public service motivation: public employees and charitable gifts of time, blood, and money," *Journal of Public Administration Research and Theory*, vol. 16, no. 1, pp. 67–86, 2006.

[15] W. Vandenabeele, "Toward a public administration theory of public service motivation," *Public Management Review*, vol. 9, no. 4, pp. 545–556, 2007.

[16] H. G. Rainey and P. Steinbauer, "Galloping elephants: developing elements of a theory of effective government organizations," *Journal of Public Administration Research and Theory*, vol. 9, no. 1, pp. 1–32, 1999.

[17] Y. Choi and I. H. Chung, "Effects of public service motivation on turnover and job satisfaction in the u.S. Teacher labor market," *International Journal of Public Administration*, vol. 41, no. 3, pp. 172–180, 2018.

[18] A. M. Kjeldsen and L. B. Andersen, "How pro-social motivation affects job satisfaction: an international analysis of countries with different welfare state regimes," *Scandinavian Political Studies*, vol. 36, no. 2, pp. 153–176, 2013.

[19] B. Liu, T. L.-P. Tang, and K. Yang, "When does public service motivation fuel the job satisfaction fire? The joint moderation of person–organization fit and needs–supplies fit," *Public Management Review*, vol. 17, no. 6, pp. 876–900, 2015.

[20] A. M. Kjeldsen and J. R. Hansen, "Sector differences in the public service motivation–job satisfaction relationship: exploring the role of organizational characteristics," *Review of Public Personnel Administration*, vol. 38, no. 1, pp. 24–48, 2016.

[21] Y. Ma and X. Leng, "The efffect of public service motivation on job satisfaction of tax service officers," *Taxation Research*, vol. 1, pp. 106–113, 2018.

[22] W. Mao, D. Zheng, and Y. Chen, "The effect of public service motivation on job satisfaction of civil servants," *Journal of Shandong Academy of Governance*, no. 4, pp. 15–20, 2017.

[23] X. Liu, X. Wang, and W. Jiang, "The impact of Chinese civil servants' work values on job satisfaction," *Chinese Public Administration*, vol. 12, pp. 83–88, 2016.

[24] N. J. Allen and J. P. Meyer, "Organizational commitment: evidence of career stage effects?" *Journal of Business Research*, vol. 26, no. 1, pp. 49–61, 1993.

[25] W. Potipiroon and M. T. Ford, "Does public service motivation always lead to organizational commitment? Examining the moderating roles of intrinsic motivation and ethical leadership," *Public Personnel Management*, vol. 46, no. 3, pp. 211–238, 2017.

[26] A. J. Wefald and R. G. Downey, "Construct dimensionality of engagement and its relation with satisfaction," *Journal of Psychology*, vol. 143, no. 1, pp. 91–112, 2009.

[27] S. Mahiro, N. Takashi, and N. Satoko, "Work engagement and attitudes toward caring for dying patients and families among home-visiting nurses in Japan," *International Journal of Palliative Nursing*, vol. 20, no. 7, pp. 343–348, 2014.

[28] R. Zeffane and S. J. B. Melhem, "Trust, job satisfaction, perceived organizational performance and turnover intention," *Employee Relations*, vol. 39, no. 7, pp. 1148–1167, 2017.

[29] Y. Wang and J. Qin, "The influence of kindergarten teacher's work engagement on their work attitude and mental health," *Studies in Early Childhold Education*, no. 2, pp. 56–63, 2015.

[30] W. A. Kahn, "Psychological conditions of personal engagement and disengagement at work," *Academy of Management Journal*, vol. 33, no. 4, pp. 692–724, 1990.

[31] S. Park, "Determinants of the job satisfaction of public officials: testing the mediation effect of organizational commitment," *Public Organization Review*, vol. 20, no. 4, pp. 665–684, 2020.
[32] Z. Guo, B. Xie, and Y. Guo, "The effects of supervisor' work engagement on followers' job attitudes: examining a multilevel model," *Human Resources Development of China*, no. 5, pp. 67–75, 2016.
[33] J. L. Perry, "Measuring public service motivation: an assessment of construct reliability and validity," *Journal of Public Administration Research and Theory*, vol. 6, no. 1, pp. 5–22, 1996.
[34] R. N. Kanungo, "Measurement of job and work involvement," *Journal of Applied Psychology*, vol. 67, no. 3, pp. 341–349, 1982.
[35] F. D. Boateng and M.-L. Hsieh, "Explaining job satisfaction and commitment among prison officers: the role of organizational justice," *The Prison Journal*, vol. 99, no. 2, pp. 172–193, 2019.
[36] J. P. Meyer, N. J. Allen, and C. A. Smith, "Commitment to organizations and occupations: extension and test of a three-component conceptualization," *Journal of Applied Psychology*, vol. 78, no. 4, pp. 538–551, 1993.
[37] K.-P. Gan, Y. Lin, and Q. Wang, "Public service motivation and turnover intention: testing the mediating effects of job attitudes," *Frontiers in Psychology*, vol. 111289 pages, 2020.
[38] L. L. Zhou Hao, "Statistical remedies for common method biases," *Advances in Psychological Science*, vol. 12, no. 6, p. 942~950, 2004.
[39] S. Liu, K. Liu, T. Li, and L. Lu, "The effect of mindfulness on subjective well-being of colleage students: the mediating effects of emotion regulation and resilience," *Journal of Psychological Science*, vol. 38, no. 4, pp. 889–995, 2015.
[40] X. Tan, B. Liu, and Y. Wang, "Effect of inducements-contributions oriented psychological contracts' difference on public servants' turnover intention based upon the comprehensive analysis of public service motivation and attitudes to change," *Journal of Public Management*, vol. 14, no. 4, pp. 27–43, 2017.
[41] W. Wilson, *The Study of Administration*, CPC Central Party Press, Beijing, China, 1997.
[42] S. Zhang and F. Kong, "On the shaping of the spirit of public service from the perspective of public service motivation," *Nanjing Social Sience*, vol. 11, pp. 75–80, 2016.
[43] L. Ge, "The effect of transformational leadership on the work attitude of civil servants—a study of the mediating effect of public service motivation," *Journal of Yantai University*, vol. 29, no. 3, pp. 111–120, 2016.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 4152637, 15 pages
https://doi.org/10.1155/2022/4152637

Research Article

Correspondence Model of Human Resource Management and Marketing Based on Genetic Algorithm

Yulin Bi[1] and Ming Tang [2]

[1] *Catholic University, Daegu 42708, Republic of Korea*
[2] *Yeungnam University, Daegu 42708, Republic of Korea*

Correspondence should be addressed to Ming Tang; tangming717@163.com

Received 13 May 2022; Accepted 29 July 2022; Published 22 August 2022

Academic Editor: Jiafu Su

In order to realize the optimal allocation of human resources and avoid the waste or relative shortage of human resources, human resources and marketing based on the genetic algorithm are combined together and a combined evaluation management model is established in this study. By constructing a weight vector to represent the evaluation performance of different evaluation models for talent evaluation, the combination problem of multiple evaluation models is transformed into an optimization problem of weights. Based on the evaluation accuracy, the fitness function is designed, the weight vector is optimized by the genetic algorithm, and the individual selection strategy is designed to avoid falling into local optimization. The validity of the model and algorithm is verified by a numerical example, and the calculation results show that the proposed method can gradually improve the average working ability of employed employees by reasonably controlling the number of employment at different periods and the number of dismissals, and the average working ability of employees can be improved by 41%, thus realizing the optimization of human resources for the project.

1. Introduction

Human resource management (referred to as HRM), the upgrading of personnel management, refers to the effective use of relevant human resources inside and outside the organization through recruitment, selection, training, compensation, and other management forms under the guidance of economics and humanistic thinking. It is a general term for a series of activities to meet the needs of the current and future development of the organization, and to ensure the realization of the organization's goals and the maximization of member development. In the era of knowledge economy, human resources have gradually become an important competitive factor after natural resources and capital. The research on human resource management theories and methods has been deepened, and the development and management of human resources have been comprehensively carried out, including human resource costs, value and quality analysis, talent forecast, allocation, and planning. [1]. Among them, human resource evaluation is an important link and basic work. With the intensification of competition and the continuous in-depth application of artificial intelligence technology, experts and scholars have established a series of evaluation technologies and models for different application environments, as shown in Figure 1. Statistical investigation and analysis, AHP, neural network, expert system, fuzzy logic, gray system, and other methods are used to improve management efficiency and enhance the competitiveness of the organization. From the perspective of managers who can manage human resources, it is not easy to find that there are many uncertainties in the management process. Especially compared with the traditional humanized management, the modern humanized management follows the general pattern of the times. We use scientific management to give full play to people's potential, improve efficiency, and maximize goals [2]. It is the whole process of predicting the human resource needs of an organization and making a manpower demand

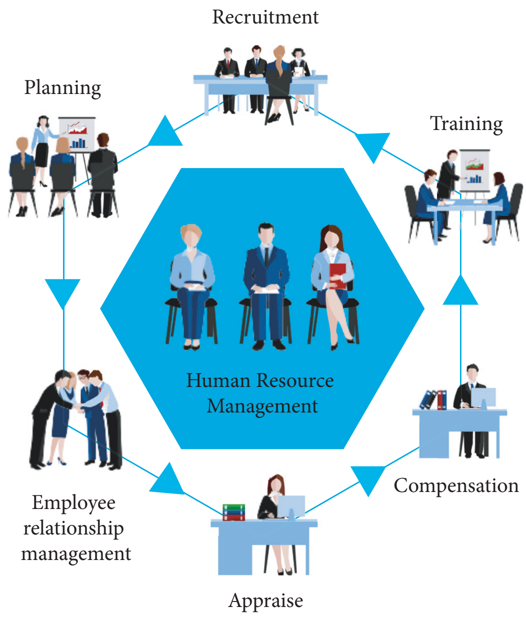


Figure 1: Human resource model.

plan, recruiting and selecting personnel and conducting effective organization, evaluating performance and paying compensation and conducting effective incentives, and combining organizational and individual needs for effective development in order to achieve optimal organizational performance. It is also an important position in the company. Previous human resource management experience may fall short of expectations, and a mathematical model based on multiple assessments is an indicator of future management performance improvement. Human management algorithms include the following: ant colony algorithm, decision tree algorithm, and genetic algorithm. But this process is based on static analysis of simulated HR data. In the face of difficult data and variable heterogeneity of business processes, the efficient operation of human resources cannot be achieved using the same algorithms. Talent evaluation is a multiobjective and multiattribute comprehensive evaluation. Various factors influence and restrict each other. Different evaluation models and methods have different adaptability. Talent assessment is "the activity of measuring and evaluating people's basic quality and performance through a series of scientific means and methods," and it is applied in the fields of enterprise management such as organizational development and talent management. The specific object of talent evaluation is not an abstract person, but the inherent quality and performance of a person who exists as an individual. The essence of human resource portfolio assessment in this study is as follows: we construct a comprehensive model based on various observation data to obtain a comprehensive assessment of each talent. The index system and weighted score are the commonly used methods [3]. However, in actual work, each indicator and its weight cannot be dynamically adjusted, and the preset indicator system and weight cannot reflect the best combination model for each individual under different circumstances. Therefore, it is necessary to construct a dynamic combination model that can adaptively search for appropriate combination weights, thereby improving the accuracy of evaluation. In this study, a combined testing model based on the genetic algorithm is proposed to improve the accuracy of testing, and experiments are designed to test the performance of model construction.

2. Literature Review

Chen et al. believe that the traditional human resource management model is based on job analysis, which can play a role in a simple management model, but with the development of the times, human resource management has been endowed with more functions. With attributes, the traditional model has been unable to provide a sustainable guarantee for human resource management in the new situation. At present, the limitations of the human resource management model based on the job analysis model widely used by enterprises have gradually become prominent [4]. In order to adapt to the modern enterprise human resource management model and the enterprise's demand for talents, through continuous research and exploration by Emiroglu and Uyaroglu et al., a new type of competency model emerges as the times require, which fully and closely combines human resources and organizational relationships to the quality requirements of the employees of the enterprise as the composition of the model, through a series of qualities required by human resource managers, modeling by a combination of various factors, to develop a set of competency models suitable for different enterprises [5]. Martowibowo and Kemala Damanik put forward different concepts of internal marketing in the research process of service marketing. With the continuous development and innovation of internal marketing practice and theory, the meaning of internal marketing is also constantly evolving, becoming richer and more perfect [6]. Chen stated that the evolution of the meaning of internal marketing has mainly gone through the following three stages: in the first stage, internal marketing is defined from the viewpoint that employees are customers. Its main point of view is to regard employees as internal customers, emphasizing employee motivation and employee satisfaction [7]. Manita et al. believe that since employees are the key factor in improving service quality, in order to improve service quality, companies need to motivate and satisfy employees, and treating employees as customers is an effective way to improve employee satisfaction [8]. Nayana et al. first advocated that enterprises should treat employees as customers and believed that employees are an important market for service enterprises. The second stage is the customer-oriented internal marketing concept. Internal marketing means that the service company must effectively train and motivate the staff and all auxiliary service personnel who are in direct contact with customers, so that they can work together to provide customers with satisfactory services. For a company that consistently delivers high-quality service, marketers must get everyone in the company to execute a customer-focused strategy. In the research on internal marketing, some

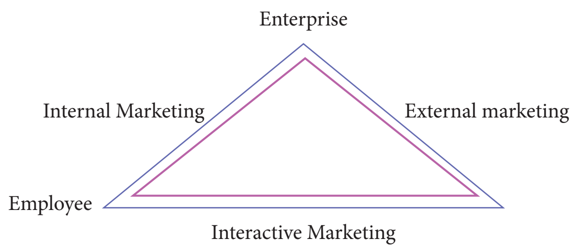


Figure 2: Service marketing triangle.

scholars realize that only focusing on the internal market will lead to "marketing myopia," and point out that the purpose of internal marketing is to cultivate employees' customer orientation and service awareness to achieve customer satisfaction, and internal marketing is to serve external marketing, so as to organically link internal marketing and external marketing [9]. Xing and Zhang believe that in interactive marketing, frontline employees can not only influence consumers' current and future purchasing behavior through direct contact with customers, but also conduct marketing activities to consumers to increase the market share of enterprises. It is not enough for enterprises to make employees complete their own work, but also to motivate employees, so that they have a customer-oriented concept and a sense of active sales, and become part-time marketers. In the third stage, internal marketing is a tool to promote corporate strategy implementation and change management [10]. Samanta et al. found in their research that internal marketing has the function of multifunctional integration and reduces resistance to organizational change, and proposed that internal marketing should be used as a management tool to achieve organizational strategic goals [11]. Al-Amin et al. pointed out the role of internal marketing in integrating HR functions and marketing [12]. Indian and Bhatia believe that internal marketing is a planned effort to overcome organizational change obstacles and guide and integrate employees to effectively implement company strategies [13]. Internal marketing is a process of continuously sharing information with employees and recognizing their contributions. This continuous process is the basis for building a healthy corporate culture in which employees follow the philosophy of "I am for everyone, everyone is for the customer." Continuous internal marketing is also the cornerstone of building a world-class company.

3. Methods

3.1. Introduction to Human Resource Management and Marketing. In the process of research on service marketing, someone proposed the service marketing triangle theory, as shown in Figure 2. This theory believes that service marketing is composed of three parts, namely, external marketing, interactive marketing, and internal marketing, and the three must be closely coordinated to achieve the effective design and smooth delivery of services to build loyal customer relationships [14]. Among them, external marketing is a variety of activities in which enterprises make service-related commitments to customers, including service products, service pricing, service channels, service integration and communication, service processes, service personnel, and tangible display. Through external marketing, the company indicates to the customer that the company will provide the customer with a specific service product at a certain price and through a certain distribution channel. These commitments set the customer's expectations. Interactive cooperative marketing is an original new marketing model, referred to as ICM. Interactive marketing refers to the skills of employees in serving customers. Interactive marketing is a process in which service personnel and customers interact and interact with each other in the process of contact. In interactive marketing, employees become marketers and are required to have service awareness and active sales awareness, to provide customers with high-quality services, to fulfill their promises to customers, and to maintain long-term relationships with customers. Whether employees can fulfill their promises depends on their willingness and ability to serve [15]. In order to enable employees to establish correct concepts and have the necessary skills to provide customers with corresponding services to fulfill the company's commitment, companies must conduct internal marketing to employees. The service marketing triangle theory shows that internal marketing is a prerequisite for the success of external marketing and interactive marketing, and reveals that enterprises should conduct internal marketing before external marketing.

A professor of a business school proposed the service profit chain theory, showing a concise and incomparably clear method to increase the profit of service enterprises. As shown in Figure 3, the service profit chain shows that if the enterprise can provide employees with good internal service quality, employee satisfaction will increase, employees who are satisfied with the enterprise will provide more service value to customers, the increase in service value will win more customer satisfaction, and improved customer loyalty will ultimately increase corporate profits and enhance corporate profitability [16]. In short, the service profit chain shows that employee satisfaction will bring more satisfied customers and ultimately more profits to the enterprise. This has been confirmed by the practice of some foreign enterprises. Financial services benefit from the idea of proposing the idea that business value and trust are increased by improving customer satisfaction and efficiency, and that customer satisfaction and loyalty are increased through service value and service value from customers. Design and supply are based on employee satisfaction and professionalism, which are determined by the quality of the company's internal services to its employees [17]. The service profit chain can be vividly understood as a link between "profitability, customer loyalty, employee satisfaction, and loyalty and productivity. It is a closed chain of circular action, in which the implementation quality of each link is the same." It will directly affect the subsequent links, and the ultimate goal is to make the enterprise profitable. Internal marketing can satisfy employees by improving the quality of internal services, and ultimately achieve corporate profit growth and capacity enhancement through a series of transmission mechanisms. The theory reveals that employees are the key

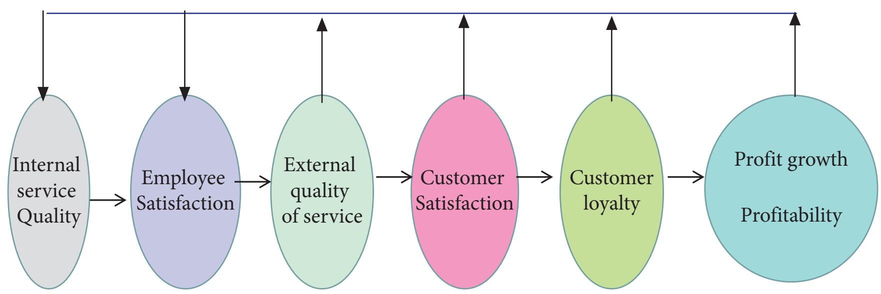


Figure 3: Service profit chain.

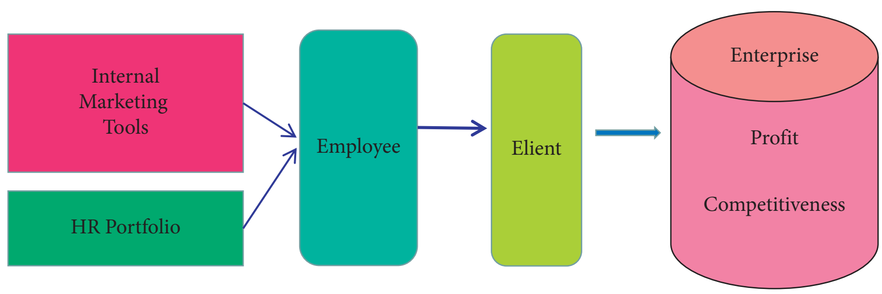


Figure 4: Integration model of internal marketing and human resource management.

for enterprises to increase profits and enhance their profitability. In order to improve employee satisfaction, it is necessary for enterprises to carry out internal marketing.

When enterprises carry out internal marketing activities, there are mainly two types of internal marketing tools used: first, the use of similar marketing tools. That is to say, the company regards employees as internal customers, and uses a set of marketing techniques and means for external customers as internal marketing tools [18]. Internal marketing is to apply the concepts, technologies, and methods of external marketing to the internal management process of the enterprise. In particular, when using similar marketing tools in internal marketing, the company first conducts market research on internal employees, then uses STP strategy to conduct market segmentation and internal positioning of all employees, and then uses a marketing mix to meet different internal target markets, which needs to satisfy employees. The second is to use human resource management tools [19]. Human resource management tools include enterprise organizational structure design, job analysis and evaluation, job description template, competency model, human resource planning, personnel recruitment management, training operation management, salary performance management, employee handbook, etc. When such tools are used in internal marketing, companies empower their employees and empower them by hiring the right people, training, motivating, and empowering those people, and providing the equipment, and technical and managerial support needed in employee service delivery. The enterprise fulfills its commitment to provide high quality service. Based on these two types of internal marketing tools, the theory of service profit chain, and the content and means of human resource management, an integrated model of internal marketing and human resource management is constructed, as shown in Figure 4.

As shown in Figure 4, the model illustrates that through the use of effective internal marketing tools, companies can improve employee satisfaction, loyalty, and productivity, which in turn can enhance customer satisfaction and loyalty, which ultimately leads to increased corporate profits and corporate profits' enhanced profitability. In the integration model of internal marketing and human resource management, internal marketing tools include quasimarketing tools and human resource management tools. It combines the advantages of marketing-like tools and human resource management tools, and integrates the two tools to better achieve the goal of employee satisfaction [20]. When carrying out internal marketing, companies first use market research and STP strategies in similar marketing tools, and then formulate corresponding human resource combinations for different internal markets, that is, design different trainings according to the needs and preferences of various employee groups, a mix of human resources such as projects, incentives, empowerment, and communication. Since this human resource combination is formulated on the basis of market research and market segmentation, and fully considers the different needs and preferences of employees, the effective implementation of this combination can better meet the needs of employees. Therefore, similar to marketing tools and human resources the integrated use of management tools makes it easier for employees to be satisfied [21].

3.2. Overview of Genetic Algorithm Use. The genetic algorithm (GA) is designed and proposed according to the evolutionary laws of organisms in nature. It is a computational model that simulates the natural selection and genetic mechanism of Darwin's theory of biological evolution. The genetic algorithm is an important part of global equation research and can be obtained during the research process. By targeting the genetics of the study, the research process can be modified to obtain the best solution. The genetic algorithm has two characteristics:

(1) *Group Search*. Genetic algorithms simulate the learning process of a population consisting of a sequence of individuals, where each individual represents a possible solution in the search space. Starting from an initialized group, through random search process, crossover, and mutation operations, individuals are evolved, and the group gradually evolves to an optimized area in the search space to obtain the optimal solution [22].

(2) *Information Exchange between Individuals*. The crossover operation represents the social information interaction within the group, and the mutation operation introduces new individuals into the group, that is, introduces a new feasible solution, and maintains the diversity of information in the group, so as not to be trapped in local optimization.

Therefore, it is an effective way to apply the genetic algorithm, through the self-adaptive learning process and parallel processing, to deal with the human resource evaluation problem with less computational cost.

Assuming that there are K evaluation models, the result of the k-th evaluation model is a vector $E_K = \{e_{1k}, e_{2k}, \ldots, e_{Nk}\}$, e_{ik} is the evaluation value of the evaluation model for a certain talent, and N is the number of all personnel to be evaluated.

In the traditional evaluation system, since there is only one evaluation model, talents can be sorted according to the evaluation value. However, it is difficult to guarantee the objectivity and reliability of the evaluation results [23]. Under the condition of the coexistence of multiple evaluation models, since the results produced by the multiple evaluation models are different or even in conflict with each other, effective synthesis is required to obtain the final evaluation result. Therefore, let $W_k = \{w_{1k}, w_{2k}, \ldots, w_{Nk}\}$ be the weight vector of the first evaluation model, and w_{ik} represents the evaluation performance of the k th evaluation model for talent i. The larger the value, the better the evaluation model is at predicting the i-th talent. That is, the talent is more suitable for evaluation using this model, and the conditions are as shown in the following equation:

$$\sum_{k=1}^{K} w_{ik} = 1, \quad (1)$$

$W = \{W_1, W_2, \ldots, W_K\}$ is the combination of each weight vector. In order to obtain the final evaluation value o_i for talent i, the measurement value of each evaluation model can be synthesized by the following equation:

$$o_i(W) = \sum_{k=1}^{K} w_{ik} e_{ik}. \quad (2)$$

Therefore, the combined assessment problem can be described as a two-layer network, as shown in Figure 5.

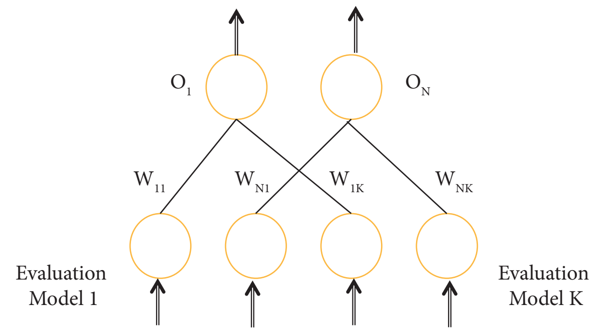


Figure 5: Structure of the combined evaluation model.

The input layer has $K \times N$ nodes, including the result vectors of each evaluation model. The output layer has N nodes, which represent N evaluation results, respectively, and are used to represent the comprehensive evaluation results of talents. Then, under the condition that the performance of each evaluation model is certain, the connection weight of the network becomes a factor that affects the accuracy of the final classification result. Thus, the combination problem of multiassessment models is transformed into the optimization problem of connection weights [24]. The genetic algorithm is an optimization method that applies the statistical search algorithm on the basis of the biological evolution process and is often used in the optimization of system parameters. In a genetic algorithm, the problem is encoded into strings called chromosomes, the solutions are called humans, and the authors are called citizens. Since the genetic algorithm cannot directly deal with the parameters of the problem space, the problem to be solved must be expressed as chromosomes or individuals in the genetic space through coding. This conversion operation is called encoding. Populations are constantly evolving, and better and better solutions will gradually emerge according to the principle of survival of the fittest. Here, the combination $W = \{w_{11}, w_{21}, \ldots, w_{N1}, w_{12}, w_{22}, \ldots, w_{N2}, \ldots, w_{1K}, w_{2K}, \ldots, w_{NK}\}$ of each weight vector becomes the problem to be solved. An initial value can be set, or the first public can be created with heavy equipment obtained from offline work, and then repeated, each iteration creating a new one whose performance is measured by physical activity, recent residents. Through rotation and transformation, individuals with disabilities are selected from the current population for development [25]. The fitness function is an evaluation function, which is used to analyze the performance of an individual and evaluate whether it can make better predictions. The current individual W is used to combine various evaluation models. Then, the final evaluation result is consistent with the actual score value. It indicates the accuracy of its evaluation, which is the basis for evaluating the applicability of the current individual W. We construct the accuracy function $f(W, x)$, as shown in the following equation:

$$f(W) \begin{cases} 1, & \text{If assessed correctly,} \\ \dfrac{o(W)}{\sum_{i=1}^{N} o_i(W)}, & \text{Otherwise.} \end{cases} \quad (3)$$

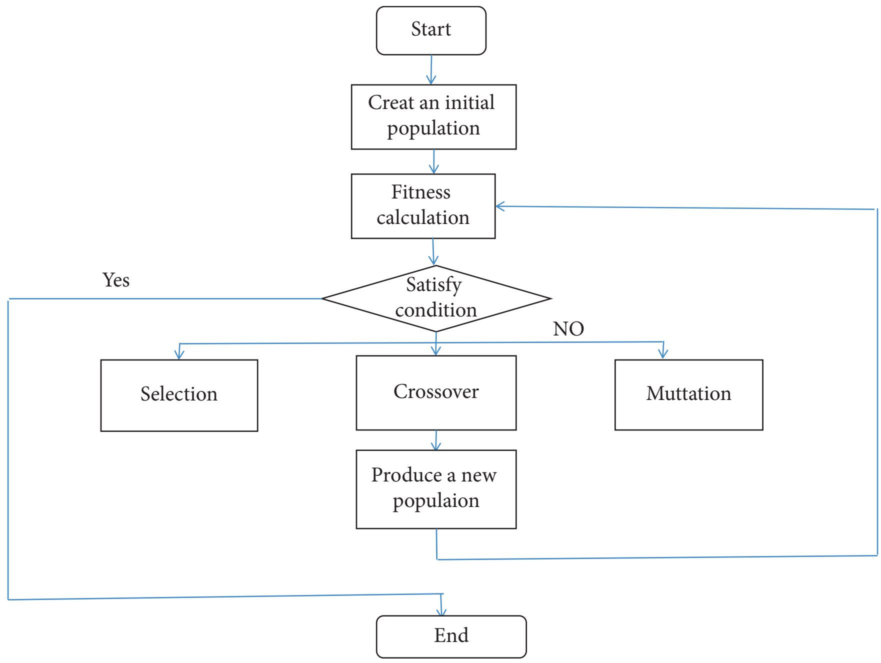


Figure 6: Flowchart of GA optimization methods.

In the formula, $o_i(W)$ represents the comprehensive measure of the current talent i score when the weight value is W; $o(W)$ is the true score of the talent. Therefore, a training set is needed, that is, the evaluation historical data that have been actually verified. Obviously, the larger the $f(W)$ is, the higher the prediction accuracy is [26]. The accuracy function represents the accuracy of a talent evaluation. When there are multiple talents, the accuracy of its prediction needs to be further analyzed to represent the fitness of the individual W, thus representing the performance of the combined evaluation model, as shown in the following equation:

$$\text{fitness}(W) = \frac{\sum_{x\in C} f(W)}{|C|}. \tag{4}$$

In the formula, C is the training set; $|C|$ is the number of elements in the training set.

Based on health, individuals are selected from the current population to create the next generation. The outcome model is used to select individuals, and the outcome of each selected person is shown as follows:

$$p(W) = \frac{\text{fitness}(W)}{\sum_{i=1}^{M} \text{fitness}(W_i)}, \tag{5}$$

where M is the number of individuals in the population. Since the sum of the fitness of the population is constant, the greater the fitness of an individual, the greater the probability of it being selected. Figure 6 shows the flowchart of the genetic algorithm optimization method.

3.3. Theoretical Basis of Personnel Demand Forecast. Exponential smoothing, also known as exponential smoothing, is the critical time of the estimation method and is a special weighted moving average estimation method pioneered by centrifugal smoothing. Exponential smoothing uses weights from historical data to estimate future outcomes over a period of time. Exponential smoothing, consistent with full-time and moving averages, focuses on the effect of long-term values on a time series of future estimates. The larger the weight, the larger the weight ratio, which is equal to 1, because the weight coefficient follows the exponential law and has the function of exponential smoothing, which is called exponential smoothing. According to the different exponential smoothing methods, it can be divided into the following: first exponential smoothing method, second exponential smoothing method, and third exponential smoothing method; first exponential smoothing method is suitable for time series and linear right, and third exponential smoothing method is suitable for the second exponential smoothing method. The time series of the first polynomial order changes.

3.3.1. First Exponential Model. Let the time series be $X_t, t = 1, 2, \ldots, n$; we use S to represent the exponential smoothing value, α is the smoothing coefficient $0 < \alpha < 1$, the first exponential smoothing value in the t period is recorded as $S_t^{(1)}$, and the formula for calculating the first exponential smoothing value is shown as follows:

$$S_t^{(1)} = \alpha X_1 + (1-\alpha) S_{t-1}^{(1)}. \tag{6}$$

3.3.2. Quadratic Exponential Model. The secondary exponential smoothing value of the t period is denoted as $S_t^{(2)}$, and the calculation formula of the secondary exponential smoothing value is shown as follows:

$$S_t^{(2)} = \alpha S_t^{(1)} + (1-\alpha) S_{t-1}^{(2)}, \tag{7}$$

$\widehat{X}_{t+T}$ represents the forecast value with the forecast period T starting from time t, and the mathematical model of the quadratic exponential smoothing method forecast is shown as follows:

$$\widehat{X}_{t+T} = a_t + b_t T. \tag{8}$$

The calculation formulas of parameters a_t and b_t are shown as follows:

$$a_t = 2S_t^{(1)} - S_t^{(2)}, \tag{9}$$

$$b_t = \frac{\alpha}{1-\alpha}\left(S_t^{(1)} - S_t^{(2)}\right). \tag{10}$$

3.3.3. Triple Exponential Model. When the changes in the time series show a quadratic parabolic trend, the triple exponential smoothing method is required. The triple exponential smoothing method is to perform another smoothing on the basis of the secondary smoothing, $S_t^{(3)}$ is the value of the third exponential smoothing, and the calculation formula is shown in as follows:

$$S_t^{(3)} = \alpha S_t^{(2)} + (1-\alpha) S_{t-1}^{(3)}, \tag{11}$$

$\widehat{X}_{t+T}$ represents the predicted value with the prediction period T starting from time t, and the mathematical model of the triple exponential smoothing method is shown as follows:

$$\widehat{X}_{t+T} = a + bT + cT^2. \tag{12}$$

The calculation formulas of parameters a_t, b_t, and c_t are shown as follows:

$$a_t = 3S_t^{(1)} - 3S_t^{(2)} + S_t^{(3)}, \tag{13}$$

$$b_t = \frac{\alpha}{2(1-\alpha)^2}\left((6-5\alpha)S_t^{(1)} - 2(5-4\alpha)S_t^{(2)} + (4-3\alpha)S_t^{(3)}\right), \tag{14}$$

$$c_t = \frac{\alpha^2}{2(1-\alpha)^2}\left(S_t^{(1)} - 2S_t^{(2)} + S_t^{(3)}\right). \tag{15}$$

3.3.4. Smooth Coefficient. The smoothing coefficient is a coefficient related to the exponential smoothing method, and the value of the exponential smoothing coefficient is very important. The smoothing constant determines the level of smoothing and how quickly it responds to the difference between the predicted value and the actual result. When the exponential smoothing method is used for trend forecasting, the value of the smoothing coefficient α needs to be reasonably determined. The determination of the α value can be obtained according to the empirical judgment method. For example, when the time series shows a relatively stable horizontal trend, a smaller α value is selected, so that the forecast values of each period have a similar impact on the forecast results; when the time series fluctuates, however, when the long-term trend does not change much, a slightly larger α value can be selected; when the time series greatly fluctuates, the long-term trend greatly changes, it shows an obvious and rapid upward or downward trend, and a larger α value should be selected to avoid and improve the sensitivity of predictive models.

However, according to the empirical method, there are certain limitations. In most cases, the α value is not easy to judge. For the sake of accuracy, different α values need to be used for trial calculation. According to the actual situation, the minimum error of the most recent periods is generally considered as the criterion. Error analysis indicators generally use the standard deviation of error (SDE) and mean relative error (MAPE), as shown in the following formulas:

$$SDE = \frac{1}{n-1}\left[\sum_{i=1}^{n}\left(X_i - \widehat{X}_i\right)^2\right], \tag{16}$$

$$\text{MAPE} = \frac{1}{n}\left[\sum_{i=1}^{n}\left|\frac{X_i - \widehat{X}_i}{X_i}\right|\right]. \tag{17}$$

The advantage of the exponential smoothing method for forecasting is that it can carry out the self-weight correction and monitoring model for the forecast error generated by each operation, so it can effectively judge the impact of each traffic cycle on future traffic.

3.4. Human Resource Management and Marketing Data Survey. We take a call center as an example to study human resource management and marketing. In order to reduce the heavy investment in labor costs, call center systems generally tend to operate at a high utilization rate. However, high utilization is not safe for call centers with heavy system load, because if the actual arrival traffic slightly exceeds the forecast value, it may generate a long call queue, which will increase the customer abandonment rate. The increasing queuing queue will increase the average delay time of customers in the waiting process, which in turn will increase the customer abandonment rate. Efficiency has important implications [27]. Traffic forecasting refers to the estimation and expectation of the traffic that may appear in the future network by analyzing the historical data statistics of the telecommunication network traffic or related factors. Traffic forecasting is not a simple task; it is inseparable from the customer analysis and business type analysis of the call center. The call center traffic has the following characteristics:

(1) When the service level increases, the utilization rate of the operator decreases.

   To improve service levels, call centers need more operators to answer customer questions in a timely manner, so the workload of each operator will be reduced. It is foreseeable that when the service level increases by one percentage point, the utilization rate of man-hours will decrease accordingly. These two

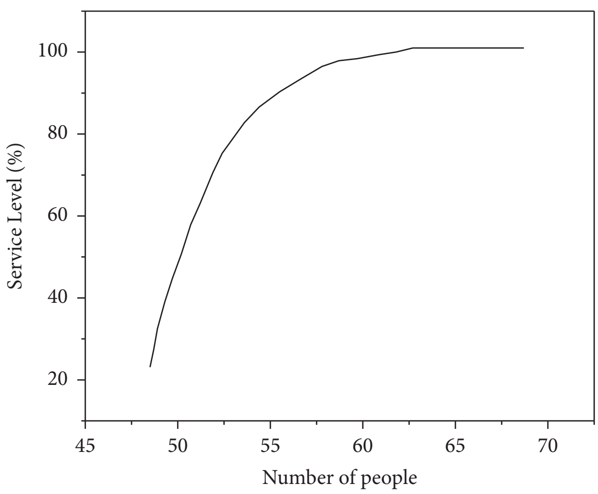


Figure 7: The relationship between the number of operators and service level.

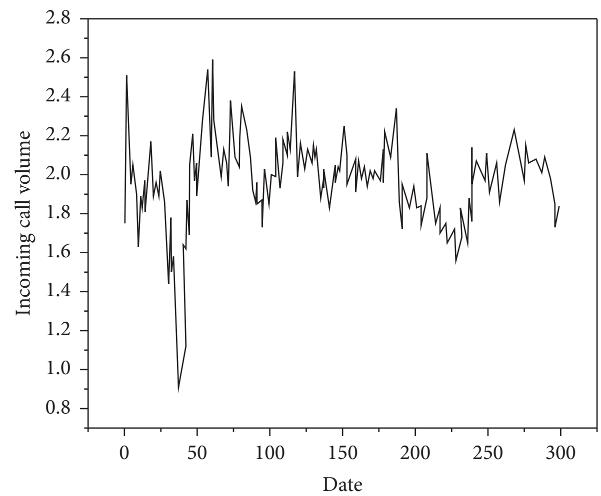


Figure 8: Graph of incoming calls from January 1, 2021, to October 31, 2021.

indicators are negatively correlated in the call center management system.

(2) When the number of operators increases, the average answering speed and trunk load decrease.

When the number of call center operators increases, calls waiting to be answered are answered as quickly as possible, so the average speed of answering decreases, and thus, the time of the trunk is occupied. These three indicators are also negatively correlated in the call center management system.

(3) The law of diminishing returns is as follows:

When the service level of a call center is low, adding operators will have a significant effect on improving the service level. But as service levels continue to rise, the effect of increasing the number of operators will become less and less effective, to zero. When the traffic volume per unit time is 0.2 calls l s and the average service duration is 240 s, the relationship between the number of operators and the service level is shown in Figure 7.

(i) Traffic volume is affected by special events. Under normal circumstances, the traffic volume shows a regular trend of change. The historical traffic data automatically imported through the CTI system interface can be used as the basis for future traffic forecasting. However, the traffic volume is also affected by special events. Events that have an impact on traffic volume can be divided into two categories; one refers to regular events such as weekends and billing days at the beginning of the month and the end of the month; the other refers to special events such as promotional activities, holidays, system upgrades, and system failures.

The change trend of traffic volume is closely related to the work and living habits of call center customers. The traffic volume changes with time and has the relevant characteristics of time series. Therefore, the traffic volume corresponding to different times can be called traffic time. From the long-term change trend of traffic time series, it can be intuitively seen that there is a certain period of traffic time series. When establishing a traffic time-series prediction model, it is necessary to consider the periodic characteristics of the sequence. The following is the analysis of historical traffic volume laws that change over time.

(ii) Monthly variation of traffic volume is as follows: Figure 8 shows the daily traffic volume of a company's call center from January 1, 2021, to October 31, 2021. The horizontal axis represents hours, and the vertical axis represents traffic volume. It can be seen from the figure that the change trend of traffic has a typical monthly cycle. Near the billing period at the beginning of the month and at the end of the month, the traffic volume is large and fluctuates in monthly units. There are some obvious abnormal points in the figure, such as February 2021. At the beginning of the month, the traffic volume dropped significantly compared with the usual one. This is due to the influence of the Spring Festival holiday. Most of the abnormal traffic curves are caused by special events. From the data analysis and research, it can be seen that the traffic volume is measured in monthly terms. It presents a regular change for the period, which is called the monthly change rule of the traffic volume.

Figure 9 shows the traffic volume change curve in hours on a certain day in June 2021. The curve in the figure represents the traffic volume change on a certain day in June, and the corresponding relationship between the curve and the date is

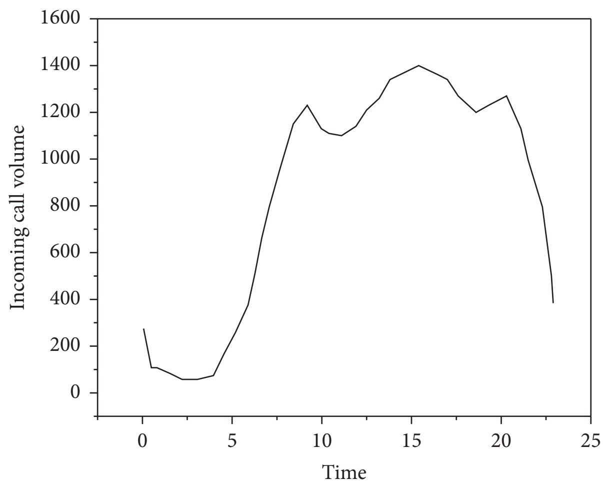


FIGURE 9: Graph of incoming calls on a certain day in June 2021.

shown in the subscript in the figure. It can be seen that there are two relatively fixed traffic peak periods in June. The morning traffic peak occurs around 12:00, the evening peak occurs from 18:00 to 21:00, and the traffic occurs between 0:00 and 6:00 in the morning. The volume trough, because the daily time period of the traffic volume is obvious and fixed, is called the daily change rule of the traffic volume.

Figure 10 is a graph of the change trend of in-call volume in weeks from January to October 2021. It can be seen from the figure that most of the time traffic distribution has obvious characteristics, the traffic volume on Monday is slightly more than other working days, and the traffic volume on Saturday and Sunday is relatively slightly less. It is formed by working and living habits, but other curves do not conform to this law very well, and cannot show regular changes. The change rule in weeks is relatively insignificant, but it can still be used as one of the analysis and reference bases for traffic forecasting.

To sum up, although there is randomness in the arrival of traffic, the change trend conforms to a certain law, so it is completely feasible to make a more accurate prediction.

The historical traffic volume of a company's call center studied in this research is in hours. The following is a one-month historical traffic data of a company's call center from April 15 to May 14, 2021, by the exponential smoothing method. We predict the traffic volume of each time period within 24 hours on May 15, 2021.

When using exponential smoothing method, according to the above to determine the coefficient of exponential smoothing method, the historical data of different period with different weights, using type (17) shows the average relative error of calculation in the last ten days the sum of the error of predicted value, as the final measurement standards, selection of smoothing coefficient of minimum error, with the coefficient of computing time.

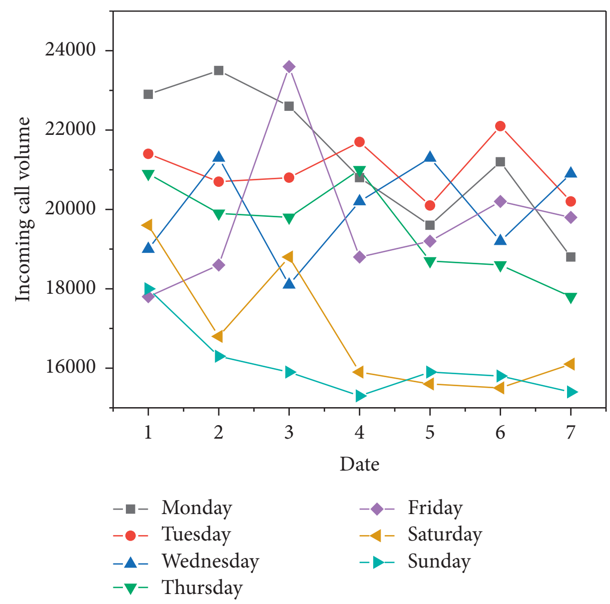


FIGURE 10: Curve of weekly inbound call volume from January to October 2021.

The secondary and tertiary exponential smoothing methods are used to predict the traffic volume, the traffic volume forecast value of each time period on May 15, 2021, is calculated, and the error and accuracy of the two methods are compared and analyzed. The selection of the smoothing index is based on the minimum average relative error, and the traffic volume in hours is calculated according to the quadratic and triple exponential smoothing methods, and is compared and analyzed with the actual value, as shown in Table 1 and Figure 11. Among them, the optimal smoothing coefficient 2 refers to the coefficient with the smallest average error when the quadratic exponential smoothing method is used, and the optimal smoothing coefficient 3 refers to the coefficient with the smallest average error when the triple exponential smoothing method is used.

It can be seen from Figure 12 that the change in the projected curve of the quadratic exponential smoothing path is gradual at the peak running time. In other periods, the two methods are similar to the observed values, but the average error of the quadratic exponential smoothing method is similar to that of the cubic exponential smoothing method. The exponential smoothing method is relatively small, and the change trend is more consistent with the observed value, so the second exponential smoothing method has more advantages than the third exponential smoothing method in traffic forecasting.

As shown in Figure 13, the service duration includes call-in delay, waiting time, and operator working time, which is the sum of customer delay reminder time, waiting time, actual call time, and operator finishing time. Service duration is an important part of the Erlang C queuing model. The input parameter is one of the important indicators to measure the service quality of the call center operation. It can be used to calculate the system load and the demand of personnel in each time period, and finally decide the

Table 1: Predicted data of the quadratic and triple exponential smoothing method.

Period	Observations	Optimal smoothing factor 2	Quadratic exponential smoothing	Mean relative error 2	Optimal smoothing factor 3	Triple exponential smoothing	Mean relative error 3
0 o'clock	365	0.1	332.22	0.023	1	156	0.1
1 o'clock	411	0.2	356.79	0.12	0.1	279.78	0.1
2 o'clock	97	0.1	114	0.079	1	100	0
3 o'clock	132	0	126	0.123	0.2	379.99	1
4 o'clock	349	0	397.67	0.089	0.3	465.12	0
5 o'clock	623	0	720	0.117	0	776.81	1
6 o'clock	798	0.3	598.91	0.113	0.1	684.89	0
7 o'clock	1034	1	1132.24	0.123	0.1	1777	0.1
8 o'clock	1058	0	1250.18	0.2	0.1	1546.76	0.1
9 o'clock	1177	0	1278.61	0.218	0.1	1443.52	0.1
10 o'clock	1288	0.1	1300	0.098	0	1323.01	0.3
11 o'clock	1358	0.2	1319.11	0.078	0.1	1433.89	0
12 o'clock	1369	0.1	1495.67	0.056	0.2	1221.51	0.1
13 o'clock	1451	1	1588.34	0.091	0.1	1589	0.1
14 o'clock	1436	1	1597.40	0.045	0	1563.21	1
15 o'clock	1479	0	1518	0.026	0.3	1479.08	0
16 o'clock	1498	0	1478.89	0.178	1	1322.6	0.1
17 o'clock	1266	0	1789.60	0.195	0.1	1345.8	0
18 o'clock	1049	0.1	1123.5	0.123	0.1	1209.5	0
19 o'clock	954	0.1	943.98	0.2	0	1092.3	1
20 o'clock	723	0.2	669.99	0.1	0	926.1	0.1
21 o'clock	651	0.2	723.12	0.028	0.2	765.22	0.3
22 o'clock	543	0	503.1	0.052	0.1	591.84	0.1
23 o'clock	425	0.1	298.7	0.071	0	335.26	1

arrangement of the call center operators in each time period. Therefore, it is very important to estimate the average service duration. It directly affects the call center call queue length and user waiting time, and has a strong correlation with organizational capacity planning issues such as personnel arrangement and recruitment training, thereby affecting the service level, quality, and user satisfaction of the call center. Figure 13 shows the change trend of the average service time of a company's call center within 24 hours of a day on April 15, 2021. It can be seen that the average service time in a day is shorter from 2:00 a.m. to 6:00 a.m., and during the peak traffic period, it is relatively long, and the changes in other time periods are relatively gentle.

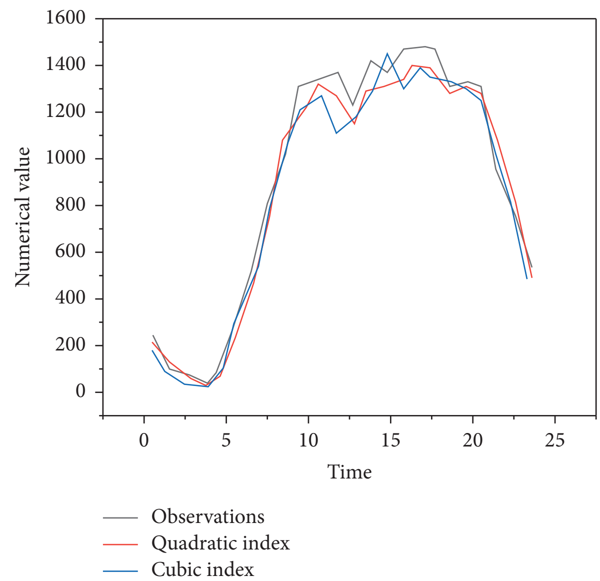


Figure 11: Comparison of prediction results between the second exponential smoothing method and the third exponential smoothing method.

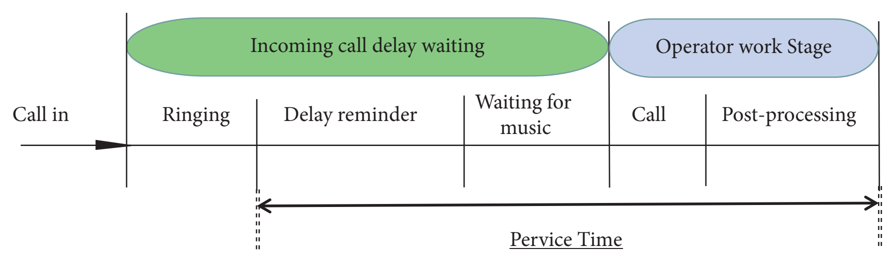


Figure 12: Schematic diagram of the duration of a phone call.

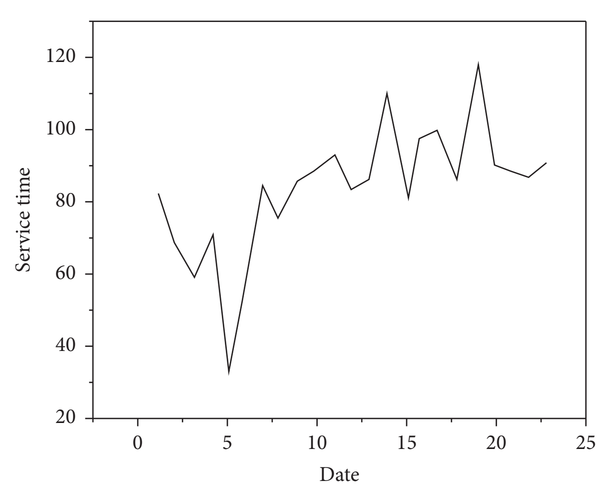


Figure 13: Variation trend of service hours on April 15, 2021.

4. Results and Analysis

In the early stage of the genetic algorithm operation, the individual differences are large, and the roulette method can be used for selection. The number of offspring produced is proportional to the fitness of the parent individual. It is easy to make the population contain too many offspring of good individuals, resulting in premature maturity (premature). In the late stage of algorithm operation, the fitness of individuals gradually tends to be consistent, and the advantages of good individuals are no longer obvious when producing offspring, which makes the evolution of the entire population stagnate, which is not conducive to obtaining better offspring. Therefore, a selection strategy needs to be designed to avoid getting stuck in local optimization [28]. The simulated annealing (simulated annealing) algorithm is a commonly used combinatorial optimization improvement method. It is a heuristic random search process based on the Monte Carlo iterative solution method. The simulated annealing algorithm is an optimization algorithm that can effectively avoid falling into a local minimum and eventually tend to the global optimal serial structure by giving the search process a time-varying probability that eventually tends to zero. We construct the individual selection strategy, as shown in the following equation:

$$P(W) = \frac{e^{\text{fitness}(W)/T}}{\sum_{W \in \text{Generation}} e^{\text{fitness}(W)/T}}, \tag{18}$$

$$T = \frac{T_0}{\log(1+t)}. \tag{19}$$

In the formula, generation is the collection of individuals, namely, the population; T is the temperature; T_0 is the set larger initial temperature; and t is the genetic algebra. With the increase in the genetic algebra, the temperature T gradually decreases, and the fitness of the individual W is gradually stretched, so that the fitness difference of the individuals with similar fitness is enlarged, so that the advantages of the excellent individual become more obvious. A large probability is selected, which overcomes the problem of early maturity of the population.

Suppose the number of individuals in the population is 50, the crossover rate is 0.55, the mutation rate is 0.001, and the initial temperature T_0 is 100. The experimental results are shown in Figure 14.

Through the analysis of the experimental results in Figure 14, the following conclusions can be drawn:

(1) As the number of people being evaluated increases, the error will increase. However, the three weighted means have no obvious advantage, and their mean errors are significantly higher than the results of the combined model.

(2) The combined evaluation model based on the genetic algorithm can automatically find reasonable connection weights and form a dynamic weight combination, which has a better evaluation effect than the fixed weighted average.

(3) The increase in the number of people being evaluated means an increase in the types of talents, which leads to an increase in the error of the evaluation results. At this time, the combined evaluation model based on the genetic algorithm can adaptively obtain a new weight combination so that the average error of the evaluation can be significantly improved.

We assume that a project performed by a company has 4 tasks that can be completed at the same time, and the values of each project are shown in Table 2. For each project, the

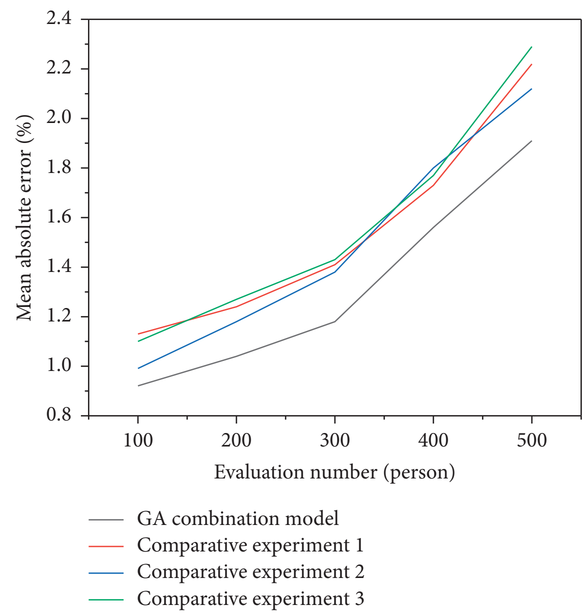


Figure 14: Comparison of mean absolute errors.

operational capabilities of employees follow a similar distribution, the upper limit of the probability density function is $e_1^{\max} \in [0.8, 0.9]$, the lower limit is $e_1^{\min} \in [0.2, 0.4]$, and the project parameters are shown in Table 2.

To solve this problem using the above model, first we divide the whole project into 4 cycles of 1 month each and create the best model for the situation. We set up three cases of different lengths for comparison as follows: in all cases, the onboarding period for new employees is 1 month, i.e., new employees can be onboarded at the beginning of each period. The three scenarios were solved by the method proposed above, and the calculation results are shown in Table 3.

In Scenario 2, the time to fire an employee is 2 months, so an employee can only be fired once during the work period, which is at the beginning of the third period. The average working ability changes of employees at different times are shown in Table 4.

The results show that the completion time of the project scenario is 4 months, and then, there will be layoffs during the work period. The workload of the second project is 157 people/day, and we will not complete it on time; the remaining equipment is outsourced, and the total project cost is 920,000 yuan. As can be seen from Scenario 1 in Table 4, there is no change in the average work capacity of employees at different times. In Scenario 2, the dismissal delay period is 2 months. So it is possible to remove the active worker once at the start of the 3rd period, which will start position 3 for the 3rd time, and cannot remove the functionality of task 3. Completing on time can reduce 41 people/d. Compared with scenario 1, the total project cost is also increased by 20,000 yuan. The average work capacity of 2 employees in period 3 increased from 0.55 person/day in period 2 to 0.63 person/day. Scenario 3 fires at shorter intervals, so when calculating benefits, Job 1, Job 3, and Job 4 are fired once, and Job 2 is fired twice. The average work

Table 2: Main parameters of the project task number task time.

Task number	Mission time (month)	Total workload (person d-1)	Employee monthly salary (yuan)	Minimum number of people	Outsourcing cost (yuan)
1	3	698	2500	6	300
2	2	552	3000	8	250
3	2	719	3500	8	280
4	4	980	4000	5	400

Table 3: Calculation result tasks for three scenarios.

Task number	Calculation results for scenario 1		Calculation results for scenario 2		Calculation results for scenario 3	
	Workload (person d-1)	Total cost/ten thousand	Workload (person d-1)	Total cost/ten thousand	Workload (person d-1)	Total cost/ten thousand
1	0	100	0	98	0	83
2	0		0		0	
3	210		168		110	
4	0		0		0	

Table 4: Relationship between average working ability and time of employees in each scenario.

Period	Scenario 1				Scenario 2				Scenario 3			
	Task 1	Task 2	Task 3	Task 4	Task 1	Task 2	Task 3	Task 4	Task 1	Task 2	Task 3	Task 4
1	0.6	0.55			0.6	0.55			0.6	0.56		
2	0.6	0.56			0.6	0.55			0.68	0.62		
3	0.55			0.65		0.63	0.5	0.65		0.79	0.5	0.65
4			0.5	0.65			0.5	0.65			0.54	0.71

capacity of employees has increased several times due to multiple posting times, which allows all tasks to be completed within the allotted time. The total cost of the project dropped to 850,000 yuan, 90,000 yuan is less than the second target, and the average working ability of various staff members showed an upward trend. Among them, the average working ability of 2 employees improved the most, from 0.56 people/day in the first stage to 0.79 people/day in the third stage, with an increase of 41%.

5. Conclusions

Portfolio assessment is an important work in human resource management. With existing assessment tools to assess human resources, the results are feasible. However, in the unified assessment of different types of teachers, the index system and weight need to be set in advance, which cannot reflect the best combination model of each individual under different circumstances, and the result has a large error. The combined evaluation model designed in this research can adaptively obtain a dynamic weight combination, thereby effectively reducing the evaluation error. In the specific implementation process, the school can scientifically design the corresponding team or job evaluation focus and index system according to the positioning and development goals, and combine the self-adaptive combined evaluation model to effectively carry out unified evaluation for different types of personnel, so as to provide basis and reference for improving performance management. It can be seen from the calculation example that by reasonably controlling the timing and quantity of the employment and dismissal of employees, the overall working ability of the human resources of the project can be gradually improved, thereby reducing the human resource cost of the project. The frequency of employee dismissal has a significant impact on the optimization effect of the employment plan; on the premise that the time interval of newly hired employees remains unchanged, the shorter the time interval for dismissal of employees, the greater the improvement in the average work efficiency of employees. The optimization effect of human resources is more obvious. In a word, using the hierarchical genetic algorithm to optimize supply chain platform, look from the inside, can effectively coordinate the relationship between the human resource department and other departments, look from the outside can also help enterprises and realize the maximization of the interests of stakeholders for the construction of human resources in the supply chain system, improve the efficiency of companies and institutions should the human resources department with the screening has positive significance and talents.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported by the Catholic University and Yeungnam University.

References

[1] D. Peng, G. Tan, K. Fang, L. Chen, and Y. Zhang, "Multi-objective optimization of an off-road vehicle suspension parameter through a genetic algorithm based on the particle swarm optimization," *Mathematical Problems in Engineering*, vol. 2021, no. 9, Article ID 9640928, 14 pages, 2021.

[2] I. Ibrahim, "Human resource management of bureaucracy in Indonesia: is it true that job satisfaction increases employee performance," *Psychology*, vol. 58, no. 1, pp. 3218–3231, 2021.

[3] A. Latif, N. U. Hadi, and I. A. Malik, "Impact of csr on service innovation performance in hospitality industry: the strategic role of human resource management," *The Pacific Review*, vol. 13, no. 8, pp. 20–33, 2021.

[4] L. Chen, W. Lin, P. Chen, S. Jiang, L. Liu, and H. Hu, "Porosity prediction from well logs using back propagation neural network optimized by genetic algorithm in one heterogeneous oil reservoirs of ordos basin, China," *Journal of Earth Sciences*, vol. 32, no. 4, pp. 828–838, 2021.

[5] S. Emiroglu and Y. Uyaroglu, "Genetic algorithm (ga)–based delay feedback control of chaotic behavior in the voltage mode controlled direct current (dc) drive system," *Zeitschrift für Naturforschung A*, vol. 76, no. 1, pp. 13–21, 2021.

[6] S. Y. Martowibowo and B. Kemala Damanik, "Optimization of material removal rate and surface roughness of aisi 316l under dry turning process using genetic algorithm," *Manufacturing Technology*, vol. 21, no. 3, pp. 373–380, 2021.

[7] Y. Chen, "Location and path optimization of green cold chain logistics based on improved genetic algorithm from the perspective of low carbon and environmental protection," *Fresenius Environmental Bulletin*, vol. 30, no. 6, pp. 5961–5973, 2021.

[8] K. Manita, Z. Fuqiang, and R. Maroy Prisca, "Evidence for mediating role of affective commitment on green human resource management and organisational citizenship behaviour," *Asian Academy of Management Journal*, vol. 26, no. 1, pp. 113–141, 2021.

[9] Nayana, "Electric vehicle charging with battery scheduling and multicriteria optimization using genetic algorithm," *Journal of Electrical Engineering and Automation*, vol. 2, no. 3, pp. 123–128, 2021.

[10] J. Xing and Z. Zhang, "Prediction model of network security situation based on genetic algorithm and support vector machine," *Journal of Intelligent and Fuzzy Systems*, no. 3, pp. 1–9, 2021.

[11] S. Samanta, A. Ojha, B. Das, and S. K. Mondal, "A profit maximisation solid transportation problem using genetic algorithm in fuzzy environment," *Fuzzy Information and Engineering*, vol. 13, no. 1, pp. 40–57, 2021.

[12] M. Al-Amin, R. Akter, A. Akter, M. Uddin, and A. Mamun, "Socially responsible human resource management and voluntary environmental behavior: moderating the effect of ecocentric leadership," *Central European Management Journal*, vol. 29, no. 2, pp. 147–168, 2021.

[13] A. Indian and K. Bhatia, "An approach to recognize handwritten Hindi characters using substantial zernike moments with genetic algorithm," *International Journal of Computer Vision and Image Processing*, vol. 11, no. 2, pp. 66–81, 2021.

[14] S. C. Liu, Z. G. Chen, Z. H. Zhan, S. W. Jeon, S. Kwong, and J Zhang, "Many-objective job-shop scheduling: a multiple populations for multiple objectives-based genetic algorithm approach," *IEEE Transactions on Cybernetics*, no. 99, pp. 1–15, 2021.

[15] S. V. Shet, T. Poddar, F. Wamba Samuel, and Y. K. Dwivedi, "Examining the determinants of successful adoption of data analytics in human resource management -a framework for implications," *Journal of Business Research*, vol. 131, no. 3, pp. 311–326, 2021.

[16] A. Saputro and L. C. Nawangsari, "The effect of green human resource management on organization citizenship behaviour for environment (ocbe) and its implications on employee performance at pt andalan bakti niaga," *European Journal of Business and Management Research*, vol. 6, no. 1, pp. 174–181, 2021.

[17] O. A. Awad, "A switching gain controller in drsigning optimum constraint control problems based on genetic algorithm," *Iraqi Journal of Information & Communications Technology*, vol. 4, no. 1, pp. 31–44, 2021.

[18] M. Alshheri, P. Sharma, R. Sharma, and O. Alfarraj, "Motion-based activities monitoring through biometric sensors using genetic algorithm," *Computers, Materials & Continua*, vol. 66, no. 3, pp. 2525–2538, 2021.

[19] M. Gamal, A. Elsawy, and A. AbuEl-Atta, "Hybrid algorithm based on chicken swarm optimization and genetic algorithm for text summarization," *International Journal of Intelligent Engineering and Systems*, vol. 14, no. 3, pp. 319–331, 2021.

[20] D. B. Kumar and M. R. Panchal, "Climate parameter based electric power requirement forecasting by fish genetic algorithm," *Design Engineering*, vol. 2021, no. 7, pp. 1728–1740, 2021.

[21] I. V. Kulikova, "Application of the variation operator in a genetic algorithm for the synthesis of fuzzy controllers," *Herald of Dagestan State Technical University Technical Sciences*, vol. 47, no. 4, pp. 92–100, 2021.

[22] S. Guliyeva, Y. Sadigov, N. Guliyeva, L. Isayeva, and S. Aliyeva, "Person-centered approach effectiveness in human resource management in the agriculture of Azerbaijan," *Journal of Eastern European and Central Asian Research (JEECAR)*, vol. 8, no. 2, pp. 267–279, 2021.

[23] Q. Li and X. Zhang, "How SMEs can create value for customers through emergency logistics and human resource management under unexpected events
—service-dominant logic perspective," *Open Journal of Business and Management*, vol. 09, no. 01, pp. 255–267, 2021.

[24] X. Liu, N. Shi, Y. Wang, Z. Ji, and S. He, "Data-driven boolean network inference using a genetic algorithm with marker-based encoding," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 19, no. 3, pp. 1558–1569, 2022 May-Jun.

[25] F. O Albalawi and M. S Maashi, “Selection and optimization of software development life cycles using a genetic algorithm,” *Intelligent Automation & Soft Computing*, vol. 28, no. 1, pp. 39–52, 2021.

[26] A. A. Masud, M. A. Hossain, and S. Biswas, “Impact of human resource management practices on the performance of the textile employees in barishal region,” *Annals of the Romanian Society for Cell Biology*, vol. 25, no. 5, pp. 5291–5304, 2021.

[27] Kengatharan, “Does national culture matter? the influence of human resource management practices on business strategy and firm performance,” *Journal of Critical Reviews*, vol. 7, no. 19, pp. 6900–6913, 2021.

[28] R. D. J. Dayrit, “Nurses’ work motivation and their demographics: basis for human resource management,” *Advances in Social Sciences Research Journal*, vol. 8, no. 1, pp. 170–185, 2021.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 1923386, 10 pages
https://doi.org/10.1155/2022/1923386

Research Article

Design of Emergency Intelligent Terminals for Field Exploration Based on Intelligent Internet of Things Technology

Chengbai Qiu[1] **Jun Zhang**[1] **and Jie Yu**[2]

[1]*Sports Institute, Hubei University of Science and Technology, Xianning 437100, China*
[2]*College of Art and Design, Hubei University of Science and Technology, Xianning 437100, China*

Correspondence should be addressed to Jie Yu; yujie@hbust.edu.cn

Received 14 June 2022; Accepted 14 July 2022; Published 21 August 2022

Academic Editor: Jiafu Su

To solve the intelligent rescue design of field exploration, the design of field intelligent terminal is explored based on intelligent Internet of things technology. Through the system's platform management, integration of internal information, and strengthening technology research and development, the system has successfully completed the deployment and implementation, which strongly supports the design of intelligent terminals for field exploration emergency rescue. Relevant technicians need to recognize the application and development trend of Internet of things technology in intelligent field emergency rescue scenarios and take various measures to promote it, so as to make it a better support and enrich intelligent field emergency rescue functions and provide better impetus for intelligent field emergency rescue and the construction of digital city. For the construction of the current intelligent field emergency rescue scene, the Internet of things technology plays a major role, which also shows that the system has better realized the predetermined function. The field intelligent rescue system makes good use of the shared location function, which can avoid accidents such as team members falling behind and also well identify the location of team members.

1. Introduction

Today, with the social development technology, technicians should clarify the advantages of Internet of things technology and combine online things with all social and skills, especially wireless transmission technology and other communication technologies. The things need to rely on a relatively complete system, which should have good communication function, wireless information transmission function, and data processing ability. In addition, with the support of big data technology and artificial intelligence technology, the Internet of things system also has a certain ability of intelligent identification and automatic processing, so that the social online things can have the ability of perception, network transmission, and personalized application at the same time. Figure 1 shows the planning of the Internet in artificial intelligence. Field exploration is a high-risk sport. With the increasing types, regions, and participants of field exploration activities, there is an increasing trend of distress and rescue events. The lack of field exploration norms and rescue system needs to be solved urgently. According to the experience of relevant countries and regions, an effective field exploration and rescue mechanism should be based on legal norms, protected by prevention and early warning, led by government rescue, and assisted by folk rescue. Taking this as a lesson, from a legal perspective, we should establish China's field exploration and rescue mechanism as soon as possible, focusing on establishing and improving the legal standard system of field exploration and rescue in distress, taking effective arrangements for field exploration prevention and early warning system, and building a field exploration and rescue organization system in line with national conditions. We are keen on visiting scenic spots everywhere and even explore deep mountains, wild areas, or virgin forests. Therefore, safety is particularly important. At present, there are a wide range of navigation products, such as Baidu, Gaode, and other map navigation software. Although it can provide convenient and considerate services for people living in cities and towns, it is lacking in the desolate field.

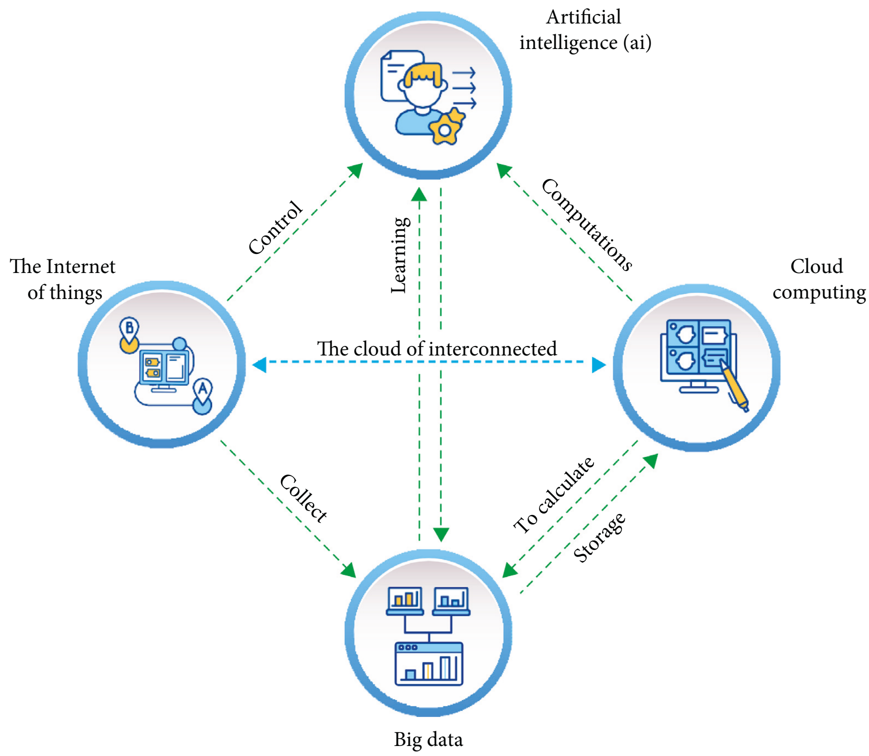


FIGURE 1: Planning of internet in artificial intelligence.

2. Literature Review

Sebastian said that physical network teaching social project is a humor type of social medicine, involving sensor technology, signal recognition technology, positioning technology, infrared sensing, and other technologies, and needs the support of hardware systems corresponding to various technologies [1]. Song et al. believe that, at this stage, with the online technology and social technology, the information age has come. For the online of things online social technology, which is extremely important in the modern information technology system, it can be applied to many fields such as intelligent field emergency rescue, intelligent medical treatment, and intelligent transportation, which can not only provide sufficient living convenience for people but also promote the coordinated and innovative development of many industries [2]. Zhou et al. believe that the look on the combined social things of emerge and online emergency rescue should become the research focus of current technicians [3]. Atiqur said that field exploration usually refers to exploration activities with certain risks in field areas with the purpose of self-development or self-challenge, using or partially using nontraditional equipment and facilities [4]. Therefore, Rathee et al. believe that the sense of achievement and stimulation brought by enhancing self-awareness are the essence and temptation to encourage field explorers to take risks. However, some risks during the exploration are huge or even fatal [5]. Hansaraj et al. believe that, in intelligent emergency rescue, such equipment is added to the specific Internet of things system, and the equipment becomes a part of the system. It should have good system adaptability and provide basic conditions for information sharing and transmission in the system, so as to enrich the data on the server side and provide more content data for user [6]. Zhen et al. reported that this is an important issue and is related to this question [7]. Xu also recognized that the intelligent rescue technology is widely used in the field of intelligent rescue, and they said that it is the best in the field of intelligent rescue. The reason why Internet of things technology can achieve relatively rapid development in the field of intelligence is that, on the one hand, in order to meet the new needs of people's intelligence in the area, exploration has continuously created and innovated in the field of social things, enriching the social things functions of products [8]. Iyer and Doraiswamyiyer believe that, at this stage, the local government is the first person responsible for the rescue of field explorers, which is first determined by its legal responsibilities, rights, and obligations [9]. Duishonbaeva et al. believe that the rehabilitation work mainly includes the resettlement and comfort of people in distress, the recognition and reward of rescue personnel, and the maintenance of rescue equipment and facilities. It can be said that the current rescue mechanism for field explorers in China is a compulsory emergency rescue mechanism led by local governments, coordinated by relevant departments and personnel, regardless of cost and conditions [10].

3. Method

3.1. Platform Management. The platform management problem is mainly related to the constituent elements of the Internet of things system. Generally speaking, all kinds of equipment in the Internet of things system have their specific attributes, which are set at the time of delivery. The corresponding manufacturers will personalize the appearance or function of the equipment according to the corresponding efficient application [11]. However, when such equipment is added to the specific Internet of things system, the equipment becomes a part of the system. It should have good system adaptability and provide basic conditions for information sharing and transmission in the system. However, in the normal application of online develop social technology, due to the obvious differences of product design standards implemented in different industries, the information quality produced by different devices is different, and there is a certain interaction lag between them. As shown in Figure 2, the connection flow chart between Internet system and equipment is shown. Secondly, undertaking the rescue obligation is the internal core mission of the government to serve the society. The reports of the 17th and 18th National Congress of the Communist Party of China clearly put forward the need to develop just order miss administrative project and build effect social government [12]. Modern social assistance system of such equipment is added to the specific Internet of things system, and the equipment becomes a part of the system. It should have good system adaptability and provide basic conditions for information sharing and transmission in the system. From the perspective of culture and humanitarianism, one view of China's traditional Confucian culture is to publicize nationalism, emphasizing that citizens are loyal to the state, the state protects citizens, and rescue vulnerable citizens project message in distress. From the perspective of social ethics and morality, everyone has the less message and responsibility to reduce lives in survival crisis. As a primary organization, local governments also have the social project and just of human social and morality to actively emerge advice and ensure the personal safety of field explorers to the greatest extent [13]. As long as it is an intelligent exploration device, it can use the compass to accurately determine the direction of the current position. The new route memory function can effectively remember the current user's route and help the user get out of the wrong way. As shown in Figure 3, the design flow chart of intelligent system is shown.

Even though the devices of different manufacturers and platforms can exchange information with each other, there will often be delayed information feedback or wrong execution of system commands. For example, the interoperability between rescue devices based on Android system and devices based on IOS system is ignored [14]. The interoperability is shown in Figures 4 and 5: comparison of Android and IOS interoperability. Therefore, technicians should develop a unified system management platform and promote all kinds of equipment to perform unified operation under the guidance of unified management protocol. Technicians should develop a unified system with high efficiency and high quality. In order to achieve high efficiency and high quality IoT technology application process, the average time formula of Internet of things under intelligent mobile is as follows:

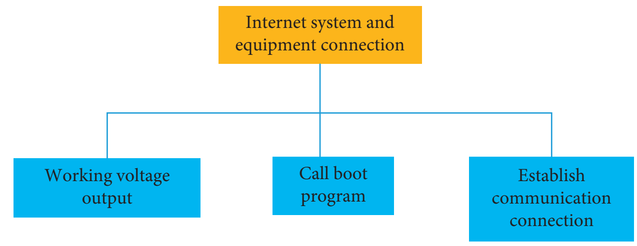


Figure 2: Connection flow chart of internet system and equipment.

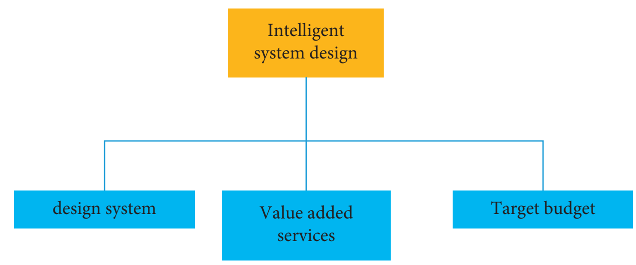


Figure 3: Design flow chart of intelligent system.

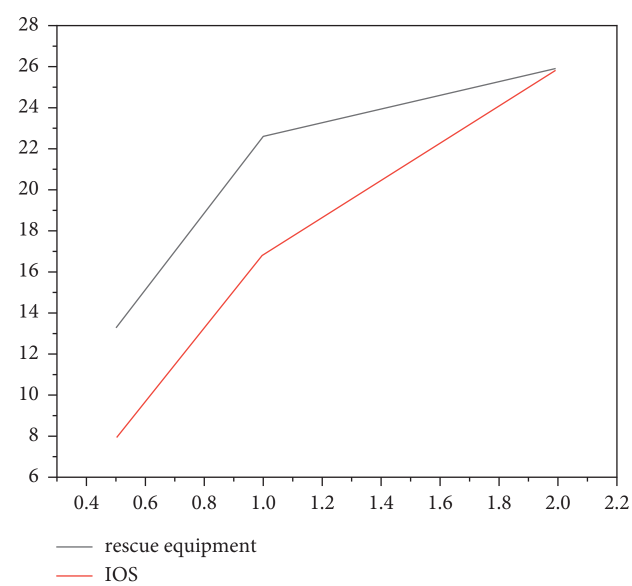


Figure 4: Interoperability between rescue equipment and IOS system equipment.

$$T = a + b\log_2\left(1 + \frac{D}{W}\right), \quad (1)$$

T represents the average time required to complete the movement (in some cases, MT can also be used to represent the movement time), a represents the start or stop time of the operation, b represents the movement or operation speed, D represents the distance from the starting point to the end of the target, and W represents the width of the target [15].

Starting from the input terminals a and B of the Internet of things, the logic formula of the output terminal is expressed in logic formula one by one, and

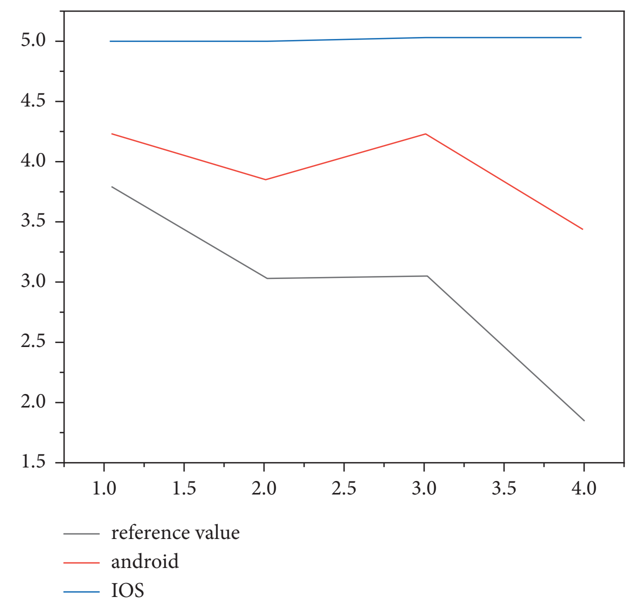


Figure 5: Comparison of interoperability between android and IOS.

$Y = \overline{\overline{A+B}+\overline{A+B}}$ is obtained. Then, the formula after exchange is

$$Y = (A+B)(\overline{A}+\overline{B}). \tag{2}$$

According to the input layer of IOT technology, the output formula of IOT intelligent design terminal can be as follows:

$$E = \frac{1}{2}(d-o)^2 = \frac{1}{2}\sum (d_k - o_k)^2. \tag{3}$$

3.2. Integration of Internal Information. With the support of Internet of things technology, many aspects including commodity supply, production, logistics, and demand have been effectively integrated, which can form a complete industrial supply chain, improve the service quality for customers based on clarifying the relationship between supply and demand, and ensure customers' product satisfaction [16]. At the long medicine time, things information management system can accurately control the market in many aspects, including raw materials, products, and finished products, avoid the uncertain factors in the relationship between supply and demand on the basis of timely taking risk control measures, and establish an enterprise interactive mechanism to achieve the timely communication of market information, promote the closer integration of supply chain links, and ensure the service quality of logistics. Finally, promote the long-term development of the whole online development project reducing unnecessary enterprise capital waste and cost waste [17].

Although most field exploration will not bring direct economic benefits to the territorial Government, as an increasingly prosperous special outdoor activity, it has comprehensive economic benefits. From the perspective of rights and obligations, local governments also have the obligation of management to undertake the rescue work of field explorers. Figure 6 shows the crisis management module. The intelligent rescue system is based on the open-source mobile application platform. In order to realize accurate positioning and navigation, Baidu API, cloud database, etc. are adopted, combined with software and hardware to realize the functions of real-time monitoring direction, human body temperature, ambient temperature, positioning and navigation, route memory, sharing location, and scenic spot information [18]. It is mainly aimed at people who like to carry out field activities. It can help users monitor their own situation and positioning information in real time and can effectively avoid accidents such as team members' loss and disorientation. It is an intelligent navigation system with strong practicability and great market potential [19].

The crisis management module is the main module of the system, which mainly realizes a series of processes when the field personnel encounter danger to ensure their own safety. The realization of this module is the main process that directly affects the safety of the system in case of crisis [20]. Baidu API large-scale construction module is shown in Figure 7.

Application programming interface (API) is translated as "application programming interface." Open the positioning and click social. The purpose is to provide applications and developers with the ability to access a set of routines based on a piece of software or hardware. When starting from a fixed point and constantly moving the position, the longitude and latitude information will be continuously updated in a short time and connected with online learning record management: record students' learning time, login time, and online situation, so that teachers can timely grasp students' learning dynamics and learning process. During field navigation, in order to identify the correct direction and not to get easily lost, the memory route function needs to be used [21]. When using it, open the positioning and click "start recording." When starting from a fixed point and constantly moving the position, the longitude and latitude information will be continuously updated in a short time and connected with the previous point to form a route. As long as the location is updated, keep connecting until you click "stop recording," and the obtained route information will be saved in the SD card file. The cloud database management module is shown in Figure 8.

The test and evaluation module are mainly used to test the rescue effect and are an intuitive evaluation of the field rescue results. The main functions of this module are phased inspection: the phased inspection is to ensure the safety of field explorers [22].

The test evaluation module is shown in Figure 9.

User communication module is an important functional module to realize the interaction between Internet management and field explorers in the process of teaching activities. It is mainly used for information exchange activities between users. The module mainly includes real-time interactive online Q and A and BBS online message. Before answering questions, rescuers need to create a Q and a room,

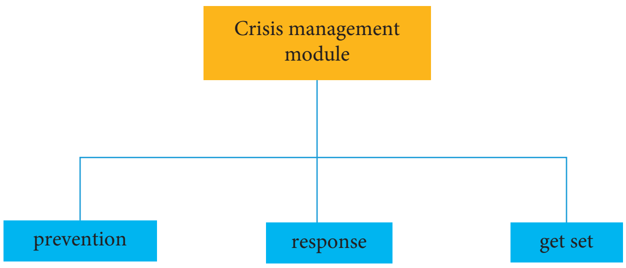


Figure 6: Crisis management module diagram.

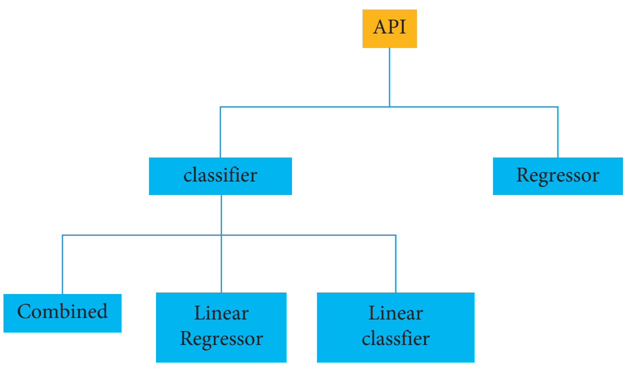


Figure 7: Large scale construction module diagram of Baidu API.

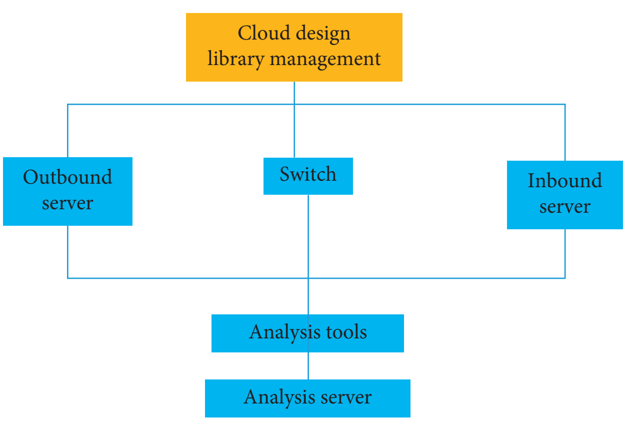


Figure 8: Cloud design library management module.

and enterprise personnel can answer questions in real-time only after entering the corresponding Q and a room, so that they can answer questions in groups in different places and classes at the same time; BBS message realizes non-real-time online communication. The user communication module is shown in Figure 10.

The system management module mainly manages user accounts, including adding apps and deleting social and model app information, and can assign permissions to users at the same time. System parameters can be maintained, and system data can be backed up [23]. The function module of the background management module is shown in Figure 11.

3.3. Strengthen Technology Research and Development. Facing the broad development space of logistics information management system from the online social project app, we

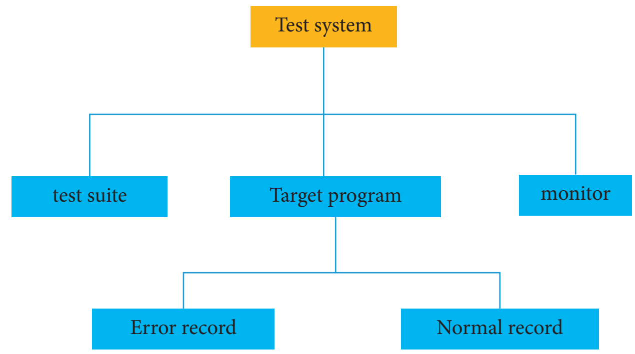


Figure 9: Schematic diagram of test and evaluation module.

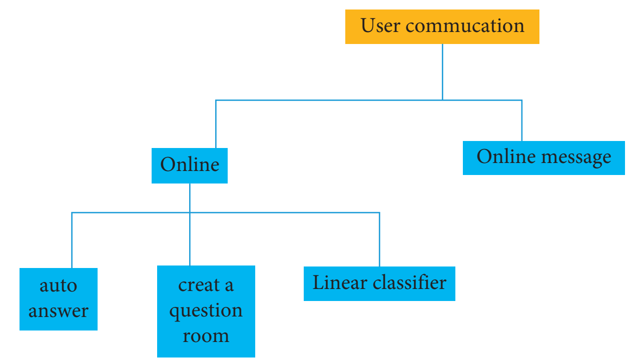


Figure 10: Schematic diagram of user communication module.

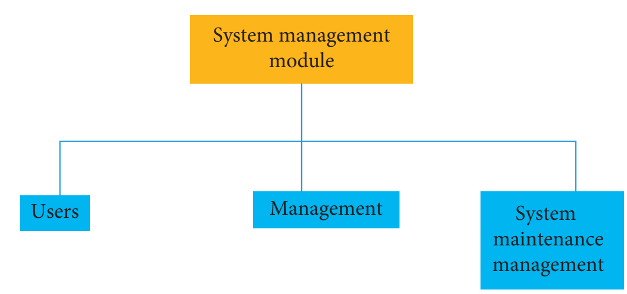


Figure 11: Function module diagram of background management module.

should take advanced online social education as the core, build a technical system as soon as possible, ensure the innovation and development of logistics industry in the process of accelerating technology research and development, and improve the technical level of Internet of things of logistics enterprises. Among them, based on the current national policies, we should pay more attention to the research and application of Internet of things technology, form a high-quality Internet of things technology talent team in the process of integrating and concentrating a variety of technical resources, as logistics enterprises, on the one hand, and promote the further enhancement of the intelligent level of logistics information management system on the other hand. The human pulse is measured by the photoelectric reflective analog sensor, and the user wears it on the finger or earlobe. The change cycle of the electrical signal of the photoelectric converter is the pulse rate. Through wire connection, the collected analog signal can be transmitted to MCU, the pulse value can be calculated simply, and the general situation of the body can be obtained according to the pulse data. Another index: body temperature, tn901 infrared thermometer without contact with human body that is used to measure the ambient temperature and the full wavelength thermal radiation of the radiation object through thermopile, determine the radiation temperature of the object, and quickly and accurately measure human body temperature by temperature compensation technology [24]. Compared with traditional Mercury and electronic thermometer, it is easy to use, and it can measure temperature without touching the body. It is not only hygienic but also safe and has the advantages of high sensitivity, high precision, and low power consumption. In addition, using the compass of modern smart phones can quickly realize the direction guidance without additional hardware as shown in Figure 12.

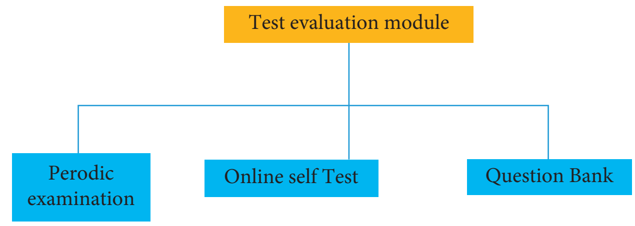


Figure 12: Test system.

3.4. Management and Rescue Mechanism. Field exploration activities are the inevitable product of social and economic development to a specific stage. They are related to the promotion of national spirit and the self-development of citizens. They should be supported and encouraged from the legal system level, and their healthy and standardized development should be ensured through appropriate control measures and rescue mechanisms. First, we should focus on prevention and early warning, supplemented by rescue. Figure 13 shows the emergency accident rate of field exploration in recent years. The legal norms related to field exploration should focus on establishing and improving the prevention and early warning system, including the implementation of preventive measures such as the cultivation of risk awareness of field exploration, safety skill training, and exploration route planning, as well as early warning measures such as the release of unsuitable weather information and sudden natural disaster information. Rescue is not the purpose, but it is also indispensable. The legal norms related to field exploration should also stipulate the basic framework of rescue mechanism, including rescue subjects, rescue organizations, rescue funds, and other relevant contents [25]. Second, encourage and support, supplemented by management and control constraints. That is, the citizens' field exploration should be liberalized from the activity area as much as possible, the participants should be relaxed as much as possible, and the participation form should be supported as much as possible. At the same time, in line with the principle of appropriate control and restraint, the government departments, industry associations, or relevant clubs that must be managed should be appropriately restrained and managed according to the situation. The combination of the two can complete the rescue task in the fastest time, the greatest effect, and the least cost. The location formula displayed by the online things according to the signal sent by the intelligent system is shown in the following formula:

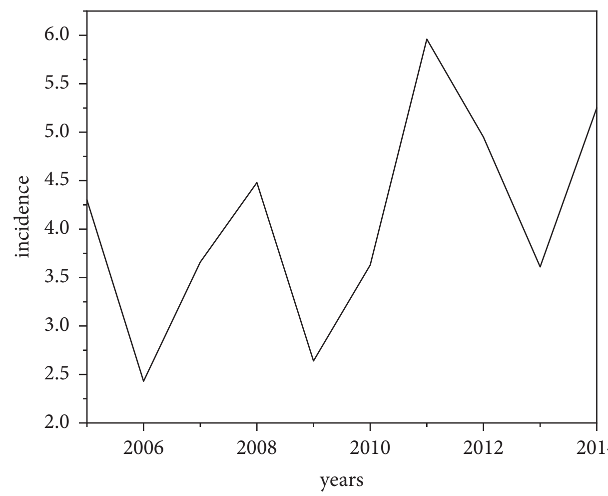


Figure 13: Emergency accident rate of field exploration in recent years.

$$V_i^d = Wv_i^{d-1} + c_1 r_1\left(p\text{best}_i^d - x_i^d\right) + c_2 r_2\left(g\text{best}^d - x_i^d\right). \quad (4)$$

Among them, C_1 and C_2 are individual learning factors. During the d-th iteration of V_i^d, that is, the i-th speed of the Internet of things, during the d-th iteration of X_i^d, the i-th position is expressed, $p\text{best}_i^d$ is the best position up to the d-th iteration, and gbest^d is the position of all personnel up to the d-th iteration. As shown in Figure 14, system quality management flow chart is shown.

3.5. Standardize Personnel Behavior. First, establish and improve the approval or filing system for specific exploration activities. Necessary preconditions can be set for administrative approval or filing, such as requiring field explorers to submit the exploration qualification or experience certificate of the leader and its members, exploration equipment, experience certificate of guides or collaborators, and the proposed travel route. If the examination and approval or filing department considers that the conditions are not met, or the degree of danger is too high, it can require field explorers to improve the conditions or change the route to reduce the possibility of risk in advance. The equipment and success rate of emergency rescue are shown in Figures 15 and 16. Professional qualification management can improve the risk identification and response level of field explorers, such as establishing and improving the professional

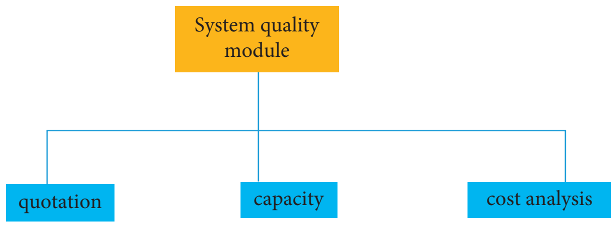


Figure 14: System quality management flow chart.

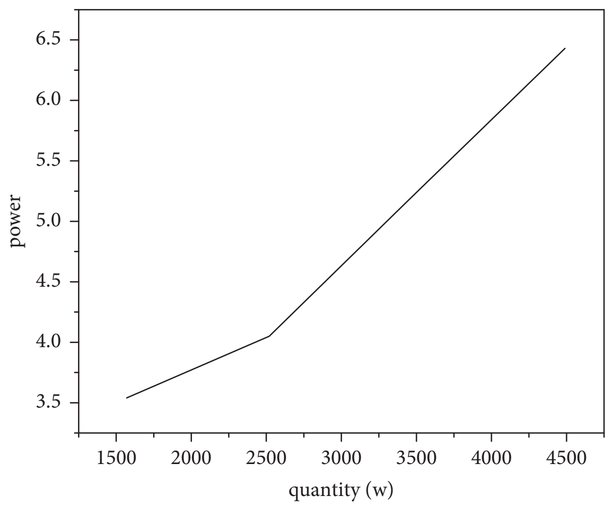


Figure 15: Number of IOT devices used in field exploration.

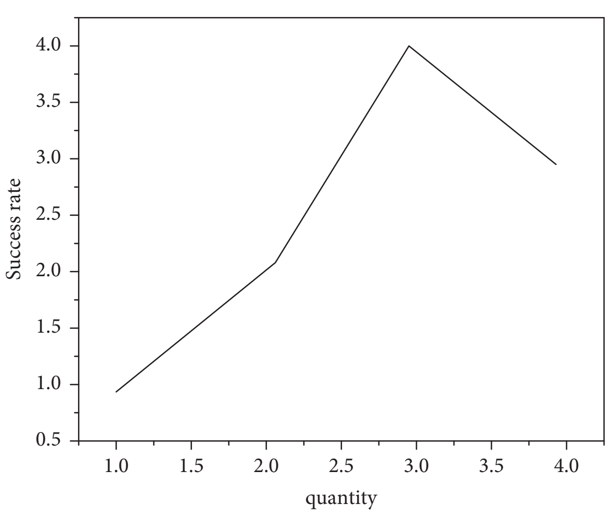


Figure 16: Success rate of intelligent emergency rescue.

qualification management system of exploration leaders, alpine collaborators, and divers by specifying physical fitness, skills, qualifications, and other conditions. Third, establish and improve the punishment system for illegal exploration [26]. Punishment is not an end, but it is an effective management means. In order to standardize the exploration behavior of field explorers in China, appropriate administrative penalties for violations should be set. Explore and establish a "blacklist" system for field exploration, implement blacklist management for field exploration organizers and their participants who do not report for approval and filing according to the regulations and change the approved route or behavior without authorization, give certain restrictions in terms of qualification access, fee payment, training, and education, and impose a lifelong ban for serious cases. When the project's requirements are all met, relevant data can also be considered to be included in the national personal credit investigation system to increase its violation cost.

At the same time, in the definition of civil legal liability, we should clarify the safety and security obligations of organizers and the principle of self-responsibility of participants. With the city and the closer connection between people, there are obvious potential safety hazards in some dense field exploration areas. To solve the problem of living safety of field personnel, technicians can use online social things technology to arrange various sensors in the exploration area. In case of danger in the residential area, the alarm will give an alarm in time to remind people to escape from the dangerous area; Meanwhile, the security system based on Internet of things technology will automatically alarm and transmit the corresponding distress signal to the police and rescue units. After receiving the signal, the rescue personnel will formulate the corresponding emergency rescue plan in time according to the dangerous situation. Through the online social of things, the position error of rescuers can be reduced. The hidden formula is as follows:

$$\mathrm{net}_j = \sum_{i=0}^{i=n} v_{ji} x_j, \text{among}, \ j = 1, 2, \ldots\ldots, m. \quad (5)$$

4. Results and Analysis

4.1. Performance and Safety Analysis. The construction of intelligent field emergency rescue scene has become a major demand of current social development. When constructing intelligent field emergency rescue scene, Internet of things technology, 5g communication technology, and artificial intelligence technology should be scientifically applied. For the construction of the current intelligent field emergency rescue scene, the Internet of things technology plays a major role. Relevant technicians need to recognize the application and development trend of Internet of things technology in intelligent field emergency rescue scenarios and take various measures to promote it, so as to make Internet of things technology better support, enrich intelligent field emergency rescue functions, and better provide impetus for intelligent field emergency rescue and the construction of digital city. Field exploration and its rescue legislation can learn from the experience of relevant countries and regions, but it is more important to explore the path model suitable for China's field exploration and rescue mechanism legislation based on Chinese characteristics. The analysis results are shown in Figures 17 and 18. First, some pilot projects will be carried out first, and then the model will be comprehensively promoted. That is, first carry out the legislative pilot of field exploration and rescue mechanism in some areas and

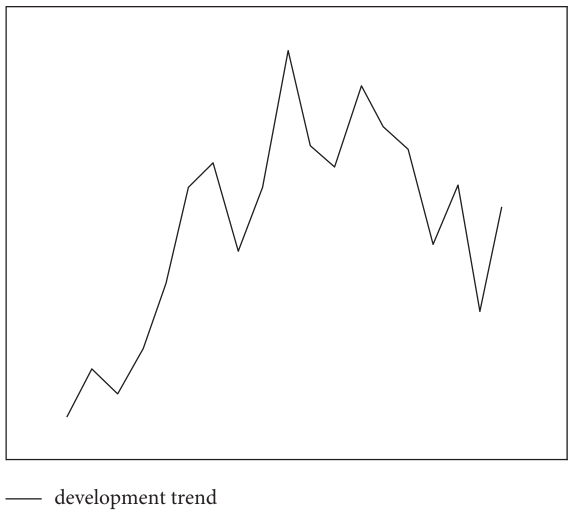


Figure 17: Application and development trend of internet of things technology in intelligent field emergency rescue scenario.

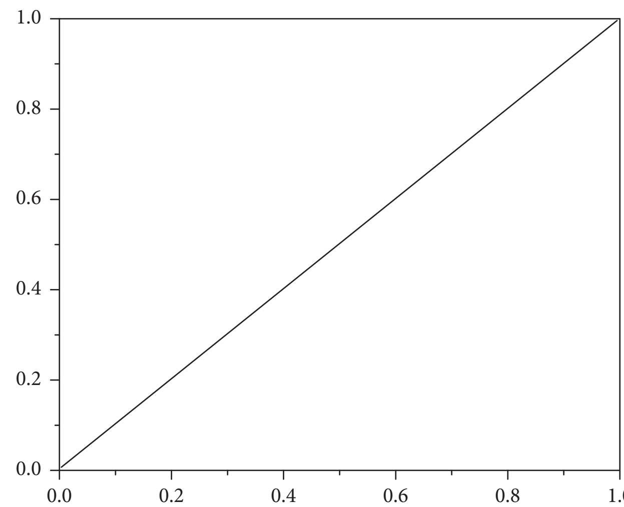


Figure 19: Privacy of intelligent terminal equipment.

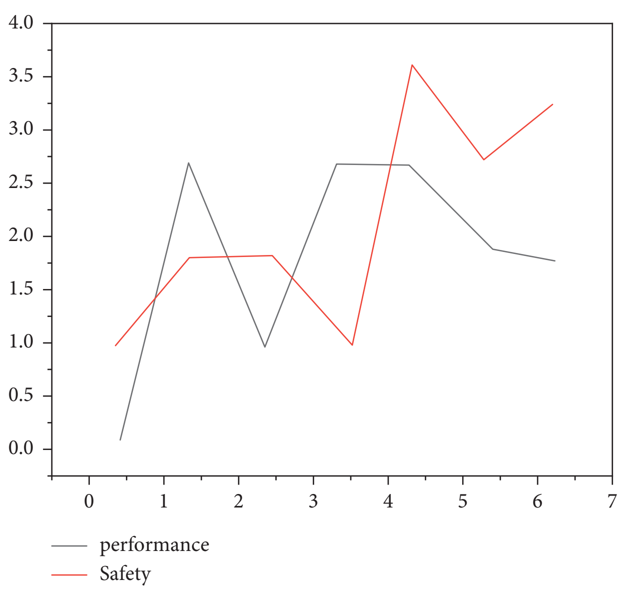


Figure 18: Performance and safety test of intelligent emergency system.

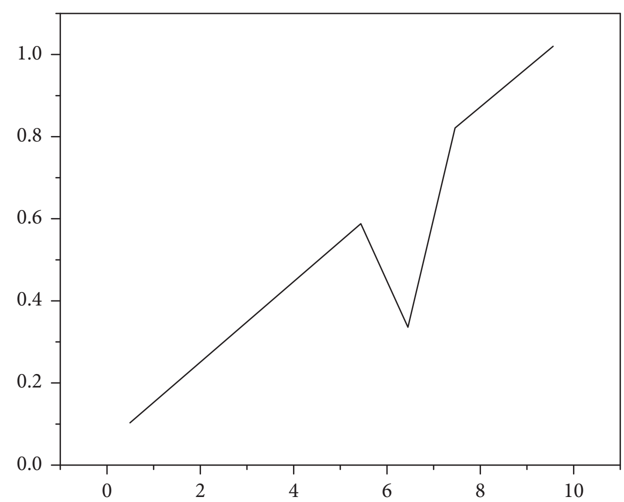


Figure 20: Location accuracy of terminal system.

industries with relatively developed exploration activities. After the development of field exploration activities to a certain extent, formulate national legal norms of field exploration and rescue mechanism in combination with the legislative experience of pilot areas and industries according to the needs. With the continuous development of social economy and computer technology, various industries are widely using computer technology. However, in the process of use, the security of online information is a problem in the currently project. The online pen social has become more complex. In the computer industry, network security is an important problem. The field intelligent rescue system makes good use of the shared position function, which can avoid accidents such as team members falling behind and can also well identify the position of team members.

4.2. Platform Design Analysis and Test. Internet of things technology aims to realize the connection between things and even the interconnection of all things, which requires that when applying Internet of things technology, firstly, technicians should use very stable wireless interconnection technology to do a good job in the connection between things. Secondly, technicians should clarify the essence of Internet of things technology. The essence of Internet of things technology is related to the Internet. Internet of things is an Internet based equipment control network. Moreover, when designing the Internet of things system, technicians must pay attention to strengthening the security of the system. This is because when users use the Internet of things technology, they are often required to upload some sensitive information of users, such as fingerprint information or face, voice, and other information, which has extremely strong privacy. The test results are shown in Figures 19 and 20. Any smartphone can use a compass to determine the direction of its current location. The new route memory function can effectively remember the current user's route and help the user get out of the wrong way. In addition, it also has the function of teammates forming teams and sharing team members' positions, which can effectively solve the problem of teammates falling behind and always query the whereabouts of teammates, so as to facilitate carrying out land exploration activities by the team. The new scenic spot information promotion function can provide nearby known scenic spots for the exploration crowd and provide an exploration destination for the crowd. The above is what most mainstream navigation devices lack at present, which is the significance of designing this system.

5. Conclusion

In short, based on the analysis of the basic connotation and application requirements of online of things technology, this paper further analyzes the problems and corresponding solutions of Internet of things technology in field emergency rescue and mainly discusses the problems of platform management, technology application cost, and security. When designing the online project social system, technicians must combine the actual needs of users and should refer to the application standards of social things technology to protect user information and improve the application quality of things technology. In addition, in order to optimize the user's application experience, when selecting intelligent rescue equipment, designers should also select intelligent rescue equipment with unified operation platform, which can eliminate the information sharing problems caused by platform mismatch to a certain extent. In addition, intelligent equipment manufacturers should also actively apply social development things, collect and analyze users' feedback and opinions, optimize the system operation interface based on users' actual needs, humanize the application process of Internet of things technology, and then effectively improve the application quality of Internet of things technology. Combined with the use of software and hardware, it can realize the functions of real-time monitoring direction, human body temperature, ambient temperature, positioning and navigation, route memory, shared location, and scenic spot information. It focuses on the two functions of route memory and shared location, which can effectively solve the defects of single function and inflexibility of the current navigation system. This intelligent system has high practical value. It is extremely practical for field explorers who like outdoor sports. It is easy to carry, is simple to operate, and has great commercial potentials.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] Z. Sebastian, "Zulssigkeit politischer artikel und einer umfrage zu bevorstehender wahl in internetmedium," *AfP*, vol. 52, no. 4, pp. 298-299, 2021.

[2] Q. Song, N. Zhang, and H. Liang, "Review of the Chinese internet philanthropy research (2006-2020): analysis based on citespace," *The China Nonprofit Review*, vol. 13, no. 2, pp. 3-4, 2021.

[3] L. Zhou, M. Ying, and J. Wu, "Conceptualising China's approach to 'internet plus government services': a content analysis of government working plans," *Information Development*, vol. 37, no. 4, pp. 633–646, 2021.

[4] R. Atiqur, "Automated smart car parking system for smart cities demand employs internet of things technology," *International Journal of Informatics and Communication Technology*, vol. 10, no. 1, pp. 46–53, 2021.

[5] G. Rathee, R. Iqbal, and A. Khelifi, "Decision making in internet of vehicles using pervasive trusted computing scheme," *Computers, Materials & Continua*, vol. 68, no. 2, pp. 2755–2769, 2021.

[6] D. Hansaraj, S. Wankhede, N. V. Gowri, and P Sharma, "An internet of things for prevention of security attack on cloud medical data using artifical intelligence," *International Journal of Grid and Distributed Computing*, vol. 14, no. 1, pp. 1143–1161, 2021.

[7] L. Zhen, Y. Zhang, K. Yu, N. Kumar, A. Barnawi, and Y. Xie, "Early collision detection for massive random access in satellite-based internet of things," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 5, pp. 5184–5189, 2021.

[8] Y. Xu, "Application of green building design based on the internet of things in the landscape planning of characteristic towns," *Advances in Civil Engineering*, vol. 2021, no. 3, pp. 1–11, Article ID 6317073, 2021.

[9] G. Iyer and R. Doraiswamyiyer, "Internet addiction in freshmen engineering students in India," *Journal of Management & Public Policy*, vol. 12, no. 1, pp. 45–58, 2021.

[10] A. Duishonbaeva, A. Apysheva, G. Bekmyrzaeva, T. Aitbaev, E. Salieva, and Z. Abdullaeva, "Main trends in development of media and internet resources in the independent Kyrgyzstan," *Advances in Journalism and Communication*, vol. 09, no. 02, pp. 74–83, 2021.

[11] H Heriansyah and S. Istiqphara, "Optimasi sistem cerdas pada pengering tanaman obat berbasis internet of thing dengan memanfaatkan sumber energi terbarukan," *Electrician*, vol. 14, no. 1, pp. 7–13, 2020.

[12] Z. Li, X. Zhang, Y. Wang, and X. Su, "Predicting the sequential behavior of mobile internet users based on msm model," *International Journal of Market Research*, vol. 62, no. 6, pp. 743–757, 2020.

[13] Q. Li, "Research on postmodernism educational trend of thought in the era of internet," *Advances in Education*, vol. 10, no. 04, pp. 577–584, 2020.

[14] V. K. Singh, H. Chandna, A. Kumar, S. Kumar, N. Upadhyay, and K. Utkarsh, "Iot-q-band: a low cost internet of things based wearable band to detect and track absconding covid-19 quarantine subjects," *EAI Endorsed Transactions on Internet of Things*, vol. 6, no. 21, pp. 163997–164197, 2020.

[15] E. I. Galyashina, "The concept of information (worldview) security of internet-media communication in the aspect of forensic speech examination," *Courier of Kutafin Moscow State Law University*, vol. 22, no. 6, pp. 33–43, 2020.

[16] K. T. Mursi and Y. Zhuang, "Experimental study of component-differentially-challenged xor pufs as security primitives for internet-of-things," *Journal of Communications*, vol. 15, no. 10, pp. 714–721, 2020.
[17] B. K. Sihotang, S. Sumarno, and B. E. Damanik, "Implementasi access control list pada mikrotik dalam mengamankan koneksi internet koperasi sumber dana mutiara," *JURIKOM (Jurnal Riset Komputer)*, vol. 7, no. 2, p. 229, 2020.
[18] H. I. Kusuma, E. Harnelly, Z. Thomy, M. A. Fitra, and S. Samingan, "Exploration of potential wild medicinal mushrooms from pocut meurah intan forest park, aceh, Indonesia," *Jurnal Natural*, vol. 20, no. 3, pp. 66–73, 2020.
[19] F. Santicchia, L. A. Wauters, B. Dantzer et al., "Relationships between personality traits and the physiological stress response in a wild mammal," *Current Zoology*, vol. 66, no. 2, pp. 197–204, 2020.
[20] L. Harten, A. Katz, A. Goldshtein, M. Handel, and Y. Yovel, "The ontogeny of a mammalian cognitive map in the real world," *Science*, vol. 369, no. 6500, pp. 194–197, 2020.
[21] S. Chavan and A. Jondhale, "Exploration and documentation of some more wild edible food plants from tribal region of peth tehsil, nashik district (Maharashtra.) India," *International Journal of Botany and Research*, vol. 6, no. 2, pp. 360–366, 2021.
[22] M. Oun, A. A. Majeed, and S. R. Oleiwi, "Biotherapy exploration for gallbladder stones removing in vitro," *International Journal of Drug Delivery Technology*, vol. 11, no. 1, pp. 147–152, 2021.
[23] A. Nakhforoosh, K. A. Nagel, F. Fiorani, and G. Bodner, "Deep soil exploration vs. topsoil exploitation: distinctive rooting strategies between wheat landraces and wild relatives," *Plant and Soil*, vol. 459, no. 1-2, pp. 397–421, 2021.
[24] G. A. Mayer, "Anthropology and cryptozoology: exploring encounters with mysterious creatures edited by samantha hurn and chris wilbert (in the series multispecies encounters)," *Journal of Scientific Exploration*, vol. 34, no. 1, pp. 146–154, 2020.
[25] C. Martina, G. Cowlishaw, and A. J. Carter, "Individual differences in task participation in wild chacma baboons," *Animal Behaviour*, vol. 172, no. 1, pp. 73–91, 2021.
[26] S. Shamsi, L. Rogers, E. Sales, R. K. Kopf, and R. Freire, "Do parasites influence behavioural traits of wild and hatchery-reared murray cod, maccullochella peelii?" *Parasitology Research*, vol. 120, no. 2, pp. 515–523, 2021.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9790252, 1 page
https://doi.org/10.1155/2023/9790252

Retraction

Retracted: Integration and Recommendation of Multimedia Network-Assisted English Instructional Resources Based on Association Rules Mining

Mobile Information Systems

Received 22 August 2023; Accepted 22 August 2023; Published 23 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] L. Liu, "Integration and Recommendation of Multimedia Network-Assisted English Instructional Resources Based on Association Rules Mining," *Mobile Information Systems*, vol. 2022, Article ID 8806525, 10 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 8806525, 10 pages
https://doi.org/10.1155/2022/8806525

Hindawi

Research Article

Integration and Recommendation of Multimedia Network-Assisted English Instructional Resources Based on Association Rules Mining

Lihong Liu

Business School, Guangxi Modern Polytechnic College, Hechi 547000, Guangxi, China

Correspondence should be addressed to Lihong Liu; rainbownvren02@163.com

Received 27 May 2022; Revised 11 July 2022; Accepted 18 July 2022; Published 13 August 2022

Academic Editor: Jiafu Su

The integration of multimedia teaching refers to a novel teaching mode that combines multimedia, information resources, and information methods with content and teaching process in the teaching process with the guidance of contemporary educational ideas and organically unifies them on a spatial and temporal platform, so as to better accomplish the teaching tasks. However, with the deepening of network teaching and school connotation construction, various resources such as media materials, courseware, network courses, and teaching cases generated by education and teaching, as well as academic works, papers, patents, project results, and lectures of experts and professors gathered by scientific research, are increasing. In this situation, how to integrate teaching resources and courses to provide students with an independent learning platform and an information-based learning environment has become an urgent issue. For this problem, data mining has shown a strong vitality, and data mining can find out the potential connection between datasets. Association rule mining is the most researched and widely used data mining method. In this paper, we propose a multimedia network-assisted English integration of educational resources method based on association rule algorithm to give full play to the role of "multimedia technology" and realize the integration of traditional English teaching and information network culture, so as to achieve the purpose of optimizing English teaching. The experimental results show that with the increase of data volume, the average testing time of Apriori is reduced by 56.49 s compared with MapReduce, so the advantage of Apriori algorithm is more obvious. Therefore, this research can effectively provide students with rich and colorful personalized resources, realize high-performance interactive learning, and help students cultivate their independent learning ability as well as their lifelong learning habits.

1. Introduction

For the purpose of meeting the needs of knowledge-based society and constructing an education system of lifelong education and learning, China has envisioned the development of distance education using the advantageous resources of general schools since early years [1]. Modern IT represented by networking, digitalization, multimedia, and intelligence has become the main tool to expand people's ability and quality [2]. It has changed people's traditional way of living, learning, and working and has brought a profound impact on modern education content and teaching methods, which has also triggered a boom in the integration of IT and curriculum [3]. However, the application of multimedia requires certain technology and skills, and in reality, the application level of teachers varies, and the effect of multimedia technology is not uniform [4]. Therefore, we should broaden the field of English learning and application and pay attention to interdisciplinary learning and the use of modern technology, so that students can broaden their horizons and improve their learning efficiency in the intersection, penetration, and integration of different contents and methods and initially acquire the practical English skills required by modern society [5].

As human beings enter the era of knowledge economy, knowledge is being updated and accumulated at an extremely fast rate [6]. The heterogeneity and tight coupling of development platforms and tools have led to a large number of educational resource systems that are not interoperable, resource sharing, and software reuse, and the phenomenon of "information silos" still prevails in schools [7]. Association rule mining comes as a major branch of data mining research, and association rule is the most typical type of knowledge among many types of data mining [8]. Most teachers have no idea how to search for the required teaching resources and how to integrate the useful resources they find together, resulting in a vast amount of high-quality resources not being fully utilized [9]. Therefore, the integration of IT and teaching curriculum is necessary to deepen the curriculum content and teaching reform, and it is also an effective way to improve teaching efficiency, which has attracted widespread attention and exploration research of educators [10]. Association rule mining can discover interesting relationships between items or attributes that exist in the database, which are unknown and hidden in advance; that is, they cannot be derived by logical operations of the database or by statistical methods [11]. In order to give better play to the advantages of multimedia and improve the level of English teachers' application of multimedia teaching, not only is it necessary to further clarify the basic theoretical issues such as the connotation of integration of multimedia and curriculum teaching, the ways of integration, and the strategies of integration, but there is also an urgent need to solve various problems and confusions that arise in the real integration of practical teaching.

This is because multimedia have the advantage of being graphic, audio and visual, and large capacity, which can present teaching content in a timely manner, show materials for classroom language practice activities, link teaching sessions, and supplement the background corpus of relevant topics [12]. At the same time, teaching resources are the basic conditions for training innovative talents and conducting research activities in higher education [13]. After all, the excellent educational resources of a school are limited, and the integration of the excellent educational resources has great limitations [14]. It is necessary to integrate the existing educational resources in order to meet the needs of higher education as much as possible under the existing conditions, from the point of view of both the current situation of educational resources and the current challenges of education [15].

The innovation points of this paper are as follows:

(1) The multimedia online platform is adopted as the basic platform for the integration of teaching resources to achieve the integration of existing teaching resources effectively.

(2) Using the association rule mining to study and evaluate the information of the existing teaching resources, we propose some ways to integrate the resources through scientific methods so that we can evaluate the teaching quality and students' learning quality.

(3) The research of multimedia network-assisted English integration of educational resources system based on association rule algorithm is to improve the efficiency and accuracy of English teachers in schools, free them from tedious and boring work, and then improve their management level teaching quality.

The research framework of this paper contains five major parts, which are organized as follows.

The first part of this paper introduces the background and significance of the research and then introduces the main work of this paper. The second part introduces the work related to multimedia network-assisted English integration of educational resources, association rules algorithm. The third part of the paper introduces the design of the computer network integration of educational resources system and the implementation of the computer network integration of educational resources system, so that the readers of this paper can have a more comprehensive understanding of the idea of multimedia network-assisted English integration of educational resources based on the association rule algorithm. The fourth part is the core of the thesis, from the analysis of the integration of educational resources system using Apriori algorithm and the solution space analysis of the optimized association rules, to completing the description of the application analysis of the association rule algorithm in the integration of multimedia network-assisted teaching resources. The last part of the paper is the summary of the whole work.

2. Related Work

2.1. Multimedia-Assisted English Integration of Educational Resources. During the implementation of school informatization construction, the construction of an integrated environment is a top priority. Especially in the current period when most schools are still in the application integration and information integration, the effective sharing and development of educational information resources, which is an important part of the whole informatization system, is becoming a key issue. With the help of educational information resource integration, it is greatly convenient for teachers to integrate various useful resources and make rational use of them, thus effectively solving the problem of insufficient technical personnel and teachers' lack of IT skills.

Tian [16] proposed a composite digital object composed by Structural Kernel and Disseminator Layer to separate data and operations on data [16]. Zhang and Wang [17] proposed a data integration strategy for enterprise information applications in isolation and analyzed the application integration and data combination to describe the standardization of information in the data integration process [17]. Liu et al. [18] proposed the integration of IT and subject teaching, and the integration of IT and curriculum has been included in the outline of the new curriculum reform [18]. Peng [19] defines "curriculum and integration of educational resources platform" as follows: a curriculum and integration of educational resources platform is a platform for improving the quality of education and teaching by using information technology to organically integrate various resources, technologies, and environments that help and facilitate students' learning with the curriculum content they are learning [19]. Wang and Gao [20] published the idea of optimizing students' English reading through multimedia network teaching resources, and they wanted to use concrete image resources to make the abstract information in reading concrete and thus facilitate the learning of English reading for the majority of students [20].

The application of IT in subject teaching not only inherits the advantages of traditional subject teaching but also can assist subject teaching to create a vivid context, with the relatively independent characteristics and advantages of the information age. Integration can make the elements in the system achieve overall coordination and interpenetration, so that each element of the system can maximize its effectiveness, and this process will lead to the generation of a new thing. Therefore, the efficient sharing, integration, and development of educational resources to ensure the rational, orderly, and healthy in-depth development of education informatization construction are an important research direction for school informatization construction.

2.2. Association Rule Algorithm. Most countries in the world, represented by the United States, pay great attention to the planning and standards of IT integration applied to teaching and learning when formulating curriculum standards and take IT and teaching curriculum integration as an important element in the new curriculum reform. However, the traditional way used to integrate the existing education resource system will cause problems such as long cycle time and large investment, and at the same time there is a lack of unified standards, which makes it difficult to fully ensure the orderly, compact, and optimal resource integration services. Therefore, the successful application of association rule mining in business and other fields has made it the most mature, important, and active research content in data mining.

Zhao et al. [21] first proposed in their analysis of the market basket problem to discover customer buying patterns in merchandising, which can be used to guide merchants in the scientific arrangement of stocking, inventory, and shelf design [21]. Yang [22] proposed algorithm that uses a division-based technique to partition the database processing effectively reducing the number of database scans during the mining process and reducing the burden [22]. Yu et al. [23] argue that the study of visualization methods for data mining process enables or makes the knowledge discovery process easy to understand for users on the one hand and improves the human-computer interaction in the knowledge discovery process on the other hand [23]. Xie [24] used sampling from the transactional database to get some rules that may hold in the whole database and then verified this result for the remaining part of the database, significantly reducing the cost of input and output [24]. Bai and Zhang [25] proposed an association rule algorithm based on the MapReduce computing framework under Hadoop; it takes its candidate item set for each transaction record by Map function and then uses the Reduce function for normalization to obtain the frequent item set [25].

Association rule mining can discover rules or patterns that cannot be obtained by traditional artificial intelligence and statistical methods and therefore has important research value. The integration of educational resources with association rule mining is hoped to satisfy the current demand for quality and quantity of teaching resources in higher education by making full use of existing teaching resources.

3. The Idea of Integrating Multimedia Web-Assisted English Education Resources on the Basis of Association Rule

3.1. Computer Network Integration of Educational Resources System Design. The general idea of the design of the computer network integration of educational resources system is to adhere to the people-oriented and establish a comprehensive, coordinated, and sustainable development concept as the overall leader. The aim is to comprehensively improve the level of information construction and application of the school and fulfill the objectives of talent training [26]. The integration of English teaching resources in the IT environment requires the use of modern IT, with computers and networks as the core, as a cognitive tool for promoting students' independent learning and a tool for creating a rich teaching environment under the guidance of advanced educational thought and theory [27]. The overall structure of the computer network integration of educational resources system is shown in Figure 1.

First of all, the system platform is designed. The computer network integration of educational resources system based on B/S mode adopts a multilayer application architecture, and the user working interface is all realized through a web browser. Given a transaction database with a minimum support of $w\min\mathrm{Sup}$ for user input, if the t attribute set X is frequent, then its support number and minimum support number should, respectively, satisfy

$$
\begin{aligned}
SC(X) &\geq \frac{w\min\mathrm{Sup}\times T}{\mathrm{MAX}(w_i)},\\
B(X) &= \left\lceil \frac{w\min\mathrm{Sup}\times T}{\mathrm{MAX}(w_i)} \right\rceil.
\end{aligned}
\tag{1}
$$

The process of schema creation is the process of data mining [28]. XML is a new standard language used to automatically describe information, which enables computer communication to extend the function of the Internet from information transfer to a wide variety of other human activities [29]. The interestingness measure of XML was utilized to measure the interestingness of association rules, which is defined as the product of the joint probability density of two variables divided by the expected probability of these two variables:

$$
I(A,B) = \frac{P(A,B)}{P(A)P(B)} = \frac{f_{11N}}{f_{1+}f_{+1}}. \tag{2}
$$

When there are two m dimensional vectors $X_1 = (x_{11}, x_{12}, \ldots, x_{1n})$ and $X_2 = (x_{21}, x_{22}, \ldots, x_{2n})$, the Euclidean distance between them is

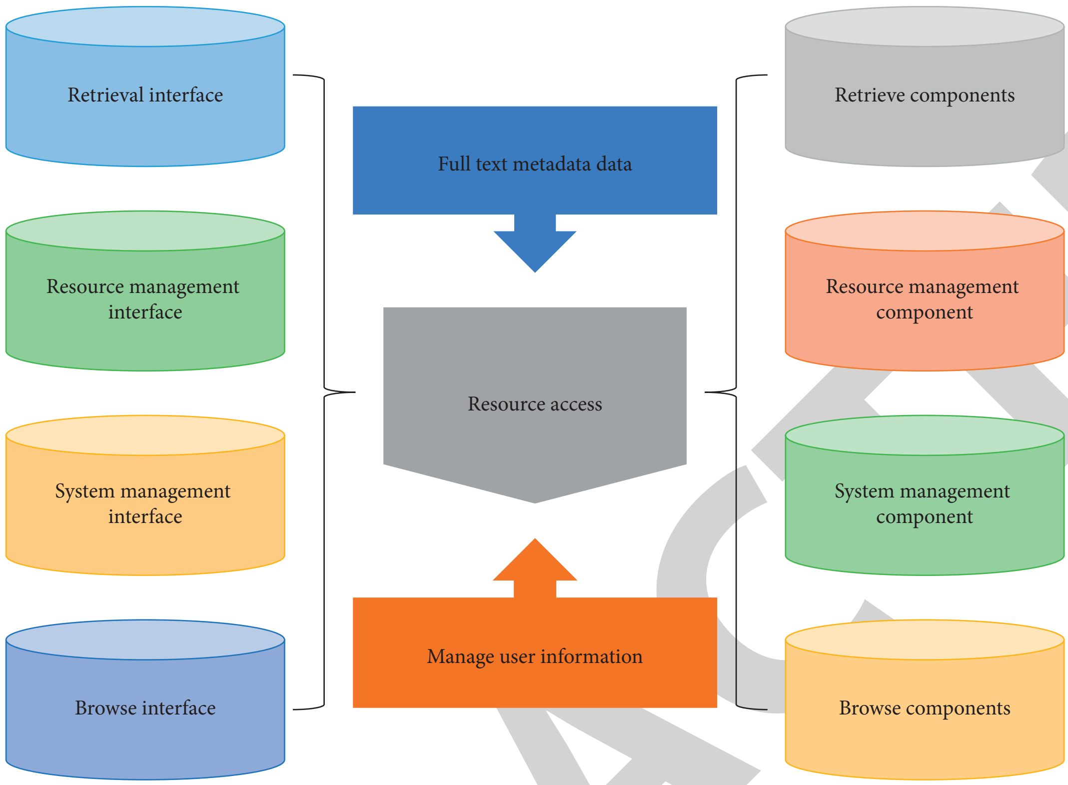


Figure 1: The overall structure of the computer network integration of educational resources system.

$$\operatorname{dist}(X_1, X_2) = \sqrt{\sum_{i=1}^{n} (x_{1i} - x_{2i})^2}. \quad (3)$$

According to the platform of operation, it can be divided into online version and stand-alone courseware [30]. To integrate teaching information resources, online education is used as the basic platform to integrate all teaching information resources that are scattered in libraries, resource rooms, and information resource centers without effective planning. We first standardize them and then reorganize them to form systematic teaching information resources that can be shared on a large scale. The workflow of the computer network integration of educational resources system is shown in Figure 2.

Next is the functional design of the system. In the collection and submission module of this resource integration system, in addition to the system administrator, resource maintenance personnel need to register to fill in personal information and add it to the corresponding faculty and topic. After logging into the system, they enter "My Workspace" and complete the description, citation, and upload of resources through the submission interface. Therefore, data mining technology is required to respond quickly to changes in the data to provide decision support. The level-weighted support of the item set $X = \{i_1, i_2, ..., i_k\}$ is defined as

$$\mathrm{Sup}_h = \max\{h_1, h_2, \ldots, h_k\} \times \mathrm{Sup}(X). \quad (4)$$

$\mathrm{Sup}(X)$ is the traditional support count of item X and $\mathrm{Max}\{h_1, h_2 \ldots, h_k\}$ is the weight of items.

The sum of teaching contents and implemented teaching activities of a course expressed through the network include teaching contents and network teaching support environment organized according to certain teaching objectives and teaching strategies. However, when multiple machines work together to process data, various problems can occur; for example, local node crashes can lead to errors in the operation of applications, and the failure of switches and routers can bring communication failures to the local network. Therefore, by taking advantage of the technical advantages and high compatibility of multimedia network assistance, we establish a unified search interface for teaching information resources with cross-platform search and optimize fuzzy query, so that users can retrieve resources that meet the needs in all databases by only one search. The system calculates the fuzzy support degree based on the results of the fuzzy query, which is calculated as follows:

$$\mathrm{prop} = \frac{\sum_{i=1}^{n} T\left(\mu_{C\text{-much}}(T_i), T\left(\mu_{A\text{-much}}(T_i), T\left(\mu_{B\text{-low}}(T_i)\right)\right)\right.}{\sum_{i=1}^{n} T\left(\left(\mu_{A\text{-much}}(T_i), \mu_{B\text{-low}}(T_i)\right)\right)}. \quad (5)$$

i is the article i record, T is the T-norm operation, and T_i is the record.

Finally, the database is designed. The system uses Postgre SQL, an open-source relational database product, and JDBC as the database driver. The web-based courseware needs to run in a standard browser and can be shared through the

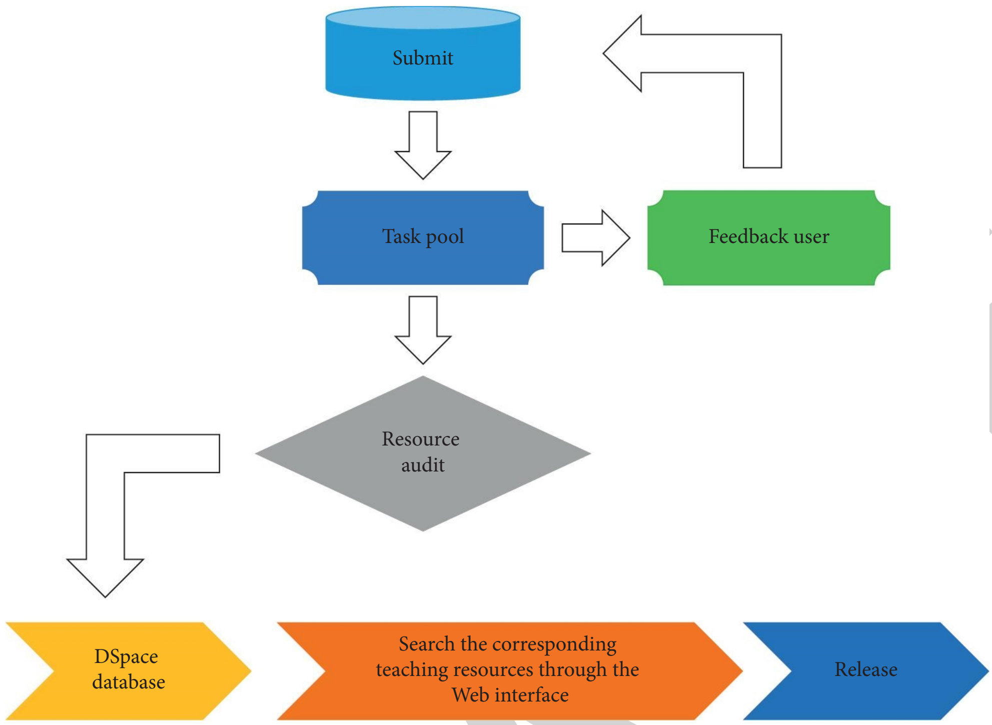


Figure 2: The overall process of resource integration system.

web-based teaching environment, while the stand-alone courseware can be downloaded from the web and run on a local computer. Because of the richness and diversity of resources, it is not easy for learners to find the resources they need, and providing a good in-site search can improve efficiency and increase the utilization of resources at the same time. The mining of association rules focuses on records with certain common characteristics, making the goal of mining closer to the needs of users. By making full use of the good communication environment of multimedia network and the function of recording the whole teaching process, we can obtain timely feedback information about the content and quality of teaching information resources and optimize them continuously.

3.2. The Realization of Network Integration of Educational Resources System. At present, the main task of the computer network integration of educational resources system is to integrate the four categories of various catalog resources in the university's collection, educational teaching affiliated resources, scientific research affiliated resources, and various free network resources. At the same time, all kinds of resources added to the college in the future, including the external digital resource library, are within the scope of integration of this system. The association rule mining includes the process of understanding and proposing the problem, data collection, data processing, data transformation, data mining, pattern evaluation, and knowledge representation, as shown in Figure 3.

Firstly, we collect relevant resources and plan them in a unified way and implement the digitalization of resources. Then, according to the system's resource classification system, the resource maintenance personnel are responsible for uploading and submitting teaching resources to the corresponding faculty or subject specialties. Students can submit electronic assignments to teachers through the network, and the approved assignments can be freely accessed by students, and feedback can be given to students quickly. Integrated management is a proven method to improve the efficiency of resource management and utilization of school teaching resources. The required association rules are generated by using frequent item sets, and then the strong association rules are obtained by rounding according to the minimum confidence level set by the user. Thus, the feature vector is used to describe the related data with strong similarity in the database. The following equation enables the calculation of variation parameters for different data attribute features:

$$e(y,z) = \left[\sum_{j=1}^{p} \left|y_j - z_j\right|^s\right]^{1/s}. \tag{6}$$

y is the amount of data in the database, z is the number of attributes, and s is the differences of data characteristics.

The language teaching process is a dynamic process, which should follow certain rules. However, teaching methods are not static and should be used for different teaching tasks and objects, which requires teachers to adopt different teaching strategies according to different situations. The operations of access and the request/response messages used for access are abstractly described and then bound to specific transport protocols and message formats to finally define the service access points for specific deployments. For comparative analysis of factor data with different

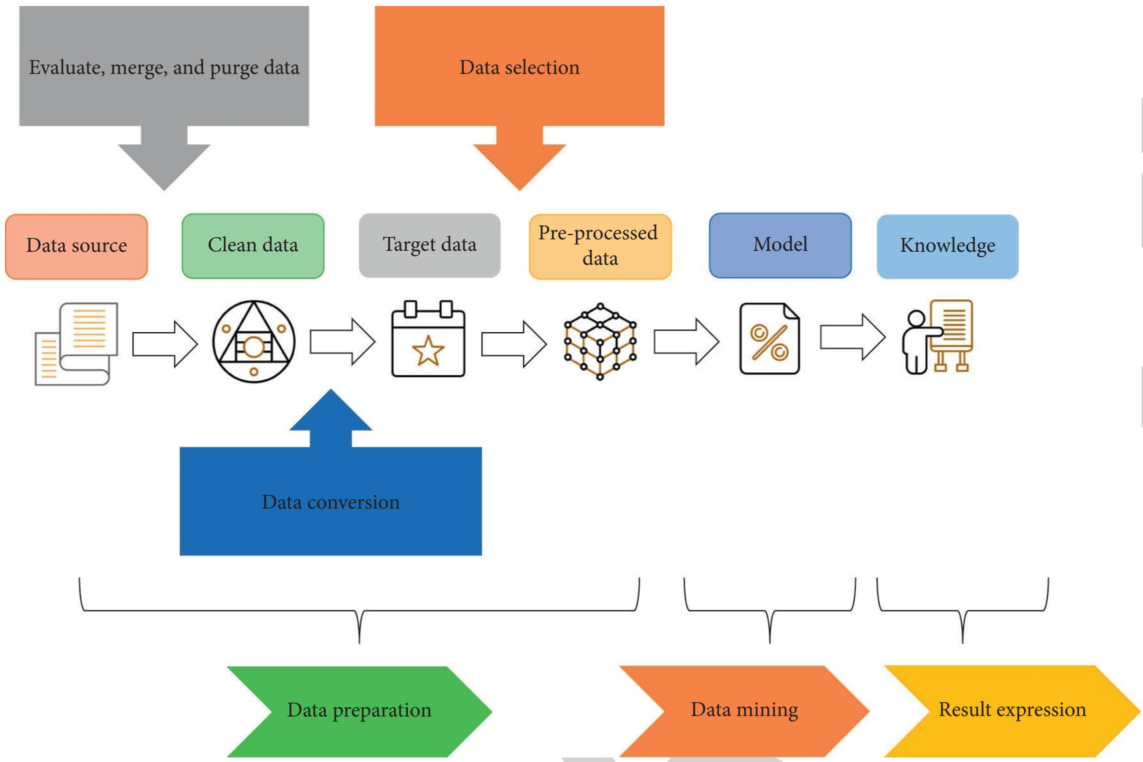


Figure 3: Process diagram of association rule mining.

magnitudes at the same time, they need to be normalized. The value of i factor j in all samples is

$$Z_{i,j} = \frac{X_{i,j} - \mu_j}{\sigma_j}. \tag{7}$$

μ_j is the mean value of i th factor and σ_j is the standard deviation of j th factor.

Second, when the user needs to access the full-text data in the resource, the system first takes out the full-text description information from the database and then points to the file system to read the corresponding full-text data. Then, the new mean value of each cluster is calculated. This process is repeated until the criterion function converges. The squared error criterion is used, which is defined as follows:

$$E = \sum_{i=1}^{k} \sum_{p \in C_i} \left| p - m_i \right|^2. \tag{8}$$

E is the sum of squared errors of all objects in the dataset, p is the point in space, i is the given object, and m_i is the mean value of cluster C_i.

For teachers and students in the network, real-time exchange of views is very necessary; it is an important form of teaching interaction, and in the electronic chat, subject teachers can answer students' questions in a timely manner. At present, the teaching resource management in schools is scattered in various departments and the basic management information is scattered in different management systems, and the integrated management of teaching resources is actually the integration of the management systems of these teaching resources. Since fuzzy association rule mining is still the mining of the laws between items in the database, but at the same time, the numerical information corresponding to these items is considered, fuzzy association rule mining is essentially an extension of association rule mining. Each item set needs to carry a decision attribute d, and when performing a join to generate a candidate $k+1$ item set, two frequent k item sets that can perform the join operation must satisfy

$$\left(l_a, l_h\right) \in \left\{\left(l_1, l_2\right) \middle| \operatorname{diff}\left(l_1, l_2\right) = 1, d_1 = d_2\right\}. \tag{9}$$

Finally, users can enter any of the English or Chinese search terms in the search box and the system will automatically match the title, author, subject term, abstract, series number, sponsor, and identifier records in the resource and display the search results that meet the requirements. In addition, resources related to the content of the course under study include electronic books, literature, etc. These resources can enable students to supplement their subject knowledge, expand their horizons, and increase their knowledge. If the resource base of multimedia network teaching can develop and integrate teaching courseware and software modules under these learning modes, combine them with the characteristics of language teaching, starting from the heuristic teaching and the dynamic teaching process, and realize the optimization and sharing of network media information resources in depth and breadth, it can realize the full utilization of these resources. Therefore, it is necessary to realize integrated management, and to allocate, utilize, and plan further construction of teaching resources from a global perspective. As for how to evaluate the validity of the output fuzzy rules, it is necessary to define a fuzzy rule truth degree.

4. Application Analysis of Association Rule Algorithm in Multimedia Network-Assisted Integration of Educational Resources

4.1. Analysis of Integration of Educational Resources System Using Apriori Algorithm. The Apriori algorithm is a width-first algorithm that finds all frequent item sets by multitrip scanning of a database. We will discuss the problem of association rule mining with constraints using the Apriori algorithm as a prototype, and the following briefly describes the analysis of the integration of educational resources system using the Apriori algorithm. The scalability of the algorithm in this chapter is compared under different database sizes using two types of fuzzy rules, namely, $t = 0.1, t = 0.2$. The comparison results are shown in Figure 4.

First, the candidate frequent set is pruned, and the strategy of bottom-up or top-down search of the database is used several times. Through computers and the Internet, students can choose their own content and mode of learning based on their needs. Interactive computers and web-based learning resources also provide timely feedback to students. During the course of multimedia-assisted English teaching, teachers can obtain a great deal of knowledge points related to the text to be taught through network information resources and, through certain screening, put the most important and meaningful contents into the courseware to guide the text and lay the foundation for the teaching of the text afterwards. Then, we choose the model with neither constant nor temporal terms, and the time spent by Apriori algorithm with different support degrees when the support degrees are 30, 60, and 90, respectively, is shown in Figure 5.

The research of data mining methods in the network environment can, on the one hand, study distributed data mining algorithms with the help of the network to improve the mining efficiency, and on the other hand, a data mining server can be established on the network to cooperate with the database server to realize data mining. It implements a set of publicly accessible interfaces for Web Services to register their service information with UDDI's web service information repository. The acceleration ratio is the ratio of the time consumed by the same task running in a uniprocessor system and a parallel processor system. Table 1 shows the Apriori test of 20,000 records on a cluster with different number of nodes, with the same amount of data and the same degree of support in the test.

Secondly, the database is projected onto a frequent pattern tree, FP-tree, to keep the association information of item sets, and then the frequent item sets are searched by scanning the count of path nodes in the FP-tree. The wide application of IT in the integration of educational resources system provides a good technical support to integrate the scattered teaching resource management information from different management systems. For those knowledge points that are difficult for teachers to express in words, or abstract concepts, students can get good learning in multimedia classroom in various forms. Some data or applications in a particular domain may require specific algorithms to find

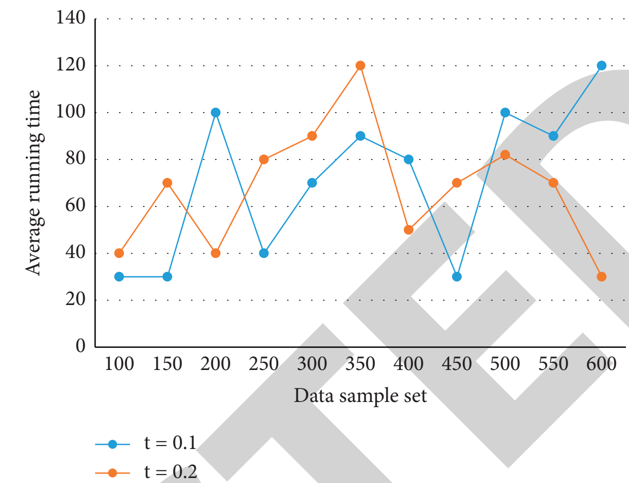


Figure 4: Scalability of Apriori algorithm.

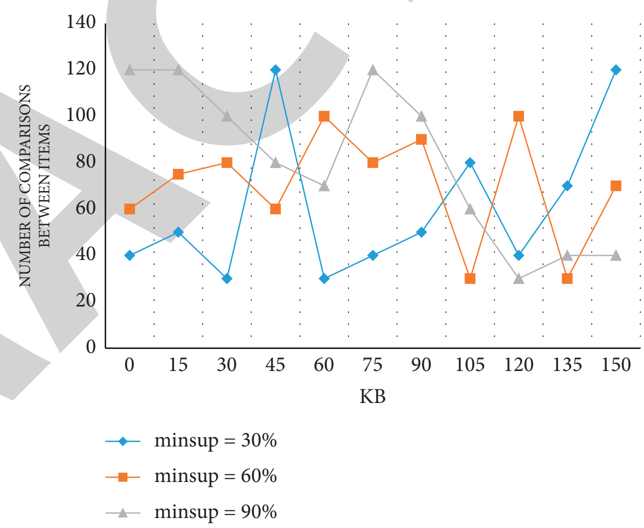


Figure 5: Comparison of time spent under different support levels.

Table 1: Tests under different nodes.

Number of nodes	Elapsed time (s)	Speed-up ratio
2	1382	1.9038
4	3827	2.3988
6	7643	3.2815

patterns, and general-purpose data mining systems have inherent limitations for these domain-specific data and may not be able to meet the requirements. In the performance comparison experiments between DL and association rule algorithms, we use a minimum support of 15%. The database size for the experiments ranges from 0 to 120 tuples, and the sequence of items climbs. As the number of tuples increases, the performance comparison results of DL and association rule algorithms are shown in Figure 6.

Finally, each time the support of a subset is calculated, the transaction records in the transaction database are traversed from top to bottom, and all the records are scanned and compared. Therefore, it is easy to use Analysisi Services

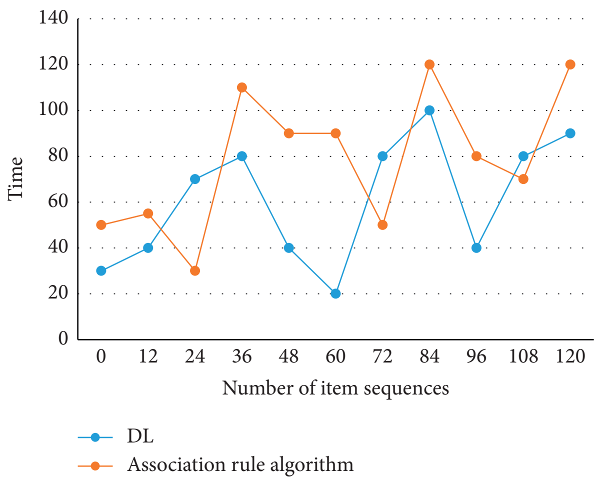


Figure 6: Performance comparison experiment between DL and association rule algorithm.

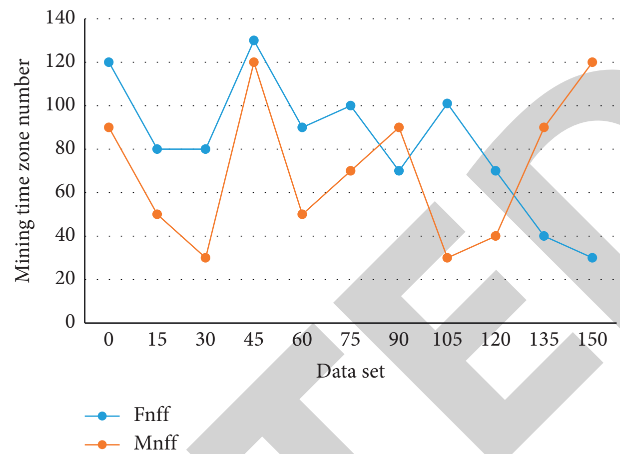


Figure 7: Experiment of filtration efficiency feff and merging efficiency meff.

to build the dimensional and multidimensional record sets of the data warehouse designed in the previous section, by following the various wizards in Analysisi Services step by step, integrating business logic and data mining systems in specific areas, and combining data analysis techniques with domain-specific knowledge to accomplish specific tasks. The multimedia system is used to select the appropriate question and answer format, find the right entry point for each student, and more effectively help students consolidate their knowledge in the classroom. It also enlivens the dull classroom atmosphere to a certain extent and allows students to learn new content more efficiently. At the same time, the integration of educational resources system itself is a collection of many information systems and information resources, with powerful data processing functions, which can well handle different standards and formats of data between different teaching resource management systems.

4.2. Solution Space Analysis of Optimization Association Rules. From subject knowledge to reading materials, from learning new knowledge to practice quizzes, from group solutions to real-time answers, from sharing resources between teachers and students to links with other platforms, from teachers' classification and management of knowledge to the provision of search engines, from students' learning to teacher-student interaction, the platform provides a three-dimensional, student-serving support environment. Due to completeness, the result of association rule mining is often a huge number of rule sets. These redundant rules can be summarized into two parts: one part of the rules lacks statistical relevance, while the other part of the rules does not meet the "novelty" requirement. Thus, optimizing the solution space of association rules is particularly important. In order to verify the filtering and merging efficiency of optimizing the solution space of association rules, we conducted an experiment on a transaction database of about 20 MB containing temporal interval attributes, and the results are shown in Figure 7.

First, the given set of data is mined using the Apriori algorithm and the domain knowledge (given in the form of template rules) provided by the user. It can be used in conjunction with many existing Internet communication protocols, including HTTP, SMTP, and FTP, and also provides a simple and more convenient mechanism for exchanging structured and unstructured information flows between endpoints in an interconnected network. It also supports a large number of applications ranging from messaging systems to remote procedure calls (RPC). Instructional resource integration systems must be autonomous; that is, they need to have automation technologies embedded to alleviate or eliminate manual deployment and management tasks yet allow the platform to respond intelligently to the requirements of the application. Based on the integration, teaching resource managers can build an integrated model of teaching resource list and process list for teaching and research programs, unify static resource management information and dynamic teaching and research processes, and configure existing teaching resources with the actual needs of teaching and research.

Secondly, for the rules that differ from the latter items of the template rules, we use the principle of information theory to sort them in order to identify those unexpected rules. The basic situation that teaching resource management information and teaching research information are independent of each other, and that teaching resource management information flows are independent of each other among different departments, still exists after the integration of the management system. To illustrate the memory occupation of the solution space of association rules, we track the size of TISS and TISS* for different transaction database sizes, and the experimental results are shown in Figure 8.

The main goal of Web Services is to build a common platform-independent and language-independent technology layer on top of various existing heterogeneous platforms, and applications on top of different platforms rely on this technology layer to implement mutual connection and integration. Users can submit requests for

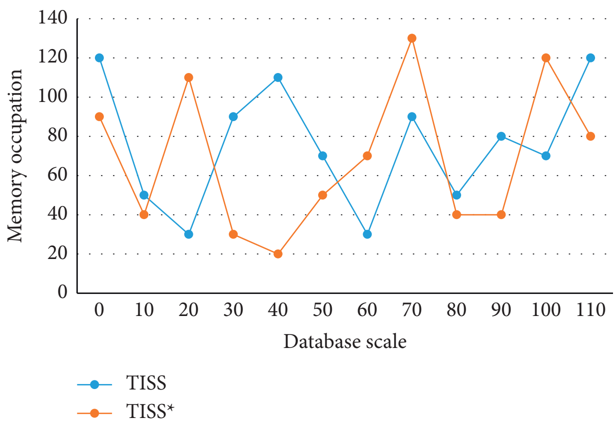


Figure 8: Experimental results of memory occupancy.

Table 2: Test data and time.

Number of records	MapReduce test time (s)	Apriori test time (s)
120	94.28	54.27
240	124.27	73.87
360	172.36	93.29
Average test time	130.30	73.81

resources and applications at any time, and then the cloud environment manager is responsible for allocating resources and deploying services. For the Apriori algorithm, we verify the efficiency of the algorithm by means of node comparison. The experimental results are shown in Table 2, which shows the amount of data used for testing and the time used.

As the amount of data increases, the test time gap between MapReduce and Apriori gradually increases. The average test time of Apriori is 56.49 s less than that of MapReduce, so the advantage of Apriori algorithm is more obvious.

Finally, the rule that is optimal for a given posterior term is selected, and this part of the operation leads to a local update of the rule template, and if the mining result is unsatisfactory, the adjusted rule template can be used to remine the dataset. The use of Web Services technology to build a data mining platform enables the interconnection and interoperability of information between heterogeneous information systems, making it easier to achieve data mining on distributed and heterogeneous databases. Parallel deployment refers to changing the traditional sequential deployment method to parallel execution, performing multiple deployment tasks simultaneously, and deploying virtual machines to multiple physical machines at the same time. Through the communication environment and teaching process recording function of the integration of educational resources system, we can get timely and comprehensive feedback information about the utilization status of teaching resources. In turn, we can make appropriate adjustments to the integration and construction of teaching resources, so as to make better use of teaching resources.

5. Conclusions

With the deepening of education informatization and the increase of integration of curriculum teaching resources, the demand for curriculum integration of educational resources platform is bound to increase. The integration of IT and English teaching aims to fully mobilize students' initiative and enthusiasm, so that students' creative thinking and practical ability can be effectively exercised in the process of teaching integration, thus creating an ideal English learning environment for international students. Through multimedia technology, teachers can create many realistic English learning scenarios, so that students can devote themselves to them, adapt to the English environment, and learn English better. The computer network integration of educational resources system provides richer functions to meet the needs of various teaching resources submission, saving, management, and publishing applications. Since data mining software can discover useful information hidden in a large amount of data, it can provide powerful support to specific fields; particularly, the complex process of data mining is invisible to users in the application field. Therefore, in this paper, we design and implement a computer network integration of educational resources system by linking association rule mining in data mining with multimedia network English integration of educational resources and analyze the association between Apriori algorithm and integration of educational resources system. By using this paper's association rule algorithm-based multimedia network-assisted English integration of educational resources method, the existing resources can be integrated to realize the full sharing of high-quality resources within and among schools, and the advantages are complementary.

Data Availability

The labeled datasets used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by Guangxi Department of Education, the key project of Guangxi Vocational Education Teaching Reform Research Project of 2018 (no.GXGZJG2018A026).

References

[1] R. Liu, "Design of ideological and political multimedia network teaching resources integration system based on wireless network," *Scientific Programming*, vol. 2021, no. 3, pp. 1–15, Article ID 4293771, 2021.

[2] Y. Du and T. Zhao, "Network teaching technology based on big data mining and information fusion," *Security and*

Hindawi
Mobile Information Systems
Volume 2022, Article ID 5115184, 12 pages
https://doi.org/10.1155/2022/5115184

Research Article

Multisource Analysis of Big Data Technology: Accessing Data Sources for Teacher Management of Sports Training Institutions

Yiting Wang and Le Yu

Department of Sports, Sichuan International Studies University, Chongqing 400030, China

Correspondence should be addressed to Le Yu; 99002766@sisu.edu.cn

Received 26 May 2022; Revised 21 June 2022; Accepted 25 June 2022; Published 13 August 2022

Academic Editor: Jiafu Su

In the information age, "mobile Internet," "cloud computing," "Internet of Things," and "data mining" concepts are emerging at the same time, as well as other fields of related data-based applications. The mobile application will be born as a result. Therefore, in the information age, big data, which involves information in a specific key or specialized field, has gradually begun to receive a lot of attention in recent years. In 2011, the US consulting firm McKinsey and Company first proposed the arrival of the "era of big data" and in August 2015 in China's State Council issued a notice of action outline "to promote the development of big data." Meanwhile, big data has gradually become an important factor in driving national reform and innovation, promoting scientific and technological progress, improving the way society is managed, and guiding changes in education and research. Big data is driving a very influential shift in thinking in an era where big data is changing the way we live, becoming the way we understand the world, and gradually becoming the source of new inventions and services. At the same time, the rapid development of big data technology for physical education teachers needs big data for management and training and other institutional managers to provide more effective ways and means of education management, but up to now, the status of big data for management is still another serious challenge, sports and training and other institutions of big data and processing process of data nonintelligent, nonclosed-loop processing, data nonlinked processing, etc. Many problems are also still very obvious. According to the new characteristics of sports big data refinement management, the current situation of sports professional training institutions teacher management, combined with sports training institutions to find some more practical sports training institutions teachers big data management methods can effectively improve the efficiency of management, teacher team building, strengthen sports training institutions to improve the quality of teaching teachers, and promote the overall quality of students have a positive impact. In this paper, we combine the characteristics of "big data" and the construction of teachers in sports training institutions, and put forward some suggestions on how to improve the level of teachers in sports training institutions in the era of big data and conclude that the construction of teachers in sports training institutions should seize the key era now and enter the "|big data era." We conclude that the construction of teachers in sports training institutions should seize the critical era and enter the "big data era," so as to rely on science and technology to improve the construction system of teachers in sports training institutions.

1. Research Background

Today's world is an information world in which the scale of information resources is rapidly expanding and changing, big data technology can bring about a wide range of social concerns, and social influence will increase. The international market and the whole domestic market are always maintaining the continuous high growth of big data market scale, which can also explain the future development of big data industry investment will also continue to grow at a high speed, the benefits will also get a continuous high-speed rapid increase, and the future development of enterprises in big data-related industries will also be very good prospects. The arrival of the "big data era" has completely changed the work habits of modern people, and the company's new business model and industry division of labor. In short, it will make people start to rethink and redefine everything. Many Internet companies have begun to realize that big data is a real asset that can really play a more important role in driving their business forward [1]. The large amount of data collected will be classified, stored, analyzed, and displayed through various information terminals in order to become the basis for business decisions, customer behavior, predicting market trends for the company, achieving precision marketing, and improving work efficiency, and innovative business models will be improved and optimized across the board [2].

In recent years, with the rapid development of social productivity, people's living standards and quality of life continue to improve, and the importance of education is also increasing, thus making the market demand for education grow. According to the statistics of the institute, the market scale of the national education and training industry is increasing in 2018 [3]. Generally speaking, the education and training industry is mainly divided into public schools and private education and training institutions outside the system. Therefore, more and more businesses realize the huge potential of the education and training industry, and have begun to deploy in the market. According to statistics, by the end of 2018, there were as many as 81,100 private education and training institutions outside the system in China. The private education and training institutions outside the system mainly include children's training, IT training, civil service training, vocational training, etc. [4]. As an important part of China's education system, private education and training institutions outside the system have attracted particularly high attention and are playing a very important role in improving people's knowledge. The scale of education training market is getting bigger and bigger, and the market competition is getting more and more fierce. It can be seen that the big data wave and education big data applications are an essential force for social development. With the booming development of education and training in China, more and more education and training institutions, the competition is getting more and more brutal, and the innovation and development of big data will be promoted by in-depth research and exploration of education and training [5].

The idea of "human resource is the first resource to promote social and economic development" widely recognized, and the overall condition of teachers will directly affect the quality of the talents cultivated. The traditional teacher management model is no longer able to meet the development needs of today's training institutions. Therefore, there is an urgent need to apply the concept of multisource analysis of big data to teacher management and establish a modern teacher management model. The times have placed new demands on educators and managers. Educational administrators should respond to the times, investigate, compare, summarize, and conclude the current situation and development of teacher management in sports training institutions at home and abroad, explore suitable teacher management models, and contribute to the benign development of teacher management in sports institutions [6].

1.1. Big Data. In recent years, the rapid development and popularity of big data seems to have become a common phenomenon. From national strategies to corporate strategies, big data seems to exist as if no one knows about it [7]. There is no uniformity in the concept of big data. The famous scientific research institute Gartner defines big data as follows: "big data" is a huge, high growth, and differentiated digital information asset that requires innovative processing for scientific decision-making, insight, and process optimization [8]. Amazon (AWS) data scientist John Rauser mentioned a simple definition: big data goes beyond the ability of computers to process large amounts of data. R&D team Big Data: "Big data is the most valuable advocacy technology and the most fashionable and cutting-edge technology. In these particular cases above, the definition is very confusing." Big data analytics is one of the more popular online vocabulary technologies currently being used in the IT industry in China. According to McKinsey, one of the management consulting and management consulting firms, "big data analytics," has gradually begun to enter into the application and service functions of today's society and many other management industries, and has begun to become an information-producing resource with great potential. Through the in-depth theoretical research and comprehensive analysis and application of these large amounts of consumer data, a new wave of consumer data with new characteristics reflecting the development of social productivity and meeting the surplus needs of consumers has emerged [9]. More and more advanced and mature mobile Internet companies and technology research and development companies have begun to slowly realize that the ability to analyze and process big data is slowly becoming the most critical strategic value of the core assets of Internet companies, and the ability to analyze and integrate data mining will become an increasingly important core competency of enterprise organizations in the future.

Compared with traditional structured data, big data is highly abstract and has not yet been defined in academic circles. The McKinsey Global Institute (MGI) defines big data as a set of data used to capture, store, manage, and analyze larger datasets than traditional database software tools [10]. The National Science Foundation (NSF) defines large-scale data in terms of data sources and characteristics. It considers big data as a large, diverse, long-term, and complex distributed dataset generated by a variety of data sources (e.g., scientific instruments, Internet transactions, sensor devices, audio and video software, etc.). Gartner, a leading IT consulting group, considers computer-processable computing power as a core element in defining big data capabilities and defines big data as an extreme information management and processing problem. One or more dimensions are beyond the traditional information technology processing capabilities. These different organizations have different perspectives and representations on the definition of big data, but the views on data volume, data source types, data transfer methods, data processing speed, and low data value density are essentially the same [11]. Based on the characteristics of large-scale data storage, transmission, and computation, the authors believe that the processing capability of big data mainly refers to the large and complex data structure system and various complex types of data collections built based on the cloud computing model in big data processing and applications. Through the integration of systems, data sharing and a variety of data processing methods, such as cross-filtering useful data assist in the development of strategic solutions and assist in decision-making.

Big data has typical "four" characteristics: diversification, mass, value, and speed [12].

(1) The data storage and calculation of massive data reflect the large scale of big data. At present, the scale of big data has evolved from GB, TB, PB, EB to ZB. According to the official documents released by technology companies, the data era 2025 released by the International Data Corporation (IDC), the global data circle will expand to 163 ZB in 2025, equivalent to 16.1 ZB data obtained in 2016 increased by 10 times; the total amount of global data will grow 50 times to 5.2 ZB.

(2) Diversity refers to various types of data sources, including not only structured data, but also unstructured data such as text, web pages, audio and video, comment statements, and semantic analysis. With the rapid development of mobile Internet, e-commerce, and social networks, after actual research, unstructured data show an explosive growth trend[13].

(3) Fast means that the generation speed, collection speed, transmission speed, processing speed, and analysis speed of big data are faster than traditional data processing methods, and the analysis dimension is wider.

(4) This also indicates that the use of big data development and application research will have its very high potential socioeconomic value and applied scientific research and development value, and more and more excellent Chinese enterprises have begun to gradually realize the meaning behind "data." Data asset management has gradually formed a huge potential core strategic resource advantage and enterprise core competitiveness for the new development of enterprises, fully utilizing and effectively using enterprise big data resources to contribute to the healthy and sustainable rapid development of enterprises in various business areas.

The concept of big data represents not only its rich data content, but also big data technology. Among them, various types of training institutions are the concentration of integrated management of big data [14]. There are not only various types of organizations, but also more responsibilities and processes [15]. Therefore, in this complex environment, there is such a wide range of information data source channels between different departments, various information channels of management organization, and information flow of interlocking departments, all of which undoubtedly provide a challenge to the way of information management organization of various education and training management organizations [16]. For all kinds of education and training service institutions in the country to carry out the management of training teachers for big data management, can be more effective use of big data and powerful professional network technology support, revitalize training data resources, improve the data flow between the training service institutions teachers management data operability, enhance the accuracy of training teachers information flow management data. This is also a new challenge for education and training institutions in the new era.

Technical analysis is a kind of explanation of the development tools, data management systems, and advanced development technologies that need to be mastered and operated to develop the information management of the training institution's faculty, and the analysis of these advanced technologies provides theoretical support for the smooth completion of the system's research. The main technologies used to develop the professional management platform are ASP.NET, AJAX, SQL Server, etc.

Data source is the database or database application using database server. Data source is the premise of all analysis to mine the data source, and access and maintain continuously updatable data through real time and nonreal time. Provide a comprehensive, friendly interface to configure data source information, through the task scheduling data collection, support custom-sensitive information encryption processing, support a comprehensive range of data sources, such as MySQL, Oracle, and SQL Server.

SQL (Structured Query Language) is a structured database query language, which can be used to build the interconnection between various database system statements for better and more effective data communication, and the standard statement pattern of various database system statements can be unified to achieve data sharing [17]. The standard statement pattern of SQL can also be used in the process of querying various data types, including updating various data information queries and data extraction related to various data types stored in various database system statements. Microsoft SQL Server is also a server model that can be used in a language environment that supports certain languages, and is a database system designed with a server architecture (C/S). It is a database system designed with a server architecture (C/S). Users are also able to achieve simple and perfect search and local operation in this model environment [18]. There is also a very important part of the database management system that can improve the database process, and can further ensure the high security of data storage and maintain the high integrity of stored data information, and it also hopes to further improve the real effective data operation mechanism in the database to achieve a more efficient way of data analysis and processing. This means that for the database data management itself, a very critical management function is how to achieve the data management and eventually become a major tool in database management [19]. Whether you are an organization, enterprise whether internal unified enterprise e-commerce system management or external to the enterprise organization of office management and enterprise administrative management of all the applications of the database system for the overall unified management is for a business when necessary, the need to establish a safe, efficient and reliable unified and long-term stable operation can be used to permanently save all the enterprise system data The drill-down database management system and database management software [20]. And all the data information can be saved as a direct file form in a computer database system needs to use such a computer database system with special functions for large and complex computer system services to help complete the daily management and maintenance of the unified database and system planning. At present, there are relatively more database software applied in the Chinese market, and the MySQL Server database of Microsoft and Oracle database of Oracle are more frequently used in China. These two major databases have their own advantages and disadvantages, for the use of Microsoft's application development tools to develop software; of course, the use of SQL Server database is the first. Compared with the Oracle database, its installation and operation are relatively easy, and the Microsoft exit database can improve the performance and functionality of the system.

Web 2.0 website technology is one of its greatest advantages to allow users to participate directly in the creation of website content [21]. It can also be said that it has a large degree of technical difference with the general website information release processing method; first of all, some of the main content information in the website is directly by the website users to participate in the creation of content information release processing; to some extent, it has increased a lot of website user interactivity and increased a great degree of website user interactivity, so that the website users at the same time also. In this way, website users are both the main viewers and the main producers of web content information, which shows that the Web 2.0 website will bring more new opportunities for website users to participate in the creation of website content information; for example, we now see some social networks and For example, we now see some social networks and Sina Weibo that is a typical user participation in the creation of information content information is a guiding idea, through such a new Internet technology can be these user-created content information to re-categorize [22]. Moreover, Web 2.0 sites pay attention to interactivity, so that users can not only use the site's server, but also interact with our prewritten programs, and it is possible to communicate with other different users. In this way, people will be able to give birth to many different types of network-related applications at the same time, which will greatly facilitate the convenience of network applications and increase the mutual adhesion between network users, which is not simple on the one hand, but also conducive to the innovation of the network, and more importantly, it provides a further development of the whole network-related industry in the future. Web 2.0 has a standardized approach to the development of the Web [23]. Web 2.0 has a standardized web design, standard is a very important web specification, which is also a very important technical key, he will help design and develop a better compatibility between different browser languages, to improve the user experience is very critical, the general so-called Web standards are usually referred to the construction of the web page of the language based on the language, in fact, In fact, Web standards have not actually become a standard for anything. It is inaccurate to say that standards are actually a collection of standards that allow us to access the web across browsers and platforms because of the existence of these standards [24]. Its advantages are obvious, and it makes the website design code specification, so it greatly reduces the workload of the code, but also greatly reduces the waste of network bandwidth and improves the speed of network access; in fact, the final benefit of the majority of users, the user experience will be greatly enhanced because of these changes. The most important point is that a web standard compliant website is more user- and search engine-friendly.

The ASP.NET language has undergone revolutionary changes and is a compiled programming language, which is a great improvement and breakthrough for the ASP programming language, which on the surface seems to be an upgraded version of, but in fact, there is nothing directly related to the two. There is a large degree of difference between the basic working principle of the internal system and the various programming languages used and their internal operation mechanism [25]. The ASP.NET language is used to implement programming features, and the Web application environment used for development does not refer to a language program built entirely from a built-in common language programming framework that can be applied to the Web application development server and can implement and create a more comprehensive architecture-based application. The development platform provides the advantage of a more powerful model of application in a web programming environment. ASP.NET applications are based on Common Operating Language (CRL) programs built to run on a Web application server [26]. NET no longer performs the traditional function of just explaining the program, but compiles the current running application on the service when it is first requested. For the development of the program for the compiled program compared to the explanatory program, its efficiency has been greatly improved and increased; the current better object-oriented programming tools such as ASP.NET and JSP are compiled.

1.2. Faculty Management. Faculty management is teacher management. The content of teacher management has been clearly pointed out in the "Teachers Law" promulgated and implemented in 1995, mainly including: (1) teachers' rights and obligations; (2) teachers' qualification and appointment; (3) teachers' training and improvement; (4) teachers' treatment and reward; and (5) teachers' assessment and evaluation [27].

Education and education are the most important of many school tasks. Teachers are the main workers in accomplishing educational tasks and are the actual managers and operators of the educational process. The manager of the training institution manages the school and the management, rationalizes the resources of teachers, and effectively implements the management functions of teachers, professional support facilities, united and coordinated, well-structured, appropriate number of professional teams of teachers with high personal and overall quality, and strong level of scientific competence, which is inevitably a necessary condition and an important guarantee for the educational prosperity and development of the training institution [28]. As one of the main forces of the training institution, teachers are both educators and managers. Therefore, the importance of teacher management is that teachers effectively regulate the management of the institution. Two quantities need to be mastered; namely, teachers cannot be completely passive, blind, and simply accept management. Instead, they need to be proactive in the management of the institution and participate proactively in the management of the entire training institution. In classroom management, teachers are managers as opposed to students, and teachers need to accept rational suggestions from students to manage the curriculum. Therefore, the openness of teacher management is positive, which allows for the validity of teachers' opinions and suggestions, and the practical management of teacher management using advanced concepts in human resource management. It can contribute to the long-term development of the institution.

Teacher resource management in sports training institutions refers to the process of planning, organizing, coordinating, and controlling the recruitment, deployment, use, evaluation and development of teaching staff in training institutions, and ultimately building a team of teachers with good quality, excellent quality, and unity, so as to maximize the overall work goals of training institutions [29]. Foreign countries generally consider education as a supplier of human resources, and the human resource management of service education includes issues such as attracting, training, and retaining effective teachers. Scientific, rational, and effective human resource management of training institutions can promote the faster and better development of training institution education.

2. Problem Solving

The teacher management system of sports training institutions is to solve the problems of teacher training, teacher ethics, and teacher assessment in training institutions, and to improve the information management ability of teachers. Before developing the system, we first communicate with the teachers' affairs department and personnel department of training institutions to familiarize ourselves with the policies of teachers' training institutions, determine the basic business of teachers' information management, and determine that the information management of teachers' training institutions includes the information management of teachers' files, teachers' training management, teachers' recruitment, teachers' ethics management, and teachers' training management. Information management, teacher moral and ethical construction management, performance assessment management of teachers and staff, comprehensive information query management of teacher construction, and user rights and role management are the seven core businesses [30].

2.1. System Management Business Analysis. Teacher information management system is a module for system setting of basic parameters in the system and initialization data for system operation. Teachers and employees can manage teachers' assessment information only if the categories of assessment parameters and related parameters are set first. This business module allows only administrators to log in and operate, add teacher assessment types, and then add their corresponding assessment parameters and configure index scores for the relevant types. The specific faculty system management business activities are shown in Figure 1.

2.2. Analysis of Faculty and Staff File Information Management. Faculty and staff file management is the user object in the information management system of faculty construction. All faculty and staff data are firstly obtained uniformly through the external personnel management system to ensure that the information of faculty and staff is unified with the faculty and staff of training institutions and other systems. Therefore, in order to realize the data connection with external systems, we need to set up an external interface in the personnel management system to obtain the information of faculty and staff. By accessing the interface, the faculty construction management system can get the faculty data and determine the information of the training institution's faculty members' seniority and titles based on the information of the faculty members in the personnel system, laying the foundation for the establishment of the training institution's faculty file and the assignment of user rights. In addition to importing data from external systems, the administrator can also add new staff information that does not exist. The specific faculty and staff file information management business activities are shown in Figure 2.

2.3. Business Analysis of Faculty Recruitment Management. This module enables the release of part-time faculty requirements and collects all external applicants for faculty positions to fill in the information related to annual faculty recruitment through the faculty construction management platform within the specified recruitment resume reporting time period, including the filling in of job applications and the uploading of original scanned copies of supporting materials. Once the information is approved by the registrar, no candidate can modify or delete the information. External staff can first download the unified information reporting module and fill in the information according to different types and then upload it in batches or add new application information one by one. Teacher application management module is mainly responsible for the unified management of the application information and resume of each teacher applicant, which is an auxiliary module in the information management system for teacher construction and provides the basis for training institutions to establish a resource base of external teachers. Only the administrator can review and manage the recruitment information, as shown in the diagram of the specific teacher recruitment information management business activities (Figure 3).

2.4. Business Analysis of Teacher Training Management. Teacher training is established by training institutions to improve teachers' business ability and provide more learning opportunities for young teachers. Firstly, each

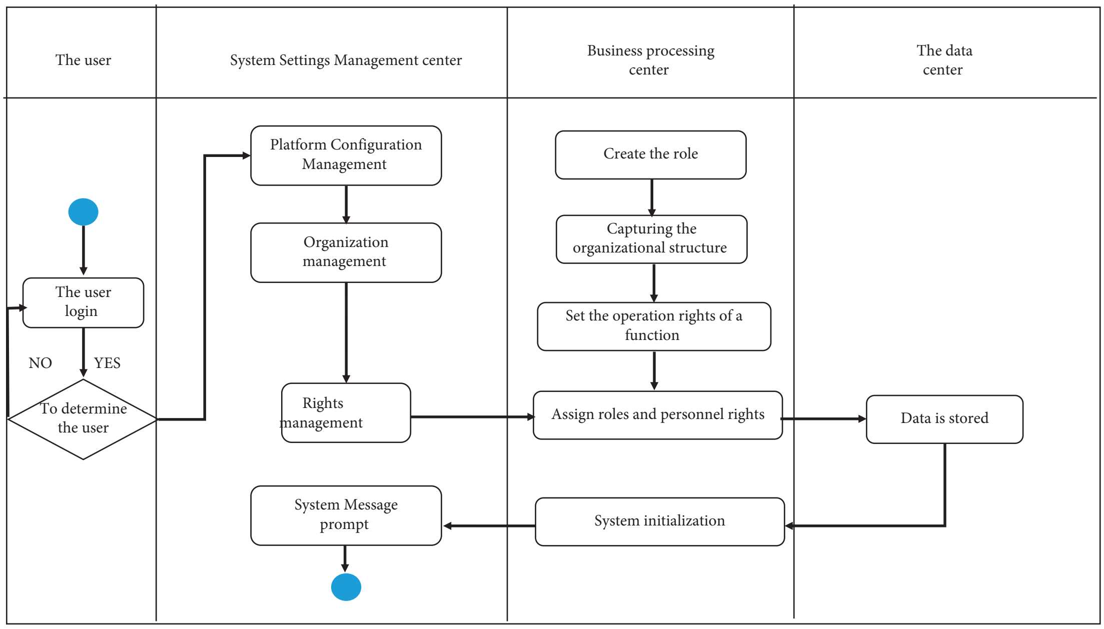


Figure 1: System management business activity diagram.

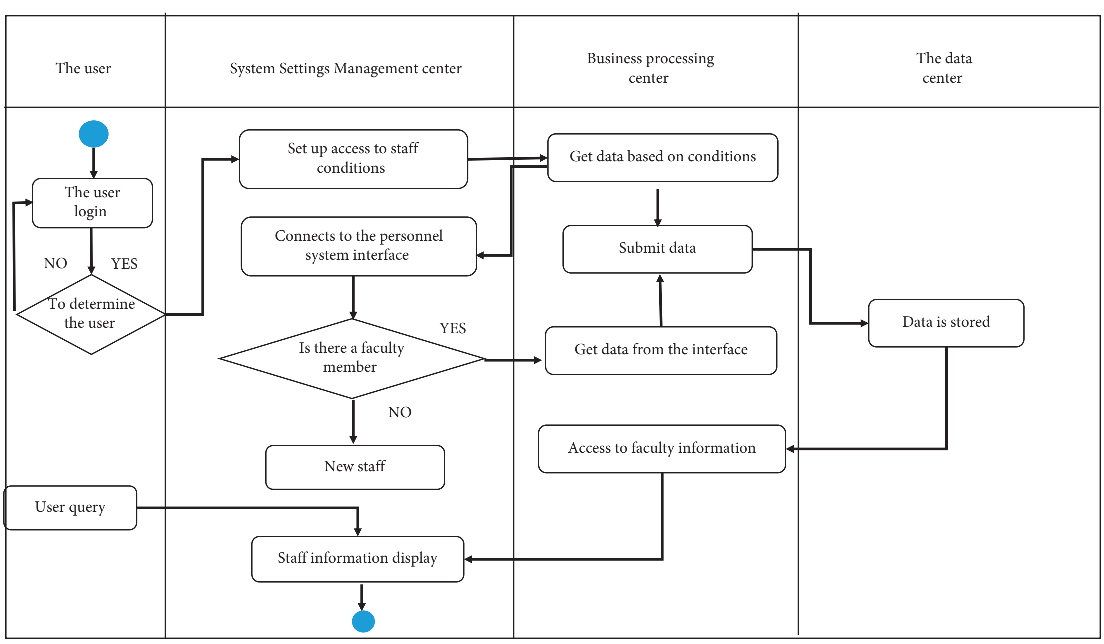


Figure 2: Business flowchart of faculty and staff file information management.

teaching department proposes annual training plan, etc., and then the Academic Affairs Office unifies to review and make specific annual training plan and release it, and teachers can apply for it according to the released training status. Only teachers' training information that has passed the audit can be aggregated. The administrator first sets up the conditions for obtaining training management process and then obtains the system teachers' training data according to the conditions and makes classification and aggregation. Then, the training analysis report is formed by comparing the teacher training plan made by the training institution with the actual training completed. The administrator can track the teachers' daily training through this module, and after the training is finished, the teachers need to fill in the training summary and upload it instantly through the network. Only teachers who pass the final review of the administrator can indicate that their training is really finished; otherwise, the teacher's training will be recorded as invalid, as shown in the specific teacher training management business activity diagram (Figure 4).

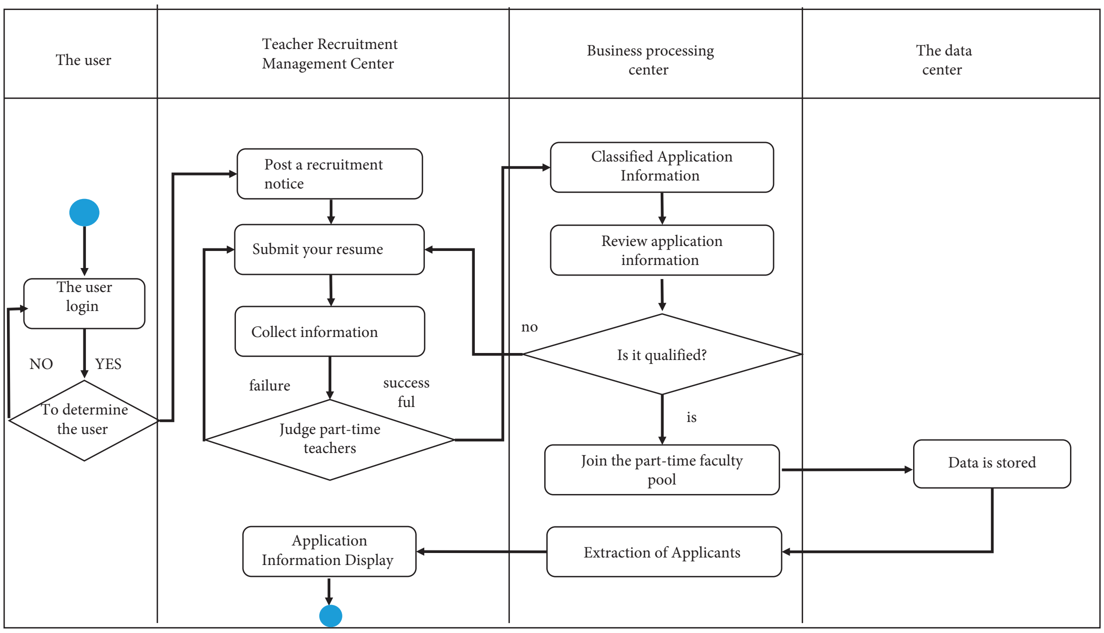


Figure 3: Teacher recruitment management business flowchart.

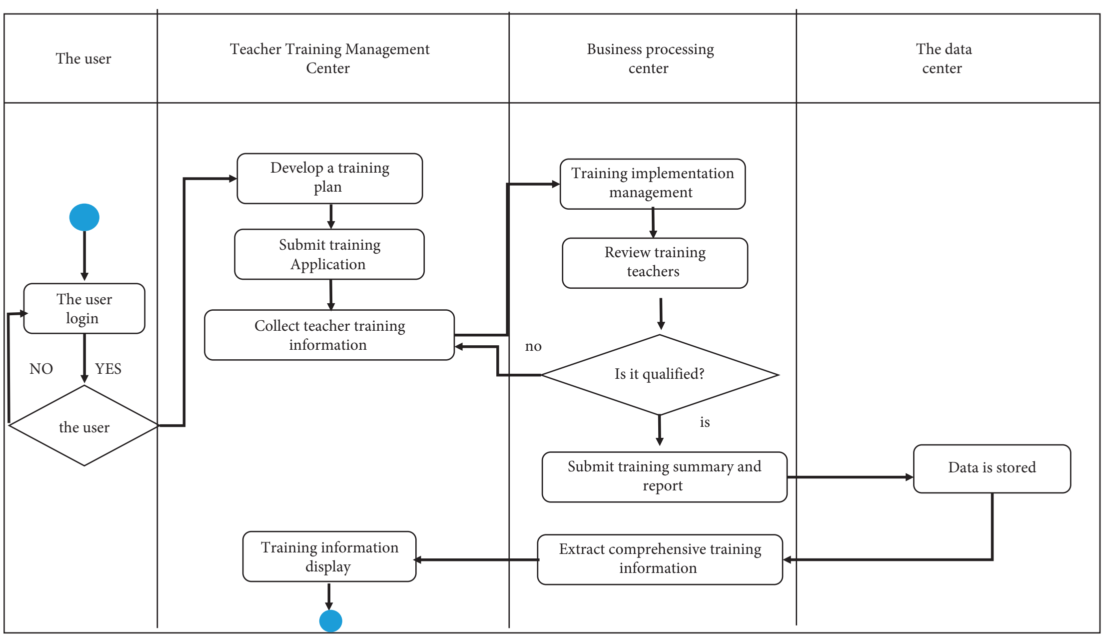


Figure 4: Teacher training management business activity diagram.

2.5. Business Analysis of Teacher Moral and Ethical Assessment Management. Teacher moral and moral assessment is an important assessment for all full-time teachers, which reflects their specific performance in teaching. Through this assessment, leaders of the Academic Affairs Office can track the teaching situation of full-time teachers in the teaching departments of training institutions. Assessment settings are mainly the detailed settings of forms, test indicators, scores, processes, and related parameters for various forms and types of tests, and the system performs assessment based on these settings. Assessment management is mainly to create, publish, and start relevant assessments by relevant assessment managers, and to view the assessment results and summarize and analyze the results at any time, which is the core function of assessment work. The evaluation of full-time teachers' moral character is mainly composed of students' evaluation, teachers' mutual evaluation, and branch leaders' assessment, which provides basic data for teachers' performance evaluation. All evaluation indexes and scores can be set arbitrarily, and the corresponding evaluation table and summary table are automatically generated after the evaluation. The data and results of the evaluation of teachers'

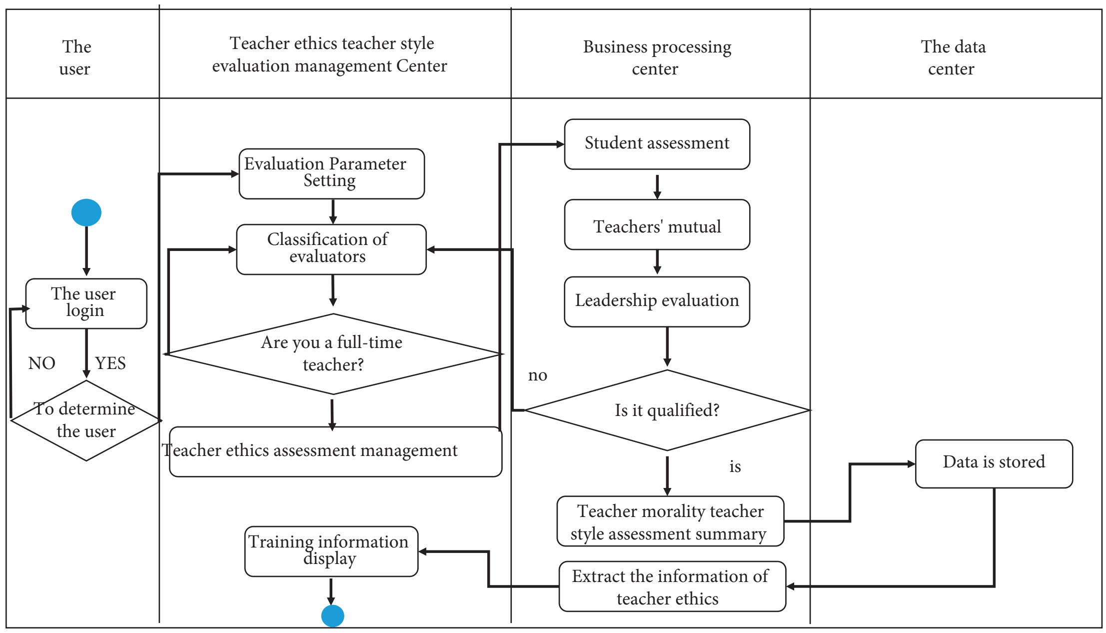


Figure 5: Diagram of teacher moral and ethical evaluation management activities.

moral and ethical conduct are statistically analyzed, and various charts are formed. The specific teacher ethics assessment management activity diagram is shown in Figure 5.

2.6. Business Analysis of Teacher Performance Evaluation Management. In order to establish teaching ability, practical ability, scientific research ability, and ability to serve enterprises and society as important contents of teachers' performance appraisal, the evaluation combined with multiple parties such as students' evaluation, teachers' mutual evaluation, supervisors' evaluation, and industry enterprises' evaluation is taken as an important basis of quality appraisal index, and the same performance appraisal and the same performance reward are implemented for full-time and part-time teachers. The appraisal settings are mainly the detailed settings of templates, indicators, scores, processes, and related parameters for various forms and categories of appraisals, and the system performs appraisals based on these settings. Appraisal management is done by relevant personnel to create and release appraisals in the system, view the appraisal results at any time, and summarize the appraisal results. The performance appraisal assessment is used for each appraisal subject to conduct self-assessment, the superior receives and participates in the appraisal of subordinates, gives feedback, and publishes the results of the appraisal, and the indicators and scores of the appraisal can be set arbitrarily in the system. This module is public and can be called a public module, which can be used by all people who log in to the information management system of faculty construction, but there are also permissions. Teaching staff can only view the information about their performance in each semester, and they can check their basic information through this module. The teaching managers of each teaching department of the training institution can check the summary ranking of the performance assessment information of the faculty (department) through this module, and they can export the queried performance assessment information, while the system administrator can check all the performance information of the teaching staff in school. This facilitates the management of faculty performance information and improves the quality of faculty information management in training institutions. The specific teacher performance appraisal activity diagram is shown in Figure 6.

Based on this, the architecture of the sports training institution's faculty construction management system is designed using a three-layer browser network database architecture, with the browser providing a browsing mode, the service providing requests and processing, and the database realizing the storage of system data. The user of the system can operate the system by providing a browser on the client side to add, modify, delete, and query the data in real time. The three basic layers of ASP.NET are the representation layer, the business logic layer, and the data layer. The representation layer contains the components that define the user interface consisting of forms, menus, and control panels, and provides an effective channel for users to interact with external system users directly and make calls to the external business layer. The service logic layer consists of the following two parts: the service logic layer and the access layer of service data. The business logic layer is to achieve control of the user interface layer, data interaction between the data and access object layer, and automatic acquisition of business data in the database layer, so as to realize the logical management of the service access object layer and the logical management of business data for the implementation of the service. The data access layer can be used to realize the operation of data retrieval, data addition, change, and clearing. The data layer stores the data in the system. The

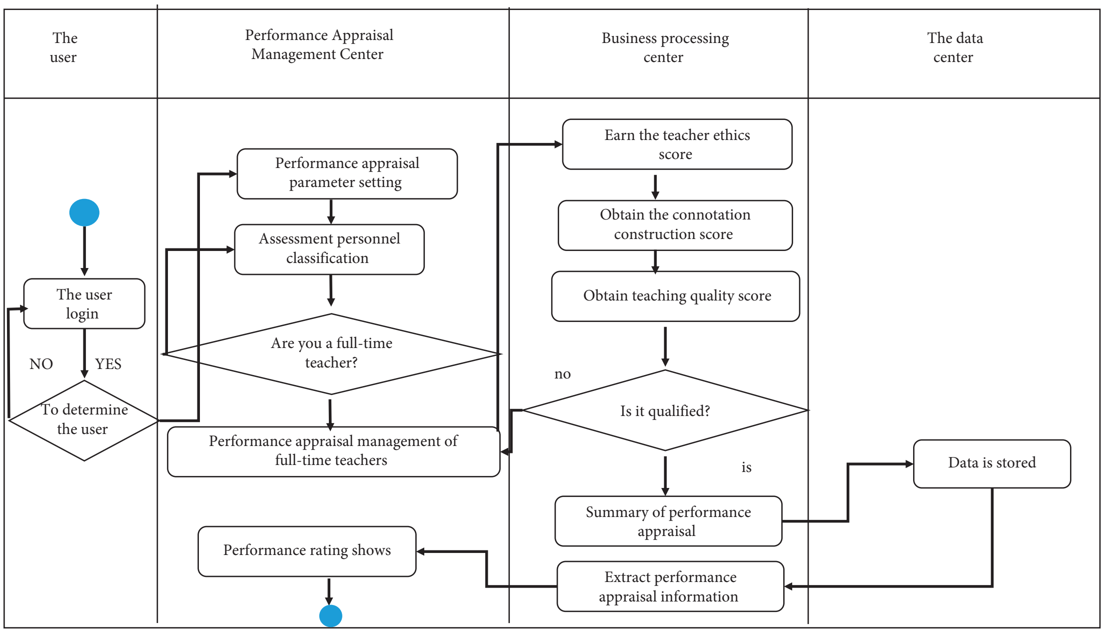


Figure 6: Teacher performance appraisal activity diagram.

three-tier framework system is a relatively mature, easy to use, and easy to popularize the use of the application framework, which can be divided into an application framework independent of at least three relatively separate packages. The design ideas of each layer and the specific role are as follows.

The main function of the representation layer in the application is responsible for the actual operation of the dialogue system between the users, the human-computer dialogue window directly interacts with the user interface, and the representation layer can be used for the system to provide the operation of the user to display data or input and output interfaces. Change the interface of the operating user, you only need to display control of system functions and data inspection, and you can rewrite the logic processing without affecting the other two. For the business logic layer, the system plays a role in the functional layer for the exchange of data, access to database data returned to the client, and the data submitted by the client back to the database, which plays a top-down function of the Jonquil, the main business achievements need to develop business rules, through this interface can send user input and output information data layer, the data layer request and response back to the business logic layer. Data layer requests and responses are returned to the business logic layer.

The data layer is mainly responsible for reading, writing, updating, deleting, and adding data to the database, and performs physical processing of data, which requires a large number of SQL statements.

The system adopts the current mainstream browser business layer database three-layer architecture for design, its browser is a browser mode, and the user side of the operation to submit data on the system and the system layer is mainly responsible for the final processing results. The Web layer, also called the interlayer, is responsible for converting the data format between the business processing and the implementation system. In order to further improve the stability and efficiency of the system, the system uses JSON format for operating system data. To complete the business logic of the class and database operations of business processing and data transfer will be the traditional development of data feedback to the client's data control, this mode of operation has a huge impact on system performance, and here the submission of data formatting, JSON format to the client generated data, the client can submit the JSON format rich client technology to achieve the receipt of data, and therefore further improve the performance of the information management system, thus achieving efficiency in processing service requests. The database layer is responsible for providing a data storage system for the NET architecture-based content management system using the database management system SQL Server 2008 to ensure data storage (Figure 7).

3. System Design

The implementation of the faculty building platform management system requires strict adherence to the software architecture design and functional module design of the previous system. The implementation of the system is based on the design concept of software engineering. Before the implementation of the software project, it is necessary to determine the planning of the network system, to build the network of each functional module of the system using the program code, and then to test the performance and functions of the system network management system. Finally, the programming language for the functions implemented in the system is given. In the development of the

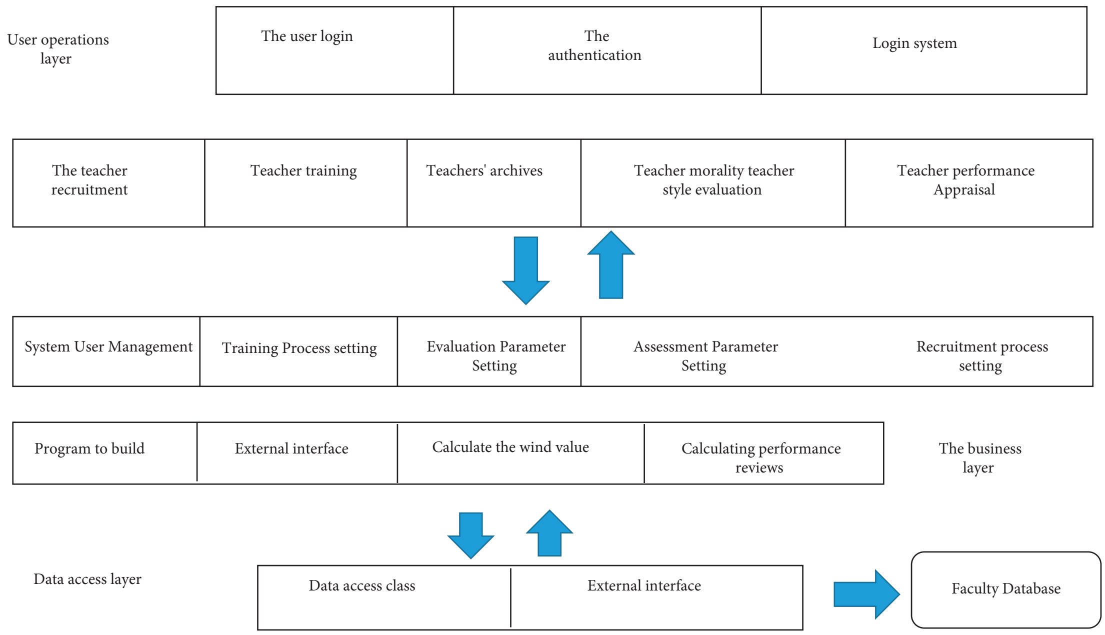


Figure 7: System architecture diagram.

faculty development management system, C# and SQL Server database were also used to develop the Web application in order to facilitate the expansion with other systems.

A data interface was established between the Faculty Development Management System and the Academic Affairs Management System and Personnel Management System developed by the training institution to share data between the systems and ensure data consistency. The information of teaching affairs can be obtained directly from the teaching affairs management system, and the information of teaching staff's files can be obtained from the personnel management system. The interface between the different systems can be implemented in various ways, but since the systems are deployed on the same server and use the same development tools, the data exchange between them can be done in many different ways. Therefore, it is relatively easy to exchange data and interface between them. We created a dynamic connection file in the academic affairs management system and personnel management system, and wrote all the methods to be opened to the public into the file, so that we can get the required data by directly referring to the file in the faculty construction management system.

Class access open interface code

```
//public DataTable getClassInfo(){...}
```

Faculty and staff access open interface code

```
//public DataTable getTeachers(){...}
```

Considering the cross-platform data interaction and processing efficiency, the external interface of teaching management designed in this system is more practical than the open interface mentioned above. The specific approach is to automatically generate files of the data to be opened to the public, which can be accessed by other systems.

4. Summary

With the development of science and technology, the teacher management of sports training institutions in the era of big data and personnel management institutions teacher construction and development are facing opportunities, but also accompanied by challenges. The big data application of sports training institutions should be regarded not only as a resource but also as a tool. While improving the ability of physical education teachers to manage training institutions, it is more important to use big data intelligent management technology to effectively transform teacher management. The rapid development of big data technology provides more opportunities for the management of teachers in sports training institutions. At the same time, there are increasingly obvious contradictions that need to be resolved. The informatization of faculty management process can match information accurately and quickly with the help of large databases, but the processing of big data information base cannot be fully intelligent due to overlapping information and other reasons, which requires opening the faculty management information base, opening multiple information sources, and strengthening faculty information interaction.

Faculty management information is available in all system files. Even if the faculty management teaching process can be carried out smoothly in a relatively independent and complete teaching system environment, it may still be nonclosed-loop management in terms of compatibility and effective connection between various teaching system environments. Therefore, strengthening the teaching process and matching it with the endpoints can be done to make the best use of big data in education and better manage the faculty effectively. Moreover, in the era of big data in sports, although faculty management in sports and training management organizations can access data and other applications simultaneously on the same regional base data, all these data are still not linked from the data distribution across provinces and municipalities and the flow to the outside. Open channels of communication and cooperation among multiple parties and setting uniform requirements for information items are important tools for open linkages. Big data technology can help sports training institutions to provide scientific and reasonable data analysis in talent selection and play an important role in the teacher evaluation system. When providing services to teachers, sports training institutions should also focus on service quality and provide teachers with the most needed and best quality services through data analysis [31].

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The author declares no conflicts of interests.

Acknowledgments

This research was supported by Research on the Operation Status and Governance of Chongqing Youth Sports Training Market in the New Media Era (sisu202130).

References

[1] M. A. Griffiths, V. A. Goodyear, and K. Armour, "Massive open online courses (MOOCs) for professional development: meeting the needs and expectations of physical education teachers and youth sport coaches," *Physical Education and Sport Pedagogy*, vol. 27, no. 3, pp. 276–290, 2022.

[2] M. Catherine Scheeler, A. Markelz, J. C. Taylor, D. S. Deshpande, and P. Wolfe, "Teacher workplace bullying: how pervasive is the problem?" *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children*, vol. 45, no. 2, pp. 123–140, 2022.

[3] D. Fang and M. Zhang, "Research on the reform path of Chinese school PE teacher training during," *International Journal of Educational Management*, vol. 7, no. 1, 2022.

[4] C. Xiang, "Research on the training mode of physical education practice and physical education professionals," *International Journal of Education and Economics*, vol. 5, no. 1, 2022.

[5] L. Wen, "Research on classroom teaching reform of physical education specialty under the background of teacher training specialty certification in colleges and universities," *International Journal of Higher Education Teaching Theory*, vol. 3, no. 1, 2022.

[6] G. Li, "Practice of inquiry teaching method in higher vocational physical education," *International Journal of Higher Education Teaching Theory*, vol. 3, no. 1, 2022.

[7] H. J. Elvis, Q. Frank, S. K. Anin et al., "COVID-19-Related knowledge and anxiety response among physical education teachers during practical in-person lessons: effects of potential moderators," *Behavioral Sciences*, vol. 12, no. 3, p. 83, 2022.

[8] Z. Anna, K. Karolina, B. Rosołek, and E. Tomińska-Conte, "Validation of the perceptions of inclusion questionnaire including PE teachers' opinion as part of an innovative use of the tool," *Sustainability*, vol. 14, no. 6, p. 3241, 2022.

[9] A. Girard, N. Gal-Petitfaux, and O. Vors, "The student's experience of teacher support in French vocational high-school classes with difficulties in school engagement in physical education: interest of mixed methods research," *Physical Education and Sport Pedagogy*, vol. 27, no. 2, pp. 156–171, 2022.

[10] T. Yu and Z. Du, "An empirical study on factors affecting the professional well-being of physical education teachers in colleges based on multiple regression," *Frontiers in Sport Research*, vol. 4, no. 1, 2022.

[11] Hao, "The current situation and suggestions of physical education curriculum for all primary school teachers in rural chongqing," *International Science Research*, vol. 2, no. 1, 2022.

[12] L. L. Gonçalves, M. Parker, C. Luguetti, and M. Carbinatto, "The facilitator's role in supporting physical education teachers' empowerment in a professional learning community," *Sport, Education and Society*, vol. 27, no. 3, pp. 272–285, 2022.

[13] S. Preece and R. Bullingham, "Gender stereotypes: the impact upon perceived roles and practice of in-service teachers in physical education," *Sport, Education and Society*, vol. 27, no. 3, pp. 259–271, 2022.

[14] J. Su, X. Pu, K. Yadav, and M. Subramnaiyan, "A physical education teacher motivation from the self-evaluation framework," *Computers & Electrical Engineering*, vol. 98, Article ID 107772, 2022.

[15] N. Bechir, K. Abdelmotaleb, N. Souissi, and M. Rouissi, "The effects of the socialization of physical education teachers on their modes of interaction with students in Tunisian schools," *Frontiers in Sociology*, vol. 6, Article ID 747092, 2022.

[16] E. H. Kwon, J. Park, and T. Kim, "Investigating the impact of adapted physical education service-learning projects on preservice teachers' attitudes toward individuals with disabilities," *Journal of exercise rehabilitation*, vol. 18, no. 1, pp. 74–78, 2022.

[17] A. Prieto-Ayuso, M. P. León, O. R. Contreras-Jordan, and D. Morley, "Spanish physical education teachers' perceptions of talent development," *European Physical Education Review*, vol. 28, no. 1, pp. 3–19, 2022.

[18] J. Liu and X. D. Keating, "Development of the pre-service physical education teachers' teacher identity scale," *European Physical Education Review*, vol. 28, no. 1, pp. 186–204, 2022.

[19] G. David, G. López Luis, and P. A. Hastie, "Adoption and fidelity of sport education in Spanish schools," *European Physical Education Review*, vol. 28, no. 1, 2022.

[20] D. Hortigüela-Alcalá, D. Bores-Garcia, R. Barba-Martin, and G. Gonzalez-Calvo, "Bores at al. Yes, I have a disability. Does it prevent me from being a PE teacher? A qualitative approach from the point of view of future PE teachers and their families," *Sport, Education and Society*, vol. 27, no. 2, pp. 196–209, 2022.

[21] J. S. Deborah, R. Jeanette, W. Erika, P. Missy, and C. R. Julie, "Program evaluation of healthy Moves™: a community-based trainer in residence professional development program to support generalist teachers with physical education instruction," *Journal of Physical Activity and Health*, vol. 19, no. 2, pp. 125–131, 2022.

[22] M. Arteaga-Checa, M. V. Palop-Montoro, and D. Manzano-Sanchez, “Application of a program to improve personal development in future physical education teachers of the degree in education and its relationship with wisdom,” *Sustainability*, vol. 14, no. 3, p. 1188, 2022.

[23] M. B. Chaifa and A. M. Naceur, “Effect of gender of Tunisian teachers on teaching practice during gymnastics sessions,” *Creative Education*, vol. 13, no. 04, pp. 1368–1383, 2022.

[24] R. R. Jorge, M. R. Fernando, C. A. Jose, A. D. Angel, G. P. Santiago, and B. F. Sabina, “Spanish physical education teachers’ perceptions about their preparation for inclusive education,” *Children*, vol. 9, no. 1, p. 108, 2022.

[25] H. Mostafa, L. Said, and T. Mohammed, “Design and implementation of a distance training device of physical and sports education’s future teachers: application of ADDIE model,” *International Journal of Sciences: Basic and Applied Research*, vol. 61, no. 1, pp. 70–86, 2022.

[26] M. Duggan, “Instilling positive attitudes to physical activity in childhood – challenges and opportunities for non-specialist PE teachers,” *Education*, vol. 50, no. 1, pp. 129–143, 2022.

[27] H. Larsson, M. Torun, and F. Magnus, “(Non-)Diversity and cultural (re)production in physical education teacher education: a Swedish example,” *Curriculum Studies in Health and Physical Education*, vol. 13, no. 1, pp. 3–18, 2022.

[28] C. Valérian, G. D. Emma, and L. K. Vanessa, “Teacher well-being and perceived school climate during COVID-19 school closure: the case of physical education in Switzerland,” *The Teacher Educator*, vol. 57, no. 1, pp. 45–60, 2022.

[29] R. Eirín-Nemiña, A. Sanmiguel-Rodríguez, and J. Rodriguez-Rodriguez, “Professional satisfaction of physical education teachers,” *Sport, Education and Society*, vol. 27, no. 1, pp. 85–98, 2022.

[30] M. Atiquzzaman, N. Yen, and Z. Xu, “Big data analytics for cyber-physical system in smart city,” *Advances in Intelligent Systems and Computing*, Shenyang, China, 2019.

[31] W. Barber, W. Walters, P. Chartier, and C. Temertzoglou, “Examining self-Confidence and self-Perceived competence in Canadian pre-service teachers (PSTs): the role of biographies in physical education teacher education (PETE),” *Sport, Education and Society*, vol. 27, no. 3, pp. 347–360, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9796507, 1 page
https://doi.org/10.1155/2023/9796507

Retraction

Retracted: Exploring the Application of Online Teaching Platform in the PBL Teaching Mode of Film and Television Courses for Teacher Trainees

Mobile Information Systems

Received 22 August 2023; Accepted 22 August 2023; Published 23 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Peng and M. Yuan, "Exploring the Application of Online Teaching Platform in the PBL Teaching Mode of Film and Television Courses for Teacher Trainees," *Mobile Information Systems*, vol. 2022, Article ID 2289981, 13 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 2289981, 13 pages
https://doi.org/10.1155/2022/2289981

Research Article

Exploring the Application of Online Teaching Platform in the PBL Teaching Mode of Film and Television Courses for Teacher Trainees

Jian Peng and Meng Yuan

College of Education, Hunan First Normal University, Changsha 410205, Hunan, China

Correspondence should be addressed to Jian Peng; dysfpj@hnfnu.edu.cn

Received 21 May 2022; Revised 27 June 2022; Accepted 4 July 2022; Published 10 August 2022

Academic Editor: Xuefeng Zhang

The aim of the study is to deeply study the effect of the PBL teaching mode in improving students' problem-solving ability, learning interest and autonomous learning ability. This study takes the practice teaching of film and television courses for normal students as the research object, constructs a general network teaching platform based on B/s and PBL teaching mode, deeply investigates the teaching status of film and television production practice courses, and selects the experimental class and control board to test the PBL teaching mode. The test results show that 13 students in the experimental class have a general learning attitude, 12 students have a good learning attitude, 7 students have a good learning attitude, and 5 students have a very correct learning attitude. This shows that 24 students have a positive learning attitude in the PBL teaching mode, which shows that they are very interested in learning information technology courses. At the same time, it also verifies the feasibility of the PBL teaching mode. At the same time, on this basis, this study puts forward some suggestions for the optimization of PBL teaching design under the network teaching platform.

1. Introduction

At this stage, with the development of information technology, China's information technology curriculum has also developed to a certain extent. However, there are still some problems in the current information technology curriculum, such as single teaching mode, relatively poor teaching effect, and low students' interest in learning. The proposal of the PBL teaching mode is helpful to improve these problems [1]. Of course, the PBL teaching mode is mainly used in the medical field. The so-called PBL teaching mode is a set of teaching methods for designing learning situations, which can also be regarded as project-based teaching methods or problem-based learning methods. The earliest application in the field of medical education is a problem-oriented and student-centered teaching method. This teaching model can make students gradually become the role in the problem situation, and teachers can also design and arrange the content of the course around a complete problem, actively encourage students to learn knowledge related to the problem, and then better solve the problem. Of course, by creating a specific learning environment, teachers also help to stimulate students' independent thinking, actively encourage students to ask questions, and constantly guide students to solve problems [2].

2. Literature Review

Some scholars believe that PBL is a teaching method. For example, PBL is a situational teaching method that takes problems as the starting point of teaching activities, so as to teach courses, take students as the center, and solve real problems; Royle believes that PBL is a student-centered teaching method that "anchors" learning in problem situations. Some scholars define it as a teaching method. For example, Saville believes that PBL is a teaching method that takes learners as the center and learners apply their knowledge and skills to develop solutions to a specific problem [3]; Qi et al. believe that PBL is a teaching method based on the design of problems close to reality to cultivate students' knowledge system, problem-solving ability, high-order thinking ability, and cooperative communication ability. Most domestic scholars define PBL as a teaching mode [17]. For example, Li et al. believe that PBL is a teaching mode that focuses on problems, allows students to construct knowledge according to problems, guides students to master basic knowledge, and develops high-level thinking skills, problem-solving ability, and autonomous learning ability [4]; Hu believes that PBL is a teaching model that puts students in specific problem situations and encourages students to carry out organized and in-depth inquiry learning on problems through autonomous and cooperative learning [5]; Sanjeev et al. put forward that PBL is a teaching model worthy of reference. It takes problems as the basis. Learners think and analyze problems, query materials, actively learn knowledge, and improve students' thinking skills [6].

With the deepening of PBL research, PBL has obviously developed into a teaching model. In the problem environment, according to the existing knowledge and experience and the help of the outside world, students can independently study and explore the hidden knowledge in the problem in the form of group cooperation. In the PBL classroom, the role of teachers has changed greatly. Teachers are not only rule makers, but also instructors and supervisors in the classroom. Students have become the main body of the PBL classroom and their status in the classroom has been improved.

3. Network Teaching Platform Based on *B/S*

3.1. Framework Design of Network Teaching Platform. After logging into the system, you can enter three subsystems: test question bank management subsystem, online examination subsystem, and online learning subsystem. In order to clearly explain the system framework and have a preliminary understanding of the system process as a whole, the main process of the system is given as shown in Figure 1.

The system is divided into a task manager module and a student work module according to the task. The main functions of the system are shown in Figure 2.

The system is composed of a test question bank management subsystem, online examination subsystem, and online learning subsystem (basic teaching subsystem of Computer Culture).

3.2. Database Design of Network Teaching Platform

3.2.1. Database Planning. First, establish the access database, store the examinee's personal information and the test questions of each course in the database "test_database. MDB," and store it in the data directory of the system the database "test_database. MDB" contains the following 8 data tables: administrator table (admin), examinee file table (student), examinee examination record table (report), examination paper table (test), judgment question table (Type1), single choice question table (type2), blank filling question table (type3), and multiple topic table (type4).

3.2.2. E-R Diagram Analysis. According to the total data structure of the designed test question bank system, using an E-R diagram, and according to the system analysis and module design of the system, we can make various entities that can meet the needs of users and the relationship between them. According to the above plan, the system plans out: the administrator information entity, student information entity, test paper information entity, examination information entity, judgment question entity, single choice question entity, multiple choice question entity, and blank filling question entity. Figure 3 shows the E-R diagram of the system database.

The entity diagram of platform system administrator information is shown in Figure 4.

The student information entity diagram of the platform system is shown in Figure 5.

The examination information entity diagram of the platform system is shown in Figure 6.

The physical diagram of test paper information of the platform system is shown in Figure 7.

The entity diagram of platform system judgment question is as shown in Figure 8.

The entity diagram of single choice questions of the platform system is shown in Figure 9.

The multi-topic entity diagram of the platform system is shown in Figure 10.

The entity diagram of filling in the blanks of the platform system is shown in Figure 11.

3.2.3. Database Table Design. There are three fields in the administrator table: administrator ID, administrator account, and administrator password. The contents are shown in Table 1.

This form includes the candidate's personal information, name, student number, and registration date. The contents of the student information table are shown in Table 2.

This form records the examinee's number, test date, test score, etc. If a make-up examination is required, the make-up examination field is true, and the make-up examination date and score are recorded at the same time. The contents of the candidate information table are shown in Table 3.

You can specify the number of each question type of the test paper and randomly select the test questions from the test question library to generate an original test paper. The contents of the test paper information table are shown in Table 4.

The judgment question table records the judgment question stem and answer set by the administrator, as well as the addition date. Its design table is shown in Table 5.

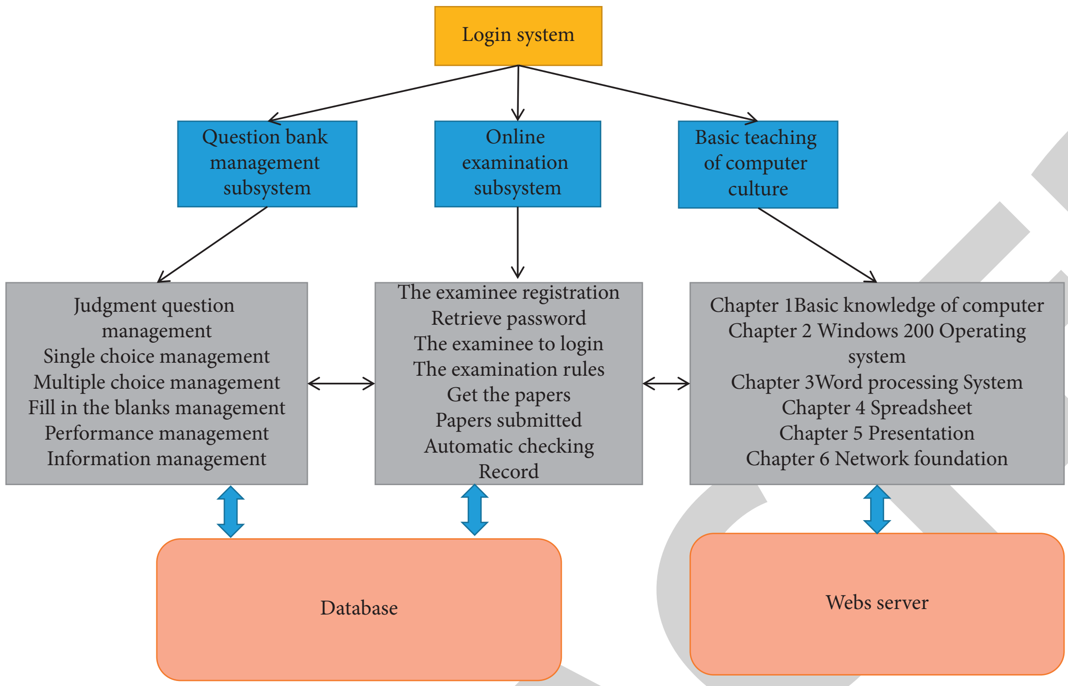


Figure 1: Main process of system design.

The single choice question table records the single choice question stem and answer set by the administrator, as well as the joining date. Its design is shown in Table 6.

The multi-choice question table records the multi-choice question stems and answers set by the administrator, as well as the joining date. Its design is shown in Table 7.

The blank filling form records the blank filling stem and answers set by the administrator, as well as the addition date. Its design is shown in Table 8.

This section studies and develops a web-based network teaching platform. The teaching platform is composed of test question bank management subsystem, online examination subsystem, and online learning subsystem. It mainly studies test question management, test paper management, online examination, score management, information management, online learning subsystem, and so on. There are four types of examination questions: judgment questions, single choice questions, multiple-choice questions, and blank filling questions, which are authorized by administrators and students.

4. Experimental Research on PBL Teaching Mode in Curriculum

4.1. Preliminary Preparation of the Study. Before carrying out PBL teaching, this study conducted a questionnaire survey on each class of Grade 8 in junior middle school. The questionnaire is mainly composed of five dimensions: problem-solving ability, learning interest, cooperative learning ability, autonomous learning ability, and information ability. According to the analysis of the survey results, it is found that there are significant differences between class 8 of Grade 8 and class 5 of Grade 8 in the dimension of autonomous learning ability, but there are no significant differences in problem-solving ability, learning interest, cooperative learning ability, and information ability. Therefore, this study takes all the students in these two classes as the research object [7, 8]. The quasi-experimental research of PBL teaching is carried out with the unit of animation design in the first volume of information technology of Grade 8 published by Tsinghua University Press as the teaching content. In the early stage of the experiment, this study takes class 5, Grade 8 of s middle school as the experimental class and class 8, Grade 8 as the control class. In order not to affect the normal teaching activities of the school, this study selects two classes in a normal way, and the researcher carries out the experiment as a teacher.

The specific experimental scheme is shown in Table 9:

The independent transitions in this experiment are multidisciplinary, and the main differences are the students' problem-solving skills, collaborative learning, independent learning, enjoys learning, and literacy. The content of the two classes is the same, and the classes are two hours a week.

4.2. Research Process

4.2.1. Early Stage of PBL Teaching. According to the content of animation design, create problem situations, ask questions, and guide students to analyze problems. Then, the teaching of the first module "getting to know components" was carried out. According to the content of this module, this study has made the following teaching design and implemented classroom teaching. The specific target framework and contents are shown in Table 10:

Before PBL teaching, learners have learned the basic operation of Flash software, including basic animation production skills such as importing image modification,

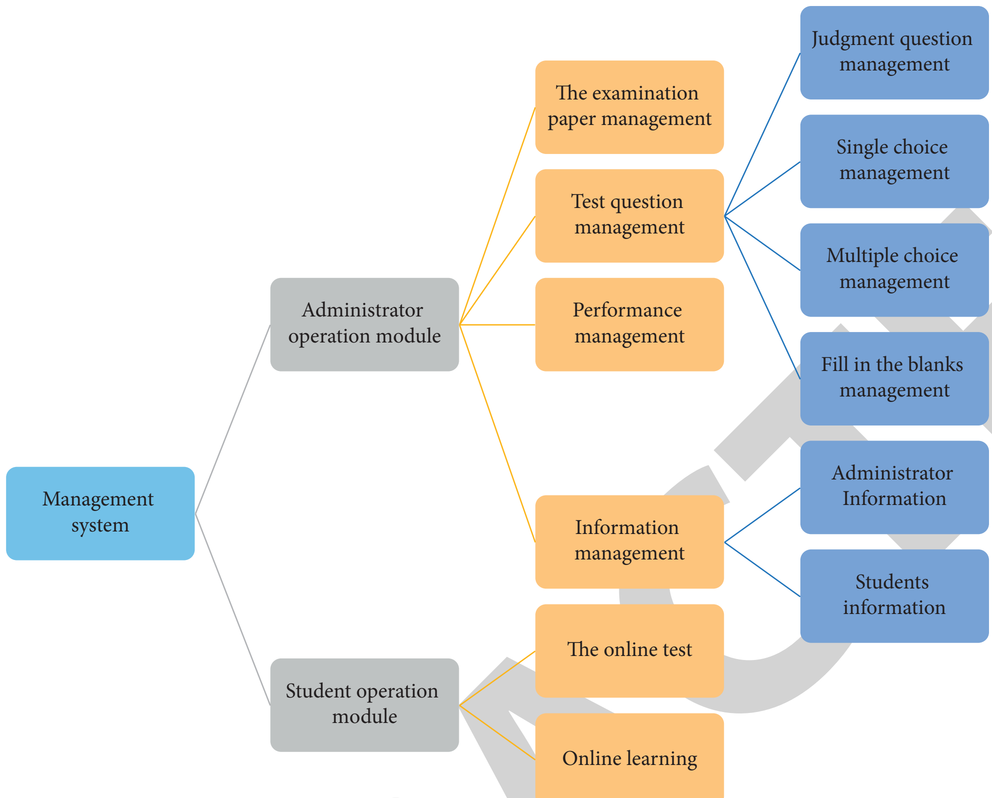


Figure 2: System function division.

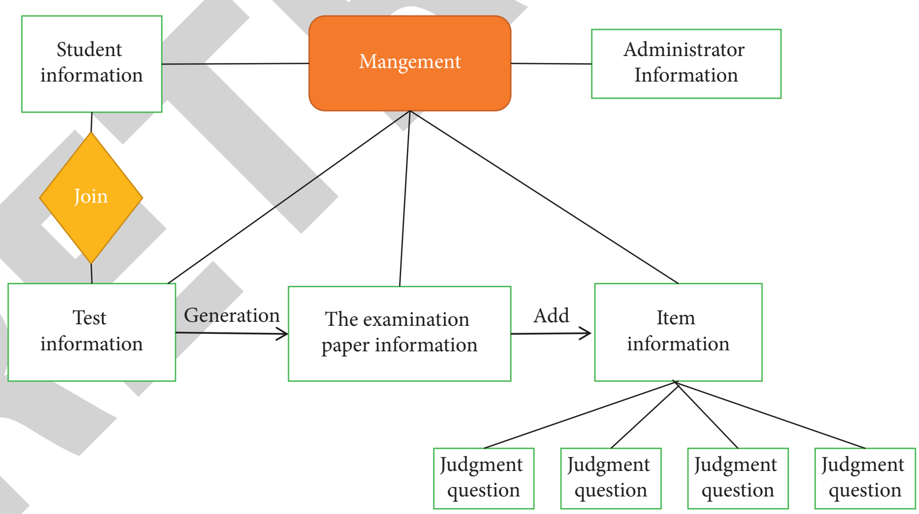


Figure 3: E-R diagram of system database.

frame-by-frame animation, creating gap shape, and drawing image, which lays a foundation for the learning of components. In the past, learners only need to imitate the teacher's operation steps to complete the exercises. In the PBL teaching mode, learners mainly rely on self-study and group discussion to learn new knowledge and skills. For the eighth-grade students, the content of the component module is not difficult. Therefore, in this module teaching, teachers will teach according to the characteristics of PBL teaching and students' current learning situation, so that students can gradually adapt to the PBL teaching mode.

4.2.2. Middle Term of PBL Teaching. Guide layer animation is the learning content of classes 9–10 in PBL teaching this semester. In animation, there is often an animation effect that an object moves along the preset path. We call it to guide

Figure 4: Entity diagram of administrator information.

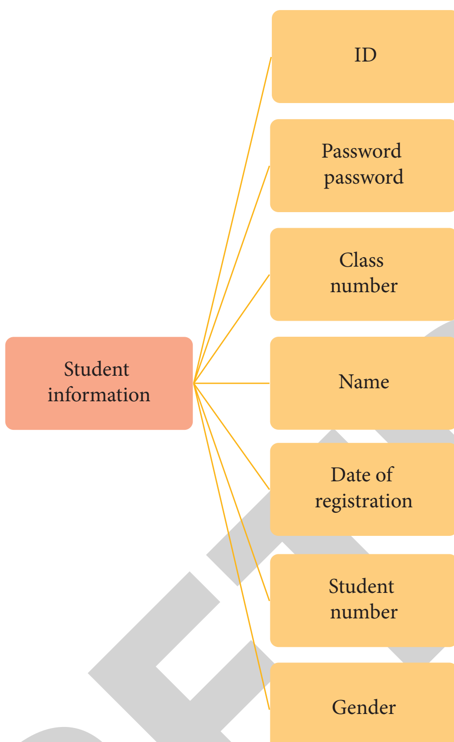


Figure 5: Entity diagram of student information.

layer animation. It is widely used in flash animation. The guide layer animation is an extension of the traditional patching animation. Therefore, learning the guide layer animation well also lays a foundation for the comprehensive application of students' flash works [9]. Teaching objectives are shown in Table 11:

4.2.3. Later Stage of PBL Teaching. In the later stage of PBL teaching, the teacher spent two class hours to complete it, mainly for each group to display the animation greeting card works of the group, the team leaders and teachers of each group to score the works, and the students give scores for their cooperative learning and learning attitude with the group members according to the corresponding evaluation criteria [10]. The specific process is shown in Table 12.

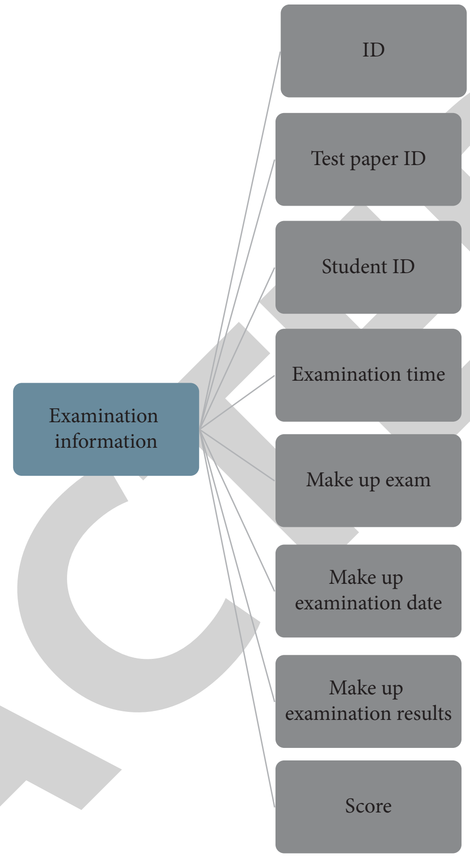


Figure 6: Entity diagram of examination information.

4.3. Experimental Effect Analysis of PBL Teaching Mode. As shown in Table 13, prior to the experiment, Sig (two sides) value of the four dimensions of the laboratory and control room, problem-solving skills, learning satisfaction, collaborative learning, and capacity of data, was greater than 0.05. There is no significant difference between the test room and the control room in these four dimensions. The Sig (two-way) measure of independent learning ability was less than 0.05 and higher than 0.01, indicating differences in self-study ability between the experimental room and the control room.

4.3.1. Effect Analysis of Problem Capability Dimension. To understand the possible changes in the solutions of the two students after the intervention experiment, we analyzed the pre and posttest data to test students' problem-solving skills in Grade 2, as described in Table 14. Prior to the test, the accuracy of the test room was 3.44 and the average score of the control room was 3.54, and the average of the test center was slightly lower than that of the control room. From the results of the two-variable solution potential *t*-test, we can see that the Sig value (two tails) is greater than 0.05 before the test. This shows that there is no significant difference in the density of the pretest laboratory and the control room. After the test, the average density of the test room was 4.04 and the average control room was 3.69. The

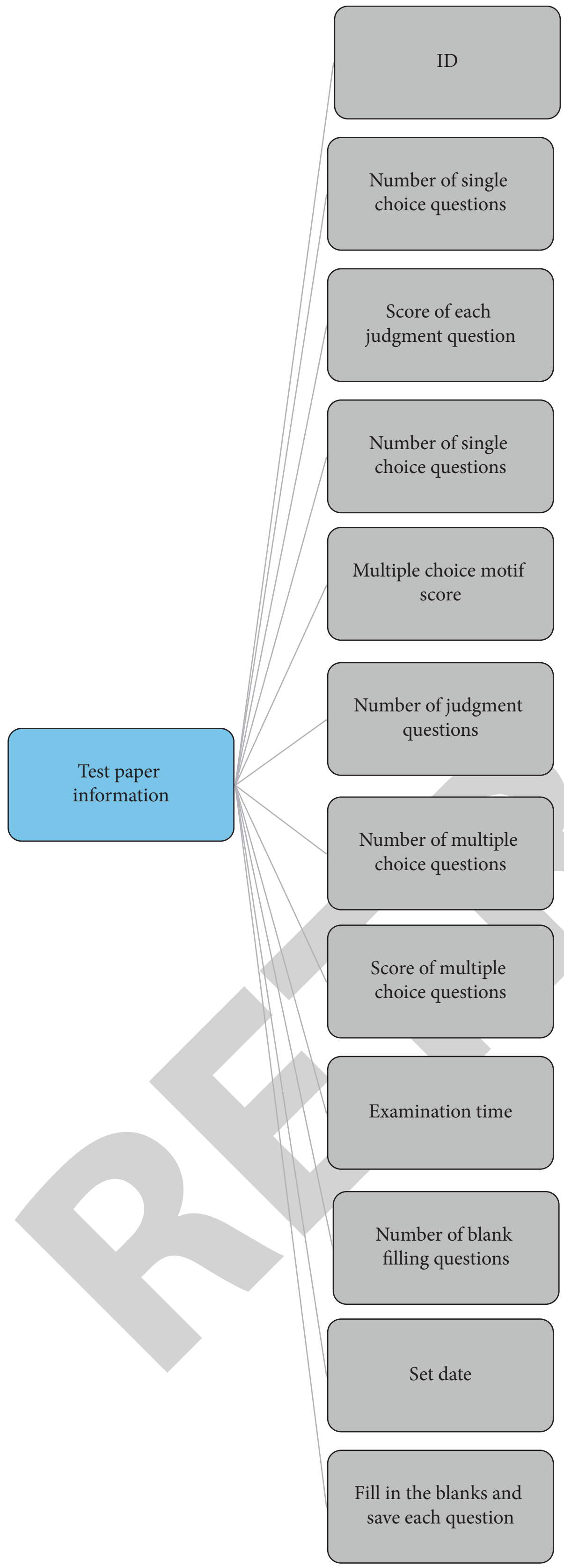


Figure 7: Entity diagram of test paper information.

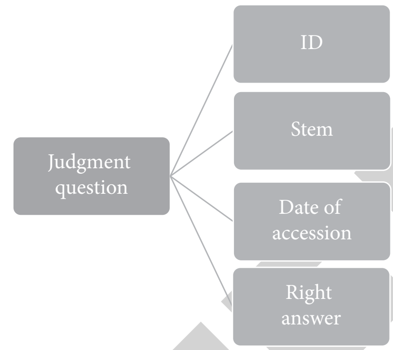


Figure 8: Entity diagram of judgment question.

average laboratory score is higher than that of the control room. Although the compactness of the two classes improved compared to the pretest results, the improvement of the experimental classes was much higher than that of the control classes [11]. After solving the two classifications, the Sig value in the standard difference of the experimental data is less than 0.05, which indicates that the experimental data after the capacity difference show that the solutions of the two classes are not the same and have significant differences. There is no other way to determine between the two classes. Therefore, based on the *t*-test scores of the test and control classes (posttest): If the Sig (double-tailed) value is less than 0.01, it can be concluded that there is a difference in the laboratory's problem-solving capacity. Control room before and after the test. This is shown in Table 15.

Problem-solving ability is the comprehensive ability of students in solving problems. The reason why there is a significant difference between the control class and the experimental class is mainly because in the experimental class, teachers will guide students' main tasks in each link step by step according to the PBL teaching mode process. PBL teaching mode mainly formulates a set of detailed steps to solve problems based on problem solving. Although in the early stage of the experiment, students are still unclear about the process of PBL teaching mode. However, in each module teaching, teachers will design corresponding problem situations according to the teaching content [12]. Under the guidance of teachers, students find problems, analyze and solve problems, and gradually adapt to and get used to this set of the problem-solving process and form a problem-solving thinking mode. In the control class, teachers mainly use lecture teaching methods and teach according to the content of teaching materials. Although most students can master the operation skills demonstrated by teachers in class, once they encounter problems and cannot find the answer, most students will choose to give up and solve this problem. Therefore, once the hypothesis is established, compared with the traditional teaching mode, the PBL teaching mode is more conducive to improve students' problem-solving ability [13]. This is shown in Table 16.

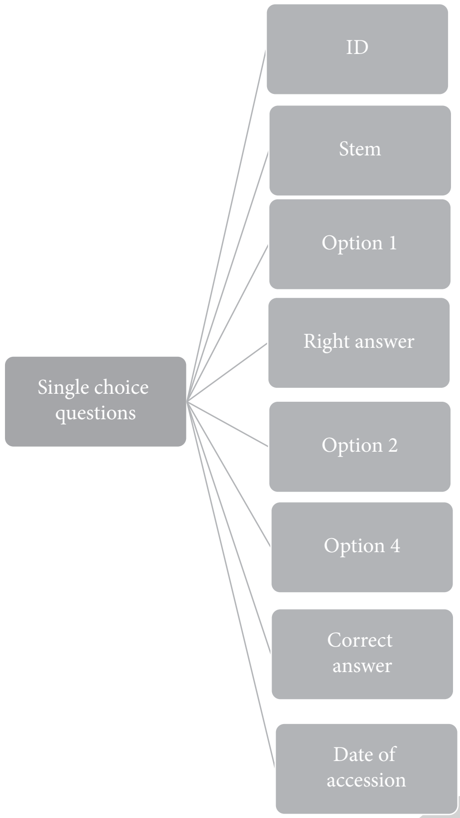


Figure 9: Entity diagram of single choice questions.

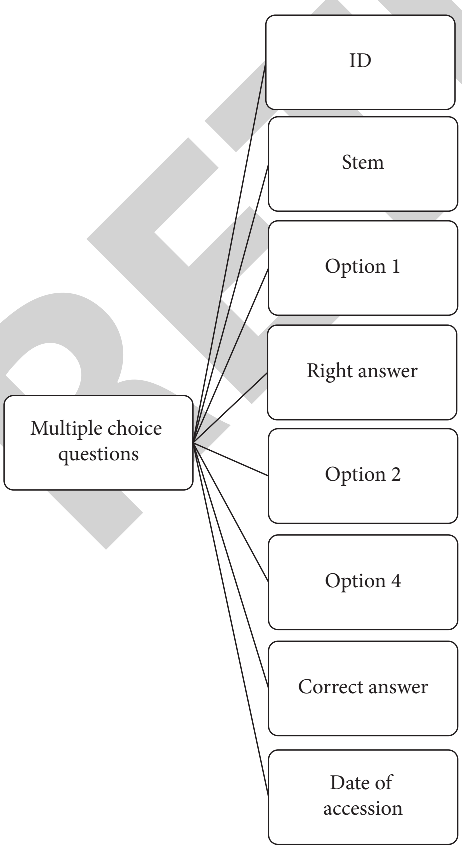


Figure 10: Entity diagram of multiple topics.

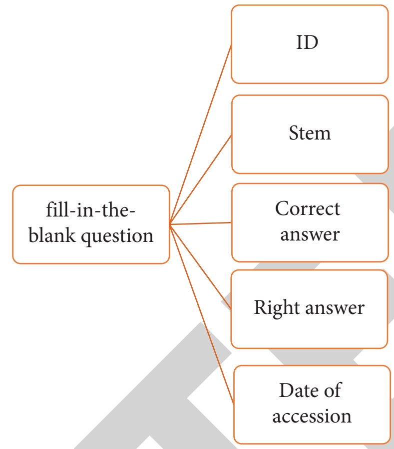


Figure 11: Entity diagram of blank filling questions.

Table 1: Administrator table.

Field name	Meaning	Data type
ID	Administrators	Auto number
Account number	Administrator account	Text
Password	Administrator password	Text

Table 2: Student information.

Field name	Meaning	Data type
ID	Student number	Auto number
Full name	Student name	Text
Password	Password	Text
Gender	Gender	Text
Class number	Class number	Text
Student number	Student number	Text
Date of registration	Date of registration	Date/time

4.3.2. Analysis on the Effect of Learning Interest Dimension. To understand changes in student interest in both classes after the experiment, we analyzed pretest and posttest data from information on student interest in both classes. Prior to the experiment, the laboratory average score was 3.97, the control room score was 4.06, and the laboratory average score was slightly lower. The pretest t-test students' interest in learning showed that the pretest Sig (double-tailed) value was >0.05, indicating that there was no difference between student satisfaction in the laboratory and in the control room Before the test. In the event of a test discrepancy after the survey data have been provided, the Sig value >0.05 and the difference between the two samples can be detected in the same way [1]. After eliminating the effects on students' early learning interest after the experiment, the average of the pilot classes was 4.14 and the average of the control classes was 3.92. The average laboratory score is higher than that of the control room. Based on a one-sided comparison of learning satisfaction, $0.01 < \mathrm{Sig} = 0.037 < 0.05$ indicates that learning in experimental and control classes eliminates the effects of the previous testing. This is shown in Table 17.

Students' learning attitude in class can also reflect students' learning interest to a great extent. In order to make the research results convincing, this study

Table 3: Examination information.

Field name	Meaning	Data type
ID	Examination record No	Auto number
Test paper ID	Test paper ID	Number
Student ID	Student ID	Number
Fraction	Test score	Number
Examination date	Examination date	Date/time
Make-up exam	Make-up exam	Yes/no
Make-up examination results	Make-up examination results	Number
Examination date	Make-up examination date	Date/time

Table 4: Test paper information.

Field name	Meaning	Data type
ID	Test paper No	Auto number
Number of judgment questions	Number of judgment questions	Number
Score of each judgment question	Score of each judgment question	Number
Number of single choice questions	Number of single choice questions	Number
Number of single choice questions and score of each question	Score of single choice questions	Number
Number of multiple choice questions	Number of multiple choice questions	Number
Score of multiple choice questions	Multiple choice media score	Number
Number of blank filling questions	Number of blank filling questions	Number
Score of each blank filling question	Score of single blank filling question	Number

Table 5: List of judgment questions.

Field name	Meaning	Data type
ID	Judgment question number	Auto number
Stem	Stem	Text
Right key	Right key	Yes/no
Date of accession	Date of accession	Date/time

Table 6: Single topic selection table.

Field name	Meaning	Data type
ID	No. of single choice questions	Auto number
Stem	Stem	Text
Option 1	Single choice option 1	Text
Option 2	Single choice option 2	Text
Option 3	Single choice option 3	Text
Option 4	Single choice option 4	Text
Right key	Right key	Text
Date of accession	Date of accession	Date/time

Table 7: Multiple topics.

Field name	Meaning	Data type
ID	Multiple choice number	Auto number
Stem	Stem	Text
Option 1	Multiple choice option 1	Text
Option 2	Multiple choice option 2	Text
Option 3	Multiple choice option 3	Text
Option 4	Multiple choice option 4	Text
Right key	Right key	Text
Date of accession	Date of accession	Date/time

Table 8: Blank filling form.

Field name	Meaning	Data type
ID	Fill in the blanks No	Auto number
Stem	Stem	Text
Right key	Right key	Text
Date of accession	Date of accession	Date/time

statistically analyzed 50 learning attitude evaluation scales distributed to the students of the experimental class and obtained the scores. Among them, the scale mainly evaluates and measures student's learning attitude from the three aspects of participation, preparation and attendance, and then grades each student's learning attitude. In the reliability analysis of the scale, the calculated Kendall harmony coefficient (W) value is 0.675, and the evaluated objects are 50 people, $n > 7$. It is also necessary to convert the W value into x^2 value. The test shows that $x^2 = 99.24 > X^2$ (49) 0.01, indicating that the obtained w value has reached a very significant level, and the grade consistency assessed by raters is very high. The KMO value of the scale is 0.860, indicating that the scale has high validity. It is suitable for measuring students' learning attitude [5].

Among them, the learning attitude in this study is divided into five levels. Among them, 1 indicates that the attitude is not correct, 2 indicates that the attitude is general, 3 indicates that the learning attitude is good, 4 indicates that the learning attitude is very good, and 5 indicates that the learning attitude is very correct, as shown in Figure 12. Among them, 13 students in the experimental class have an incorrect learning attitude in the PBL teaching mode this semester. 13 students have an average learning attitude, 12 students have a good learning attitude, 7 students have a good learning attitude, and 5 students have a very correct learning attitude [14]. This shows that 24 students have a positive learning attitude in the PBL teaching mode,

Table 9: Quasi-experimental research scheme.

	Pretest	Independent variable	Posttest (dependent variable)
G1 (experimental class)	O_1	Teaching mode 1	O_2
G2 (control team)	O_3	Teaching mode 2	O_4
		Experiment time: 20 class hours	

Table 10: Framework of teaching objectives of "new components."

Knowledge and skills	(1) Master and understand the different knowledge and skills of graphic elements, film editing elements and button elements (2) Master the method of creating and editing components (3) Understand the library and learn to apply the library to manage components
Process and method	(1) In autonomous learning, students can learn the principle, manufacturing method, process and method of components (2) Through the form of group discussion, students can learn new knowledge and skills and begin to adapt to new learning forms
Emotional attitude and values	(1) Through the introduction of problem situation, improve students' desire to explore component knowledge (2) Initially cultivate students' autonomous learning ability and sense of cooperation (3) Through the guidance of teachers, students' emotional attitude and values ability and habits of discovering, analyzing and solving problems are preliminarily cultivated (4) After students learn new knowledge, through the guidance of teachers, cultivate students' application and innovation ability of new knowledge

Table 11: Framework of teaching objectives of "guide layer animation."

Knowledge and skills	(1) Master the meaning and characteristics of the guideline (2) Master the relationship between the guided layer and the guided layer (3) Learn to create guide layers and draw guide line paths
Process and method	(1) Through the creation of problem situations, stimulate students' thirst for knowledge and arouse students' thinking and discussion (2) Through the cooperation of team members, we can cultivate the awareness of team members to communicate with the old work and enrich the new work according to the group knowledge
Emotional attitude and values	(1) Improve students' interest in learning information technology courses, so that students can obtain a sense of achievement when realizing the animation effect of the guide layer (2) By creating problem situations, stimulate students' enthusiasm to explore problems, so as to cultivate students' problem-solving ability (3) Cultivate students' cooperative learning ability and autonomous learning ability in group discussion and communication (4) enhance the friendship between students when making animation works in groups

indicating that they have a high interest in learning information technology courses, and 13 students have a general learning attitude, which also shows that their interest in learning information technology courses is not high to a certain extent, and the other 13 students have an incorrect learning attitude, indicating that the teaching effect of PBL teaching mode to stimulate their interest in learning is low.

In the experimental class, one or two students in each group have low learning enthusiasm. Generally, they do not speak actively in group discussion and study by themselves. The teacher's energy is limited and cannot pay attention to the students with low learning enthusiasm in each class. But on the whole, the students in each group are serious and active in group discussion and autonomous learning. In the control class, the students only operate according to the teacher's presentation and teaching materials. Some students will talk and discuss when the teacher gives a speech. When they practice independently, when they encounter problems that cannot be solved, some students will choose to browse the web to spend their time. After finishing the exercises, only a few students in the control class will independently learn other information technology disciplines, while other students surf the Internet and play games. Generally speaking, the control class is not enthusiastic about the exploration of information technology knowledge [15].

The main reason is that in the PBL teaching mode, teachers create interesting problem situations for students, which are closely related to reality, and students are easy to understand. Moreover, flash animation production is mainly to train students' practical ability. In the PBL teaching classroom, teachers have little teaching time. The whole classroom is mainly dominated by students' self-study and cooperative communication, which gives students enough time for hands-on practice. In addition, teachers provide students with various learning resources to facilitate students' learning [16]. PBL teaching mode not only mobilizes students' learning initiative and gives students sufficient learning resources and a good learning atmosphere, but also enables students to obtain a sense of achievement in hands-on practice and exploration, so as to stimulate students' interest in learning information technology. In the traditional teaching mode, teachers only talk about the animation

Table 12: Post PBL teaching content.

Teaching process	Teacher behavior	Student behavior	Design intent
Exhibition of works	Organize each group to arrange a reporter to report the idea of the work and the production process of the group work	(1) Submission of works by each group (2) Each group will arrange a student to introduce the production idea, production process and division of labor of each group member	Students practice their language organization ability by presenting their works in the whole class
Summary evaluation	(1) Make oral comments on the works of each group, point out the advantages and disadvantages of the works, and guide students to make comments on the works of other groups and reflect. At this stage, both students and teachers are evaluators and reflectors. (2) Fill in the evaluation scale (Appendix III) and evaluate the works of each group and the students' learning attitude and cooperation. (3) After class, ask two other IT teachers to evaluate the students' works and fill in the work evaluation scale	(1) The team leader of each group shall evaluate the works of this group and other groups, and fill in the work evaluation scale. (2) Students' self-evaluation and mutual evaluation of learning attitude and cooperation, and fill in the evaluation scale	In order to obtain the evaluation results of each group's works objectively and fairly. Therefore, this study will calculate the average score of three teachers' evaluation group works according to the evaluation scores of three teachers. In the stage of students' self-evaluation and mutual evaluation, students can initially realize their learning ability

Table 13: *T*-test of independent samples before test.

	Variance homogeneity test		*T*-test of mean variance		
	F	Sig	*T*	df	Sig (bilateral)
Problem-solving ability	0.076	0.783	1.047	96	0.298
			1.045	93.472	0.299
Learning interest	2.525	0.115	0.602	96	0.549
Cooperative learning ability	1.765	0.187	−0.743	96	0.459
			−0.739	88.436	0.462
Autonomous Learning Ability	0.413	0.522	2.088	96	0.993
			2.092	94.102	0.993
Ability information	0.557	0.457	0.008	96	0.993
			0.008	94.102	0.993

Table 14: *T*-test data of independent samples before and after the test of problem-solving ability of experimental class–control class.

		Class	*N*	Variance homogeneity test			*T*-test of mean equation		
				Average	*F*	Sig	*T*	df	Sig (bilateral)
Problem-solving ability	Pretest	Control class	48	3.5365	0.076	0.783	1.047	96	0.298
		Experimental class	50	3.4400			1.045	93.472	0.299
	Posttest	Control class	48	3.6917	4.343	0.040	−2.858	96	0.005
		Experimental class	50	4.0360			−2.839	84.690	0.006

Table 15: *T*-test results of problem-solving ability in experimental class and control class (posttest–pretest).

	Variance homogeneity test		*T*-test of mean equation		
	F	Sig	*T*	df	Sig (bilateral)
Problem-solving ability	0.126	0.723	−3.445	96	0.001
			−3.436	93.392	0.001

Table 16: Statistical data before and after the experimental class — the control class's learning interest.

					Variance homogeneity test	
		Class	N	Average	F	Sig
Learning interest	Pre-test	Control class	48	4.0573	2.525	0.115
	Posttest (eliminate the influence of pretest)	Control class	48	3.921	0.024	0.818
			50	4.140		

Table 17: One-way covariance analysis of learning interest.

Source	Class III sum of squares	Df	Mean square	F	Significance	Local ETA square
Modified model	1.170[a]	2	0.585	2.231	0.113	0.045
Intercept	52.927	1	52.927	201.852	0.000	0.680
Class	1.170	1	1.170	4.463	0.037	0.045

production steps for students according to the teaching materials. Students imitate the teacher's steps to make works, but when they encounter difficulties in animation production, they lack the awareness of active help seeking, and do not feel the sense of achievement of making works in their study, so students will be afraid and bored of information technology courses. Therefore, hypothesis 2 is established. Compared with the traditional teaching mode, the PBL teaching mode is more conducive to maintaining and improving students' learning interests [17].

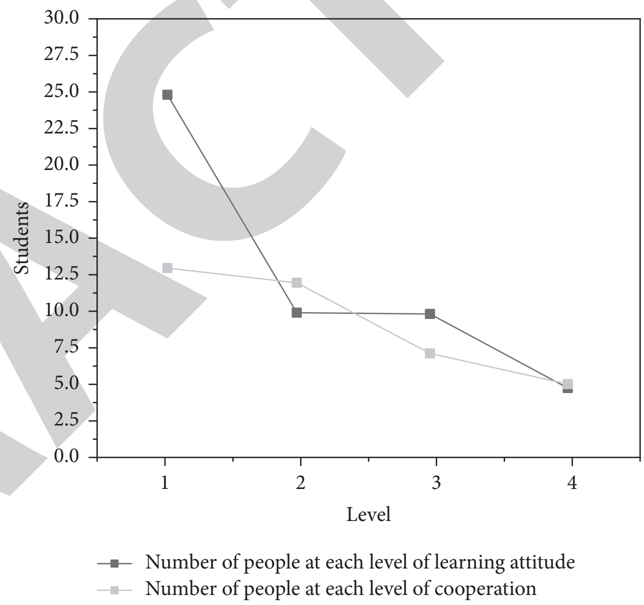


Figure 12: Statistical chart of learning scale.

4.4. PBL Teaching Mode Design

4.4.1. Curriculum Teaching Goal Design. In the online course designed based on PBL, knowledge is not taught directly by teachers to students, but acquired and formed by students in building the problem framework, exploring solutions, and making effective decisions. Therefore, the design of knowledge objectives is very important and difficult. Teachers should design the unit knowledge to be learned into problem situations. When designing, we should consider how to closely connect the learning of knowledge with problem situations and learning tasks according to the teaching materials. The ability to solve problems is called a skill goal, including thinking ability, cooperation ability, application information, evaluation information ability, and so on. The design focuses on allowing the diversity and hierarchy of problem solving [18].

4.4.2. Course Content Design. PBL is based on problems, so the design of problems is very important and should contain the basic structure of the discipline.

According to the knowledge and skill objectives of the course, it is proposed to take the "class student achievement management system" as the teaching example at the early stage of the course teaching, which runs through the whole teaching process. PBL usually encourages active learning and self-directed development, so corresponding to the multiple goals that PBL can achieve, PBL emphasizes integrated evaluation and situational evaluation, and suggests combining student evaluation with tutorial evaluation. Such an evaluation process is integrated with the learning process and implemented in the way of authenticity. It is embedded, transparent, dynamic, continuous, nontraditional and nonclosed orientation [19]. The teaching process is usually evaluated by the combination of group member evaluation, self-evaluation, and teacher evaluation. Teachers should guide students to summarize their experience or harvest in the whole problem-solving process and reflect on the shortcomings in the problem-solving process, so as to enable students to connect relevant concepts, specific skills, and strategies with the current problem categories and form a more coordinated understanding of this problem.

4.4.3. Embedded Design Course. According to the arrangement of time, teachers provide problem situations on the network teaching platform in advance before class, so that the learning group can discuss problems after class. In the process of classroom teaching, let students realize specific operation, analyze problems onsite, design algorithms, and write programs. Create a better classroom discussion atmosphere and improve students' interest in classroom learning. Students adopt collaborative learning to carry out after-school learning. The group includes 4 members to explore and determine problem solutions. The performance of group members in cooperation and their contributions to the group are important reference factors in the formation of group mutual evaluation. Teachers arrange a fixed time for counseling every week after class, participate in group discussions, and timely correct and guide students to solve problems [20].

5. Conclusion

The teaching content selected in this experiment is practical flash animation production, focusing on cultivating students' practical ability. Eighth-grade students have mature logical thinking ability. Influenced by the traditional teaching mode, students are used to learning knowledge and solving problems through teachers' explanation. Therefore, for the students in the experimental class, in the "asking questions" stage, teachers guide students' learning initiative by creating real problem situations. In the "problem analysis" stage, under the guidance of teachers, students analyze the problems to be solved, stimulate students' desire to explore and cultivate students' analytical ability. In the "problem solving" stage, students can find ways to solve problems in autonomous learning and cooperative communication according to the analyzed problems, so as to create works. And in the process of solving this problem, students will encounter new problems, and the PBL teaching mode provides students with freedom of communication and discussion and learning resources. Learners improve their problem-solving ability in the process of discovering, analyzing, and solving problems. In the control class, students learn new knowledge through teachers' teaching and understanding of teaching materials. They only master the production methods and skills, and lack the process of hands-on exploration. According to the analysis of the data obtained in this experiment, it can be seen that after the experiment, the students in the experimental class are significantly more interested in information technology than the students in the control class. Through PBL teaching mode, teachers can create a good learning atmosphere for students and provide rich learning resources. Therefore, the PBL teaching mode in junior middle school information technology curriculum is conducive to improving students' interest in learning.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by Hunan First Normal University. This work is supported by the 2019 National Social Science Foundation Project: the development and research of the moral education function of the research base of colleges and universities in the new era (fund no. 19bks149).

References

[1] Y. Yu, T. Hao, and H. Zhang, "Research and practice of hybrid teaching mode of "film and television production technology and art" course based on obe concept," *Creative Education*, vol. 12, no. 09, pp. 2066–2073, 2021.

[2] N. Chang, Z. Wang, and S. H. Hsu, "A comparison of the learning outcomes for a pbl-based information literacy course in three different innovative teaching environments," *Libri - International Journal of Libraries and Information Services*, vol. 70, no. 3, pp. 213–225, 2020.

[3] K. Royle, "What's good what's bad? conceptualising teaching and learning methods as technologies using actor network theory in the context of palestinian higher education," *Postdigital Science and Education*, vol. 3, no. 1, pp. 120–143, 2021.

[4] G. Li, W. Zhang, and S. Lou, "The exploration of pbl mixed teaching mode in secondary vocational classes," *Journal of Physics: Conference Series*, vol. 1976, no. 1, p. 012076, 2021.

[5] Y. Hu, "Research on the implementation of happy teaching in film and television professional courses of college," *Creative Education*, vol. 11, no. 03, pp. 321–327, 2020.

[6] R. Sanjeev, R. V. Vrath, P. Sreedhar, A. Ambala, and V. Jagannadham, "Carbon credits: project based learning to enhance conceptual understanding," *Resonance*, vol. 27, no. 4, pp. 667–671, 2022.

[7] C. Li, H. Gu, and Q. Zhang, "Research on the training mode of film and television cultural talents under the background of barrier free information development," *Art and Design Review*, vol. 10, no. 01, pp. 136–143, 2022.

[8] J. Zhang, Z. Lin, W. Tang, and X. Wang, "Research on teaching mode of introduction to mechanical engineering course," *Journal of Physics: Conference Series*, vol. 1707, no. 1, p. 012021, 2020.

[9] Z. Hongqiang, L. Li, S. Weiqi et al., "Research and practice of course mixed teaching mode based on mooc and flipped classroom," *E3S Web of Conferences*, vol. 236, no. 1, p. 05054, 2021.

[10] H. Cong, "Personalized recommendation of film and television culture based on an intelligent classification algorithm," *Personal and Ubiquitous Computing*, vol. 24, no. 2, pp. 165–176, 2020.

[11] Z. Guangyao and W. Yandong, "The practice and theory of China's national image in Hong Kong region a case study of film and television," *E3S Web of Conferences*, vol. 236, no. 1, p. 05058, 2021.

[12] F. Liu, "Era of big data is based on the study of physical education teaching mode in mooc," *Journal of Physics: Conference Series*, vol. 1744, no. 3, p. 032008, 2021.

[13] W. Sun, "Research on the English mixed listening and speaking teaching mode based on the new mobile media on internet," *Journal of Physics: Conference Series*, vol. 1648, no. 3, p. 032079, 2020.

[14] Z. Guo and Z. Xiao, "Research on online calibration of lidar and camera for intelligent connected vehicles based on depth-edge matching," *Nonlinear Engineering*, vol. 10, no. 1, pp. 469–476, 2021.

[15] M. Yu, "Film and television culture dissemination based on zynq embedded digital image processing," *Microprocessors and Microsystems*, vol. 82, no. 9, p. 103921, 2021.

[16] B. Qi, "Teaching mode of Chinese and western art history course based on big data," *Journal of Physics: Conference Series*, vol. 1852, no. 2, p. 022021, 2021.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 3949597, 12 pages
https://doi.org/10.1155/2022/3949597

Research Article

A Multi-Level Fuzzy Comprehensive Evaluation Method for Knowledge Transfer Efficiency in Innovation Cluster

Xiaoli Zhang[1,2] and Rui Xu[3]

[1]*Anhui Technical College of Mechanical and Electrical Engineering, Wuhu 241000, China*
[2]*International College, National Institute of Development Administration, Bangkok 10240, Thailand*
[3]*Anhui Business College, Wuhu 241002, China*

Correspondence should be addressed to Rui Xu; 0125000366@ahcme.edu.cn

Received 4 May 2022; Revised 18 June 2022; Accepted 30 June 2022; Published 31 July 2022

Academic Editor: Xuefeng Zhang

Knowledge transfer is the essential requirement for innovation clusters to carry out collaborative innovation, and it is a necessary process for innovation clusters to realize the knowledge value enhancement. The evaluation of knowledge transfer efficiency in innovation cluster can effectively reflect the knowledge gap, environment, and whether it is effectively coordinated among members of the innovation cluster. In order to evaluate the knowledge transfer efficiency in innovation clusters more scientifically and accurately, this paper analyzes the main factors affecting the efficiency of knowledge transfer based on the characteristics of innovation clusters and establishes a multi-level comprehensive evaluation system including knowledge transfer subject features, knowledge content features, knowledge transfer environment, and knowledge transfer coordination behavior. Furthermore, a set of AHP-Entropy index weight determination method and multi-level fuzzy comprehensive evaluation method are proposed to evaluate the knowledge transfer efficiency in innovation cluster. The results of the case study show that the evaluation system and method of knowledge transfer efficiency established in this paper are effective, and they can provide valuable reference for the management of knowledge transfer activities in innovation clusters.

1. Introduction

With the continuous development of information technology, the increasingly fierce market competition makes the internal and external environment of enterprise greater complexity and dynamism, and the enterprise boundary becomes increasing more blurred and flexible [1]. In the above context, enterprises must break through the original organizational boundaries and scale restrictions in the utilization and management of intellectual capital such as information and knowledge, and break down interorganizational information and knowledge barriers by means of extensive and in-depth knowledge collaboration with external organizations and enterprises [2]. Innovation clusters have emerged from this development context. In a specific regional scope or industrial field, innovation clusters are formed on the premise of effective aggregation of human resources, information resources, and knowledge resources [3, 4].

In the process of innovation clusters collaboration, the innovation cluster subjects realize transfer, sharing and innovation of knowledge through knowledge transfer, and use it to solve the problems encountered in engineering practice, and finally achieve the collaborative innovation [5, 6]. In this process, how to realize effective knowledge transfer among innovation subjects and improve knowledge transfer efficiency is one of the most important knowledge management issues of innovation clusters. Therefore, it is of great theoretical and practical significance to investigate the knowledge transfer efficiency in the innovation cluster. The evaluation of knowledge transfer efficiency in innovation cluster is a complex decision problem, which needs to consider numerous knowledge transfer efficiency influencing factors and indicators. For the influence of knowledge characteristics on knowledge transfer efficiency, Zander and Kogut [7] conducted a pioneering study, and they stated that the explicit degree of knowledge determines the efficiency of knowledge transfer to a large extent. Simonin [8] further sublimated Zander and Kogut's findings by proposing multiple internal and external factors of knowledge transfer. Among them, the internal factors are the knowledge transfer subject and the own characteristic attributes of knowledge, and the external factors are the relevant environmental factors of knowledge transfer. Quigley et al. [9] identified team-oriented incentives, member self-efficacy, and self-goal setting and trusting relationship among members as the important factors affecting knowledge sharing efficiency. Luo et al. [10] investigated the co-evolution of complex networks and knowledge sharing based on a multi-intelligence model, and their simulation results showed that factors such as inter-subject knowledge distance, close association, and network cohesiveness had important effects on knowledge sharing efficiency. Regarding the influence of open innovation network features on knowledge transfer efficiency, Su et al. [11] proposed a new measurement method for knowledge transfer efficiency of open innovation network using the weighted complex network theory.

On the other hand, the stream of evaluation method of knowledge transfer efficiency is also very significant. Chen et al. [12] proposed an evaluation system of inter-enterprise knowledge sharing efficiency from two levels of knowledge authorization scope and depth. Wu and Pang [13] evaluated the static knowledge exchange efficiency of academic communities based on the SBM model, and investigated the dynamic evolution of knowledge exchange in virtual academic communities. Zhu et al. [14] constructed an evaluation system of knowledge flow efficiency in practice communities from four aspects: knowledge flow level, knowledge innovation level, knowledge application level, and knowledge perception level. Cowan and Jonard [15], Yang et al. [16], and Li et al. [17] used the average knowledge stock, variation coefficient of knowledge stock, and knowledge diffusion rate to evaluate the knowledge sharing efficiency in the complex network contexts. Regarding the evaluation of knowledge transfer efficiency in the context of innovation clusters, Gai and Dong [18] constructed an evaluation index system of knowledge management efficiency and measured it using the super-efficiency DEA method. From the perspective of knowledge potential difference, Li et al. [19] constructed an evaluation index system of knowledge transfer performance of manufacturing industry innovation clusters, and used AHP-fuzzy set method to comprehensively evaluate the knowledge transfer performance.

Based on the above research, it can be easily found that the current research on knowledge transfer mainly focuses on knowledge transfer models, knowledge transfer influencing factors, and quantitative evaluation methods, while there is a lack of systematic and in-depth research on knowledge transfer efficiency evaluation in the context of innovation cluster. Therefore, this paper intends to conduct an in-depth study on knowledge transfer efficiency evaluation in innovation clusters, systematically analyze the knowledge transfer efficiency evaluation index system under innovation clusters collaboration, and propose the corresponding quantitative evaluation method of knowledge transfer efficiency, thus providing theoretical basis and decision support for innovation clusters and cluster enterprises to effectively improve knowledge transfer efficiency.

2. Evaluation Index System

The selection of knowledge transfer efficiency evaluation indexes under innovation clusters collaboration is a complex systemic issue, which requires the adoption of scientific and reasonable selection principles and methods to select the most important knowledge transfer efficiency evaluation indexes within a reasonable range of evaluation accuracy and cost [20, 21]. Knowledge transfer is the process of knowledge subjects exchanging, acquiring, learning, and utilizing knowledge to knowledge sources through certain transfer environment or medium, and then realizing knowledge increment and knowledge innovation. Szulanski [22] believed that the influencing factors of knowledge transfer performance should contain five elements, including knowledge transfer source, knowledge transfer recipient, knowledge transfer content, knowledge transfer path, and knowledge transfer scenario. Hu [23] proposed that knowledge sharing evaluation indicators in network organizations should be analyzed from four dimensions: cognitive gap among network members, knowledge sharing environment, knowledge sharing coordination behavior, and knowledge sharing results. Drawing on the above research, this paper constructs the evaluation index system of knowledge transfer efficiency in innovation clusters from four dimensions, including knowledge transfer subject features, knowledge content features, knowledge transfer environment, and knowledge transfer coordination behavior. The details of the four dimensions are as follows:

In the process of innovation clusters collaboration, the knowledge transfer subject refers to the knowledge sender and the knowledge receiver involved in knowledge transfer activities, and knowledge transfer is the process of knowledge exchange and interaction between the knowledge sender and the knowledge receiver [24]. For specific knowledge, knowledge senders and knowledge receivers can switch to each other. In innovation cluster, the difference in the types and stocks of knowledge possessed by knowledge transfer subjects leads to knowledge potential differences. Knowledge potential difference is the original driving force of knowledge transfer [25]. Knowledge transfer willingness of knowledge subjects is an important factor for smooth knowledge transfer, and knowledge transfer willingness has a significantly positive effect on knowledge transfer efficiency [26]. The stronger the knowledge transfer willingness, the more proactively, actively, and effectively the knowledge transfer subjects can communicate and share each other's knowledge resources [27, 28]. Knowledge transfer capability likewise contributes positively to knowledge transfer efficiency, which can be further subdivided into knowledge sending capability of the knowledge sender and knowledge absorbing capability of the knowledge receiver. The stronger the knowledge transfer ability of both sides of knowledge transfer makes knowledge transfer less difficult and sticky, and thus can effectively improve the efficiency of knowledge transfer [29]. On the other hand, the degree of trust and reciprocity among knowledge transfer subjects has a positive contribution to knowledge transfer efficiency. Researchers have shown that the degree of trust and reciprocity among knowledge subjects facilitates the acquisition of new information and knowledge, and reduces opportunistic behavior and free-riding behavior among subjects [30, 31]. Finally, in innovation clusters, the cluster embeddedness of knowledge transfer subjects has a positive impact on the formation of good knowledge cooperation norms among knowledge subjects, which can help knowledge subjects acquire more heterogeneous knowledge [32].

Knowledge content refers to the data, information, and knowledge exchanged and transferred between knowledge transfer subjects. Knowledge in innovation clusters can likewise be divided into two categories, that is, explicit knowledge and tacit knowledge. Explicit degree of knowledge largely determines the difficulty of knowledge transfer among knowledge subjects, and there is a significant positive correlation between the explicit degree of knowledge and knowledge transfer efficiency [33]. Systematization degree of knowledge refers to the extent of knowledge embedding in organizational processes and norms based on knowledge preservation in the organization. The higher systematization degree of knowledge indicates the higher ability of the organization to absorb and integrate knowledge, and the corresponding higher knowledge transfer efficiency among knowledge subjects [34]. On the other hand, the sources and uses of knowledge also have an important influence on knowledge transfer. The source of knowledge will determine the content of knowledge to a certain extent, and the difficulty of acquiring knowledge sources will determine the difficulty of knowledge transfer, thereby affecting the efficiency of knowledge transfer. The usage of knowledge determines the knowledge transfer subject's seeking of specific knowledge and the judgment and cognition of the value of knowledge content to a certain extent, which makes the knowledge subject carry certain purpose in the process of knowledge seeking and acquisition [29, 35].

Knowledge transfer occurs in a specific environment, and the knowledge transfer environment is an important collaborative element to realize knowledge transfer. Organizational culture is a most important environmental factor of knowledge transfer, and whether the cluster culture encourages open and deep knowledge exchange within innovation cluster has a great impact on knowledge transfer efficiency [36, 37]. Each side of the knowledge transfer subject has its own institutional and cultural background, and the compatibility and matching of cognitive structure and management system directly affect the efficiency of knowledge transfer. Similarly, the incentive mechanism of knowledge transfer activities within the cluster plays an important role in mobilizing the motivation of knowledge transfer activities and improving the performance of knowledge transfer. On its basis, the fairness of knowledge collaboration procedures and benefit distribution among knowledge transfer subjects is the institutional guarantee to ensure the deep knowledge collaboration of both parties, and it also has a significant impact on knowledge transfer efficiency [38]. An open and smooth knowledge exchange platform and diversified knowledge transfer media are important guarantees for the smooth implementation of knowledge transfer activities, which have positive effects on reducing the uncertainty and ambiguity of knowledge transfer and ensuring the quality and effect of knowledge transfer [39].

Knowledge transfer focuses on the knowledge behavior activities and interactive coordination characteristics among cluster internal enterprises, and enterprises can improve the efficiency of knowledge transfer only by conducting mutual knowledge interaction and coordination behaviors. In the process of innovation cluster collaboration, there are dynamic and complex knowledge exchange relationships among cluster enterprises, so enterprises need to apply scientific and reasonable coordination mechanisms to cope with the uncertain knowledge exchange environment. Firstly, communication between cluster enterprise managers helps enterprises better discover the strengths and weaknesses of both sides to better utilize their knowledge advantages, and then form the complementary advantages of knowledge collaboration. Therefore, communication among managers is an effective means to improve knowledge transfer efficiency [40, 41]. Secondly, since a large amount of knowledge in the process of cluster collaborative innovation is tacit knowledge, it requires in-depth knowledge exchange and communication between employees from different enterprises. Only through extensive and close communication among employees in common cooperative tasks can the system of knowledge exchange and transfer be put into practice, and a good atmosphere of knowledge transfer and sharing can be created, thus improving knowledge transfer efficiency, especially the transfer efficiency of tacit knowledge [42]. Finally, due to the problems of insufficient and asymmetric information between the two sides of cooperative enterprises in cluster collaboration, there are cognitive biases about the knowledge transfer problem in cooperation, which requires the constraint and adjustment of the cooperation contract to realize the continuous improvement and perfection of knowledge transfer behavior, thereby achieving the purpose of improving the efficiency of knowledge transfer under cluster collaboration [43].

Based on the above comprehensive analysis, the evaluation index system of knowledge transfer efficiency in innovation cluster is constructed as shown in Table 1.

3. Knowledge Transfer Efficiency Evaluation Methods

On the premise that the evaluation index system of knowledge transfer efficiency has been determined, the validity and accuracy of the evaluation results of knowledge transfer efficiency mainly depend on two major factors: one is the determination of each evaluation index weight of knowledge transfer efficiency, and the other is the selection of comprehensive evaluation methods.

Table 1: Knowledge transfer efficiency evaluation index system in innovation cluster.

Level 1 indicator	Level 2 indicators	Level 3 indicators
Knowledge transfer efficiency A	Knowledge transfer subject features B_1	Knowledge potential difference among knowledge transfer subjects $\mathbf{C_{11}}$
		Knowledge transfer willingness $\mathbf{C_{12}}$
		Knowledge transfer capability $\mathbf{C_{13}}$
		Trust degree among knowledge transfer subjects $\mathbf{C_{14}}$
		Reciprocity degree among knowledge transfer subjects $\mathbf{C_{15}}$
		Cluster embeddedness of knowledge transfer subjects $\mathbf{C_{16}}$
	Knowledge content features B_2	Explicit degree of knowledge $\mathbf{C_{21}}$
		Systematization degree of knowledge $\mathbf{C_{22}}$
		Sources of knowledge $\mathbf{C_{23}}$
		Usage of knowledge $\mathbf{C_{24}}$
	Knowledge transfer environment B_3	Knowledge exchange culture within the cluster $\mathbf{C_{31}}$
		Institutional compatibility among knowledge transfer subjects $\mathbf{C_{32}}$
		Fairness of collaboration process and benefit distribution $\mathbf{C_{33}}$
		Knowledge exchange platform $\mathbf{C_{34}}$
		Knowledge transfer media and channels $\mathbf{C_{35}}$
	Knowledge transfer coordination behavior B_4	Communication between cluster enterprise managers $\mathbf{C_{41}}$
		Communication between cluster enterprise employees $\mathbf{C_{42}}$
		Design and adjustment of cooperation contract $\mathbf{C_{43}}$

3.1. Determination of Index Weights. In the process of knowledge transfer efficiency evaluation, the determination of index weights is the most important step to ensure that knowledge transfer efficiency evaluation can be successful. Currently, the methods of determining index weights can be divided into two main categories [44, 45]: one is the subjective weighting methods, including Delphi method, AHP (Analytic Hierarchy Process), and fuzzy comprehensive evaluation method. The other is the objective weighting method, including maximum deviation method, mean difference method, and threshold method. Both objective and subjective weighting methods have their advantages and disadvantages and areas of application. The subjective weighting method has advantages to evaluate the subjective preference of the subject, but because there are often differences in the subjective judgment of individuals, the index weights confirmed by this type of method lack smoothness. In contrast, because the weights confirmed by the objective weighting method can only have small amount of information based on the main data of the indicators, sometimes there is a problem that the indicator weights are different from the true importance level of the indicators. Another drawback is that the confirmation of the weights suffers from the randomness of the sample data.

Based on the above analysis, this paper intends to use the AHP and entropy weight method jointly with each other to determine the index weights of knowledge transfer efficiency, which is an objective and subjective composite method. The evaluation index system of knowledge transfer efficiency in innovation cluster has the characteristics of multi-objective and multi-level, and the evaluation elements have great fuzzy and qualitative characteristics. The alone application of AHP has the following shortcomings [46]: first, as a subjective weighting method, the AHP method often determines the weight values according to the appraiser's subjective judgment when constructing the judgment matrix, so the appraisal results may vary greatly due to the appraiser's experience and perception; second, the AHP method ignores the situation that it is assumed that all evaluators think that a certain indicator is critical, and the value given to this indicator is relatively similar, and the weight value given by the AHP method is also relatively high, which makes the recognition of this indicator greatly reduced, and finally leads to the decrease of the effectiveness of this evaluation indicator. Tosolve the above problems of AHP method, this paper introduces the entropy weight method, an objective weighting method, to amend the AHP method, reduce the subjectivity of the weights determined by the AHP method, and lower the weights of those indicators with low recognition power, so that the subjective and objective weighting methods can be combined with each other to improvethe rationality and effectiveness of the evaluation index weights.

3.1.1. Overview of AHP Method

(1) Constructing the hierarchical structure of evaluation index system: on the premise of comprehensively mastering the knowledge transfer efficiency evaluation index system, AHP method firstly analyzes the structure of the index system and the relationship between indicators at each level, and divides the index system into several levels, mainly including the target level, the standard level, and the indicator level.

(2) Constructing pairwise comparison decision matrix: when constructing the pairwise comparison judgment matrix, the evaluator firstly needs to assign a certain scale value to the relative importance of each evaluation index. As shown in Table 2, this paper applies a scale of 1–7. The results obtained from the pairwise importance comparisons between the elements as shown in Table 3.

The judgment matrix $A = (a_{ij})_{m\times n}$ has the following properties:

Table 2: The definition of judgment matrix.

1	Represents a comparison of 2 indicators that have consistent importance
3	Represents a comparison of 2 indicators, where 1 indicator is more important than the other
5	Represents a comparison of 2 indicators, where 1 indicator is much more important than the other
7	Represents a comparison of 2 indicators, where 1 indicator is extremely important than the other
2, 4, 6	The median value between the above 2 intervening judgment values

Table 3: The judgment matrix.

U	A_1	A_2	……	A_n
A_1	α_{11}	α_{12}	……	α_{1n}
A_2	α_{21}	α_{22}	……	α_{1n}
……	……	……	……	
A_n	α_{n1}	α_{n2}	……	α_{nn}

$$a_{ij}>0 \quad a_{ij} = \frac{1}{a_{ji}} a_{ij} \cdot a_{jk} = a_{ik}. \tag{1}$$

(3) Calculating the relative importance of evaluation index: the relative importance vector $W = (W_1, W_2, \cdots, W_n)^T$ of the evaluation index is calculated by the following:

(i) Arithmetic average method:

$$W_i = \frac{1}{n}\sum_{j=1}^{n}\frac{a_{ij}}{\sum_{k=1}^{n} a_{kj}}, \quad i = 1, 2, \cdots, n. \tag{2}$$

(ii) Geometric average method:

$$W_i = \frac{\left(\prod_{j=1}^{n} a_{ij}\right)^{1/n}}{\sum_{i=1}^{n}\left(\prod_{j=1}^{n} a_{ij}\right)^{1/n}}, \quad i = 1, 2, \ldots, n. \tag{3}$$

(4) Coherence test:

The coherence index $C.I.$ is calculated based on the following equations:

$$C.I. = \frac{\lambda_{\max} - n}{(n-1)}, \tag{4}$$

$$\lambda_{\max} \approx \frac{1}{n}\sum_{i=1}^{n}\frac{(AW)_i}{W_i} = \frac{1}{n}\sum_{i=1}^{n}\frac{\sum_{j=1}^{n} a_{ij}W_j}{W_i}. \tag{5}$$

Then, it needs to find the corresponding average random coherence index R.I. Table 4 shows the average random coherence indicators obtained by calculating the 1 ~ 14 rank positive reciprocal matrices 1000 times.

The average random coherence indicator R.I. is the mean value of the coherence indicators of the same rank random judgment matrix, and the introduction of R.I. can circumvent the drawback that the coherence judgment indicator increases significantly with the increase of n to some extent.

Finally, the coherence ratio C.R. = /R.I.ˇ0.1, it passes the coherence test, at which time the obtained judgment result can be considered reasonable.

3.1.2. Overview of Entropy Weight Method. The concept of entropy is originated from thermodynamics and later introduced to information theory by Shannon. According to the definition and principle of entropy, the entropy value can be used as a measure of the amount of effective information provided by a system and represents the degree of disorder of a system. The entropy weight method is an objective weighting method that combines qualitative and quantitative analysis. The entropy weight method determines the indicator weights based on the amount of information that each indicator conveys to the decision maker [47]. For the evaluation problem, with m evaluation objects and n evaluation indicators, the original evaluation matrix $X = (x_{ij})_{m\times n}$ is obtained, and x_{ij} denotes the value of the j evaluation indicator of the i evaluation object. So, the entropy value of the j evaluation index x_j can be denoted as follows:

$$\eta_j = -\frac{1}{\ln m}\sum_{i=1}^{m}\kappa_{ij}\ln\kappa_{ij}, \tag{6}$$

where $\kappa_{ij} = x_{ij}/\sum_{i=1}^{m} x_{ij}$, κ_{ij} denotes the proportion of the i participant under the j indicator. According to the definition and principle of entropy, when the entropy value of an indicator is smaller, it means that the less effective information provided by the indicator, indicating the smaller the function of the indicator in the system evaluation, and the smaller its weight accordingly; on the contrary, the larger the entropy value, the more effective information provided by the indicator, the larger the function in the comprehensive evaluation, and the larger its weight. The correction process of the entropy to the AHP method is shown as follows:

(1) A dimensionless treatment of the X matrix yields the matrix $Y = (y_{ij})_{m\times n}$, that is,

$$y_{ij} = \frac{x_{ij}}{\left[\sum_{i=1}^{m} {x_{ij}}^2\right]^{1/2}}, \qquad i = 1, 2, 3, \ldots, m; \quad j = 1, 2, 3, \ldots, n. \tag{7}$$

(2) Calculate κ_{ij}, which is the weight of the j indicator of the i participant.

$$\kappa_{ij} = \frac{y_{ij}}{\sum_{i=1}^{m} y_{ij}}. \tag{8}$$

(3) Calculate the entropy value η_j of the j indicator.

$$\eta_j = -\frac{1}{\ln m}\sum_{i=1}^{m}\kappa_{ij}\ln\kappa_{ij}, \quad (j = 1, 2, 3 \ldots, n), \tag{9}$$

where $0 \le \eta_j \le 1$.

(4) Calculate the difference coefficient χ_j of the j indicator.

Table 4: The average random coherence indexes.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14
R.I.	0	0	0.52	0.89	1.12	1.26	1.36	1.41	1.46	1.49	1.52	1.54	1.56	1.58

$$\chi_j = \left(1 - \eta_j\right). \tag{10}$$

For the j indicator, the greater the χ_j, the greater the role of the indicator for evaluation; conversely, the smaller the χ_j, the smaller the role of the indicator for evaluation.

(5) Calculate the weight w_j of the j indicator.

$$w_j = \frac{\chi_j}{\sum_{j=1}^{n} \chi_j}. \tag{11}$$

3.1.3. The AHP-Entropy Weight Method. For the subjective weights of each index obtained by AHP, the weight w_j obtained by entropy weight method is used to adjust the existing weight.

$$w_j' = w_j' \cdot w_j, \tag{12}$$

where w_j' represents the index weight value obtained by AHP method.

The final adjusted weights W_j are obtained by normalizing w_j''.

$$W_j = \frac{w_j''}{\sum_{j=1}^{n} w_j''}, \quad j = 1, 2, 3 \ldots n. \tag{13}$$

3.2. Fuzzy Comprehensive Evaluation Method of Knowledge Transfer Efficiency. Knowledge transfer efficiency evaluation in innovation cluster is a system engineering of great complexity and fuzziness, which contains many issues and factors with fuzziness and difficult to quantify accurately, and it is often difficult to obtain complete and sufficient data and information about the evaluation process. According to this fuzzy characteristic of knowledge transfer efficiency evaluation, this paper intends to apply the fuzzy comprehensive evaluation method to comprehensively evaluate the knowledge transfer efficiency in innovation cluster [48]. In addition, when evaluators evaluate knowledge transfer efficiency under innovation cluster collaboration, the rubric used is often somewhat ambiguous. Therefore, this paper proposes the fuzzy comprehensive evaluation method of knowledge transfer efficiency in innovation cluster: based on determining knowledge transfer efficiency evaluation index weights by using AHP-entropy weight method, establishing fuzzy evaluation matrix, and finally conducting comprehensive evaluation of knowledge transfer efficiency under innovation cluster collaboration by using fuzzy comprehensive evaluation method.

The calculation process of the fuzzy comprehensive evaluation method is as follows.

3.2.1. Determining the Evaluation First-Level Model

(1) Establishing the factor set of the evaluation object U: the factor set is the set of evaluation indicators, set as $U = (U_1, U_2 \cdots U_n)$.

(2) Determining the evaluation set V: the evaluation set is the set of evaluation levels given by the evaluation subject, set as $V = (V_1, V_2, \cdots V_q)$.

In general, the rubric level number q is taken as an integer between $[3, 7]$. If q is too large, then it is difficult to describe the evaluation level and to determine the grade of the rubric; if q is too small, then the quality requirements of fuzzy comprehensive evaluation cannot be achieved. Usually q is taken as an odd number, so that there is an intermediate grade, which is easy to distinguish the grade of the evaluation object. The specific rank can be determined by the evaluation experts according to the content and characteristics of the evaluation object, and described in appropriate language.

(3) Establishing the fuzzy mapping relationship between the factor set and the evaluation set: establish a fuzzy mapping from U to V, that is,

$$\begin{aligned} f&: U \longrightarrow F(V), \\ u_i &= f(u_i) \\ &= m_i \\ &= \left(m_{i1}, m_{i2} \ldots m_{iq}\right). \end{aligned} \tag{14}$$

Then, the single-factor judgment matrix M is obtained.

$$M = \begin{pmatrix} m_{11} & \cdots & m_{1q} \\ \vdots & \ddots & \vdots \\ m_{n1} & \cdots & m_{nq} \end{pmatrix}, \tag{15}$$

where m_{ij} is the affiliation of factor U_i in U corresponding to the rank V_j in V. m_{ij} is the number of people choosing levelv_i for the i indicator/ total number of people involved in the evaluation

(4) Determining the evaluation factor weight vector W: since each factor in the evaluation factor set U has different importance to the evaluation object, it is necessary to assign different weights to each factor, that is, $W = (w_1, w_2, \ldots, w_n)$.

The regulations are as follows:

$$\sum_{i=1}^{n} w_1 = 1, \quad w_i \geq 0, (i = 1, 2, 3, \ldots, n). \tag{16}$$

(5) Selecting composite operator for comprehensive evaluation.

The basic model of the fuzzy comprehensive evaluation method can be expressed by the formula as follows:

$$R = W \cdot M. \tag{17}$$

In the basic formula $R = W \cdot M$ of the fuzzy comprehensive evaluation model, the combination of W and M has a very important influence on the final evaluation result, so the selection of fuzzy composite operator "$\cdot$" is very important. The composite operators frequently used in fuzzy comprehensive evaluation include: the principal factor determinant type, the principal factor prominent type, the unbalanced average type, and the weighted average type. Knowledge transfer efficiency evaluation under innovation cluster collaboration is a comprehensive evaluation problem with multiple indicators and multiple levels, which requires a balanced consideration of the relative importance of each factor and its influence on the overall evaluation results. Therefore, based on the above analysis, it is appropriate to choose the weighted average type of composite operator in this paper.

3.2.2. Multi-Level Fuzzy Comprehensive Evaluation Model. Based on the comprehensive evaluation of the lower-level factors, the evaluation results of the lower-level factors are used to comprehensively evaluate the higher-level factors.

The evaluation factor set U is divided into P subsets, which is denoted as $U = (U_1, U_2, \ldots, U_p)$, and the i subset is defined as $U_i = (U_{i1}, U_{i2}, \ldots, U_{ik})$, $(i = 1, 2, 3, \ldots, p)$.

For each subset U_i, the comprehensive evaluation is conducted by the first-level model separately. Suppose the corresponding weight set of U_i is W_i and the corresponding fuzzy evaluation matrix of U_i is M_i. There are

$$\begin{aligned} R_i &= W_i \cdot M_i \\ &= \left(r_{i1}, r_{i2}, \ldots, r_{\text{im}}\right)(i = 1, 2, 3, \ldots, p). \end{aligned} \tag{18}$$

Suppose R_i , which is obtained from the evaluation of each subset U_i in the factor set U, as P single-level evaluations in U . Then, suppose the weight assignment set is W , so the total fuzzy evaluation matrix is

$$\begin{aligned} R &= \begin{bmatrix} R_1 \\ R_2 \\ \cdots \\ R_p \end{bmatrix} \\ &= \left(r_{ij}\right)_{pm}. \end{aligned} \tag{19}$$

Finally, the second-level evaluation results are

$$R = W \cdot M. \tag{20}$$

4. Case Study

This section takes mobile phone industry innovation cluster of Chongqing, China as the research object to evaluate its knowledge transfer efficiency under the cluster innovation collaboration. At present, there are 116 mobile-phone enterprises and 138 mobile-phone supporting enterprises in mobile phone industry innovation cluster of Chongqing, and the output value of the mobile-phone industry exceeds 100 billion yuan. Through data collection and on-site research on mobile phone manufacturer enterprises in Chongqing mobile phone industry cluster, this paper collects the first-hand data and information of knowledge transfer efficiency evaluation. Based on the index system and comprehensive evaluation method of knowledge transfer efficiency under innovation cluster collaboration proposed in this paper, the process of evaluating and analyzing the knowledge transfer efficiency of the innovation cluster of Chongqing mobile phone industry is shown as follows.

4.1. Application of AHP Method to Determine Subjective Weights. Based on the evaluation index system of knowledge transfer efficiency, senior leaders of backbone enterprises of Chongqing mobile phone industry cluster (7 persons) and experts in innovation cluster and knowledge management (3 persons) are invited to make pairwise comparison of the importance of evaluation indexes at the same level, judge the relative importance of each index using the Delphi method, and then evaluate the relative importance of the indexes, and establish a judgment matrix at all levels from high-level indexes to low-level indexes as shown below:

Layer A-Layer B (First level judgment matrix):

$$A = \begin{bmatrix} A & B_1 & B_2 & B_3 & B_4 \\ B_1 & 1 & 3 & 2 & 4 \\ B_2 & \frac{1}{3} & 1 & \frac{1}{2} & 2 \\ B_3 & \frac{1}{2} & \frac{1}{2} & 1 & 1 \\ B_4 & \frac{1}{4} & \frac{1}{2} & 1 & 1 \end{bmatrix}. \tag{21}$$

Layer B-Layer C (Second level judgment matrix):

$$
\begin{aligned}
B_1 &= \begin{bmatrix} B_1 & C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{11} & 1 & 1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{3} & 2 \\ C_{12} & 1 & 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{2} & 3 \\ C_{13} & 2 & 2 & 1 & 2 & 2 & 5 \\ C_{14} & 2 & 3 & \frac{1}{2} & 1 & 2 & 4 \\ C_{15} & 3 & 2 & \frac{1}{4} & \frac{1}{2} & 1 & 6 \\ C_{16} & \frac{1}{2} & \frac{1}{3} & \frac{1}{5} & \frac{1}{4} & \frac{1}{6} & 1 \end{bmatrix}, \\
B_2 &= \begin{bmatrix} B_2 & C_{21} & C_{22} & C_{23} & C_{24} \\ C_{21} & 1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{2} \\ C_{22} & 2 & 1 & \frac{1}{3} & 3 \\ C_{23} & 4 & 3 & 1 & 4 \\ C_{24} & 2 & \frac{1}{3} & 4 & 1 \end{bmatrix}, \\
B_3 &= \begin{bmatrix} B_3 & C_{41} & C_{42} & C_{43} & C_{44} & C_{45} \\ C_{41} & 1 & 2 & \frac{1}{3} & \frac{1}{2} & \frac{1}{2} \\ C_{42} & \frac{1}{2} & 1 & \frac{1}{4} & 2 & \frac{1}{2} \\ C_{43} & 3 & 4 & 1 & 2 & 2 \\ C_{44} & 2 & \frac{1}{2} & \frac{1}{2} & 1 & \frac{1}{2} \\ C_{45} & 2 & 2 & \frac{1}{2} & 2 & 1 \end{bmatrix}, \\
B_4 &= \begin{bmatrix} B_4 & C_{41} & C_{42} & C_{43} \\ C_{41} & 1 & 2 & 1 \\ C_{42} & \frac{1}{2} & 1 & \frac{1}{2} \\ C_{43} & 1 & 2 & 1 \end{bmatrix}.
\end{aligned}
\tag{22}
$$

Using AHP, the weight sets of first-level indicators can be obtained as

Wu = {0.470, 0.172, 0.219, 0.139} ∘

Further, second-level indicator weight sets can be obtained as follows:

Wu1 = {0.104, 0.112, 0.242, 0.239, 0.204, 0.049},
Wu2 = {0.100, 0.248, 0.531, 0121},
Wu3 = {0.125, 0.118, 0.387, 0.135, 0.235},
Wu4 = {0.413, 0.260, 0.327}.

The above judgment matrix is tested to meet the consistency requirements, thus ensuring the reliability of the weight vector results.

4.2. Application of Entropy Weight Method to Determine Objective Weights. This paper selects the mobile phone industry clusters in other four provinces and cities which are similar to mobile phone industry cluster of Chongqing, and express them as A, B, C, and D, respectively. Several experts in the related field are organized to set up an expert panel, and the expert panel scores the evaluation indicator system with the score range of 1–5. The higher the score indicates, the higher the development level of a specific innovation cluster on a certain indicator. Finally, the scoring results of each expert on the indicator system are combined to obtain the raw data in Table 5.

According to AHP method, the objective weights of the first-level evaluation indicators are first determined, and the raw data of the first-level evaluation indicators can be seen from Table 6.

The weight p_{ij} of the j indicator of the i innovation cluster is calculated, and the weight table P_{ij} is obtained as shown in Table 7.

Furthermore, the entropy value, variation coefficient, and objective weight of each evaluation indicator are obtained in Table 8.

From the above table, the objective weights of the first-level evaluation indicators wu = {0.417,0.194,0.250,0.139}.

Further, the objective weights of the second-level evaluation indicators can be obtained as follows:

Wu1 = {0.394, 0.081, 0.212, 0.252, 0.061},
Wu2 = {0.241, 0.057, 0.231, 0.161, 0.069, 0.241},
Wu3 = {0.049, 0.138, 0.317, 0.317, 0.114, 0.065},
Wu4 = {0.259, 0.309, 0.061, 0.272, 0.099}.

4.3. Using AHP-Entropy Weight Method to Calculate the Comprehensive Weight. The comprehensive weights of the first-level indicators can be obtained by (13).

$$\overline{W}_u = \{0.646, 0.109, 0.180, 0.065\}. \tag{23}$$

Repeating the above steps, the comprehensive indicator weights of the second-level evaluation indicators can be obtained as follows:

Table 5: The raw data of expert scoring.

			A		B		C		D	
Knowledge transfer efficiency evaluation A	$\mathbf{B_1}$	$\mathbf{C_{11}}$	3	3	4	3	3	2	3	3
		$\mathbf{C_{12}}$		3		3		3		2
		$\mathbf{C_{13}}$		5		4		4		3
		$\mathbf{C_{14}}$		4		3		4		4
		$\mathbf{C_{15}}$		5		4		5		4
		$\mathbf{C_{16}}$		2		2		3		2
	$\mathbf{B_2}$	$\mathbf{C_{21}}$	3	3	4	3	3	4	4	2
		$\mathbf{C_{22}}$		4		2		4		3
		$\mathbf{C_{23}}$		4		4		3		3
		$\mathbf{C_{24}}$		4		3		5		5
	$\mathbf{B_3}$	$\mathbf{C_{31}}$	5	3	4	2	4	3	5	4
		$\mathbf{C_{32}}$		3		4		4		3
		$\mathbf{C_{33}}$		3		3		4		3
		$\mathbf{C_{34}}$		3		3		2		4
		$\mathbf{C_{35}}$		4		3		5		3
	$\mathbf{B_4}$	$\mathbf{C_{41}}$	3	3	2	3	3	5	3	3
		$\mathbf{C_{42}}$		2		4		2		4
		$\mathbf{C_{43}}$		3		4		3		4

Table 6: The raw data of first-level evaluation index.

	B_1	B_2	B_3	B_4
A	3	3	5	3
B	4	4	4	2
C	3	3	4	3
D	3	4	5	3

Table 7: The proportion of the j index in the i innovation cluster.

	B_1	B_2	B_3	B_4
A	0.214	0.214	0.358	0.214
B	0.286	0.286	0.286	0.142
C	0.231	0.231	0.307	0.231
D	0.200	0.267	0.333	0.200

Table 8: The entropy value, variation coefficient, and weight of evaluation index.

	B_1	B_2	B_3	B_4
Entropy value	0.985	0.993	0.991	0.995
Variation coefficient	0.015	0.007	0.009	0.005
Objective weight	0.417	0.194	0.250	0.139

$$\begin{aligned}\overline{W}_{u1} &= \{0.110, 0.125, 0.280, 0.225, 0.198, 0.062\},\\ \overline{W}_{u2} &= \{0.110, 0.236, 0.527, 0.127\},\\ \overline{W}_{u3} &= \{0.139, 0.117, 0.368, 0.142, 0.234\},\\ \overline{W}_{u4} &= \{0.436, 0.217, 0.337\}.\end{aligned} \tag{24}$$

4.4. Multi-Level Fuzzy Comprehensive Evaluation of Knowledge Transfer Efficiency. Based on the multi-level fuzzy comprehensive evaluation method, this paper determines the first-level index set as $U = (U_1, U_2, U_3, U_4)$ and the second-level index set as $U_1 = (U_{11}, U_{12}, U_{13}, U_{14}, U_{15}, U_{16})$, $U_2 = (U_{21}, U_{22}, U_{23}, U_{24})$, $U_3 = (U_{31}, U_{32}, U_{33}, U_{34}, U_{35})$, $U_4 = (U_{41}, U_{42}, U_{43})$. In this paper, the judgment set is defined as $V = \{$excellent (4), good (3), qualified (2), unqualified (1)$\}$, and senior leaders of the backbone enterprises of Chongqing mobile phone industry cluster (7 people) and experts of innovation cluster and knowledge management field (3 people) are again invited to participate in the evaluation. Then, the fuzzy judgment matrix of "current situation of knowledge transfer is obtained as follows:

$$M_{u1} = \begin{bmatrix} \frac{3}{10} & \frac{4}{10} & \frac{2}{10} & \frac{1}{10} \\ \frac{3}{10} & \frac{4}{10} & \frac{3}{10} & 0 \\ \frac{4}{10} & \frac{3}{10} & \frac{3}{10} & 0 \\ \frac{4}{10} & \frac{5}{10} & \frac{1}{10} & 0 \\ \frac{3}{10} & \frac{3}{10} & \frac{2}{10} & \frac{2}{10} \\ \frac{4}{10} & \frac{3}{10} & \frac{2}{10} & \frac{1}{10} \end{bmatrix}. \tag{25}$$

The fuzzy relationship vector R_{u1} is obtained by the first-level fuzzy comprehensive evaluation as follows:

$$R_{u1} = \overline{W}_{u1} \cdot M_{u1}$$
$$= \{0.110, 0.125, 0.280, 0.225, 0.198, 0.062\}$$
$$\cdot \begin{bmatrix} \frac{3}{10} & \frac{4}{10} & \frac{2}{10} & \frac{1}{10} \\ \frac{3}{10} & \frac{4}{10} & \frac{3}{10} & 0 \\ \frac{4}{10} & \frac{3}{10} & \frac{3}{10} & 0 \\ \frac{4}{10} & \frac{5}{10} & \frac{1}{10} & 0 \\ \frac{3}{10} & \frac{3}{10} & \frac{2}{10} & \frac{2}{10} \\ \frac{4}{10} & \frac{3}{10} & \frac{2}{10} & \frac{1}{10} \end{bmatrix} \quad (26)$$
$$= \{0.303, 0.368, 0.218, 0.056\}.$$

Similarly, we can obtain

$$R_{u2} = \{0.271, 0.311, 0.342, 0.087\},$$
$$R_{u3} = \{0.299, 0.361, 0.238, 0.098\}, \quad (27)$$
$$R_{u4} = \{0.263, 0.275, 0.318, 0.133\}.$$

Therefore, the affiliation matrix R_u of the second-level indicators can be obtained as follows:

$$R_u = (R_{u1}, R_{u2}, R_{u3}, R_{u4}) = \begin{bmatrix} 0.303 & 0.368 & 0.218 & 0.056 \\ 0.271 & 0.311 & 0.342 & 0.087 \\ 0.299 & 0.361 & 0.238 & 0.098 \\ 0.263 & 0.275 & 0.318 & 0.133 \end{bmatrix}. \quad (28)$$

A second-level fuzzy comprehensive evaluation is performed to determine the fuzzy comprehensive evaluation vector for the first-level objective.

$$R = W_u \cdot R_u$$
$$= \{0.646, 0.109, 0.180, 0.065\}$$
$$\cdot \begin{bmatrix} 0.303 & 0.368 & 0.218 & 0.056 \\ 0.271 & 0.311 & 0.342 & 0.087 \\ 0.299 & 0.361 & 0.238 & 0.098 \\ 0.263 & 0.275 & 0.318 & 0.133 \end{bmatrix} \quad (29)$$
$$= \{0.297, 0.355, 0.242, 0.072\}.$$

Finally, the total evaluation score of knowledge transfer efficiency of Chongqing mobile phone industry innovation cluster is obtained:

$$S = R \cdot V^T$$
$$= \{0.296, 0.354, 0.241, 0.071\} \cdot \{4, 3, 2, 1\}^T \quad (30)$$
$$\approx 2.91.$$

Based on the above results, the knowledge transfer efficiency level of the Chongqing mobile phone industry innovation cluster belongs to the good level. It should be noted that the total knowledge transfer efficiency evaluation score only reflects knowledge transfer efficiency level of Chongqing mobile phone industry innovation cluster on a whole, but it does not mean that the innovation cluster has achieved good level in all knowledge transfer efficiency indicators. Therefore, innovation clusters should not only be satisfied with knowledge transfer efficiency evaluation scores, but also review and analyze the advantages and disadvantages of each second-level and even third-level indicator scores in the process of knowledge transfer efficiency evaluation step by step, find and summarize the excellent experiences and problematic shortcomings among them, and form knowledge transfer systems, strategies and methods from the excellent experiences, and carry out special remediation and improvement for the problems and weak links among them, so as to continuously improve the knowledge management level and innovation competitiveness of innovation clusters by taking the knowledge transfer efficiency evaluation as an opportunity.

5. Conclusion

Efficiency evaluation of knowledge transfer is an important issue of knowledge management of innovation clusters, which can enable the managers of innovation clusters to accurately recognize the knowledge transfer status. This paper evaluated the knowledge transfer efficiency in innovation clusters from the perspective of systems engineering, and comprehensively considered the characteristics of the environment, subject, relationship, and knowledge of innovation clusters. Then, a multi-level comprehensive evaluation system is constructed, which includes Knowledge transfer subject features, Knowledge content features, Knowledge transfer environment, and Knowledge transfer coordination behavior. As for the evaluation method, according to the characteristics of the combination of qualitative and quantitative evaluation indicators, an AHP-entropy method of index weight is further proposed. Moreover, considering the uncertainty and ambiguity of the evaluation of knowledge transfer efficiency in innovation clusters, a comprehensive fuzzy evaluation method of multi-level model is used to comprehensively evaluate the knowledge transfer efficiency in innovation clusters. Through the empirical analysis on the knowledge efficiency in Chongqing smartphone innovation cluster, the validity and practicability of the evaluation system and evaluation method proposed in this paper are tested. This paper can provide a feasible research reference for investigating the mechanism of knowledge transfer behavior in innovation clusters from the perspective of systematic knowledge management, and evaluating and measuring the development trend of knowledge transfer in innovation clusters. However, how to reflect the individual differences of knowledge subjects and the transfer time delay in the knowledge transfer of innovation clusters is still a problem that needs to be further solved, and it is also an important direction for further exploration.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the Anhui Humanities and Social Science Key research project (SK2019A0981).

References

[1] D. Santistevan, "Boundary-spanning coordination: insights into lateral collaboration and lateral alignment in multinational enterprises," *Journal of World Business*, vol. 57, no. 3, Article ID 101291, 2022.

[2] A. Castillo, J. Benitez, J. Liorens, and J. Braojos, "Impact of social media on the firm's knowledge exploration and knowledge exploitation: the role of business analytics talent," *Journal of the Association for Information Systems*, vol. 22, no. 5, pp. 1472–1508, 2021.

[3] S. De Brito and J. Leitão, "Mapping and defining entrepreneurial ecosystems: a systematic literature review," *Knowledge Management Research and Practice*, vol. 19, no. 1, pp. 21–42, 2021.

[4] C. Liu and L. Li, "Place-based techno-industrial policy and innovation: government responses to the information revolution in China," *China Economic Review*, vol. 66, Article ID 101600, 2021.

[5] Z. Gu, F. Meng, and M. Farrukh, "Mapping the research on knowledge transfer: a scientometrics approach," *IEEE Access*, vol. 9, Article ID 34647, 2021.

[6] X. Hu, Y. Tang, and K. Motohashi, "Varied university-industry knowledge transfer channels and product innovation performance in Guangdong manufacturing firms," *Knowledge Management Research and Practice*, vol. 19, no. 2, pp. 197–207, 2021.

[7] U. Zander and B. Kogut, "Knowledge and the speed of the transfer and imitation of organizational capabilities: an empirical test," *Organization Science*, vol. 6, no. 1, pp. 76–92, 1995.

[8] B. L. Simonin, "Transfer of marketing know-how in international strategic alliances: an empirical investigation of the role and antecedents of knowledge ambiguity," *Journal of International Business Studies*, vol. 30, no. 3, pp. 463–490, 1999.

[9] N. R. Quigley, P. E. Tesluk, E. A. Locke, and K. M. Bartol, "A multilevel investigation of the motivational mechanisms underlying knowledge sharing and performance," *Organization Science*, vol. 18, no. 1, pp. 71–88, 2007.

[10] S. Luo, Y. Du, P. Liu, Z. Xuan, and Y. Wang, "A study on coevolutionary dynamics of knowledge diffusion and social network structure," *Expert Systems with Applications*, vol. 42, no. 7, pp. 3619–3633, 2015.

[11] J. Yang Su, Y. Zhang, and X. Su, "Knowledge transfer efficiency measurement with application for open innovation networks," *International Journal of Technology Management*, vol. 81, no. 1/2, pp. 118–142, 2019.

[12] T.-Y. Chen, Y. M. Chen, P. Y. Chen, and C. J. Lin, "A two-dimensional knowledge authorization evaluation method enabling inter-enterprise knowledge sharing," *Computers & Industrial Engineering*, vol. 108, pp. 124–135, 2017.

[13] J. L. Wu and J. G. Pang, "Evaluation of knowledge exchange efficiency in virtual academic community based on SBM model," *Information Science*, vol. 35, no. 9, pp. 125–130, 2017.

[14] X. Zhu, C. Peng, J. Luo, and J. Jiang, "The research on knowledge flow performance evaluation for communities of practice based on two-tuple linguistic multiple decision method," *Science of Science and Management of S.&.T.* vol. 38, no. 5, pp. 101–116, 2017.

[15] R. Cowan, N. Jonard, and J. B. Zimmermann, "Bilateral collaboration and the emergence of innovation networks," *Management Science*, vol. 53, no. 7, pp. 1051–1067, 2007.

[16] G.-Y. Yang, Z.-L. Hu, and J.-G. Liu, "Knowledge diffusion in the collaboration hypernetwork," *Physica A: Statistical Mechanics and Its Applications*, vol. 419, pp. 429–436, 2015.

[17] B. Z. Li, J. W. Zeng, H. W. Dong, and D. Wang, "A model for knowledge diffusion in communities of practice and its simulation," *Chinese Journal of Management*, vol. 30, no. 1, pp. 130–139, 2021.

[18] L. S. Gai and J. Dong, "Measurement of the knowledge management efficiency of rational hi-tech industries study of jiangsu province-case," *Science and Technology Management Research*, vol. 33, no. 21, pp. 72–75, 2013.

[19] C. Li, L. I. Chen, J. Zhang, and Y. Ni, "Research on knowledge sharing model of manufacturing industry cluster based on AHP-fuzzy mathematics evaluation method," *Science and Technology Management Research*, vol. 36, no. 14, pp. 139–144, 2016.

[20] A. Kochenkova, R. Grimaldi, and F. Munari, "Public policy measures in support of knowledge transfer activities: a review of academic literature," *The Journal of Technology Transfer*, vol. 41, no. 3, pp. 407–429, 2016.

[21] X. Li, "An English translation quality evaluation model integrating knowledge transfer and wireless network," *Mobile Information Systems*, vol. 2022, pp. 1–10, Article ID 2086486, 2022.

[22] G. Szulanski, "The process of knowledge transfer: a diachronic analysis of stickiness," *Organizational Behavior and Human Decision Processes*, vol. 82, no. 1, pp. 9–27, 2000.

[23] P. B. Hu, "The indicator system construction of on the efficiency evaluation of knowledge sharing in network organizations," *Journal of Intelligence*, vol. 28, no. 1, pp. 68–71, 2009.

[24] A. Burmeister, U. Fasbender, and J. Deller, "Being perceived as a knowledge sender or knowledge receiver: a multistudy investigation of the effect of age on knowledge transfer," *Journal of Occupational and Organizational Psychology*, vol. 91, no. 3, pp. 518–545, 2018.

[25] G. N. Thompson, C. A. Estabrooks, and L. F. Degner, "Clarifying the concepts in knowledge transfer: a literature

review," *Journal of Advanced Nursing*, vol. 53, no. 6, pp. 691–701, 2006.
[26] L. Kong, F. Ciabuschi, and O. Martín Martín, "Expatriate managers' relationships and reverse knowledge transfer within emerging market MNCs: the mediating role of subsidiary willingness," *Journal of Business Research*, vol. 93, pp. 216–229, 2018.
[27] Y. Héliot and M. Riley, "A study of indicators of willingness in the knowledge transfer process," *Journal of Management and Organization*, vol. 16, no. 3, pp. 399–410, 2010.
[28] S. S. Levine and M. J. Prietula, "How knowledge transfer impacts performance: a multilevel model of benefits and liabilities," *Organization Science*, vol. 23, no. 6, pp. 1748–1766, 2012.
[29] J. L. Cummings and B.-S. Teng, "Transferring R&D knowledge: the key factors affecting knowledge transfer success," *Journal of Engineering and Technology Management*, vol. 20, no. 1-2, pp. 39–68, 2003.
[30] M. H.-W. Ho and F. Wang, "Unpacking knowledge transfer and learning paradoxes in international strategic alliances: contextual differences matter," *International Business Review*, vol. 24, no. 2, pp. 287–297, 2015.
[31] R. Filieri, R. C. McNally, M. O'Dwyer, and L. O'Malley, "Structural social capital evolution and knowledge transfer: evidence from an Irish pharmaceutical network," *Industrial Marketing Management*, vol. 43, no. 3, pp. 429–440, 2014.
[32] F. J. Froese, S. Stoermer, B. S. Reiche, and S. Klar, "Best of both worlds: how embeddedness fit in the host unit and the headquarters improve repatriate knowledge transfer," *Journal of International Business Studies*, vol. 52, no. 7, pp. 1331–1349, 2021.
[33] E. Bacon, M. D. Williams, and G. H. Davies, "Recipes for success: conditions for knowledge transfer across open innovation ecosystems," *International Journal of Information Management*, vol. 49, pp. 377–387, 2019.
[34] F. Karnani, "How can industry use the innovation potential of universities and research institutes? Systematization, evaluation, and new development of transfer strategies based on the theory of knowledge at the interface between companies and science," *International Journal of Innovation Science*, vol. 6, no. 4, pp. 205–212, 2014.
[35] X. Xie and H. Wang, "How to bridge the gap between innovation niches and exploratory and exploitative innovations in open innovation ecosystems," *Journal of Business Research*, vol. 124, pp. 299–311, 2021.
[36] Y. Wei and S. Miraglia, "Organizational culture and knowledge transfer in project-based organizations: theoretical insights from a Chinese construction firm," *International Journal of Project Management*, vol. 35, no. 4, pp. 571–585, 2017.
[37] S.-H. Liao, D. Hu, C. C. Chen, and Y. L. Lin, "Comparison of competing models and multi-group analysis of organizational culture, knowledge transfer, and innovation capability: an empirical study of the Taiwan semiconductor industry," *Knowledge Management Research and Practice*, vol. 13, no. 3, pp. 248–260, 2015.
[38] M. Zhou, K. Govindan, and X. Xie, "How fairness perceptions, embeddedness, and knowledge sharing drive green innovation in sustainable supply chains: an equity theory and network perspective to achieve sustainable development goals," *Journal of Cleaner Production*, vol. 260, Article ID 120950, 2020.
[39] L. Kuang, N. Huang, Y. Hong, and Z. Yan, "Spillover effects of financial incentives on non-incentivized user engagement: evidence from an online knowledge exchange platform," *Journal of Management Information Systems*, vol. 36, no. 1, pp. 289–320, 2019.
[40] D.-G. Ko, L. J. Kirsch, and W. R. King, "Antecedents of knowledge transfer from consultants to clients in enterprise system implementations," *MIS Quarterly*, vol. 29, no. 1, pp. 59–85, 2005.
[41] G.-C. Wu, Y.-H. Cheng, and S.-Y. Huang, "The study of knowledge transfer and green management performance in green supply chain management," *African Journal of Business Management*, vol. 4, no. 1, pp. 044–048, 2010.
[42] A. Thomas and J. Paul, "Knowledge transfer and innovation through university-industry partnership: an integrated theoretical view," *Knowledge Management Research and Practice*, vol. 17, no. 4, pp. 436–448, 2019.
[43] M. Franco and C. Pinho, "A case study about cooperation between University Research Centres: knowledge transfer perspective," *Journal of Innovation & Knowledge*, vol. 4, no. 1, pp. 62–69, 2019.
[44] E. Koksalmis and Ö. Kabak, "Deriving decision makers' weights in group decision making: an overview of objective methods," *Information Fusion*, vol. 49, pp. 146–160, 2019.
[45] R. Sadiq and S. Tesfamariam, "Probability density functions based weights for ordered weighted averaging (OWA) operators: an example of water quality indices," *European Journal of Operational Research*, vol. 182, no. 3, pp. 1350–1368, 2007.
[46] S.-M. Chen, "Evaluating weapon systems using fuzzy arithmetic operations," *Fuzzy Sets and Systems*, vol. 77, no. 3, pp. 265–276, 1996.
[47] Y. Zhu, D. Tian, and F. Yan, "Effectiveness of entropy weight method in decision-making," *Mathematical Problems in Engineering*, vol. 2020, Article ID 3564835, 5 pages, 2020.
[48] W. Chu, Y. Li, C. Liu, W. Mou, and L. Tang, "A manufacturing resource allocation method with knowledge-based fuzzy comprehensive evaluation for aircraft structural parts," *International Journal of Production Research*, vol. 52, no. 11, pp. 3239–3258, 2014.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 3473716, 12 pages
https://doi.org/10.1155/2022/3473716

Research Article

The Improvement Effect of Online Teaching Based on Intelligent Speech Recognition Technology on the Teaching Management Mode in Colleges and Universities

Tian Tian

Wuxi Vocational College of Science and Technology, Wuxi 214028, China

Correspondence should be addressed to Tian Tian; 3101050@wxsc.edu.cn

Received 23 May 2022; Revised 12 June 2022; Accepted 25 June 2022; Published 23 July 2022

Academic Editor: Jiafu Su

In order to improve the advanced nature of teaching management mode in colleges and universities, this paper analyzes the teaching management process in colleges and universities combined with intelligent speech recognition technology. For the dataset of linear teaching speech information distribution, in the case of overlapping of various datasets, this paper finds the overlapping points of various datasets through an algorithm. Moreover, this paper makes full use of the known information to find a vector that can equivalently represent the linear teaching speech information where each data point is located. After that, this paper takes it as the feature vector of each data point and then performs clustering through spectral clustering to obtain the clustering result. The experimental research results show that the model in this paper has a good effect on speech recognition in college teaching, and the model in this paper has a good role in promoting teaching management in colleges and universities.

1. Introduction

As a first-level administrative management organization in a university, a college is a collection of multiple or single disciplines and majors, and it is a very important part of the internal management system of colleges and universities. At present, the college-level management system of most colleges and universities in China is still very imperfect, and each college has not yet formulated an effective management system according to the school's system and the actual situation of the college. In addition, the management relationship between universities and colleges has not been completely straightened out, which also affects the release of energy from college-level units and makes the teaching management work of colleges lack vitality. Therefore, the implementation of two-level management of the school and the school needs to be based on the school, and the focus should be placed on the management system of the school. At the same time, it is necessary to establish a college-level teaching management mode that is fast and efficient, and is conducive to the full play of the functions of each college. First of all, it is necessary to solve the problems existing in the current process of college-level teaching management in Chinese universities.

In the traditional college management process, colleges and universities will emphasize the function of the organizational structure and its standard procedures. The "external control" management mainly achieves unified goals through standard, procedural, and general standard methods and procedures [1]. It is required that the goals of each college are consistent, while ignoring the unyielding characteristics and methods of each college, adopting a centralized management principle, and emphasizing the control of the process, in order to avoid problems in each college, so, no matter big or small, the school is responsible for details. With careful control, the college has little autonomy, and everything has to be reported to the school for instructions, which often results in the phenomenon of "external control" management, which makes it impossible for the college to solve problems in time [2]. From the perspective of the college as a whole, the college is often regarded as a tool or way to achieve the school's goals. It is a passive system that must be tightly controlled and has no clear autonomy, responsibility, and obligation at all; "external control" management emphasizes the management of the organization, as long as the organization. The goal is clear. People can work effectively, so they will be more inclined to supervise the college and constantly introduce new regulations to control the school. This has resulted in the expansion of institutions at the school level, the college's dependence on the school has become stronger, and the college has become ineffective. In the process of college-based teaching management, colleges and universities emphasize that there are many different ways for each college to achieve its goals, highlighting flexibility, and each college should be based on its own different conditions [3]. Take a variety of approaches. Rather than adopting a standard method for college management, this allows each college to have a larger space for activities to develop and formulate teaching management methods suitable for the college; the power and responsibility are moved down to the college, which is a prominent feature of modern college management phenomenon, because the daily goal of college teaching management is efficiency and problem solving rather than escaping or avoiding the problem [4]. The college-based management should effectively discover the problems of the college itself, solve them in time, and make due efforts to improve the teaching efficiency: the college-based management does not deny that the college is a tool or way to achieve goals, but believes that to achieve the school. There are many ways to achieve the goals of the school. Therefore, on the premise of realizing the school goals, the college has greater autonomy, assumes certain responsibilities and obligations, conducts effective management by itself, develops teaching, allocates teachers and materials, solves problems according to the characteristics of the school problems, and achieve goals [5]; the college-based teaching management emphasizes the establishment of a suitable environment, allowing members of the college to participate in the realization of the goal, fully excavating and developing the human resources in the college, and giving full play to their enthusiasm, so as to improve the teaching quality more effectively [6].

In the external control management model, college managers often think that the goals are clear, simple, and static, the college is only a tool and means to achieve these goals, and the value of teachers is the value of the tool; the college-based teaching management model believes that the college is not only a student. The place of learning and growth is also the place where the faculty and staff learn and grow. If the potential of students and faculty cannot be continuously developed, the potential of the college or even the entire school will not be developed [7]. Decision-making is a participatory process. All members of an organization have an equal opportunity to influence decisions and actions. Under the strict control of external control management, the decision-making of the college is mostly made by the administrative staff of the school and then handed over to each college for implementation. The faculty and staff of the college rarely have the opportunity to participate in the decision-making and are regarded as unimportant; in a complex educational environment, teaching tasks are diverse and require the wisdom, imagination, and efforts of many people to complete. Therefore, the decision-making method at the school level is gradually transformed into "decentralized" or "participatory," so that more people, including the faculty, staff, students, and even the parents of the various colleges, involved, and they will be more actively involved in the teaching and management of the college [8]. In terms of leadership style, the leadership style of external control management focuses on the lower level, thinking that technology and interpersonal relationships are very important to achieve goals, while ignoring cultural leadership [9]; when the background and ideas are more and more divergent, the dean should lead by example, help and guide the members of the college to understand the meaning of various activities of the college, eliminate differences and misunderstandings between individuals, and clarify unstable and uncertain factors [10].

The primary task of the project management implementation of the teaching management information system is to clarify the goals of the project. Only correct and clear goals can ensure the success of the project. The goal of the implementation of the teaching management information system project of the School of Mechanical and Electrical Engineering is to deliver a complete set of information systems that can meet the teaching management of the college, and each functional module can realize sub-functions to meet the functions of the entire system. The realization of system functions can save manpower, standardize teaching management, realize auxiliary analysis, and provide basis for high-level decision-making. This goal is consistent with the long-term development strategy of the college [11]. Quantitative benefits include building a school management platform, reducing labor costs, improving work efficiency, and improving service levels for teachers and students [12]. Qualitative benefits include giving full play to the role of people, improving service quality, realizing the informatization of the educational process, establishing a more effective teaching management mechanism, scientifically and uniformly allocating teaching resources, and ensuring the healthy and sustainable development of the school [13].

When the actual situation in the project implementation process deviates from the project management plan, contingency measures must be taken to reduce the project risk caused by not considering the impact of the change on the overall project goal or plan. Change control runs through the project, and the project manager is ultimately responsible for it [14]. Changes need to be managed carefully and continuously to maintain the project management plan. During the implementation of the teaching management information system project, each stakeholder can make a change request orally, but all change requests must strictly implement the change management process to prevent scope creep caused by randomness [15]. The focus should be to assess the time and cost impact of the change and provide the results of the assessment. Change requests are issued during the execution of an instructional management information system project including corrective actions, preventive actions, and defect remediation [16]. During the implementation of the project, the output of the teaching management information system often does not conform to the management system of our school, and the result is not what we need. At this time, it needs to be corrected. Therefore, before implementing certain modules, some preventive measures will be taken to exclude the factors that will affect the output to ensure the desired output result [17].

This paper combines the intelligent speech recognition technology to analyze the teaching management process in colleges and universities, improve the teaching management mode in colleges and universities, and improve the efficiency of teaching management in colleges and universities.

2. Speech Recognition Technology

2.1. A New Linear Speech Information Clustering Algorithm. Generalized principal component analysis (GPCA) algorithm clusters linear speech information data by partitioning the dataset into corresponding subspaces by using algebraic geometric methods. Marc Pollefeys et al. proposed a spectral clustering algorithm based on local subspace affinity (LSA).

GPCA is a method of partitioning subspaces, which uses algebraic geometric methods to partition a dataset into corresponding subspaces. The joint distribution of k linear subspaces can be fitted by a set of polynomials of degree k, and the derivative of these polynomials at a point is a normal vector of the subspace at that point. We assume that any point in the dataset is $x \in R^k$, the dataset is in n different linear subspaces Ω_c, and any point $x \in R^k$ in the dataset satisfies $b_i^T x = 0$. Among them, b_i is the vertical vector of the subspace to which x belongs. The original dataset is $X \in R^K$, and the data in the dataset are in n different linear subspaces X_c, $\cup X_c = X, c = 1 \ldots n$. We set the vertical vector of these n linear subspaces to be $b_c, c = 1 \ldots n$, and for $x_i \in X_c$, there is $b_c \cdot x_i = 0$. Then, for any $x \in X$, there is $\prod_{c=1}^{n} b_c \cdot x = 0$, which is $\prod_{c=1}^{n} (b_{c1} \cdot x_1 + \ldots + b_{cK} x_K) = 0$. Among them, K is the dimension of the subspace, and this polynomial is expanded to get

$$c_1 x_1^n + c_2 x_1^{n-1} x_2 + \ldots + c_{M_n} x_K^n = V_n(x)^T C. \tag{1}$$

Among them, $V_n(x)$ is the Veronese map of x. C can be composed of left singular vectors corresponding to several minimum singular values of $V_n(x)$, and then b_c is obtained by factoring, and the vertical vector of each subspace is obtained. Then, by clustering, each data point is divided into its nearest linear subspace.

The LSA method mainly includes the following steps:

(1) The arc cosine angle between each sample point is calculated and sorted to obtain each K sample point closest to the sample data_i.

(2) For each sample point, itself and the K nearest sample points are formed into a matrix, and a local affine subspace Ωi is fitted to the matrix by singular value decomposition (SVD).

(3) The similarity matrix is computed: the distance between the local affine subspaces can be measured by the principal angle between them. The principal angle μ_{ij} between the local affine subspace Ω_i of data_i and the local affine subspace Ω_j of data_j is calculated as

$$\mu_{ij} = e^{\left(-\Sigma_{m=1,\ldots,M} \sin^2\left(\theta_{ij}^m\right)\right)}. \tag{2}$$

Among them, θ_{ij}^m is the m-th principal angle between the local affine subspaces Ω_i and data_j of the local affine subspace Ω_j.

(4) Spectral clustering: Through the spectral clustering method, the obtained similarity matrix is used as input, and the clustering result is obtained.

Spectral curvature clustering linear speech information clustering algorithm is based on multi-channel clustering technology. Usually, at least $d+1$ data points are required to define a d-dimensional affine subspace. Spectral curvature clustering uses a multi-channel similarity measure to describe the possibility that $d+2$ data points come from the same latent linear speech information, so as to construct the similarity measure matrix W between data points. We set $m\Omega_{d+2} = \{x_{il}\}_{l=1}^{d+2}$ to be $d+2$ randomly selected data points, and spectral curvature clustering is based on the probability of polar curvature to define a multi-way similarity $S_{i_1,i_2,\ldots,i_{d+2}}$:

$$S_{i_1,i_2,\ldots,i_{d+2}} = \exp\left(-\frac{1}{2\sigma^2} diam^2(\Omega_{d+2}) \sum_{l=1}^{d+2} \frac{(d+1)!^2 vol^2(\Omega_{d+2})}{\prod_{1\le m\le d+2, m\ne l} \|x_{im} - x_{il}\|^2}\right). \tag{3}$$

If and only if $i_1, i_2, \ldots, i_{d+2}$ is different, there is $S_{i_1,i_2,\ldots,i_{d+2}} = 0$. Among them, $\text{diam}(\Omega)$ represents the radius of the set Ω, and $\text{vol}(\Omega)$ represents the volume of the $d+1$-dimensional simplex formed by the points in the set Ω. The similarity matrix W is defined as

$$w_{ij} = \sum_{k_2,\ldots,k_{d+2}} S_{i,\mathbf{k}_2,\ldots,\mathbf{k}_{d+2}} S_{j,k_2,\ldots,k_{d+2}}. \tag{4}$$

In spectral curvature clustering, the initial sampling in the iterative sampling will significantly affect the final clustering result, it is easy to fall into the local optimal solution, and the algorithm is unstable.

2.2. A New Linear Speech Information Clustering Algorithm. For two straight lines in the same plane, in general, there is an intersection (there is overlap). For three straight lines on the same plane, in general, there are three intersections, as shown in Figure 1. Among them, $p1$, $p2$, and $p3$ are 3 overlapping points. For n straight lines, there are $n(n-1/2)$ intersections.

We assume that there are n types of datasets, each type of data conforms to the linear speech information distribution, and there is overlap between various types of data. First, the method of finding density peaks is used to obtain the overlapping points of various data, and the overlapping point set $P = \{p1, p2, \ldots, pk\}$ between various types of data is obtained. The algorithm to find the set of overlapping points in the sample is as follows:

Input: set of data points X

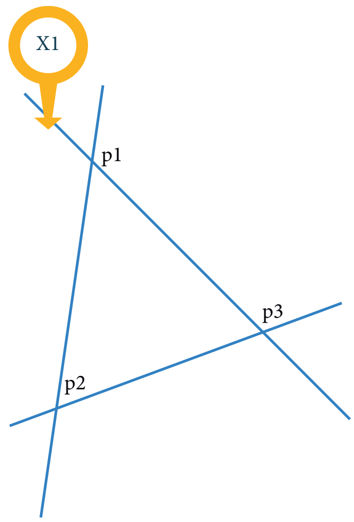


Figure 1: Linear speech information data with overlapping points.

Output: set of overlapping points $P = \{p1, p2, \ldots, pk\}$

step: 1: The algorithm is based on the assumption that the sample points at the overlap of each linear subspace are denser. For each sample point x_i, feature attribute ρ_i is calculated. Among them, ρ_i represents the local density of sample point x_i, and δ_i represents the minimum distance between sample point x_i and all points with higher local density than x_i. The calculation method of ρ_i is$\rho_i = \sum_j \exp[-(d_{ij}/d_c)^2]$.

Among them, d_{ij} represents the Euclidean distance between the sample point x_i and the sample point x_j, and d_c represents the truncation distance.

2: The calculation method of the feature attribute δ_i, δ_i is $\delta_i = \min_{j:\rho_j > \rho_i}(d_{ij})$.

δ_i is the minimum distance between sample point x_i and all points with higher local density than x_i. For the point with the highest density, there is $\delta_i = \max_j (d_{ij})$.

3: Only ρ_i and δ_i of the overlapping point are relatively large. For the point concentration with high density, the point near the overlapping point is only ρ_i which is larger, δ_i is equal to the distance between it and the overlapping point, the distance between the two points is very close, and δ_i is smaller. From this, we can draw a two-dimensional decision diagram by finding the ρ_i and δ_i of each sample point and obtain the overlapping point.

As shown in Figure 2(a), there are two types of linear data with overlapping points. According to the above method, the decision diagram is drawn, as shown in Figure 2(b), and the point in the upper right corner of the diagram is the overlap point.

We set P_i, P_j to represent two vectors and define the degree of parallelism s_{ij} between the two vectors as

$$s_{ij} = abs\left(\frac{P_i \cdot P_j}{\|P_i\|\|P_j\|}\right). \tag{5}$$

The closer the degree of parallelism is to 1, the more parallel the two vectors are.

The new linear speech information clustering algorithm consists of the following steps:

(1) The overlapping point is found: according to Algorithm 1, the overlapping point set $P = \{p1, p2, \ldots, pk\}$ is found.

(2) The feature vector is calculated: for the sample point, $x_i \in R^d, d$ is the sample space dimension. For the overlapping point set $P = \{p1, p2, \ldots, pk\}$, we find $x_i p_k = (x_i - p_k)/\|x_i - p_k\|$ separately, and then there must be m vectors that are nearly parallel (the angle is small or close to 180°). According to formula (7), the degree of parallelism of each vector is calculated, the two vectors $x_i p_1^*$ and $x_i p_2^*$ with the highest degree of parallelism are taken, and the calculation method to obtain its average value as a eigenvector V_i, V_i of the sample point x_i is

$$V_i = \frac{x_i p_1^* + x_i p_2^*}{\|x_i p_1^* + x_i p_2^*\|}. \tag{6}$$

Among them, if the subtraction operation between vectors just results in a zero vector, the negative vector of the subtraction vector can be subtracted. The eigenvectors of a set of points belonging to a linear subspace are parallel to each other.

(3) The similarity matrix W is calculated: the calculation formula of the similarity w_{ij} between the two sample points is

$$w_{ij} = abs\left(\cos\left(\theta_{ij}\right)\right). \tag{7}$$

Among them, θ_{ij} represents the angle between the two corresponding eigenvectors of sample point x_i and sample point x_j:

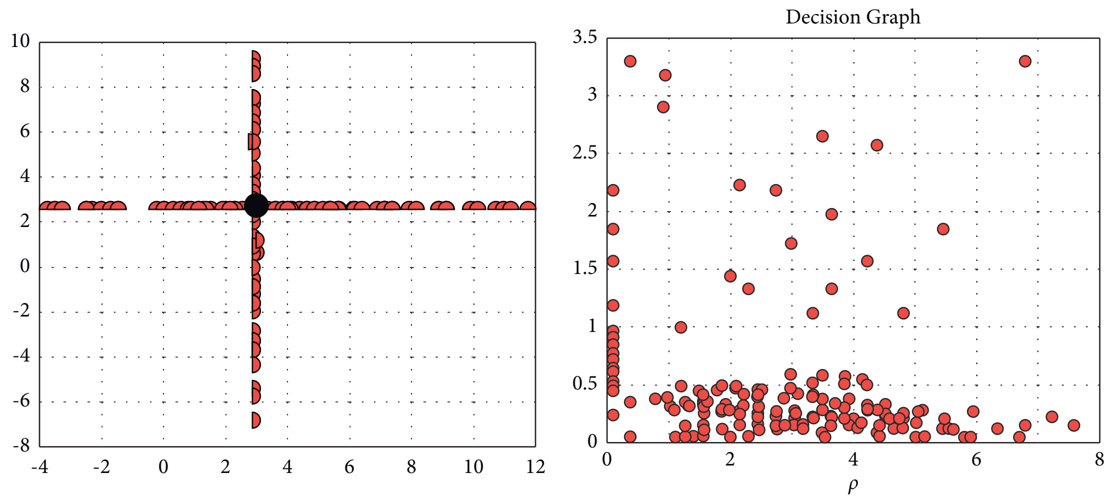
$$\cos\left(\theta_{ij}\right) = \frac{V_i \cdot V_j}{\|V_i\| \cdot \|V_j\|}. \tag{8}$$

For two sample points in a linear subspace, their eigenvectors are nearly parallel, that is, the sample point similarity $abs(\cos(\theta_{ij})) \approx 1$. The two sample points that are not in a linear subspace have a larger angle between their eigenvectors. When the eigenvectors are vertical, there is $abs(\cos(\theta_{ij})) \approx 0$. As shown in Figure 3, it is the angle between the characteristic vectors and the corresponding cosine value.

(4) Spectral clustering: The clustering result of the data is obtained by applying the spectral clustering method to the constructed similarity matrix W.

3. Experimental Results and Analysis

In all the experiments in this paper, the clustering accuracy is used as the criterion to judge the clustering performance. At the same time, the performance of the algorithm is verified



(a)

(b)

Figure 2: Overlapping points are found. (a) Two types of linear data with overlapping points. (b) Decision diagram.

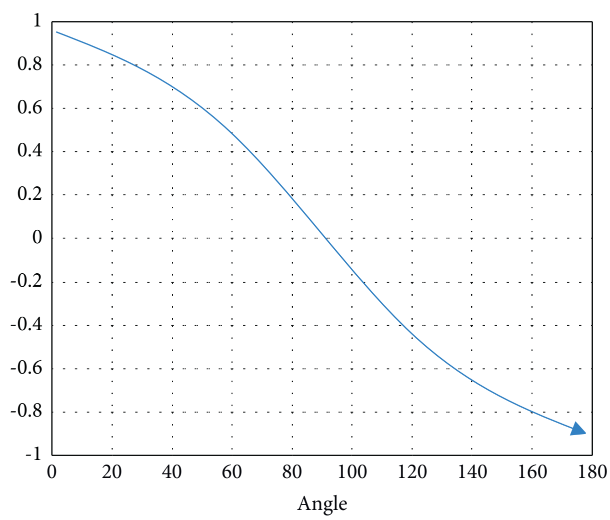


Figure 3: The cosine value corresponding to the angle.

and evaluated in the absence of speech noise and in the presence of speech noise. Among all possible labels, the classification accuracy with the largest alignment with the true label category is the final clustering accuracy, and the formula is

$$\text{clustering accuracy} = \max \sum_{i=1}^{N} \frac{\delta\left(\text{labels}_i = \text{truelabels}_i\right)}{N}. \tag{9}$$

Among them, labels represent the clustering label of the sample point x_i in the algorithm, true labels represent the true label of the sample point x_i, and $\delta(\cdot)$ represents the delta function. The higher the clustering accuracy, the better the clustering performance of the algorithm.

First, experiments are carried out on simulated data to verify the clustering performance of the new linear speech information clustering algorithm. The simulation generates 4 types of data that conform to the distribution of linear speech information, and each type of data has 150 sample points and a total of 600 sample points. At the same time, the 4 sets of data overlap each other, and there are 6 intersections.

The original dataset is shown in Figure 4. As can be seen from the figure, there are four types of data in the original dataset, and each type of data conforms to the linear distribution of speech information. Moreover, there are overlaps between various types of data, and the data to be clustered at the overlaps are generally difficult to process.

In the absence of noise and speech noise, the clustering results of GPCA, LSA, K-flats, and the algorithm in this chapter are shown in Figure 5. It can be seen from Figures 5(a)–5(c) that in the absence of speech noise, the clustering effects of the GPCA algorithm, the K-flats algorithm, and the algorithm in this chapter are all ideal. In Figure 5(d), the overlapping points are clustered into one category, and there are two types of data that are mistakenly clustered into one category. The clustering effect of the LSA algorithm is poor.

On the basis of the original dataset, some speech noises are randomly generated, as shown by the star-shaped points in Figure 6. In the presence of speech noise, the clustering performance of some linear speech information clustering algorithms will be affected because the premise of many linear speech information clustering algorithms is based on the premise that there is no influence of speech noise, such as the least squares method. If there is speech noise, the 2 paradigms of the error will increase significantly, which will increase the error and cause the result to be too biased.

The clustering results of GPCA, LSA, K-flats, and the algorithm in this chapter are shown in Figure 7. It can be seen from the comparison that although there are speech noise and outliers, the algorithm in this paper still has a good clustering effect. First of all, speech noise will not affect the search for overlapping point sets in Algorithm 1. The characteristic attribute δi of speech noise is large, while the characteristic attribute ρi is small. Because the speech noise is relatively scattered, there are no other sample points around, and the local density is small.

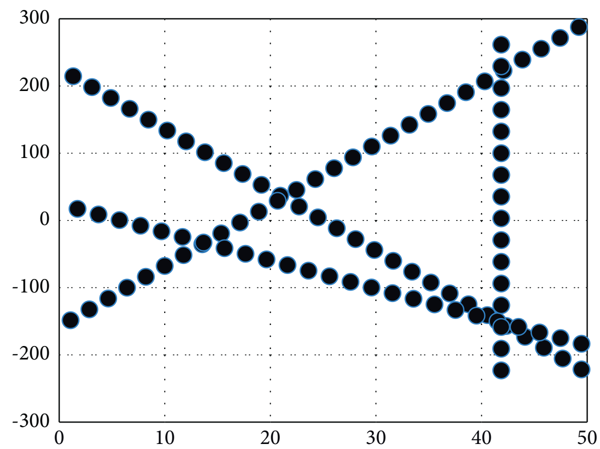


Figure 4: Original dataset.

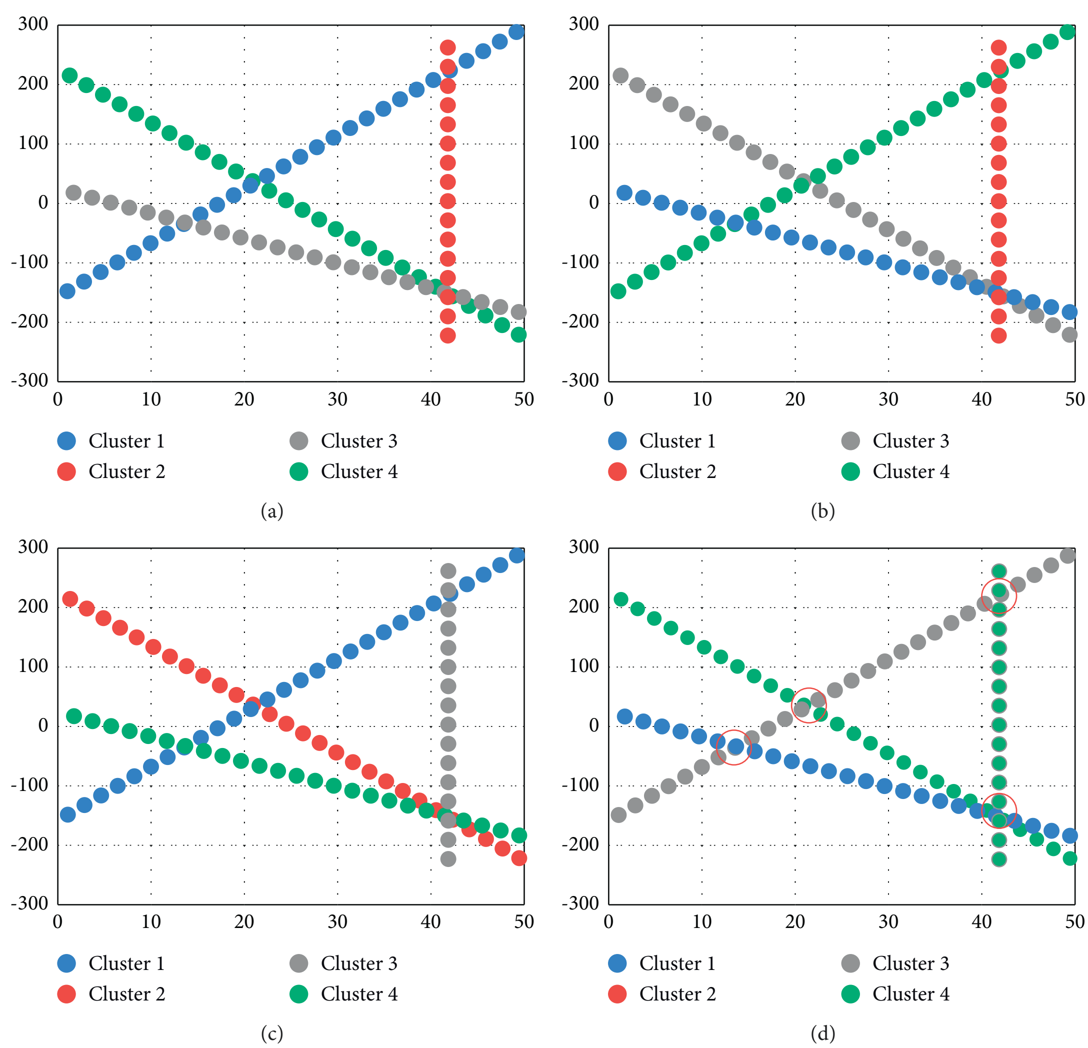


Figure 5: Clustering results of each algorithm without speech noise. (a) Clustering results of the algorithm in this paper. (b) GPCA clustering results. (c) K-flats clustering results. (d) LSA clustering results.

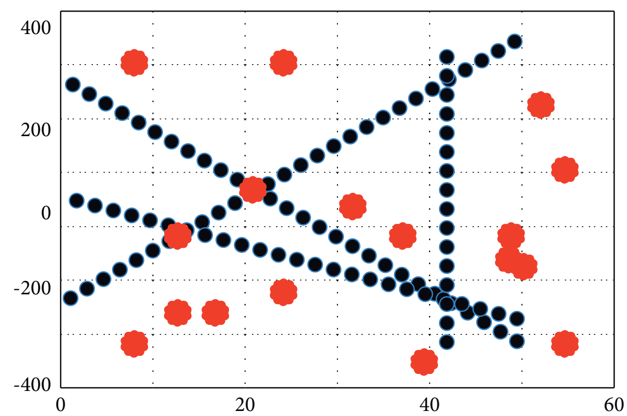


Figure 6: Original dataset with speech noise.

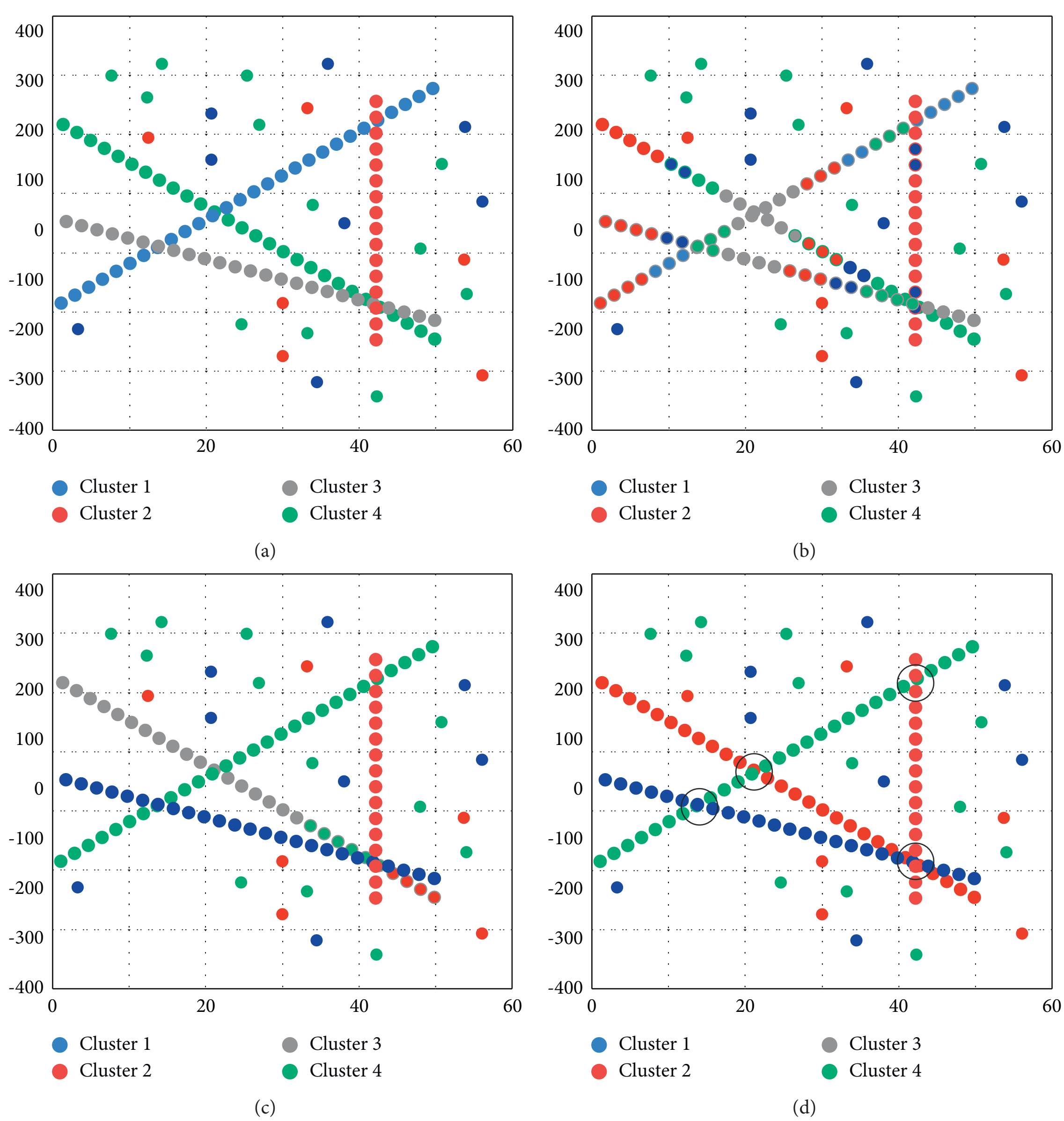


Figure 7: Clustering results of each algorithm in the presence of speech noise. (a) Clustering results of the algorithm in this paper. (b) GPCA clustering results. (c) K-flats clustering results. (d) LSA clustering results.

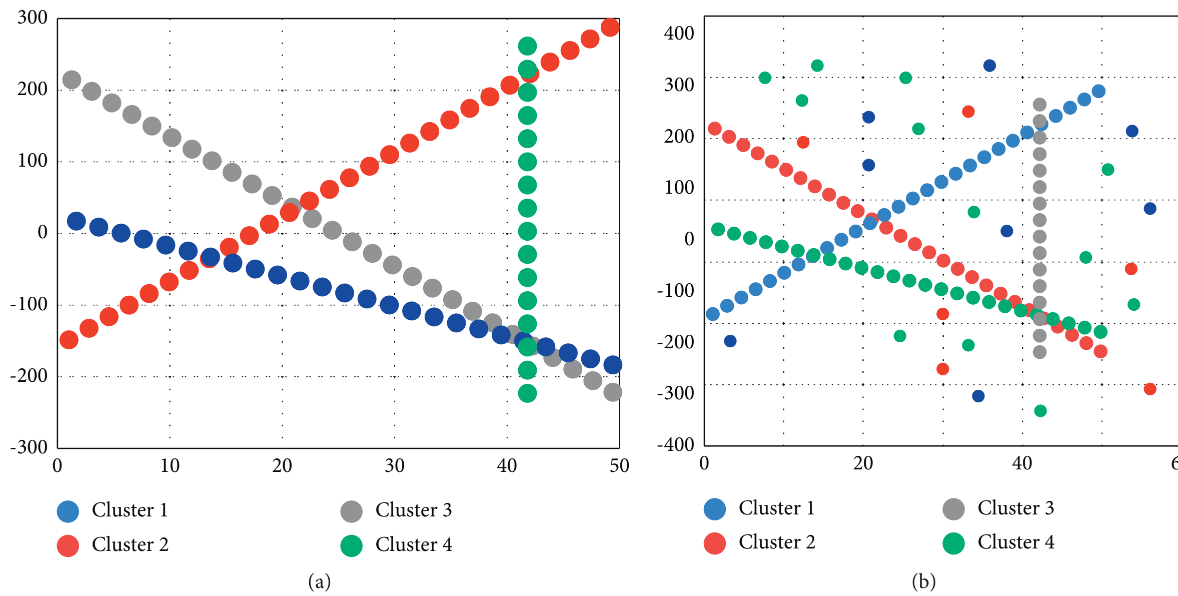


FIGURE 8: Clustering results of spectral curvature clustering algorithm. (a) In the case of no interference point. (b) In the presence of interference points.

The spectral curvature clustering algorithm generally has an ideal clustering effect in the absence of speech noise and in the presence of speech noise, as shown in Figure 8.

However, the spectral curvature clustering algorithm suffers from a disadvantage. That is, the algorithm adopts a random sampling strategy in the initial sampling in the iterative sampling. In some cases, the randomly sampled sample points are all in different linear speech information distributions. As a result, the calculation of multi-channel clustering similarity is biased, which significantly affects the final clustering result, the algorithm often falls into the local optimal solution, and the algorithm is unstable. For the same dataset, 200 experiments are performed, and the accuracy rate will occasionally be very low. As shown in Figure 9, the clustering accuracy rate of multiple experiments will be lower than 70%, all of which are trapped in the local optimal solution. However, after the data points of the algorithm proposed in this paper are determined, the search for overlapping points is fixed, and there is no randomness problem in the subsequent process, so there is no stability problem in the algorithm in this chapter.

In the dataset experiments with and without speech noise, the clustering accuracy of the four methods is shown in Table 1. It can be seen from Table 1 that in the absence of speech noise and outliers, the clustering effect of the GPCA algorithm is the best and better than other algorithms. The clustering effect of the algorithm proposed in this paper is relatively ideal, but for the LSA algorithm, the clustering effect is poor. In the presence of speech noise and outliers, the GPCA and K-flats algorithms have very unsatisfactory clustering effects due to the limitations of their own algorithms, while LSA does not properly handle the data points where the speech information overlaps. Its clustering effect is not ideal whether there is speech noise or not. The algorithm in this paper is not affected by speech noise and outliers, and can also process data points where speech information overlaps. In both cases, the clustering effect is ideal.

In this paper, the time-frequency distribution of the rearranged spectrum is obtained from the multi-component signal, and the method of clustering is used to sort the multi-component signal.

For example, for the multi-component signal composed of 3 signals in Table 2, its expression is

$$s(t) = \exp\left[j2\pi\left(0.05t + 0.4t^2\right)\right] + \exp\left[j2\pi\left(0.45t - 0.4t^2\right)\right] + \exp\left[j2\pi(0.32t)\right]. \tag{10}$$

The time-frequency distribution diagram of the obtained rearrangement spectrum is shown in Figure 10. It can be seen from the figure that the time-frequency distribution sample points of each single-component signal are in line with the linear speech information distribution. However, there is overlap between each single-component signal, so it is suitable to use the algorithm of this chapter for clustering and obtain the time-frequency distribution point set of each single-component signal.

When the sampling frequency is too large, there are too many sample points in the time-frequency distribution, which will cause the calculation of the clustering algorithm to be too slow. Therefore, for each time point, it only extracts n points with the highest energy peaks as sample points, where n is the number of signal components. Then, it clusters the extracted main sample points and then sorts the signals. The sample points obtained are shown in Figure 11. Among them, in order to make the sample point components meet the requirements of the algorithm in this paper, the energy value is the same value.

This algorithm is used to cluster the sample points, and the clustering results are shown in Figure 12. The

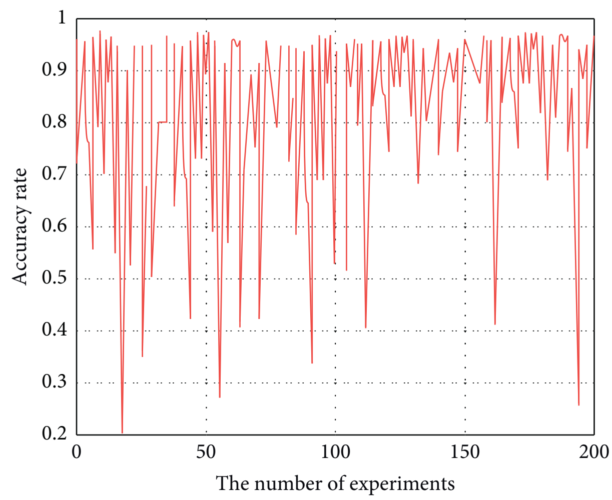


Figure 9: Accuracy rate of 200 experiments of SCC algorithm.

Table 1: Comparison of accuracy rates of various algorithms.

Method	GPCA	LSA	K-flats	The algorithm proposed in this paper
Accuracy (no speech noise)	98.51	68.55	96.23	96.34
Accuracy (presence of speech noise)	46.02	68.16	69.14	95.83

Table 2: Parameters of multi-component signal.

	Modulation	Initial frequency	Frequency modulation
Signal component 1	Linear FM	0.05	0.40
Signal component 2	Linear FM	0.45	−0.40
Signal component 3	Constant frequency	0.32	0.0

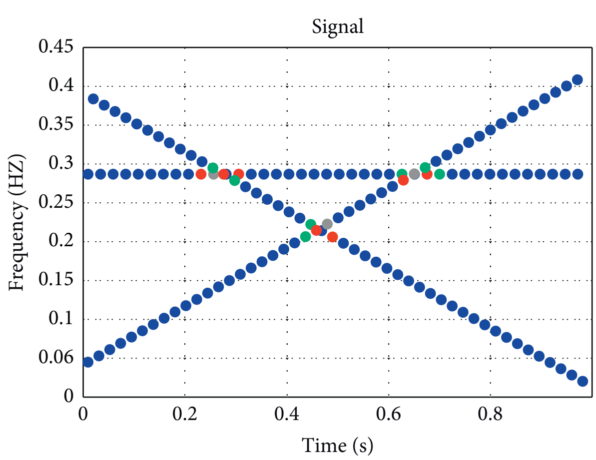


Figure 10: Time-frequency distribution diagram of the rearranged spectrum of the multi-component signal (10).

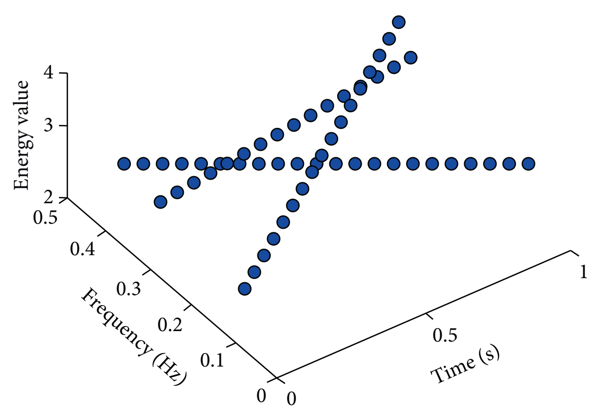


Figure 11: Original sample point set.

parameter estimation is shown in Table 3. It can be seen from the table that the parameter estimation error of the multi-component signal is small. Since the processing of the time-frequency distribution of multi-component signals at the overlapping point has not been well resolved, the time-frequency distribution at the overlapping point has some deviations, and the clustering results will be affected to some extent. However, the overall clustering effect is ideal.

In this paper, a new clustering method of linear speech information is presented by exploiting the geometric properties between data points. In this paper, the overlapping points between each linear subspace are obtained by finding the peak density of data points. Then, using the geometric properties between data points and

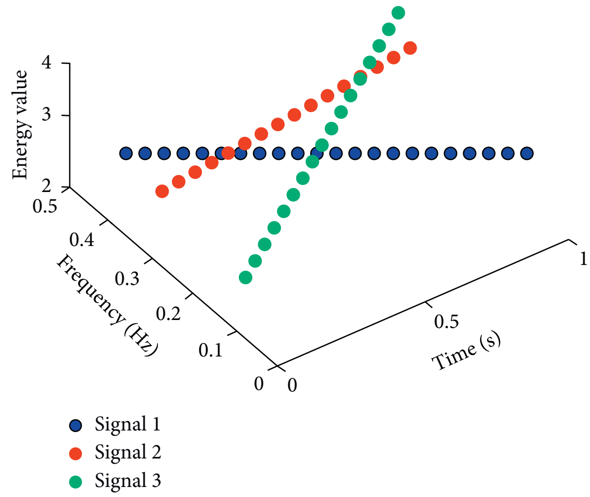


Figure 12: Clustering results.

Table 3: Parameter estimation of multi-component signals (10).

	Actual value	Estimated value
Carrier frequency of component signal 1	0.050	0.053
Frequency modulation of component signal 1	0.400	0.393
Carrier frequency of component signal 2	0.450	0.443
Frequency modulation of component signal 2	−0.400	−0.405
Carrier frequency of component signal 3	0.320	0.318
Frequency modulation of component signal 3	0.0	0.002

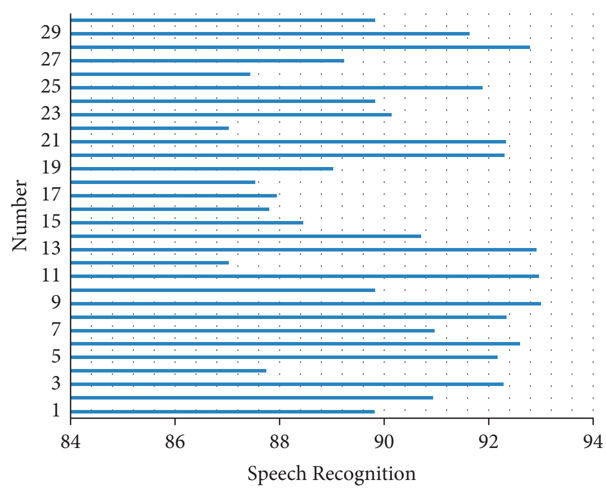


Figure 13: The effect of the model proposed in this paper on speech recognition in college teaching.

overlapping points, the constructed data points can be equivalently represented as eigenvectors of the linear subspace in which they reside. After that, the absolute value of the cosine value of the angle of the eigenvectors between the data points is obtained as the similarity between the data points, the similarity matrix is constructed, and finally the clustering result is obtained by spectral clustering. The experimental results show that the clustering effect of this method is ideal in both simulated data and real data. However, the algorithm has certain limitations; that is, the distribution of data points needs to be uniform, and there are overlapping points between various types of data.

On this basis, this paper explores the effect of this model in speech recognition in college teaching and obtains the results shown in Figure 13.

Through the above research, it is verified that the model in this paper has a good effect in college teaching speech recognition. On this basis, the promotion effect of the model proposed in this paper on the teaching management of colleges and universities is explored, and the results shown in Figure 14 are obtained.

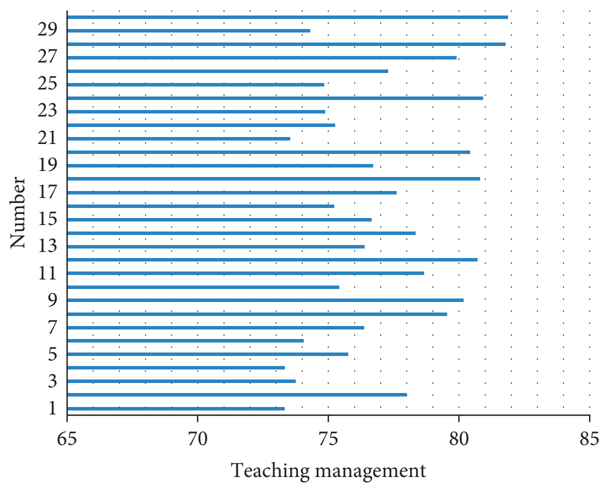


Figure 14: The promotion effect of the model proposed in this paper on teaching management in colleges and universities.

The above research verifies that the model proposed in this paper has a good role in promoting teaching management in colleges and universities.

4. Conclusion

The goal of teaching management information system project implementation is the smooth use of the system. In order to deliver the teaching management information system, the various scopes of work in the implementation process must be completed, and aspects such as technical support and human resources must be evaluated. At the same time, for the smooth implementation of the project, the school should also clarify the measures taken and the resources to be invested to ensure that the project is completed on time. At the same time, the teaching management information system can bring quantitative and qualitative benefits to schools. This paper combines the intelligent speech recognition technology to analyze the teaching management process in colleges and universities, and improve the teaching management mode of colleges and universities. The experimental research results show that the model proposed in this paper has a good effect on speech recognition in college teaching, and the model proposed in this paper has a good role in promoting teaching management in colleges and universities.

Data Availability

The labeled dataset used to support the findings of this study is available from the author upon request.

Conflicts of Interest

The author declares no conflicts of interest.

Acknowledgments

This study was sponsored by Wuxi Vocational College of Science and Technology.

References

[1] G. Cooper, H. Park, Z. Nasr, L. P. Thong, and R. Johnson, "Using virtual reality in the classroom: preservice teachers' perceptions of its use as a teaching and learning tool," *Educational Media International*, vol. 56, no. 1, pp. 1–13, 2019.

[2] J. Zhao, X. Xu, H. Jiang, and Y. Ding, "The effectiveness of virtual reality-based technology on anatomy teaching: a meta-analysis of randomized controlled studies," *BMC Medical Education*, vol. 20, no. 1, p. 127, 2020.

[3] S. J. Bennie, K. E. Ranaghan, H. Deeks et al., "Teaching enzyme catalysis using interactive molecular dynamics in virtual reality," *Journal of Chemical Education*, vol. 96, no. 11, pp. 2488–2496, 2019.

[4] S. F. M. AlfalahAlfalah, J. F. M. FalahFalah, T. Alfalah, M. Elfalah, N. Muhaidat, and O. Falah, "A comparative study between a virtual reality heart anatomy system and traditional medical teaching modalities," *Virtual Reality*, vol. 23, no. 3, pp. 229–234, 2019.

[5] M. Reymus, A. Liebermann, and C. Diegritz, "Virtual reality: an effective tool for teaching root canal anatomy to undergraduate dental students - a preliminary study," *International Endodontic Journal*, vol. 53, no. 11, pp. 1581–1587, 2020.

[6] V. L. Dayarathna, S. Karam, R. Jaradat et al., "Assessment of the efficacy and effectiveness of virtual reality teaching module: a gender-based comparison," *International Journal of Engineering Education*, vol. 36, no. 6, pp. 1938–1955, 2020.

[7] O. Hernandez-Pozas and H. Carreon-Flores, "Teaching international business using virtual reality," *Journal of Teaching in International Business*, vol. 30, no. 2, pp. 196–212, 2019.

[8] V. Andrunyk, T. Shestakevych, and V. Pasichnyk, "The technology of augmented and virtual reality in teaching children with ASD," *Econtechmod: Scientific Journal*, vol. 7, no. 4, pp. 59–64, 2018.

[9] R. Mayne and H. Green, "Virtual reality for teaching and learning in crime scene investigation," *Science & Justice*, vol. 60, no. 5, pp. 466–472, 2020.

[10] M. Taubert, L. Webber, T. Hamilton, M. Carr, and M. Harvey, "Virtual reality videos used in undergraduate palliative and oncology medical teaching: results of a pilot study," *BMJ Supportive & Palliative Care*, vol. 9, no. 3, pp. 281–285, 2019.

[11] K. E. McCool, S. A. Bissett, T. L. Hill, L. A. Degernes, and E. C. Hawkins, "Evaluation of a human virtual-reality endoscopy trainer for teaching early endoscopy skills to veterinarians," *Journal of Veterinary Medical Education*, vol. 47, no. 1, pp. 106–116, 2020.

[12] K. Kim and Y. Park, "A development and application of the teaching and learning model of artificial intelligence education for elementary students," *Journal of The Korean Association of Information Education*, vol. 21, no. 1, pp. 137–147, 2017.

[13] O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education–where are the educators?"

International Journal of Educational Technology in Higher Education, vol. 16, no. 1, p. 39, 2019.
[14] S. C. H. Yang, W. K. Vong, R. B. Sojitra, T. Folke, and P. Shafto, “Mitigating belief projection in explainable artificial intelligence via Bayesian teaching,” *Scientific Reports*, vol. 11, no. 1, p. 9863, 2021.
[15] Y. Lee, “An analysis of the influence of block-type programming language-based artificial intelligence education on the learner’s attitude in artificial intelligence,” *Journal of The Korean Association of Information Education*, vol. 23, no. 2, pp. 189–196, 2019.
[16] J. M. Alonso, “Teaching explainable artificial intelligence to high school students,” *International Journal of Computational Intelligence Systems*, vol. 13, no. 1, p. 974, 2020.
[17] R. Yang, “Artificial intelligence-based strategies for improving the teaching effect of art major courses in colleges,” *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, no. 22, p. 146, 2020.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9805461, 1 page
https://doi.org/10.1155/2023/9805461

Retraction

Retracted: The Construction of Green Building Integrated Evaluation System Based on BIM Technology

Mobile Information Systems

Received 22 August 2023; Accepted 22 August 2023; Published 23 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] H. Li and C. Wang, "The Construction of Green Building Integrated Evaluation System Based on BIM Technology," *Mobile Information Systems*, vol. 2022, Article ID 5906827, 12 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 5906827, 12 pages
https://doi.org/10.1155/2022/5906827

Research Article

The Construction of Green Building Integrated Evaluation System Based on BIM Technology

Hongwei Li and Chongyu Wang

Urban and Rural Construction Institute, Hebei Agricultural University, Baoding 071000, Hebei, China

Correspondence should be addressed to Hongwei Li; lihongwei@hebau.edu.cn

Received 23 May 2022; Revised 25 June 2022; Accepted 30 June 2022; Published 18 July 2022

Academic Editor: Xuefeng Zhang

In order to study the problems of information disconnection, poor coordination of project participants, design evaluation, and so on, in the process of green building evaluation in China, this paper proposes a construction method of green building integrated evaluation system based on BIM Technology. This method constructs the research framework through the research on green building evaluation and BIM (building information modeling), analyzes the whole life cycle of BIM in green building, solves the problems existing in green building evaluation in China through the application of BIM, expounds the green building evaluation process under BIM environment, and studies the information interaction of green building evaluation under BIM environment and the application of BIM in green building design evaluation and operation evaluation. The results show that by studying the application of BIM in green building evaluation, it is demonstrated that the green building evaluation based on BIM can effectively improve the evaluation efficiency and realize the informatization, integration, and high efficiency of green building evaluation.

1. Introduction

Since the rapid development of the construction industry, the development of the construction industry has driven the economy, but also caused environmental pollution, sharp increase in resource consumption and other problems. Architecture has relatively complex construction characteristics [1]. No matter any link of construction, it will consume a lot of resources. Green building advocates the use of new materials and processes, which can ensure the harmony and unity of people and society. The core of BIM Technology lies in informatization, which can create a core database based on green buildings, build an evaluation system, eliminate the information island effect, and ensure the quality of building construction. The core value subject of the evaluation is dynamic [2]. Only based on BIM Technology, we can achieve accurate evaluation and maximize the utility of BIM Technology in the building evaluation system.

2. Research Background

2.1. Green Building and Construction Industry. In today's increasingly severe situation of energy and environmental problems, the state adheres to the concept of green development. "Green development + ecological priority" has become the general direction of China's economic development [3]. As a pillar industry of the national economy, the construction industry has a huge negative impact on resources and environment. As shown in Figures 1(a)–1(c), the situation of building energy consumption is serious [4]. Therefore, developing green buildings; thoroughly implementing the concept of "four sections and one environmental protection"; providing healthy, comfortable, and efficient living space and living environment for mankind; and realizing the harmonious coexistence between man and nature are the fundamental ways to effectively improve China's living environment, reduce building energy consumption, solve energy

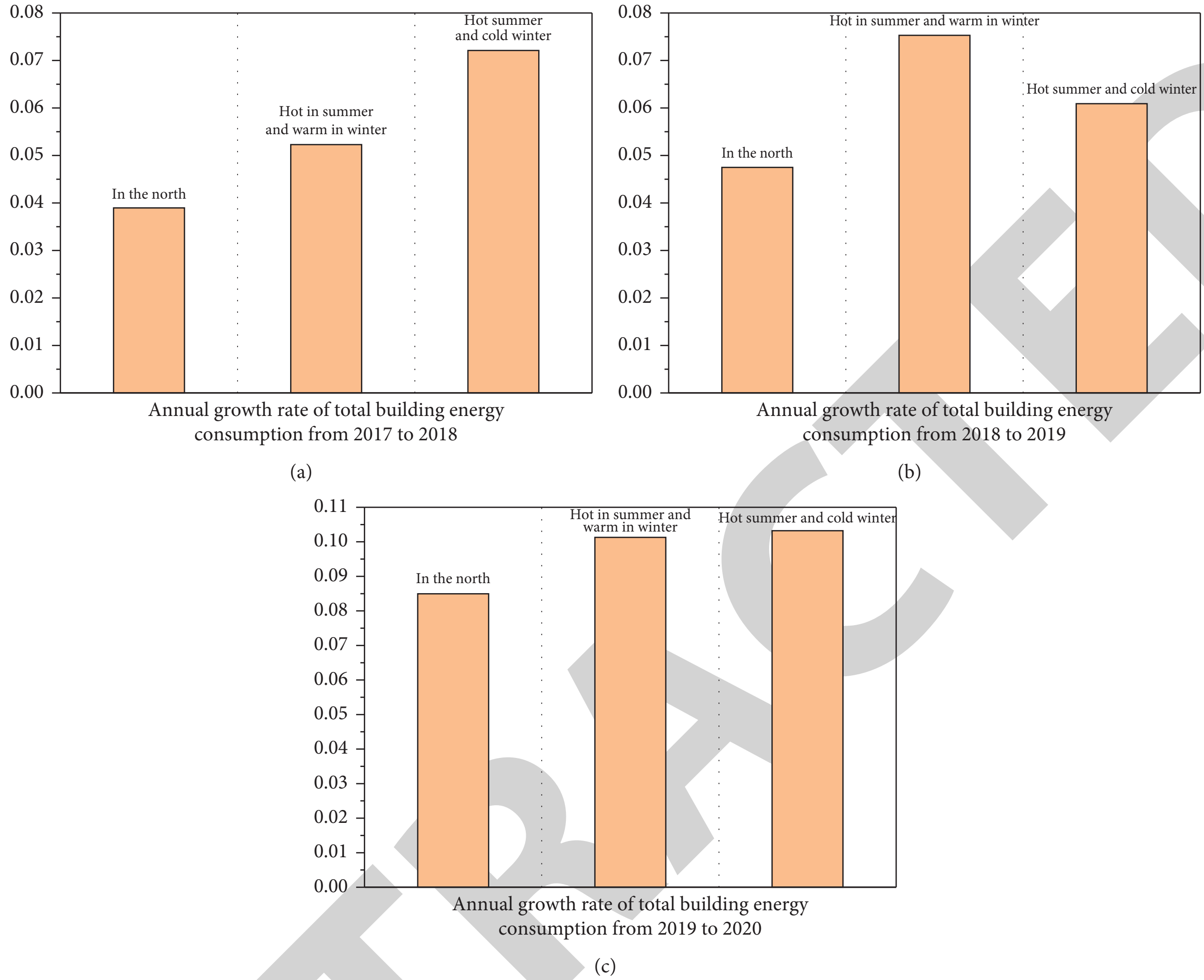


Figure 1: Average annual growth rate of total building energy consumption from 2017 to 2020. (a) 2017–2018. (b) 2018–2019. (c) 2019–2020.

problems, and realize the sustainable development of the construction industry [5].

2.2. BIM and Construction. The promotion of information technology in China's construction industry is slow, and the upgrading and transformation of industrial structure is seriously restricted [6]. Since the 1940s, most reports from the government and professional institutions show that the construction industry has the characteristics of decentralization, the participants lack efficient communication, the construction process lacks formal procedures, and the structure is chaotic [7]. Promoting industrialization through informatization can be an important way to transform and enhance the traditional construction industry. It can be seen that changing the way of information sharing and transmission in the construction industry is one of the important problems to be solved in the construction industry [8].

2.3. Green Building and BIM. Green building has the characteristics of large amount of information and complex types [9]. As an interactive platform for efficient transmission of information and management data, BIM Technology ensures the unity of information in the stages of building planning, design, and construction with digital technology. The application of BIM in green building can effectively eliminate the problem of "information island" between various stages and realize the balance between green management and green technology. It is embodied in the seamless connection of space dimension and time dimension. In terms of time dimension, the collection, sorting, summary, and analysis of green building design information can be realized through the parametric model of BIM. At the same time, it provides an intelligent data platform for later project post evaluation and even operation evaluation, so as to provide strong data support for green building performance simulation and be used in green building environmental performance analysis [10]. As shown in Figure 2, it

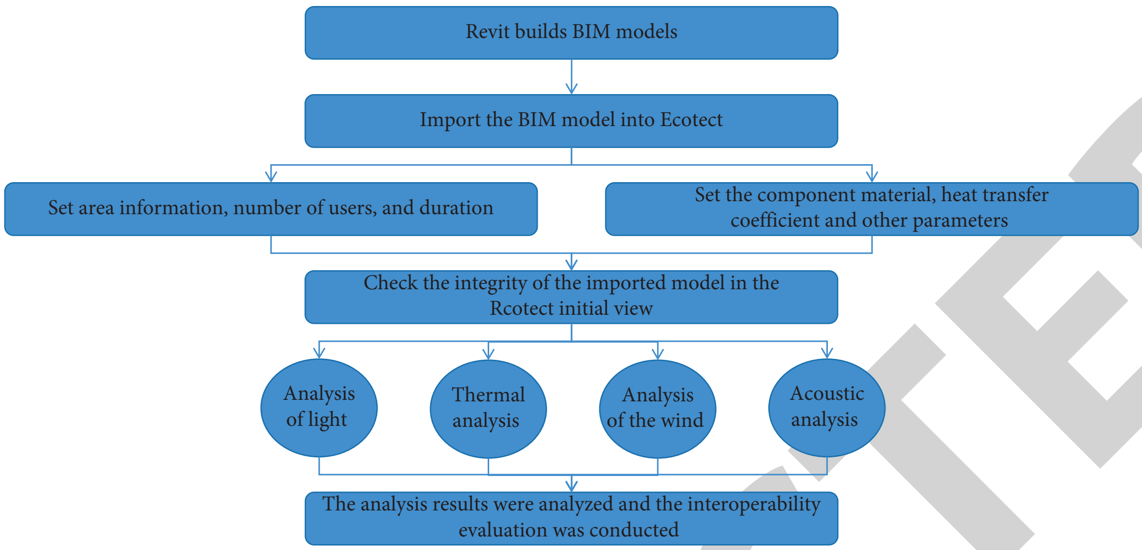


Figure 2: BIM simulation analysis data transmission process.

can realize the effective link between technology and building evaluation system.

3. BIM and Green Building Design

3.1. Life Cycle Theory and Green Building. Life cycle is important in construction engineering. On the premise of ensuring the construction objectives and meeting the mutual coordination between the functions of the building itself and the surrounding environment, it is necessary to ensure that all links of the life cycle are scientifically supervised [11]. The construction life cycle needs to be accurately controlled, so as to save the overall construction cost and ensure sustainability through technology and cost. At present, the energy consumption of building construction makes the ecological environment problem serious. In order to better implement the sustainable development strategy, it is necessary to change the previous development concept and promote the long-term development of the industry from the aspects of building ecological environmental protection and full cycle life [12].

3.2. Advantages of BIM Technology in Green Buildings. As an important supporting tool of informatization in the construction industry, BIM can realize technology integration with the help of software and ensure the engineering informatization level to a great extent and improve the economic benefits of construction projects [13]. BIM Technology can build an evaluation system around the whole building, analyze and integrate the existing problems, and obtain an ideal model. In addition, without the support of BIM Technology, green building analysis will only become a visual expression, which lacks some design relevance in practical engineering. It is not just a professional problem to modify the design of the problems existing in the analysis results of green buildings [14]. Usually, all majors should deepen the design of the problems that do not meet the standards. In this case, the integration of computer-aided simulation and architectural design will be more intuitive. BIM tool can not only provide more in-depth quantitative analysis for green buildings but also solve design problems more accurately, intuitively, and efficiently.

4. Theoretical Basis of BIM Technology Application Value in Green Building

By combing the theoretical basis of the whole process application value of BIM in green building, green building, BIM, value and its value evaluation, and the definition of the whole process application scope of BIM, this paper expounds the development and evolution of various theories in detail, pays attention to the systematicness of theoretical research, and lays a theoretical foundation for BIM in green building to build an evaluation system and ensure building quality [15].

4.1. Green Building Concept. Green building is a kind of high-quality building. Based on the people-oriented concept, it can better deal with the relationship between man and nature. It can protect the environment to the greatest extent, so as to meet people's pursuit of life [16]. Compared with ecological buildings and sustainable buildings, green buildings save resources to the greatest extent in their life cycle, so as to protect the environment and reduce pollution. Ecological architecture is the architectural system of group and single buildings and their surrounding environment based on the principle of ecology. Sustainable architecture is an environmental protection living architectural culture that needs to be completely shaped from the Earth scale such as ozone layer destruction, global warming, and biodiversity. Sustainable buildings, ecological buildings, and green buildings are energy-saving buildings. The three building types have different emphases. The relationship and comparison between them are shown in Figure 3.

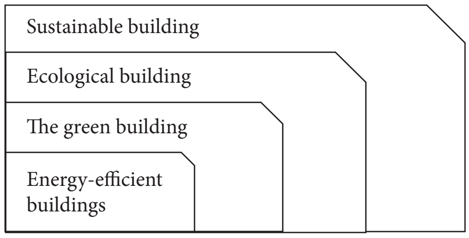


Figure 3: Relationship among sustainable buildings, ecological buildings, green buildings, and energy-saving buildings.

4.2. Value Evaluation Method. Value evaluation, also known as value cognition or evaluative cognition, refers to the cognition and evaluation of the value subject on the value attribute and value relationship with the value object. The value evaluation method refers to the means, tools, and methods used in the value evaluation of complex problems or integration problems. The main research object of this paper is the green building project. Considering its diversity, integration, and uncertainty, this paper selects the following four value evaluation methods to quantitatively evaluate the whole process application value of BIM in green building.

4.2.1. Factor Analysis. Factor analysis (FA) is based on the idea of multivariate statistical analysis. Its core is to find the relationship between variables and extract class variables through data dimensionality reduction. Its advantage is to reflect the evaluation content with a small number of comprehensive public factors after refining and processing, which is simple and objective.

(1) Factor Analysis Model. For factor analysis modeling, it is necessary to set up n sample companies and m indicators to build an original data matrix with $n * m$ order:

$$X = \begin{bmatrix} X_{11} & X_{12} & \ldots & X_{1m} \\ X_{21} & X_{22} & \ldots & X_{2m} \\ \ldots & \cdots & \ldots & \ldots \\ X_{n1} & X_{n1} & \ldots & X_{nm} \end{bmatrix}. \tag{1}$$

When m is large, it is necessary to reduce the dimension of the formula and extract P common factors for modeling and analysis. Assuming $Y = (Y1,\ Y2,\ Yp)$, the linear relationship is expressed as:

$$\begin{aligned} Y_1 &= a_{11}X_{11} + a_{12}X_{12} + \ldots a_{1P}X_{1P}, \\ Y_2 &= a_{21}X_{21} + a_{22}X_{22} + \ldots a_{2P}X_{2P}, \\ Y_m &= a_{m1}X_{m1} + a_{m2}X_{m2} + \ldots a_{mP}X_{mP}. \end{aligned} \tag{2}$$

(2) Factor Analysis Method. The indicators are treated in the same direction, and the original treatment is treated in the following ways:

$$X_{ij}' = \left|X_{ij} - k\right|, \tag{3}$$

where k is the mean value of the original data. Indicators need to be selected for indicator processing, but there will be differences in indicators, resulting in reduced comparability. Select:

$$X_{ij}' = \frac{X_{ij} - \overline{X}}{S_i}. \tag{4}$$

The correlation index verification will use SPSS software to extract the principal components and extract the common factors for the initial eigenvalues and variance contribution rate of various data [17]. After rotating the common factor, if the clear meaning of each factor cannot be obtained in a single rotation, the rotation process can be repeated until the ideal result is obtained, and the obtained common factor is named. The specific steps are shown in Figure 4.

4.2.2. Basic Structure of Convolutional Neural Network. Convolutional neural networks are derived from feedforward neural networks and generally include an input layer, a convolution layer, a pooling layer, an excitation layer, a fully connected layer, and an output layer. As shown in Figure 5, after inputting the original image data, after convolution, pooling, nonlinear activation function mapping, and other operations, the high-level semantic information is abstracted layer by layer, and then through the full connection layer and output layer, the image category is finally obtained.

Activation function is a nonlinear mapping, which is used to simulate complex nonlinear functions and enhance the expression ability of network model. The output of convolution layer and full connection layer usually needs to connect the activation function. ReLU function, Sigmid function, and Tanh function are nonlinear activation functions widely used in neural networks.

The expression formula of Sigmid function is: $\text{Sigmid}(x) = 1/(1 + e^{-x})$. The output response range of Sigmid function is [0, 1], which corresponds to neurons respectively. Tanh function is based on Sigmid function, and the output response range is [−1, 1]. In order to avoid gradient saturation effect, ReLU is introduced into the formula:

$$\begin{aligned} \text{ReLU} &= \max\{0, x\} \\ &= \begin{cases} 0, & x \langle 0, \\ x, & x \geq 0. \end{cases} \end{aligned} \tag{5}$$

As shown in Figure 6, based on the mapping from input to output of the algorithm, the error can be transmitted to the forward layer of the network by using directional propagation, and the actual value of the model can be obtained through calculation.

By combing the theoretical basis of BIM application value in green buildings, the basic theoretical system of BIM application value evaluation system in green buildings includes green building theory, value theory, the definition system of BIM application scope, and value evaluation methods. This paper makes in-depth research on the above contents; analyzes the existing relevant research results; and expounds the core concepts, development, and evolution of various theories and their application in the field of

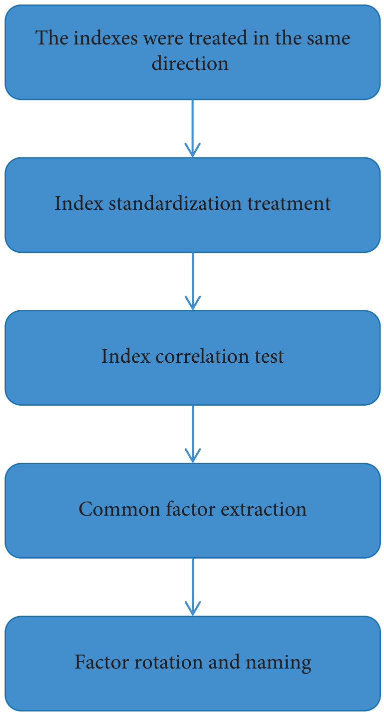


Figure 4: Steps of factor analysis.

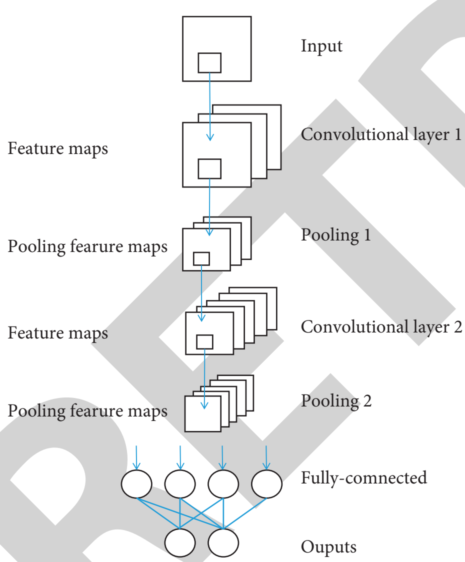


Figure 5: Basic structure of convolutional neural network.

construction engineering [18]. The relevant research results of this chapter can lay a theoretical foundation for the quantification of the whole process application value, economic benefit analysis, and value promotion of BIM in green building.

5. BIM Technology Green Building Assessment

5.1. Green Building Assessment. Green building evaluation system is the basis for the implementation and promotion of green building. A perfect evaluation process and system can better promote the development of green building and realize the application of green building from theory to practice. The development process of foreign green building evaluation is mainly divided into three stages: the first stage is the overview and evaluation of the building itself and construction technology; the second stage is the software simulation and evaluation of the building environment design stage, such as lighting, wind environment, building heat; the third stage is to integrate the concept of the whole life cycle of the building, and comprehensively review and pre evaluate the building environment, the surrounding environment, and the operation status of the building. Through these three stages of the evaluation process, each country has compiled similar evaluation tools on its basis and further improved the evaluation system through simulation auxiliary tools and network information technology [19]. The content of green building evaluation involves many fields, a wide range of specialties, and complex contents, so it is necessary to formulate scientific and advanced evaluation standards and systems, form an integrated evaluation system, play an evaluation role in actual projects and lay a foundation for improving building quality.

5.1.1. Contents of Green Building Evaluation Indicators. Due to the different conditions of various countries and the different understanding and requirements of architecture and environment, the content and weight of green building evaluation are also different. At present, based on the comprehensive contents of national evaluation systems, they can be divided into five categories of indicators, namely:

(1) The planning includes project site selection, surrounding environment, and traffic conditions.

(2) The design comprehensively considers the building performance indicators and architectural design characteristics.

(3) The impact of environmental buildings on the surrounding environment, including considering the pollution of buildings to water resources, land and air, and the damage to the surrounding ecology [20].

(4) Social and economic benefits, the impact of social development and construction, and so on.

(5) Healthy indoor environment quality, living comfort, and so on.

5.1.2. Green Building Evaluation Mechanism. First, determine the type of building and the use of the local environment (including the local climate) according to the evaluation of the project's development conditions, the construction practice, and other factors. The second is to determine the evaluation criteria for the above established index projects. Generally, the current national or regional norms and recognized international standards are used as the most important references and guidelines. At the same time, in some evaluation tools, the evaluation criteria are also set as a scale to dynamically reflect the best level and latest progress of regional time. Finally, the relevant projects are evaluated according to the standards.

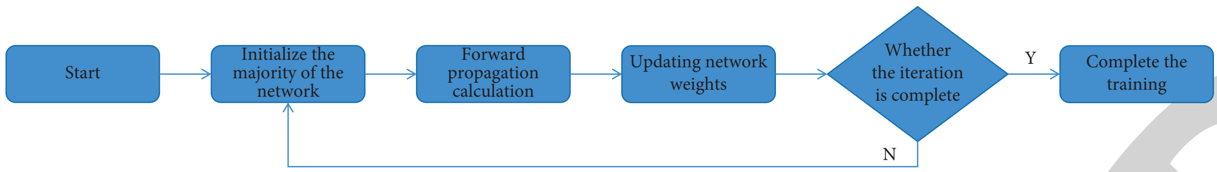


Figure 6: Training flow chart of convolutional neural network model.

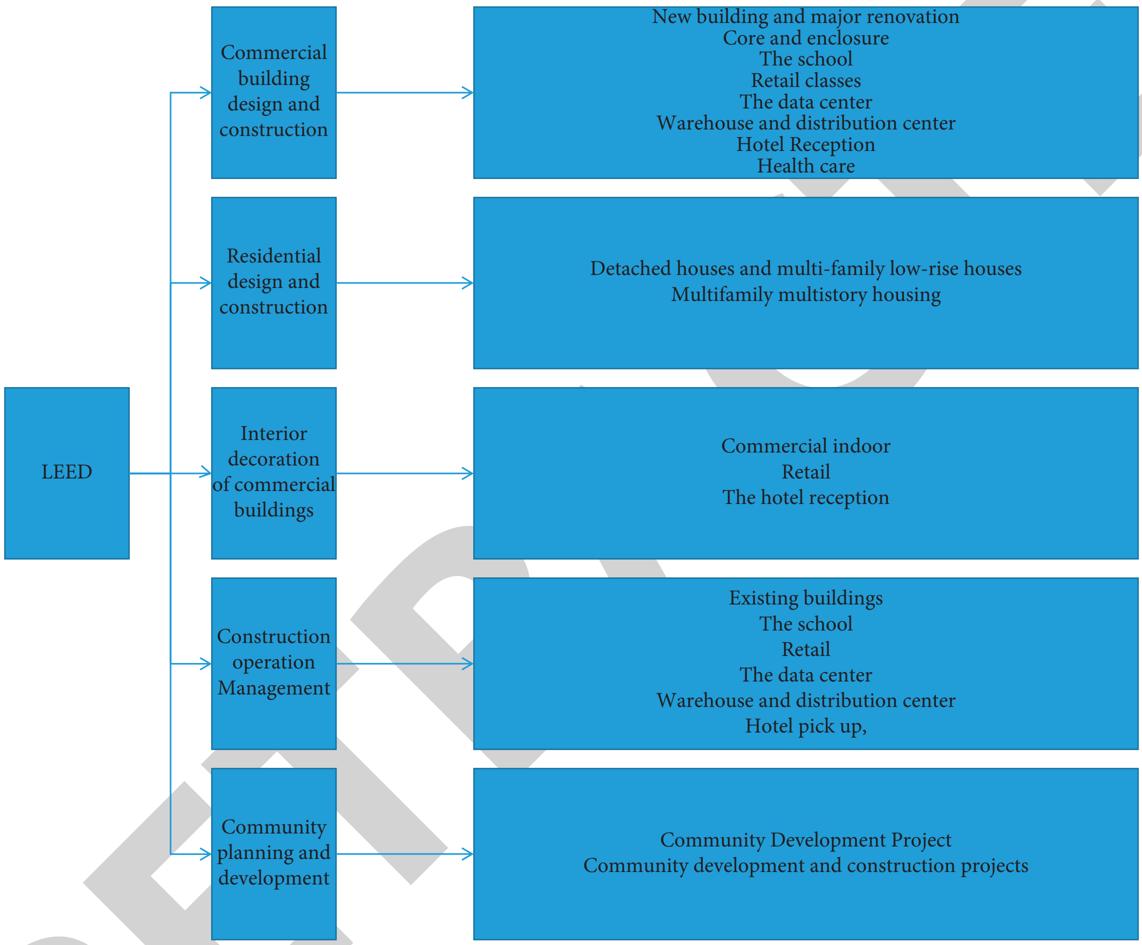


Figure 7: LEED evaluation standard classification.

5.1.3. Evaluation Process. The existing data of the building are substituted through the internal calculation method of green building evaluation, and the relevant design, planning, management, and operation data of the building are input, and then combined with its detailed documents to form the initial data [21]. The second is the comprehensive score, which is scientifically scored by special reviewers according to the green building evaluation standard manual, and the final score is obtained by weighted accumulation. Finally, according to its score, determine the level and issue the certification certificate.

5.2. Green Building Evaluation System

5.2.1. Foreign Green Building Evaluation System. Among many global evaluation systems, the LEED evaluation system of the United States is a representative system. This system is described below. LEED evaluation criteria include many types, covering not only the whole type of buildings but also different stages of the whole life cycle of buildings. Small categories are also set in different categories, and special scoring points are given according to the characteristics of building use functions, as shown in Figure 7.

LEED has been developed to version 4.0. In LEEDv4.0; there are nine evaluation indicators, each of which contains 1–20 evaluation subitems [22]. The specific scores are obtained according to the evaluation of subitems, and the specific evaluation information can be obtained by accumulating the scores of all subitems as shown in Table 1.

LEED does not adopt the weight system. For LEED certification of construction projects, it is necessary to accumulate the scores of various evaluation indicators to obtain the total score and judge the building grade according to the range of scores (see Table 2 for details).

Table 1: Score of LEEDv4.0 evaluation classification and number of scoring terms.

Index		Integrate project planning and design	Site selection and transportation	Water efficiency	Energy and atmosphere	Materials and resources	Indoor environmental quality	Innovate	Geographical priority
Commercial building design and construction	New buildings and major renovation	1	16	11	33	13	16	6	4
	School	1	20	12	31	13	16	6	4
	Data center	1	16	11	33	13	16	6	4
	Medical care	1	9	11	35	19	16	6	4
Residential design and construction	Independent residence and multifamily stratum residence	2	15	12	38	10	16	6	4
	Multifamily multistorey residence	2	15	12	37	9	18	6	4
Interior decoration of commercial buildings	Commercial indoor	2	18	12	38	13	17	6	4
	Retail	2	18	12	38	14	16	6	4
	Hotel reception	2	18	12	38	13	17	6	4
Construction operation management		-	15	12	38	8	17	6	4

Table 2: LEED certification level.

Fraction	40–49	50–59	60–79	80–110
Grade	Certification level	Silver grade	Gold grade	Platinum grade

5.3. Domestic Green Building Evaluation Indicators. In 2006, the Ministry of housing and urban rural development issued the green building evaluation standard GB/T 50378–2006. By the end of 2014, 2538 green building projects had been evaluated, with a total construction area of 290 million m^2. This standard is an important part of guiding the practice of green building in China. However, since the 12th Five Year Plan, China's green buildings have developed rapidly, and the evaluation standard for green buildings (GB/*T* 50378–2006) cannot fully meet the needs of green building evaluation. Therefore, the evaluation standard for green buildings (GB/T50378-2014) (hereinafter referred to as the standard) was prepared by China Academy of Building Sciences in conjunction with relevant units and implemented on January 1, 2015.

(1) Due to the different use functions of the buildings to be evaluated, there are great differences in their resource consumption and impact on the environment. The standard adopts specific provisions to divide the evaluated buildings into residential buildings and public buildings and also defines the evaluation of comprehensive single buildings [23].

(2) In the evaluation stage, the evaluation stage of the standard is mainly composed of two stages: design and operation. The design evaluation is carried out after the construction drawing is reviewed and approved, focusing on the green design and effect of the building; the operation evaluation is carried out one year after the building is put into use, focusing on the implementation and actual effect of building green measures. The division of stages is in line with the objective process of the construction industry and the actual needs of the evaluation work. The score calculation method of domestic green building evaluation indicators is to set corresponding weights according to the project scores specified in the standard and the importance of various indicators. Calculation formula of total score:

$$\sum Q = w1Q1 + w2Q2 + w3Q3 + w4Q4 + w5Q5 + w6Q6 + w7Q7 + Q8, \tag{6}$$

where Q1-7 = (actual score value/actual full score) ∗ 100 points. According to the actual full score = theoretical Full Score - ∑ the score of articles not participating in the evaluation = ∑ the score of articles participating in the evaluation, and the weight is taken according to Table 3.

5.4. Comparison at Home and abroad. American LEED and domestic evaluation standards can find that American LEED has a more sound evaluation system for buildings, while domestic evaluation standards cover few contents. Moreover, in terms of the provisions of the evaluation object, LEED has a complete system, which can ensure the connection between various elements from technology to architecture to the overall area. However, the domestic green building evaluation system lacks relevant standards, and its application and promotion are poor. It is necessary to draw lessons from foreign standards and formulate a domestic integrated evaluation system.

6. Research on the Application of BIM Technology in Green Building Evaluation

6.1. Green Building Evaluation Process

6.1.1. Traditional Green Building Evaluation Process. The traditional green building evaluation is mainly composed of the application unit and the evaluation unit. The application unit issues the relevant preparation materials; the evaluation unit invites the experts of relevant disciplines for evaluation; and finally, the evaluation experts determine the evaluation results, publicize, and issue the registration certificate and logo according to the evaluation results. The traditional green building evaluation process is through offline application, and expert review is conducted according to the application materials composed of design drawings, on-site monitoring data, and calculation and analysis reports assisted by three-dimensional models [24]. The specific evaluation process is shown in Figure 8. There are some disadvantages in the evaluation process, such as cumbersome evaluation process, single evaluation method, poor model auxiliary effect, and so on.

6.1.2. Green Building Evaluation Process Based on BIM Technology. The green building evaluation based on BIM Technology can greatly optimize the evaluation process and integrate all building related information as the basis of evaluation. The whole process of green building evaluation can be completed with the help of BIM systematic information integration platform, and the whole process will become more compact and efficient. After using BIM Technology, all information related to the evaluation will be integrated into the BIM model. Both sides of the evaluation can analyze and process the data through the BIM model to reduce the evaluation error and improve the accuracy. The green building evaluation process based on BIM is shown in Figure 9.

6.2. Information Exchange of Green Building Evaluation in BIM Environment

6.2.1. Data Interaction Provided by IM Technology. The standard specifies the specific requirements and standards for buildings. From the development trend of information integration, although the models of each stage are established based on the same project, the single model cannot realize the information processing in the whole project cycle.

Table 3: Weight of various evaluation indexes of green buildings.

		Land saving and outdoor environment $w1$	Energy conservation and energy utilization $w2$	Water saving and water resources utilization $w3$	Material saving and material resource utilization $w4$	Indoor environmental quality $w5$	Construction management $w6$	Operation management $w7$
Design evaluation	Residential building	0.21	0.24	0.20	0.17	0.18	—	—
	Public buildings	0.16	0.28	0.18	0.19	0.19	—	—
Operation evaluation	Residential building	0.17	0.19	0.16	0.14	0.14	0.10	0.10
	Public buildings	0.13	0.23	0.14	0.15	0.15	0.10	0.10

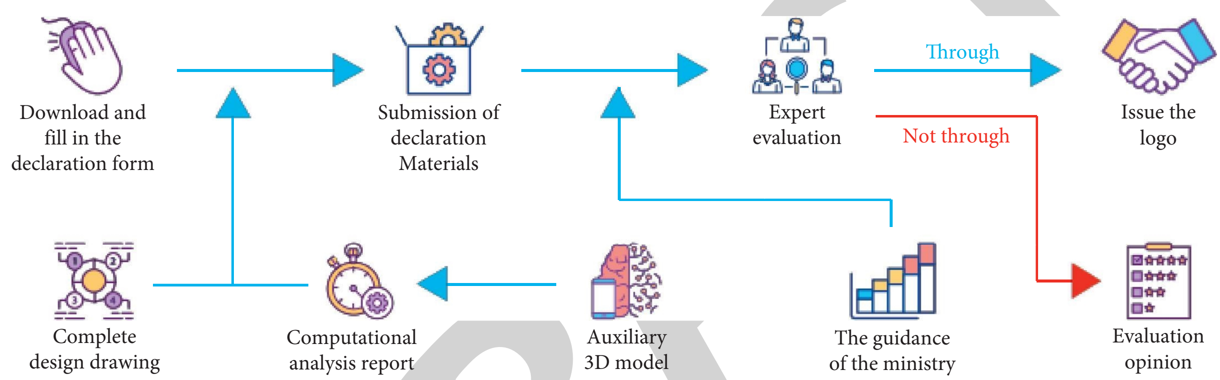


Figure 8: Evaluation process of traditional green building.

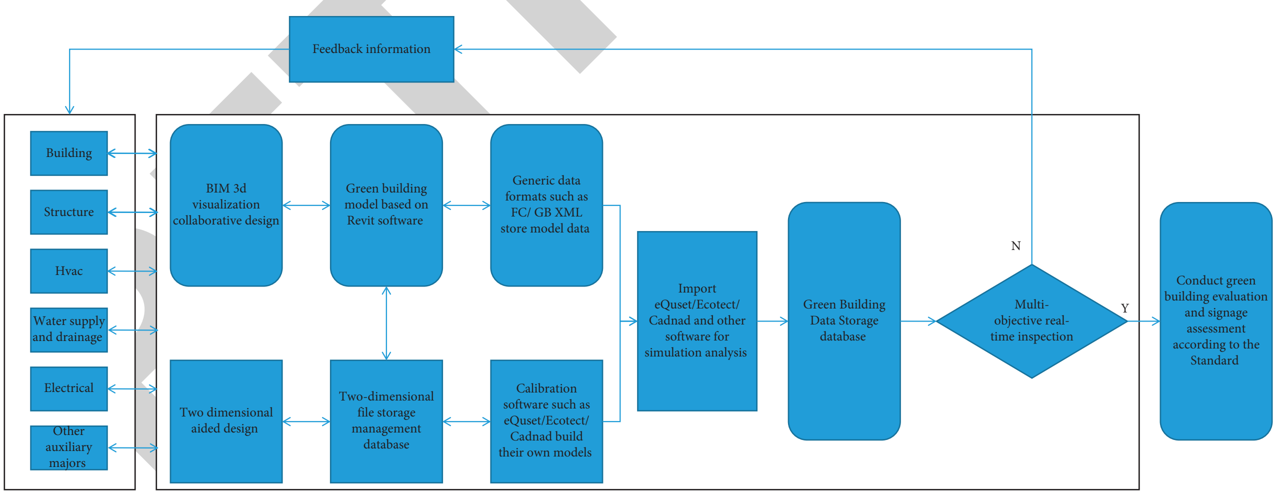


Figure 9: Green building evaluation process based on BIM.

If the model can realize data interaction, it can more effectively reduce repeated labor and improve the efficiency of information utilization. The data interaction between BIM and software can ensure the timeliness of information through the analysis of building environmental performance. BIM Technology can also provide various simulation data according to the building to create a real green building.

6.2.2. Collaborative Work of Project Participants in BIM Environment. "Owner," "design," "construction," and "property" as the four participants in the whole life cycle of green building, are important parts of the implementation of green building and shoulder different tasks and responsibilities. Due to the different objectives of the four participants, the degree of participation in the project is also different. Therefore, the collaborative work between the project participants is particularly important. Based on the collaborative work of participants, BIM Technology is used to build integrated management, design each stage of the building, and ensure that there are more scientific schemes in the stages of building decision-making, design, construction, and operation and maintenance. Through the application of BIM in green building, it is ensured that the owner can linkage with the design unit in the decision-making stage, simulate the site situation, and conduct feasibility analysis. Before the implementation of the project, use the simulation technology to complete the virtual operation of the whole process of the building and further carry out the layout of streamline organization and functional zoning [25]. After the construction party is determined in the later stage, it can analyze the key and difficult points of construction through the simulation effect of BIM model and feed back to the design unit to improve the implementable design and reduce the risk of change. Through BIM information management platform, the three participants can monitor and inspect various indicators of green buildings before construction, eliminate potential safety hazards, and reduce construction risks. The linkage management of engineering information and project decision-making ensures that all parties can quickly obtain accurate information and make correct response, enhance interactive management, and carry out rational allocation and timely dynamic management of production factors.

6.2.3. Integrated Management of Green Building Evaluation Information Based on BIM. Green building projects need to be completed by multiple disciplines. Therefore, the integrated management of project information of various disciplines can effectively improve work efficiency, and BIM can provide good technical support for the integrated management of green building evaluation information. The evaluation information of green building can be divided into decision information, design information, construction information, and operation information according to the construction stage. The green building evaluation information based on BIM is shown in Figure 10.

In the decision-making stage, the owner should have a full understanding of the site selection, surrounding environment and internal conditions of the project, and BIM Technology can assist in green building ecological environment simulation, preliminary modeling, and preliminary planning, so that the owner can have a preliminary budget for the cost and construction period of the project. In the design stage, the main task is to transform the owner's construction concept into a specific description of the implementable object, and its design documents are the basis for later implementation. The key control point of green building evaluation in this stage is to integrate all building information management; extract comprehensive, clear, and accurate green building design scheme information; and import such information into BIM information management platform database for later use by all participants. Green building design is an integrated design process

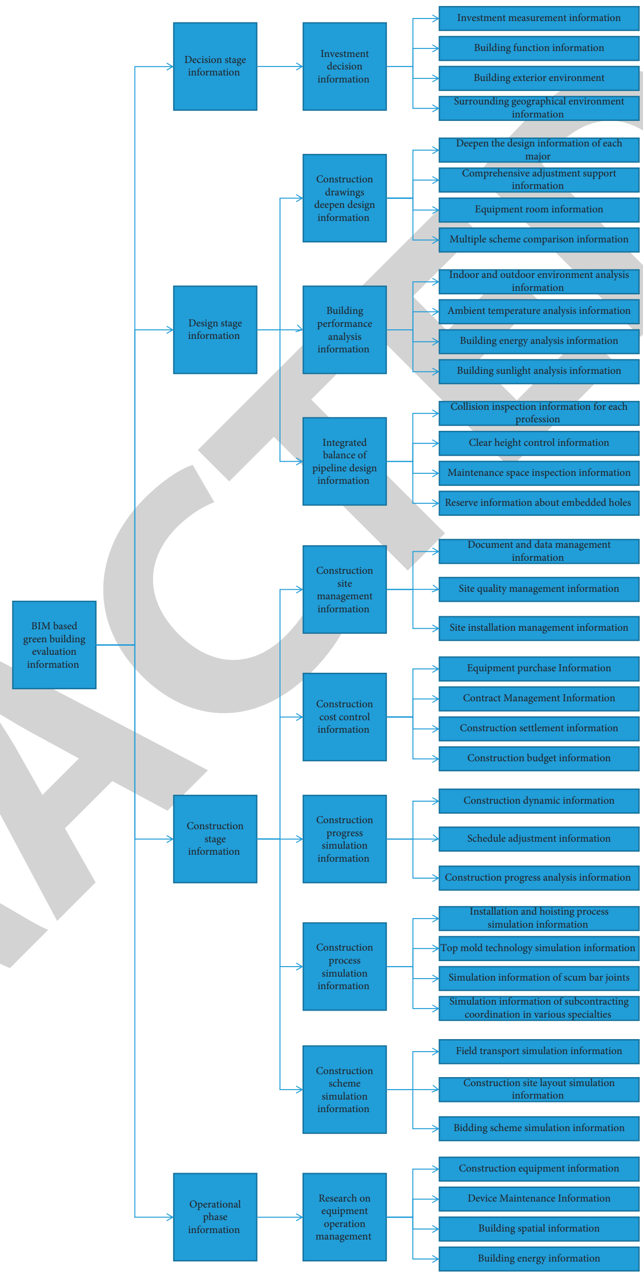


Figure 10: Green building evaluation information based on BIM.

integrated by multiple disciplines, and BIM model can unify information on the data platform and link all parties to work together to form the final three-dimensional collaborative management platform. BIM model carries all kinds of accurate and detailed project information. It is the operation basis of green building simulation analysis software and the accurate guarantee of analysis results. The parameterization and component library of BIM can effectively feed back the later design results, making the analysis results more representative.

In the construction stage, the main task is to transform the abstract integrated information into concrete green buildings. This stage has many uncertain factors, such as heavy workload, complex process, many organizational interventions, and so on. In order to eliminate adverse conditions, BIM Technology can realize virtual simulation sustainable management according to construction scheme, construction progress, construction quality, and construction cost and ensure more efficient information management. The split management concept of the construction unit will be transformed into the overall information sharing management based on BIM Technology to realize the emerging information management mode of information integration. In the operation stage, the contents involved include facilities and equipment management, energy management, monitoring management, space management, and other comprehensive management. The operation model is established by combining the data with the model, and the BIM operation model is used as the main basis for the daily operation and maintenance management of green buildings.

7. Conclusion and Prospect

This paper puts forward the viewpoint of using BIM Technology in green building evaluation, summarizes the shortcomings of green building evaluation system at this stage, and discusses the feasibility and necessity of BIM application in green building evaluation. This paper compares and analyzes the green building evaluation system at home and abroad from the aspects of evaluation object, evaluation process, evaluation method, and clause setting and puts forward modification suggestions for the integrated evaluation system of green building in China. This paper comprehensively expounds the application of BIM in green building evaluation, forms a complete research route, puts forward the green building evaluation process based on BIM, takes the evaluation information interaction under BIM environment as the starting point, expounds the necessity and specific methods of realizing collaborative work, and analyzes the application of BIM in specific evaluation provisions. Verify the theoretical analysis through actual cases. Based on BIM model and various analysis software as auxiliary tools, select representative green building provisions for statistical analysis and simulation analysis, test the effect of applying BIM to green building evaluation, and complete the evaluation of relevant provisions. From the perspective of application value, the application of BIM in the evaluation process shows that it plays an important role in improving building performance. However, as a new technology, it cannot guarantee the popularization and application of technology in the industry, and there are certain social challenges in the development process.

The characteristics of green building determine that it needs multistage and multidisciplinary building data information. In the future, a multidisciplinary collaborative work system covering all building information can be established to divide the BIM model for green building evaluation, construct a full cycle evaluation system, and conduct different in-depth research on the construction of various disciplines of green building. The application of this paper should be limited in length, lack of building types, and cannot perfect the green building evaluation system. However, we believe that with the high development of domestic informatization, various evaluation methods and relevant provisions and standards of green building evaluation standards will be further improved. In the future, we can try to introduce artificial intelligence and other technologies into green building evaluation to further improve the green building evaluation system.

Data Availability

The labeled data set used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported by the Hebei Agricultural University.

References

[1] C. Liu, “Energy consumption simulation of green building based on BIM system and improved neural network,” *Journal of Intelligent and Fuzzy Systems*, vol. 29, no. 2, pp. 1–12, 2021.
[2] W. Sun, “Research on the construction of smart tourism system based on wireless sensor network,” *Mathematical Problems in Engineering*, vol. 2021, no. 18, 8 pages, Article ID 9950752, 2021.
[3] Y. Sun, X. Yuan, and Q. Chen, “the construction and empirical research on the dynamic evaluation model of university science and technology output,” *IEEE Access*, vol. 9, no. 99, pp. 57280–57290, 2021.
[4] E. Kamel and A. M. Memari, “Review of bim’s application in energy simulation: tools, issues, and solutions,” *Automation in Construction*, vol. 97, pp. 164–180, 2019.
[5] H. Bapat, D. Sarkar, and R. Gujar, “A sustainable approach to reduce embodied and operational cooling energy for an elevated metro rail station of ahmedabad, India, using building information modelling (bim) and factor comparison method,” *Journal of the Institution of Engineers: Series A*, vol. 103, no. 1, pp. 115–128, 2022.
[6] Q. H. Han, “Research on the construction of cold chain logistics intelligent system based on 5g ubiquitous internet of things,” *Journal of Sensors*, vol. 2021, no. 1, 11 pages, Article ID 6558394, 2021.
[7] R. Abdulla, S. Duraikannan, and V. Thiruchelvam, “Integrated BIM and energy tools with building certification system for sustainable building design and energy certification,” *Solid State Technology*, vol. 63, no. 1s, pp. 981–992, 2020.
[8] N. Yang, Y. Huang, B. Dong, P. Xin, and H. Li, “Research on the joint planning method of electricity-gas integrated energy system based on multi-agent game,” *Zhongguo Dianji Gongcheng Xuebao/Proceedings of the Chinese Society of Electrical Engineering*, vol. 39, no. 22, pp. 6521–6533, 2019.
[9] S. Gao and L. K. Bhagi, “Design and research on caddcam system of plane based on nc machining technology,” *Computer-Aided Design and Applications*, vol. 19, no. S2, pp. 64–73, 2021.
[10] P. Wang, S. Gao, L. Li, S. Cheng, and H. Zhao, “Research on driving behavior decision making system of autonomous

Hindawi
Mobile Information Systems
Volume 2022, Article ID 1950877, 8 pages
https://doi.org/10.1155/2022/1950877

Research Article

Investigation and Analysis of Classroom Teaching Quality in Universities in an Interval Intuitionistic Fuzzy Environment

Quansong Qi

Southwest University of Political Science and Law, Chongqing 401120, China

Correspondence should be addressed to Quansong Qi; qiquansong@swupl.edu.cn

Received 16 May 2022; Revised 7 June 2022; Accepted 30 June 2022; Published 18 July 2022

Academic Editor: Xuefeng Zhang

In recent years, China's economic strength has been developing rapidly. At the same time, people's daily living standards have also been increasingly rising. In this context, the authorities are paying more and more attention to the training of students as well as talents. As a result, the quality of classroom teaching in universities has become a major concern. As an essential evaluation tool for teaching quality, evaluation of classroom teaching has been widely recognised and implemented. With the rapid development of computer technology and information technology, teaching evaluation based on various algorithms has been widely used, thus becoming an increasingly important teaching evaluation method. Thus, classroom teaching assessment has become one of the hot issues discussed in the field of education both at home and abroad. A great number of universities are now fully aware that the effectiveness of teachers is central to their survival and development and can have a direct impact on the training of students. The criteria for evaluating the effectiveness of teaching are highly subjective, but at the same time, there is a considerable degree of objectivity. As a result, the application of more scientific methods and computational data can make the results of classroom assessment more reliable. However, there is a lack of reliable reference data for evaluating the quality of classroom teaching in higher education. In addition, it is difficult to develop an evaluation data management and evaluation management model based on multiple data sources, which may cause a negative effect on the impartiality and objectivity of evaluation results. What is worse, many teachers do not have enough time to interact with their students. After all, most teachers in universities now have a heavy teaching load, and much of their time is spent on lesson preparation. Hence, they usually lack active interaction with their students. In the current education system, the objectives of higher education are mainly test-taking, and practical application is lacking. Therefore, based on the existing teaching basis, how to improve the effectiveness of university classroom teaching and improve the overall quality of students on the basis of test-taking objectives will be a key issue in teaching research. In order to solve this problem, this paper mainly researches the evaluation of classroom teaching quality in universities based on the interval intuitionistic fuzzy theory. This study is based on a detailed analysis of the needs of teaching quality assessment management, in terms of comprehensiveness of assessment and data sources. Then, the teaching evaluation system is designed and developed based on an analysis of the roles of students and teachers. This system can be applied to speed up the efficiency and fairness of teaching evaluation in universities, thereby greatly improving the management of teaching in universities.

1. Introduction

In recent years, people's quality of life has become better and better, driven by China's rapid economic development. Against this backdrop, the education of children has become increasingly important to the government and the people [1]. As an essential vehicle for training the future talents of our country, schools carry the dream of a strong country and the high hopes of parents [2]. In the development of schools, the quality of teaching and learning is a direct measure of the quality of the school. To be specific, it can reflect the reputation of the school and has a direct impact on the future of the students [3]. As a result, how to strengthen the management of teaching quality and gradually improve the quality of the teaching team and the level of teaching is the key to cultivating high quality talents in the new century. In the management of teaching, the evaluation of teachers' quality is an essential means of improving teaching quality and is the core of teaching management [4]. There is a growing consensus that schools are only as excellent as the students they produce. As a result, many scholars have begun to study and develop measures to improve the quality of teaching and learning, with the aim of maintaining order and improving the quality of teaching and learning [5, 6]. As the scale of education continues to expand and the number of students increases, the quality of teaching and learning becomes an increasing concern for society. The quality of education is crucial to the survival and development of the school and is the centrepiece and constant theme of all schoolwork [7].

However, there are still many problems with the existing university teaching quality evaluation system in China. Although many universities have discussed the effectiveness of classroom teaching, the question of how to build a fair and objective teaching evaluation system is still an essential issue for most universities [8]. To be specific, the following problems exist in the existing system for evaluating the quality of classroom teaching. First and foremost, the evaluation indicators fail to reflect the characteristics of each evaluator and therefore make it difficult to realistically reflect the standard of the subject of the evaluation [9]. In the actual evaluation of teaching and learning, most methods focus only on the behaviour and measurability of indicators [10]. However, many indicators and factors cannot be studied quantitatively and are therefore often neglected. Effectiveness evaluation, on the other hand, usually treats some specific indicators as value orientations, and the process is only concerned with whether the indicators are achieved or not [11]. Therefore, such systems can directly lead to evaluation results that make it difficult to reasonably identify the standard of the evaluated object. Second, most systems for evaluating the quality of teaching are fixed and hardly serve as an incentive. Nowadays, most universities have developed a fixed system of relevant teaching evaluation indicators based on the actual situation. For the evaluation of teaching, whether professor, lecturer, or assistant teacher, similar evaluation criteria are applied [12]. This phenomenon leads to a vague selection of indicators for the evaluation of teaching. As a result, the rigidity of the indicators makes it difficult to evaluate the effectiveness of teaching in a way that stimulates and enhances it [13]. In addition, in most universities, teaching quality is still assessed adopting traditional paper-based approaches, which can lead to some drawbacks such as low effectiveness. This paper-based approach to evaluation makes it a tricky project [14]. It can be time-consuming and inefficient, seriously affecting the efficiency of the school and ultimately leading to a formal evaluation of teaching and learning that does not improve outcomes. Therefore, it is necessary to design an excellent teaching evaluation system in order to strengthen the management of teaching quality and to better improve the quality of education in universities.

At present, many universities' teaching quality evaluation systems are based on student grades and practical activities as the core data [15]. Generally speaking, the evaluation indicators for practical activities are the same for all students, so the evaluation of teaching quality is based on students' performance in school. However, this approach cannot truly characterise the quality of teaching and learning [16]. The evaluation data are entered directly by the class tutor and do not go through the online evaluation and scoring process. As a result, the final results of the evaluation are often contested as criteria for teacher evaluation and selection and fail to reflect the principles of fairness and objectivity [17]. It is clear that the assessment of teaching quality is not the kind of assessment that is entered at the end of the semester but rather a dynamic way of managing the performance of teachers and student counselling during the school year [18]. The evaluation process is cumbersome, and the indicators are therefore vague. Generally, at the end of the semester, each academic department will give a token comment as a conclusion to the evaluation of teaching quality [19]. This model is not conducive to a comprehensive evaluation of teaching quality and is not based on a fixed standard, let alone a mature evaluation system. As a result, it becomes formalised and does not constitute a comprehensive evaluation system based on information-based indicators [20]. At present, there is no recognised evaluation system for the quality of teaching in our schools. In addition to the main evaluation, the evaluation criteria for teaching quality should consist of other evaluations, self-evaluations, and comprehensive evaluations of various kinds [21]. This is the only way to better reflect the comprehensive evaluation of teaching quality and to measure and evaluate the quality of a teacher's teaching in a comprehensive manner.

Although some universities have developed software that can be used to evaluate the quality of teaching, these software systems are generally limited to a particular department, such as the Academic Affairs Office, which mainly evaluates the quality of teaching through student performance [22]. To a certain extent, these software systems can be used to evaluate the quality of teaching and provide decision support for teaching quality management. However, the data cannot be shared across departments. If other departments need to know these data, they can only do so by asking the Academic Affairs Office or obtaining the relevant reports [23]. As a result, there is a lack of information and high communication costs between departments, which is not in line with the current requirement for data sharing between the relevant departments of the university. In recent years, the rapid development of computer technology has brought significant development to the information construction of universities. Many computer technologies such as machine learning [24, 25], deep learning [26, 27], and life-cycle assessment [28, 29] have been fully developed. The rapid advances in technology have led to a gradual reduction in the cost of information in universities. A number of universities have implemented digital university systems that enable comprehensive management of all aspects of student registration, student work management and teaching and learning assessment management. These systems can provide a complete set of solutions for the normal teaching and research operations of universities. At the same time, each department can complete the operation of the corresponding modules through the assigned user roles. The application of the system can improve the competitiveness of the university and thus improve the efficiency of the university management.

With the strengthening of teaching evaluation management in overseas universities, more and more university administrators have gained a deeper and more comprehensive understanding of teaching evaluation management [30]. In terms of information systems, they are no longer limited to the management of basic information on teaching and learning assessment. Instead, these systems are designed to meet the needs of the various teaching and evaluation departments in universities and to provide a software platform for teaching and evaluation management in the current context. The management of teaching evaluation in higher education is a dynamic process that will continue to improve as the management mechanisms of higher education change [31]. In particular, with the application of new technologies, such as mobile Internet and mobile phone clients, the original B/S-based system has evolved into a richer and more comprehensive platform. The use of mobile phones for teaching evaluation management is becoming more and more common, changing the way it is managed and thus allowing for flexibility in the office. At present, teaching evaluation offices in universities have corresponding teaching evaluation management systems, which cover all aspects of teaching evaluation management, including teaching evaluation information management, course management, and other process management.

The development of the teaching evaluation system in this paper is based on research of the design models and ideas of excellent teaching systems at home and abroad. The system proposed in this study is designed and developed with the advantages of other systems, and the shortcomings of other systems are avoided, making the system more suitable for the task of teaching evaluation in schools. By developing a teaching evaluation system based on interval intuition fuzzy theory, this research provides a suitable teaching management platform for universities to meet the functional requirements of teaching evaluation and improve the quality of teaching in universities.

2. Technology Related to Evaluation of Teaching Quality

2.1. Net Platform. The .Net platform is a complete application development solution based on a component-based approach. Although the development languages applied are divergent, the base class libraries are the same. To be specific, applications developed in each language can call on a common library of components to implement the underlying functionality. In this case, the .Net platform-based applications therefore have the advantage of sharing data and technical frameworks. By using the .Net platform for development, the technical staff involved can significantly shorten the development cycle and reuse different components, thus increasing the efficiency of system design and development.

The .Net platform defines a set of enterprise-level standards for application development. This standard allows technicians to quickly develop applications in their own programming language. In other words, the .Net framework enables the classic layering of data, business, and functionality, which can greatly improve the ease of operation and scalability of later systems. For example, when the business changes, the logic definition of the business layer can be modified. In this case, when the system database platform is changed, only the data interface needs to be adjusted in the data layer. If the layout of the operating pages needs to be adjusted in the final functional layer, the programmer only needs to make changes to the functional layout, not the business logic or the data access interface.

Therefore, the .Net platform can provide a complete solution for application development in all business areas. Figure 1 shows the basic framework of the .Net platform. The .Net database can access interface uses data access interfaces, including ADO, ODBC, and so on. These database drivers are installed during initialisation and can be upgraded from the relevant website. In addition to the computer's own database drivers, the .Net platform also supports other types of database drivers, such as Oracle, MySQL, and so on.

2.2. UML. UML is a unified modelling language that graphically models the analysis and design phases of a software system, using object-oriented design thinking. In the analysis phase of a software system, after the analysis of the system requirements has been completed, the functional analysis of the system needs to be completed using UML. In general, the functional analysis of a system should be done in terms of a functional model, a static model and a dynamic model, which are shown in Figure 2. The functional model is described by a use case diagram, which includes the participants, the use cases, and the relationships between the use cases. In the use case diagram, the participants are the final users of the system and the use cases represent the operational functions. In addition to this, functions include primary and secondary functions. The static model of the system is represented by the entity class diagram of the system, which does not include the methods of the class but only the properties of the class. The dynamic model is described by the activity diagram, which is a sequential, selection, and juxtaposition structure to complete the operation of a use case.

In the design phase of a software system, class diagrams and sequence diagrams are adopted to design the functionality of the system. The class diagram is designed within the entity class diagram in the system analysis phase and allows for further expansion and design of the solution space. Sequence diagrams are used to represent the timing of calls between classes in the design of a function. In the design phase, algorithms or operations with complex state changes can be described by means of state diagrams. For large applications or integrated applications, a package diagram can be used to describe the composition of the packages. To be specific, a unit function can be designed as a package, which can include several classes or interfaces. In summary, the deployment of a software system can be described by means of component diagrams during the programming implementation

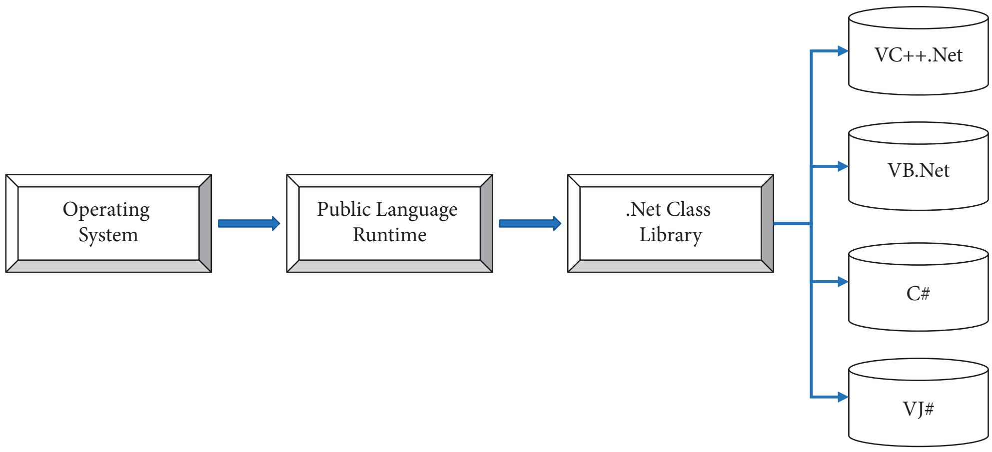


Figure 1: Basic framework of the.Net platform.

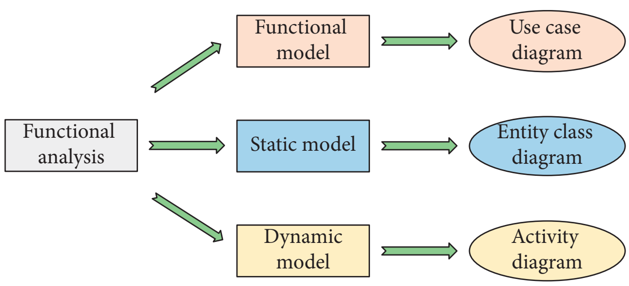


Figure 2: Functional analysis of a system.

phase. During the programming implementation, the implementation process of the system can also be described through activity diagrams and for some core algorithms can be described through sequence diagrams and activity diagrams.

2.3. Interval Intuitionistic Fuzzy Theory. Based on the research of fuzzy theory, interval intuitionistic fuzzy theory has been widely studied by scholars at home and abroad and has been applied to many fields such as multiattribute decision-making, pattern recognition, and data mining. Similarity measures are an important information tool in interval intuitionistic fuzzy theory, aiming to measure the closeness of different interval intuitionistic fuzzy sets to each other. In addition, similarity measures play a decisive role in fuzzy pattern recognition. Therefore, it is of great interest to find a reasonable and effective similarity measure for interval intuitionistic fuzzy sets. At present, a great number of scholars have conducted a lot of research on the similarity measures of interval intuitionistic fuzzy sets. However, it is found that most of the similarity measures of intuitionistic fuzzy sets do not satisfy the axioms of similarity measures. Although the similarity measure of interval intuitionistic fuzzy sets satisfies the similarity measure axiom, the method is more complicated. The detailed steps are shown as follows:

2.3.1. Determine Set of Evaluation Factors. The set of evaluation factors is a set of evaluation factors of the subject to be evaluated as elements. The set of evaluation factors is generally based on the characteristics of the subject to be evaluated and the evaluation criteria known from the needs analysis. In this study, the set of evaluation factors is divided into two levels of evaluation indicators. Thus, the first-level evaluation indicators' set can be obtained as

$$F = \{F_1, F_2, \ldots, F_n\}, \tag{1}$$

where F refers to the evaluation indicator.

Next, the first-level evaluation indicators can be subdivided, so that the second-level evaluation indicators can be identified:

$$F_n = \{F_{n1}, F_{n2}, \ldots, F_{nm}\}. \tag{2}$$

In the teaching evaluation system, the first-level evaluation indicators are divided into four areas: preparation, lectures, assignments, and education. Each first-level indicator can be divided into the second-level indicators. The specific division criteria are shown in Table 1.

2.3.2. Determine Judging Set. After determining the set of evaluation factors, the judging set can be determined as follows:

$$J = \{J_1, J_2, \ldots, J_k\}, \tag{3}$$

where J refers to the judging set, and each level can be compared to a fuzzy subset.

2.3.3. Determine Weight of Evaluation Indicator. Based on the statistical analysis of the evaluation factors, the weight vectors for each evaluation factor were obtained as shown in Table 2. The weight table corresponds to the evaluation factor in Table 1, respectively.

2.3.4. Determine Affiliation of Evaluation Indicator. For each evaluation factor in the set of evaluation factors, its affiliation can be determined from the data collected from the students' evaluations. The fuzzy matrix can be obtained by putting all the students' evaluations into a matrix.

Table 1: Specific division criteria.

First-level indicator	Second-level indicator
Preparation	Well prepared for teaching Teaching content enrichment
Lecture	Enlightenment Important and difficult points
Assignment	Attitude Timeliness
Education	Work ethic Caring

Table 2: Weight of each evaluation indicator.

First-level indicator	Second-level indicator
Preparation (0.2)	Well prepared for teaching (0.1) Teaching content enrichment (0.1)
Lecture (0.4)	Enlightenment (0.2) Important and difficult points (0.2)
Assignment (0.3)	Attitude (0.15) Timeliness (0.15)
Education (0.1)	Work ethic (0.05) Caring (0.05)

$$FM = \begin{bmatrix} F_{11} & \cdots & F_{15} \\ \vdots & \ddots & \vdots \\ F_{n1} & \cdots & F_{n5} \end{bmatrix}, \tag{4}$$

where F_{ij} refers to the affiliation of each evaluation indicator.

2.3.5. Determine Fuzzy Operator. The weighting of the evaluation factors identified in the teaching evaluation system is even. There are no situations where one weight dominates over another, so the weighted average method is chosen. This method is more effective in cases where the weights of the evaluation factors are not dominant and considers the weights of all evaluation factors, so that the results reflect the effects of all evaluation factors.

3. Teaching Quality Evaluation System

As the evaluation of teacher teaching quality is a comprehensive exercise, it is necessary to use sources of evaluation such as supervision data from the Academic Office. This approach makes the evaluation results more impartial and objective. In addition, the evaluation of teaching quality is often subject to changes in the priorities of the school and faculty from year to year and from semester to semester. As a result, the weighting of each factor can be used to calculate the final overall grade, which can be adjusted by the academic staff and the staff concerned.

3.1. Function Analysis. As the system is designed to evaluate teachers, teacher information needs to be managed. In addition to basic teacher information, this includes information about majors and classes. As a result, the maintenance of basic information also includes the management of majors, classes as well as teachers' information. This can be achieved through data entry and bulk import by the Academic Affairs Office. The basic data management use case diagram can be obtained as shown in Figure 3.

The teacher information addition activity diagram is shown in Figure 4. In the classroom quality assessment system, the objects involved include the Registrar's Office, the system, and the data processing object. The data processing object is used to complete database operations. Specifically, the Registrar's Office first logs into the system and enters the teacher information registration page. You will then need to select the course and class that the teacher is teaching and enter the teacher's job number. Once entered, the system will check the teacher's work number to determine if the teacher exists. If it is a duplicate of a record in the database, the teacher will be prompted. Conversely, if the teacher's work number is not duplicated, the Registrar will complete the entry of the other information. Once the entry is complete, the data are saved to the database by performing a save operation and the system indicates that the save is complete.

3.2. Selection of System Model. As software systems continue to improve in functionality and scale, the choice of system architecture becomes the most important factor in constraining system performance. The architectural requirements of software systems vary considerably from application to application. The architecture is the foundation of the system, and the choice and design of the system architecture is more important than the design of the functional structure of the system, the design of the database structure, and the writing of the program code. The following three types of system architecture are commonly used today.

The single-user architecture is the earliest version of a stand-alone management information system. Most of the current tools and software still follow this model. Each stand-alone system can only run on one computer and users cannot share data and work together directly.

C/S architecture, i.e., client and server architecture. C/s simulation can make use of the resource configuration of the client. The structure is shown in Figure 5.

The B/S architecture is a three-tier system based on the Internet. Under the B/S structure, the system is divided into three layers: the database service layer, the application service layer, and the user layer. The database management system is deployed on the database server to manage the data files. The system application is installed on the application server and the system program is run on the application layer, which generates the web interface for the user browser. The B/S architecture is currently the mainstream architecture for system development and is used by large enterprise groups and customers for decentralised applications.

Combining the many advantages of B/S structure design and the practical needs of this study, the design system of the teaching evaluation system adopts the B/S model. Figure 6 shows the schematic diagram of the B/S model structure.

The user representation layer is the interface that the user sees. It runs on the client computer and distributes

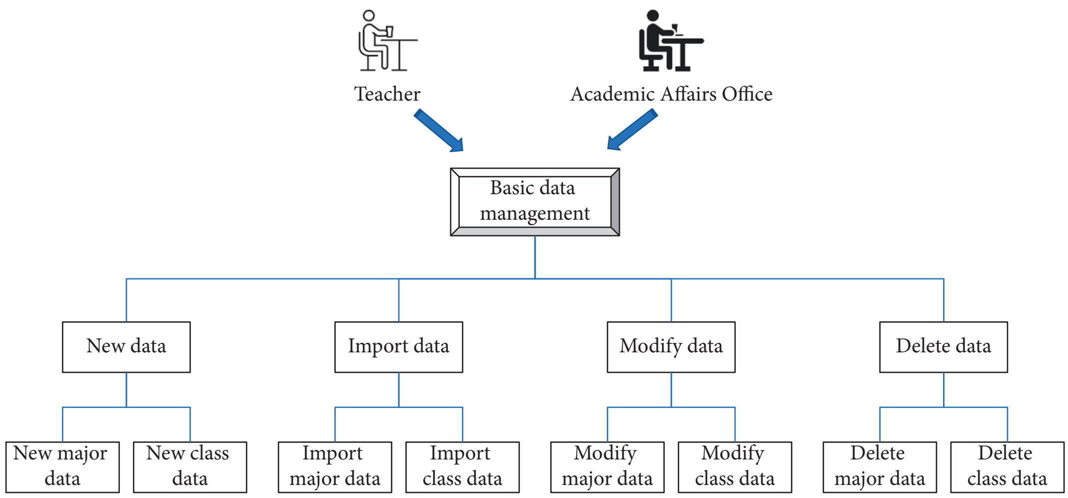


Figure 3: Basic data management use case diagram.

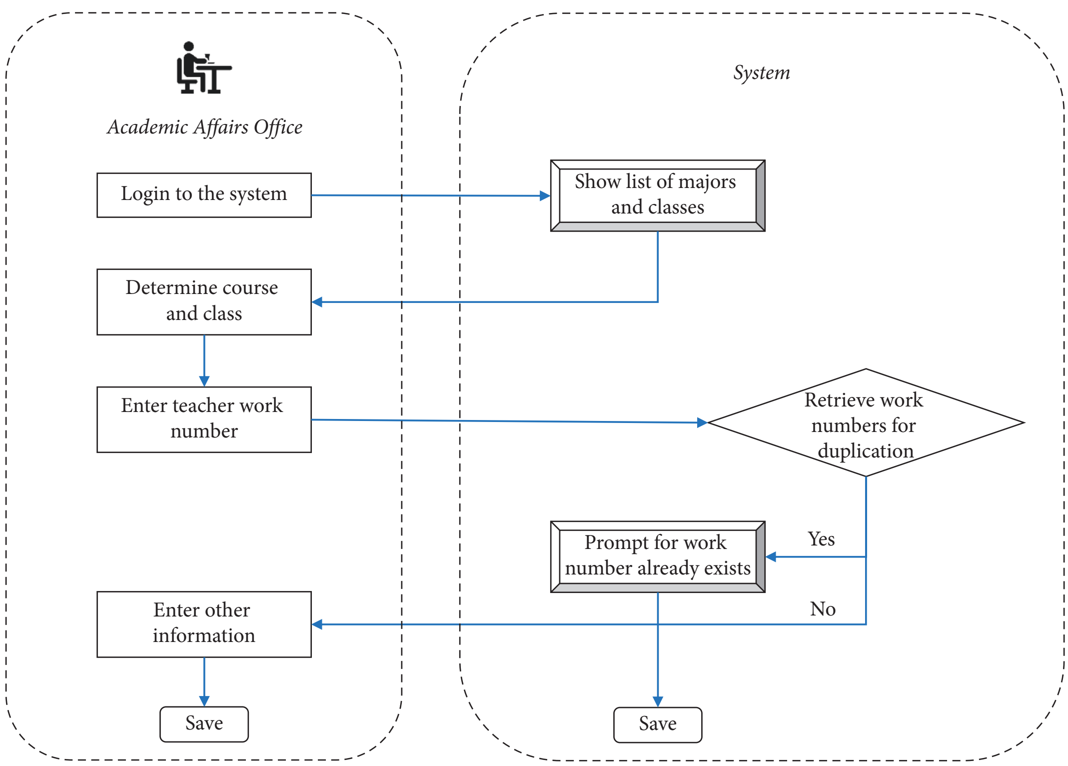


Figure 4: Teacher information addition activity diagram.

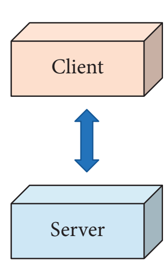


Figure 5: Structure of C/S architecture.

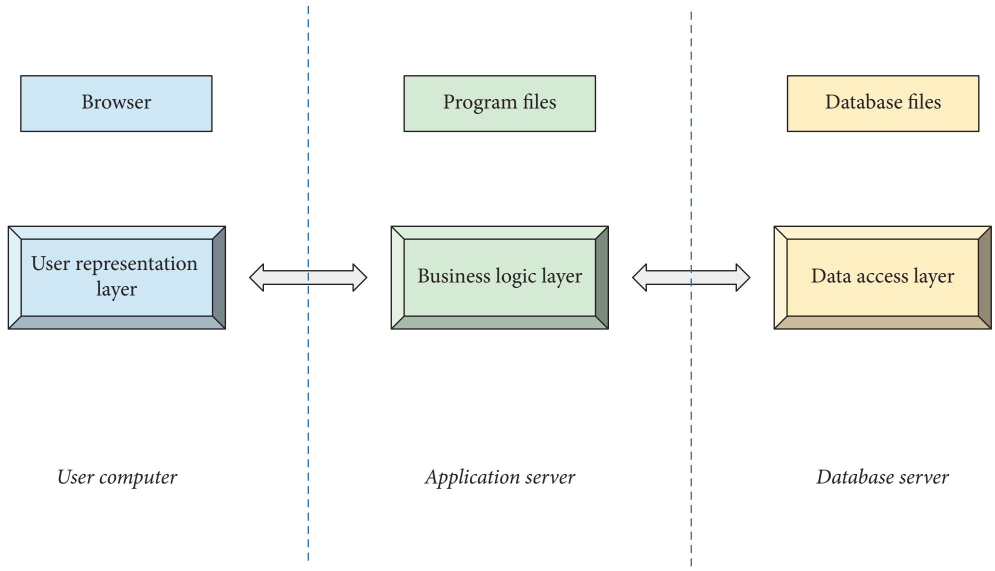


Figure 6: Schematic diagram of the B/S model structure.

information through a browser. The user can manipulate the business menus, send business requests to the server, and display the results returned by the server. The user representation layer does not perform any actual data processing but only transmits the user's instructions to the business logic layer.

The business logic layer receives processing instructions from the representation layer and mobilizes the program files to complete the business processing. At the same time, it generates data processing requests to the data access layer and generates a user interface with the data returned from the database, which is fed back to the user's computer browser.

The database management system and database files are deployed on the database server. The data access layer responds to data processing requests from the program files, writes, reads, and deletes data to the database and feeds the data processing results to the business logic layer.

4. Conclusion

The teaching evaluation system based on interval intuition fuzzy theory was developed mainly to meet the task of teaching evaluation in universities. The system can serve the teaching quality of the university in a fair and objective way. At the same time, the system can exchange data well with other management systems of the school such as the financial system to meet the practical daily needs of the school. The teaching evaluation system focuses on the problem of the large amount of data and the difficulty of statistical analysis, which is traditionally difficult to deal with in school teaching quality evaluation. The system also enables data sharing with other systems and solves the long-standing problem of low evaluation efficiency in schools. The paper proposes a four-layer structure based on the presentation layer, presentation control layer, business logic layer, and data layer. The detailed design of modules such as basic data management, evaluation scheme management, online marking and evaluation, teaching supervision data management, and evaluation process is completed through class diagrams and sequence diagrams, and the database design is completed through the database conceptual model and physical model. However, this study does not cover user security audit content, and this part of the work will be carried out in the next step to enhance the security audit design of user login logs.

Data Availability

The labeled data set used to support the findings of this study is available from the author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This work was supported by the Southwest University of Political Science and Law.

References

[1] S. Otto and P. Pensini, "Nature-based environmental education of children: environmental knowledge and connectedness to nature, together, are related to ecological behaviour," *Global Environmental Change*, vol. 47, pp. 88–94, 2017.

[2] D. Berridge, "The education of children in care: agency and resilience," *Children and Youth Services Review*, vol. 77, pp. 86–93, 2017.

[3] L. Coelho, N. Torres, C. Fernandes, and A. J. Santos, "Quality of play, social acceptance and reciprocal friendship in preschool children," *European Early Childhood Education Research Journal*, vol. 25, no. 6, pp. 812–823, 2017.

[4] F. Hammonds, G. J. Mariano, G. Ammons, and S. Chambers, "Student evaluations of teaching: improving teaching quality in higher education," *Perspectives: Policy and Practice in Higher Education*, vol. 21, no. 1, pp. 26–33, 2017.

[5] Q. W. Dong, S. M. Wang, F. J. Han, and R. D. Zhang, "Innovative research and practice of teachers teaching quality evaluation under the guidance of innovation and entrepreneurship," *Procedia Computer Science*, vol. 154, pp. 770–776, 2019.

[6] S. L. Campbell and M. Ronfeldt, "Observational evaluation of teachers: measuring more than we bargained for?" *American Educational Research Journal*, vol. 55, no. 6, pp. 1233–1267, 2018.

[7] A. F. Gilmour, C. E. Majeika, A. W. Sheaffer, and J. H. Wehby, "The coverage of classroom management in teacher evaluation rubrics," *Teacher Education and Special Education*, vol. 42, no. 2, pp. 161–174, 2019.

[8] M. A. Flores and M. L. Derrington, "School principals' views of teacher evaluation policy: lessons learned from two empirical studies," *International Journal of Leadership in Education*, vol. 20, no. 4, pp. 416–431, 2017.

[9] C. Palermo and M. M. Thomson, "Teacher implementation of self-regulated strategy development with an automated writing evaluation system: effects on the argumentative writing performance of middle school students," *Contemporary Educational Psychology*, vol. 54, pp. 255–270, 2018.

[10] T. G. Ford, M. E. Van Sickle, L. V. Clark, M. Fazio-Brunson, and D. C. Schween, "Teacher self-efficacy, professional commitment, and high-stakes teacher evaluation policy in Louisiana," *Educational Policy*, vol. 31, no. 2, pp. 202–248, 2017.

[11] J. A. Grissom and B. Bartanen, "Strategic retention: principal effectiveness and teacher turnover in multiple-measure teacher evaluation systems," *American Educational Research Journal*, vol. 56, no. 2, pp. 514–555, 2019.

[12] C. Aldeman, "The teacher evaluation revamp, in hindsight: what the Obama administration's signature reform got wrong," *Education Next*, vol. 17, no. 2, pp. 60–68, 2017.

[13] X. Xu, Y. Wang, and S. Yu, "Teaching performance evaluation in smart campus," *IEEE Access*, vol. 6, pp. 77754–77766, 2018.

[14] A. Valentine, I. Belski, and M. Hamilton, "Developing creativity and problem-solving skills of engineering students: a comparison of web-and pen-and-paper-based approaches," *European Journal of Engineering Education*, vol. 42, no. 6, pp. 1309–1329, 2017.

[15] N. Huang, "Analysis and design of University teaching evaluation system based on JSP platform," *International Journal of Education and Management Engineering*, vol. 7, no. 3, pp. 43–50, 2017.

[16] N. A. Paufler and E. F. Sloat, "Using standards to evaluate accountability policy in context: school administrator and teacher perceptions of a teacher evaluation system," *Studies In Educational Evaluation*, vol. 64, Article ID 100806, 2020.

[17] A. Al Ansari, M. Rushdi Arekat, and A. H. Salem, "Validating the modified system for evaluation of teaching qualities: a teaching quality assessment instrument," *Advances in Medical Education and Practice*, vol. 9, pp. 881–886, 2018.

[18] S. Cadez, V. Dimovski, and M. Zaman Groff, "Research, teaching and performance evaluation in academia: the salience of quality," *Studies in Higher Education*, vol. 42, no. 8, pp. 1455–1473, 2017.

[19] A. Y. Noaman, A. H. M. Ragab, A. I. Madbouly, A. M. Khedra, and A. G. Fayoumi, "Higher education quality assessment model: towards achieving educational quality standard," *Studies in Higher Education*, vol. 42, no. 1, pp. 23–46, 2017.

[20] H. Lin, J. You, and T. Xu, "Evaluation of online teaching quality: an extended linguistic MAGDM framework based on risk preferences and unknown weight information," *Symmetry*, vol. 13, no. 2, p. 192, 2021.

[21] J. Ye, "Modeling of performance evaluation of educational information based on big data deep learning and cloud platform," *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 6, pp. 7155–7165, 2020.

[22] K. Royal, "A guide for making valid interpretations of student evaluation of teaching (SET) results," *Journal of Veterinary Medical Education*, vol. 44, no. 2, pp. 316–322, 2017.

[23] D. E. Clayson, "Student evaluation of teaching and matters of reliability," *Assessment & Evaluation in Higher Education*, vol. 43, no. 4, pp. 666–681, 2018.

[24] B. Cheng, C. Fan, H. Fu, J. Huang, H. Chen, and X. Luo, "Measuring and computing cognitive statuses of construction workers based on electroencephalogram: a critical review," *IEEE Transactions on Computational Social Systems*, pp. 1–16, 2022.

[25] Y. Qian, S. Chen, J. Li et al., "A decision-making model using machine learning for improving dispatching efficiency in Chengdu Shuangliu airport," *Complexity*, vol. 2020, Article ID 6626937, 9 pages, 2020.

[26] T. Higaki, Y. Nakamura, J. Zhou et al., "Deep learning reconstruction at CT: phantom study of the image characteristics," *Academic Radiology*, vol. 27, no. 1, pp. 82–87, 2020.

[27] R. M. Filius, R. A. de Kleijn, S. G. Uijl, F. J. Prins, H. V. Rijen, and D. E. Grobbee, "Promoting deep learning through online feedback in SPOCs," *Frontline Learning Research*, vol. 6, no. 2, pp. 92–113, 2018.

[28] B. Cheng, K. Lu, J. Li, H. Chen, X. Luo, and M. Shafique, "Comprehensive assessment of embodied environmental impacts of buildings using normalized environmental impact factors," *Journal of Cleaner Production*, vol. 334, Article ID 130083, 2022.

[29] B. Cheng, J. Huang, J. Li, S. Chen, and H. Chen, "Improving contractors' participation of resource utilization in construction and demolition waste through government incentives and punishments," *Environmental Management*, 2022.

[30] Y. W. Hou, C. W. Lee, and M. G. Gunzenhauser, "Student evaluation of teaching as a disciplinary mechanism: a foucauldian analysis," *The Review of Higher Education*, vol. 40, no. 3, pp. 325–352, 2017.

[31] G. Pan, V. Shankararaman, K. Koh, and S. Gan, "Students' evaluation of teaching in the project-based learning programme: an instrument and a development process," *International Journal of Management in Education*, vol. 19, no. 2, Article ID 100501, 2021.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 6079473, 11 pages
https://doi.org/10.1155/2022/6079473

Research Article

Recommendation of Online Business English Learning Resources Integrating Attention Mechanism and Collaborative Filtering Model

Zhemei Yang [1] and Dongbin Lv [2]

[1] *Department of Foreign Languages, Guangxi Normal University for Nationalities, Chongzuo 532200, China*
[2] *Guangxi Expressway Investment Co Ltd, Nanning 530000, China*

Correspondence should be addressed to Zhemei Yang; yangzhemei@gxnun.edu.cn

Received 17 May 2022; Accepted 10 June 2022; Published 12 July 2022

Academic Editor: Jiafu Su

Faced with massive resources, many learners find it difficult to quickly screen out useful content for themselves. In order to help learners acquire the required network resources quickly and accurately, the birth of a personalized recommendation system solves this problem perfectly. A collaborative filtering algorithm has been widely used in the field of personalized recommendation. However, due to the limitation of the model, the recommendation effect has not been further improved. The single weakness of a collaborative filtering algorithm to recommend learning resources is difficult to meet the needs of learners to acquire personalized resources. This paper proposes a recommendation algorithm for business English online learning resources based on an attention mechanism and collaborative filtering model. The learner vector and learning resource vector are mapped to multispace, and the learner-learning resource interaction is done from multiple angles. The final learner representation vector and learning resource representation vector are aggregated by a two-level attention mechanism to predict scores. Through teaching practice in student associations, it is found that students from different backgrounds have different preferences for business English online learning resources. This method has a positive impact on online learning. This study aims to provide some references for English education resource recommendations. The results at Precision@K and Recall@K prove that the proposed model has better recommendation ability.

1. Introduction

With the rapid development of online education, digital learning resources show the characteristics of massive resources. When learners have many choices, they also inevitably face serious knowledge overload and learning loss [1]. To solve this problem, learners need to rely on personalized learning and adaptive recommendation to navigate [2]. The adaptive recommendation is the core task in the process of personalized learning. Classical recommendation algorithms explore learners' potential interest preferences based on the historical behaviour information and similarity relationship of groups [3]. However, such algorithms only take the interaction information between learners and resources as input, and the sparsity of data makes the recommendation have certain defects.

The goal of the recommendation algorithm is to extract the information that learners are interested in from massive data, and it is one of the effective tools to solve the problem of "information overload." Literature [4] converts learners' learning behaviours into learners' scoring of resources and improves learners' similarity calculation to solve potential data sparsity and cold start problems in the recommendation system. Literature [5] uses learning materials to build an e-learning resource knowledge base based on domain ontology and combines content-based and rule-based methods to provide mixed recommendations for learners. A feature model based on domain ontology and learner attribute information was proposed in literature [6]. Based on the learning feature model, a collaborative filtering recommendation method integrating similarity was designed. The learning resource recommendation algorithm based on behaviour analysis was proposed in literature [7], which is used to mine the behavioural data of learners and format it into collaborative filtering recommendations. In the above literature, in order to overcome the defects of traditional recommendation algorithms and improve recommendation performance, researchers introduced different types of auxiliary information. However, these auxiliary information contain only isolated characteristics of learners or learning resources. In fact, there are abundant connections between learners and learning resources and between learning resources and learning resources.

The reform of English teaching in universities proposes to explore the deep integration of information technology and English education under the background of "Internet + education," so as to bring about changes in education methods and learning methods [8]. Online English learning is a new learning mode under the background of education informatization. It breaks the space-time limit and reconstructs the learning process and teacher-student relationship by relying on mobile devices [9]. Online business English learning in the context of English learning not only can create a real language communication environment with the help of modern information technology so that students can appreciate diverse cultures but also can provide the possibility to change the "teacher-cantered" education [10].

In actual learning, due to the weakened role of teachers as instructors, students lack the correct positioning of learning activities and themselves when facing massive resources independently. They are unable to construct their own knowledge networks and choose appropriate learning methods as needed, resulting in information overload [11]. Many students do not have a deep understanding of "what to learn and how to learn," and cannot become the active acquirer and constructor of knowledge. The Ten-year Development Plan for Education Informatization proposes to "build an intelligent teaching environment and provide personalized learning information environment and services for learners" [12].

With the widespread use of online learning platforms, the number of English online learning resources has also increased rapidly, and it is difficult for learners to quickly locate the resources they need among the huge number of English online learning resources [13]. How to recommend valuable information to interested learners from massive English resources has always been a core issue in online education services [14]. As an important solution to this problem, the educational resource recommendation system has attracted more and more researchers' attention [15]. However, at present, most educational resource recommendation systems are only for college students and are used by their own online systems. However, a wider range of off-campus students and online systems outside universities are unable to obtain educational resource recommendation services, which greatly reduces the utilization rate of English online learning resources [16].

In order to solve the problem of resource rate limitation existing in existing models, this paper introduces the concept of multispace interactive feature extraction and proposes a resource recommendation model for business English online learning based on the attention mechanism and collaborative filtering model, namely Multispace Interactive Collaborative Filtering (MSICF). The proposed model maps the learner vector and the learning resource vector to multiple spaces to extract interactive features for scoring prediction. Multispace can consider the interaction of learners' learning resources from multiple perspectives, and more comprehensive features can better fit the learner-learning resource scoring, thus improving the recommendation ability of the model.

The innovations and contributions of this paper are listed below:

(1) The multispace concept is introduced into the collaborative filtering recommendation system, which enhances the interpretability of recommendations and refines the granularity of learner-learning resource interaction feature extraction.

(2) A novel multispace interactive collaborative filtering recommendation model (MSICF) is proposed, which maps the learner vector and the learning resource vector to multiple spaces and extracts the interactive features of learners and learning resources from multiple perspectives.

(3) For the top-K recommendation task, the top-K learning resources are recommended for learners in the test set, and the performance of the model is evaluated by using Precision@K and Recall@K indicators. The results of the MSICF model are better than other comparison models under multiple evaluation indexes.

This paper consists of four main parts: the first part is the introduction, the second part is methodology, the third part is result analysis and discussion, and the fourth part presents the conclusion.

2. Methodology

Figure 1 shows a recommended example of a learner's access to a business English online learning resource. Figure 1 provides a learner-learning resource interaction matrix G. When the element is 1, it means that learners like the independent learning resource. When the element is 0, it means that the corresponding learners have not visited the corresponding independent learning resource. For example, G11 = 1 means that learners p1 likes the independent learning resource x1. According to the learner-learning resource interaction matrix, learners p3 and p1 both like independent learning resources x2 and x5, and learners p3 and p2 both like independent learning resources x4. According to the traditional collaborative filtering idea, learners p3 and p1 are more similar. It is considered to recommend the independent learning resources favoured by learners p1 to p3, but it is not clear whether p3 prefers x1 or x3. However, learners' access to business English online learning resources can be divided into two situations: (1) interaction occurs from the perspective of types of learning

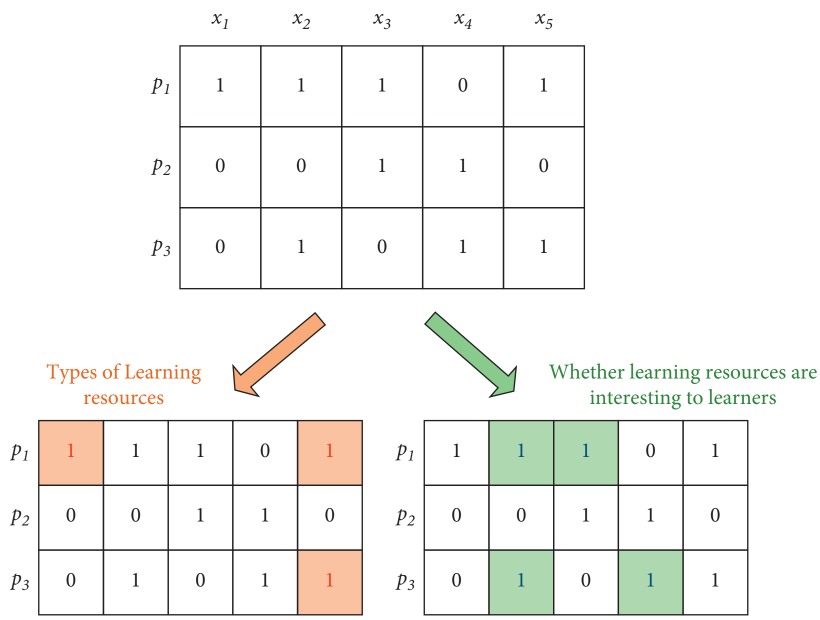


Figure 1: Examples of learner-learning resource interaction from different perspectives.

resources and (2) interaction from the perspective of whether learning resources are interesting to learners. From the perspective of the types of learning resources, learner p1 only likes x1 and x5, and both learners p1 and p3 like x5, so the x1 favoured by learner p1 is recommended to p3. Similarly, from the perspective of whether learning resources cater to learners' interests, learners p1 and p3 like x2 at the same time, so they recommend x3 that learners p1 likes p3. The traditional collaborative filtering model fails to analyze the interaction between learners and learning resources from multiple perspectives, and the recommendation performance is limited. However, feature interaction extraction from different perspectives can discover learners' preferences more effectively.

Learner-learning resource interaction from different perspectives is different, so learner-learning resource interaction from multiple perspectives is described. In this paper, the full connection layer is used to map the learner's embedding vector and learning resource embedding vector to multiple spaces. Due to the difference in the full connection weight, the learner's embedding vector and learning resource embedding vector after mapping also contain different features, indicating learner-learning resource interaction from different perspectives.

The connotation of the recommendation system is to get learners' scores of learning resources through the model. Figure 2 is a frame diagram of the proposed model. This model uses row data and column data of the learner-learning resource interaction matrix as input for the learner module and learning resource module, respectively. After the learner embedding vector and the learning resource embedding vector are mapped to multiple spaces, the learner representation vector and learning resource representation vector are obtained through a multispace interaction module. Then, the learner representation vector and the learning resource representation vector are splicing and sent into the multilayer perceptron (MLP) to obtain the learner's score on the learning resource. The learner part and the learning resource part have similar structures. The learner part and the learning resource part's multispace interaction modules aggregate the learning resource embedding vector and the learner embedding vector, respectively, to generate the learner representation vector and the learning resource representation vector. As the key innovation point of this paper, the multispace interaction module mainly does the following operations:

(1) Through N full connection layers with different weights, the embedded vectors are mapped to N different vector spaces
(2) The magnitude attention mechanism assigns different weights to vectors in a single subspace and aggregates them into a representation vector of learners (or learning resources)
(3) Spatial attention mechanism assigns different weights to different subspace vectors and aggregates them into the final representation vector of learners (or learning resources)

2.1. Input Module. The learner-learning resource interaction matrix G features with original data W and T; the number of rows in the matrix is the number of learners W; and the number of columns is the number of learning resources T. The interaction matrix represents the element value G_{yz} of row y and column z as whether learner y has interaction with learning resource z. The number 1 indicates interaction, and 0 indicates no interaction. For row y, 2, 4, 5, . . ., column data is 1, indicating learning resources 2, 4, 5, ... There was interaction with learner y. A full connection layer is used to embed all learners and learning resources into the low-dimensional and dense vector space to obtain the embedding

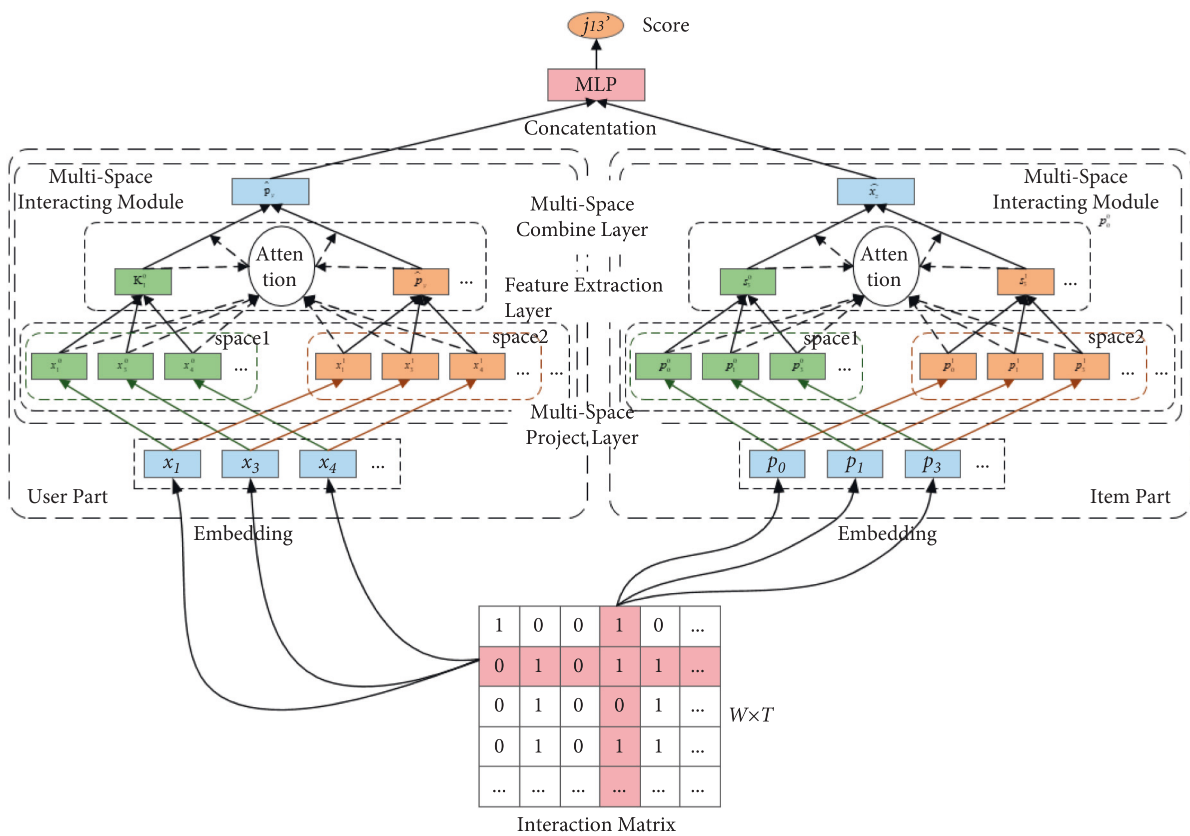


Figure 2: Framework of the proposed model.

vectors of all learners and learning resources. And y produced interactive learning resources, as well as corresponding learning resources; vector $\{x_z | z \in x_y\}$ are embedded to learners. Learning resources will work with the same z corresponding to the interaction of learners; vector $\{p_y | y \in p_z\}$ are embedded into the learning resources section as the input of the model.

2.2. Multispace Interaction Module. The learner part and the learning resources part have a similar structure of a multispace interaction module. As to learners, for example, many spatial interaction modules are input to learners' history, interactive learning resources embedded with vector $\{x_z | z \in x_y\}$ the output of the final learners is said by vector $\widehat{p}_y$. Specifically, the multispace interaction module consists of three parts: multispace mapping layer, feature extraction layer and multispace combination layer.

In order to extract learner-learning resource interaction features comprehensively from multiple perspectives, this paper introduces the concept of multispace feature extraction. Through n full connection layers, learners and learning resource data are mapped to corresponding n different spaces, and the mapping process of learners y to space n is as follows:

$$p_y^n = \text{Dense}_t(p_y) = U_n p_y, \tag{1}$$

where $U_n \in U = \{U_1, U_2, \ldots, U_N\}$ that is the weight matrix of the full connection layer and p_y^n represents the new embedding vector after the embedding vector of learner y is mapped to space n. Different spatial mappings adopt different weight matrices. After mapping, learner embedding vectors in different spaces contain different elements, which can describe learner characteristics from different perspectives.

For learning resource z, a similar operation is performed to map learning resource to space n.

$$x_z^n = \text{Dense}_n(x_z) = V_n x_z, \tag{2}$$

where $V_n \in V = \{V_1, V_2, \ldots, V_N\}$ that is the weight matrix of the full connection layer and x_z^n represents the new embedding vector after the embedding vector of learning resource z is mapped to space n.

After learning resource vectors are mapped to multiple spaces, learners' preferences can be obtained by aggregating their representation vectors in each space. In the learner part, x_y, a set of learning resources that interact with learner y, is used as a feature to aggregate and generate the representation vector of learner y in a single space. In each space, when learning resource vectors are aggregated to generate learner representation vectors, learners' interest in each learning resource in this space is inconsistent. Therefore, when aggregation generates a learner representation vector, different weight values should be assigned to each learning resource vector. A directional attention mechanism is used in the feature extraction layer to assign different weights to each learning resource vector in a single space.

In space n and j produced interactive learning resources of embedded vector for $\{x_z^n | z \in x_y\}$, in this space, learning resources z's contribution to the learners y vector generated is calculated by the following formulas:

$$\alpha_{yz}^n = \frac{\exp(\Psi(p_y, x_z^n))}{\sum_{z\in X_y} \exp(\Psi(p_y, x_z^n))},$$
$$\Psi(p_y, x_z^n) = M_{\text{Query}} p_y, M_{\text{Key}} x_z^n, \tag{3}$$

where $\Psi(p_y, x_z^n)$ is the scoring function used to score the learning resource z. It can be defined as a neural network or other similarity calculation function. Because the inner product is simple and efficient, this paper directly uses the inner product to calculate the similarity between the representation vector p_y of learner y and the representation vector x_z^n of learning resource z in space n. M_{Query} and $M_{\text{Key}} \in R^{d\times d}$ are mapping matrices, which map the representation vector p_y of learner y and x_z^n of learning resource Z in space N from R^d to $R^{d'}$.

After obtaining the weights of all learning resource vectors in space n, they are aggregated into the representation vector k_y^n of learner y in space n.

$$k_y^n = \sum_{z\in X_y} \alpha_{yz}^n \left(M_{\text{Value}} x_z^n\right), \tag{4}$$

where $M_{\text{Value}} \in R^{d\times d}$ is the mapping matrix. So, every space layer of feature extraction from input for learners to interactive learning resources is embedded with vector y $\{x_z^n | z \in x_y\}$. In the space of the output of said vector k_y^n polymerization of learners. N spaces can aggregate n learner representation vectors $\left\{k_{\sigma_y}^n\right\}_{n=1}^N$.

The structure of learning resources is similar to that of learners, and the representation vector of learning resources in N is spaces $\{S_z^n\}_{n=1}^N$.

In the learner part, n learner representation vectors $\{k_y^n\}_{n=1}^N$ are synthesized from n different spaces from different angles. When the representation vectors of n learners are aggregated into the final representation vector of learner y because learners have different preferences for different angles, vectors in different spaces should make different contributions to the aggregation. Therefore, a spatial attention mechanism is used to assign different weights to different spaces.

This layer will be the output of all feature extraction layers $\{K_y^n\}_{n=1}^N$ as input; learn different weight values for them.

$$\beta_y^n = \frac{\exp(\Phi(k_y^n))}{\sum_{n=1}^N \exp(\Phi(k_y^n))},$$
$$\Phi(k_y^n) = \sigma(v k_y^n + h), \tag{5}$$

where $\Phi(k_y^n)$ is the scoring function, and a layer of the neural network is used to calculate the score of k_y^n. Parameter v is the weight of the neural network. h is the offset. σ is ReLU function, which is the activation function of the neural network.

After obtaining the weight values of all space vectors $\{\beta_y^n\}_{n=1}^N$, the representation vectors of each space are aggregated according to different weights.

$$p_y = \sum_{n=1}^N \beta_y^n \cdot k_y^n. \tag{6}$$

The learner representation vector of the multispace was weighted and aggregated, and the learner representation vector p_y containing multiple feature interaction information was obtained as the output of the multispace combination layer. The learning resource representation vector x_z containing multiple feature interaction information can also be obtained by using the same method.

2.3. Output Module. The output layer splices the learner representation vector p_y obtained by the learner module and the learning resource representation vector x_z obtained by the learning resource module into the Multilayer Perceptron (MLP), so as to obtain the score values of learner y and the learning resource z.

Specifically, the learner representation vector p_y and learning resource representation vector x_z are spliced into a vector.

$$g_0 = \text{Concat}(p_y, x_z). \tag{7}$$

Then, the splicing vector is sent into the feedforward neural network F with D hidden layers. The d hidden layer of feedforward neural network F is f^d, and its nonlinear function with the previous hidden layer f^{d-1} is expressed as follows:

$$f^d = \sigma(M^d f^{d-1} + h^d), \tag{8}$$

where M^d and h^d are the parameters of layer d, $f^1(i) = g_0$, and σ Is the nonlinear activation function ReLU. Combining formulas (7) and (8), the following can be obtained:

$$j'_{yz} = f^D\left(\cdots f^2\left(f^1, (p_y, x_z)\right)\cdots\right), \tag{9}$$

where D is the total number of hidden layers. Finally, the predicted score j_{yz}' of learner y on learning resource z is obtained.

2.4. Model Training. The final output of the MSICF model is learners' rating of learning resources. For the scoring prediction problem, the commonly used objective function is the square loss function.

$$L_r = \sum_{(y,z)\in T} \left(j'_{yz} - j_{yz}\right)^2, \tag{10}$$

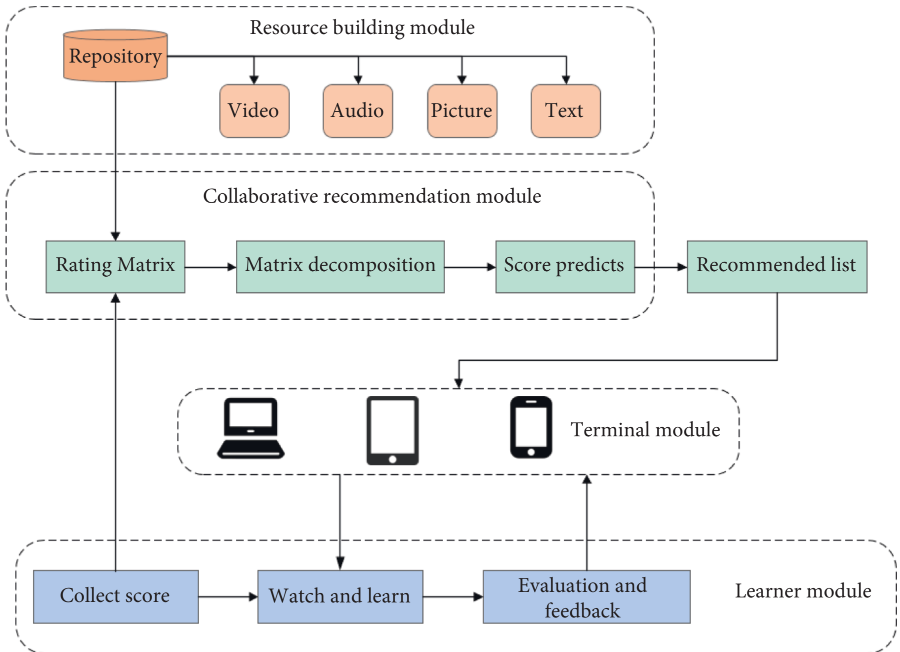


FIGURE 3: Personalized collaborative recommendation framework.

where T is the interactive set of all users and projects and j_{yz} is the real score of learner y on learning resource z.

This paper uses the adaptive gradient algorithm (Adam) to minimize the objective function. Adam algorithm is different from the traditional gradient algorithm. The learning rate in the traditional gradient algorithm is fixed. Adam algorithm can design independent adaptive learning rates for different parameters.

As the number of spaces increases, the model parameters also increase, and the model is more prone to overfitting, which leads to a decrease in the generalization ability of the model. To alleviate the problem of overfitting, dropout and L2 regularization techniques are introduced.

The idea of dropout is to randomly drop connections between neurons during training so that each training model is somewhat different. It has been proved to alleviate the overfitting problem of complex models. The introduction of the L_2 regular term can punish the high-frequency parameters and alleviate the overfitting problem.

After introducing L_2 regularization, the actual objective function used for training is as follows:

$$L = L_r + \lambda \|M\|^2 = \sum_{(y,z)\in 7} \left(j'_{yz} - j_{yz}\right)^2 + \lambda \|M\|^2, \tag{11}$$

where λ is the regularization coefficient, which controls the intensity of regularization, and M is the parameter set of the model.

This research adopts qualitative and quantitative research methods, and the framework is shown in Figure 3, mainly involving four modules.

First is resource construction module. Business English online learning aims to cultivate students' English communication skills in work and social life. In view of the above factors, this study selected four popular English online

TABLE 1: Distribution of learners' scores.

Number of scores	Number of learners	Proportion (%)
[6, 10)	10	33
≥10	14	47
<3	2	6.67
[3, 6)	4	13.33

learning websites in teaching practice according to the characteristics of English online learning scenarios. These online resource sites include Tianya Xiaozhu, Akasuo oral English, One 100 easy multimedia textbook library, and Cocoa English.

Second is collaborative recommendation module. Here, we use matrix factorization. R represents the original scoring matrix. P represents the number of learners. X represents the number of resources. Each row in the matrix represents a learner; each column represents a teaching resource; and each element value in the matrix represents the score of the corresponding learner on the corresponding resource.

Third is terminal module. Data will be collected through laptops, mobile phones, Pads, and other terminal devices, and business English online learning resources will be displayed to target learners.

Fourth is learner module. Learn the knowledge reserve and preferences of learners based on the collected data. When collecting data in the teaching practice, the options of each question are set at five levels from "very good," "not bad," "average" to "not very good," and "very bad." The collected scoring data is quantified into a scoring matrix and sent to the collaborative recommendation module. Then, the resources recommended by the system are shown to students for learning. After watching, students give feedback and score as post-test data.

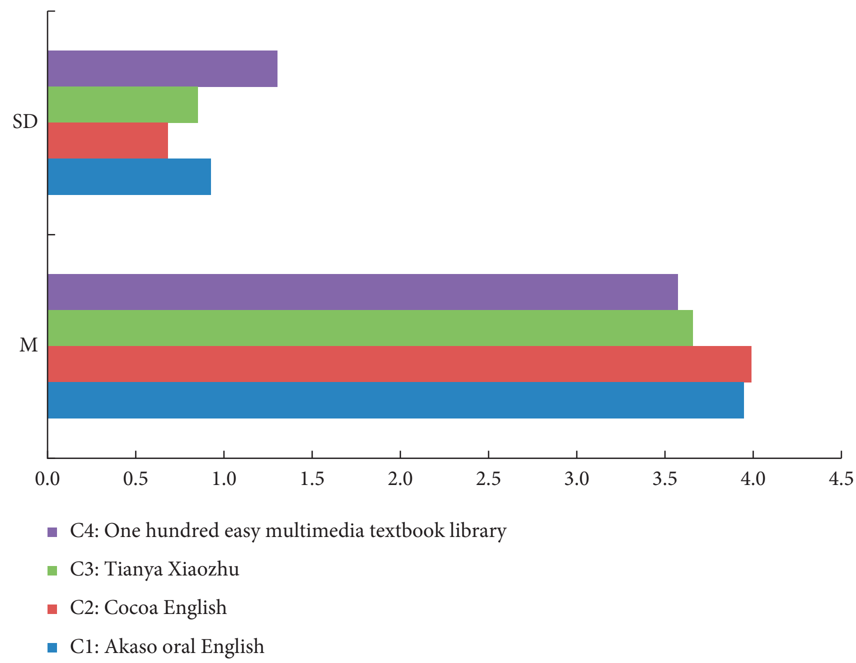


FIGURE 4: Descriptive analysis of different resource types.

3. Result Analysis and Discussion

The teaching practice was carried out for students in the community, and 30 students had different degrees of understanding of the 4 online business English learning websites provided by the platform, as shown in Table 1. There are 2 people who use Tianya Xiaozhu, 4 people who use One hundred easy multimedia textbook library, 10 people who use Akasuo oral English, and 14 people who use Cocoa English. It can be seen that most students are familiar with the selection of materials, which provides a good foundation for the follow-up implementation of online independent learning. All data with a score of 0 (i.e., "not used") were filtered out, and a total of 260 pieces of rating data were collected. The density of the scoring matrix reaches $260/(30 * 12) \sim 72.22\%$, and the scoring data per capita also reaches 8.6, indicating that the matrix sparsity problem faced in this case is not obvious.

In this study, business English online learning resources are divided into four types: Tianya Xiaozhu, Akasuo oral English, One hundred easy multimedia textbook library, and Cocoa English. First, descriptive analysis is carried out. The mean and standard deviation of scores of different types of resources are shown in Figure 4. As can be seen from Figure 4, the average score of Cocoa English resources is the highest at 3.99. Akasuo followed with an average of 3.95. The average score of Tianya Xiaozhu and One hundred easy multimedia textbook library is relatively low. It can be seen from Table 1 and Figure 4 that in general, the students in this English club are familiar with the networked and information-based learning environment and have much contact with online business English learning resources in daily life. At the same time, the resources are more inclined toward Cocoa English and Akasuo oral English.

TABLE 2: Analysis of gender differences in business English online learning resources.

Type	Gender	*N*	M	Sd	*t*	Sig.
C1	Female	24	4.23	0.8	2.004	0.067
	Male	6	3.34	1.63		
C2	Female	24	4.07	0.62	3.635	0.012
	Male	6	2.73	1.44		
C3	Female	24	4.08	0.84	−0.233	0.819
	Male	6	4.17	0.73		
C4	Female	24	3.74	1.53	−0.052	0.961
	Male	6	3.78	1.05		

The independent sample t-test was used to analyze the preference differences of students of different genders in different business English online learning resources (see Table 2). As can be seen from Table 2, the significance on C1, C3, and C4 were all greater than 0.05, without a significant gender difference. The significance on C2 was less than 0.05, indicating a significant difference between genders.

Variance analysis was conducted on the preference differences of students of different grades in different business English online learning resources, and the results are shown in Table 3. In this study, there are 18 sophomores, accounting for 60%. There are 6 juniors, accounting for 20%. There are 6 senior students, accounting for 20%. Because the English club requires members to have a certain level of professional English literacy, the club mainly recruits students of sophomore and above grade. As can be seen from Table 3, the significance was less than 0.05 in C1 and C2, indicating that there were significant differences among different grades, while there were no significant differences in C3 and C4.

The post hoc comparison of C1 and C2 is shown in Figure 5. In C1 and C2, the significance of sophomore and

Table 3: Analysis of differences in online learning resources of business English for learners of different grades.

Type	Grade	M	Sd	F	Sig.
	Sophomore	4.08	0.65		
C1	Junior	4.01	0.42	5.23	0.011
	Senior	2.62	1.65		
	Sophomore	3.71	0.41		
C2	Junior	4.34	0.73	8.16	0.001
	Senior	2.35	1.57		
	Sophomore	4.04	0.68		
C3	Junior	3.55	0.42	1.59	0.208
	Senior	3.42	1.04		
	Sophomore	3.66	1.3		
C4	Junior	2.64	1.8	0.838	0.418
	Senior	3.62	1.31		

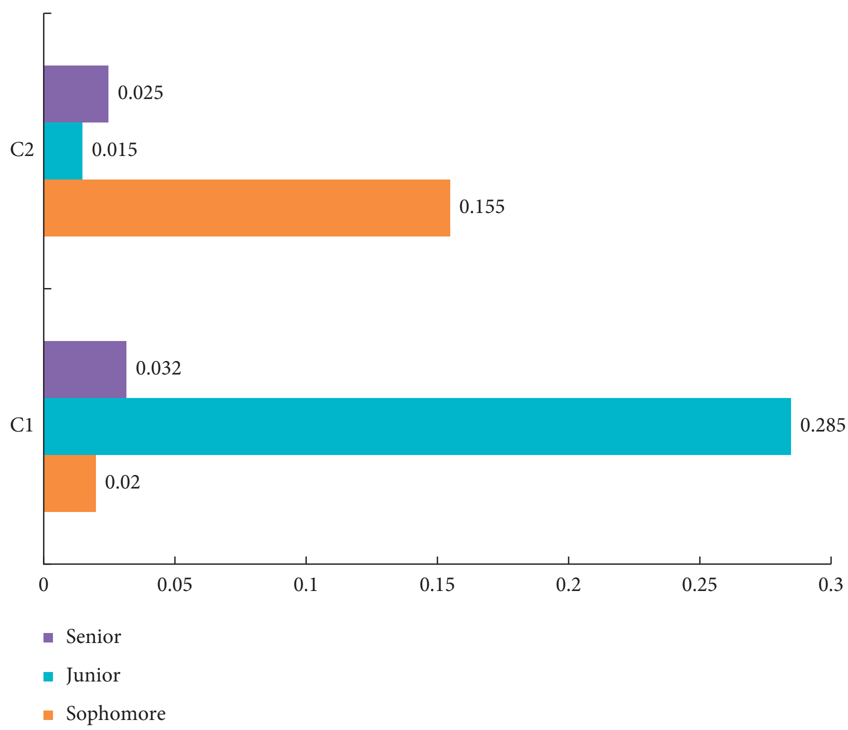


Figure 5: Comparisons of C1 and C2 at grade level.

senior, junior, and senior were all less than 0.05, indicating significant differences. Compared with seniors, sophomores and juniors preferred Akasuo oral English and Cocoa English, indicating that these resources are more attractive to middle- and lower-grade students.

Analyze students' satisfaction with the recommended resources. The collected data is used for collaborative recommendations. Score 0 indicates that the student has not used the resource, and the resource is added to the set of candidate resources to be recommended. When the predicted score reaches more than 3.5, it corresponds to "not bad" and "very good," thinking that the resource is likely to be useful to the learner, so it will be pushed to students. Urge them to watch and learn online and score comprehensively. Click the "Submit" button to indicate the end of this round of learning. The behavioural data of learners were recorded throughout the experiment as the basis for experimental analysis.

Compare the feedback received with the system's predictions, as shown in Figure 6. Blue data indicates that the prediction score recommended by this resource is 4 or 5 points, a total of 46 items. Students' feedback after learning was marked as read data. When the student also scores 4 or 5, the student is considered to have approved the resource recommended by the system. Students rated the resource 3 and below as not resonating with students. Among the 46 recommended data, 33 received positive feedback from learners, with a satisfaction rate of 71.74%.

Based on the above quantitative analysis and teaching practice, it can be concluded that the collaborative filtering recommendation model based on the attention mechanism used in this paper can effectively recommend business English online learning in the context of online learning and meet the needs of students' personalized learning.

In interaction rate prediction, the top-K learning resources were recommended for learners in the test set for the

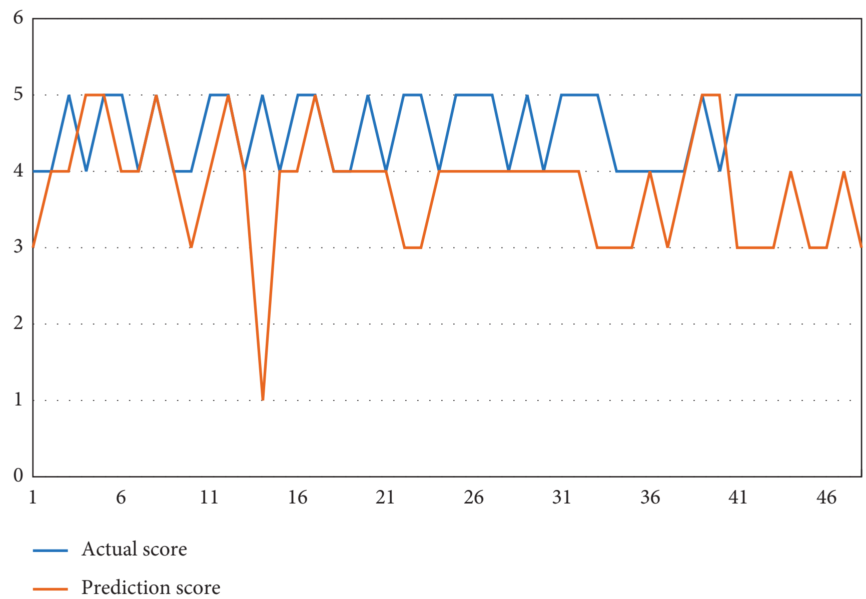


Figure 6: Statistics on the effect of pushing business English online learning resources.

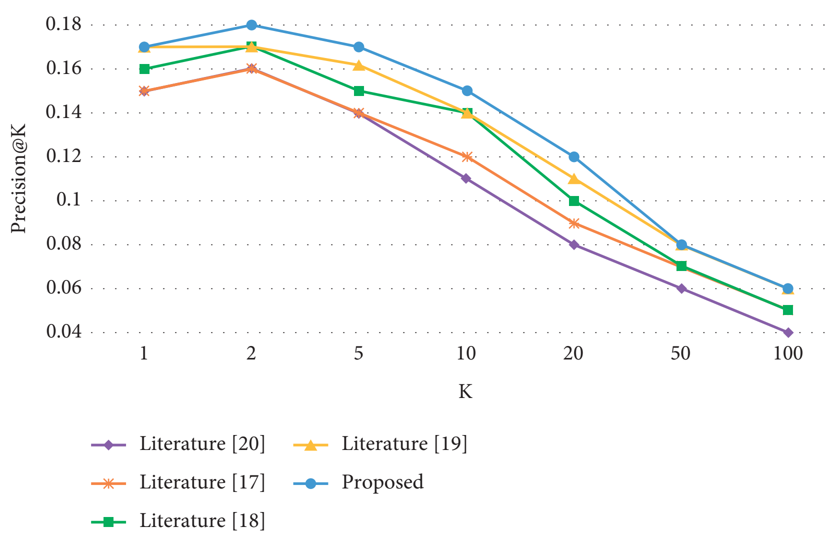


Figure 7: Precision of Top-K.

top-K recommendation task, and the model performance was evaluated by using Precision@K and Recall@K indicators. The algorithm in this paper was compared with recent literature [17–20]. Figures 7 and 8 show the comparison between Precision@K and Recall@K in top-K services.

As can be seen from Figures 7 and 8, when $K = 5$, Precision@K and Recall@K of literature [19] perform best. Compared with literature [19], the proposed model has improved by 8% and 15% in Precision@K and Recall@K. By comparing the experimental data, the analysis shows that the three baseline models referred to in literature [17] are superior to literature [20], indicating that entity and relationship information is conducive to improving recommendation performance after the introduction of attention mechanism. Among them, literature [17], starting from the learner end, uses the entities around learning resources to spread the preference information of learners to calculate the vector representation of learners. Its deficiency lies in that it does not use the attention mechanism to improve the information quality at the learning resource end. Similar to literature [17], literature [18] focuses on the learning resource end, integrates the neighbour nodes of learning resources to obtain its embedded representation, and enriches the learner embedded representation without utilizing the information of the knowledge graph. The advantage of literature [19] is that both the learner end and the learning resource end are taken into account. However, when aggregating information on the learner end, the demographic information of learners is aggregated by constructing learner attribute maps. This results in the lack of knowledge characteristic information on the learner side, which leads to the lack of semantic richness of the learner's embedded

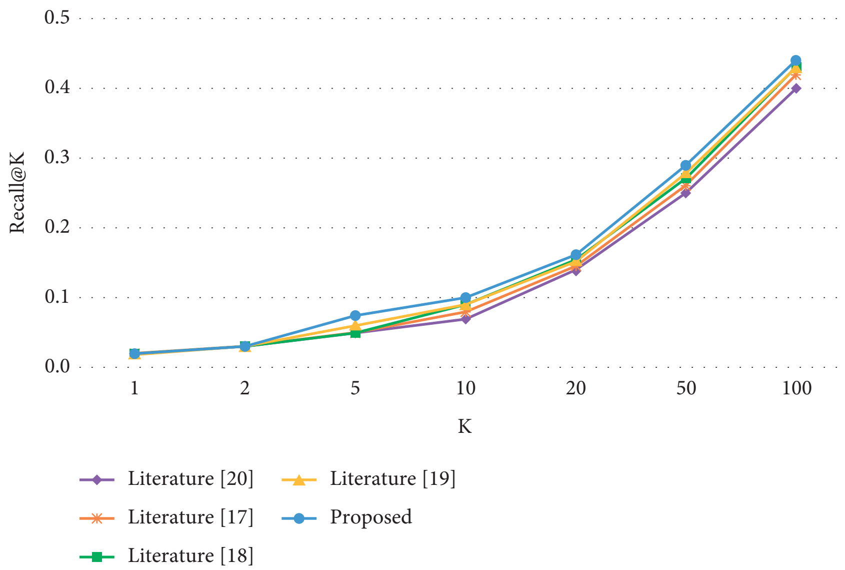


Figure 8: Recall of Top@K.

representation. The model proposed in this paper makes full use of the heterogeneous information of attention mechanism at both the learner side and the learning resource side and integrates the entity information between the learner's interactive learning resource and learning target and its neighbour information into the vector embedded representation of the learner, thus resulting in a significant improvement in performance.

4. Conclusion

Most of the existing improvement methods based on the collaborative filtering model introduce representation learning methods to get better representation vectors of learners and learning resources. The recommendation ability is enhanced by improving the learner-learning resource matching function. However, such work is focused on extracting learner-learning resource interaction information from a single interaction. In order to solve the problem of limited feature extraction of existing models, this paper proposes a recommendation algorithm for online business English learning resources based on the attention mechanism and collaborative filtering model. The model maps the learner vector and learning resource vector to multiple spaces to extract interactive features for score prediction. Multispace can consider the interaction of learner-learning resources from multiple perspectives, and more comprehensive features can better fit learner-learning resource scoring. The teaching practice in student associations proves that the model has a good recommendation ability in business English online learning resources recommendation. The data set in this paper is the interaction matrix between learners and learning resources, with only the index of learning-learning resources as the input of the model. In future work, we try to add more contextual information, social network information, and other auxiliary information to enhance the expressive force of the model.

Data Availability

The labeled data sets used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This study was supported by 2021 First-Class Course Program of Guangxi Normal University for Nationalities Integrated Business English, the authors(YLHHKC202102); 2021 Improvement Project for Young and Middle-aged Teachers' scientific Research Ability in Guangxi Universities Research on the Compilation of Chinese and English Picture Books for Children in the Direction of Zhuang Culture Communication (2021KY0764); and 2019 Project of Basic Education Research Center of Guangxi Ethnic Region A Pilot Study of English Enlightenment for Children of Zhuang Nationality Culture Dissemination (msyjjzx2019B07)

References

[1] W. Wang, L. Guo, and L. He, "Effects of social-interactive engagement on the dropout ratio in online learning: insights from MOOC," *Behaviour & Information Technology*, vol. 38, no. 6, pp. 621–636, 2019.

[2] Y. Xiong and H. K. Suen, "Assessment approaches in massive open online courses: possibilities, challenges and future directions," *International Review of Education*, vol. 64, no. 2, pp. 241–263, 2018.

[3] J. Guo, W. Zhang, W. Fan, and W. Li, "Combining geographical and social influences with deep learning for personalized point-of-interest recommendation," *Journal of*

Management Information Systems, vol. 35, no. 4, pp. 1121–1153, 2018.

[4] J. Jeevamol and V. G. Renumol, "An ontology-based hybrid e-learning content recommender system for alleviating the cold-start problem," *Education and Information Technologies*, vol. 26, no. 4, pp. 4993–5022, 2021.

[5] U. Javed, K. Shaukat, I. A. Hameed, F. Iqbal, T. M. Alam, and S. Luo, "A review of content-based and context-based recommendation systems," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 16, no. 3, pp. 274–306, 2021.

[6] M. K. Najafabadi, A. Mohamed, and C. W. Onn, "An impact of time and item influencer in collaborative filtering recommendations using graph-based model[J]," *Information Processing & Management*, vol. 56, no. 3, pp. 526–540, 2019.

[7] N. A. Albatayneh, K. I. Ghauth, and F. F. Chua, "Utilizing learners' negative ratings in semantic content-based recommender system for e-learning forum," *Journal of Educational Technology & Society*, vol. 21, no. 1, pp. 112–125, 2018.

[8] Z. Sun, M. Anbarasan, and D. Praveen Kumar, "Design of online intelligent English teaching platform based on artificial intelligence techniques," *Computational Intelligence*, vol. 37, no. 3, pp. 1166–1180, 2021.

[9] C. H. Lai, H. W. Lin, R. M. Lin, and P. D. Tho, "Effect of peer interaction among online learning community on learning engagement and achievement," *International Journal of Distance Education Technologies*, vol. 17, no. 1, pp. 66–77, 2019.

[10] A. Tratnik, M. Urh, and E. Jereb, "Student satisfaction with an online and a face-to-face Business English course in a higher education context," *Innovations in Education & Teaching International*, vol. 56, no. 1, pp. 36–45, 2019.

[11] S. D. M. Chethiyar, M. Asad, M. R. U. Kamaluddin, and A. Ali, "Impact of information and communication overload syndrome on the performance of students," *Opción: Revista de Ciencias Humanas y Sociales*, vol. 35, no. 24, pp. 390–405, 2019.

[12] L. Zhou, S. Wu, M. Zhou, and F. Li, "School's out, but class' on', the largest online education in the world today: taking China's practical exploration during the COVID-19 epidemic prevention and control as an example," *Best evid chin edu*, vol. 4, no. 2, pp. 501–519, 2020.

[13] S. Dhawan, "Online learning: a panacea in the time of COVID-19 crisis," *Journal of Educational Technology Systems*, vol. 49, no. 1, pp. 5–22, 2020.

[14] J. K. Tarus, Z. Niu, and G. Mustafa, "Knowledge-based recommendation: a review of ontology-based recommender systems for e-learning," *Artificial Intelligence Review*, vol. 50, no. 1, pp. 21–48, 2018.

[15] J. Xiao, M. Wang, B. Jiang, and J. Li, "A personalized recommendation system with combinational algorithm for online learning[J]," *Journal of Ambient Intelligence and Humanized Computing*, vol. 9, no. 3, pp. 667–677, 2018.

[16] R. R. F. Sinaga and R. Pustika, "Exploring STUDENTS'ATTITUDE towards English online learning using moodle during COVID-19 pandemic at smk yadika bandarlampung," *Journal of English Language Teaching and Learning*, vol. 2, no. 1, pp. 8–15, 2021.

[17] Y. Luo, T. Xu, and Z. Xu, "Recommended method study based on incorporating complex network ripple net[J]," *Journal of Northwestern Polytechnical University*, vol. 39, no. 5, pp. 1070–1076, 2021.

[18] H. Zhang, Y. Wang, C. Chen, and R. Liu, "Enhancing knowledge of propagation-perception-based attention recommender systems," *Electronics*, vol. 11, no. 4, p. 547, 2022.

[19] L. I. Xiang, Y. Xingyao, and Y. U. Jiong, "Double end knowledge graph convolutional networks for recommender systems," *Journal of Frontiers of Computer Science & Technology*, vol. 16, no. 1, p. 176, 2022.

[20] D. Jannach, A. Manzoor, W. Cai, and C. Li, "A survey on conversational recommender systems," *ACM Computing Surveys*, vol. 54, no. 5, pp. 1–36, 2021.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 9382890, 13 pages
https://doi.org/10.1155/2022/9382890

Research Article

English Text Analysis System Based on Genetic Algorithm

Jian Yao and Jin Xu

School of Foreign Languages and Literature, Tianjin University, Tianjin 300350, China

Correspondence should be addressed to Jin Xu; amandaxujin@tju.edu.cn

Received 2 June 2022; Accepted 24 June 2022; Published 11 July 2022

Academic Editor: Jiafu Su

In order to accurately extract the useful information in English, this paper studies English text analysis combined with a genetic algorithm and establishes a text analysis system. In this method, a text tendency analysis algorithm based on a genetic algorithm language model is proposed, and a Doc2vec text feature representation algorithm integrating the LDA model is designed; the parallelization technology of text analysis algorithm is studied, and the parallelization model of the algorithm by using spark big data platform is designed; the process of English text tendency analysis is studied, and a Chinese text analysis system is designed and implemented based on big data platform, including corpus intake, corpus annotation, corpus storage, model training, model verification, and other modules. In order to verify the feasibility of this subject, the accuracy of the Doc2vec text feature representation algorithm of the fused LDA model designed in the prototype system is tested. The experimental results show that the fused text representation model has high recognition degree, and the AUC value of the ROC curve reaches 0.95. At the same time, this paper tests the text analysis-related algorithms involved in the system. The test results show that the parallel algorithm can greatly improve the efficiency of the system.

1. Introduction

Since the birth of the genetic algorithm in the 1970s, many institutions and researchers have conducted extensive and in-depth research on it, achieved many important research results, and rapidly extended its application fields to optimization, search, machine learning, and other aspects. It has gradually developed into a calculation model to solve optimization problems by simulating the natural evolution process [1]. Content-based text information filtering is an important part of machine learning. Genetic algorithm was first applied to machine learning to solve some simple learning problems. For example, the CS-1 system proposed by Holland and Reitman applies the genetic algorithm to solve maze problems for the first time while Goldberg applies the genetic algorithm to engineering control. These studies have produced genetic-based machine learning (GBML) [2]. Text analysis is about the representation of text and the selection of its product features. Text analysis is a key problem in text retrieval and archiving. It counts words extracted from text to represent information as shown in Figure 1. Text has much the same meaning as text. It refers to a data format that contains symbols or numbers. These templates are available in multiple languages such as text and graphics. Words are created by special people, and the content of a book should reflect people's work, thoughts, values, and interests. Thus, the reader's purpose and intent can be determined by identifying the content of the text and converting them from nonstandard texts into documents that computers can recognize and process, i.e., study the texts and develop their mathematical models to interpret and alter the texts. Through the calculation and operation of the model, the computer recognizes the text [3]. Since text is data-intensive, in order to mine useful data from multiple texts, the text must first be converted into process code. Most people use a vector space model to describe vector text, but if the product features come from word segmentation algorithms and word frequency statistics used to represent vector text of different lengths, the length of this vector will be large. This unfinished business not only brings huge overhead to the next task, making the whole process inefficient, but also increases the accuracy of the distribution and grouping algorithms, resulting in unsatisfactory results [4]. Therefore, we need to further refine the vector text by refining the

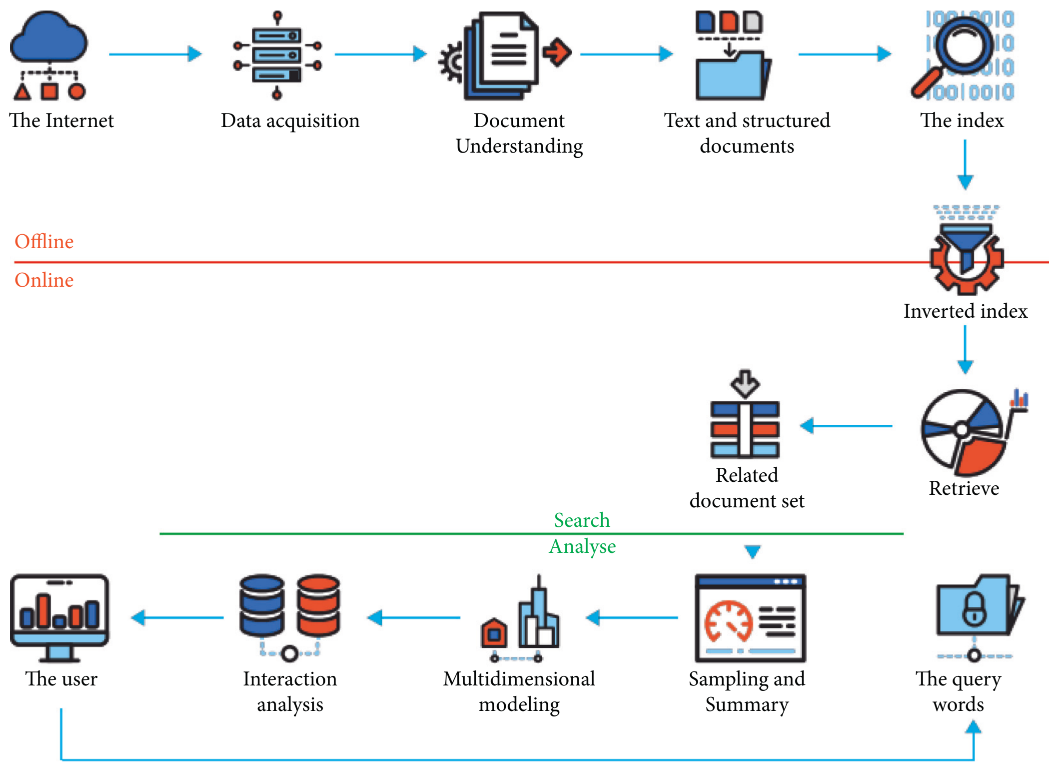


FIGURE 1: This analysis system.

content of the original text to find the most representative text for the text. To solve this problem, the best way is to reduce the dimensionality by selecting a feature.

2. Literature Review

Genetic algorithm is rarely applied to the research of text information generation, especially the application of the genetic algorithm to text information selection. In 2000, Wang and others first applied the genetic algorithm to feature selection [5]. Then, Liu and others introduce the text distribution region selection feature based on the genetic algorithm. Since then, many scientists have devised various improvements [6]. Bai et al. also discuss mutation algorithms and use them to select options [7]. According to Sheluhin et al., the distribution of text is the process of dividing the text into one or more subgroups according to the content. It is an expression. As a set of information training tools, find the relationship model between data features and data categories. These relational models are then used to determine categories of anonymous data [8]. Izrailova and Badaeva believe that in English text classification, the text set becomes a word set after word segmentation, and then, the feature set is obtained by removing the stop words and roughly reducing the dimension. However, the feature set is still a high-dimensional feature space, which is too large for all classification algorithms. Therefore, we are faced with finding an efficient feature extraction method to reduce the dimensionality of feature regions and improve the efficiency and accuracy of distribution [9]. Mufti and others said that the purpose of feature selection is to remove the features that cannot better represent effective information in the feature set so as to improve the classification accuracy and reduce the computational complexity [10]. Meng and others said that in text classification, generally speaking, when the text is expressed in vector form, there may be tens of thousands of feature items in the training text set. It is generally believed that any one of these features has its contribution to the realization of correct classification. However, these large number of features must also contain many interrelated features, which are redundant and can be removed [11]. Tominaga and others believe that too large feature space will make the evaluation of sample statistical characteristics more difficult, thus reducing the generalization ability of the classifier and causing the phenomenon of "over learning." Moreover, the processing of this high-dimensional vector has extremely high computational complexity, especially the so-called "dimension disaster" problem [12]. Therefore, Yue and Wang said that how to retain those features that play an important role in classification and remove redundant features in order to reduce the total number of features, that is, how to carry out dimension reduction, has become an increasingly important research field [13]. Shi described that the representation of text distribution as a process in data filtering, data recovery, archival, digital library, and e-mail distribution has wide application reliability [14].

3. Method

3.1. Feature Selection Dimension Reduction. Feature selection dimensionality reduction is classified according to the concept of algorithms and can be divided into three categories: filtering feature selection, encapsulation feature selection, and embedding feature selection. The filtering feature selection algorithm is independent of the classification algorithm. It directly judges the advantages and disadvantages of text features according to the characteristics of text data and finally selects the excellent text features as the final feature subset. The encapsulated feature selection algorithm needs to preset the classification algorithm to obtain the classification model and indirectly evaluate the classification efficiency of the feature subset by detecting the final effect of the model on the test set. Embedded feature selection algorithm is used to automatically select features in the process of training classification model [15]. The commonly used filtering text feature dimensionality reduction algorithms are described in detail below.

3.1.1. Term Frequency-Inverse Document Frequency (TF-IDF). TF (term frequency) means word frequency in Chinese and IDF (inverted document frequency) means inverse text frequency index in Chinese. The theoretical focus of its algorithm is applied to text feature selection, which is to calculate the reciprocal product of the number of times a single text feature appears in the overall text set and the number of documents [16]. The calculation is shown in the following formula:

$$TF-\mathrm{IDF}_a = \frac{N_{a,A}}{N_A}\cdot \ \log\frac{Z}{Z_a+1}, \tag{1}$$

where $N_{a,A}$ represents the number of times the text feature a appears in document A, N_A represents the total number of words in document A, Z refers to the total number of documents in the corpus data, and Z_a refers to the number of documents including text feature a.

When using the term frequency-inverse document frequency as the method of text feature selection, there are mainly the following two disadvantages: (1) when the data set is skewed, the calculation method of inverse document frequency will be affected by the imbalance of the number of documents, which is difficult to achieve our desired goal, and (2) the scoring standard only takes the contribution of text features to the whole as the weight, ignores the performance ability of text features in a single category, and the ability to distinguish between categories is weak [17].

3.1.2. CHI Square Statistics. After calculating the CHI squared value for letters and categories, the CHI squared value is calculated from large to small. The higher the value, the better the relationship. The CHI square value of text feature a and category P is calculated as shown in the following formula:

$$x^2(a,P) = \frac{N*(AD-BC)^2}{(A+C)*(B+D)*(A+B)*(C+D)}, \tag{2}$$

where A is the number of documents with text feature a and belonging to category p; B is the number of documents with text feature a but not belonging to category P, C is the number of documents without text feature a but belonging to category P, D is the number of documents without text feature a and not belonging to category P, and N is the number of total documents [18]. For multiple distributions, count the squared CHI values for each format contained in the corpus data, and then, assign a mean or higher value based on the squared CHI value of text a. Studies have shown that the squared CHI statistic for the largest breast in multiple distributions is superior to the squared CHI statistic in terms of time and effect [19].

The CHI square statistical algorithm for text feature selection mainly has the following two disadvantages: (1) because the CHI square statistical algorithm based on interclass discrimination does not consider the competition between similar feature words, for example, it does not consider the interference of word frequency distribution between each type of feature words, which reduces the accuracy of its evaluation features and exaggerates the role of low-frequency words; (2) because there is the factor $(AD-BC)^2$ in the formula, if $AD<BC$ occurs in multi-classification, the characteristic words with poor classification effect will be wrongly given high score evaluation, which will interfere with the evaluation expressiveness of characteristic words.

3.1.3. Mutual Information (MI). Mutual information algorithm is a statistical algorithm that shows the correlation between two subjects. In text feature dimensionality reduction, it calculates the relationship between text features and corpus categories. It is used for text feature selection. It is usually used to judge the amount of information associated with text features and various categories as shown in the following formula:

$$\begin{aligned} MI(a,c_j) &= \log\frac{p(a,c_j)}{p(a)\cdot p(c_j)},\\ &= \log\frac{p(a|c_j)}{p(a)}, \end{aligned} \tag{3}$$

where $p(a,c_j)$ represents the probability of the existence of text feature a in Category c_j, $p(a)$ represents the probability of the existence of text feature a in the total number of documents, $p(c_j)$ represents the probability of the existence c_j of category in the total number of documents, and $p(a|c_j)$ represents the probability of the existence of text feature a in category c_j. Let $\{c_1,c_2,c_3,\ldots,c_n\}$ represent the collection of categories in the document, then the mutual information calculation of text features in corpus data is shown in the following formula:

$$MI(a) = \sum_{i=1}^{n} P(c_i)MI(a,c_i). \tag{4}$$

Computational approaches to data sharing are always algorithms with the following disadvantages: (1) ignoring the influence of word frequency factor on features, the formula focuses on the selection of low-frequency words, resulting in the loss of some feature words with high word frequency and strong classification; (2) the feature may have a negative value for the mutual information calculation value of a single category, indicating that the text feature does not exist or exists less in this category, which plays an important role in category judgment, and the value of the data cannot be affected in the computational model of the data sharing algorithm; (3) when the resources of various documents in the data set are unbalanced, the evaluation of text features is not accurate.

3.1.4. Information Gain (IG). The information gain algorithm takes the value brought by the feature to the whole as the evaluation standard and represents the amount of information brought by the feature to the system according to the difference between the amount of information when the system includes feature a and does not include feature a [20]. When calculating the information gain of a single text feature, calculate the difference between the direct line of the classification system when the text feature a is included and the direct line of the classification system when the text feature a is not included, which represents the information gain brought by the text feature a to the classification system and the contribution value of the text feature a. There are two cases without feature a: the first case is that feature a does not exist in the classification system, and the second case is that feature a exists but a has been fixed in the classification system. In the actual calculation process, we use the second method to calculate. At this time, the amount of information is also called "conditional Di," and the condition is that the feature a has been fixed. In the Chinese text classification system, when feature a is fixed, there are two situations: occurrence and nonoccurrence. We use a to represent the occurrence of feature a and $\overline{a}$ to represent the nonoccurrence of feature a. Let $\{c_1, c_2, c_3, \ldots, c_n\}$ represent the collection of categories in the document and $p(c_i)$ represent the distribution probability of various texts. The calculation is shown in the following formula:

$$\begin{aligned} IG(a) = &-\sum_{i=1}^{n} P(C_i)\log_2 P(C_i) \\ &+ P(a)\sum_{i=1}^{n} P(C_i|a)\log_2 P(C_i|a). \end{aligned} \tag{5}$$

The disadvantages of using the information gain method to evaluate the weight of text features are as follows: (1) paying too much attention to document frequency and weakening the attention to word frequency; (2) when the resources of various documents in the data set are unbalanced, the actual evaluation of text features will deviate from the expectation, resulting in inaccurate evaluation [21].

3.2. Language Model and Text Representation Method. Language model is used to model natural language. The traditional language model is a statistical language model, which is a probability distribution function representing language fragments. Its mathematical expression is as follows:

$$p(W) = p\left(w_1^T\right) = \Pi_{t=1}^{T} p(w_t|\text{Context}), \tag{6}$$

where $W = w_1^T = (w_1, w_2, \ldots, w_T)$ represents the language fragment composed of T words $w_1, w_2, \ldots, w_T$ in order and $p(W)$ represents the probability of these words being combined. According to the Bayesian formula, the $p(W)$ chain can be decomposed into the following equation:

$$p\left(w_1^T\right) = p(w_1)\cdot p(w_2|w_1)\cdot p\left(w_3|w_1^2\right)\cdots p\left(w_T|w_1^{T-1}\right). \tag{7}$$

If the context of each word is uniformly recorded as context, the expression of equation (6) can be obtained. Different language models can be formed according to different context division strategies. The design usually includes the N-gram structure, n-pos structure, decision model, maximum entropy model, maximum entropy Markov model, and neural network language model. Different text modeling methods will have different effects. At the same time, each language model also has its own characteristics. The following is an introduction to each language model:

3.2.1. N-Gram Model. The N-gram model determines the occurrence of a word relative to the n words that precede it. The biggest is when $n = 1$, that is, the occurrence of a word only interacts with the word itself [22]. This language model is called context-free model as shown in the following equation:

$$\begin{aligned} p(w_t|\text{Context}) &= p(w_t), \\ &= \frac{N_{wt}}{N}. \end{aligned} \tag{8}$$

When $n \geq 2$, it is called context-dependent model. In general application, it takes $n = 2$ or $n = 3$, that is, Bigram or Trigram. The advantage of N-gram model is that it takes into account the factor of the first $n-1$ words, which have strong meaning in natural semantics. At the same time, using Bigram or Trigram can greatly simplify the calculation scale and improve the efficiency [23]. However, the n-gram language model itself has some limitations. For example, the N-gram model is based on the relationship between the corpus, the corpus is insufficient, and the training result is generally not ideal. Moreover, this model ignores the similarity relationship between words, only considers the relationship between words and context, but does not consider the relationship between words. Secondly, the N-gram model will have a statistical probability of 0 when some tuples have not appeared. This will lead to the probability of the whole language sequence to be 0. In this case, it often needs to be corrected to obtain accurate results.

3.2.2. N-Pos Model. The n-pos model is a language model derived from the N-gram model. The n-pos model is based on the following assumptions: considering the first n words alone is not enough to represent the characteristics of the current word, and the word collocation in natural language is often determined according to the grammatical function of the word. Therefore, n-pos classifies the first n words of the current word according to the grammatical function, and these words determine the probability of the current word. This classification is called Part-Of-Speech, that is, the origin of the n-pos algorithm. The conditional probability formula of the n-pos model is shown in the following equation:

$$\begin{aligned} p(w) &= p\left(w_1^T\right), \\ &= \prod_{t=1}^{T} p\left(w_t | c\left(w_{t-n+1}\right), c\left(w_{t-n+2}\right), \ldots, c\left(w_{t-1}\right)\right). \end{aligned} \tag{9}$$

c is the part-of-speech mapping function, and $c(w_t)$ means to map the word w_t to its part-of-speech category. If a language sequence has T words and K part-of-speech classifications, the conditional probability solution process of n-gram can be changed from T^{n-1} to K^{n-1}, which greatly improves the calculation efficiency, and the improvement of this efficiency will not affect the decline of accuracy.

3.2.3. Maximum Entropy Model. The main idea of the maximum entropy model is as follows: when estimating the probability of random events, if the probability model satisfies certain constraints, then when the constraints are met, the unknowns are meaningless. In this case, the resulting distributions are usually similar, and the entropy of the received distribution is the largest [24]. The probability distribution formula of the maximum direct language model is shown in the following equation:

$$p\left(w_t | \text{Context}\right) = \frac{e^{\Sigma_i \lambda_i f_i(\text{context}, w)}}{Z(\text{context})}, \tag{10}$$

where λ_i is the parameter and $Z(\text{context})$ is the normalization factor.

3.3. Overall Framework Design of Text Tendency Analysis System. This system is a text analysis system designed for the analysis of film review tendency. Its main process design consists of four parts: text preprocessing stage, text storage stage, text analysis stage, and result display stage. The processing flow chart is shown in Figure 2.

Firstly, the system obtains the comment information of the film as the initial corpus, then saves it as the training corpus after preprocessing the film information, and then enters the text analysis stage, including the training and classification of the model. The output is the trained text tendency classification model, shows the accuracy of the model through the display interface, and provides the interface to demonstrate the judgment of the tendency of the text [25]. According to the overall process of the system described above, the overall framework of the system is designed as shown in Figure 3.

The system is divided into three main modules: text preprocessing module, text storage module, and text analysis module. The following is the function introduction of each module:

(1) Text preprocessing module: the text preprocessing module is mainly responsible for text acquisition and processing. The text acquisition uses the customized spider crawler to obtain the text corpus. Compared with the general crawler, the indiscriminate crawling training corpus will produce a lot of noise for the extraction of text features. These noises will greatly reduce the accuracy of text training, thus affecting the final effect. Using customized crawlers, we can crawl the required corpus for analysis according to the characteristics of web pages. For example, for film comment information, the content to crawl includes comment subject, number of likes, comment scoring, comment label, and other information. The general crawler often simply removes the web page tag, leaving the text part as the training corpus. The granularity of such text corpus is very coarse, and the effect after word segmentation and filtering is often difficult to meet the requirements [26]. The differential classification of these information can not only remove the noise influence of the corpus but also facilitate the extraction of text features.

    Another important function of the text preprocessing module is to segment and label the crawling corpus. For text tendency analysis, it is necessary to label the emotional words in the text. The tagging of emotional words is very helpful for the weight calculation of feature extraction. Emotional words are words stored in the emotional dictionary. The emotional dictionary often contains the part of speech, level, subjective, and objective attributes of these words. By judging the emotional words of the text, we can obtain the emotional level of the sentence in the text, and by judging the emotional level of the sentence, we can obtain the emotional level of the text segment, which is of great help to judge the emotional level of the whole text.

(2) Text enclosure: the text storage module saves the text into a special format and saves it into the distributed file system as the input of text analysis. In order to achieve fairness, the processed text needs to be divided into two categories: one is the training corpus with emotional level, and the other is the test corpus with implicit emotional level. Finally, the accuracy of the system is measured by comparing the scores obtained from the analysis of the test corpus with its actual scores.

(3) Text analysis module: the text analysis module uses two methods to represent text from the text. One is to vectorize the language according to standard neural network languages, and the other is to extract the content using grammar. Based on the text vector combined with word vector features and topics, the classification model is trained, and the classification model is finally obtained for analysis.

3.3.1. Design of Text Preprocessing Module. The text preprocessing module includes customized crawler module and word segmentation module, and its module composition diagram is shown in Figure 4.

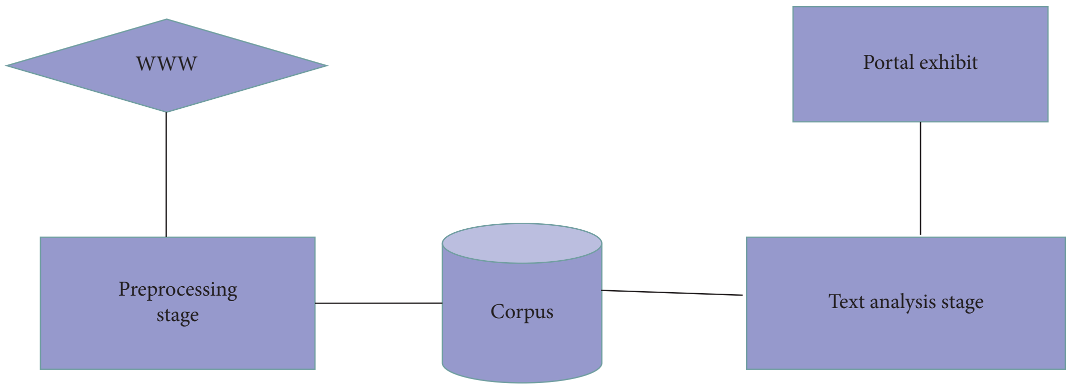


Figure 2: Schematic diagram of text analysis and processing flow.

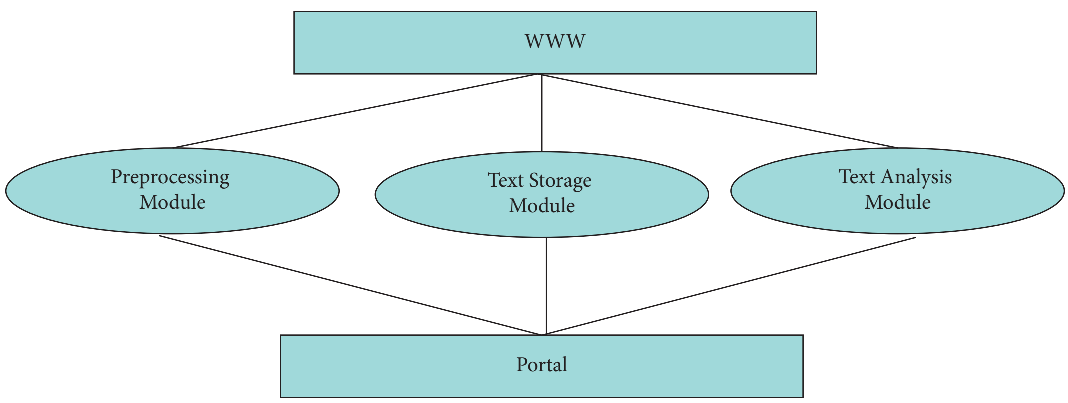


Figure 3: Overall framework of text tendency analysis system.

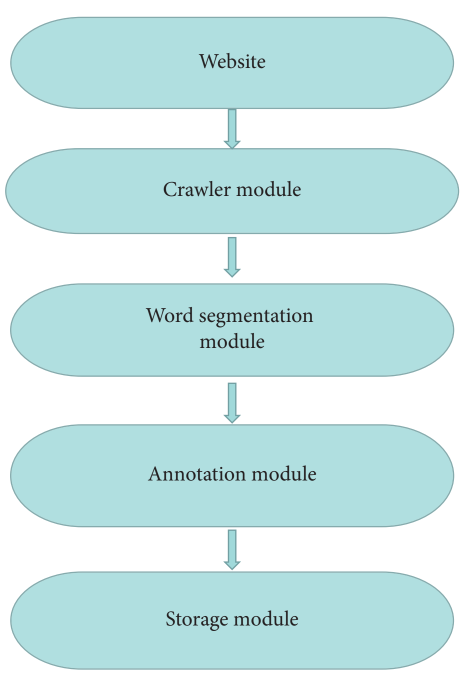


Figure 4: Text preprocessing flow chart.

3.3.2. Design of Text Storage Module. The main function of the text storage module is to store the crawled comment corpus and the preprocessed text. In order to prepare for text analysis, the system uses HDFS as the file system for text storage, saves the positive and negative text to HDFS, and saves the meta-information of the file in the database. When used to generate the model, the text analysis module reads the text meta-information from the database and then downloads the text to be trained from IDFS for local training. HDFS is generally used as the underlying storage system of the spark big data platform. Spark has a special interface to read text from HDFS and convert it into RDD. At the same time, RDD can also be persisted to HDFS as intermediate results. HDFS is a master-slave architecture. HDFS groups contain a personal name and some file nodes. As the owner of the node, the name node is mainly responsible for managing the names of system files and providing scheduling time for users to access data. Data nodes are storage nodes. Usually, each data node corresponds to a physical node in the cluster to manage the file data stored on it [27]. HDFS file system provides a name space and allows users to store data in files. Files are stored on a group of data nodes in the form of data blocks. Figure 5 is the architecture diagram of HDFS.

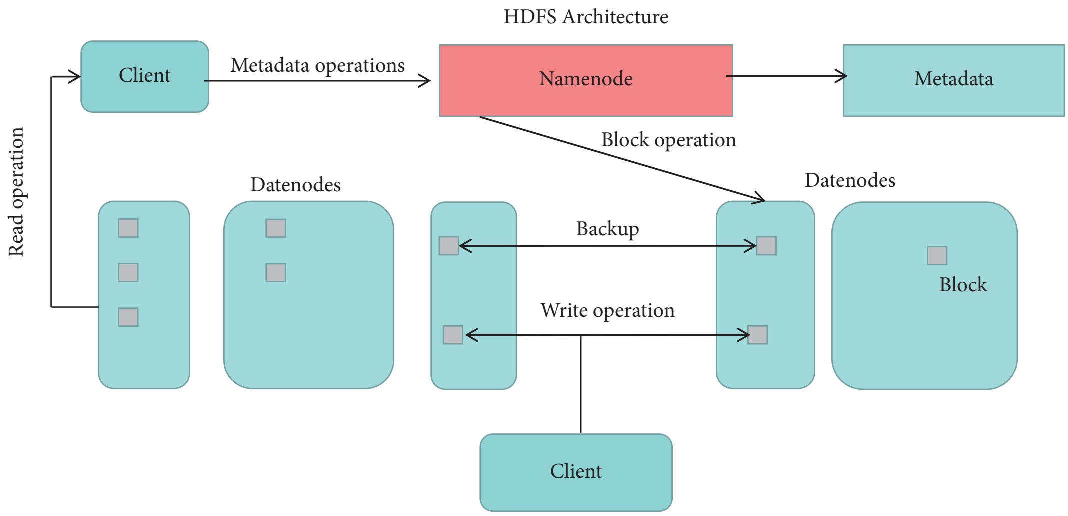


Figure 5: HDFS architecture.

The text storage module is a storage module designed based on HDFS. It adopts the way of small file storage. The crawler classifies the crawled comment text by movie category and saves it to the IDFS file system, then saves the path meta-information of the comment file on HDFS to the database, and saves other information of the file in the database, such as movie name, category, crawl time, and the number of comments. At the same time, the text preprocessing module will obtain the crawled comment information through the database and construct a corpus for text tendency analysis. The constructed corpus will be saved to the text storage system in a fixed directory structure. The text storage module is shown in Figure 6.

Take MySQL as the database information of comment text, and the database table storing comment information is shown in Table 1.

In the crawler crawling process, the text preprocessing module will process the saved comment text at the same time and save the processed comment text to the corpus. The corpus directory structure is organized according to certain rules. Its purpose is to provide the training module with available training corpus and use it to persist the trained model data, which is saved on HDFS.

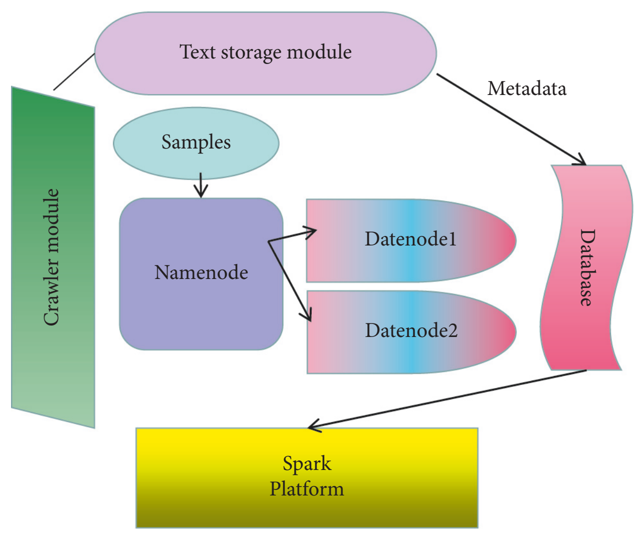


Figure 6: Structure diagram of text storage module.

3.3.3. Design of Text Analysis Module. The text analysis module is used to model the text. The text modeling method adopted by the system is to train the text model by integrating the text vector representation and text topic representation, which absorbs the advantages of both the language model based on statistics and the language model based on neural network. The module flow chart of the text analysis module is shown in Figure 7.

Among them, the Doc2vec algorithm is used for text vector representation, LDA Algorithm is used for text subject representation, and the Doc2vec algorithm is the vectorization of text fragments, which can be used to represent the characteristics of the text. The text topic representation adopts LDA Algorithm. LDA Algorithm is a text topic model algorithm. It extracts the topic of the text, and sparse represents the text. It can also be used as the feature model of the text. The feature vector of the text can be obtained by fusing the two vectors, and then, the text classification model can be obtained by using the SGD classification model. This model is the language model finally used for propensity analysis.

4. Results and Analysis

4.1. Experimental Design and Analysis

4.1.1. Experimental Design and Process. In the overall process design of the experiment, due to more preparations, the overall process is divided into two parts. The first part is the preparation process before text dimensionality reduction using the genetic algorithm, and the second part is the execution process of text feature dimensionality reduction using the genetic algorithm. The preparation process is shown in Figure 8.

Table 1: Comment information database.

Column name	Type	Primary key	Is empty	Explanation
ID	Int	Yes	No	Primary key
MovieName	Varchar	No	No	Movie name
Type	Int	No	No	Movie types
CrawlDate	DateTime	No	No	Crawl time
AveScore	Int	No	No	Star
CommentsCount	Int	No	No	Number of comments
CommentsFilePath	Varchar	No	No	The comment text is in road on HDFS path
Tag	Int	No	No	Has been dealt with

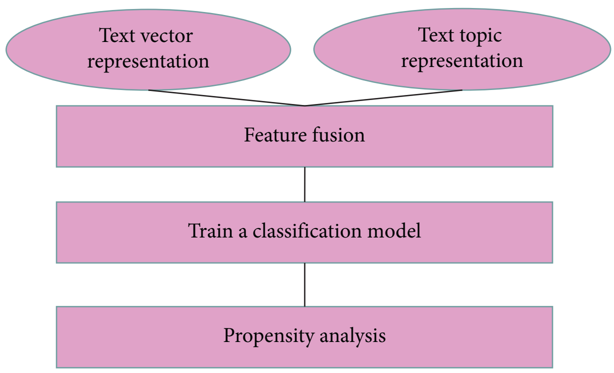


Figure 7: Structure diagram of text analysis module.

The specific operations in the preparation process are as follows:

The first thing to be decided is the classification algorithm used to generate the classification model. According to the characteristics of small data set, high text feature dimension, and secondary classification, the support vector machine with solid theoretical foundation is decided to be used as the classification algorithm of the training data generation model. We also include the preprocessing part of text data into the preparation process. In this step, we need to process punctuation, segment words, and remove stop words. The specific contents are as follows: firstly, the symbol table is made according to the common symbols and special symbols in the corpus, and the symbols in the text are accurately removed. Then, the precise mode in Jieba word segmentation introduced above is used to realize word segmentation and obtain the text feature set. Finally, the stop words in the text feature set are removed according to the stop word table of Harbin Institute of technology. Then, the text representation methods are selected. According to the comparative experimental analysis of three text representation methods in Chapter 3, the text data set used in the experiment has the best classification effect when using text Boolean representation method [28]. Therefore, the text Boolean representation method is selected to represent the text structurally. Finally, the text features need to be preliminarily filtered to obtain the individual gene group in the genetic algorithm because it is found that the classification effect is often poor when using full features to process text data. It shows that there are many redundant or irrelevant item features in the text feature set. Affected by these item features, the expected classification accuracy cannot be achieved. In order to improve the classification performance of the classification model and reduce the spatial complexity of the next genetic algorithm, a step is added before using the genetic algorithm. TF-IDF filtered feature selection dimensionality reduction algorithm is used to preliminarily filter the text features, and the first 10000 of the 19036 text features obtained after word segmentation are selected as the gene set of chromosomes.

The execution flow of dimensionality reduction of text features using the genetic algorithm is shown in Figure 9.

4.1.2. Experimental Data and Parameters. The data used in this experiment is the hotel psychiatric examination written by Professor Tan Songbo, with a total of 10,000 corpora, including 3,000 good corpus and 7,000 negative corpus. In order to reduce the influence of unreliable data distribution, 2000 positive and negative data are used in the training process, and then, 100 positive and negative corpora are selected from the corpus according to the experimental process. This experiment runs on Linux 413 system. The code language of the experiment is Python 3, and the classification algorithm is the support vector machine. Because in support vector machine, RBF kernel function is suitable for high-dimensional nonlinear classification, and the effect is the best in the comparison experiment with other kernel functions, RBF kernel function is adopted in the experiment, the penalty coefficient is set to 400, and gamma is $1/k$ ($k =$ the number of text features selected by individuals). In the parameter setting of the genetic algorithm, we set the number of individuals in the population to 20, the population termination algebra to 160 generations, the selection operator to RWS (roulette selection algorithm), the crossover probability to 0.7, and the mutation probability to 0.0001 as shown in Table 2.

4.1.3. Experimental Results and Analysis. As shown in Figure 10, the black line is the line graph of fitness function of the best individual in each generation of population, and the red line is the average line graph of fitness function of all individuals in each generation of population. From the broken line diagram, it can be found that both the optimal value and the mean value show a gradual upward trend in the early stage of population reproduction. We regard the difference between the two broken lines as the distribution of individuals in the solution space. It can be found that when using the traditional genetic algorithm to reduce the

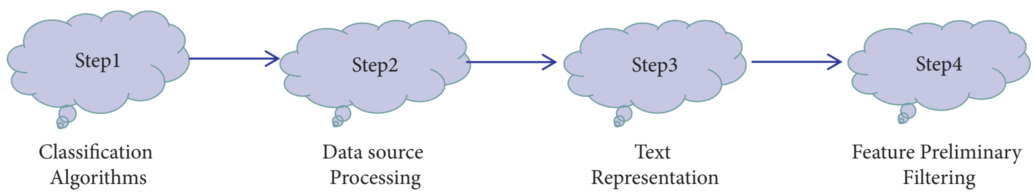


Figure 8: Preparation process.

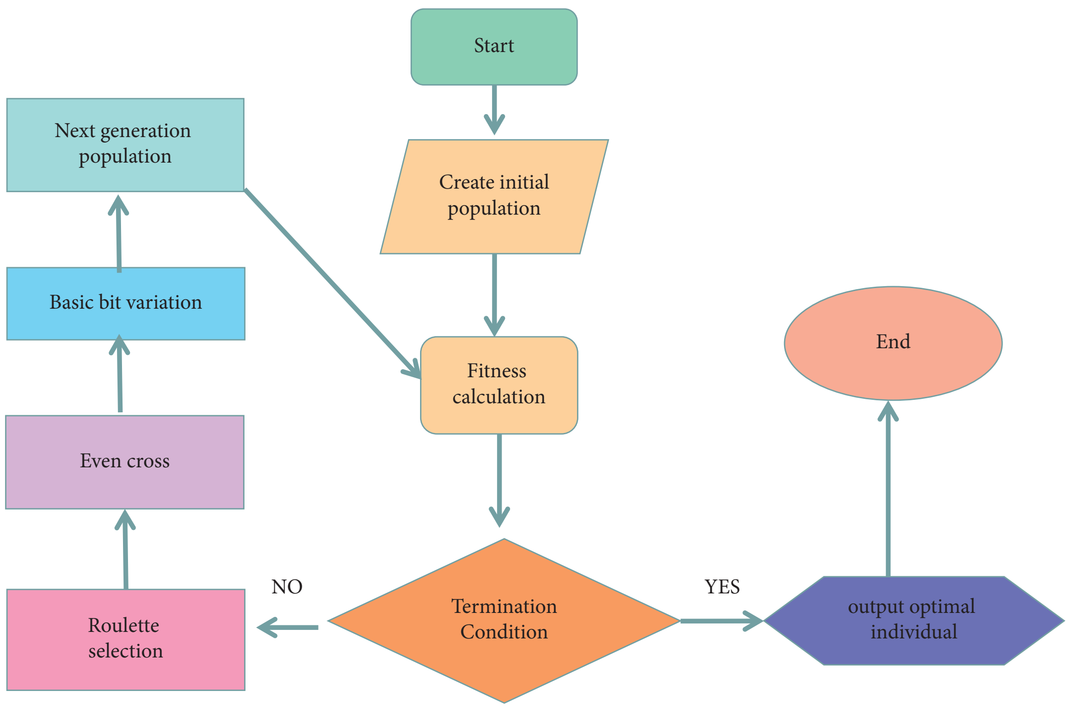


Figure 9: Execution algorithm flow.

Table 2: Experimental parameters.

Parameter name	Parameter value
Kernel function	RBF
Penalty coefficient	350
Gamma	$1/d$ (d = feature dimension)
Population size	20
Iteration number	150
Selection technique	RWS (roulette selection algorithm)
Crossover type	UC (even cross)
Crossover rate	0.8
Mutation type	SM (basic bit variation)
Mutation rate	0.0002

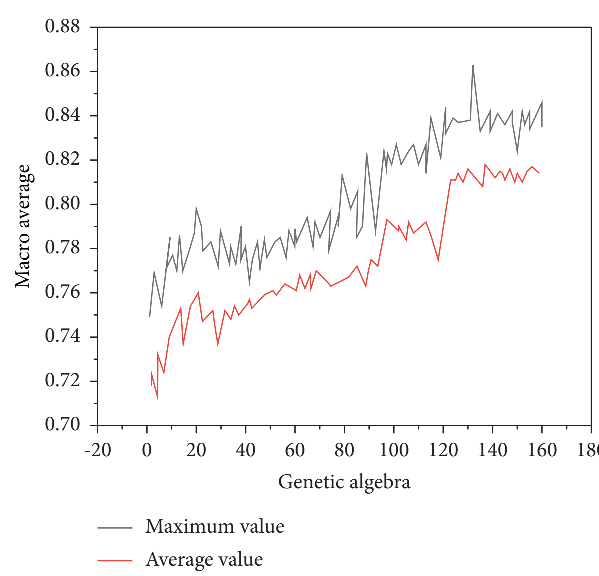


Figure 10: Population performance of different generations.

dimension of text features, the distribution of individual solutions formed by different text feature subsets in the solution space changes constantly between dispersion and concentration, and the distribution of individual solutions in the population is relatively scattered in the periods of 1 to 35 generations and 93 to 125 generations. During the 36–92 generation and 126–160 generation, the distribution of individual solutions is more concentrated. The dispersion after each concentration is a process of jumping out of the local optimum. Under the condition that the maximum reproduction algebra is 160, we can regard the last 126 to 160 generations as the last population convergence part. In the last population convergence part, with the help of the green broken line, we can easily find that the individual with the maximum fitness function value in the reproduction process appears in the 132 generation population, and this individual is the optimal solution we are looking for. Comparing it with the text feature subset obtained by the filter algorithm TF-IDF and CHI at the optimal macro average value, it can be found that the optimal individual obtained by the genetic algorithm has better recall rate, accuracy, and macro average value than a single filter algorithm, in which the recall rate is significantly improved.

Table 3 shows the comparison of recall rate, accuracy rate, macro average value, and characteristic dimension when the macro average value is obtained based on CHI, TF-IDF, and GA. It can be found that GA obtains the highest macro average value when the dimension is 5648, CHI obtains the highest macro average value when the dimension is 8000, and TF-IDF obtains the highest macro average value when the dimension is 9000 (CHI and TF-IDF obtain the highest macro average value through multiple comparisons by dichotomy). The experimental results show that the highest macro average value of CHI is similar to that of GA, the macro average values of both are higher than those obtained by TF-IDF method, and the recall rate of GA is significantly higher than that of CHI and TF-IDF of filter algorithm. Compared with the dimensionality reduction effect at this time, it is obvious that the dimensionality reduction effect of GA is much better than that of CHI and TF-IDF. As shown in Figure 11, the optimal dimension numbers of GA, TF-IDF, and CHI are 5648, 9000, and 8000, respectively. Compared with the dimension number 19033 after word segmentation, the dimension reduction rates of GA are 70.3%, 52.7%, and 58%, respectively. The dimension reduction rates of GA are 17.6% and 11.7% higher than those of TF-IDF and CHI, respectively. The dimension reduction effect is remarkable. By dividing the experiment from the hotel review data, the experiment shows that this paper is based on the CHI filtering algorithm, TF-IDF filtering algorithm, and feature selection algorithm based on genetics and the text feature selection algorithm based on genetic algorithm. In the case of macroscopic media (such as distribution effects), the dimensionality reduction ability is the best, and the distribution efficiency is good.

4.2. Research and Analysis of Text Feature Dimensionality Reduction Based on Improved Genetic Algorithm. In order to solve the problems of the TF-IDF algorithm and enhance the class discrimination ability of the algorithm, the mutual information filtering feature selection algorithm is introduced to calculate the text feature weight, and the traditional TF-IDF algorithm is improved. Through mutual information algorithm, we can get the information weight value between feature words and single category documents. It is planned to replace the calculation significance of the number of documents with feature words in TF-IDF with the dispersion of feature words and the weight value of each category information. Formula (11) calculates the average value of mutual information of feature word a among various texts.

Table 3: Comparison of GA, CHI, and TF-IDF when macro average is the best.

	Precision (%)	Recall (%)	Macro-F (%)	Dimension
GA	85.50	93.29	87.30	5545
CHI	85.50	87.07	87.3	8220
TF-IDF	85	87.87	88.9	9560

$$\overline{MI} = \frac{1}{j}\sum_{1}^{j} MI\left(a, c_j\right). \tag{11}$$

In formula (12), by calculating the standard deviation of the mutual information weight between the feature word and various texts, the dispersion degree of the weight value of the feature word and each category of information is expressed.

$$s = \sqrt{\frac{1}{j}\sum_{1}^{j}\left(MI\left(a, c_j\right) - \overline{MI}\right)^2}. \tag{12}$$

Finally, based on TF-IDF and MI, a new text feature weight calculation method with the overall contribution ability and category contribution ability of text features is obtained as shown in the following formula:

$$\mathrm{FUSION}_a = \frac{N_{a,A}}{N_A}\log\left(1 + e^s\right). \tag{13}$$

In order to verify that the text features selected by the filtering selection algorithm of multirule fusion have better classification effect, the following experiments are carried out. TF-IDF, MI, and fusion algorithms are used to calculate the weight value of the experimental data, respectively, and the top 10000 text features with high weight value are filtered. The text Boolean representation method is used for text representation, and support vector machine is used as the classification algorithm to generate the model for experimental effect comparison. The data used in the experiment is the same as the experimental data used in Chapter 4. It is the hotel comment emotion analysis corpus collected by Professor Tan Songbo. There are 4000 training sets and 200 test sets. The kernel function of support vector machine uses RBF, the penalty coefficient is set to 400, and the gamma is 1/10000.

4.2.1. Experimental Results. The experimental results are shown in Table 4. The first 10000 features are obtained by using the new filtering algorithm of multirule fusion. The effect of the model after training on the test set is better than that of MI and TF-DF. The accuracy, recall, and macro average are more than 80%. Compared with MI and TF-IDF, the accuracy, recall, and macro average are 8% and 3% higher, respectively, 5% and 4% higher, respectively; and 5.9% and 2.9% higher, respectively. The feasibility of the algorithm in the experimental data is proved. Therefore, in the preparation process, the multirule fusion filtered feature

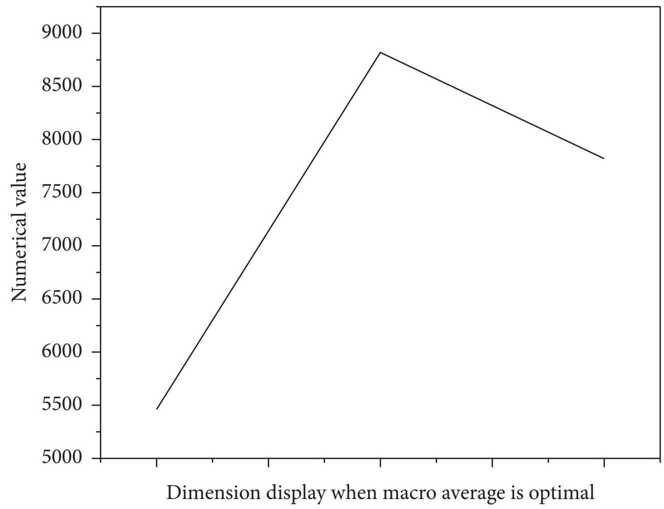


Figure 11: Best dimension display.

Table 4: Effect comparison of gene group selection algorithm.

	Precision (%)	Recall (%)	Macro-F (%)
MI	70	73	76
TF-IDF	75	77	75.4
FUSION	80	81	81.9

selection algorithm is used to filter the text features to form the individual gene group to be selected.

4.2.2. Design and Process. In order to improve the text feature selection ability of the genetic algorithm, improve the convergence speed, and achieve the effect of dimensionality reduction, the preparation process and execution process of text feature dimensionality reduction using the traditional genetic algorithm in the previous chapter are improved step by step. The overall flow chart of the improved algorithm is shown in Figure 12.

In the preparation process stage, in order to make the text features obtained after preliminary filtering have better performance ability and reduce the omission of excellent text features, TF-IDF algorithm is improved. According to the problems of the TF-IDF algorithm and the advantages of the mutual information algorithm, the two are fused to form a new multirule fusion filtering feature selection algorithm, which makes the text feature gene group obtained after preliminary filtering have better expressiveness and persuasion.

4.2.3. Experimental Analysis. Figure 13 is a broken line diagram of individual performance of different generations in the improved genetic algorithm, in which the black broken line represents the maximum value of individual fitness function in each generation, and the red broken line represents the average value of all individual fitness functions in each generation. It can be found from the red broken line that when the population continues to multiply, the

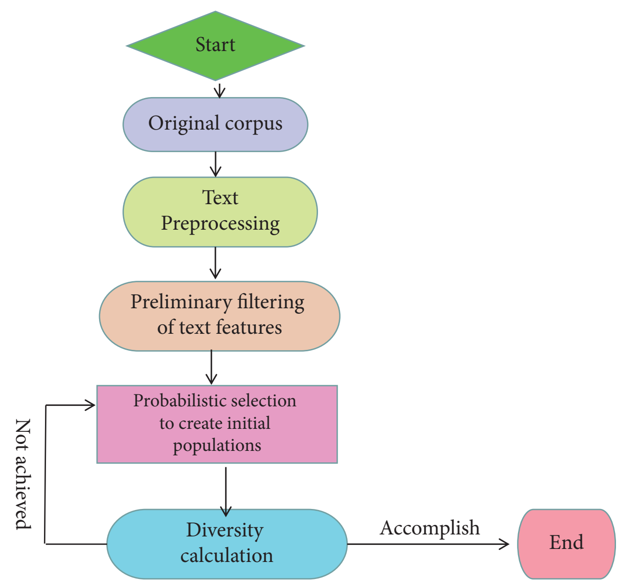


Figure 12: Improved algorithm flow.

individual's external performance is constantly improving, and the improvement is relatively gentle most of the time. Measuring the difference between the black broken line and the red broken line represents the dispersion degree of the individual population. It can be found that the general difference within the stage range is gradually shrinking, which is in line with our expectations for population reproduction, that is, the population gradually converges in the reproduction process. When the population reproduces to 86 generations, the optimal individual appears, and the fitness function value is 0.8925.

Figure 14 is a comparison diagram of the energy function of the genetic algorithm and the genetic improvement algorithm. By comparing the average physical activity cost of each population, it can be seen that the genetic algorithm and the improved algorithm in this form of physical activity

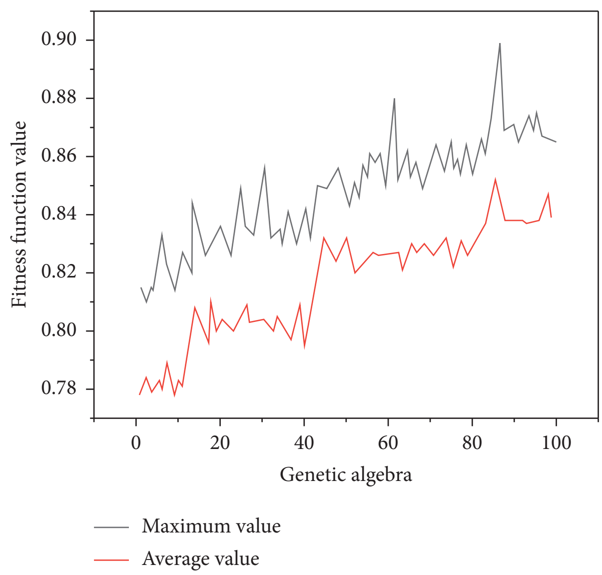


Figure 13: Individual performance of improved genetic algorithm.

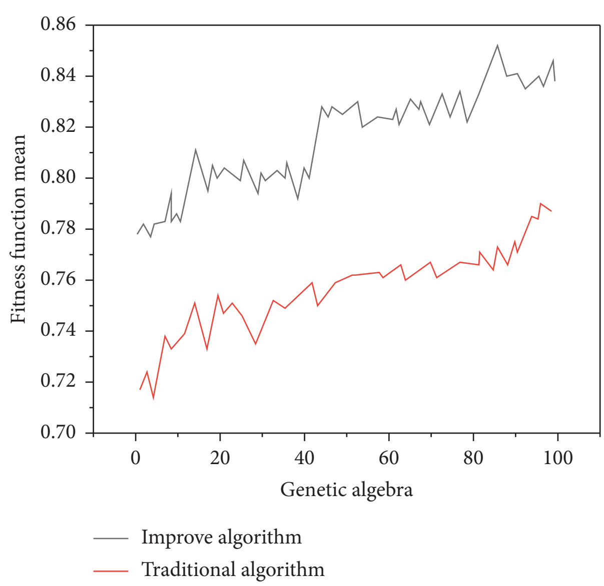


Figure 14: Comparison of fitness function between simple genetic algorithm and improved genetic algorithm.

usually show a significant increase in the growth rate of 100 generations. In the first stage of population breeding, all the benefits of genetic improvement algorithms over traditional genetic algorithms are the result of many different fusion algorithms and methods that select workers in repair, and therefore improve the performance of individuals in the first population. In the genetic algorithm, the energy function increases gradually, and the increase is not obvious, and the frequency of the population mean decreases, while in the improved genetic algorithm, the population mean increases by 3 times. It can be seen that the improved genetic algorithm is better than the traditional genetic algorithm.

The experimental results show that the features screened by the improved genetic algorithm have better feature screening ability, larger dimension reduction range and higher accuracy, recall, and macro average than the single TF-IDF algorithm and CH algorithm. It is faster than the traditional genetic algorithm and can find the optimal individual with better dimension reduction effect and classification performance in a shorter time.

5. Conclusion

The development of the Internet is more and more rapid. With the continuous progress of science and technology, mankind will slowly enter the intelligent era. A series of applications based on text analysis will gradually affect people's way of life. However, the field of text analysis is still under exploration, and the old research methods are no longer suitable for the current needs. Based on the latest research results and combined with the genetic algorithm, this paper makes a systematic research on the field of text analysis. In this paper, the experimental process is clarified in a phased way. After the relevant operations of the preparation process, the dimension of text features is reduced by the genetic algorithm, so that the dimension-reduced text features can achieve better results in the application of text classification, and is based on the traditional genetic algorithm. According to the problems of convergence speed and local optimization, the relevant steps are improved, and the effectiveness of the improvement is proved by experiments. The main improvements include the following two aspects:

(1) Improvement in the preparation process stage: the filter feature selection algorithm based on the multirule fusion is used to select gene groups, which avoids the problems of the TF-IDF filter feature selection algorithm in gene group selection, and then, the rationality of the improvement is verified by classification performance evaluation.

(2) Improvement of execution process stage: the generation mode of individuals in the initial population is changed, the operation of population diversity calculation is increased, and the fitness function is modified to increase the influence ability of dimension. Finally, in order to speed up the convergence speed, the probability of the constant crossover operator and variation operator is changed into an adaptive mode. The experimental results show that the improved genetic algorithm improves the convergence speed and local optimization.

Data Availability

The labeled data set used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported by the School of Foreign Languages and Literature, Tianjin University.

References

[1] A. A. Liednikova, D. V. Shypik, and P. I. Bidyuk, "Project risk analysis using text data mining of comments in project management system jira," *System Research and Information Technologies*, vol. 2020, no. 2, pp. 121–136, 2020.

[2] Z. Li, K. Guo, M. Liao, A. Zhao, M. Tian, and Y. Wang, "Micro-hybrid energy storage system capacity based on genetic algorithm optimization configuration research," *International Core Journal of Engineering*, vol. 6, no. 2, pp. 78–83, 2020.

[3] C. C. Hsieh and J. Z. Chiu, "System review: a text analysis on supply chain finance," *Universal Journal of Management*, vol. 8, no. 2, pp. 29–32, 2020.

[4] T. Rizvi, "Identification of mind-set of students through web based basic psychological text and graphical analysis system," *International Journal of Management and Humanities*, vol. 4, no. 7, pp. 33–36, 2020.

[5] P. Wang, Y. Zhang, and H. Yang, "Research on economic optimization of microgrid cluster based on chaos sparrow search algorithm," *Computational Intelligence and Neuroscience*, vol. 2021, no. 3, 18 pages, Article ID 5556780, 2021.

[6] H. Liu, H. Liu, X. Wang, W. Shao, X. Wang, and J. Du, "Smartmeeting: an novel mobile voice meeting minutes generation and analysis system," *Mobile Networks and Applications*, vol. 25, no. 2, pp. 521–536, 2020.

[7] H. Bai, H. Yu, G. Yu, L. Rocha, and X. Huang, "Analysis on an auto increment detection system of Chinese disaster weibo text," *Journal of Universal Computer Science*, vol. 27, no. 2, pp. 230–252, 2021.

[8] O. I. Sheluhin, D. V. Kostin, and M. G. Gorodnichev, "Multiclass classification of anomalous states of computer systems by means of intellectual analysis of system journals," *Automatic Control and Computer Sciences*, vol. 54, no. 6, pp. 549–559, 2021.

[9] E. S. Izrailova and A. S. Badaeva, "Analysis of the speech signal quality of the Chechen speech synthesis system," *Automatic Documentation and Mathematical Linguistics*, vol. 55, no. 2, pp. 74–78, 2021.

[10] M. D. Mufti, M. Y. Zargar, and A. W. Kumar, "Adaptive predictive control of flywheel storage for transient stability enhancement of a wind penetrated power system," *International Journal of Energy Research*, vol. 46, no. 5, pp. 6654–6671, 2022.

[11] Y. Meng, Y. Liang, Q. Zhao, and J. Qin, "Research on torsional property of body-in-white based on square box model and multiobjective genetic algorithm," *Mathematical Problems in Engineering*, vol. 2021, no. 41, pp. 1–13, Article ID 7826496, 2021.

[12] M. Tominaga, M. Okajima, M. Yamagishi, M. Shinagawa, J. Katsuyama, Y. Matsumoto et al., "Noise analysis of an electro-optic sensor system," *Optical Review*, vol. 28, no. 6, pp. 704–715, 2021.

[13] M. Yue and X. Wang, "Research on control strategy of ship energy management system based on hybrid ga and pso," *International Core Journal of Engineering*, vol. 6, no. 5, pp. 185–193, 2020.

[14] R. Shi, "Research on data mining system based on artificial intelligence and improved genetic algorithm," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 8, pp. 1–12, 2020.

[15] Y. M. Yeshilbashian, A. A. Asatryan, and T. G. Ghukasyan, "Plagiarism detection in Armenian texts using intrinsic stylometric analysis," *Proceedings of the Institute for System Programming of RAS*, vol. 33, no. 1, pp. 209–224, 2021.

[16] D. M. Karshiyevich and H. H. Syed, "Basic principles of creating software system to control and correct errors in text," *SSRN Electronic Journal*, vol. 7, no. 11, pp. 8–14, 2020.

[17] Y. Chen, H. Jiao, H. Zhou, J. Zheng, and T. Pu, "Security analysis of qam quantum-noise randomized cipher system," *IEEE Photonics Journal*, vol. 12, no. 99, p. 1, 2020.

[18] Z. Kafadar, "A geophone-based and low-cost data acquisition and analysis system designed for microtremor measurements," *Geoscientific Instrumentation Methods and Data Systems*, vol. 9, no. 1, pp. 365–373, 2020.

[19] B. J. Galli and L. C. Wood, "How to apply system analysis and system thinking to lean six sigma initiatives," *International Journal of Service Science, Management, Engineering, and Technology*, vol. 12, no. 4, pp. 1–25, 2021.

[20] A. Chaudhuri, N. Sinhababu, M. Sarma, and D. Samanta, "Hidden features identification for designing an efficient research article recommendation system," *International Journal on Digital Libraries*, vol. 22, no. 6, pp. 1–17, 2021.

[21] P. Chang and H. J. Tsai, "Text-image complementarity and genre in English as foreign language textbooks," *Semiotica*, vol. 2022, no. 244, pp. 53–80, 2022.

[22] O. Agarwal, Y. Yang, B. C. Wallace, and A. Nenkova, "Interpretability analysis for named entity recognition to understand system predictions and how they can improve," *Computational Linguistics*, vol. 47, no. 1, pp. 1–24, 2021.

[23] P. J. Younse, J. E. Cameron, and T. H. Bradley, "Comparative analysis of model-based and traditional systems engineering approaches for arCHItecting a robotic space system through automatic information transfer," *IEEE Access*, vol. 9, no. 99, p. 1, 2021.

[24] Z. N. Al-Kateeb and M. Jader, "Encryption and hiding text using dna coding and hyperchaotic system," *TELKOMNIKA Indonesian Journal of Electrical Engineering*, vol. 19, no. 2, pp. 766–774, 2020.

[25] A. L. Mahmood, A. M. Shakir, and B. A. Numan, "Design and performance analysis of stand-alone pv system at al-nahrain university, baghdad, Iraq," *International Journal of Power Electronics and Drive Systems*, vol. 11, no. 2, pp. 921–930, 2020.

[26] F. N. Al-Wesabi, S. Alzahrani, F. Alyarimi, M. Abdul, and M. M. Almazah, "A reliable nlp scheme for English text watermarking based on contents interrelationship," *Computer Systems Science and Engineering*, vol. 37, no. 3, pp. 297–311, 2021.

[27] C. Payant and P. Bell, "Very easy, it's an English class, therefore they should not rely on a French text: English language teachers' beliefs regarding l1 use for literacy instruction," *Language TeaCHIng for Young Learners*, vol. 4, no. 1, pp. 143–170, 2022.

[28] A. Kusuma, M. H. Santosa, and I. Myartawan, "Exploring the influence of blended learning method in English recount text writing for senior high school students," *Jet*, vol. 6, no. 3, pp. 193–203, 2020.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 2339845, 10 pages
https://doi.org/10.1155/2022/2339845

Research Article

The Impact of Entrepreneurial Orientation on Firm Performance: The Multiple Mediating Roles of Competitive Strategy and Knowledge Creation Process

Liping Yang and Vesarach Aumeboonsuke

International College, National Institute of Development Administration, Bangkok 10240, Thailand

Correspondence should be addressed to Vesarach Aumeboonsuke; vesarach.a@nida.ac.th

Received 10 May 2022; Revised 12 June 2022; Accepted 21 June 2022; Published 6 July 2022

Academic Editor: Jiafu Su

The strategy and knowledge creation process help entrepreneurial SME managers achieve sustainable competitive advantage and firm performance. This study aims to investigate the multiple mediating roles of competitive strategies and knowledge creation processes implemented by Chinese SMEs in Thailand between their entrepreneurial orientation and firm performance. Structural equation model is used to analyze data collected from Chinese SMEs in Thailand. Results showed that (1) there is a positive relationship between perceptions of entrepreneurial orientation and perceptions of firm performance, (2) competitive strategy and knowledge creation process individually play a mediating role between entrepreneurial orientation and firm performance, and (3) competitive strategy and knowledge creation process play a chain mediating effect between entrepreneurial orientation and firm performance. This hypothesis is verified by a chained multiple mediation model, and this model has theoretical significance for understanding the relationship between perceived entrepreneurial orientation and firm performance. The competitive strategy has positive effects on SME performance and knowledge creation expands the firm's valuable resources and actively updates firm performance.

1. Introduction

During the global COVID-19, the economies of countries and regions around the world have experienced a continuous decline. International trade and investment have been drastically reduced, and uncertainty factors have continued to increase. This is a great challenge for China's foreign investment and cooperation. Thailand's society is relatively stable, with clear policy formulations, low wage costs for employees, an open and inclusive business environment, and its level of economic development ranks high among Association of Southeast Asian Nations countries, and can enjoy zero-tariff treatment in ASEAN countries.

To enhance national competitiveness, the Thai government proposed the "Thailand 4.0" strategy and the national economic development plan of the "Eastern Economic Corridor" in 2016, which are highly compatible with the "Belt and Road" initiative promoted by China. To promote cooperation between the two countries, Thailand and China cofounded and established the Amata City Rayong Industrial in 2006. It is also one of the first overseas economic and trade cooperation zones in China. Now there are more than 180 Chinese firms in the Amata City Rayong Industrial to build factories in Thailand with a local staff of 40,000 people, China's core technology administrators and workers numbering more than 400 people. It led to the Chinese firms' investment exceeding $4 billion in Thailand , and the cumulative gross industrial output value exceeding $19 billion. It has created opportunities for Chinese firms to develop in Thailand. The Thai Investment Promotion Board (BOI) released investment data for the first quarter of 2021, showing that firms in the Amata City Rayong Industrial accounted for 40% of China's total direct investment (FDI) in this quarter. The Amata City Rayong Industrial development and construction planning area of 12 square kilometers has developed more than 8 square kilometers. It has become the largest industry in China in Thailand and ASEAN cluster centers and manufacturing export base. The optimization of industrial structures, improving the level of economic development, and local for local economic and social development in Thailand has injected strong driving force and promoted economic prosperity and development of the country. During the epidemic prevention and control period, the Amata City Rayong Industrial explored a new model of "cloud investment" and "cloud service" with the use of digital technology. 20 new firms were admitted to the Amata City Rayong Industrial in 2020, and 7 Chinese firms entered the park from January to August 2021. Most of the 187 firms in the Amata City Rayong Industrial are entrepreneurial SMEs established less than 10 years ago, mainly in auto parts, machinery, manufacturing, new energy, and other industries. Most of the factories of the entrepreneurial Chinese SMEs in Thailand are mainly composed of foreign workers from Thailand, Myanmar, Cambodia, and other countries.

The firm's resource-based view theory (RBVT) describes the firm as a heterogeneous bundle of specific, difficult to imitate abilities and resources [1–4]. Amit and Schoemaker [5] thought that resources are the stock of available elements owned or controlled by a firm, and capability is based on information, visible or invisible process, which is unique to a firm through the complex interaction between firm resources developed over time. Previous research literature has clearly and consistently found that firms face a huge risk of failure when they are young SMEs. From a firm's RBVT, if the firm cannot yield self-sustaining levels to organize rent, it will fail. The key challenge for new firms is to build real resources or abilities before the initial assets are exhausted [6].

The firm's resource-based view theory (RBVT) and dynamic abilities (knowledge creation process) to explain firm resource contribution to firm performance (FP) have been outstanding [7]. They believe that a single firm that owns resources can determine its market status within the forces of a dynamic environment [8]. This dynamic capability is mainly embedded in formal activity aimed at updating firm performance [8, 9].

Top managers use purposeful knowledge flows to accelerate internal innovation to improve competitive advantage. Measuring firm performance through entrepreneurial orientation (EO), competitive strategy, and knowledge creation processes is an important issue for researchers and entrepreneurs in today's business environment. Entrepreneurial Chinese SMEs in Thailand often face a problem when investing in the low productivity and low motivation of their local employees, which may lead to a decline in firm performance (FP). Therefore, based on resource-based view theory (RBVT), this study aims to investigate whether Chinese SMEs in Thailand improve corporate performance through the process of competitive strategy and knowledge creation to maintain market share and continue strategic investment in the future, thereby improving the performance of entrepreneurial Chinese SMEs overseas.

2. Literature Review

The resource advantage theory claims that EO is a resource that helps firms gain a competitive advantage [10, 11]. EO development needs members of the organization to join in extensive knowledge action. With resource advantage theory, knowledge has the features of tacit and immovability, and it is not easy to transfer and disperse [10, 12].

Grounded on knowledge creation theory, knowledge is a process of socialization, externalization, combination, and internalization, namely SECI [13–15]. This dynamic knowledge transfer in SECI can improve the firm's ability to implement strategic objectives and fulfill firm performance for instance product renovation or process amelioration [16, 17]. Therefore, the knowledge creation process (KCP) of the formation of EO and activation plays an important role. It can promote the transformation of EO into knowledge assets shared by organization members, thus improving firm performance.

Firm employees can study together and exchange knowledge to better comprehend the entrepreneurial type and mission expressed by their conceptions and ideas. Integrate and disseminate entrepreneurial practice and activities everywhere in the firm to acquire additional knowledge applications. Firms can turn entrepreneurial orientation into actual actions while turning that knowledge into valuable assets to drive new product development or marketing campaigns [13, 18]. From the knowledge creation theory [14], the purpose of this study is to test how EO affects FP through the knowledge creation process. Nonaka [15] confirmed organizations from the perspective of how organizations create information and knowledge and develop the basic elements of organizational knowledge creation theory.

2.1. The Influence of Entrepreneurial Orientation on Firm Performance. In this study, EO improves firm performance [19]. Although EO on performance growth of short-dated organizations gradually decreases over time, it is related to short-term or long-term firm performance growth [20]. EO can improve a firm's capability to improve learning orientation [21]. Firm performance (FP) is seriously affected by the depth as well as breadth of management associations [22]. Firms take the lead in designing and introducing new projects, technologies, or innovativeness that will be operated by the firms practicing EO, and typically achieving outstanding performance [23]. In a highly competitive business circumstance, firms adopt EO to improve performance [24].

Lumpkin and Dess [25] explained the nature of EO; they analyzed an alternative model that included tests of mediating, interaction, moderation, and independent effects. Lumpkin and Dess [26] found that aggressive firms perform better than competitive and aggressive firms in a dynamic environment of rapid change and uncertainty. Competitive firms perform better in a hostile environment where resources are limited and competition is intense. The dimension of EO is independent, so the effectiveness of a firm's strategic plan is often based on the organization's environment and the organization itself. EO positively contributes to performance improvement in any setting, but EO may not imply immediate success. Wiklund and Shepherd [27] found EO can strengthen the role of knowledge resources in promoting firm performance. Rezaei and Ortt [28] found that enterprises need to develop the internal logical sequence of design, sales, and production, and this logic is very good for the improvement of firm performance. Wiklund and Shepherd [27] found that empirical support that entrepreneurial orientation may be performance and differentiation improve strategic orientation.

Hypothesis 1. Entrepreneurial orientation is positively associated with the firm performance of entrepreneurial Chinese SMEs in Thailand.

2.2. Mediating Role of Competitive Strategy. In today's competitive society, lazy and risk-averse firms will soon be eliminated by society [29]. When entering a new market, a firm must take risks and avoid some unnecessary risks so as not to lose market share and misposition [30]. How to choose the right competitive strategy is the key to the success of a small business [31, 32]. There is a significant relationship between high-return strategy and competitive advantage and performance [30]. Although differences in the framework can be used in the classification of firm strategy [33, 34], but the widely accepted model is Porter's model [35–37]. It includes cost leadership strategy (CLS), differentiation strategy (DS), cost focus strategy (CFS), and differentiation focus strategy (DFS), in fact, CLS and DS belong to the core of Porter's model, they together define the competitive advantage. And CFS and DFS together define the market or product [38], although firms can add them to other strategies by searching deeply for customers or markets. But they may also fail to provide above-average returns [39]. Therefore, we should focus on these two dimensions of competition [40]: CLS and DS.

While past research has pointed to the importance of using DS, Porter explained that a competitive strategy in the Asia-Europe region is still useful and is an important factor in business success [41, 42]. This is considered an essential condition for securing market position and good performance [43]. Firms seek greater market share and competitive position as this strategic purpose greatly affects the level of a firm's competitive advantage [44]. Therefore, this research is limited to taking the implementation of CLS and DS as the core competitive strategies of this research.

Firms must consider a cost leadership strategy in their strategic planning because it may effectively stimulate the firm's operational activities, which in turn contributes to the firm's performance [45], and it can also help firms improve and increase their financial success [46].

A competitive strategy can significantly increase performance. A cost leadership strategy is also critical to the development of a firm [36, 46]. Fundamentally, both strategies can significantly improve small business performance. In addition, the differentiation strategy enables SMEs to achieve sustainable performance, but the implementation of differentiation may mean unstable performance and various systemic risks [47], and a business firm's competitive strategy can help them in volatile markets to find a sustainable competitive position. It is very important to high market performance [48]. Pursuit of DS small- and medium-sized enterprise for a higher value than only the pursuit of CLS [46, 49]. Lechner and Gudmundsson [42] investigated 335 small Icelandic firms. They found that each dimension of EO has a different degree of impact on strategy. [50] found that new business models promote competitive advantage and SME performance.

Therefore, we hypothesize that competitive strategy will mediate the association between entrepreneurial orientation and firm performance.

Hypothesis 2. Competitive strategy mediates the relationship between entrepreneurial orientation and firm performance of entrepreneurial Chinese SMEs in Thailand.

2.3. Mediating Role of Knowledge Creation Process. Hina et al. [22] found that EO and organizational learning can affect business performance. Li et al. [51] investigated those 165 entrepreneurs. They found that the indirect effects of EO through the knowledge creation process were added to the total effects model. Therefore, EO has positively correlated with FP: the knowledge creation process acts as a mediator. Nonaka et al. [52] found that using knowledge to create entities is the knowledge concept of firms. They believe that an important guarantee for the sustainable advantage of an enterprise is the process of knowledge creation and utilization. Nonaka [14] shows that organizational knowledge creation processes have dynamic patterns, which, as a core theme of the organization, result from the combined effects of explicit and tacit knowledge sustainability.

Dröge et al. [17] found that knowledge is a complete mediator of the impact of context and performance. Li Sa et al. [7] showed that customer orientation is directly correlated with FP. Therefore, consumers' behavioral orientations and plans influence the KCP, which is determined by the dynamic capabilities of employers. They also illustrated that the customer orientation shown by small hotels is partly mediated through the knowledge creation process and has an effect on FP. Therefore, we hypothesize that the knowledge creation process will mediate the association between entrepreneurial orientation and firm performance.

Hypothesis 3. Knowledge creation process mediates the relationship between entrepreneurial orientation and firm performance of entrepreneurial Chinese SMEs in Thailand.

2.4. Multiple Mediating Role of Competitive Strategy and Knowledge Creation Process. Competitive strategy is usually used to discuss how to help a business run better [53] so as to improve FP [30]. Entrepreneurial orientation and strategy are concepts of the strategic unit [54]. The former is seen as a strategic creative process, and the latter is seen as an elaboration of the content. From this point of view, the entrepreneurial-oriented competition strategy adds new content and guides [42]. On the other hand, EO and competitive strategy interact to form complementary advantages, and research on them is necessary [55]. Use of competitive strategy views mediating effects in entrepreneurial orientation and performance relationship. Thus, a problem appears around how entrepreneurial orientation and competitive strategy are correlated. Resource constraints may strengthen small firms' strategical lock-in. Thus, both structures, entrepreneurial orientation and competitive strategy are important to FP and have long-lasting effects. In fact, competitive strategy needs innovation capacity to create new products that help customers generate new value and fully justify the premium [30, 39].

Entrepreneurial orientation (EO) has always been highly concerning to researchers [56]. Entrepreneurial orientation (EO) is the firm's strategy formulation process, structure, and behavior [57], because the EO can partly explain why a few firms have stay ahead in the competition of the management process [25]. EO has garnered quite a lot of concepts and experience, it represents the business accumulated a lot of knowledge in the study of one of the rare fields. The correct use of the entrepreneurial orientation dimension can help firms gain competitive advantages and improve performance [56, 58, 59]. Entrepreneurship orientation is also the main direction of literature on strategic management and entrepreneurship [60].

To accommodate the dynamic competitive circumstances, firms continuously transform EO into practical strategical activities by concentrating on taking advantage of the knowledge creation process to achieve firm goals and achieve excellent performance. Knowledge creation process, transferring it to operate the activity to increase efficiency and produce commercial value [52].

Hypothesis 4. Competitive strategy and knowledge creation processes play a continuous intermediary role in the associations between entrepreneurial orientation and firm performance of entrepreneurial Chinese SMEs in Thailand as shown in Figure 1.

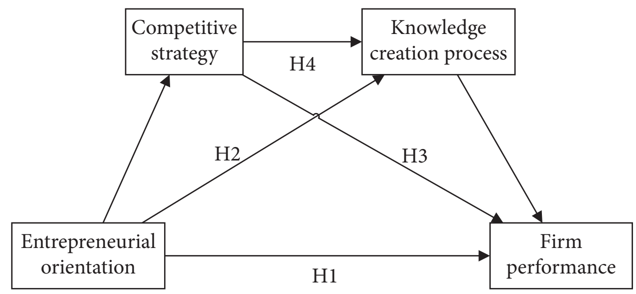


Figure 1: Hypothesized model.

3. Materials and Methods

3.1. Participants. This study selected Chinese SMEs located in Bangkok, Thailand. We distributed 352 questionnaires to survey senior managers of Chinese SMEs in Thailand, and 333 valid questionnaires were retained, giving a response rate of 94.60%.

Among these participants, 65.5% were males and 34.5% were females, which means more male managers than females. Samples from Rojana Industrial Park, Amata City Rayong Industrial Estate, Pinthong Industrial Park, and WHA areas were 36%, 52.3%, 2.7%, and 9%, respectively. The respondents were all types of senior managers of Chinese SMEs; high school and below accounted for 4.5%; bachelor's and master's accounted for 85.9% and Ph.D. accounted for 9.6%; owner and chief executive officer (CEO) accounted for 39.6%, followed by top manager accounting for 52.3% and finance manager for 8.1%. Among the subjects, 77.8% were under 45 years old, and 22.2% were over 45 years old.

3.2. Procedure. This study selects Chinese SMEs in Thailand as the survey object. These Chinese SMEs come to Thailand to participate in entrepreneurial activities. Participants received support from the Thai government and industrial parks in terms of funding, site use, tariff relief, and other needs. Respondents can know the items in this research, to meet the standards of empirical analysis. In April 2022, we contacted senior managers of Chinese SMEs starting businesses in industrial parks in Thailand and sent an invitation to participate in this study to their senior managers in the form of a questionnaire. Corporate executives are told that the data is used for study only, that we will maintain secrecy of their privacy, and they can choose to join voluntarily or withdraw or refuse to join at any point in time. We asked them to return the questionnaire only when they wanted to join in the study.

3.3. Measures. This study refers to important relevant literature that has been published internationally and selects a very mature scale to ensure the accuracy of empirical research. Based on deep communication with senior managers of Chinese SMEs, this research accurately translates the scales of competitive strategy and knowledge creation processes based on the specific scenarios of Chinese SMEs starting businesses in Thailand and makes the expressions of the items on the scale more in line with the understanding of Chinese SMEs in Thailand. Finally, a research scale is formed. We use a five-point Likert scale, in which "1" on the entrepreneurial orientation, competitive strategy, and knowledge creation process scale means "strongly disagree," "5" means "strongly agree," and "1" on the firm performance scale means "very dissatisfied" and "5" means "very satisfied." Small- and medium-sized enterprises that start businesses should evaluate the corresponding projects according to their actual situation.

3.3.1. Entrepreneurial Orientation. In order to measure the participation of Chinese SMEs in entrepreneurship in Thailand, this research adopts the five-point Likert-style scale originally developed by Lumpkin and Dess [25, 26]. The entrepreneurial orientation scale consists of 15 items which measure five in the following dimensions: 3 items for

Table 1: Confirmatory factor analysis results of variable discriminant validity.

Model	χ^2	df	χ^2/df	IFI	CFI	TLI	AIC	BIC	RMSEA	SRMR
Four-factor model	2458.358	896	2.744	0.912	0.912	0.907	2646.358	3004.323	0.072	0.0368
Three-factor model	3095.812	899	3.444	0.877	0.876	0.870	3277.812	3624.353	0.086	0.0492
Two-factor model	4092.820	901	4.543	0.821	0.820	0.811	4270.820	4609.744	0.103	0.0572
Single-factor model	4605.311	902	5.106	0.792	0.791	0.781	4781.311	5116.428	0.111	0.0599

Four-factor model: entrepreneurial orientation, competitive strategy, knowledge creation process, and firm performance. Three-factor model: entrepreneurial orientation + competitive strategy, knowledge creation process, and firm performance. Two-factor model: entrepreneurial orientation + competitive strategy + knowledge creation process and firm performance. Single-factor model: entrepreneurial orientation + competitive strategy + knowledge creation process + firm performance.

innovativeness, 4 items for risk-taking, 3 items for proactiveness, 2 items for competitive aggressiveness, and 3 items for autonomy. Respondents will be asked to rate the items on a 5-point Likert scale, ranging from 1 = "strongly disagree," through 3 = "neutral," to 5 = "strongly agree." This measurement scale has been previously used in other studies, for example by Li et al. [51]. The Cronbach's alpha coefficient of this scale is 0.966, and it shows good reliability.

3.3.2. Competitive Strategy. In order to measure the participation of Chinese SME's use of competitive strategies in Thailand, this study adopts the five-point Likert-style scale originally developed. The instrument consists of 7 items which measure two in the following dimensions: 4 items for differentiation strategy and 3 items for cost leadership strategy. Respondents will be asked to rate the items on a 5-point Likert scale, ranging from 1 = "strongly disagree," through 3 = "neutral," to 5 = "strongly agree." This measurement scale has been previously used in other studies such as by Anwar et al. [61]. The Cronbach's alpha coefficient of this scale is 0.976, and it shows good reliability.

3.3.3. Knowledge Creation Process. In order to measure the participation of Chinese SME's using the knowledge creation process in Thailand, this study adopts the five-point Likert-style scale originally developed by Nonaka [14]. The instrument consists of 16 items which measure five in the following dimensions: 4 items for socialization, 5 items for externalization, 4 items for combination, and 3 items for internalization. Respondents will be asked to rate the items on a 5-point Likert scale, ranging from 1 = "strongly disagree," through 3 = "neutral," to 5 = "strongly agree." This measurement scale has been previously used in other studies such as by Li et al. [51]. The Cronbach's alpha coefficient of this scale is 0.978, which shows good reliability.

3.3.4. Firm Performance. In order to measure the Chinese SME's firm performance in Thailand, this study used the firm performance scale. This scale includes three dimensions of efficiency, growth, and profits on which the firm performance variable is measured. Compared with competitors, the respondents were graded with the five-point scale on firm performance, ranging from 1 = "strongly dissatisfied," through 3 = "neutral," and 5 = "strongly satisfied" [51]. The Cronbach's alpha coefficient of this scale is 0.874, and it shows good reliability.

3.3.5. Control Variables. A control variable is a variable that is not considered a major factor in experiments, observational studies, and data analysis, so it reflects an external or a third factor whose influence needs to be controlled or excluded [62]. The key objective of considering control variables is to ensure that estimates of the influence of dependent variables are independent of the influence of external variables [62]. The control variables that are considered based on previous studies include job title, firm size, education level, and the firm's location [63]. When a firm is small and young, its survival and sustainability are lower [42].

4. Results Analysis

This study takes SPSS 23.0, AMOS 24.0, and Mplus v8.3 as tools. The analysis mainly has three aspects: (1) the test measurement model mainly has model fit, reliability, and validity; (2) descriptive statistics; and (3) multimediation tests and using the bootstrap method is used to repeatedly sample 5000 times to test the mediation effect and select the (Model 6) developed by [64].

4.1. Common Method Deviation Test. This study uses Harman's univariate method to test for common method bias in the study, loading all measured items with a common latent factor, and the results show a poor model fit ($\chi^2 = 5176.658$, df = 1034, χ^2/df = 5.006, GFI = 0.469, AGFI = 0.421, PGFI = 0.430, CFI = 0.783, NFI = 0.743, RFI = 0.731, IFI = 0.783, TLI = 0.773, CFI = 0.783 RMSEA = 0.110, SRMR = 0.0578) [65]. It shows that there are no methodological factors that can explain most of the variation in this study [66]. Therefore, we did not find a common method bias effect.

4.2. Confirmatory Factor Analysis. A CFA was performed on all variables using AMOS 24.0 software. The results in Table 1 show that the four-factor model used is the most appropriate compared to the one-, two-, and three-factor models. The combination effect is good, the fit index of the four-factor model meets the standard, and the model fit is good.

4.3. Descriptive Statistics. Statistical analysis was performed by SPSS 23.0. As shown in Table 2, the mean and standard deviation of every variable are within acceptable limits. There is a significant correlation between entrepreneurial

Table 2: Means, standard deviations, and correlations for variables ($N = 333$).

Variable	Mean	SD	1	2	3	4	5	6	7	8
(1) Education level	2.420	0.727	1							
(2) Your job title in this firm	2.590	0.769	−0.331**	1						
(3) The location of your firm	1.850	0.852	0.052	−0.027	1					
(4)The size of your firm	2.500	0.735	0.222**	−0.201**	−0.026	1				
(5) Entrepreneurial orientation	3.266	0.518	0.020	0.042	−0.057	0.038	1			
(6) Competitive strategy	3.296	0.515	0.045	−0.005	−0.096	0.060	0.810**	1		
(7) Knowledge creation process	3.326	0.510	−0.014	0.035	−0.043	0.080	0.786**	0.877**	1	
(8) Firm performance	3.296	0.552	0.026	0.013	−0.080	0.053	0.784**	0.850**	0.862**	1

$^{*}p < 0.05$; $^{**}p < 0.01$.

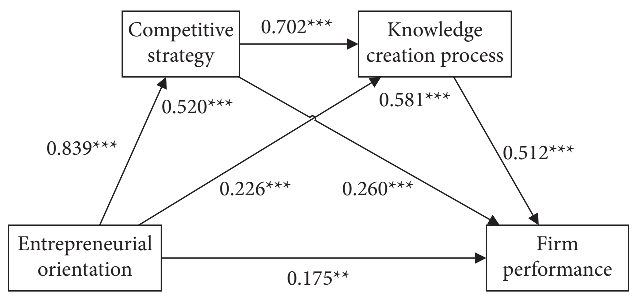


Figure 2: The standardized path coefficients in model testing. $^{**}p < 0.01$, $^{***}p < 0.001$.

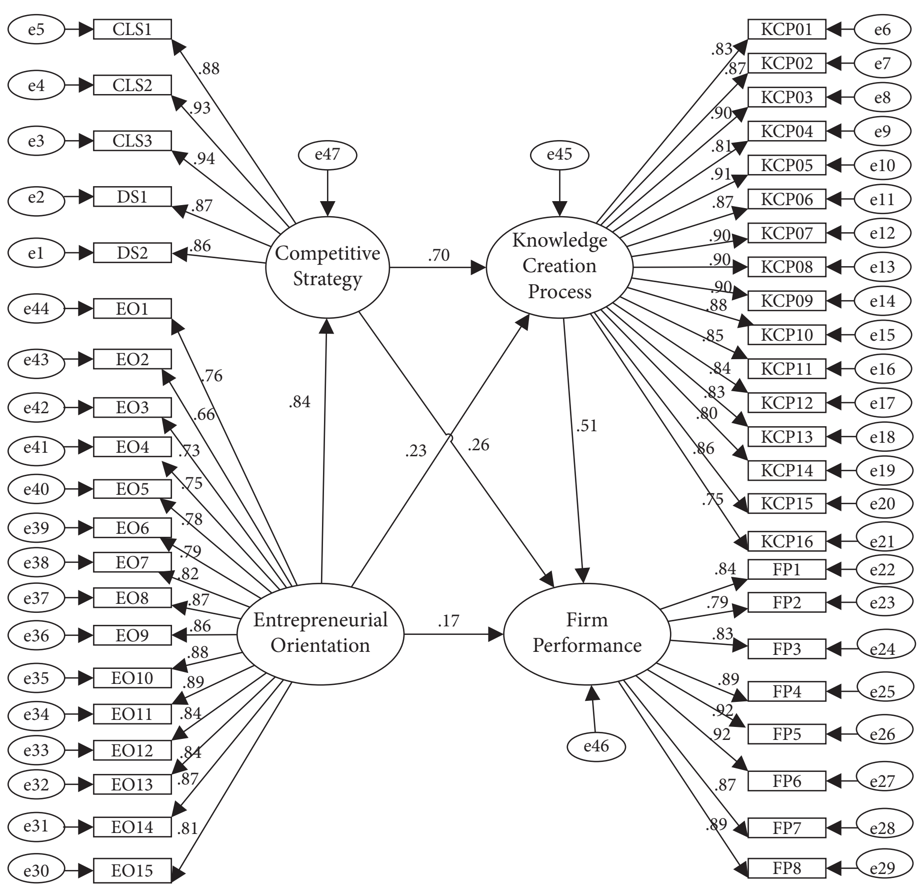


Figure 3: Structural equation model.

TABLE 3: Mediating effect and 95% confidence interval estimated by bootstrap method.

	Path	Estimation	CI at 95% level	
	Total indirect effect	0.673	0.546	0.805
Total effect	EO ⟶ FP	0.858	0.741	0.987
	EO ⟶ CS ⟶ FP	0.231	0.073	0.475
Indirect effect	EO ⟶ KCP ⟶ FP	0.123	0.034	0.295
	EO ⟶ CS ⟶ KCP ⟶ FP	0.319	0.219	0.473

Note: CI, confidence interval; EO, entrepreneurial orientation; CS, competitive strategy; KCP, knowledge creation process; FP, firm performance.

orientation, competitive strategy, the knowledge creation process, and firm performance. These results show the relationship between variables and lay the foundation for the next data analysis

4.4. Structural Equation Model Analyses. First, we tested the main effect, structural equation model 1 consists of EO as an independent variable and FP as a dependent variable. Model 1 fit index up to standard ($\chi2/\mathrm{df} = 2.288$, CFI = 0.962, TLI = 0.958, IFI = 0.962, NFI = 0.935, RMSEA = 0.062, and SRMR = 0.0292). The results show that entrepreneurial orientation positively influences firm performance ($\beta = 0.842$, $p < 0.001$), and H1 is supported.

Then, we construct model 2 and model 3 using competition strategy and the knowledge creation process as single mediators, and the results show that the models fit well. (Model 2: $\chi^2/df = 2.789$, CFI = 0.939, TLI = 0.934, IFI = 0.940, NFI = 0.909, RMSEA = 0.073 and SRMR = 0.037; Model 3: $\chi^2/df = 2.703$, CFI = 0.921, TLI = 0.916, IFI = 0.921, RMSEA = 0.072, and SRMR = 0.032).

Through Mplus v8.3, the bootstrap method is used to repeatedly sample 5000 times to test the mediation effect as shown in Figure 2.

In addition, through AMOS 24.0, the bootstrap method is used to repeatedly sample 5000 times to test the mediation effect as shown in Figure 3.

The mediating effect of the competitive strategy was 0.520, with 95% confidence interval (0.356, 0.706), excluding 0, based on the assumption that H2 was verified. The mediating effect of the knowledge creation process is 0.581, with a 95% confidence interval (0.463, 0.719), excluding 0, based on the assumption that H3 is verified.

Finally, we tested for the existence of chained multiple mediation effects and found a significant correlation between competitive strategies and knowledge creation processes. This study hypothesizes that these two variables mediate the effect of EO on FP. Therefore, we used Mplus to test for multiple mediation effects, and the results are shown in Table 3. Entrepreneurial orientation ⟶ competitive strategy ⟶ firm performance mediating effect is 0.231, the 95% confidence interval is (0.073, 0.475), excluding 0, and the mediating effect is significant. EO ⟶ KCP ⟶ FP, the mediating effect is 0.123, the 95% confidence interval is [0.034, 0.295], excluding 0, and the mediating effect is significant. EO ⟶ CS ⟶ KCP ⟶ FP chain multi-mediating effect is 0.319, 95% confidence interval (0.219, 0.473), excluding 0, showing that competitive strategy and knowledge creation process are between entrepreneurial orientation and firm performance, and H4 is verified.

5. Conclusions

Based on the resource-based view theory (RBVT), this study explores the impact mechanism of EO on FP. Structural equation modeling is used to examine the single and successive mediating roles of competitive strategy and knowledge creation processes, and to verify the competitiveness of Chinese SMEs in Thailand and the knowledge creation process capability. Chain-based multiple intermediary roles provide a new way to consider the impact of EO on firm performance. The empirical study shows the following results: (1) *The Main Effect Test*. The results show that there is a positive correlation between the perception of entrepreneurial orientation and the perception of corporate performance. (2) *Mediating Effect Test*. The test results show that competitive strategy and the knowledge creation process play a mediating role in entrepreneurial orientation and firm performance, respectively. Competitive strategy enhances the ability of knowledge creation and plays a continuous mediating role in the effect of EO on FP.

5.1. Theoretical Implications. This research has made contributions in the following aspects: first, this research, through the integration of entrepreneurial orientation, knowledge creation process, competitive strategy, and firm performance four key conceptions, contributes to the current hot entrepreneurship areas. Second, this study treats entrepreneurial orientation as a structure that affects other variables that affect performance. Competitive strategy and the knowledge creation process act as mediators between EO and firm performance, and EO can indirectly affect firm performance through competitive strategy and the knowledge creation process. This research is based on the RBVT, which includes all unique resources [1] and supplements porter's [30] view of the competitive strategic structure. Shift the focus from the competitive environment of the firm to the resources developed within that environment [67]. Thirdly, this study is very important for the development and empirical research of entrepreneurial orientation, competitive strategy, knowledge processing, theoretical description of firm performance, and strengthening of norms. Finally, the knowledge creation process expands the internal capabilities of the firm.

This research expands our understanding of the possible conditions that entrepreneurial orientation may be related to firm performance. Competitive strategy and knowledge creation process have theoretical implications for researchers to understand the relationship between EO and firm performance. At the same time, this study allows researchers to deeply know the relationship between EO and FP from the perspective of knowledge creation theory [14].

5.2. Managerial Implications. We use Thailand as the research background. Thailand has an open and inclusive business environment. Its economic development level ranks high among ASEAN countries, and it can enjoy zero-tariff treatment in ASEAN countries and expand contingency in entrepreneurial orientation research. There have been many explorations of the influence of EO on firm performance in the entrepreneurial field. However, there are few relevant theoretical or empirical studies on whether competitive strategy and knowledge creation processes are related to EO and firm performance. Therefore, building on previous literature, this research more accurately reveals the conditions under which EO affects firm performance, but the degree to which firm performance improves or does not improve may depend on the kind of competitive strategy employed.

This study explores the RBVT as the subject to give prominence to the significance of a firm's entrepreneurial activities, knowledge transfer, and competitive strategy on entrepreneurial SME firm performance. The results can benefit entrepreneurial SME firms in Thailand and owners, CEOs, top managers, and policymakers liable for firm performance in entrepreneurial SMEs in China. This study reveals how entrepreneurial SME EO, knowledge creation processes, and strategies can help firms gain a sustainable competitive edge, a core goal of RBVT. Entrepreneurial SMEs adopting entrepreneurial orientation, cost leadership strategy, and differentiation strategy may assist owners, CEOs, top managers, and policymakers to fund in even more appropriate and steady projects, rather than unsure and high-stake projects. The study's conceptual framework and hypotheses shown in a simple and overall manner, effectively promoted the existing literature about entrepreneurial orientation, competitive strategy, knowledge creation processing, entrepreneurial SME firm performance, and RBVT. This study treats EO, competitive strategy, and knowledge creation processing as internal capabilities of the firm that can improve performance. Existing literature on EO, competitive strategy, and knowledge processing is enhanced by gathering empirical information on entrepreneurial SMEs.

Managers can find more effective ways to compete so that the firm can continue to have a competitive edge. This research studies the auto parts, machinery, manufacturing, new energy, and other industries in Rojana, Amata City Rayong Industrial, Estate, Pingtong, and WHA. This study refines the understanding of the impact of EO on firm performance by researchers, executives, and entrepreneurs alike, it can not only help Chinese and other entrepreneurial SMEs entering Thailand in the future to improve their firm performance but also help entrepreneurial SME gain a competitive advantage when competing with their competitors.

The study found that some Chinese entrepreneurial SMEs in Thailand, which means that they have overseas strategic or low-cost strategies. It may be because it has a cheap labor force or easy access to raw materials in Thailand. Therefore, it is more meaningful and valuable to study Chinese entrepreneurial SMEs in Thailand than Chinese local firms.

5.3. Limitations and Future Study Directions. In terms of study samples, due to the limitation of study subjects, this study only judged the entrepreneurial orientation evaluation of Chinese SME entrepreneurs in Thailand unilaterally and failed to collect relevant data on Chinese SMEs who started businesses in other countries. Second, considering that perceptions in different periods have different effects on human behavior and choices, future research may be conducted from different countries and regions and consider dynamic tracking; research techniques can use qualitative or meta-analysis research methods. In addition, the impact of entrepreneurial orientation on firm performance is multidimensional, further enriching research models and research conclusions so as to establish a good theoretical framework for SMEs to set up a business environment overseas. This study only analyzes the mediating factors between entrepreneurial orientation and firm performance. Therefore, future research should add other mediating or moderating factors to the research framework.

Data Availability

All data that support the findings of this study are available upon reasonable request to the corresponding author.

Conflicts of Interest

The authors declare that there are no conflicts of interest in the publication of this article.

References

[1] B. Wernerfelt, "A resource-based view of the firm," *Strategic Management Journal*, vol. 5, no. 2, pp. 171–180, 1984.
[2] R. P. Rumelt, "How much does industry matter?" *Strategic Management Journal*, vol. 12, no. 3, pp. 167–185, 1991.
[3] J. Barney, "Firm resources and sustained competitive advantage," *Journal of Management*, vol. 17, no. 1, pp. 99–120, 1991.
[4] K. R. Conner, "A historical comparison of resource-based theory and five schools of thought within industrial organization economics: do we have a new theory of the firm?" *Journal of Management*, vol. 17, no. 1, pp. 121–154, 1991.
[5] R. Amit and P. J. Schoemaker, "Strategic assets and organizational rent," *Strategic Management Journal*, vol. 14, no. 1, pp. 33–46, 1993.
[6] S. Thornhill and R. Amit, "Learning about failure: bankruptcy, firm age, and the resource-based view," *Organization Science*, vol. 14, no. 5, pp. 497–509, 2003.
[7] M. L. Li Sa, S. Choon-Yin, Y. K. Chai, and J. H. Aik Joo, "Knowledge creation process, customer orientation and firm performance: evidence from small hotels in Malaysia," *Asia Pacific Management Review*, vol. 25, no. 2, pp. 65–74, 2020.

[8] D. J. Teece, G. Pisano, and A. Shuen, "Dynamic capabilities and strategic management," *Strategic Management Journal*, vol. 18, no. 7, pp. 509–533, 1997.

[9] S. G. Winter, "Understanding dynamic capabilities," *Strategic Management Journal*, vol. 24, no. 10, pp. 991–995, 2003.

[10] S. D. Hunt and R. M. Morgan, "The resource-advantage theory of competition: dynamics, path dependencies, and evolutionary dimensions," *Journal of Marketing*, vol. 60, no. 4, pp. 107–114, 1996.

[11] S. D. Hunt and R. M. Morgan, "Resource-advantage theory: a snake swallowing its tail or a general theory of competition?" *Journal of Marketing*, vol. 61, no. 4, pp. 74–82, 1997.

[12] R. M. Grant, "Toward a knowledge-based theory of the firm," *Strategic Management Journal*, vol. 17, no. SUPPL, pp. 109–122, 1996.

[13] I. Nonaka and N. Konno, "The concept of "Ba": building a foundation for knowledge creation," *California Management Review*, vol. 40, no. 3, pp. 40–54, 1998.

[14] I. Nonaka, "A dynamic theory of organizational knowledge creation," *Organization Science*, vol. 5, no. 1, pp. 14–37, 1994.

[15] I. Nonaka, "A dynamic theory of organizational knowledge creation," *The strategic management of intellectual capital and organizational knowledge*, vol. 437, p. 462, 2002.

[16] D. J. Teece, "Capturing value from knowledge assets: the new economy, markets for know-how, and intangible assets," *California Management Review*, vol. 40, no. 3, pp. 55–79, 1998.

[17] C. Dröge, C. Claycomb, and R. Germain, "Does knowledge mediate the effect of context on performance? Some initial evidence," *Decision Sciences*, vol. 34, no. 3, pp. 541–568, 2003.

[18] I. Nonaka and R. Toyama, "The theory of the knowledge-creating firm: subjectivity, objectivity and synthesis," *Industrial and Corporate Change*, vol. 14, no. 3, pp. 419–436, 2005.

[19] S. H. Haider, M. Asad, and M. Fatima, "Entrepreneurial orientation and business performance of manufacturing sector small and medium scale enterprises of Punjab Pakistan," *European Business and Management*, vol. 3, no. 2, pp. 21–28, 2017.

[20] V. K. Gupta and A. Gupta, "Relationship between entrepreneurial orientation and firm performance in large organizations over time," *Journal of International Entrepreneurship*, vol. 13, no. 1, pp. 7–27, 2015.

[21] S. S. Liu, X. Luo, and Y. Z. Shi, "Integrating customer orientation, corporate entrepreneurship, and learning orientation in organizations-in-transition: an empirical study," *International Journal of Research in Marketing*, vol. 19, no. 4, pp. 367–382, 2002.

[22] S. M. Hina, G. Hassan, M. Parveen, and S. Arooj, "Impact of entrepreneurial orientation on firm performance through organizational learning: the moderating role of environmental turbulence," *Performance Improvement Quarterly*, vol. 34, no. 1, pp. 77–104, 2021.

[23] P. K. Chaney, T. M. Devinney, and R. S. Winer, "The impact of new product introductions on the market value of firms," *Journal of Business*, vol. 64, no. 4, pp. 573–610, 1991.

[24] L. Cui, D. Fan, F. Guo, and Y. Fan, "Explicating the relationship of entrepreneurial orientation and firm performance: underlying mechanisms in the context of an emerging market," *Industrial Marketing Management*, vol. 71, pp. 27–40, 2018.

[25] G. T. Lumpkin and G. G. Dess, "Clarifying the entrepreneurial orientation construct and linking it to performance," *Academy of Management Review*, vol. 21, no. 1, pp. 135–172, 1996.

[26] G. T. Lumpkin and G. G. Dess, "Linking two dimensions of entrepreneurial orientation to firm performance," *Journal of Business Venturing*, vol. 16, no. 5, pp. 429–451, 2001.

[27] J. Wiklund and D. A. Shepherd, "Where to from here? EO-as-experimentation, failure, and distribution of outcomes," *Entrepreneurship: Theory and Practice*, vol. 35, no. 5, pp. 925–946, 2011.

[28] J. Rezaei and R. Ortt, "Entrepreneurial orientation and firm performance: the mediating role of functional performances," *Management Research Review*, vol. 41, no. 7, pp. 878–900, 2018.

[29] D. Kantur, *Strategic Entrepreneurship: Mediating the Entrepreneurial Orientation-Performance Link*, Management Decision, New York, 2016.

[30] M. E. Porter, *Competitive Strategy: Techniques for Analyzing Industries and Competitors*, Free Press, New York, 1980.

[31] R. R. G. Javalgi, A. C. Gross, W. B. Joseph, and E. Granot, "Assessing competitive advantage of emerging markets in knowledge intensive business services," *Journal of Business & Industrial Marketing*, vol. 1, 2011.

[32] S. O. Mokaya, "Corporate entrepreneurship and organizational performance theoretical perspectives, approaches and outcomes," *International Journal of Arts and Commerce*, vol. 1, no. 4, pp. 133–143, 2012.

[33] R. E. Miles, C. C. Snow, A. D. Meyer, and H. J. Coleman, "Organizational strategy, structure, and process," *Academy of Management Review*, vol. 3, no. 3, pp. 546–562, 1978.

[34] J. J. Chrisman, C. W. Hofer, and W. B. Boulton, "Toward a system for classifying business strategies," *Academy of Management Review*, vol. 13, no. 3, pp. 413–428, 1988.

[35] R. Chaganti, R. G. Cook, and W. J. Smeltz, "Effects of styles, strategies, and systems on the growth of small businesses," *Journal of Developmental Entrepreneurship*, vol. 7, no. 2, p. 175, 2002.

[36] R. S. Allen, M. M. Helms, M. B. Takeda, C. S. White, and C. White, "A comparison of competitive strategies in Japan and the United States," *SAM Advanced Management Journal*, vol. 71, no. 1, pp. 24–35, 2006.

[37] L. Teeratansirikool, S. Siengthai, Y. Badir, and C. Charoenngam, "Competitive strategies and firm performance: the mediating role of performance measurement," *International Journal of Productivity and Performance Management*, 2013.

[38] R. M. Grant, *Contemporary Strategy Analysis*, John Wiley & Sons, China, 2021.

[39] M. E. Porter, *Competitive Advantage: Creating and Sustaining Superior Performance*, Free Press N. Y, New York NY, 1985.

[40] R. Torgovicky, A. Goldberg, S. Shvarts et al., "Application of Porter's generic strategies in ambulatory health care," *Health Care Management Review*, vol. 30, no. 1, pp. 17–23, 2005.

[41] M. Acquaah, "Business strategy and competitive advantage in family businesses in Ghana: the role of social networking relationships," *Journal of Developmental Entrepreneurship*, vol. 16, no. 1, pp. 103–126, 2011.

[42] C. Lechner and S. V. Gudmundsson, "Entrepreneurial orientation, firm strategy and small firm performance," *International Small Business Journal: Researching Entrepreneurship*, vol. 32, no. 1, pp. 36–60, 2014.

[43] H. E. Salavou, "Competitive strategies and their shift to the future," *European Business Review*, vol. 27, no. 1, pp. 80–99, 2015.

[44] S. A. Schulz and R. L. Flanigan, "Developing competitive advantage using the triple bottom line: a conceptual

framework," *Journal of Business & Industrial Marketing*, vol. 31, no. 4, pp. 449–458, 2016.

[45] J. R. F. Lorenzo, M. T. M. Rubio, and S. A. Garcés, "The competitive advantage in business, capabilities and strategy. What general performance factors are found in the Spanish wine industry?" *Wine Economics and Policy*, vol. 7, no. 2, pp. 94–108, 2018.

[46] L. O. Oyewobi, A. O. Windapo, and R. O. James, "An empirical analysis of construction organisations' competitive strategies and performance," *Built Environment Project and Asset Management*, vol. 5, no. 4, pp. 417–431, 2015.

[47] R. D. Banker, R. Mashruwala, and A. Tripathy, *Does a Differentiation Strategy lead to More Sustainable Financial Performance than a Cost Leadership Strategy? Management Decision*, NY City, 2014.

[48] M. A. Hitt, C. M. Carnes, and K. Xu, "A current view of resource based theory in operations management: a response to Bromiley and Rau," *Journal of Operations Management*, vol. 41, no. 1, pp. 107–109, 2016.

[49] E. Teti, F. Perrini, and L. Tirapelle, "Competitive strategies and value creation: a twofold perspective analysis," *The Journal of Management Development*, vol. 33, no. 10, pp. 949–976, 2014.

[50] M. Anwar, "Business model innovation and SMEs performance-Does competitive advantage mediate?" *International Journal of Innovation Management*, vol. 22, no. 7, 2018.

[51] Y. H. Li, J. W. Huang, and M. T. Tsai, "Entrepreneurial orientation and firm performance: the role of knowledge creation process," *Industrial Marketing Management*, vol. 38, no. 4, pp. 440–449, 2009.

[52] I. Nonaka, R. Toyama, and A. Nagata, "A firm as a knowledge-creating entity: a new perspective on the theory of the firm," *Industrial and Corporate Change*, vol. 9, no. 1, pp. 1–20, 2000.

[53] D. E. Schendel and C. W. Hofer, *A New View of Business Policy and Planning*, Strategic Management, Boston, 1979.

[54] J. G. Covin and G. T. Lumpkin, "Entrepreneurial orientation theory and research: reflections on a needed construct," *Entrepreneurship: Theory and Practice*, vol. 35, no. 5, pp. 855–872, 2011.

[55] W. J. Wales, V. K. Gupta, and F. T. Mousa, "Empirical research on entrepreneurial orientation: an assessment and suggestions for future research," *International Small Business Journal: Researching Entrepreneurship*, vol. 31, no. 4, pp. 357–383, 2013.

[56] M. Anwar and S. Z. A. Shah, "Entrepreneurial orientation and generic competitive strategies for emerging SMEs: financial and nonfinancial performance perspective," *Journal of Public Affairs*, vol. 21, no. 1, 2021.

[57] W. B. Gartner and S. A. Shane, "Measuring entrepreneurship over time," *Journal of Business Venturing*, vol. 10, no. 4, pp. 283–301, 1995.

[58] T. Quince and H. Whittaker, *Entrepreneurial Orientation and Entrepreneurs' Intentions and Objectives*, University of Cambridge, ESRC Centre for Business Research, Brown Boston, 2003.

[59] F. Akbar, A. T. B. Bon, and F. Wadood, "Open innovation mediates the relationship between entrepreneurial orientation and firm performance: a preliminary survey," *Proceedings of the International Conference on Industrial Engineering and Operations Management*, , NY China, June 2020.

[60] R. A. Aziz, R. Mahmood, A. Tajudin, and M. H. Abdullah, "The relationship between entrepreneurial orientation and business performance of SMEs in Malaysia," *International Journal of Management Excellence*, vol. 2, no. 3, pp. 221–226, 2014.

[61] M. Anwar, S. Z. Khan, and N. U. Khan, "Intellectual capital, entrepreneurial strategy and new ventures performance: mediating role of competitive advantage," *Business & Economic Review*, vol. 10, no. 1, pp. 63–94, 2018.

[62] N. J. Salkind, "Encyclopedia of research design," *sage*, vol. 1, 2010.

[63] W. Stam and T. Elfring, "Entrepreneurial orientation and new venture performance: the moderating role of intra- and extraindustry social capital," *Academy of Management Journal*, vol. 51, no. 1, pp. 97–111, 2008.

[64] A. F. Hayes, "Introduction to mediation, moderation, and conditional process analysis," *A Regression-Based Approach*, Guilford Publications, China, Second Edition, 2017.

[65] L. T. Hu and P. M. Bentler, "Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives," *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 6, no. 1, pp. 1–55, 1999.

[66] L. L. Zhou Hao, "Statistical remedies for common method biases," *Advances in Psychological Science*, vol. 12, no. 06, pp. 942–942~950, 2004.

[67] D. Miller and J. Shamsie, "The resource-based view of the firm in two environments: the hollywood film studios from 1936 to 1965," *Academy of Management Journal*, vol. 39, no. 3, pp. 519–543, 1996.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 7363221, 11 pages
https://doi.org/10.1155/2022/7363221

Research Article

Development and Design Case Function Comparison of Panoramic Roaming System of Virtual Museum Based on Pano2VR

Kai Cao

College of Art and Design, Wuhan Textile University, Wuhan 430079, China

Correspondence should be addressed to Kai Cao; 2001029@wtu.edu.cn

Received 24 May 2022; Revised 16 June 2022; Accepted 18 June 2022; Published 5 July 2022

Academic Editor: Jiafu Su

The development of modern information technology gives birth to virtual museums. The exhibition of physical museums is evolving into online virtual exhibition, allowing users to have an immersive experience of museums without leaving home. Through content analysis, this study analyzes the status quo of online virtual museums in China and discovers two major problems: the exhibits are mainly brief introductory graphs and texts, and the exhibition module is independent of the panoramic roaming module. To solve the problems, Pano2VR was used to develop and design a virtual museum panorama roaming system for Hubei Museum. After realizing the map navigation of the museum, photos of the museum were taken on the site and prepared into a panorama. Then, the panorama roaming of the virtual museum was achieved based on Pano2VR. Finally, the results were displayed on the virtual web page of the museum, and real interactive virtual exhibitions were conducted on the website platform. The proposed system enhances the research value of the cultural relics of the Hubei Museum and increases the interaction with users.

1. Introduction

With the development of the computer industry, web page construction and virtual roaming technologies are increasingly mature. Virtual roaming provides an important means for cultural dissemination in modern society, owing to its interactivity, imaginativeness, and immersiveness [1]. It is now a popular practice to fully apply virtual roaming to museum exhibitions. Many domestic and foreign scholars have studied virtual roaming technology, trying to offer users an immersive experience without leaving home [2–6]. At present, no camera can take panoramic photos of a huge scene. This gives rise to various panorama stitching algorithms [7, 8]. Chen et al. [9] took photos with an ordinary camera and created a 360° panoramic photo in the horizontal direction. But viewers cannot observe freely in the vertical direction. Wu and Hu [10] designed a novel algorithm to obtain a relatively complete panorama. Some results have been achieved by combining virtual roaming with panoramic photos on virtual campus [11, 12]. Although the virtual roaming technology involves disadvantages such as high cost, poor security, and many restrictions, the biggest advantage is that the objects to be roamed are real objects in the real world. Based on the data related to the real objects, a virtual roaming scene is created and realized remote virtual roaming. In the multidimensional information space, virtual roaming technology creates a virtual information environment based on real objects. After logging in to the system, users can watch various virtual scenes as if they were there and interact with the virtual scenes through corresponding operations. Receive the feedback information of the virtual scene, prompt the user to mobilize the subjective consciousness to think during the viewing process, and form a memory or impression of the relevant information of the virtual scene.

Focusing on the most simple and feasible practical methods, this study adopts Pano2VR, a panoramic roaming production tool Pano2VR, to design and realize a three-dimensional (3D) virtual museum roaming system through image modeling. The first part of this paper introduces the current situation of the construction of online virtual museums. Through comparative analysis, it is concluded that the virtual museum needs to be equipped with interactive modules for top-level design and transformation. The second part analyzes the realization basis of virtual museum panorama roaming through four aspects: virtual panorama roaming technology, panorama stitching software, characteristics and image stitching algorithm, Pano2VR, and HTML5 web page production. The third part takes Hubei Museum as an example, uses Pano2VR software to design the system, and finally realizes the panoramic roaming of the virtual museum. In the fourth part, the experimental analysis is carried out, and it is concluded that the Pano2VR technology development museum roaming system has good three-dimensional display and interactive functions.

2. Status Quo of Online Virtual Museums

Content analysis was performed to reveal the status quo of online virtual museums in China. The friendly links to other museums were collected from the websites of 12 national and provincial museums, and the overlapping ones were analyzed in Window 7. 360 Explorer and Internet Explorer (IE) 9 were tested separately at the normal speed of the campus network. Table 1 lists the status quo of online virtual museums. Out of the 12 museums, 10 have established virtual museums.

(1) From the perspective of the expression form of museum exhibits, the Palace Museum, Shanghai Museum, and Anhui Museum introduce the details of the exhibits purely in graphs and texts. The other museums display the exhibits with the help of virtual reality (VR) technology. In most cases, panoramic roaming is adopted, and introductory graphs and texts are embedded in interactive hotspots.

(2) From the perspective of smoothness, most virtual museums can be browsed very smoothly. Only two virtual museums are loaded slowly. But the browsing is smooth as well after the webpages are fully loaded.

(3) From the perspective of compatibility, the authors checked the cross-browser compatibility and the need to download and install the corresponding VR plug-ins. All virtual museums support mainstream browsers like IE, 360, Firefox, and Chrome-based browsers. In terms of plug-in support, the Palace Museum requires users to download and install a large high-definition image browser plug-in. Nanjing Museum requires users to download and install the NJMUSEUMPLAYERXP VR browser plug-in. Wuhan Museum requires users to download and install the CrystalWHMuseum 3D VR browser plug-in. But the panoramic roaming VR system cannot work after the plug-in is installed in Chrome-based browsers and the web page is refreshed. Shaanxi History Museum requires no browser plug-in, yet its web page cannot be opened in Chrome-based browsers: in the panoramic roaming system of the halls, users can merely click on an exhibit and view its information.

(4) From the perspective of interactivity, the virtual versions of Nanjing Museum, Henan Museum, Shaanxi History Museum, Shanxi Museum, and Hunan Museum are relatively well designed: all of them have designed map navigation and operational instructions. However, Hunan Museum is the only one that has set up a feedback module and embedded it in the panoramic roaming system. In addition, the message board module of most museums lacks the question and answer (Q&A) mechanism, a necessary function of virtual museums. Through content analysis, this study investigates the independent message board modules of the 12 museums and uses the ROST content mining system to capture the data from the message boards. The captured web data were subjected to semantic cluster analysis of high-frequency words. The analysis shows that most messages on the message board ask about the reservation system, venue opening hours, personal collection appraisal and consultation, venue service consultation, personal donation consultation, etc. However, the virtual exhibition is separated from the interaction module, and the message board is not related to that module. The virtual museums only provide photos of events and video playbacks of relevant theme lectures. The design of the virtual museums gives no consideration to interactivity or feedbacks. Online visitors could only browse the general information about exhibits. Overall, the virtual museums are badly in need of an additional interactive module, which could be arranged through top-level design and transformation.

(5) From a technical perspective, National Museum, Palace Museum, Henan Museum, etc. have adopted the latest Internet technology, combined with VR technology, LOD technology, etc., to maximize the sense of reality, immersion, and interaction of the exhibition hall. Shanghai Museum and Shanxi History Museum have developed virtual museum collection management systems based on Basic or dBase (Fox Base), but the technology cost for the image files in the exhibition hall is very high. Guangdong Museum, Shanxi Museum, etc. have adopted the combination of computer network technology and CMIS technology. Due to the huge amount of data resources required by this technology, the effect is very poor due to the influence of network transmission speed, and it is not suitable for online communication. Nanjing Museum and Wuhan Museum use Web3D technology to realize a 720-degree panoramic tour, which is an all-round tour mode without dead ends in the virtual exhibition hall. In addition to the several virtual museums mentioned above, Zhejiang Museum, Hunan Museum, Anhui Museum, etc. have adopted VR

Table 1: Status quo of online virtual museums.

Name of museum	Virtual museum? Yes/No	Form of online exhibits	Smoothness	Compatibility	Interactivity	Education module? Yes/No
National museum	Yes	Panoramic roaming embedded with graphs and texts	Smooth	Compatible	Navigation and operational instructions; no feedback module	No
Palace museum	Yes	Graphs and texts	Slight lag	Plug-in needed	Week navigation; no operational instructions or feedback module	No
Shanghai museum	No	Graphs and texts	—	—	—	—
Nanjing museum	Yes	VR 3D embedded with videos	Smooth and slow in loading	Plug-in needed	Clear and detailed navigation and operational instructions; no feedback module	Yes
Guangdong museum	Yes	Panoramic roaming embedded with graphs and texts	Smooth	Compatible	Navigation; no operational instructions; poor image operability; no feedback module	No
Wuhan museum	Yes	Panoramic roaming +3D	Smooth	Plug-in needed	Navigation and operational instructions; no feedback module	No
Henan museum	Yes	Panoramic roaming embedded with graphs and texts	Smooth and slow in transition	Compatible	Clear navigation, guide, and operational instructions; no feedback module	No
Zhejiang museum	Yes	Panoramic roaming	Relatively smooth	Compatible	Navigation; no operational instructions; no feedback module	No
Anhui museum	No	Pure graphs and texts	—	—	—	—
Shaanxi history museum	Yes	Links to graphs and texts and panoramic roaming	Smooth	Hall not displayed	Theme guide; graphic instructions; no feedback module	No
Shanxi museum	Yes	Panoramic roaming embedded with graphs and texts	Smooth and slow in loading	Compatible	Visual navigation map; no feedback module	No
Hunan museum	Yes	Panoramic roaming embedded with graphs and texts	Smooth	Compatible	Logical theme navigation and signs; feedback module	No

technology, Internet technology, and multimedia technology, etc., to realize the virtualization and spatialization of exhibition halls and shorten the exhibition hall and tour. The distance between the users and the excellent exhibition halls of our country will be displayed to users all over the world through the Internet.

3. Realization Basis for Panoramic Roaming of Virtual Museums

3.1. Virtual Panoramic Roaming Technology. The panoramic roaming system simulates and digitally reproduces the museum and related scenes. The virtual or real images produced by the computer are used to reproduce the real world. Then, online visitors can use the mouse and keyboard to freely browse in the virtual scenes, and engage in human-computer interactions (HCIs) [13–15]. Pano2VR is a popular panoramic roaming tool.

The VR is developed based on the virtual panorama roaming technology. Virtual panoramas are stitched together from a series of similar images on the same scene. In recent years, the VR has been widely applied, and combined with various techniques into a novel virtual environment technology.

Before the panoramic roaming of the virtual museum, the main preparatory works include shooting panoramas, understanding the mainstream panorama stitching software and their features, implementing Pano2VR, and web exhibition.

3.2. Mainstream Panorama Stitching Software and Their Features and Image Stitching Algorithm. Before realizing the virtual panoramic roaming of Hubei Museum, it is important to prepare the panorama, i.e., the panoramic photo that conforms to the normal effective viewing angle of human eyes. First, the photos with basically the same tone and small overlaps were taken consecutively in the horizonal or vertical direction. Then, these photos were stitched into a widescreen photo. The stitching of the panorama makes up for the defect of the existing imaging devices: the inability to take a 360° panoramic photo of a huge scene [12]. Table 2 summarizes the mainstream panorama stitching software and their features.

Panoramic image stitching occupies an important position in the field of image processing. Commonly used panorama image stitching algorithms include Harris, SURF, SIFT (scale-invariant feature transfer form) and other feature point detection algorithms. Among them, the stitching algorithm based on Harris corner points, due to the scale sensitivity of Harris corner points, the stability of the feature corner points is insufficient, resulting in false dislocation of the stitched images. Based on the SURF stitching algorithm, although the calculation speed of SURF is better than that of SIFT, when the image details with rich texture are stitched and stitched, the texture expression accuracy of SURF is slightly inferior to that of SIFT, which will result in a higher mismatch of matching points. Image stitching based on the SIFT algorithm, although feature points have the advantages of scale invariance and feature description accuracy, but there are problems such as long search time for feature points and insufficient optimization and filtering of incorrectly matched feature points. In view of the problems of the above methods, an improved fast panoramic image stitching algorithm based on SFIT is proposed, which improves the stitching efficiency, and purifies and improves some mismatched points, which has stronger robustness.

3.3. Pano2VR. Pano2VR is a panorama image conversion and application software based on the flash animation technology. The development of the software is reported in Figure 1. The main function of Pano2VR is to convert the panorama images of multiple formats into the corresponding image files for display. The files can be outputted as flash animation files or brand-new HTML5/CSS3 files. Besides, Pano2VR supports multiresolution progressive zooming of panoramas and is compatible with the animation display function of different platforms. Thus, it truly meets the needs of cross-platform joint development and facilitates the design and development of different virtual animation systems. Pano2VR is an important software for making flash panorama. As the mainstream format of panorama, flash can be easily played directly and easily on webpages. Pano2VR is a very professional panorama image converter, which mainly solves the problem of converting panorama images into quicktime, macromedia flash8, flash9, html5, and other formats, and whether it is a gigapixel panorama or a virtual image with hundreds of nodes Roaming can be achieved perfectly. The software also has the ability to take plane, cylinder, sphere, cube, intersection, T, strip, and quicktime VR as input formats and additions and provides an easy way to add and edit existing panoramas and Google Street View View, in addition, the software also supports multiresolution progressive zoom panorama and re-optimization of the engine to generate a more fluent flash, even compatible with JPEG, TIFF, PNG, BigTIFF, Photoshop PSD/PSB, and other image formats and Panorama output in HTML5/CSS3 format.

3.4. HTML5 Web Page Construction. HTML is a markup language for depicting web files. The language has won the favor of many users, thanks to its various excellent properties, such as ease of operation, scalability, and platform independence. HTML5 is a new HTML standard, which improves multiple aspects of web terminals, such as semantics, interactivity, system function calls, and multimedia. HTML5 solves many of the problems in previous standards, namely, HTML4. Moreover, many new features are added to HTML5, e.g., audio, video, and image embedding functions, client-side data storage, interactive files, etc. Specific rules are provided for handling all HTML elements and restoring them from errors [20]. In addition, HTML5 further enhances interactivity and effectively reduces development cost. Specific implementation steps: (1) Web page display content design, that is, the theme to be displayed on the page and the functions to be implemented. First of all, it is

Table 2: Mainstream panorama stitching software and their features.

Name	Description	Features
PTGui	PTGui [16] is a multifunctional panorama image stitching tool.	The software is powerful, supporting multiple types of stitching, without obvious stitching traces.
Canon utilities PhotoStitch 3.1	Canon utilities PhotoStitch 3.1 [17] is a panorama production software released by Canon with its digital cameras.	The operation interface is beautiful and simple, and the merging speed is fast. But the ability of hue control is weak.
Autopano giga	Autopano giga [18] is used for 360° panorama and can export panorama to flash.	The stitching accuracy is high, and the color and fusion are well controlled. But the thread and CPU usage is high during APG rendering.
Photovista	PhotoVista [19] is based on a powerful engine that automatically calculates seams and then produces seamless 360° panoramas, using various methods.	PhotoVista is a tool that can prompt the number of photos that match the lens settings, and the operation process is simple.
Adobe photoshop CS6	Adobe photoshop CS6 [16] is the 13th generation of adobe photoshop.	There are no obvious stitching traces. The tone is roughly the same. It is a relatively common mainstream software with a huge volume.

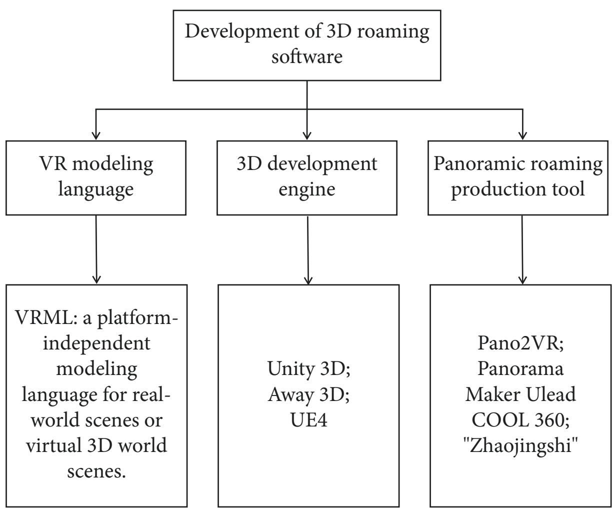


Figure 1: Development history of roaming software.

necessary to clarify the needs of customers, and second, to consider the type and nature of virtual museums on the basis of satisfying customer needs. (2) Create a web page. First, you need to plan the site, you can draw sketches on paper to complete the layout, arrange text, pictures, and even animations on the specified position of the page, determine the general outline of the page, and then compile the HTML5 script code on the computer to design the blueprint become a reality, of course, you can also use visual development tools, such as Dreamweaver, Fireworks, and so on. (3) The overall effect design of the web page. Pay attention to the collocation effect of the elements of each page. The elements on the page must construct the entire page through the combination and collocation of points, lines, and surfaces. The use of CSS technology can achieve good results. (4) Page optimization: generally speaking, text and pictures are the largest elements in a web page, so it is very necessary to use CSS style sheets to specify the style of text. As for the optimization of the picture, on the premise of ensuring the browsing quality, the size of the picture should be reduced to the minimum, which can double the download speed of the web page, or the large picture can be cut into several small pieces by using image processing software, and carried out separately. Optimized, its file type can be.GIF or.JPEG.

4. Pano2VR-Based Realization of Panoramic Roaming of Virtual Museum

Taking Hubei Museum, for example, this study designs the virtual museum by Pano2VR: (1) realizing the map navigation of the museum; (2) taking photos of the museum on the site, preparing them into a panorama, and realizing the panorama roaming of the virtual museum based on Pano2VR; and (3) displaying the results on the virtual web page of the museum and conducting real interactive virtual exhibitions on the website platform.

The virtual museum mainly includes (1) digital knowledge warehouse of cultural relics. The digital knowledge warehouse of cultural relics is the foundation of other application platforms. It contains digital cultural relics organically composed of information resources such as high-definition images, audio, video, text data, three-dimensional data, remote maps, and related research results required by each application platform. (2) Cultural relic display and dissemination platform. All data of the cultural relics display and dissemination platform comes from the digital knowledge warehouse of cultural relics, and the combination of various data adapts to the needs of different dissemination methods. It can be divided into in-library communication platform and out-of-library communication platform. The communication platform in the museum includes a large-screen multimedia cultural relics display module and a touch-screen stereoscopic display module. The communication platform outside the museum includes a virtual exhibition module of cultural relics, a mobile digital device access module, etc. (3) Cultural relics academic research platform. The platform includes the acquisition and remote interaction of digital cultural relics information resources. Experts who have obtained control can perform operations such as flipping, zooming in and out of research objects, and extracting relevant information materials.

For the virtual Hubei Museum, factors such as smooth operation mechanism, clear operation method, perfect functional system, and practical visiting experience are particularly important. To narrow the feeling gap between virtual tours and physical tours, it is necessary to restore the functions of the physical exhibition halls as much as possible in the design of the system, and on this basis, design a virtual museum system with good interactivity. At the beginning of the design, the following design elements should be considered: (1) data resource combination design. For the virtual resources of museum exhibits and information, museums should integrate and design them and store them as background resources, and at the same time, classify and manage exhibits and information by period, type, size, etc. This saves time and cost for post-data management of the virtual museum. (2) Operation design of the exhibition hall. For the tour of virtual museum projects, various service items for VR display should be determined at the beginning of design, which should include basic operations such as exhibits display, information introduction and explanation, and human-computer interaction. After the design of these functions is completed, the basic framework of the virtual museum system will be built and then expanded according to the requirements of each part of the function. (3) Visiting route and perspective design. When designing the visiting route of virtual roaming, it should be based on the structure of the physical museum and at the same time consider the habits and characteristics of users when visiting and design a variety of tour methods for users to choose. This part of the design can greatly improve the basic application functions of the museum.

4.1. Map Navigation of Virtual Museum. To realize the map navigation function of the Virtual Hubei Museum, code was written in MyEclipse using the application programming interface (API) of Baidu Map, creating a map embedded in the web page. Three kinds of maps can be generated. The first is the static GIF image, which can be directly and dynamically acquired from Baidu Map. It is more convenient than loading a complete map. The second is the dynamic map, which can be dragged and zoomed from a point. Figure 2 shows the dynamic map centering on the museum. The third is the navigation function of Baidu Map. Online visitors can identify their current positions and check the routes to the museum in three different navigation modes, completing the path planning for visiting the museum.

4.2. Panoramic Roaming of Virutal Musuem. The panomarmic roaming of the Virtual Hubei Museum can be realized in the following steps:

Step 1. Shoot panoramic photos of the museum, stitch them into a panorama of the virtual museum, and combine other technologies to prepare the panorama into a more suitable panoramic roaming map.

Step 2. Based on Pano2VR, create interactive hotspots and make the panoramic roaming map more realistic.

Step 3. Add the function of virtual roaming interaction, which enables visitors to roam in the virtual scenes of the museum. Ensure the user experience by allowing online visitors to walk forward, backward, left, and right by pressing the up, down, left, and right keys, to change the perspective through the headset, or to choose running, walking, and other modes. This function can be realized as follows:

```
void Update( )
{float movev = 0; float moveh = 0;
if(Input.GetKey(KeyCode.UpArrow))
{Movev- = m_speed * Time.deltaTime; }
If(Input.GetKey(KeyCode.DownArrow))
{Movev+ = m_speed * Time.deltaTime; }
if(Input.GetKey(KeyCode.LeftArrow))
{ Moveh+ = m_speed * Time.deltaTime; }
if(Input.GetKey(KeyCode.RightArrow))
{Moveh- = m_speed * Time.deltaTime; }
this.transform. Translate(newVector3(moveh,0,
movev)); }
```

4.2.1. Preparing 3D Panorama. The panoramic photos of Hubei Museum were captured by combining the following devices: digital single lens reflex (DSLR) camera, tripod head, tripod, ultra-wide-angle lens, etc. The photos were all shot on the site of the museum. Out of the various panorama stitching software, Autodesk 3 ds Max was adopted for 3D modeling of the museum, and the panorama output function was called to directly generate the panorama. First, the camera was configured in the 3D scene, and the camera view was chosen. Then, panorama was selected in the rendering menu. The software would automatically stitch a panorama, merge the layers, and save the photo. Note that the maximum resolution was selected for rendering output, aiming to ensure the definition and resolution of panoramic roaming prepared by Pano2VR.

4.2.2. Creating Pano2VR Project Fil. The panorama roaming map was developed by Pano2VR. Once the 3D panorama was prepared, Pano2VR was adopted to output the panoramic roaming in the format of flash. The specific steps are as follows:

Step 1. Click on the desired file ⟶ click to open the file ⟶ select the flash panoramic map(s) generated in the file or drag the image directly from the input of the window to the input box or the display area.

Step 2. Click on the desired attribute in the main interface ⟶ enter the input (select type).

There are more than 10 types of inputs, such as automatic, rectangular spherical projection, cubic mage, column chart, panel image, QuickTime VR, cross type, and T type. If automatic is chosen, the software will automatically determine the input type according to the input image.

Step 3. Click the image, zoom the pixels, drag left and right to set the initial angle of the default playback, and set the horizontal view.

Figure 2: 3D dynamic map centering on Hubei Museum (Figure 2 is from Google Maps).

If it is not a 360° column panoramic view, please set the corresponding horizontal view angle. Click OK after modifying the parameter.

Step 4. Return to the main interface, find the addition option, and select the output format.

Normally, the output format is either flash or HTML. The output format can be further modified.

Step 5. Select the name and format of the output file and the path for outputting the file.

Pano2VR is embedded with the HTML output function, which can automatically output and display the flash file on the web page.

Once all settings are complete, direct click to confirm. Then, the swf file would be generated, which can be called by JavaScript.

4.2.3. Realizing Interactive Hotspots. Interactive hotspots need to be added to the panoramic roaming process, to realize the interaction between visitors and the roaming scene, giving them an immersive feeling. By clicking the hotspots, visitors can jump from one panoramic map to another panoramic map. The specific steps are as follows:

Step 1. Open Pano2VR, import panoramic map(s), configure the parameters and process each map simply, and select the interactive hotspot.

Step 2. Open the interactive hotspot to see three tabs. The first two tabs are responsible for adding interactive hotspots, and the third one is responsible for managing the added hotspots. We mainly focused on the first tab. Select the hotspot type, double click the hotspot of the panorama, and add a hotspot URL (the panoramic map that will be opened by clicking the hotspot).

Step 3. The hotspot list mainly manages the added hotspots and displays the added hotspots. A hotspot can be added by clicking the green plus in the lower area, but the position of the additional hotspot cannot be specified. Hence, the addition function here is basically not used.

Step 4. Return to the third tab to see two main types of interactive hotspots: hot points and hot areas. On the panel, the key is the switch between the two kinds of hotspots. The polygonal interactive hot areas can be added to a map on the wall, when the picture exhibition or panorama is of low resolution. Then, a single clear image can be called from the outside and controlled by skin ID. Also, it is possible to make the following setting: when the cursor moves to a hot area, a piece of music will be played. Overall, the panel is very flexible to use.

In order to make the panorama more vivid, it is necessary to add roaming hotspots, voices, and texts to the panorama. Hotspots are equivalent to the hyperlink between two panoramas. If the hot point is chosen, a red hot point will appear by double clicking the image. Then, please enter the caption, e.g., Hubei Museum. (In the generated panorama, the caption is displayed when the cursor stops on the red roaming button.) After that, enter the URL: enter the flash file to be linked, such as fangwu.swf. To set multiple hotspots, just double click the position where we want to add them. After completing the setting, click to confirm. By clicking audio modification, the dialog box of panoramic audio editor will appear. Double click a certain area to view the interface and add the desired audio file by choosing it in the file name. After completing the setting, click to confirm. The skin editor allows users to add texts and offers some function buttons.

For interactive voice design, the string of output recognition results should be evaluated by an independent voice recognition system. In general, the recognition rate is measured by the word error rate (WER). To ensure the consistency between the recognized word series and the standard word series, it is necessary to replace, delete, or insert some words. The total number of these words divided by the total number of words in the standard word series is the WER:

$$\mathrm{WER} = 100 \times \frac{S + D + 1}{N}\%, \tag{1}$$

where S, D, and I are the number of replaced, deleted, and inserted words, respectively; N is the total number of words in the standard word series.

Step 5. Set the pattern and then set the action of the interactive hotspot. Select the action modifier. We can open the panel and set the action source (event) and the action. Then, the panorama could complete the corresponding operation. Figure 3 shows the final output of the panoramic roaming results of Virtual Hubei Museum. Click the left and right buttons to rotate the screen; click the plus and minus signs to zoom in and out. The full-screen button and refresh button are responsible for full-screen playing and replay, respectively. By clicking on the interactive hot area, we can switch to another panoramic map.

Step 6. Make reasonable use of tracking and registration technology of computer vision.

In the current VR systems, there are two main categories of tracking and registration technologies [21]. The first category is based on computer vision. The other is based on hardware-like sensors and global positioning system (GPS). The computer vision-based techniques can be further divided as augmented reality (AR) with manual labels and AR without manual labels [22–24]. In most AR systems, the tracking and registration module is realized with manual labels [25]. This paper adopts the AR with manual labels to achieve tracking and registration. As shown in Figure 4, the adopted method can be implemented in four steps.

(1) Capture video images by video collection equipment. For real timeliness and accuracy, the AR-VR system calls the camera of mobile devices to capture real-world video images.

(2) Preprocess the collected images. The collected color video images are binarized into black and white binary images. For any pixel (x, y) in a binary image, if the gray value $S(x, y) \geq t$, the pixel belongs to the background; otherwise, the pixel belongs to the region of interest (ROI). After thresholding, the image can be defined as

$$S(x, y) = \begin{cases} a0, & f(x, y) < t \\ a1, & f(x, y) \geq t \end{cases}. \tag{2}$$

By setting $a0 = 0$ and $a1 = 1$, binary images can be obtained. Then, all ROIs are searched for in these images through image segmentation and edge detection.

(3) Using the pattern recognition and matching algorithm, each labeled image is compared with the labeled templates in the database to judge if the label is legitimate. If the matching succeeds, an ROI is found.

(4) After template matching, determine the unique ID of the current object, and the object is thus recognized.

4.2.4. Generating Web Roaming System. Through web page design and production, the above realization results were displayed on the virtual web page of the museum, achieving the web page display of the virtual part of the museum. Pano2VR supports roaming systems of various formats, e.g., flash and HTML5. Take the flash format as an example. By clicking the add button in the output, the output settings dialog box will automatically open. A large pixel value is recommended for the window size, aiming to ensure the display effect of the image during roaming. The roaming experience can be improved by setting various parameters, such as automatic rotation, motion inertia, mouse sensitivity, and roaming scaling. After returning to the main interface, click the creation button to generate the final swf and HTML roaming files. Figure 2 presents the flash version of the main interface of panoramic roaming.

After the completion of the virtual system, the HCI could be realized to give online visitors an immersive experience. They would feel as if they were really walking in the museum. They can freely rotate in any direction, visit the physical objects of the virtual scene, and appreciate the artistic atmosphere of the scene. Figure 5 shows the final realization results of web page display. Some exhibits are displayed virtually on the web page. The VR technology is fully utilized, and the unique image query technology was implemented to vividly present the exhibits to visitors (Figure 6). The system relies on MyEclipse and the online API of Baidu Map to fulfil map navigation and displays the panoramic roaming flash file based on Photoshop, Pano2VR, and HTML5 [26].

By visiting the Virtual Hubei Museum, people have a deeper understanding of every aspect of Hubei, including the collections and books of the museum, the special buildings in Hubei, the history of the Chu culture, as well as the folk customs and human landscapes of the Chu state.

5. Experiments and Results Analysis

Based on the Unity3D platform, our panoramic roaming system was developed by Vuforia SDK and passed the test on an Intel Core 2 Quad CPU, Q8200@3.22 GHz, 4 GB, with NVIDIA GT9800 graphics card.

Once the Pano2VR-based roaming system starts, the VR system test was carried out. The online visitors can tour the museum as they like and appreciate the landscape outside the museum or enter the museum to view the collections. The exchanges and interactions between visitors and exhibits are very frequent. After switching to the AR system, the system needs to call the camera of the display terminal. In

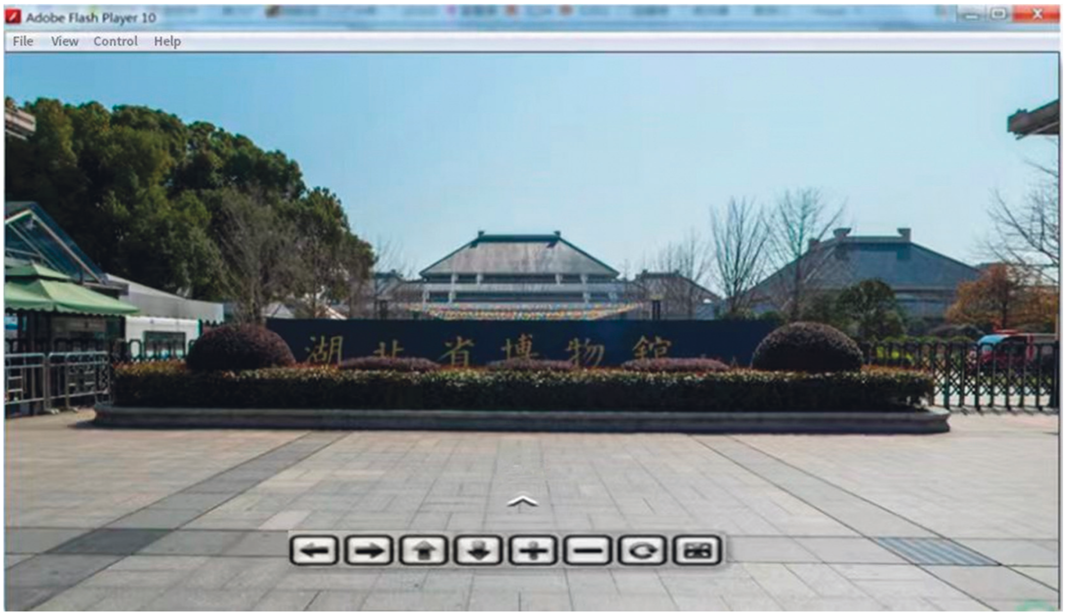


FIGURE 3: Panoramic roaming results of Virtual Hubei Musuem (Figure 3 is from Hubei Museum).

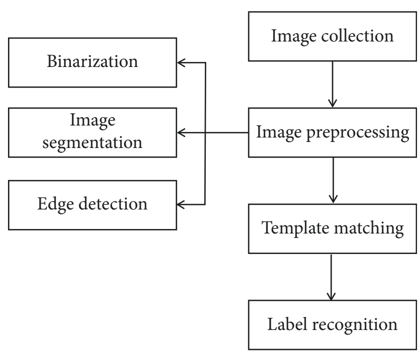


FIGURE 4: Flow of label recognition for tracking and registration.

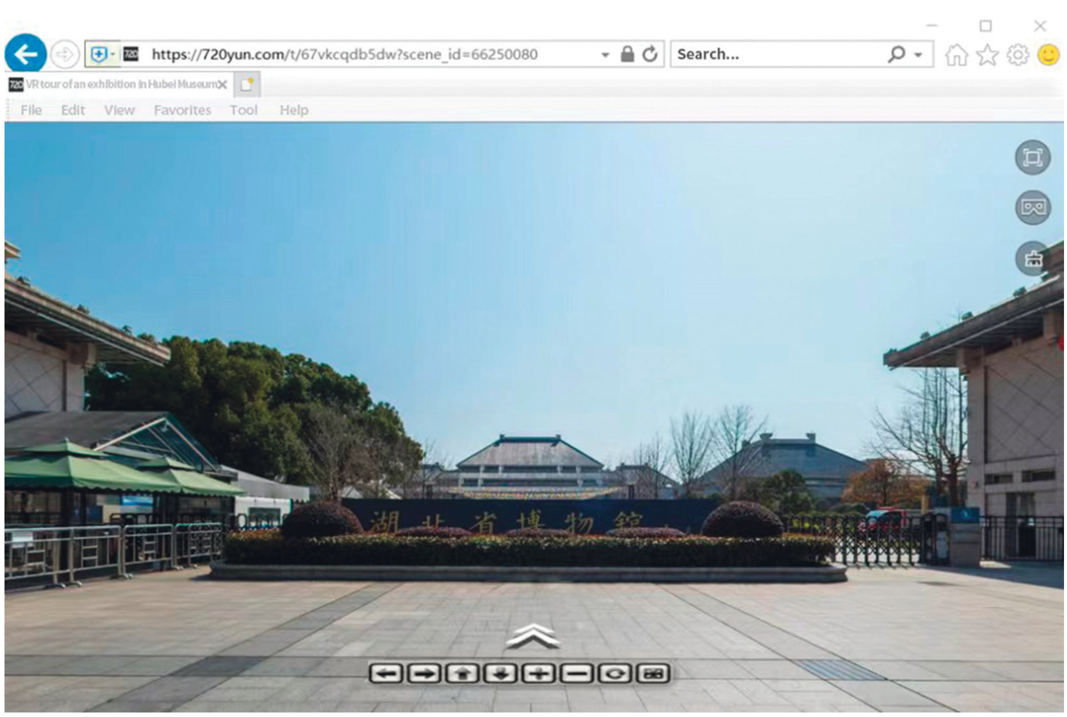


FIGURE 5: Panoramic roaming HTML5 page of Virtual Hubei Museum (Figure 5 is from Hubei Museum).

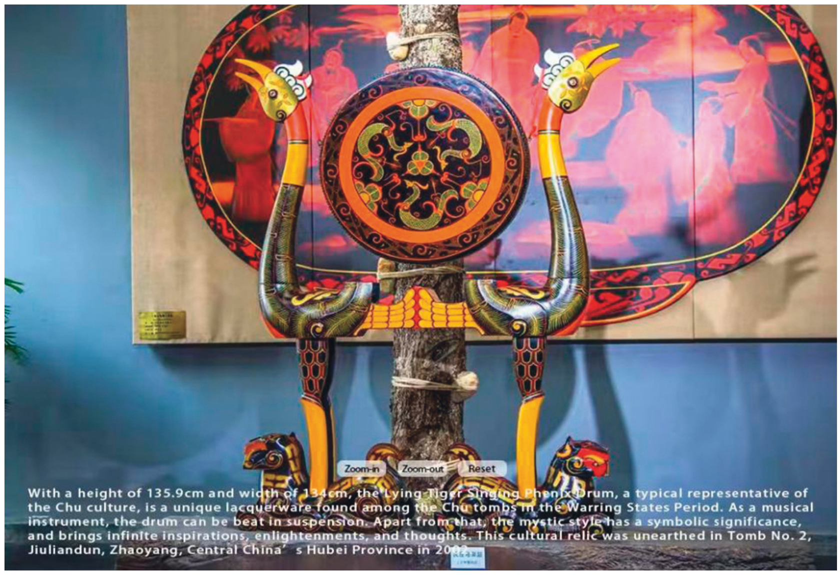


Figure 6: A virtual exhibit of Hubei Museum (Figure 6 is from Hubei Museum).

our experiments, the camera was used to shot the markers from different angles, and the generated 3D digital images were analyzed to see if the image quality is affected by the shot angle. The registration performance was verified with the extreme angles of 0° and 90°. Small images were selected for the performance test, because field AR module testing is impossible, given the special historical status and high degree of protection of Hubei Museum.

Experimental results show that, when the shot angle was 90°, the system could collect the feature information contained in the marker image; when the angle was 0°, the system could also collect that information. That is, the system realizes real-time tracking. In addition, multiple tests reveal that the proposed system can acquire accurate tracking and registration data, according to the marker information, a sign of high robustness.

Through the above experiments, it is concluded that the Pano2VR-based roaming system of Hubei Museum has a good 3D display and interactive functions. Online visitors can roam in the virtual museum. When they are interested in a certain exhibit, they can immediately switch to the AR or VR system to observe the details. The marker is not affected by complexity and can be tracked well through characteristic point detection.

6. Conclusions

This study provides website viewers with highly intuitive and realistic museum information through programming, panoramic roaming production, and virtual exhibition web production. By surfing the proposed virtual website, the public can learn about the specific location of Hubei Museum on map, interact with the panoramic roaming system using mouse and keyboard, and understand the basic structure, theme arrangement, and atmosphere by moving through the halls. The exhibits and collections of the museum are fully displayed on the webpages. In this way, our system enables Hubei Museum to hold virtual panoramic exhibitions and digitalize its official website. This paper has mastered the key technologies required to build a virtual museum system.

The main tasks are as follows: (1) implementation of multiview exhibition hall browsing: to build a three-dimensional and multiview virtual Hubei Museum, which can satisfy users' freedom of walking in the virtual museum and 360 views. View all kinds of exhibits and related materials in the museum in an all-round way. (2) Realization of exhibition hall roaming mode: a variety of exhibition hall roaming methods have been realized, and different exhibition hall roaming methods have been designed and developed according to the different needs of different audience groups. (3) Implementation of human-computer interaction in the exhibition hall: realize a series of interactive operations that users need to perform when roaming the exhibition hall. For example, if users want to watch a cultural relic exhibit in detail, they need to select, translate, and rotate the cultural relic; and pictures, video click interaction, small scene loading interaction, small game interaction, etc. (4) Implementation of collision detection algorithm: in order to solve the collision problem between dynamic objects and static objects in the virtual exhibition hall, a collision detection algorithm with good performance is researched and selected.

Through this study, the emerging technologies are fully integrated into the museum development, giving people a deeper understanding of every aspect of Hubei, including the collections and books of the museum, the special buildings in Hubei, the history of the Chu culture, as well as the folk customs and human landscapes of the Chu state. Suffice it to say that this study offers valuable experience for the spread and exploration of the Chu culture, as well as the protection of natural and cultural heritages in Hubei.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This study was supported by Ministry of Education Science and Technology Development Center University Industry-University-Research Innovation Fund Project "Application of Virtual Reality (VR) Visualization in Display Space" (2018A04005).

References

[1] X. Y. Zheng and Y. Peng, "Analysis of the application effect and application points of virtual roaming technology in industry training," *Journal of Inner Mongolia Agricultural University (Natural Science Edition)*, vol. 31, no. 4, pp. 250–254, 2010.

[2] B. Y. Yin, "Application of virtual reality technology in public facility design," *Packaging Engineering*, vol. 40, no. 16, pp. 271–274, 2019.

[3] C. Gherbi, "Internet of things and heterogeneous networks technologies: Concepts, challenges and perspectives," *Ingénierie des Systèmes d'Information*, vol. 26, no. 4, pp. 403–408, 2021.

[4] J. Luo and S. L. Deng, "Design of 3D virtual building space experience system," *Modern Electronic Technology*, vol. 41, no. 24, pp. 137–140, 2018.

[5] D. Marques and R. Costello, "Concerns and challenges developing mobile augmented reality experiences for museum exhibitions," *Curator: The Museum Journal*, vol. 61, no. 4, pp. 541–558, 2018.

[6] N. Capuano, A. Gaeta, G. Guarino, S. Miranda, and S. Tomasiello, "Enhancing augmented reality with cognitive and knowledge perspectives: A case study in museum exhibitions," *Behaviour & Information Technology*, vol. 35, no. 11, pp. 968–979, 2016.

[7] Z. Qu and X. L. Li, "Algorithm of eliminating image stitching line based on improved IGG model," *Computer Science*, vol. 44, no. 12, pp. 274–278, 2017.

[8] H. Y. Dai and F. P. Ran, "Fast panorama stitching algorithm for mobile devices," *Computer Applications*, vol. 34, no. 9, pp. 2673–2677, 2014.

[9] H. Chen, A.-Q. Long, and Y.-H. Peng, "Building panoramas from photographs taken with an uncalibrated hand-held camera," *Chinese Journal of Computers*, vol. 32, no. 2, pp. 328–335, 2009.

[10] L. P. Wu and Y. Hu, "Virtual reality technology of image smoothing in cylindrical panoramic image mosaic," *Science Technology and Engineering*, vol. 17, no. 31, pp. 271–276, 2017.

[11] Z. Ji, "Applied research on interaction design based on VR virtual roaming technology," *Modern Electronic Technology*, vol. 42, no. 15, pp. 86–89, 2019.

[12] A. Razia Sulthana, A. Arokiaraj Jovith, D. Saveetha, and A. K. Jaithunbi, "A game based virtual campus tour," *Journal of Physics: Conference Series*, vol. 1000, no. 1, p. 012131, 2018.

[13] G. Z. Huang, "Research on interactive campus roaming system based on Pano2VR," *Software Engineering*, vol. 11, no. 1, pp. 17-18, 2015.

[14] E. E. Reding, *Adobe Photoshop Cs6 Revealed*, Cengage Learning, Boston, 2012.

[15] A. Joukhadar, N. Ghneim, and G. Rebdawi, "Impact of using bidirectional encoder representations from transformers (BERT) models for Arabic dialogue acts identification," *Ingénierie des Systèmes d'Information*, vol. 26, no. 5, pp. 469–475, 2021.

[16] A. Cosentino, "Panoramic infrared reflectography. technical recommendations," *International Journal of Conservation Science*, vol. 5, no. 1, pp. 51–60, 2014.

[17] L. Rittié, D. L. Sachs, J. S. Orringer, J. J. Voorhees, and G. J. Fisher, "Eccrine sweat glands are major contributors to reepithelialization of human wounds," *American Journal Of Pathology*, vol. 182, no. 1, pp. 163–171, 2013.

[18] L. Barazzetti, M. Previtali, and F. Roncoroni, "3D modelling with the Samsung Gear 360," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII-2/W3, no. 2W3, pp. 85–90, 2017.

[19] D. Yi Xiao, "Experiencing the library in a panorama virtual reality environment," *Library Hi Tech*, vol. 18, no. 2, pp. 177–184, 2000.

[20] R. Y. Cheng, "Analysis of Responsive Web Design Based on HTML5 Technology," *Electronics World*, vol. 22, pp. 79-80, 2017.

[21] H. F. Kwok, A. Giorgi, and A. Raffone, "Improving interpretability: combined use of LVQ and ARTMAP in decision support," *Journal of Telecommunications and Information Technology*, vol. 6, pp. 129–132, 2005.

[22] X. Zhang and G. Chen, "Detection of dense small rigid targets based on convolutional neural network and synthetic images," *Traitement du Signal*, vol. 38, no. 1, pp. 61–71, 2021.

[23] S. C. Bulusu, M. Faezipour, V. Ng, M. Nourani, L. S. Tamil, and S. Banerjee, "Transient ST-segment episode detection for ECG beat classification," in *Proceedings of the IEEE/NIH Life Science Systems and Applications Workshop*, pp. 121–124, IEEE, Bethesda, MD, USA, April 2011.

[24] Y. Tadepalli, M. Kollati, S. Kuraparthi, and P. Kora, "EfficientNet-B0 based monocular dense-depth map estimation," *Traitement du Signal*, vol. 38, no. 5, pp. 1485–1493, 2021.

[25] J. S. Sahambi, S. N. Tandon, and R. K. P. Bhatt, "Wavelet based ST-segment analysis," *Medical, & Biological Engineering & Computing*, vol. 36, no. 5, pp. 568–572, 1998.

[26] Z. P. Ma, B. B. Guo, H. D. Li, and J. Zou, "Realization of Pano2VR based panoramic roaming of Ningxia virtual museum," *Modern Electronic Technology*, vol. 44, no. 8, pp. 149–153, 2021.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9871982, 1 page
https://doi.org/10.1155/2023/9871982

Retraction

Retracted: Fusion Analysis of Chinese Painting Color Teaching and Intelligent Image Color Processing Technology

Mobile Information Systems

Received 26 September 2023; Accepted 26 September 2023; Published 27 September 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] N. Zhao, "Fusion Analysis of Chinese Painting Color Teaching and Intelligent Image Color Processing Technology," *Mobile Information Systems*, vol. 2022, Article ID 3906277, 9 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 3906277, 9 pages
https://doi.org/10.1155/2022/3906277

Research Article

Fusion Analysis of Chinese Painting Color Teaching and Intelligent Image Color Processing Technology

Nan Zhao

School of Arts and Design, Zhuhai College of Science and Technology, Guangdong, Zhuhai 519041, China

Correspondence should be addressed to Nan Zhao; 12044@zcst.edu.cn

Received 11 May 2022; Revised 1 June 2022; Accepted 6 June 2022; Published 29 June 2022

Academic Editor: Jiafu Su

In order to solve the problem of traditional Chinese painting image color processing technology, meet the needs of digital museum construction, make up for the vacancy of color processing technology in traditional Chinese painting, and improve people's understanding of traditional Chinese painting. Through the analysis of the experiment, the experimental equipment adopts Epson v940 scanner, the color depth is 24 bit true color, and the scanning accuracy is 1200dpi. The main simulation methods of Chinese painting are systematically and deeply studied, from which the characteristics of Chinese painting are analyzed. Through a large number of formula analysis, it is found that the application of computer image processing technology in Chinese painting is very successful, which promotes the development of the digital museum and improves people's understanding and liking of Chinese painting.

1. Introduction

Since the birth of traditional computer graphics, its main purpose is to generate images that can imitate the effect of traditional cameras. With the continuous development of computer graphics hardware technology and graphics algorithms, people have been able to simulate various vivid graphics, such as the special effects of movies and games. In many cases, people have been unable to distinguish whether what they see is a real scene or generated by a computer [1]. The main forms of traditional Chinese painting: first, ink painting mainly appeared after the Tang and Song dynasties. From the Yuan Dynasty to the Qing Dynasty, the types and quantity of traditional Chinese painting reached its peak, with a large number and rich types. Yan BA was sincere to take off the animal husbandry lamp. Due to the long time and the limitation of the preservation conditions of paper and pigment, most of the paintings before the Yuan Dynasty had poor picture quality, the paper and silk are damaged, the color loss of pigment is serious, and the overall preservation is not good. The paintings have been preserved completely since the Yuan Dynasty [2]. Represented by Lang shining in the Qing Dynasty, traditional Chinese painting introduced western painting ideas, including the Western color system and realistic and exaggerated modeling techniques, which weakened the traditional expression of traditional Chinese painting. In order to make the selected image samples of traditional Chinese painting have certain representativeness, complete types (especially representative literati paintings), represent the mature development stage of painting theory, and consider the convenience of labeling, this paper only uses the traditional paintings of the yuan, Ming and Qing Dynasties for the selection of sample images, as shown in Figure 1:

2. Literature Review

Li and Zubrilin [3] and others believe that the rapid development of the Internet and multimedia technology makes digital image resources an effective information media. How to quickly retrieve massive digital images has become an urgent problem to be solved [3]. Guo and Qin [4] and others found that as structured data information, the retrieval method of an image is much more difficult than the text-based retrieval method [4]. Lin [5] investigated the characteristics of magnetic resonance diffusion tensor imaging (DTI) parameters and postoperative evaluation of spinal

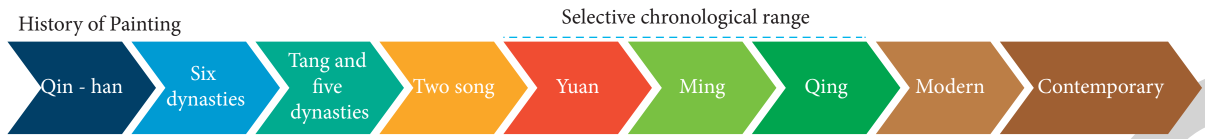


Figure 1: Selection of traditional Chinese painting samples and distribution of dynasties.

cord function in patients with high cervical myeloma. The logistic regression model was used to analyze the factors affecting the recovery of spinal cord function after surgery. The results show that abnormal DTI parameter values in patients with high cervical myeloma can better reflect the loss of spinal cord function, can effectively predict the recovery of the physical function of patients after surgery, and provide a reference for clinical diagnosis and treatment [5]. Perring et al. [6] found that with the rapid development of Chinese painting image digitization in recent years, there is an increasingly urgent demand for the establishment and management of a Chinese painting image digital library or digital museum. In particular, the processing technology of digital Chinese painting image has become the key to the urgent problem to be solved. The research on low-level feature extraction, data compression, automatic semantic annotation, retrieval, and automatic classification of digital Chinese painting image is more and more extensive [6]. Yang et al. [7] found that the research on automatic classification and annotation of digital images of Chinese painting involves the integration of computer vision, machine learning, image retrieval, cognitive psychology, and painting art. In general, because there is no direct relationship between the low-level visual features and high-level semantic concepts of Chinese painting images, the automatic semantic classification and annotation of Chinese painting images is a very challenging research topic [7]. Li et al. [8] have studied content-based image retrieval (CBIR) and image classification technology for more than ten years. Many scientific research institutions at home and abroad have done a lot of research work in general image retrieval and natural image scene classification [8]. Li [9] believe that the application of traditional Chinese painting digital images is still in its infancy [9], which is mainly due to the fact that different from the exact semantic information expressed by ordinary digital images, the characteristic of traditional Chinese painting digital images is that the semantic information reflected and represented by them is abstract. Traditional Chinese painting does not emphasize the light color change and focus perspective of objects in nature, nor does it pay attention to the realism of the appearance of objects. Traditional Chinese painting emphasizes "Writing Spirit in the form." It also has its own characteristics in composition, pen, ink, and color. Especially in the perspective method, it is not limited by the fixed field of view. It can move its foothold to paint according to the author's feelings and needs and show the seen and invisible scenery in the works. This scattered and multi-point perspective method is not available in Western painting.

The application of color based and texture based image retrieval in traditional Chinese painting is limited by the characteristics of "color based and shape based image retrieval" in the field of traditional Chinese painting. Exploratory research work has been attempted in the research field related to the automatic classification of Chinese painting images [10]. Moreover, more and more research institutions and researchers are engaged in the research in this field. Nonphotorealistic rendering is a technology that uses a computer to generate graphics with hand-painted style instead of photo like realism. It does not pursue the effect of "photo like realism" but mainly hopes to express the artistic characteristics of graphics and simulate artistic works of different styles. The research on nonphotorealistic technology first appeared in two papers by Wang [11] in the 1980s [11]. Wang [12] published two influential papers at the Siggraph conference in 1990, but the technology they demonstrated was treated in isolation at that time. Now researchers have made bold attempts on various artistic styles, giving people a refreshing feeling, which reflects that computers have great prospects in Chinese painting with intelligent image color processing technology [12]. Compared with photorealistic rendering, nonphotorealistic rendering is different in the implementation method and the scope applicable to performance. The comparison relationship between them is given. Through the comparison, we can find that nonphotorealistic rendering has more unique advantages in some specific environments. As shown in Table 1.

3. Method

The more easily acquired two-dimensional image is taken as the input, and the simulation method based on physical modeling is combined with the simulation method based on the feature. For the important elements of ink painting, paper, and pen, we establish a physical model according to its characteristics, simulate the stroke and ink diffusion effect of ink painting, and make full use of the relevant knowledge of digital image processing in the processing process. For example, in terms of edge stroke generation, by using the combination of threshold method and region generation, considering the color and spatial information of the image, the key features of the image are extracted, the redundant information is removed, and the rendering process is simplified. For the internal spatial processing of the image, the rendering method of large block coloring in the actual ink painting is simulated through filtering and color equalization [13, 14]. The general flow of the algorithm is shown in Figure 2.

K-means clustering algorithm is used to classify and extract the colors of female costumes in Dunhuang murals in various periods of the Tang Dynasty. In order to more

Table 1: Comparison between the two.

Category	Realistic rendering	Nonphotorealistic rendering
Implementation method	Simulation	Stylization
Xharacteristic	Objective	Subjective
Influence factor	Simulation of physical process	The law of artistic process and cognition
Accuracy	Accurate	Approximate
Deceptive	Unreal, the viewer will mistakenly believe that the image is true	In reality, the viewer will confirm that the image is a description of the scene
Level of detail	It is difficult to avoid redundant details, too much information and fixed level of detail	The level of detail can be changed according to the viewer's attention
Applicable to representation	Rigid surface	Natural and organic phenomena

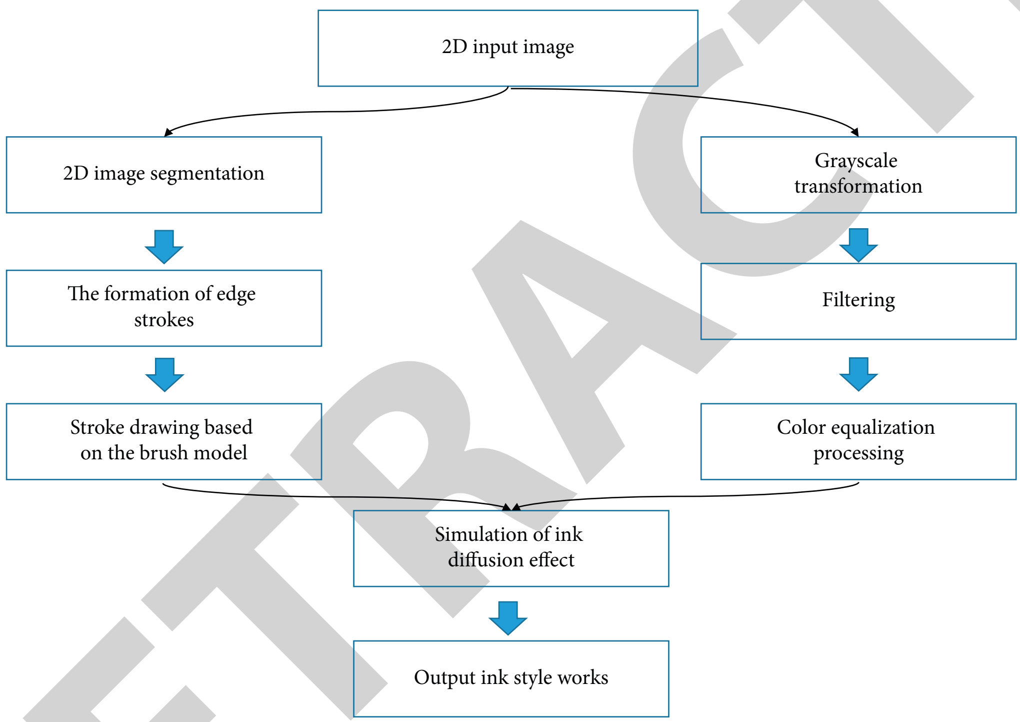


Figure 2: Drawing algorithm flow.

objectively describe the color of women's clothing in each period, through the quantitative analysis of the main color composition of women's clothing in different periods of the Tang Dynasty, the 10 main colors extracted from women's images in each period are statistically analyzed from the three dimensions of hue, saturation, and lightness. On the basis of quantitative analysis and combined with literature, this paper analyzes the causes behind the main color composition of female costumes in Dunhuang murals of the Tang Dynasty.

Through the comparison of different color spaces, the HSV color space is finally selected as the color space of this study. The three components of HSV color space are hue, saturation, and value. These three elements can affect the perception of color by human eyes. A. R. Smith created HSV color space in 1978, which can be represented by a closed inverted cone. The top view starts from red and rotates clockwise, which is composed of yellow, green, cyan, blue,

Table 2: Hue range corresponding to the color system.

Color system	Numerical range
Red system	330°–360°, 0°–30°
Yellow system	30°–90°
Green system	90°–150°
Cyan system	150°–210°
Blue system	210°–270°
Purple system	270°–330°

and magenta [15, 16]. Lateral variation from 0 to 100% of the edge constitutes saturation. The longitudinal variation from the low end *o* to the center of the top surface constitutes lightness 176. Hue H is the main feature of color. It is the most accurate to use hue to measure the difference of color. The hue starts from 0° of red and rotates clockwise. It is composed of 60° of yellow, 120° of green, 180° of cyan, 240° of blue, and 300° of magenta. Based on this, the hue ring is divided into 6 divisions. According to the above-given division basis, the hue ring is preliminarily divided into 6 equal divisions, that is, 6 color systems, as shown in Table 2. However, the visual effect of color is not only composed of hue but also affected by saturation and lightness. Therefore, the division of the color system is based on this table and slightly adjusted in combination with the visual effect [17, 18].

Saturation is the purity of color. If the purity of each hue is different, the value of saturation is also different. In the space represented by an inverted cone, the lateral variation from 0 to 100% of the edge constitutes saturation. Starting from the achromatic axis, extend the line segment horizontally to the edge of the circular surface, and divide the line segment into nine segments, i.e., 9 saturation divisions. The nearest three segments from the colorless axis are low saturation, the next three segments in the middle are medium saturation, and the farthest three segments are high saturation. Therefore, this paper divides the saturation into three ranges, as shown in Table 3.

Descriptive statistical analysis of hues in the female dress colors of Dunhuang murals of the Tang Dynasty extracted from four periods is shown in Table 4. As can be seen from the table, hue H goes clockwise for the middle Tang Dynasty, the prosperous Tang Dynasty, the early Tang Dynasty, and the late Tang Dynasty. The average hue of the early and late Tang Dynasties is around 142°, and the color on the corresponding hue ring is green with a cold tone [19, 20]. The mean hue of the prosperous Tang Dynasty is 136.37°. From the hue ring, it can be seen that the mean hue of the prosperous Tang Dynasty is less than that of the early Tang Dynasty and the late Tang Dynasty, so the corresponding color of the prosperous Tang Dynasty is warmer than that of the early Tang Dynasty and the late Tang Dynasty. The average hue of the middle Tang Dynasty is 109.30°, and the color on the corresponding hue ring is green with a warm tone, which belongs to the warmest tone of the overall tone in these four periods. Green belongs to the middle range in the warm and cold tones, that is, the middle tone. Therefore, it can be said that the overall tone of the early Tang Dynasty, the prosperous Tang Dynasty, and the late Tang Dynasty is cold, while the overall tone of the middle Tang Dynasty is warm [21, 22]. From the standard deviation of hue, the hue value fluctuated greatly in each period. Compared with other periods, the hue value fluctuated slightly in the middle Tang Dynasty, indicating that the Tang Dynasty was very bold in color use, and the colors were particularly rich, covering almost all the color systems in the whole hue ring, but the middle Tang Dynasty would focus more on the individual colors. In order to further understand the applicable law of each period on each color system, the color systems of each period are classified and analyzed, as shown in Table 4.

Saturation S refers to color purity. The higher the purity, the more distinct the performance. The lower the purity, the more elegant the performance. From the saturation S in Table 5 (the range value is 0 ~ 100%), in fact, the overall saturation difference is small, and the saturation belongs to low saturation, which is precisely the typical feature of Chinese traditional color. The saturation of early Tang and late Tang is lower than 33%, which belongs to low saturation; It is slightly higher in the prosperous Tang Dynasty and the middle Tang Dynasty, and the saturation is greater than 33%, which belongs to medium saturation. It can be seen that the colors of the early and late Tang Dynasties are more simple and elegant, and the lower saturation makes the overall style of the color show a gentle and harmonious effect. The prosperous Tang Dynasty and the middle Tang Dynasty, one is the most powerful period of the Tang Dynasty, and the other is the period of the Tubo invasion. The color saturation of the two periods is higher, even higher in the middle Tang Dynasty.

Table 3: Saturation range values.

Saturation	Numerical range
Low saturation	0–33%
Medium saturation	34%–66%
High saturation	67%–100%

The brightness and darkness of an object are perceived by the human eye from the visual experience of the light source. The brightness of the object is the brightness, which is affected by the intensity of the light source [23, 24]. If the light is stronger, it looks brighter; the weaker the light, the darker it looks. From the lightness distribution (Table 6), the overall lightness of the Tang Dynasty presents medium lightness, which gives people a more comfortable feeling in terms of visual effect. The average value of Mingdu in the middle Tang Dynasty is the highest, reaching 51.64%, which will be more beautiful visually, and the average value of Mingdu in the late Tang Dynasty is the lowest, which will be more vigorous compared with that in the late Tang Dynasty.

In order to obtain effective observation data of traditional Chinese painting, the standard observer needs to have certain conditions. The requirements and rules for selecting the standard observer are as follows:

(1) *Education Level*: high school and above, can use the mouse to operate the computer

(2) *Number of People*: 6

(3) *Age Group*: 20 ~ 40

(4) *Gender*: both male and female

We will combine the six standard observers of traditional Chinese painting images in pairs and divide them into three groups. The members of each group will observe the samples of traditional Chinese painting with the themes of flowers and birds, landscapes, and characters in the yuan, Ming, and Qing Dynasties, respectively. The personnel arrangement covers all dynasties and branches of traditional Chinese painting. This can not only eliminate the preferences caused by different people's different cognition but also balance the consistency of painting expression in different dynasties. Standardize and standardize the annotation from different angles, so as to improve the relative accuracy of the annotation of the main scene area. The combination of standard observers is shown in Figures 3 and 4.

As shown in Figures 3 and 4, each group of standard observers can mark the main scene of the painting images of

Table 4: Mean deviation of main color hue.

Color space	Period	Minimum	Maximum	Weighted mean	Standard deviation
Hue	Early Tang Dynasty	20	349	142.36	112.358
	Prosperous Tang Dynasty	17	347	136.37	117.908
	Mid-Tang Dynasty	15	345	116.29	97.258
	Late Tang Dynasty	15	345	142.98	113.158

Table 5: Mean deviation of main color saturation.

Color space	Period	Minimum	Maximum	Weighted mean	Standard deviation
Saturation S	Early Tang Dynasty	11	66	32.12	16.544
Saturation S	Prosperous Tang Dynasty	11	60	33.42	17.742
Saturation S	Mid-Tang Dynasty	14	64	34.84	15.61
Saturation S	Late Tang Dynasty	9	83	29.23	17.435

Table 6: Deviation of the mean value of main color brightness.

Color space	Period	Minimum	Maximum	Weighted mean	Standard deviation
Lightness V	Early Tang Dynasty	21	73	46.19	18.374
Lightness V	Prosperous Tang Dynasty	23	74	45.9	18.087
Lightness V	Mid-Tang Dynasty	17	73	51.64	19.424
Lightness V	Late Tang Dynasty	17	65	43.04	18.287

the three dynasties, and the three groups of members can mark the whole collection of traditional Chinese painting images, and the amount of work is similar and relatively balanced, which is convenient for personnel organization and related work.

Correlation is a common phenomenon in a large number of data. If there is some regularity between two or more data, it is called association. Finding association rules is to find the correlation of different items in the same event.

The concept of association rule mining first appeared in the Apriori algorithm proposed by Rakesh Agrawal. Its purpose is to mine whether transactions a and B have a high correlation. Because of its solid theoretical foundation and wide application background, association rules have become one of the most important research methods in data mining. In different situations and actual needs, this technology has been paid special attention by researchers in the fields of statistics, data analysis, computer science, data visualization, and so on. By combing the relevant literature, firstly, the relevant definitions of association rules are briefly described.

As shown in formula (1), an item set is a set with P different items, where the item refers to the elements contained in I. For item set I, nonempty subsets with K items are called k-item sets. The transaction set is shown in formula (2), that is, a set with n transactions. The transaction is shown in formula (3).

$$i = \{i1, i2, ..., ip\}, \quad (1)$$

$$D = \{T1, T2, ..., Tn\}, \quad (2)$$

$$T = \{i1, i2, ..., im\} an\ d\, m \le p. \quad (3)$$

Association Rules: if the item sets $A \in I$ and $B \in I$, the implication in the form of $A \longrightarrow B$ is called association rules,

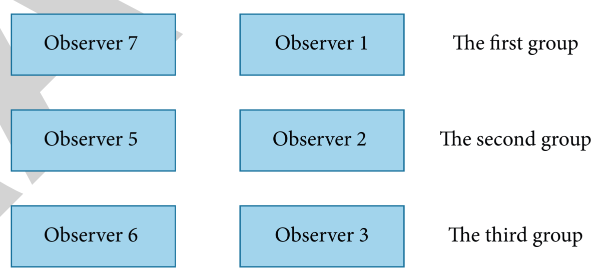


Figure 3: Standard observer combination 1.

where A is called the preceding item set and B is the subsequent item set. In order to measure the strength of this association form, confidence, and support are defined.

Support: support is used to measure the importance or scope of application of association rules. It reflects the universality of rules in all transactions. Indicates the probability of containing itemset A and itemset B in all data, that is, calculates the ratio of the number of transactions containing itemset A and itemset B in the transaction set D to the total number of transactions n contained in the transaction set. The calculation formula is shown in the following formula:

$$Sup(A \longrightarrow B) = P(A \cup B). \quad (4)$$

For the support calculation of a single itemset a, that is, the sum of the number of itemset A and the total number of transactions is included in all transactions. The calculation formula is shown in the following formula:

$$\mathrm{Sup}(A) = P(A). \quad (5)$$

Definition 5–4 (confidence): confidence is a conditional probability, which measures the accuracy of this correlation.

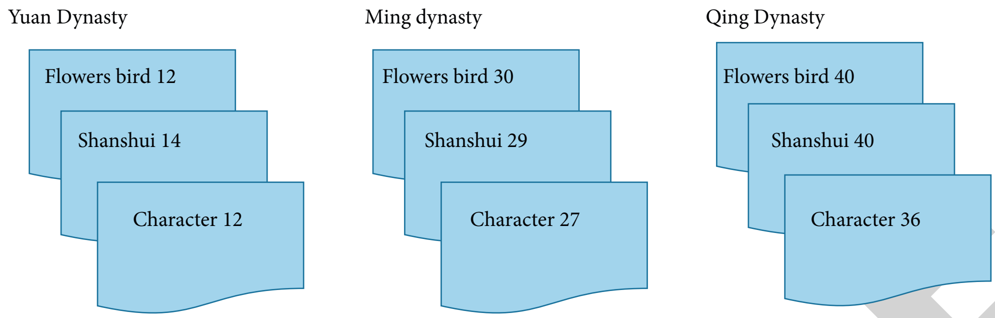


Figure 4: Standard observer combination 2.

Indicates the probability that itemset B also exists when the existence of itemset a in the transaction is known, that is, calculates the ratio of the number of transactions containing itemset A and itemset B in the transaction set D to the number of transactions containing itemset a in the transaction set. The calculation formula is shown in the following formula:

$$\mathrm{Conf}\,(A \longrightarrow B) = P\,(B/A). \tag{6}$$

When using association rules, it is necessary to give a controllable minimum range value of support and confidence according to specific application scenarios and purposes. The lowest range value of support is called the minimum support threshold minSup, which represents the lowest universality of the applicable range of association rules. The lowest range of confidence is called the minimum confidence threshold minConf, which is used to represent the lowest accuracy of the correlation.

The purpose of mining association rules is to find strong association rules that exceed the minimum support threshold minSup and minimum confidence threshold minConf specified by researchers in the transaction set D. The mining of association rules mainly consists of two steps. The first step is an important step of association rule mining. We need to find all frequent item sets with support not less than the minimum support threshold in the database. Through this step, we can quickly get all frequent item sets in the data. The second step is simpler and more direct than the first step. This step needs to build the association relationship between frequent item sets and screen out strong association rules, which belong to the process of enumeration exploration. According to the above-given steps, the association rule mining process is summarized into the model in Figure 5. When researchers input data into the program, the algorithm for finding frequent item sets begins to work, so as to search all frequent sets. Then, the output of rules with strong association relationship is obtained through the algorithm of association rule mining. In this process, the minimum support threshold minsup and the minimum confidence threshold minconf are set by the analyst to control the output results as shown in Figure 5.

The objects in the image are usually expressed in different sizes. In order to solve the disadvantage of analyzing the image signal by Fourier transform, researchers will use all the time-domain information of a signal when extracting the signal spectrum. The overall global transformation lacks the time-domain positioning function of the wavelet transform, there is no way to simultaneously the information of the time domain and frequency domain, and the resolution of the time domain and frequency domain can not be adjusted adaptively according to different signals. Therefore, wavelet transform is usually used for image analysis. The images are decomposed one by one at different scales. The larger the scale, the coarser, otherwise the more detailed. Finally, the results are compared to obtain valuable information. The multiresolution analysis is in line with the following characteristics: system closed subspace, let $\{V_j\}$ belong to $L^2\,(R)$.

(1) *Uniform Monotonicity*: as shown in formula (7)
(2) *Progressive Completeness*: as shown in formula (8)
(3) *Scaling Regularity*: as shown in formula (9)
(4) *Translation Invariance*: as shown in formula (10)

$$V_j < V_j - 1,\, j \in Z, \tag{7}$$

$$\cap V_j = \{0\},\, \cap V_j = L^2\,(R), \tag{8}$$

$$f\,(t) \in V_j, \tag{9}$$

$$f\,(t) \in V_0. \tag{10}$$

According to the compression analysis of signals and images by researchers, the matrix ¥ can be decomposed into two parts: significant area and nonsignificant area, as shown in formula (11).

$$X = XZ_0 + E_0. \tag{11}$$

An important method to solve a constrained optimization problem is to construct an auxiliary function so that the unconstrained minimum point of this function is also the minimum point of the constrained problem and then use the unconstrained optimization minimization method to find the minimum point of the auxiliary function, so as to obtain the optimal solution of the original problem.

For a constrained problem, as shown in the following formulas:

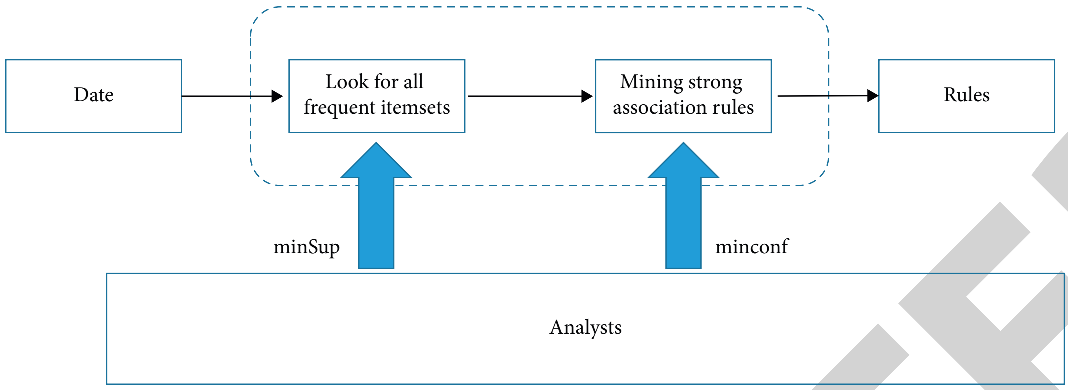


Figure 5: Basic model of association rule mining.

Table 7: Performance confusion matrix of Chinese painting image semantic classification based on the proposed algorithm.

	Landscape painting	Flower and bird painting	Figure painting
Landscape painting	0.79	0.11	0.10
Flower and bird painting	0.22	0.75	0.03
Figure painting	0.17	0.14	0.69

$$\min f(X), \tag{12}$$

$$s.t. h(X) = 0, \tag{13}$$

where $f{:}R$ “$\longrightarrow R$, $f{:}$R”$\longrightarrow R$‘. Its augmented Lagrange function is shown in the following formula:

$$L(X,Y,\mu) = f(X) + Y, h(X) + \frac{\mu}{2} h(X). \tag{14}$$

Based on the above-given feature representation and related parameters, the performance of this algorithm in the semantic classification of three types of Chinese painting images is approximately 74.4%. The element values on the diagonal in the confusion matrix shown in Table 7 represent the classification recall rate of each type of Chinese painting semantic image.

It can be seen from Table 7 that the misclassification rate of landscape and flower and bird Chinese painting images has maintained a relatively low level. In particular, landscape Chinese painting images have achieved the best classification recall rate, while Figure Chinese painting images have the lowest classification recall rate. It should be pointed out that the intraclass difference of landscape Chinese painting images is small because the intraclass texture structure is relatively similar and the color information is relatively unified; On the contrary, due to the small number of training samples and large differences within the class, some character Chinese painting images belong to close-up portraits, while others belong to distant sketches. Therefore, the classification performance of this algorithm in character Chinese painting images is worse than that of landscape, flowers, and birds. Because the visual word package model used in this algorithm makes full use of top-down supervised learning and bottom-up visual saliency information, semantic weighting is carried out for the traditional visual word package model, with special emphasis on the weight of sub-blocks corresponding to saliency regions and the importance of visual words representing specific categories in classification; Moreover, the extracted low-level visual features are color sift descriptors fused with color information, so this algorithm has achieved good results in the semantic classification of Chinese painting images.

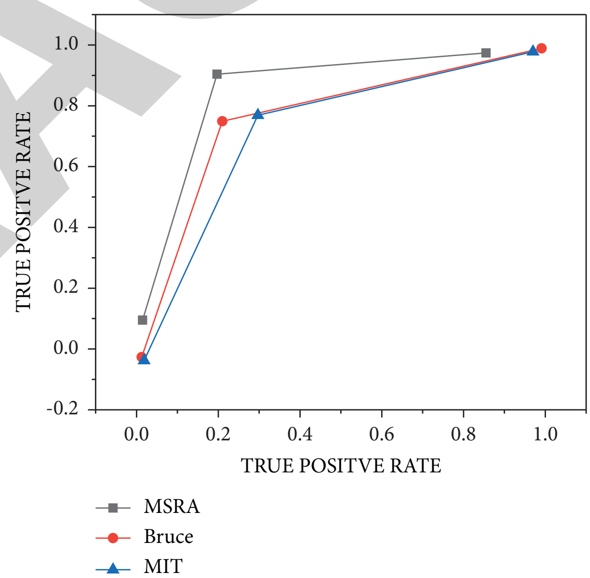


Figure 6: ROC curve of the algorithm on MIT, Bruce, and MsrA data sets.

4. Results and Analysis

Image feature extraction and representation is the most basic and key link in image classification. At present, the commonly used low-level visual features mainly include color, texture, and shape. They can be used to describe local and

Table 8: MIT database.

	IT	GBVS	SI	SR	CSD	FT	SUN	Ours
AUC	0.62	0.67	0.71	0.72	0.68	0.62	0.64	0.74
CC	0.34	0.33	0.33	0.39	0.34	0.26	0.31	0.49

Table 9: Bruce database.

	IT	GBVS	SI	SR	CSD	FT	SUN	Ours
AUC	0.61	0.68	0.73	0.69	0.70	0.55	0.64	0.70
CC	0.38	0.43	0.39	0.31	0.31	0.25	0.34	0.45

global features of images. For the visual word package representation model [25, 26].

In order to effectively measure the proposed salient region extraction algorithm, we first give the experimental results of the proposed algorithm on the current MIT and Bruce eye movement databases and MSRA database. Subjectively, we can see that the proposed algorithm has better similarity with the standard saliency map composed of eye movement data. Furthermore, Figure 6 shows the ROC curve of the algorithm. From the comparison of the three databases, the relative performance of the Bruce dataset is poor. The relationship between the image content and the performance of the algorithm is analyzed by measuring the entropy of the original image. Image entropy is a statistical form of features, which reflects the average amount of information in the graph. The one-dimensional line of an image represents the amount of information contained in the aggregation features of gray distribution in the image, while the two-dimensional line can also reflect the spatial features of image restoration distribution. It can be seen from this that the low line image often contains a core object (i.e., the significant area is more obvious), while the high line image contains more than one under different textures (the significant area is not obvious or more than one area attracts observation). In the Bruce dataset, the image has a relatively high entropy. Some image backgrounds in the image library are messy and the region of interest is fuzzy [27, 28]. The mean value in the MIT database is 0.10 and the standard deviation is 0.04, while the mean value in the Bruce database is 0.15 and the standard deviation is 0.07. Therefore, the algorithm proposed in this chapter can be better neutral in low line image as shown in Figure 6.

In order to further verify the effectiveness of this algorithm. Tables 8 and 9 show the comparison results between this algorithm and other typical algorithms. Experiments on MIT and Bruce eye movement databases show that the proposed algorithm is obviously superior to other methods. The proposed algorithm is more consistent with the process of human visual attention. These results show that the proposed algorithm has advantages in significance detection. The efficiency of the proposed algorithm is mainly due to its ability to capture information from multiple features of the unified reasoning process. In terms of the AUC index, the algorithm in the Bruce database has equivalent performance with Si and CSD, but the algorithm proposed in this chapter has significantly improved in terms of correlation with the saliency map generated by eye movement data.

5. Conclusion

It is proved that the application of color observation technology in Chinese image processing is feasible. It can effectively solve the problem of color processing of classical Chinese painting, meet the requirements of digital museum technology for the digital development of classical Chinese painting and the restoration of ancient painting, make up for the shortcomings that Chinese painting is difficult to repair, and improve the love of most people for Chinese painting. It can reasonably establish the relationship between high-level semantics and low-level features of images, so as to make computers better understand Chinese painting images and improve the performance of semantic analysis and classification of Chinese painting images. It is an important means to realize perception oriented Chinese painting image semantic processing. It is an important means to improve the semantic classification performance of Chinese painting images.

Data Availability

The labeled dataset used to support the findings of this study is available from the author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This study was supported by the Teaching Quality and Teaching Reform Project in 2019 of Guangdong: On the development of Chinese painting teaching by visual communication (Grant no. 2019014).

References

[1] J. Zhang, X. M. Li, and M. S. Miao, "Analysis of animal models of heart failure based on characteristics of clinical symptoms of traditional Chinese and western medicine," *Zhongguo Zhong yao za zhi = Zhongguo zhongyao zazhi = China journal of Chinese materia medica*, vol. 46, no. 4, pp. 786–791, 2021.

[2] M. A. Neglinskaya, ""Gunpowder painting" of Cai Guo-Qiang: Chinese artistic tradition in the era of postmodernism," Культура и искусство, vol. 35, no. 2, pp. 44–50, 2020.

[3] F. Li and K. M. Zubrilin, "Methods of teaching traditional Chinese painting style "mogufa" in the moscow institute of arts (China)," *Science and School*, vol. 2, no. 2020, pp. 217–221, 2020.

[4] Q. M. Guo and Z. H. Qin, "Development and application of vapor deposition technology in atomic manufacturing," *Acta Physica Sinica*, vol. 70, no. 2, p. 028101, 2021.

[5] W. Lin, "The value of DTI parameters in predicting postoperative spinal cord function fluctuations in patients with high cervical disc tumors," *Journal of Visual Art and Design*, vol. 13, no. 1, pp. 60–73, 2022.

[6] D. Perring, Z. Li, R. Pang, K. Wehr, and S. Huang, "Survey and digital documentation of endangered temple wall paintings in shanxi province, China," *Archaeology International*, vol. 22, no. 1, pp. 79–85, 2020.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 3401074, 9 pages
https://doi.org/10.1155/2022/3401074

Research Article

An Improved Knowledge Graph Question Answering System for English Teaching

Li Wang

College English Department, Henan Finance University, Zhengzhou, Henan 450046, China

Correspondence should be addressed to Li Wang; wangli1@hafu.edu.cn

Received 13 May 2022; Accepted 9 June 2022; Published 29 June 2022

Academic Editor: Jiafu Su

Knowledge graph, as a structured semantic knowledge base, has become an essential foundation for artificial intelligence applications with its flexible composition structure and rich semantic representation capability. This paper combines the knowledge graph embedding scoring algorithm with the link scoring algorithm to effectively solve the problem of missing answers in the current knowledge graph embedding question and answer method. This method constructs a query link while searching for the best answer and gives the answer set through the query, which effectively alleviates the omission of answers in the existing methods. The experimental results show that the F1 score of the English teaching test system on the data set is 86.85%, where the answer selection method weakly relies on a priori information such as predicates in the test data and can be trained on a test pair data set without human intervention, with good generalization performance.

1. Introduction

With the proliferation of information resources on the internet, traditional search engines are difficult to meet users' needs for accurate information search, both efficiency and accuracy. Therefore, question and answer systems have been proposed and developed rapidly, and their applications in artificial intelligence, natural language processing, and information retrieval have obtained better results, which is a research hotspot with greater development prospects at present [1]. Knowledge-Based Question Answering (KBQA) is an important component among the question and answer systems.

In recent years, intelligent question answering has made great development. Many intelligent question answering systems have entered people's life and brought great convenience to people [2]. Siri, an intelligent voice assistant developed by Apple, can not only answer questions intelligently but also control the voice of mobile phones. After that, major companies also launched their own voice assistant or question and answer system. For example, Microsoft developed a voice assistant on Windows, Microsoft Xiaona Cortana, Baidu launched its own artificial intelligence assistant Xiaodu, and chat robot Xiaobing developed by Tencent[3]. According to the source of data, intelligent Q&A can be divided into three categories: (1) knowledge-based Q&A, also known as knowledge map Q&A, that is, to retrieve answers directly from the constructed structured knowledge base [4]. (2) Text based Q&A, also known as machine reading comprehension (MRC) Q&A, each question corresponds to several unstructured text data, and the answers are retrieved and extracted from the text data [5]. (3) Based on community Q&A, the Q&A pairs generated by users constitute the data of community Q&A, such as Baidu know, Sogou Q&A, Zhihu, and other forums [6]. With the development of knowledge Atlas, knowledge Atlas Q&A has more and more important practical significance.

The development of knowledge Atlas question answering system is closely related to the development of knowledge Atlas. Knowledge map was originally designed to improve the performance of the search engine and improve users' search quality and search experience [7]. At present, the widely used storage framework of knowledge map is resource description framework (RDF), which is generally represented by SP (subject predicate object) triplet, that is, "subject predicate object" [8]. Among them, "subject" is generally entity, "predicate" is generally relationship or attribute, and "object" is generally entity or attribute value. The whole triplet represents the information between entities and their own attributes [9]. At present, the mainstream knowledge graph question answering methods are divided into two categories: semantic analysis method and information retrieval method.

The first is based on semantic parsing. In the early days, such methods used dictionaries, rules, and machine learning to directly analyze entities, relationships, and logical combinations from problems. However, this kind of method requires researchers to understand the relevant knowledge of linguistics and a large amount of annotation data. It is not easy to expand to the large-scale open domain knowledge graph question and answer task, and the generalization ability is not strong [10]. With the application of deep learning in NLP field, the combination of various neural network models and semantic parsing strategies has become the mainstream of semantic parsing methods [11]. Literature [12] introduces the graph information for semantic analysis and proposes a stage query graph generation method. This idea is also widely used in other semantic analysis generation processes. There is also a semantic analysis method based on encoder decoder. Literature [13] uses sequence to sequence model to translate the problem into multiple relational sequences. Reference [14] proposes to use the atomic operation of state transition to improve the result of problem semantic analysis. Semantic parsing based methods usually use classification models to predict relationships. However, because the knowledge graph contains hundreds of thousands of relationships, the training set is difficult to cover such large-scale relationships, so the semantic parsing based methods are limited in the question answering of the knowledge graph.

The second type is the method based on information retrieval. This kind of method first obtains several candidate entities according to the question, extracts the relationship connected with the candidate entities from the knowledge graph as the candidate query path, and then uses the text matching model to select the candidate query path with the highest similarity to the question to retrieve the answer. In the early stage, it was mainly based on the method of feature engineering. Literature [15] first analyzed the questions and extracted the candidate answers, and then generated the combined ranking of question features and candidate answer features. This method needs to customize the construction features and has a poor processing effect on complex problems. In recent years, representation based learning methods have been proposed and achieved good performance. Representation learning is to graph the candidate entities in the problem and knowledge graph to a unified semantic space for comparison. Literature [16] uses multi column convolution network to represent the semantic information of different aspects of the answer. Literature [17] proposed a translation distance model, Transe, to learn the translation invariance of entities and relations in low dimensional space and proved the effectiveness of embedding human in the knowledge graph on some related problems. Literature [18] proposed a rescal semantic scoring model, which models the latent semantics of triple facts and completes the embedded representation of the knowledge graph from the perspective of semantics. The method of information retrieval transforms the complex semantic analysis problem into a large-scale learnable problem. It focuses on calculating the similarity between the problem and the candidate relationship and has better generalization ability in relationship selection. In addition, there are also some new methods, such as complex problem decomposition, the combination of neural computing and symbolic reasoning, the use of memory network to realize question answering, and so on.

The current mainstream knowledge graph-based Q&A methods only use a single scoring mechanism to score and rank the candidate entities and then output the single entity with the highest score as the answer, which may lead to missing answers when facing multiple answer entities. Although such methods can utilize the semantic information learned during knowledge graph embedding, they do not explicitly construct knowledge graph queries. In this paper, we propose an improved multi-hop ELT knowledge test method based on the knowledge graph embedding, which introduces a relational link scoring mechanism in the answer scoring part and outputs all candidate entities on the same relational link when the best answer entity is obtained; thus, effectively solving the answer omission problem and improving the robustness of the knowledge graph embedding test method. In the question embedding model, this paper improves the embedding of sentence vectors for the ELT knowledge test domain so that the model can better understand the English semantics.

2. Methodology

2.1. English Teaching Knowledge Graph Question Answering Method. Common knowledge graph testing tasks are classified into single-hop questions and multi-hop questions. Single-hop questions apply a single fact from the knowledge base to an answer. Multi-hop questions require two or more facts to be used together as the basis for an answer. Figure 1 shows the difference between these questions. Some current research methods have achieved high accuracy in answering single-hop questions, but there are still many difficulties and challenges in answering multi-hop questions [19], which are as follows: (1) Since multi-hop questions have higher semantic complexity compared to single-hop questions, it is difficult for the model to accurately separate multiple semantic relations from the question sentences. (2) The real knowledge graph is sparse and often misses. For example, in problem 3 in Figure 1, if the relationship between (Lu Xun, Chinese name, Zhou Shuren) and if the relationship link of the entity (Lu Xun, Chinese name, Zhou Shuren) is missing in the knowledge base, it is difficult to search for the true answer by the existing methods. Some current knowledge graph embedding methods can effectively capture the potential semantic relations in the knowledge base and have better prediction ability for the missing links. Thus, this paper proposes a multi-hop ELT knowledge question and

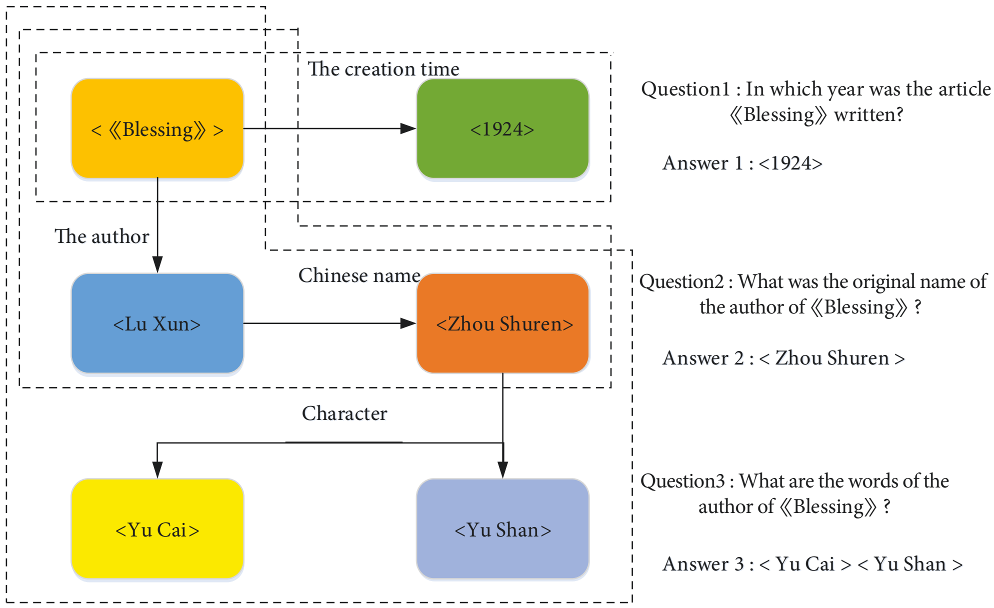


Figure 1: Examples of single-hop and multi-hop knowledge question answering.

answer method with the knowledge graph embedding model as the main body.

In this paper, a knowledge graph is defined as, given a set of entities E and a set of relations R, a knowledge graph ZA can be represented as a set of triples, i.e., $ZA \subseteq E \times R \times E$. For any triple n in ZA, it can be represented as an ordered pair (b, r, n), where $b, n \in E$ and $r \in R$, usually called b as the head entity and n as the tail entity. On this basis, the task is defined as the known knowledge graph $ZA \subseteq E \times R \times E$. Given a natural language question v_x, the central entity of the question is b, where $b \in E$. The role of the question and answer method is to give the set of answer entities $G_x \subseteq E$, for which $\forall g \in G_x$ is a reasonable answer entity for the question v_x.

The overall process can be divided into two parts as shown in Figure 2, which are the answer scoring part based on knowledge graph embedding and the link scoring and answer filtering part. The process steps of the multi-hop ELT knowledge test method based on knowledge graph embedding are as follows.

(1) For the infant question v and its head entity b_v, $\mathbf{e}_{b_v}$ is obtained by querying the embedding vector table obtained by pre-training.

(2) The embedding vector m_v of question v is computed by the sentence vector embedding model.

(3) Compute the knowledge graph embedding score for all candidate answers g , the formula is as follows:

$$S_g(g, b_v, v) = \mathrm{Re}\left(\langle m_v, e_{b_v}, \overline{e}_g\rangle\right), \tag{1}$$

where e_g can be obtained by querying the embedding vector table.

(4) For the further narrowed candidate answer g, the link score is calculated as shown in formula:

$$S_R\left(R_{(b_v, g)}, v\right). \tag{2}$$

(5) Select the best answer entity by matching the score, the formula is as follows:

$$g_{\text{best}} = \arg\max_g S_g(g, b_v, v) + \beta \times S_R\left(R_{(b_v, g)}, v\right). \tag{3}$$

(6) The query link is constructed based on g_{best}, and the answer set G is obtained, and the following query methods are available according to whether b_v and g_{best} are connected.

(a) When b_v and g_{best} are connected, $G = \mathrm{SPARQL}(b, R_{(b_v, g_{\text{best}})})$.

(b) When b_v and g_{best} are not connected, $G = \{g_{\text{best}}\}$.

(7) Return the set G as the result of the method calculation.

2.2. Knowledge Graph Embedding. A typical approach is divided into two parts. One defines the representation of entities and relations in the vector space (usually $\mathbb{R}^d$ or $\mathbb{C}^d$), and the other is to give the scoring function $f_{\text{triple}}(b, r, n)$ of a triple under this representation. The main role of the scoring function is to evaluate the rationality of the triple. In this premise, embedding models can be classified into distance transfer models and semantic scoring models according to the type of scoring functions. The former treats the plausibility of facts as the distance between vectors, while the latter evaluates the potential semantic relationships between entities.

In the knowledge graph embedding-based answer scoring, the scoring process is considered as a link prediction problem. It is hoped that the question-answer rationality

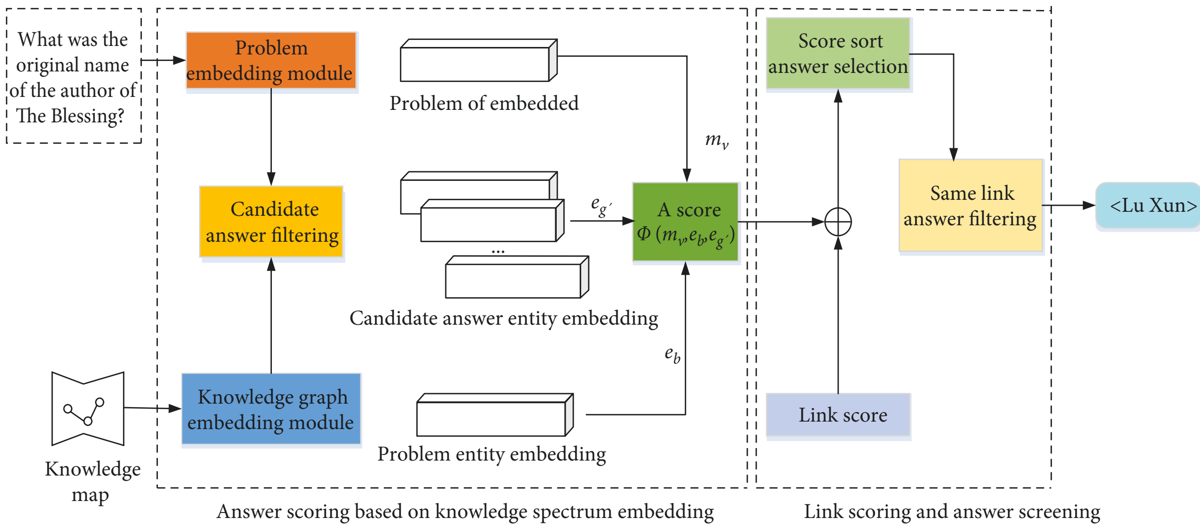


Figure 2: Overall process of multi-hop Chinese knowledge graph question answering.

evaluation can be combined with the triadic scoring in the embedding model by using the question embedding module to learn the semantics of the multi-hop relations contained in the question sentences. Literature [20] shows that the ComplEx method has excellent modeling capability for complex latent semantic relations, and the time complexity of the model pre-training algorithm is low. Therefore, the ComplEx embedding method is chosen in this paper.

2.2.1. ComplEx Embedding Method. The ComplEx embedding method extends the semantic embedding in the real space to the imaginary space, given a head entity with a tail entity $b, n \in E$ and a relation $r \in R$. ComplEx learns the vector representation $\mathbf{e}_b, \mathbf{e}_n, \mathbf{m}_r \in \mathbb{C}^d$ in the imaginary space based on the scoring function. In the embedding model, all triples considered reasonable have $\Phi(r, s, n) > 0$, all triples considered unreasonable have $\Phi(r, s, n) < 0$, and satisfies

$$\Phi(r, s, n) = \mathrm{Re}(m_r, e_b, \overline{e}_n) = \mathrm{Re}\left(\sum_{x=1}^{d} m_{rx} e_{bx} \overline{e}_{nx}\right), \quad (4)$$

where *Re* denotes the real part of the element and $\overline{\mathbf{e}}_n$ denotes the conjugate vector of $\overline{\mathbf{e}}_n$. This property allows the head and tail entity to get different fractional values when exchanging positions. Therefore, ComplEx can learn asymmetric relations, which is also more consistent with the nature of relations between entities in real knowledge graphs, making the embedding model more expressive in question and answer tasks.

2.3. Answer Scoring based on the Knowledge Graph Embedding. The answer selection module is dominated by the scoring of answers based on the knowledge graph embedding. In this paper, we use an improved question embedding model to obtain the sentence vector m_v and combine the embedding vectors e_{b_v} and e_g of related entities to obtain the answer scoring of this part by the scoring function, which is improved from the scoring function based on the ComplEx method.

2.3.1. Question Embedding. The main task of the question embedding module is to embed a question sentence v composed of natural language into a complex space to obtain a sentence vector $m_v \in \mathbb{C}^d$. In some ELT question and answer methods, the typical way to obtain the sentence vector is to fine-tune it for the downstream task based on the BERT pre-trained language model units. In this paper, the pre-trained language model is used to process the interrogative sentences, and the model incorporates $N-$gram vocabulary for representation enhancement, which preserves the semantic information of the sentences to a greater extent. The overall structure of the question embedding model is shown in Figure 3.

As seen in Figure 3, the network uses the ZEN model to embed the problem v into a 768-dimensional vector, which then passes through four fully connected layers and finally maps into the complex space $\mathbb{C}^d$. The learning is based on using the semantic scoring function in ComplEx. The composite semantic representation of multi-hop relations is obtained by approximating the relational vector m_r with the sentence vector $\mathbf{m}_v$. Given a question v and its head entity b_v, with a reference answer set C_v. The network learns based on the criteria defined in formula (5), and the loss function is the cross-entropy loss function.

$$\begin{aligned} &\mathrm{Re}\left(\langle m_v, e_{b_v}, \overline{e}_g\rangle\right) > 0, \quad \forall g \in C_v, \\ &\mathrm{Re}\left(\langle m_v, e_{b_v}, \overline{e}_{\bar{g}}\rangle\right) < 0, \quad \forall \bar{g} \in C_v. \end{aligned} \quad (5)$$

2.3.2. Scoring Function. When performing answer inference, the sentence vector w_q is obtained by question embedding, and the embedding vectors $\mathbf{e}_{b_v}$ and $\mathbf{e}_g$ of the head entity and candidate answer entities are known, the model

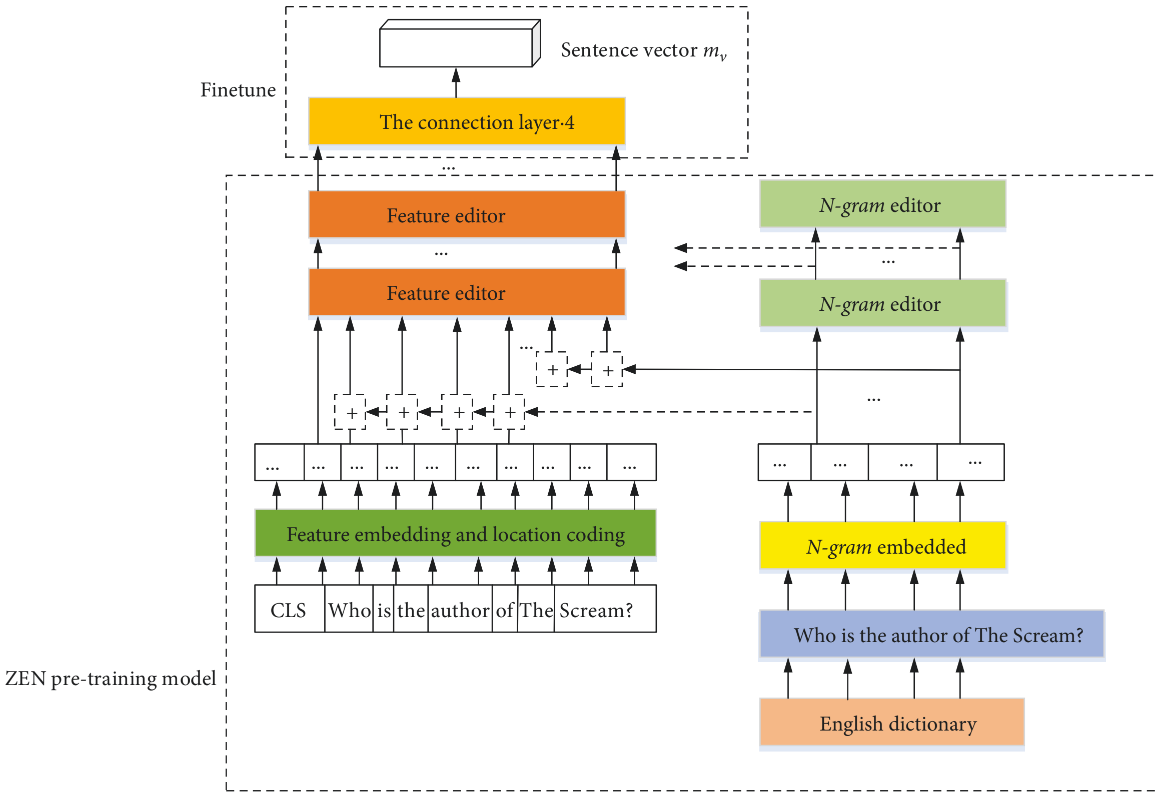


FIGURE 3: Question embedding based on ZEN.

can calculate the answer score $S_g(g, b_v, v)$ for each candidate answer, which is calculated as follows:

$$S_g(g, b_v, v) = \mathrm{Re}\left(\langle m_v, e_{b_v}, \bar{e}_g\rangle\right) = \mathrm{Re}\left(\sum_{x=1}^{d} m_{vx} e_{b_v} \bar{e}_{gx}\right). \quad (6)$$

The answer scoring function in formula (6) is formally consistent with the triadic scoring function, which is the main reason why the model can make full use of the unsupervised information in the knowledge graph embedding. In order to reduce the range of candidate answers, a scoring threshold S_{nb} can be given, and only the answers with scores above the threshold are selected as the set of candidate answers for link scoring.

2.4. Link Scoring and Answer Filtering. Unlike the query construction method, the answers obtained by the knowledge graph-based embedding method do not depend on the query path. The advantage of this method is that it has link prediction capability for sparse knowledge graphs, and can achieve a high score for the correct answer even if the query path is missing, but the disadvantage is that the robustness of the model is poor. This disadvantage is mainly reflected in the following two aspects.

(1) One of the advantages of knowledge graph Q&A is that the relational paths in the knowledge graph have higher credibility and interpretability, and simply using the knowledge graph embedding for answer scoring will weaken this advantage.

(2) The real answer set usually consists of multiple entities on the same query link, and it is difficult to obtain an accurate answer set by scoring with knowledge graph embedding.

To overcome these drawbacks, this paper introduces a link scoring mechanism based on knowledge graph embedding scoring to enhance the method's robustness.

2.4.1. Link Scoring. The query link $R_{(b_v,g)}$ between the head entity b_v of question v and the answer entity g consists of a sequence of relations $(r_1, r_2, \ldots, r_n)$. To measure the relevance between the relations and the question. In this paper, we construct a metric score $S_r(r_x, v)$ for network learning relevance based on formula (7), and use it to give a link score $S_R(R_{(b_v,g)}, v)$.

$$S_r(r_x, v) = \sigma\left(\mathrm{Re}\left(m_{r_x}^N\right) \cdot \mathrm{ZEN}(v)\right), \quad (7)$$

where m_{r_x} is the relationship vector in the knowledge graph embedding, $\mathrm{ZEN}(\cdot)$ is the English pre-trained language model, and $\sigma(\cdot)$ is the activation function, which is the Sigmoid function in this paper. Thus, different link scoring functions can be given for the two other cases of whether the head entity b_v and the answer entity g are connected or not.

When there is an optimal connected link $R_{(b_v,g)}$, the link scoring function is as follows:

$$S_R\left(R_{(b_v,g)}, v\right) = \sum_{r_x \in R_{(b_v;g)}} S_r(r_x, v). \quad (8)$$

When there is no connected link, the link scoring function is as follows:

$$S_R\left(R_{(b_v,g)}, v\right) = \underset{\substack{r \\ r_x \in \phi_g}}{\arg\max}\, S_r\left(r_x, v\right), \tag{9}$$

where ϕ_g is the set of relations of single-hop reachable answer entities g. The meaning of formula (8) gives a complementary scoring term to the answer entities that may have missing links, which is used to eliminate the bias, and the term can be set to 0 when the knowledge graph is dense.

2.4.2. Answer Filtering. To deal with the case of multiple answer entities, the same link query mechanism is used in the answer filtering. The formula is as follows:

$$g_{\text{best}} = \underset{g}{\arg\max}\, S_g\left(g, b_v, v\right) + \beta \times S_R\left(R_{(b_v,g)}, v\right). \tag{10}$$

Firstly, the answer entities are scored based on the knowledge graph embedding score and link score, and the best answer entity is given after ranking, where β is a hyperparameter.

When the optimal link $R_{(b_v,g_{\text{best}})}$ exists, the answer set G is as follows:

$$G = \text{SPARQL}\left(b, R_{(b_v,g_{\text{best}})}\right), \tag{11}$$

where $\text{SPARQL}(b, R)$ denotes the set of entities obtained from one graph database query based on the head entity b through the relational sequence R. When there is no connected link, the answer set G is as follows:

$$G = \left\{g_{\text{best}}\right\}. \tag{12}$$

The answer set G is obtained for a given question v and head entity b_v.

3. Result Analysis and Discussion

3.1. Experimental Data Set and Environment. This paper uses the knowledge base and question and answer pair data published by the DBpedia Neural Question Answering data set and selects 15050 training question and answer pairs and 9360 test question and answer pairs. In order to make the evaluation results more objective, the training Q&A pair is further randomly divided into training set and development set, and the test Q&A pair is used as the test set. The data set division is shown in Table 1.

Table 1: Division of data set.

Name	Number
Training set	13000
Development set	2050
Test set	9360

The experiments are run on a computer with Inter15-4590 CPU and 12 GB RAM, Nvidia GTX1080Ti graphics card with 11 GB video memory, CUDA10.0 and Tensorflow1.14 deep learning framework, 64-bit Windows 10 operating system, and Mysql5.6.46 for knowledge base data storage and retrieval using Mysql5.6.46.

3.2. Analysis of Named Entity Recognition Results. The training process iterates 8122 times in total, and Adam optimizer with weight attenuation is used to optimize the loss function. After completing the training on the training set, test the performance of the model on the training set, development set, and test set, respectively. The results are shown in Figure 4. Due to the size of the corpus, there is a slight overfitting phenomenon in the test results, which is basically accurate on the training set and has some errors on the development and test sets, and the overall performance is good.

3.3. Analysis of Answer Matching Results. To train the answer matching network model, we need to create the answer matching data set based on the existing question-answer pairs. Creating negative samples is similar to that of Q&A. The named entity is used as the keyword to search the knowledge base to get the set of answers related to the entity, and the nouns that are not the answer to the question are connected to the question in the same way. For entities with only one triad in the knowledge base, five answers are randomly selected as negative samples from the triads with other entities as keywords and added to the data set to obtain. The size of the answer matching data set is shown in Table 2.

The training set data are fed into the answer matching network for training. Since the feature extraction part of the network also uses BERT, the hyperparameter selection is consistent with the named entity recognition except for the absence of LSTM. The optimizer for model training also used Adam with weight decay, and the network was iterated 12120 times. Since the similarity score is a value between 0 and 1, which cannot be exactly equal to the label, the output of the network is modified as a category when calculating the test metrics, i.e., it is treated as a dichotomous problem and can only output "0" or "1." In addition to the accuracy, AUC is also an important performance index when calculating the performance index. It can more objectively measure the classification effect of the model on the answer matching data set. The test results of the answer matching model on the training set, development set, and test set are shown in Table 3. Due to the limited size of the data, the performance of the model in the development and test sets is poor but the AUC values are above 86%, which guarantees the quality of the final automatic test.

3.4. Analysis of Hyperparameter Selection Results. After completing the training of named entity recognition and answer matching model, the knowledge base Q & A can be carried out. When the super parameter selection mechanism is not added, the answer with the highest super parameter score in the triplet set containing entities in the knowledge base is directly selected as the output, and the Q & A results are shown in Table 4. Since there is only one standard answer

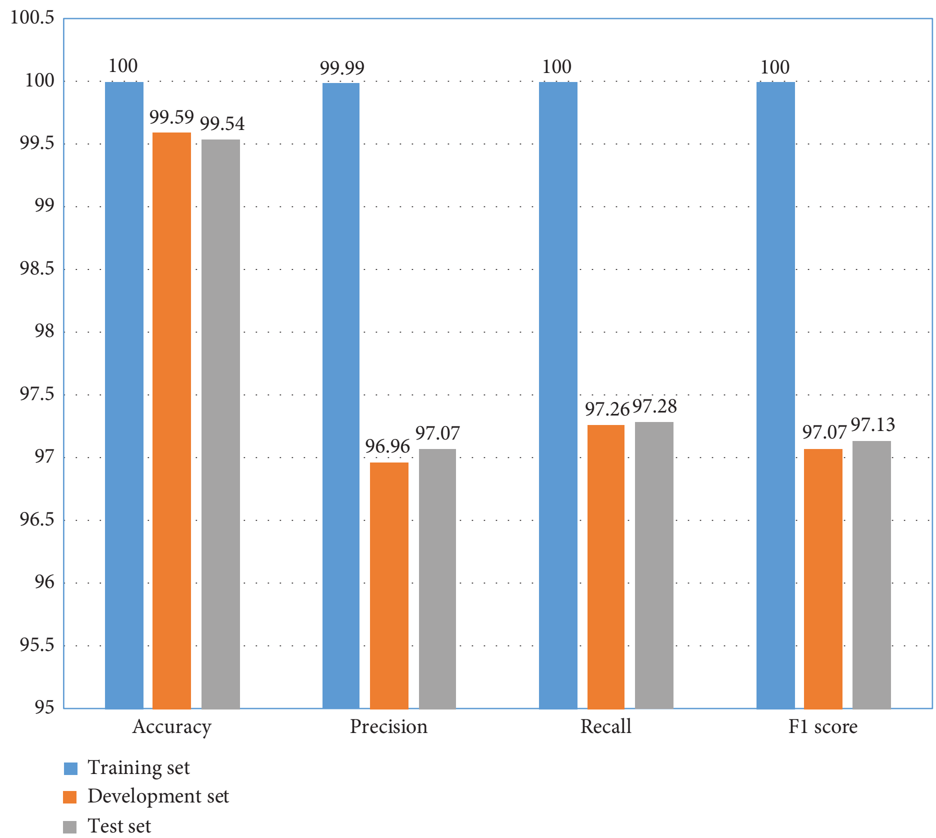


Figure 4: Test results of named entity recognition %.

Table 2: Data set size of answer matching.

Data set	Number of positive samples	Number of negative samples	Total number
Training set	13000	64225	77225
Development set	2050	31866	33916
Test set	9360	40452	49812

Table 3: Test results of answer matching %.

Data set	Accuracy	AUC
Training set	99.44	98.63
Development set	96.47	87.58
Test set	92.56	86.25

and one predicted answer and the scores are the same, only F1 scores are listed.

By recording the incorrectly answered questions and observing their hyperparameters. It was found that in addition to the noise in the data set itself, the hyperparameters among the top few of the more ambiguous answers were mainly around 0.1 to 0.9. In order to determine the best super parameters under the data set in this paper, select five super parameters of 0.1, 0.3, 0.5, 0.7, and 0.9, respectively, and call the super parameter selection mechanism on the development set for testing as shown in Table 5.

From the results of hyperparameter selection, we can see that the precision rate gradually becomes smaller and the recall rate gradually becomes larger as the selected hyperparameter decreases, which is the inevitable result of more alternative answers. When the hyperparameter is 0.5, the F1 score of the test on the development set is the highest.

Table 4: Question answering results without threshold selection mechanism %.

Data set	F1 score
Training set	95.23
Development set	87.71
Test set	86.73

Table 5: Question answering results of development set with different similarity thresholds %.

β value	Precision	Recall	F1 score
0.1	83.89	91.21	87.37
0.3	84.26	91.12	87.52
0.5	85.62	90.91	88.20
0.7	85.85	90.11	87.89
0.9	86.39	89.64	87.89

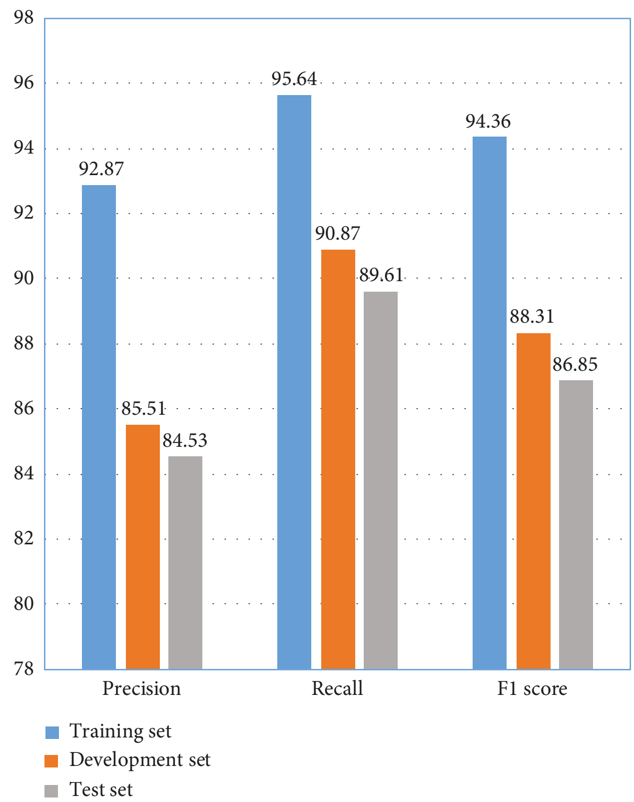


Figure 5: Final results of the knowledge base question answering.

When the hyperparameter decreases further, the precision rate decreases more due to the number of selected answers, so the F1 score decreases as well.

3.5. Comparison of Automatic Test Results. Through the experimental results of hyperparameter selection, 0.5 was selected as the β value of the knowledge base test in this paper, and it was applied to the final ELT test system as shown in Figure 5. Both the training and development sets are derived from the training set of the original test pairs of DBpedia Neural Question Answering task and the score of test set F1 is used in the public evaluation index.

In this paper, the English teaching test system uses hyperparameter selection as an optional switch in practical applications. In many application scenarios where the test task requires a single answer to be returned, the hyperparameter selection switch is turned off and the most accurate answer is presented to the user. If the user has doubts about the answer or if some scenarios allow multiple answers to be returned, the hyperparameter selection can be turned on and the set of candidate answers will be presented in order of similarity from lowest to highest.

This paper selects literature [21–28]and literature [29] as the comparison method. The automatic question and answer results are shown in Table 6. Literature [21] is based on the idea of dynamic programming. Its unsupervised idea has reference significance, but the effect of question and answer is relatively limited. Literature [22–24, 26] and literature [27] are the top five automatic question and answer methods for the evaluation results of Arts & human cities index tasks, respectively. They mainly rely on some manual rules to ensure the question and answer performance. For example, literature [27] constructs regular expressions to remove redundant information from the interrogative sentences. Literature [26] uses combinatorial features of lexicality to achieve named entity recognition, etc. Literature [25] is an automatic question and answer method constructed based on attribute mapping of predicates in knowledge base triples with a small number of artificial features. Literature [28] is an automatic question and answer method implemented by syntactic analysis, etc. Literature [29] first applied BERT for feature extraction on the DBpedia Neural Question Answering data set and achieved the best results published so far. In addition to applying BERT, the method in this paper also improves the answer selection method by decomposing it into two steps, answer matching and hyperparameter selection, which reduces the need for manual annotation and preprocessing and obtains a test set F1 score of 86.85% with the best performance.

Table 6: Automatic question answering results of ten methods %.

Automatic question and answer method	Test set F1 scores
Literature [21]	71.12
Literature [22]	72.65
Literature [23]	79.18
Literature [24]	79.52
Literature [25]	81.30
Literature [26]	81.51
Literature [27]	82.43
Literature [28]	82.87
Literature [29]	84.55
Proposed	86.85

4. Conclusion

In this paper, the multi-hop ELT knowledge test method based on knowledge graph embedding innovatively combines the knowledge graph embedding scoring algorithm and link scoring algorithm in scoring answers, and constructs a query path with high confidence through composite scoring, which effectively solves the phenomenon of missing answers in the current knowledge graph embedding-based test methods. The experimental results show that the F1 score of the English teaching test system on the DBpedia Neural Question Answering data set is 86.58%, in which the improved vector embedding model based on ZEN fits complex semantic relations and provides more accurate semantic understanding for multi-hop test tasks. At the same time, because of the introduction of knowledge graph embedding, the question-answer method has some link prediction ability, and the method still has strong inference ability on the incomplete knowledge graphs. Through the experiments, it is found that the accuracy of the English teaching test system for number type answers in this paper needs to be improved, and the subsequent methods such as representation learning will be used to filter the optimal solutions from the set of candidate answers to improve the test quality further.

Data Availability

The labeled dataset used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] K. P. Kalyanathaya, D. Akila, and P. Rajesh, "Advances in natural language processing–a survey of current research trends, development tools and industry applications," *International Journal of Recent Technology and Engineering*, vol. 7, no. 5C, pp. 199–202, 2019.

[2] Q. Zheng, B. Dong, B. Qian, and F. Tian, "The state of the art and future tendency of smart education," *Journal of Computer Research and Development*, vol. 56, no. 1, p. 209, 2019.

[3] V. A. Kotkar, "Interactive robot for automated question and answer system," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 1S, pp. 90–95, 2021.

[4] S. Hu, L. Zou, J. X. Yu, and W Haixun, "Answering natural language questions by subgraph matching over knowledge graphs," *IEEE Transactions on Knowledge and Data Engineering*, vol. 30, no. 5, pp. 824–837, 2017.

[5] S. Liu, X. Zhang, S. Zhang, H Wang, and W Zhang, "Neural machine reading comprehension: methods and trends," *Applied Sciences*, vol. 9, no. 18, p. 3698, 2019.

[6] S. Momtazi, "Unsupervised latent dirichlet allocation for supervised question classification," *Information Processing & Management*, vol. 54, no. 3, pp. 380–393, 2018.

[7] S. K. Mohamed, A. Nounu, and V. Novacek, "Biological applications of knowledge graph embedding models," *Briefings in Bioinformatics*, vol. 22, no. 2, pp. 1679–1693, 2021.

[8] J. Yuan, Z. Jin, H Guo et al., "Constructing biomedical domain-specific knowledge graph with minimum supervision," *Knowledge and Information Systems*, vol. 62, no. 1, pp. 317–336, 2020.

[9] S. Liu, H. Yang, J Li, and S Kolmanic, "Preliminary study on the knowledge graph construction of Chinese ancient history and culture," *Information*, vol. 11, no. 4, p. 186, 2020.

[10] T. Ruan, Y. Huang, and X. Liu, "QAnalysis: a question-answer driven analytic tool on knowledge graphs for leveraging electronic medical records for clinical research," *BMC Medical Informatics and Decision Making*, vol. 19, no. 1, pp. 1–13, 2019.

[11] X. Li, H. Wang, H. He, D Jixiang, C Jian, and W Jinzhun, "Intelligent diagnosis with Chinese electronic medical records based on convolutional neural networks," *BMC Bioinformatics*, vol. 20, no. 1, pp. 1–12, 2019.

[12] A. M. Rinaldi and C. Russo, "Using a multimedia semantic graph for web document visualization and summarization," *Multimedia Tools and Applications*, vol. 80, no. 3, pp. 3885–3925, 2021.

[13] H. Strobelt, S. Gehrmann, and M. Behrisch, "S eq 2s eq-v is: a visual debugging tool for sequence-to-sequence models," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 1, pp. 353–363, 2018.

[14] R. Nagamanjula and A. Pethalakshmi, "A novel framework based on bi-objective optimization and LAN2FIS for Twitter sentiment analysis," *Social Network Analysis and Mining*, vol. 10, no. 1, pp. 1–16, 2020.

[15] B. Abu-Nasser, "Medical expert systems survey," *International Journal of Engineering and Information Systems (IJEAIS)*, vol. 1, no. 7, pp. 218–224, 2017.

[16] J. Zhang, J. Sun, J Wang, and X. G Yue, "Visual object tracking based on residual network and cascaded correlation filters," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 8, pp. 8427–8440, 2021.

[17] Z. Cao, X. Qiao, S Jiang, and X Zhang, "An efficient knowledge-graph-based web service recommendation algorithm," *Symmetry*, vol. 11, no. 3, p. 392, 2019.

[18] G. A. Gesese, R. Biswas, M Alam, and H Sack, "A survey on knowledge graph embeddings with literals: which model links better literal-ly?" *Semantic Web*, vol. 12, no. 4, pp. 617–647, 2021.

[19] Y. Fang, H. Wang, L. Zhao, F Yu, and C Wang, "Dynamic knowledge graph based fake-review detection," *Applied Intelligence*, vol. 50, no. 12, pp. 4281–4295, 2020.

[20] K. Watanabe, "Latent semantic scaling: a semisupervised text analysis technique for new domains and languages," *Communication Methods and Measures*, vol. 15, no. 2, pp. 81–102, 2021.

[21] E. Dimitrakis, K. Sgontzos, and Y. Tzitzikas, "A survey on question answering systems over linked data and documents," *Journal of Intelligent Information Systems*, vol. 55, no. 2, pp. 233–259, 2020.

[22] A. Thomas and S. Sangeetha, "Semi supervised, knowledge integrated pattern learning approach for fact extraction from judicial text," *Expert Systems*, vol. 38, no. 3, Article ID 12656, 2021.

[23] Q. Lu, Z. Zhu, F Xu, D Zhang, W Wu, and Q Guo, "Bi-GRU sentiment classification for Chinese based on grammar rules and BERT," *International Journal of Computational Intelligence Systems*, vol. 13, no. 1, p. 538, 2020.

[24] R. Zheng, H. Ma, Q. Wang, J Fu, and Z Jiang, "Assessing the security of campus networks: the case of seven universities," *Sensors*, vol. 21, no. 1, p. 306, 2021.

[25] P. Wang, Q. Wu, C. Shen, A V D Hengel, and D Anthony, "Fvqa: fact-based visual question answering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 10, pp. 2413–2427, 2017.

[26] J. Li, A. Sun, J. Han, and C Li, "A survey on deep learning for named entity recognition," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 1, pp. 50–70, 2020.

[27] L. Li, M. Zhang, Z Chao, and J Xiang, "Using context information to enhance simple question answering," *World Wide Web*, vol. 24, no. 1, pp. 249–277, 2021.

[28] D. Diefenbach, V. Lopez, K. Singh, and P Maret, "Core techniques of question answering systems over knowledge bases: a survey," *Knowledge and Information Systems*, vol. 55, no. 3, pp. 529–569, 2018.

[29] F. A. Acheampong, H. Nunoo-Mensah, and W. Chen, "Transformer models for text-based emotion detection: a review of BERT-based approaches," *Artificial Intelligence Review*, vol. 54, no. 8, pp. 5789–5829, 2021.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9868629, 1 page
https://doi.org/10.1155/2023/9868629

Retraction

Retracted: Predicament and Thinking of College Students' Employment and Entrepreneurship under the Background of Supply-Side Reform

Mobile Information Systems

Received 22 August 2023; Accepted 22 August 2023; Published 23 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] M. Yu, "Predicament and Thinking of College Students' Employment and Entrepreneurship under the Background of Supply-Side Reform," *Mobile Information Systems*, vol. 2022, Article ID 1495123, 10 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 1495123, 10 pages
https://doi.org/10.1155/2022/1495123

Research Article

Predicament and Thinking of College Students' Employment and Entrepreneurship under the Background of Supply-Side Reform

Mei Yu

Xinyang Vocational and Technical College, Xinyang, Henan 464000, China

Correspondence should be addressed to Mei Yu; yumei198208@xyvtc.edu.cn

Received 11 May 2022; Revised 6 June 2022; Accepted 10 June 2022; Published 27 June 2022

Academic Editor: Jiafu Su

Based on the analysis of the current employment and entrepreneurship of college students based on the supply-side structural reform, the employment and entrepreneurship guidance of college students is far from meeting the needs of the society. Therefore, the innovation of talent training mode needs to become an important goal of college development at this stage. Based on the previously mentioned background, this paper included 853 university students as research objects and their data from 2010 to 2018. First, we used the BP neural network as the starting point to conduct in-depth research, selected the sequential model algorithm based on the Keras framework, built a one-layer network and six types of eigenvalue labels to predict the development direction of college students' employment and entrepreneurship, and evaluated the prediction accuracy of the model. This proves that the prediction effect of the model has the value of continuing in-depth research, and then, the prediction model is further optimized. Then, we added a three-layer network to the model and an SGD optimizer and used Softmax as the regression function to verify that the optimized model predicts well. The average accuracy of the prediction model constructed in this paper is 81.48%, the standard deviation is 4.34%, the Acc value of the model is stable at around 0.835, and the loss value is stable at around 0.3, which proves that the prediction model has a good prediction. It provides a set of application models that can be used for reference for the combination of BP neural network related knowledge and college students' employment and entrepreneurship development direction prediction. Based on the background of the supply-side structural reform, this paper mainly analyzes and researches the current employment and entrepreneurial paths of college students in my country, hoping to provide reference for the cultivation of talents in colleges and universities.

1. Introduction

In the report of the 19th National Congress of the Communist Party of China, General Secretary Xi Jinping proposed to "deepen supply-side structural reform." To build a modern economic system, we must focus on developing the real economy, take improving the quality of the supply system as the main direction of attack, and significantly enhance the quality of China's economy. The employment problem of contemporary college students is very prominent, but many employers still have talent shortages. On the one hand, under the influence of the country's expansionary policy, the number of college students has been increasing year by year in recent years, while the employment positions that the society can provide are minimal, and new changes have occurred in the difficulty of college students' employment. On the other hand, many employers need high-tech innovative talents, and the number of university graduates who meet this aspect is very limited.

The higher education supply system plays a vital role in the transformation and upgrading of China's economic structure and undertakes the important task of exporting comprehensive talents to society. To adapt to the transformation of China's economic and social structure, the talent output of colleges and universities must conform to the law of the operation and development of the economic market [1]. The supply-side reform of higher education should pay more attention to the quality of education and teaching of college teachers and the comprehensive quality of students, to better connect the professional setting of colleges and universities with social needs and further promote the transformation and development of China's economic field. The reform in the field of higher education should focus on the organic combination of supply and demand and cultivate comprehensive talents suitable for the development of the market economy, so as to adapt to the transformation and upgrading of China's social and economic structure [2].

Since entering the new era, the reform of higher education should be comprehensively promoted, the education of colleges and universities should focus on talent training, and colleges and universities should vigorously build a new mechanism for comprehensive, whole-process, and deep integration of social collaborative education. Under the background of higher education supply-side reform, colleges and universities should establish a talent training mechanism in close cooperation with social employment departments, clarify the talent standards required by the market, jointly formulate talent training programs, improve the level of teaching practice, transform social high-quality educational resources into educational content [3], and collaborate with relevant departments to manage talents. Higher education needs to carry out supply-side reforms to better meet the needs of the market, be guided by social needs, and cultivate high-quality talents that meet social needs [4]. In recent years, some colleges and universities have unilaterally pursued the enrollment rate and enrollment volume, ignoring their own school-running conditions and educational resources, resulting in duplication and homogeneity of majors in colleges and universities [5]. The professional setting structure of colleges and universities is not optimized, the dynamics and market adaptability are insufficient, and the degree of matching with the market is low. Colleges and universities can use big data technology to accurately predict market demand, increase communication and coordination with the market, scientifically adjust the layout and structure of disciplines, and fully realize the supply balance of higher education [6]. With the gradual increase in the number of fresh graduates in our country, the upgrading of the industrial structure has brought about an increase in the employment pressure of college students. The decline in the teaching quality of colleges and universities makes it difficult for talent training to meet the needs of the market. The employment market for college students is not perfect, which affects the employment of college students [7]. The process of urbanization is accelerating, and urban jobs are in short supply. The backward employment concept is difficult to adapt to the requirements of the times. There is gender discrimination in employment positions. The employment pressure on female college students is high. Difficulty starting a business is easy to cause blows [8]. These problems are all in the new situation. Compared with the past, the current college students are facing more and more employment problems. The employment of college students has become one of the urgent problems to be solved at present, and with the continuous changes of social and economic forms, college students' entrepreneurship has become the main means to solve this problem. College students to start a business must rely on the entrepreneurial environment. The so-called entrepreneurial environment mainly refers to the collection of all factors that may have an impact on the entrepreneurial process. The social entrepreneurship environment is the carrier of college students' entrepreneurship, and they will be affected by the entrepreneurial environment during the entire entrepreneurial process of college students. Therefore, the government attaches great importance to the work of optimizing the entrepreneurial environment. At present, many literature studies believe that the entrepreneurial behavior of college students needs various incentive policies and guarantee policies to be carried out scientifically and effectively. In other words, the success of entrepreneurship is inseparable from a suitable entrepreneurial environment.

2. State of the Art

College students have poor professional theoretical foundation, cannot master professional skills proficiently, and have weak practical ability. According to the feedback from the employer, some graduates are not proficient in mechanical memory professional knowledge, and their understanding is not thorough. The knowledge they have learned cannot be used to solve problems in actual work. They are hard to be favored. Graduates do not recognize the importance of social practice activities and participate in less social practice activities, resulting in weak hands-on ability. In the talent training model of colleges and universities, scores are still the main indicator for evaluating students, and they do not pay attention to the cultivation of practical ability to apply theory. In the assessment system, the practical ability of students' practice is not really integrated into it.

Colleges and universities have not truly integrated career development planning into the talent training system, the time is incoherent, the content is not systematic, the theoretical teaching is emphasized, and the teaching is not done according to aptitude. Graduates do not have a strong awareness of "career planning," their career goals are ambiguous, they do not have a clear career vision, and they do not plan for career development. The life of a college student is like being a monk for a day, and just passing by. In the face of fierce competition when applying for a job, college students feel confused and have no opinion, follow the herd mentality for employment, and think that the unit with many applicants is a good unit. They cannot match the personnel and jobs, resulting in low employment quality, unstable work, and high turnover rate of graduates.

Entrepreneurship education mainly trains entrepreneurs from the three levels of consciousness, thinking, and ability [9]. From the perspective of social development, entrepreneurship education is an important task entrusted to colleges and universities by this era of rapid economic development. From the perspective of colleges and universities, doing a good job in entrepreneurship education is the proof that higher education keeps pace with social development. Sexuality, but also hope, is a community of times and urgency [10, 11].

Bejerholm and Larsson point of view is that the goal of entrepreneurship education should be to help college students who are confused to find their own position, identify their goals and directions, make long-term scientific

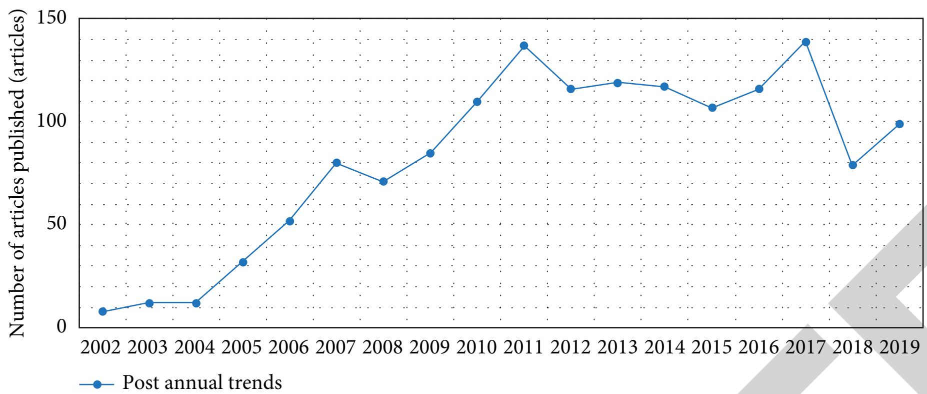


FIGURE 1: Literature search data map.

planning for their careers, and finally achieve their own entrepreneurial goals and realize their own social value [12]. Innovation is the root of social and economic development. Entrepreneurship promotes the advancement and development of society. Entrepreneurship education is the foundation of all this. Scientific and effective entrepreneurial education can cultivate many entrepreneurial talents for the society, and these talents can create more social wealth. College students are fresh blood about to enter the society, full of passion and courage. Encouraging college students to start businesses is not only conducive to improving the current situation of employment difficulties but also conducive to social stability and sustainable development [13].

Compared with Western countries, the research and practice of entrepreneurship education model in China is still short, and related projects are only in their infancy, such as (1) entrepreneurship education, (2) KAB (Know About Business, KAB), (3) SIYB training, (4) China Youth Entrepreneurship International Program, and so on. There are still many problems to be solved. The setting of entrepreneurship education courses and teaching models is still in the stage of imitation and exploration. Most of the forms are greater than the content [14]. A mature and complete entrepreneurship education system suitable for China's national conditions has not yet been formed. There is still a lack of quantitative in-depth research on the level [15].

There are still many problems to be solved in the employment and entrepreneurship of college students. The theme of "student employment and entrepreneurship" was used to manually exclude 223 documents unrelated to keywords. The trend chart of the number published in the literature is shown in Figure 1. Among them, 49.3% are involved in graduate employment and entrepreneurship guidance.

3. Methodology

3.1. Artificial Neural Network Method and Neural Network Model. Artificial neural network (ANN) was born in the late 1940s. Due to its advantages of information distribution and storage, parallel processing, and self-learning ability, it has been widely used in information processing, pattern recognition, intelligent control, and system modeling. It has been more and more widely used in other fields [16].

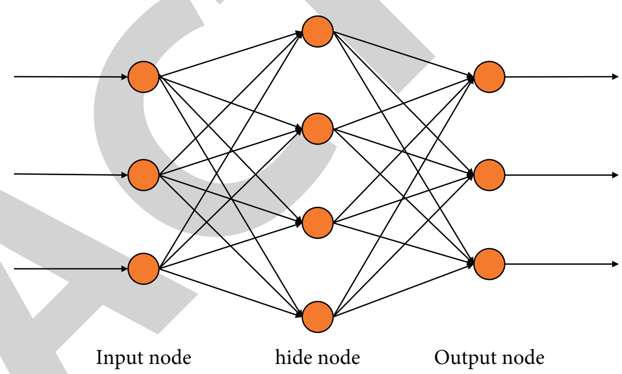


FIGURE 2: Neural network model.

BP neural network is a typical model of artificial neural network. It is a multilayer feed-forward network (multiple-layer feed-forward network, referred to as BP network) based on the error backpropagation algorithm. The network not only has input nodes and output nodes but can also have one or more hidden layer nodes [17]. A typical BP neural network model with input, output, and hidden layers is shown in Figure 2.

For the input signal of the BP neural network, it must first propagate forward to the hidden node, and after passing through the action function, the output information of the hidden node is propagated to the output layer node, and finally, the result is output [18].

If each layer of the BP neural network has N processing units and the change function of the neurons is a sigmoid function,

$$f(x) = \frac{1}{1 + e^{-x}}. \tag{1}$$

This action function makes the output quantity a continuous quantity between 0 and 1, so any nonlinear mapping rule from input to output can be realized. The artificial neuron takes the information it receives (the output of the previous layer) $0o$, o,. . ., o. Use $W1, W2, \ldots, Wn - 1$. The connection strength forms its own input I in the form of a dot product and then converts it by a function to obtain the output O of this unit [19].

In addition, the BP neural network training set includes M sample pattern pairs (X), and for the pth training sample ($p = 1, 2, \ldots, M$). If the sum of the inputs of unit j (that is, the activation function) is denoted as a, the output is recorded as op, and then,

$$\begin{aligned} a_{pj} &= \sum_{j=0}^{n} w_{pj} o_{pj}, \\ o_{pj} &= f\left(a_{pj}\right) = \frac{1}{1 + e^{-a_{\infty}}}. \end{aligned} \tag{2}$$

If the initial value of the network is set arbitrarily, then, for each input mode p, the network output generally has an error with the expected output. The network error is defined as

$$E_p = \frac{1}{2} \sum_{j} \left(d_{pj} - o_{pj}\right)^2, \tag{3}$$

where d represents the expected output of the output unit j of the pth input mode. At the same time, the essence of the network learning rule 8 is to use the gradient steepest descent method to make the weight change along the negative gradient direction of the error. The change of the weight w is recorded as $Apwj$ because

$$\begin{aligned} \frac{\partial E_p}{\partial w_{ji}} &= \frac{\partial E_p}{\partial a_{pj}} \frac{\partial a_{pj}}{\partial w_{ji}} \\ \delta_{pj} &= -\frac{\partial E_p}{\partial a_{pj}}, \\ \Delta p w_{ji} &= -\eta \frac{\partial E_p}{\partial a_{pj}} o_{pj}. \end{aligned} \tag{4}$$

In the process of network learning, the error calculation of the output layer and the hidden layer is not the same. When representing the output of the output layer unit, the error is

$$\delta_{pj} = f'\left(a_{pj}\right)\left(d_{pj} - o_{pj}\right). \tag{5}$$

When op represents the output of the hidden layer, its error is

$$\begin{aligned} \delta_{pj} &= \frac{\partial E_p}{\partial a_{pj}} = -\frac{\partial E_p}{\partial p_{pj}} \frac{\partial o_{pj}}{\partial a_{pj}} = -\frac{\partial E_p}{\partial o_{pj}} f'\left(a_{pj}\right) \\ \frac{\partial E_p}{\partial o_{pj}} &\sum_{k} \frac{\partial E_p}{\partial a_{pk}} \frac{\partial a_{pk}}{\partial o_{pj}} = \sum_{k} \partial_{pk} w_{kj}, \end{aligned} \tag{6}$$

where k represents the previous layer unit connected to the output of unit j, namely,

$$\delta_{pj} = f'\left(a_{pj}\right) \sum_{k} \delta_{pk} w_{kj}. \tag{7}$$

It reflects the error correction of the hidden layer unit 8. It reflects the error correction of the hidden layer unit by weighted summation of all connected to the output of unit j. It reflects the error correction of the hidden layer unit.

(1) Establish a network model, initialize the network, and learn parameters.

(2) Train the network by learning the training samples until the learning requirements are met.

(3) In the forward propagation process, for the input of the given training mode, calculate the output value of the network and compare it with the expected value. If the error cannot meet the accuracy requirements, the error is propagated back; otherwise, go to (2).

(4) Backpropagation process.

The learning process of BP neural network is essentially an iterative process, which consists of forward propagation and backpropagation [20]. At this time, the characteristic data of the samples to be learned are input into the trained network, and then, the network can automatically "learn" the attributes of the samples according to memory as shown in Figure 3.

3.2. Dimensional Analysis and Data Preparation of the Impact of College Students' Employment and Entrepreneurship Development Direction. Refer to the six categories of dimensions developed by the "College Student Evaluation Network" developed by the Education Management Information Center of the Ministry of Education for the comprehensive quality evaluation system of college students: academic quality evaluation, moral quality evaluation, literary and sports quality evaluation, psychological quality evaluation, innovation quality evaluation, and ability. The quality assessment and the dimension of the employment and entrepreneurship development direction of college students adopted by Yuan Jiaqian comprehensively determine the impact dimension of this paper, as shown in Table 1.

A total of 853 university students were included as research objects and their data from 2010 to 2018, and according to the six types of employment and entrepreneurship development directions previously mentioned, the comprehensive data and personal technical information data of 853 students in the school for four years were collected as the data samples of this experiment.

The data sources of this paper are the basic information database of students, the comprehensive database of four-year academic performance in school, the student education management database, and the graduate employment information database.

With the help of teachers and counselors from the relevant departments of the college, the relevant data of 853 students from 2010 to 2018 were successfully obtained, including the student's personal information form, academic record form, school activity bonus form, award evaluation publicity form, and employment information form.

(1) *Data Cleaning.* Because these data come from different management systems of the school, the initially obtained data is very messy, there are many duplicate or invalid attributes, and not all attributes in the student data table are valuable for this

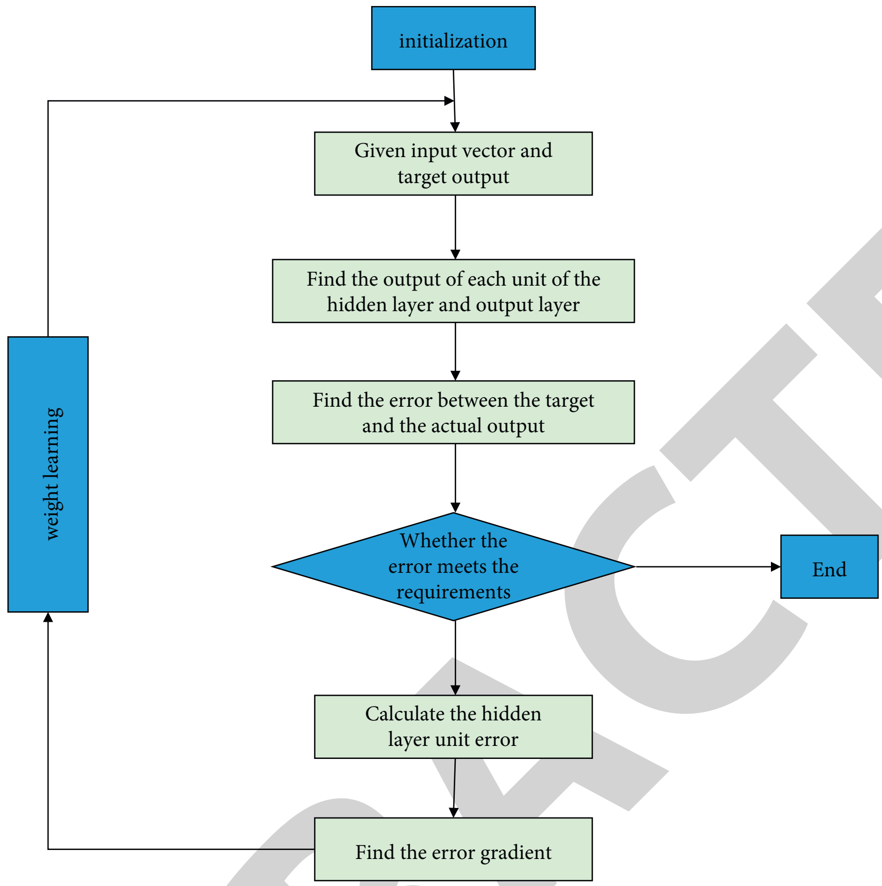


FIGURE 3: Neural network calculation flow chart.

TABLE 1: The dimension of the influence of the development direction of college students' employment and entrepreneurship.

Data name	The data shows
Academic performance	The cumulative academic performance of each subject in four years of college students
English language proficiency	It is divided into three types: pass GET4, pass GET6, and fail the GET exam
Political civilization	Including students' political status, group day activities during school, advanced class evaluation activities, and classmates' mutual evaluation index
Technological innovation	Including various scientific and technological competitions, academic seminars and innovation and entrepreneurship activities participated in during the school
Stylistic arts	Including various cultural activities and sports competitions participated in during school
Social volunteering	Including various school-organized social welfare contribution activities and personal social practice and volunteer services during school

experiment, so the obtained data must be further analyzed, including the deletion of the grades of absent students, the grades of students who did not enroll in courses, and the information of students who were suspended from school and dropped out.

(2) *Data Integration.* The scattered student information is statistically integrated into the annual data summary table of different classes of students, which contains all the impact dimension data of a certain class of students in a certain year. Some student information is listed, as shown in Table 2.

By reviewing the literature and sorting out the influencing factors of college students' employment and entrepreneurship development direction used in various research at present, the influencing dimensions of college students' employment and entrepreneurship development direction adopted in this experimental model are determined. On this basis, the source of the experimental database is determined, and the data is collected. The method proposed in this paper is to integrate, clean up, and convert the collected comprehensive data of students to form a standardized dataset, so as to lay a solid data foundation for subsequent research.

4. Result Analysis and Discussion

4.1. Implementation and Evaluation of the Prediction Model Algorithm for the Development Direction of College Students'

Table 2: Annual data information table of a certain class of students.

Serial number	Academic credits	English language proficiency	Political civilization	Technological innovation	Stylistic arts	Social volunteering	Graduation destination
001	1448	0	7.7958	0.9	2.42	0.3	Employment
002	1375	0.2	5.3765	0	0.3	0	Ascending
003	1495	0.4	8.226	7.98	3.5	0.34	Employment
004	1703	0.2	6.9803	0	1.86	0	Employment
005	1711	0.2	5.0884	0	0.5	0	Employment
006	1632	0.2	7.0121	7.42	0.3	0.15	Employment
007	1523	0.2	4.545	4.06	0.32	0	Ascending
008	1378	0.2	6.0566	0	3.92	0	Employment
009	1477	0.4	4.9982	0.2	0.35	0	Not employed
010	1743	0.2	8.6781	4.86	2.9	0.3	Employment

Employment and Entrepreneurship. Through the analysis and determination of the impact dimension of the development direction of college students' employment and entrepreneurship development and related data preparation, a normative dataset that can be used in the experiment is formed. This chapter will focus on the algorithm design ideas and processes of the prediction model for the development direction of college students' employment and entrepreneurship, the basis for selecting the evaluation indicators of the prediction model, the implementation of the basic prediction model and its evaluation, the optimization model, and the optimization model evaluation process, and the algorithm is proved through experiments. The prediction effect is good and has certain application value.

This prediction model evaluation method adopts the comprehensive evaluation of the Acc evaluation model and the loss function. Before introducing the features and benefits of the Acc evaluation model, it is important to understand the ROC curve. The output result of the two-class classifier is determined by the output probability and a predetermined probability threshold. The most used threshold is 0.5; that is, a positive sample is greater than 0.5, and a negative sample is less than 0.5. If the threshold is increased, the probability of a wrong prediction will decrease, but the probability of a correct prediction will also decrease. If the threshold is decreased, the probability of a correct prediction will increase, and the probability of a wrong prediction will also decrease and then rise. In fact, the selection of this threshold also reflects the classification ability of the classifier to a certain extent. The ROC curve is the curve used to measure the classification ability of the classifier.

As shown in Figure 4, the horizontal axis of the ROC curve is false positive rate (FPR), also called specificity, and the vertical axis is true positive rate (TPR), also called recall rate. The simplest explanation of FPR is the probability of predicting negative samples as positive. Obviously, the smaller the FPR value, the better the prediction effect. TPR is the probability of predicting positive samples as positive; the larger the TPR value, the better the prediction effect. It can be seen from the example of the ROC curve that its abscissa and ordinate are between [0, 1], and then, the area of the ROC curve is between 0 and 1.

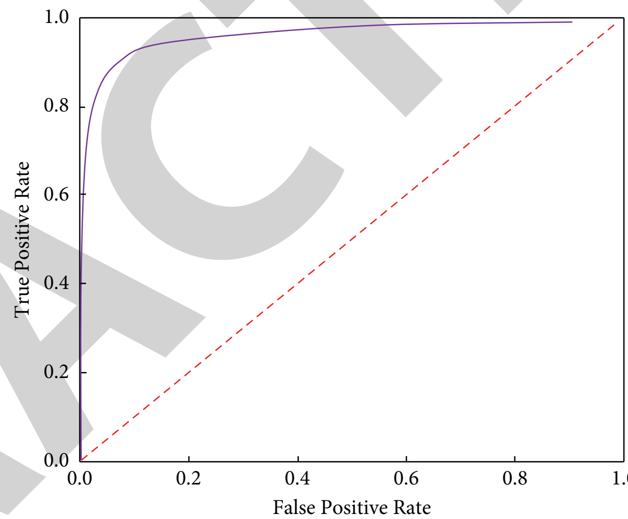


Figure 4: ROC curve example.

To better illustrate the nature of the ROC curve, we should explain the four special points and diagonal lines on the curve: (0, 0) point represents that both FPR and TPR are 0, indicating that the classifier predicts all samples to be negative, and the prediction result is meaningless; (0, 1) point represents that FPR is 0 and TPR is 1, indicating that the classifier predicts all samples correctly, and the effect is the best; (1, 0) point represents that FPR is 1 and TPR is 0, indicating that the classifier predicts all samples incorrectly, and the effect is the worst; (1, 1) point represents that both FPR and TPR are 1, indicating that the classifier predicts all samples to be positive, and the prediction results are meaningless.

The diagonal line is TPR = FPR, which means that the accuracy of the prediction result is 50%, and the prediction effect is equivalent to random classification.

Based on the previously mentioned explanations, the following conclusions are drawn: the closer the ROC curve to the upper left corner, the better the prediction accuracy of the prediction model.

The ROC curve can reflect the effect of the classifier to a certain extent, but it is not clear and intuitive enough. The Acc curve of the Acc evaluation model makes up for the deficiency of the ROC curve. As shown in Figure 5, the value of Acc is the ROC. The size of the area enclosed by the curve

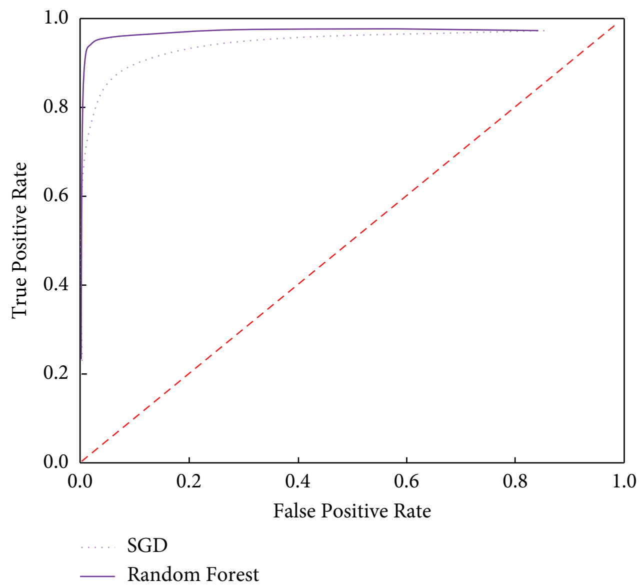


FIGURE 5: Example of Acc curve.

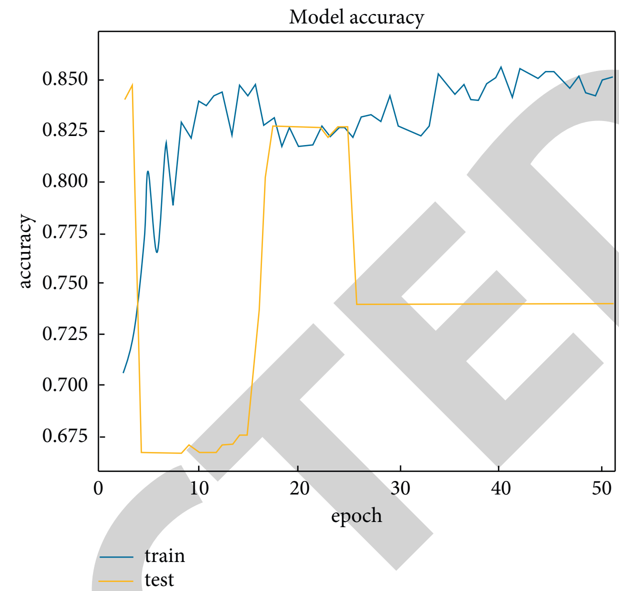


FIGURE 6: Acc value curve of prediction model.

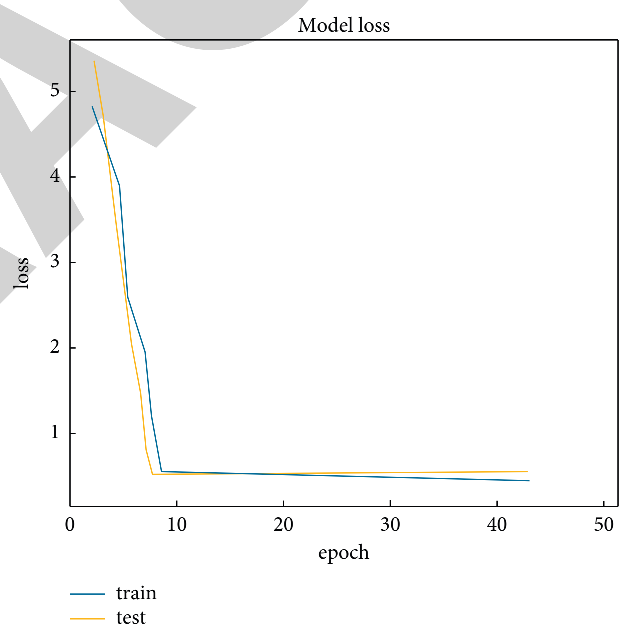


FIGURE 7: Loss value curve of prediction model.

and the X-axis, and the value can intuitively reflect the prediction effect; that is, the larger the value, the better the prediction effect, and the smaller the value, the worse the prediction effect.

The Acc value of the prediction model is shown in Figure 6. The Acc value of the test set is stable at around 0.745, and the Acc value of the training set is stable at around 0.855.

The loss value of the prediction model is shown in Figure 7. The loss value of the test set is stabilized from 4.8 to about 0.4, and the loss value of the training set is stabilized from 4.4 to about 0.3.

By comprehensively considering the Acc value and loss value of the prediction model, it can be proved that the prediction model has a good prediction effect, and further in-depth research and optimization experiments can be carried out.

4.2. Prediction Model Optimization. To further optimize the prediction model and make the prediction model have better prediction effect, the algorithm used in the experiment is further studied in this paper.

The current prediction model has established a sequential model including one layer of network and six types of eigenvalue labels. Next, the basic model will be further optimized from the following three aspects:

(1) Add a network layer to the base model. From the original one-layer network structure to a three-layer network structure, by increasing the network level and enabling each layer to cooperate, the complexity of a single network layer can be effectively reduced, thereby making the prediction results better.

(2) Increase the SGD optimizer. The main algorithm for neural network training is gradient descent, and the most common optimization method for gradient descent is to add an optimizer to the model. The optimizer selected in the experiment in this paper is the second-order convergence, fast convergence speed, and good optimization performance optimizer SGD. SGD (stochastic gradient descent) refers to microbatch gradient descent. The minibatch gradient value of each iteration is calculated, and then, the parameters are updated. Since the parameters are calculated and updated each iteration, each sample may contain a lot of noise. So, the calculated gradient noise is large, which will make the training process oscillate violently. However, it is precisely because of this oscillation that after the local optimization is achieved, the oscillation will

Table 3: Prediction result table.

Serial number	Academic credits	English language proficiency	Political civilization	Technological innovation	Stylistic arts	Social volunteering	Predict the direction of career development	Actual graduation destination	Consistent situation
001	1448	0	7.7958	0.9	2.42	0.3	Employment	Employment	Yes
002	1375	0.2	5.3765	0	0.3	0	Ascending	Ascending	Yes
003	1495	0.4	8.226	7.98	3.5	0.34	Employment	Employment	Yes
004	1703	0.2	6.9803	0	1.86	0	Ascending	Employment	No
005	1711	0.2	5.0884	0	0.5	0	Employment	Employment	Yes
006	1632	0.2	7.0121	7.42	0.3	0.15	Employment	Employment	Yes
007	1523	0.2	4.545	4.06	0.32	0	Ascending	Ascending	Yes
008	1378	0.2	6.0566	0	3.92	0	Employment	Employment	Yes
009	1477	0.4	4.9982	0.2	0.35	0	Employment	Not employed	No
010	1743	0.2	8.6781	4.86	2.9	0.3	Employment	Employment	Yes

Table 4: Individual prediction results of employment and entrepreneurship development direction.

Career development direction	Number of predicted samples	Actual number of samples	Single prediction accuracy (%)
Employment	411	391	94.9
Study	242	262	92.4
Go abroad	48	55	87.3
Freelance	29	34	85.3
Not employed	41	35	82.8
Undergraduate	11	5	52.1

enter the next better point, so the optimization effect is very good.

The Acc evaluation model and the loss function are used again to evaluate the prediction effect of the optimized model. Optimize the Acc value of the model. The Acc values of the test set and training set are both stable at around 0.835. Optimize the loss value of the model: the loss value of the test set is stable from close to 2.75 to around 0.35, and the loss value of the training set is stable from 23.8 to around 0.3. Compared with the basic model, the Acc value and loss value of the comprehensive optimization model are improved to a certain extent, so the optimization model is effective and the prediction effect is better.

4.3. Applications of Predictive Models. Import the existing student database into the prediction model and make predictions and compare and analyze the existing data samples according to the prediction results of the prediction model. There are a total of 835 pieces of student data, and 782 pieces of valid data are obtained after processing. Among them, some students' employment and entrepreneurship development direction prediction results are shown in Table 3.

After comparing all the prediction results with the actual employment situation of students, it is concluded that there are 663 correct data for the overall data sample prediction results, the overall prediction accuracy is 84.73%, and the error rate is 15.27%. The individual prediction data of various types of employment and entrepreneurship development directions are counted, and the individual prediction results are shown in Table 4. The accuracy rate of individual result prediction is calculated by the ratio of the correct predictions to the total number.

Based on the overall experimental data, the best one-way prediction effect of this prediction model is "employment", and "advanced education" ranks second. The one-way forecasting effect is the worst for "not graduated," followed by "not employed." In this regard, this paper makes the following analysis.

(1) Regarding the quality of the students, because the students are sampled from colleges and universities directly under the Ministry of Education of the People's Republic of China, which are world-class discipline construction, they are excellent in both personal quality and learning literacy, and their majors are currently popular for employment. So, it is reasonable for most students to find employment or further studies after graduation.

(2) Regarding the interference by families, the future employment and entrepreneurship development of college students is not only a matter of personal choice but also related to parents' wishes and family factors. A small number of students will follow the wishes of their parents or change their employment and entrepreneurship plans according to their own family conditions.

(3) The amount of data for individual forecast items is small. Since the algorithm used in this experiment is a convolutional neural network algorithm, the characteristic of CNN is that the larger the training set, the better the effect. In this experiment, for

example, there are only five people in "ungraded," which only accounts for 0.64% of the effective data volume. Therefore, one-way training does not work well.

(4) The prediction algorithm and model have not yet reached the ideal limit, and there is still room for optimization and improvement.

5. Conclusion

Based on the idea of convolutional neural network, this paper included 853 university students as research objects and their data from 2010 to 2018. We collected the comprehensive data and employment data of 853 students in the school for four years as data samples. We inputted six influencing factors, including academic performance, English proficiency, political civilization, technological innovation, sports and art, and social volunteer service, and set six categories of employment, including further education, going abroad, freelance entrepreneurship, unemployed, ungraduated, and employment and entrepreneurship development direction, built a prediction model for the employment and entrepreneurship development direction of college students, provided more scientific and effective decision-making basis for employment and entrepreneurship development direction for colleges and universities, and helped college students to correctly plan the employment and entrepreneurship development direction. The experimental results in this paper prove the validity of the prediction results, and to a certain extent, verify the user-friendly, modular, and easy-to-expand characteristics of the Keras framework interface. At the same time, it also proves that the sequential algorithm has convolutional properties in multilabel classification problems. This method has the advantages of high speed and high precision. The predicted student development direction can be further refined. The five types of college students' employment and entrepreneurship development directions set in this experiment include employment, further education, going abroad, freelance employment, and entrepreneurship. In future research, various development directions can be refined for more in-depth research. For example, it can predict the type of enterprise unit where students are employed, the reasons for the grades of schools where students study, and so on. Through further detailed research, the prediction function of this prediction model is increased, and the overall application degree of the model is improved.

Data Availability

The labeled datasets used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] M. Wise and O. Christine, "Henderson Amanda the contribution of paid employment for students during their preregistration nursing studies: an integrative review," *Nurse Education Today*, vol. 111, pp. 102–110, 2022.

[2] G. A. Orona, Q. Li, P. C. Bartek, and D. au, "What predicts the use of interaction-oriented pedagogies? The role of self-efficacy, motivation, and employment stability," *Computers & Education*, vol. 184, no. 12, p. 104498, 2022.

[3] M. Casciato, "Post-heart transplant employment: rates, predictors, and interventions," *The Journal of Heart and Lung Transplantation*, vol. 41, no. 4, pp. S198–S199, 2022.

[4] J. E. Bates, A. Parekh, C. R. Goodman et al., "The geography of employment outcomes for radiation oncology graduates in 2019," *International Journal of Radiation Oncology, Biology, Physics*, vol. 45, pp. 78–96, 2020.

[5] M. Joanne, "Juggling employment and studies: nursing students' perceptions of the influence of paid employment on their success," *Nurse Education Today*, vol. 92, pp. 32–45, 2020.

[6] M. Sidebotham, C. Walters, K. Baird, and J. Gamble, "Simulated employment interviews: a collaborative approach to gaining understanding of the graduate midwife employment process," *Women and Birth*, vol. 33, no. 5, pp. 455–463, 2020.

[7] P. Jung and C. Smith, "A mismatch between training preferences and employment opportunities," *Lancet Global Health*, vol. 7, pp. 287–294, 2019.

[8] A. R. Sattler and M. K. Amatayakul, "Results from the 1989 member employment survey," *Journal of*, vol. 60, no. 11, pp. 3641–4118, 2018.

[9] M. Samuel, F. Adaba, and A. Maeda, "Home parenteral nutrition and employment in patients with intestinal failure: factors associated with return to employment," *Clinical Nutrition*, vol. 38, no. 3, pp. 1211–1214, 2019.

[10] A. Warner, J. Barrow, and J. A. Williams, "The relationship among BSN students' employment, educational, and health-related characteristics and semester grades: a multi-site study," *Journal of Professional Nursing*, vol. 36, no. 5, pp. 308–316, 2020.

[11] I. M. Geisner, J. Koopmann, P. Bamberger, and P. M. E. Wang, "When the party continues: impulsivity and the effect of employment on young adults' post-college alcohol use," *Addictive Behaviors*, vol. 77, pp. 114–120, 2018.

[12] U. Bejerholm and M. E. Larsson, "Suzanne Johanson Supported employment adapted for people with affective disorders—a randomized controlled trial," *Journal of Affective Disorders*, vol. 207, pp. 627–642, 2017.

[13] C. Dillahunt-Aspillaga, M. J. Pugh, B. A. Cotner et al., "Employment stability in veterans and service members with traumatic brain injury: a veterans administration traumatic brain injury model systems study," *Archives of Physical Medicine and Rehabilitation*, vol. 99, no. 2, pp. 940–943, 2018.

[14] T. Boonekamp and Z. Joost, "Guillaume Burghouwt Determinants of air travel demand: the role of low-cost carriers, ethnic links and aviation-dependent employment," *Transportation Research Part A*, vol. 112, pp. 257–265, 2018.

[15] B. Alcott, "Should degrowth embrace the job guarantee?" *Journal of Cleaner Production*, vol. 38, pp. 675–684, 2019.

[16] K. Siyan and N. Elena, "Does an alignment of employment policies and individual preferences affect intention to stay in the profession? Evidence from Canadian Midwives," *Health Policy*, vol. 125, no. 4, pp. 383–390, 2021.