Mobile Information Systems

Big Data-Driven Mobile IoT Intelligence

Lead Guest Editor: Rong Mo
Guest Editors: Jie Liu and Dragan Pamučar

Big Data-Driven Mobile IoT Intelligence

Big Data-Driven Mobile IoT Intelligence

Lead Guest Editor: Rong Mo
Guest Editors: Jie Liu and Dragan Pamučar

Contents

Contents

Contents

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9898413, 1 page
https://doi.org/10.1155/2023/9898413

Retraction

Retracted: The Relationship between Machine Translation and Human Translation under the Influence of Artificial Intelligence Machine Translation

Mobile Information Systems

Received 13 September 2023; Accepted 13 September 2023; Published 14 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This article has been retracted by Hindawi following an investigation undertaken by the publisher [1].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] Z. Lihua, "The Relationship between Machine Translation and Human Translation under the Influence of Artificial Intelligence Machine Translation," *Mobile Information Systems*, vol. 2022, Article ID 9121636, 8 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9890363, 1 page
https://doi.org/10.1155/2023/9890363

Retraction

Retracted: Study on Employee Performance Evaluation Based on Adaptive Feature Selection Fuzzy Algorithm

Mobile Information Systems

Received 13 September 2023; Accepted 13 September 2023; Published 14 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] Y. Qian and J. Yin, "Study on Employee Performance Evaluation Based on Adaptive Feature Selection Fuzzy Algorithm," *Mobile Information Systems*, vol. 2022, Article ID 6985508, 13 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9861498, 1 page
https://doi.org/10.1155/2023/9861498

Retraction

Retracted: Deep Learning-Driven Financial Management Innovation Upgrade for Universities

Mobile Information Systems

Received 13 September 2023; Accepted 13 September 2023; Published 14 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] Y. Guan and C. Chen, "Deep Learning-Driven Financial Management Innovation Upgrade for Universities," *Mobile Information Systems*, vol. 2022, Article ID 9328712, 8 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9832807, 1 page
https://doi.org/10.1155/2023/9832807

Retraction

Retracted: The Practical Dilemma and Path Optimization of Big Data Integration into Ideological and Political Teaching in Private Colleges and Universities

Mobile Information Systems

Received 13 September 2023; Accepted 13 September 2023; Published 14 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] X. Li, "The Practical Dilemma and Path Optimization of Big Data Integration into Ideological and Political Teaching in Private Colleges and Universities," *Mobile Information Systems*, vol. 2022, Article ID 7417406, 10 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9824921, 1 page
https://doi.org/10.1155/2023/9824921

Retraction

Retracted: Construction of an Inquiry-Based Teaching Model for Ideological and Political Education in Colleges and Universities from the Perspective of Deep Learning

Mobile Information Systems

Received 13 September 2023; Accepted 13 September 2023; Published 14 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.

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] B. An and L. Gao, "Construction of an Inquiry-Based Teaching Model for Ideological and Political Education in Colleges and Universities from the Perspective of Deep Learning," *Mobile Information Systems*, vol. 2022, Article ID 9286979, 12 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9813216, 1 page
https://doi.org/10.1155/2023/9813216

Retraction

Retracted: DA-CNN-Driven Innovative Ideological Politics Education Management System

Mobile Information Systems

Received 13 September 2023; Accepted 13 September 2023; Published 14 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] S. Yingkun, C. Liang, and C. Jing, "DA-CNN-Driven Innovative Ideological Politics Education Management System," *Mobile Information Systems*, vol. 2022, Article ID 7959144, 9 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9787416, 1 page
https://doi.org/10.1155/2023/9787416

Retraction

Retracted: Quality Evaluation of College Employment Based on Fuzzy Comprehensive Evaluation and Immersive Virtual Realization Technology

Mobile Information Systems

Received 13 September 2023; Accepted 13 September 2023; Published 14 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.

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] Y. Ji and Y. Luo, "Quality Evaluation of College Employment Based on Fuzzy Comprehensive Evaluation and Immersive Virtual Realization Technology," *Mobile Information Systems*, vol. 2022, Article ID 6285591, 13 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9782353, 1 page
https://doi.org/10.1155/2023/9782353

Retraction

Retracted: The Implementation Path of Labor Education in Applied Universities Driven by Artificial Intelligence Technology

Mobile Information Systems

Received 13 September 2023; Accepted 13 September 2023; Published 14 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] Y. Luo, "The Implementation Path of Labor Education in Applied Universities Driven by Artificial Intelligence Technology," *Mobile Information Systems*, vol. 2022, Article ID 5375449, 12 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9782038, 1 page
https://doi.org/10.1155/2023/9782038

Retraction

Retracted: Artificial Intelligence Technology and Its Application in Improving Thought-Politics Education

Mobile Information Systems

Received 13 September 2023; Accepted 13 September 2023; Published 14 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] X. Bai and X. Wang, "Artificial Intelligence Technology and Its Application in Improving Thought-Politics Education," *Mobile Information Systems*, vol. 2022, Article ID 3150352, 11 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9876931, 1 page
https://doi.org/10.1155/2023/9876931

Retraction

Retracted: Analysis of the Dilemmas and Countermeasures Brought by Data Analysis Based on Short Online Videos on Civic and Political Work

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. Xie, "Analysis of the Dilemmas and Countermeasures Brought by Data Analysis Based on Short Online Videos on Civic and Political Work," *Mobile Information Systems*, vol. 2022, Article ID 3367257, 10 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9875103, 1 page
https://doi.org/10.1155/2023/9875103

Retraction

Retracted: The Modal Analysis of Multifactor Coupling of Regional Industrial Innovation

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] Y. Liu and Q. Guo, "The Modal Analysis of Multifactor Coupling of Regional Industrial Innovation," *Mobile Information Systems*, vol. 2022, Article ID 1261278, 12 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9898337, 9 pages
https://doi.org/10.1155/2023/9898337

Research Article

Design of Personalized News Recommendation System Based on an Improved User Collaborative Filtering Algorithm

Xuejiao Wang [1] and Chao Liu [2]

[1] College of Art and Design, Xi'an Mingde Institute of Technology, Xi'an 710124, Shaanxi, China
[2] School of International Economics and Trade, Universidad Complutense Madrid, Madrid 28040, Spain

Correspondence should be addressed to Xuejiao Wang; wangxj@mdit.edu.cn

Received 31 August 2022; Revised 27 September 2022; Accepted 1 October 2022; Published 22 February 2023

Academic Editor: R. Mo

To solve the problem of information overload in the field of news, this paper designs and implements a feasible news recommendation system, where the front-end web page is made by Django framework, whose performance is optimized by bootstrap and jquery, while in the back-end design, the original user similarity calculation method is improved by adding the time attenuation factor, and a news recommendation model based on user collaborative filtering (CF) algorithm is proposed. Experimental results show that the proposed algorithm achieves highest recall, accuracy, and *F*1 score ratio compared with other algorithms, which indicate that the proposed algorithm has better performance.

1. Introduction

With the advent of the information age, the Internet provides a convenient way to obtain news. However, there are millions of news reports from different channels and fields on the Internet, which makes users lost in massive data, so that they need to spend a lot of time and energy to identify the news they want to browse, that is, "information overload" problem [1, 2]. Recommendation system is a kind of information filtering system, which can predict the user's preference for items according to the analysis of users' interest characteristics and historical behavior data, so as to help users make decision and analysis, which can improve the users' experience.

In the application of news field, the huge amount of news produced makes it extremely important to recommend news efficiently and quickly. At present, many websites generally adopt popular recommendation and click through rate recommendation to deal with massive news, but these methods often do not consider the interest difference between different users, and all the recommendation results are the same, which cannot really meet the actual needs of users [3, 4]. On the one hand, the current popular recommendation system can reduce users' time to find news and improve their browsing experience. On the other hand, this system provides the output efficiency of news platform, so that more corresponding news can be timely browsed by users, which improves the utilization value of news and provides great commercial value for news platform companies.

Among them, news recommendation is a hot research field. By calculating the similarity of articles and considering topics, categories, etc., the content-based news recommendation system will generate a news list for users, which is similar to the news users as in [5]. In addition, users in the news field are more likely to be affected by popular items. At the same time, due to the timeliness of news recommendation, user interest is constantly evolving [6]. The collaborative filtering (CF) algorithm is useful for personalized news recommendation (PNC). For news recommendation based on CF, item CF is generally not adopted because the number of news is far greater than the number of users, and the update is fast, resulting in high computational complexity [7]. Therefore, a user-based CF algorithm is adopted in this paper, the basic principle of which is to find "neighbor" user groups similar to the current user's preferences (ratings) according to all users' preferences (scores) for items or information.

The purpose of this paper is to develop a solution for the news recommendation and try to make this solution have strong compatibility and applicability. Through the research of hybrid recommendation technology, a feasible news recommendation system is designed and implemented. In the algorithm design, the existing similarity calculation method does not consider the time factor. Therefore, this paper considers the time factor and improves the original user similarity calculation formula; in addition, the time decay factor is considered.

2. Research Status of News Recommendation Algorithm

The principle of CF-based recommendation algorithm is to assume that users who like similar items have similar interests [8]. It first generates a set of users with similar interests for the target users and then recommends the items that the target users have not clicked on but the users in the collection have clicked on. Because the principle of recommendation algorithm based on CF is relatively simple, it has been popular in the field of recommendation. Wu et al. [9] proposed a cooperative noise reduction coder to recommend to users, where the automatic noise reduction encoder formulates the user project feedback data to generate a distributed structure of users and projects. Han [10] improved the two stages of data preprocessing and nearest neighbor selection in the CF algorithm, filled the user project evaluation matrix, introduced the tag factor and time factor, integrated CF with dichotomy K-means, and improved the similarity calculation formula. Saranya and Sadasivam [11] adopted a CF algorithm based on rough sets to score news categories, which improve the ranking of novel news. Jiang et al. [12] combined the knowledge graph with the CF algorithm, connected web API and mashup related information through the knowledge graph, and then calculated the distance between web API vectors to make recommendations. Xue et al. [13] used the matrix decomposition model to construct user news matrix by displaying rating and implicit feedback information of users and projected users and news into low-dimensional space through neural network.

Although the recommendation algorithm based on CF can obtain satisfactory recommendation results in some cases, it often has the problem of cold start. With the development of neural network, it has achieved good results in image processing, natural language processing, and other fields. Because the traditional recommendation model is a general algorithm model, it does not combine with specific business scenarios, so it is not effective in some specific application scenarios. In this case, the recommendation algorithm needs to make some changes according to different applications, so there is a recommendation algorithm model for a certain application scenario. Veličković et al. [14] proposed a novel attention mechanism network, which assigns different weights to different nodes in the nearest neighbor through the self-attention mechanism layer. These recommended algorithm models for a certain field need to be combined with the actual business scenarios and need to design the models from the perspective of business and consider the actual application effect of the models. Neural network can find the hidden feature information in user behavior records and capture the interaction characteristics between the project and the user, which improves the accuracy of the recommendation algorithm.

3. Design of PNC System

3.1. Overall Structure. The architecture design of the system is shown in Figure 1. The whole system architecture design is divided into the following parts.

The data layer processes and stores user information, news information, and log file information, respectively. Pandas data processing framework is used in the process, and MySQL and Redis are used in the storage. There is user's historical information in the user's news platform. In addition, the historical browsing behavior of users is the ID of news, which needs to be mapped to the corresponding news content. In this system, the news refers to the news title. Log files record all operations in the system and are saved in the MySQL database.

The strategy layer includes recall strategy, sorting strategy, cold start strategy, and reordering strategy. This layer is to model the user according to the news content of user behavior sequence, so as to recommend the interested news content to the user according to the user interest of the news platform. The process steps of this layer are as follows:

(1) The candidate news is recalled in the news dataset according to the news in the user behavior. In this step, several kinds of candidate news are recalled in the mass news dataset.

(2) In the news dataset, the recommended range of news is reduced, and then the news ranking algorithm mentioned above is used to further sort the candidate news to generate the news recommendation list.

(3) When there is no history browsing record in user behavior sequence, there is no sorting strategy. In this case, we need to use cold start strategy to recommend to users.

(4) After sorting, the news needs to be reordered from different angles, and the final sorting results are returned to the user.

In the strategy layer, the model code is edited and trained through the Python framework.

The feedback layer includes feedback evaluation and recommendation list. This layer is to evaluate and adjust the results of the strategy layer. After the strategy returns the recommendation result, it evaluates the recommendation result according to the evaluation indicators such as accuracy, freshness, and popularity. When the strategy does not meet the requirements of the index, it needs to adjust the strategy again. In the evaluation process, SKlearn framework is used for evaluation.

The application layer provides interactive support for users and administrators, including user registration and

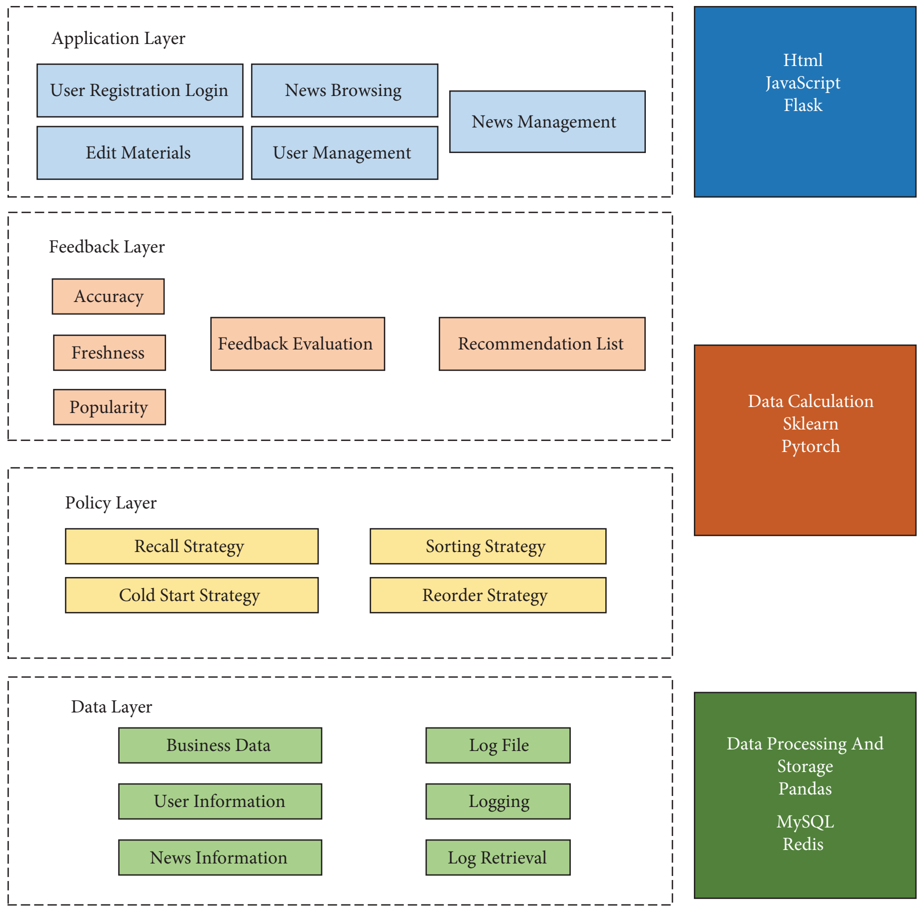


FIGURE 1: Overall system architecture.

login, news browsing, editing materials, user management, and news management. The whole framework uses Flask, HTML, and JavaScript to participate in the interaction process between the front end and back end.

3.2. Function Module. The system function module is divided into five parts, registration and login module, editing data module, news browsing module, user management module, and news management module. The system function module is shown in Figure 2. The registration and login module includes user registration and user login. The editing data module includes nickname modification, avatar modification, and password modification. News browsing module includes news browsing, category news browsing, historical news browsing, and recommended news browsing. The user management module includes user search and user deletion. News management module includes news release, news editing, and news deletion.

4. Realization of PNC System

The recommendation system designed in this paper is mainly presented in the form of web pages. Since the back-end recommendation algorithm is designed by Python, in order to facilitate the interaction between the back end and the front end, the front-end web page is planned to use Django framework, and bootstrap and jquery are used to optimize the web page effect. As a mainstream and popular web framework, Django uses a program structure similar to MVC, using MVT (Model, View, and Template) architecture. The full name of MVC is Model View Controller. M refers to Model, which mainly encapsulates the access to the database layer and adds, deletes, modifies, and checks the data in the database. V refers to View, which is used to encapsulate the results and generate the HTML content displayed on the page. C refers to Controller, which is used to receive requests, process business logic, interact with Model and View, and return results. MVT is the proper name of Django framework, and its function is the same as that of MVC. In MVT, Model is mainly used to define and interact with the database. View is similar to controller in JSP, which is used to control the behavior of the website in the background. Template is used to display the front-end effect of a web page.

4.1. Front-End Functions. The front-end design of website mainly involves the design and development of template. Different from JSP technology, under the Django framework, an HTML web page is used not only as a page but also as a template that can be reused many times.

The realization of template function: when a user logs in to a website of Django framework, the URL controller is used

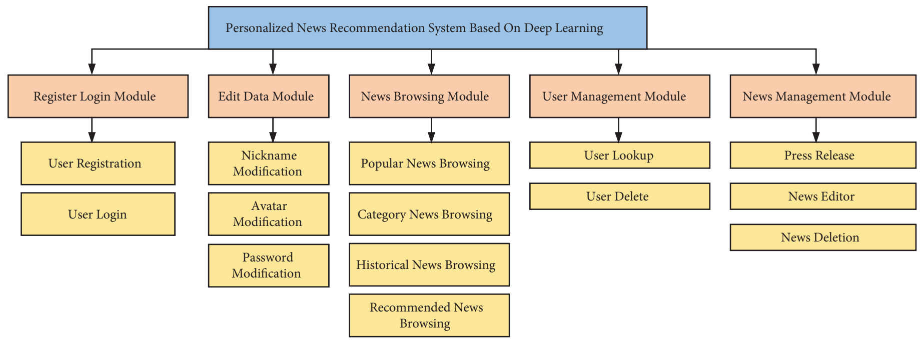


Figure 2: System function module.

for routing distribution to match the appropriate view function, and then the view function will read models or render a template directly and return to the user interface according to the process written by the programmer. Different view functions can render the same template, and the contents of the template will change according to the parameters passed by the view, which can reuse the pages with little change in style and reduce a lot of development costs. The running process of the whole framework is shown in Figure 3.

In addition to the design of templates, the front end also involves the display of back-end data. In this framework, the data are first processed by the controller and then transmitted to the template through the corresponding API. Taking the home page as an example, the parameters of the page include news title, news classification label, and user information. Among them, the news title and news tag are imported into the page as a list, and the user information only needs to pass in a user name, while in the background, the news list will be sorted according to the user's information and then output after processing.

The method of calling parameters in the template is similar to that in JSP technology. Each element in the list is displayed in the page by traversing through the parameters passed in. It is worth mentioning that all the pages of this project use bootstrap to beautify the appearance of web pages. Because bootstrap is a responsive framework, using bootstrap can avoid the problem of display dislocation in different size devices.

4.2. Back-End Functions. Assume that the user named U needs personalized recommendation. First, find the users with similar interests. Those users are called the nearest neighbor users. Then, recommend the news that the nearest neighbor users like but the user U has not seen. This is the CF algorithm based on users. The advantage of this recommendation system is that the recommended news may be completely irrelevant in content, so it can discover the potential interests of users and generate personalized recommendation results for each user.

4.2.1. Algorithm Design. The CF algorithm mainly uses the similarity of behavior to calculate the similarity of interest. Given user u and user v, formula (1) is used to calculate the similarity of interest of u and v, where $N(u)$ represents the

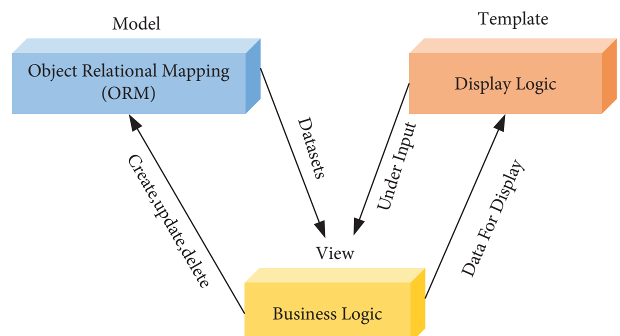


Figure 3: Operation process of MTV structure.

news set that user u has read, and $N(v)$ represents the news set that user v has browsed.

$$W_{uv} = \frac{|N(u) \cap N(v)|}{\sqrt{|N(u)||N(v)|}}. \tag{1}$$

User u has read news $\{a, b, d\}$, and user v has read news $\{a, c\}$. Formula (1) is used to calculate the similarity between user u and user v, and the calculation of W_{uv} can be rewritten by the news sets $\{a, b, d\}$ and $\{a, c\}$, which are shown in the following formula:

$$W_{uv} = \frac{|\{a, b, d\} \cap \{a, c\}|}{\sqrt{|\{a, b, d\}||\{a, c\}|}} = \frac{1}{\sqrt{6}}. \tag{2}$$

4.2.2. Improved Algorithm. Formula (1) only considers whether the user has read a certain news, but not the reading time. Due to the strong timeliness of news recommendation, in the user's reading history, the farther away from the present, the more limited the effect of user's behavior on prediction. According to the characteristics of news recommendation, we adjust the similarity formula. At the same time, the influence of hot news and news reading time on similarity is considered. In order to punish the influence of the popular news in the common interest list of user u and user v on their similarity, the popular news is downgraded to avoid the Harry Potter effect. The calculation formula considers the time attenuation factor to prevent the similarity between user u and user v who like news i from being too small within a certain time range. Formula (1) is improved by considering the characteristics of news recommendation, as shown in the following formula:

$$\frac{\sum_{i\in N(u)\cap N(v)} m^{*}\,(Ta - To/Tc - To) + (1-m)1/1 + N(i)}{\sqrt{|N(u)||N(v)|}}, \tag{3}$$

where $N(u)$ represents the reading set of user u, $N(v)$ represents the reading set of user v, $N(i)$ represents the occurrence times of the news item i, Ta represents the execution time, Tc represents the current time, To represents the initial time, and m represents the adjusting factor, which ranges from 0 to 1. The larger m is, the more factors are considered in the change of interest, while the smaller m is, the more punishment is imposed on popular news.

For the convenience of description, this improved CF algorithm is renamed as UserCF algorithm.

4.2.3. Algorithm Flow. After obtaining the user's similarity, UserCF algorithm recommends k news that users are most interested in but not read by the recommended user. Formula (3) measures user U's interest in news i in UserCF algorithm:

$$P(u,i) = \sum_{v\in S(u,k)\cap N(i)} W_{ub}R_{vi}, \tag{4}$$

where $S(u,k)$ represents the k nearest neighbor sets that are most similar to user u's interest, $N(i)$ represents the user set that has read news i, W_{uv} represents the interest similarity between user u and user v, which can be calculated by formula (3), R_{vi} represents user v's preference degree or score for news, user v has read news i, and R_{vi} is 1. The recommended news is the one that user u has not read before. The algorithm of user u's interest in news i is described as follows. for $((v = 1; v \leq$ all users; $v + 1))$

Formula (3) is used to calculate W_{uv}.

(i) *Step 1.* Sort all W_{uv} to get k nearest neighbor sets $S(u,k)$ with the most similar interests of user u.

(ii) *Step 2.* Find the users $S(u,k)\cap N(i)$ who have read news i from k nearest neighbor users.

(iii) *Step 3.* The similarity degree W_{uv} of each user $v \in S(u,k)\cap N(i)$ in user u and set $S(u,k)\cap N(i)$ is multiplied by the user's interest degree R_{vi} and the sum is the interest degree of user u in news i.

(iv) *Step 4.* According to $P(u,i) = P(u,i) + W_{uv} * R_{ui}$, sort $P(u,i)$ to find the news that k neighbors are most interested in and recommend it to user u.

When a new user registers, the user will select several categories of interest. These categories are displayed dynamically through the word segmentation results of news in the database, which are not fixed. After users choose, the back end will give priority to the categories they are interested in during the process of visiting the website. This recommendation method is mainly to deal with the problem of cold start. Before the recommendation data are formed, tag recommendation is used to deal with the situation that the recommendation system does not have sufficient data for recommendation analysis. Similarly, SQL statements are used to organize the required news objects into a linked list and then render the output. In the page, the news headlines are read in an iterative loop. The recommendation process is shown in Figure 4.

5. Experiment and Analysis

5.1. Dataset. The dataset used in this experiment is from the user data of a domestic news website, including 5000 users and nearly 120000 reading records. Each dataset contains six parts, and the details are shown in Table 1. In this experiment, the dataset is divided into two parts, in which the ratio of training set and test set is 7 : 3. The data information is shown in Table 1:

5.2. Evaluation Index. In this experiment, $F1$ score, precision, and recall were used to evaluate the algorithm proposed in this chapter. $F1$ score is the harmonic mean of accuracy and recall rate. Accuracy refers to the ratio of recommended correct news items to the recommended list, while recall ratio refers to the ratio of recommended correct news items to the list clicked by users. The calculation formulas are as follows:

$$\begin{aligned} F1 &= 2 * \frac{(\text{precision} * \text{recall})}{(\text{precision} + \text{recall})}, \\ \text{precision} &= \frac{\sum_u |R(s)\cap T(s)|}{\sum_u |R(u)|}, \\ \text{recall} &= \frac{\sum_u |R(s)\cap T(s)|}{\sum_u |T(s)|}, \end{aligned} \tag{5}$$

where $R(s)$ represents the recommended list of news titles recommended to user s and $T(s)$ represents the news headlines clicked by user s in the test set.

5.3. Results and Discussion

5.3.1. Parameter Optimization. First of all, in the fusion similarity calculation, through the weight influence factor α, traverse the value, observe the change of precision value, and measure the weight of similarity calculation method. The experimental results are shown in Figure 5.

The change of weight influence factor α can affect the prediction accuracy of news recommendation algorithm. When $\alpha = 0.5$, the news recommendation algorithm that integrates user reading records and news title similarity proposed in the model has the best performance, and when $\alpha = 0.4$, the accuracy is the highest. Therefore, α is selected as 0.4.

5.3.2. Comparison of Different Algorithms. In order to verify the effectiveness of the back-end design of the system, the experimental results are compared and analyzed with other algorithms. In this paper, three algorithms are selected for

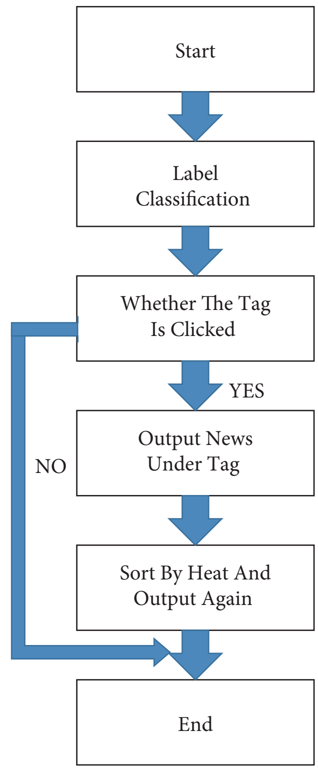


Figure 4: News recommendation process.

Table 1: Dataset information.

Attribute	Details
User_id	ID of the user
Time	When the user clicks
Click	Does the user click
Click_count	Number of user clicks
Category_id	Category of news headlines
Title	News headline text

comparison: PNR, FALS, and NR_LDA. The results are shown in Figures 6–8.

PNR is an improved PNC algorithm based on consumer click behavior [15]. Firstly, association rules between adjacent news are established in the news browsing sequence of consumers, and then the restriction of browsing time difference is added to the construction of association rules to recommend news to users.

Fused ALS (FALS) uses the weighted hybrid recommendation strategy to fuse ALS model recommendation, content-based recommendation, and improved cyclic neural network algorithm and obtains a hybrid algorithm model through logical regression training weight [16].

NR_LDA combines the user's interest and news timeliness [17] and adds the influence of news release time to the topic extraction by using the LDA model to improve the effect of the recommendation algorithm.

As can be seen from the above figures, with the increase of recommendation list N, the accuracy first increases and then decreases. When $N = 10$, the accuracy reaches the highest. Under the same N value, compared with the four algorithms, the proposed algorithm improves by 2.7% compared with PNR, 3.9% compared with NR_LDA, and 12.6% compared with FALS, with the highest accuracy among the four algorithms. With the increase of recommendation list N, the recall rate increases gradually. The recall rate of the proposed algorithm is higher than that of the other three algorithms in the same recommendation list N, indicating that the proposed recommendation algorithm has better performance.

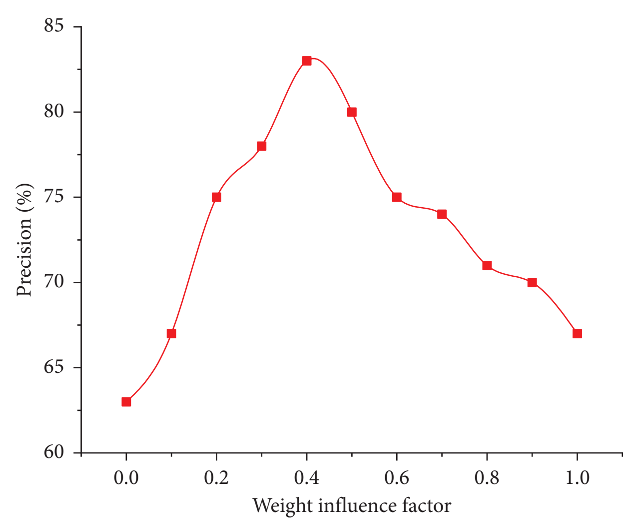


Figure 5: Influence of weight influence factors on algorithm precision.

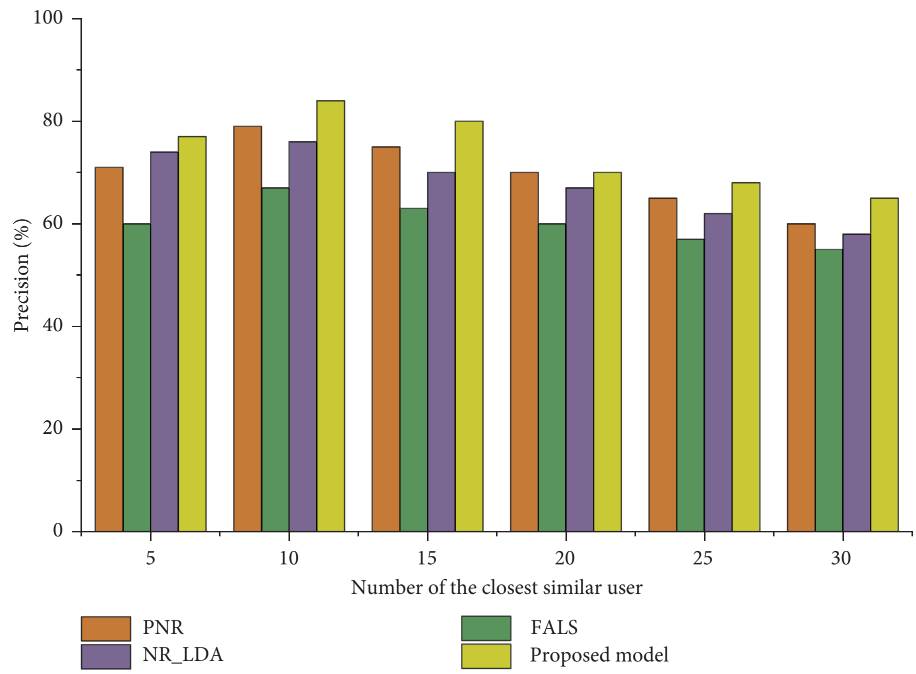


Figure 6: Precision comparison of different models.

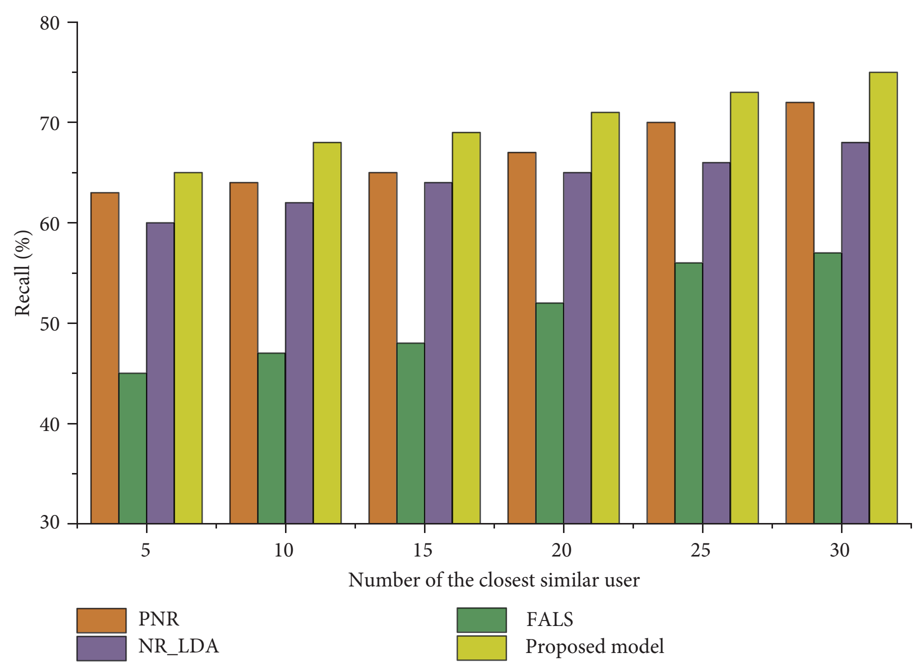


Figure 7: Recall comparison of different models.

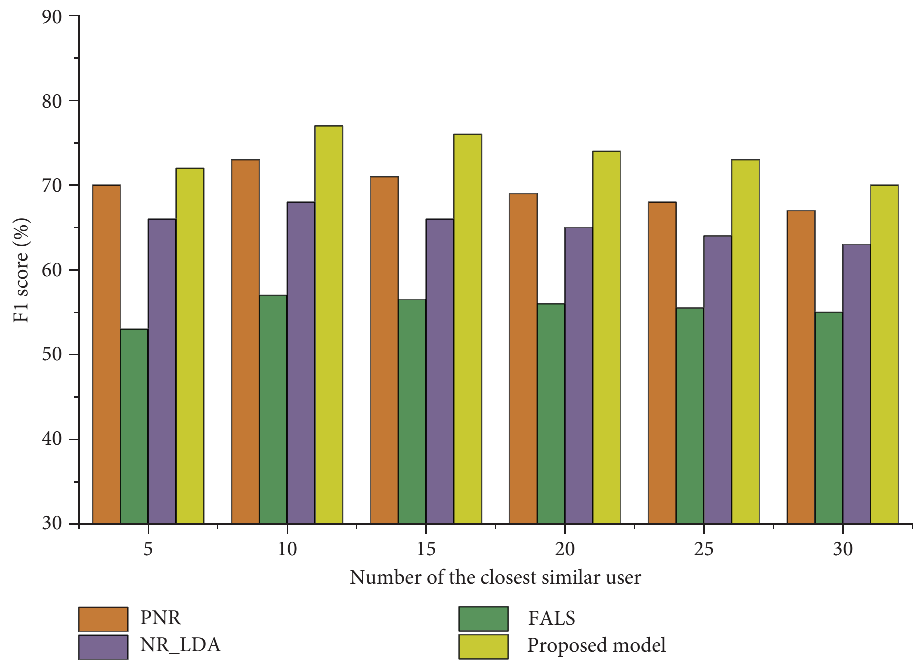


Figure 8: $F1$ score comparison of different models.

To balance recall rate and accuracy, $F1$ score was used as the evaluation standard of this experiment. As shown in Figure 8, with the increase of N, $F1$ score first increases, then decreases, and finally tends to be stable, which reaches its maximum when $N = 10$. In the case of a small number of similar users recently, the number of news recommended by users may not be comprehensive enough. However, if too many recent similar users are selected, some users with low similarity degree will be included, and the news with high popularity clicked by these users that has no interest of target users will be recommended to target users, resulting in reduced accuracy. Therefore, in the news recommendation algorithm, it is necessary to reasonably select similar users as recent similar users to obtain the ideal CF recommendation effect.

6. Conclusion

This paper designs a personalized news recommendation system. The functional requirements of this system need to meet the acquisition and storage of news data, user behavior recording, news content display, etc. Then, the architecture design of PNC system is introduced, including multilayer structure, where the front-end web uses Django framework to reduce development costs and greatly improve development efficiency, while the back-end design adopts the news recommendation model based on the UserCF algorithm. The test results show that the recall, accuracy, and $F1$ score of the proposed algorithm are higher than those of the other three algorithms in the same recommendation list N.

Data Availability

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

Conflicts of Interest

This study was conducted in the absence of any commercial or financial relationship that could be interpreted as potential conflicts of interest.

References

[1] Y. Wu, X. Wang, and Y. X. Ding, "Adaptive on-line web topic detection method for web news recommendation system," *Acta Electronica Sinica*, vol. 38, no. 11, pp. 2620–2624, 2010.

[2] M. Zihayat, A. Ayanso, X. Zhao, H. Davoudi, and A. An, "A utility-based news recommendation system," *Decision Support Systems*, vol. 117, pp. 14–27, 2019.

[3] Y. Peng, *Research and Implementation of PNC Algorithm Based on Improved LDA Topic Model*, Fuyang Normal University, Fuyang, China, 2021, in Chinese.

[4] Y. Wu, M. Qi, and R. Yang, "A news recommendation system based on improved CF algorithm," *Computer engineering and science*, vol. 39, no. 06, pp. 1179–1185, 2017, in Chinese.

[5] Y. Xu, "News recommendation analysis and exploration based on the integration of content-based recommendation and CF," *Guide to news research*, vol. 9, no. 13, p. 111, 2018, in Chinese.

[6] F. Garcin, K. Zhou, and B. Faltings, "PNC based on CF," in *Proceedings of the 2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology*, vol. 1, pp. 437–441, Macau, China, December 2012.

[7] S. Jiang and W. Hong, "A vertical news recommendation system: CCNS—an example from Chinese campus news reading system," in *Proceedings of the International Conference on Computer Science & Education*, pp. 1105–1114, British Columbia, Canada, August 2014.

[8] Z. Cui, X. Xu, F. Xue et al., "Personalized recommendation system based on collaborative filtering for IoT scenarios," *IEEE Transactions on Services Computing*, vol. 13, no. 4, pp. 685–695, 2020.

[9] Y. Wu, C. DuBois, A. X. Zheng, and M. Ester, “Collaborative denoising auto-encoders for top-N recommender systems,” in *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*, pp. 153–162, San Francisco, CA, USA, February 2016.

[10] K. Han, “Personalized news recommendation and simulation based on improved collaborative filtering algorithm,” *Complexity*, vol. 2020, Article ID 8834908, 12 pages, 2020.

[11] K. G. Saranya and G. S. Sadasivam, “Personalized news article recommendation with novelty using CF based rough set theory,” *Mobile Networks and Applications*, vol. 22, no. 4, pp. 1–11, 2017.

[12] B. Jiang, J. Yang, Y. Qin, T. Wang, M. Wang, and W. Pan, “A service recommendation algorithm based on knowledge graph and collaborative filtering,” *IEEE Access*, vol. 9, pp. 50880–50892, 2021.

[13] H. J. Xue, X. Dai, and J. Zhang, “Deep matrix factorization models for recommender systems,” *IJCAI*, vol. 17, pp. 3203–3209, 2017.

[14] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, “Graph attention networks,” 2017, https://arxiv.org/abs/1710.10903.

[15] Y. Wang and W. Shang, “PNC based on consumers’ click behavior,” in *Proceedings of the International Conference on Fuzzy Systems & Knowledge Discovery*, pp. 634–638, IEEE, Zhangjiajie, China, August 2016.

[16] Hukaida, *Design and Implementation of PNC System Based on Hybrid Recommendation Algorithm*, Beijing University of Posts and telecommunications, Beijing, China, 2021, in Chinese.

[17] Z. Ding, Y. Peng, and G. Sun, “PNC algorithm based on improved LDA topic model,” *Journal of Chifeng University (Natural Science Edition)*, vol. 37, no. 06, pp. 28–32, 2021, in Chinese.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 4213645, 12 pages
https://doi.org/10.1155/2023/4213645

Research Article

Unmanned Aerial Vehicle and Geospatial Analysis in Smart Irrigation and Crop Monitoring on IoT Platform

Wei Zhao,[1] Meini Wang ,[2] and V. T. Pham [3]

[1]*School of Digital Art and Design, Dalian Neusoft University of Information, Dalian 116023, Liaoning, China*
[2]*College of Information Engineering, Dalian Ocean University, Dalian 116023, Liaoning, China*
[3]*Saigon University, Ho Chi Minh City, Vietnam*

Correspondence should be addressed to Meini Wang; wmn@dlou.edu.cn

Received 8 September 2022; Revised 22 September 2022; Accepted 29 September 2022; Published 20 February 2023

Academic Editor: R. Mo

The geospatial analysis provides high potential for modeling, understanding, and visualizing artificial and natural ecosystems, utilizing big data analytics and the Internet of things as a pervasive sensing infrastructure. Precision agriculture, weed control, fertilizer distribution, and field management benefit from unmanned ariel vehicles (UAVs). Reduced production costs and improved crop quality are some of the benefits of using this method. Smart farming denotes geographical data utilization to identify field variability, guarantee optimal inputs, and enhance a farm's output. Hence, in this paper, an IoT-assisted Smart Farming Framework (IoT-SFF) with big data analytics has been proposed using geospatial analysis. The use of wireless sensors in IoT devices and communication methods in agricultural applications is thoroughly examined. IoT sensors are available for particular agriculture applications, such as crop status, soil preparation, insect, pest detection, and irrigation scheduled. It is now possible to view our regions in various ways and make accurate agrotechnological decisions, thanks to a computer-generated geographic information system (GIS) for crop irrigation and monitoring. Analytical and monitoring processes that yield timely and accurate decision-making add value to big data, which is a key component for intelligently managing and operating farms. Still, it is constrained by both technical and socioeconomic variables. The simulation findings show that the proposed IoT-SFF model improves the crop yield ratio by 92.4%, prediction ratio by 97.7%, accuracy ratio by 94.5%, the average error by 38.3%, and low-cost rate by 34.4%.

1. Introduction of Smart Farming Using IoT and Big Data

As a new term, "smart farming" refers to farm management that incorporates information technology with modern information and communication technologies, which increases production quantity and quality while optimizing the required human labor [1]. The objective of smart research in agriculture is to establish the farm management support decision-making system. Smart farming believes that population growth, climate change, and work must be resolved from planting and watering crops to health and harvesting [2]. This study aims to develop an IoT-based smart farming method for dealing with difficult situations by UAV. High-precision crop control and data collection can be achieved through the use of smart farming techniques for the optimization of irrigation and monitoring crops. An intelligent agricultural field monitoring system that measures soil moisture and temperature is presented here. The need to efficiently utilize natural resources, the growing use and sophistication of information and communication technologies, and the increasing demands of climate change make smart agriculture increasingly important [3]. Sustainable, smart farming methods lead to a greater diversity of feed, more efficient facilities for water preservation and drought-tolerant crops, and improved animal health. Farmers are leading advocates against climate change risks [4].

Geospatial analytics collect, manipulate, and display data and images, including GPS and satellite photographs, in geographic information systems (GISs) [5]. They are used to create geographical patterns and visualizations of data for more precise modeling and trend prediction [6]. Geospatial forecasting may help companies, due to shifting space environments or locality-based incidents, predict and plan future changes [7]. To avoid risk and disease using UAV, this smart farming sensor-based technology monitors water, monitors normal and dangerous animals through sensors, and saves and improves the farm's production time, production costs, and health for irrigation and monitoring crops. Site-based testing may help politicians understand why strategies that succeed at one location frequently fail at another [8]. It helps to consider the adequacy of soil for different land-use practices, and it is important to avoid the degradation of the atmosphere in connection with land violence [9]. GIS promotes identifying soil types and the concept of soil borders in a region [10].

The combination of GPS and GIS allows data to be correctly obtained in real-time [11]. This helps farmers use mapping devices to chart the farm's precise use of resources to improve water use and production [12]. Farmers can consider farmers' site-specific needs using remote sensing, GPS, and GIS [13]. They can devise and execute management strategies with this knowledge that ensure the optimum utilization of inputs to optimize production and income [14]. GIS will analyze soil data to assess when and how to manage soil nutrients to support the plants' growth. GIS has the potential to use plants [15]. GIS assists farmers in agriculture in increasing productivity and lowering costs by having better land resource management [16]. GIS encourages farmers in agriculture to boost productivity and cut costs by allowing enhanced land resource management [17]. Using Geomatics Technology Agricultural Geographical Information Systems, crops, precipitation, and temperatures can be mapped and forecasted by farmers [18].

Intelligent farming is a high-technology and capital-intensive, cleanly, and environmentally responsible food processing [19]. With an Internet of Things (IoT) sensors' aid (light, humidity, temperature, and soil humidity) and an automatic irrigation device, a system is developed for IoT-based intelligent agriculture to track crop area [20]. The industry will improve operating efficacy, lower costs, minimize waste, and increase its returns in the latest applications in smart agriculture and precision IoT [21]. In addition to helping farmers conserve energy and water, IoT-based systems for precise cultivation often help make agriculture greener; they greatly reduce pesticide and fertilizer use [22] in contrast to conventional farming practices, obtaining a healthier and more organic end product [23]. Trade-in agricultural and food products can be supported by digital technologies based on UAV for optimizing irrigation, which open up new markets for private sector suppliers and give governments new tools for monitoring and ensuring standard compliance, as well as providing more rapid and efficient border procedures for crop monitoring, which is critical for agricultural foods. Big data can have an important effect on intelligent agriculture and the entire supply chain [24]. Smart capabilities, data ownership and protection, and market models are the major concerns that need to be addressed in future research in order to harness the vast quantities of data that deliver unparalleled decision-making capabilities.

The major contribution of the paper is as follows:

(i) IoT-SFF is implemented to collect revenue from developed agricultural fields using UAVs

(ii) Crop monitoring, irrigation, and agricultural requirements can all be better understood with the help of big data

(iii) IoT-SFF can better plan out what crops they will be planting and harvesting times

The rest of the paper structure is as follows: Section 1 discusses the introduction of smart farming using UAV for irrigation and crop monitoring process, and Section 2 discusses literature works. In Section 3, IoT-SFF has been proposed for improved smart farming productivity. Finally, Section 4 concludes the research paper.

2. Literature Works

Saqib et al. [25] suggest smart farming applications using a low-cost information monitoring system. A low-data and low-cost solution are proposed to meet the necessity to track information on real, large-scale farmers. A small farm can be handled quickly. Measurement of sensor-based soil characteristics plays a central role in designing and delivering fully integrated agricultural farms. Remote sensing, global positioning, and geographic information systems can help farmers better understand the unique characteristics of their land. They can use these data to develop and commit to strategies that maximize their outcome and earnings by making the best use of their resources.

Sarker et al. [26] discussed sustainable farm management and digital agriculture through big data. Although it is a long-term debate on the applicability of the big data technology in agriculture, it seeks to investigate how broad data technology leads to sustainable agriculture. The research shows many available large-scale agricultural technology and methods for addressing existing and potential problems on the ground. The study showed that big data technology, that is increasing in agriculture, is still relatively poor. The study indicates that the comprehensive introduction of agricultural big data technology calls for state programs, public-private collaborations, data transparency, financial commitments, and research work on a regional basis.

Santos et al. [27] introduced a wireless sensor technology for cloud-based smart farming for crop production suitability. Agriculture plays a dominant role in the Philippines' economic growth. With more than 6% of overall exports, a total of 25% is nonconstrained, and about 75% has several problem soils, such as steep slopes, low drainage, ground texture, hard cracking clays, extreme fertility constraints, acidic sulfate soils, featuring soils, mining tailing, and contaminated fields. Integration of the wireless sensor network (WSN) technology is required to measure soil's moisture content, wetness, temperature, and pH, and evaluate its current geographic positions in 3D and 3,600 satellite views using the Global Position System (GPS). Farmers can use big data to get granular information on precipitation patterns, watercourses, fertilizer criteria, and many more. Knowing UAVs for optimizing irrigation when to plant and harvest certain crops can make more informed decisions about their business. As a result of the right decisions, farm yields will increase over time from crop monitoring.

Munz et al. [28] explored the farm management information systems (FMIS) in Germany, exploring the characteristics and utilization. Agriculture digitization is one of Germany's most ongoing trends today to address rising commercial, social, and ecological needs in the agriculture and food field. UAV for optimizing irrigation has already become a common practice in the agricultural sector, which uses ICT to collect, share, and analyze data from and within the various stakeholders and structures for crop monitoring at various stages. Based on defined characteristics and features, this paper aims to assign two stages of the digital evolution model to the "one-product" model.

Trilles et al. [29] initialized cloud computing for smart farming and a microservices-based IoT platform. This paper suggests an agnostic architecture of IoT, which emphasizes the IoT platform's role in a larger integrated environment to increase scalability, reliability, interoperability, and reusability. This idea is validated in the IoT scenario of intelligent agriculture, which deploys five IoT devices (SEnviro nodes) to improve wine production. A rigorous performance review guarantees a flexible, secure network.

Maimaitijiang et al. [30] discussed smart farming and plant morphological characteristics, as well as grain policy and food production decisions, which can benefit greatly from nondestructive crop management over huge areas with high performance. In this study, the purpose was to assess the possibility of incorporating shade structure spectral data with a tree crown individual system for crop management using unmanned aerial vehicle (UAV) big data and advanced analytics.

Sinha [31] deliberated the enhancing farmers' net benefit and aerial robot for smart farming. The developing, evaluating, and managing essential time and space factors for farming to optimize profitability, productivity, and environmental conservation is a time-consuming process of knowledge and new electronic technical advancement of the agricultural production system. In this respect, it may play an important function for the robot (aerial, land, and underwater). The existing constraints of aerial robot for the management of agricultural production are being discussed, and potential requirements and technology advancement recommendations are expected.

Based on the survey, there are some challenges in the existing model. This paper proposes the IoT-SFF model to implement smart farming and improve productivity with geospatial analysis and big data to overcome these issues using UAVs for irrigation and crop monitoring process. Section 3 discusses the proposed model briefly which is as follows.

3. IoT-Assisted Smart Farming Framework (IoT-SFF)

This paper discussed the IoT-SFF model to enhance crop yield. Intelligent agricultural research aims to develop an agricultural management decision support system [32–35]. Smart agriculture finds it appropriate to solve the population's concerns, climate change, and labor from seed planting and watering to health and harvesting, which has attracted significant interest. Based on UAV applications for optimizing irrigation, chemicals and fertilizers are commonly used to increase the yield of genetically modified crops in conventional farming. Management levels are a key difference between precision farming and traditional farming. Small areas within fields are managed rather than the entire field as a whole. This increased management level highlights the need for crop monitoring practices. Geographic information system (GIS) [36, 37] is a technology that promotes current agricultural precision methods that ensure the agricultural analytics degree and GIS implementations. This research considers GIS applications such as land adaptability, site search, discovery, allocation of services, impact measurement, land allocation, and information systems. The Internet of things (IoT) gathers geographical information from multiple sources and thus creates connectivity through the Internet to the entire world. It has been reported that a UAV-enabled process for irrigation and crop monitoring for a wide range of salinity assessment methods have been utilized, including modern electrostatic EM38, electro-optic section, and particle micrograph techniques. The knowledge would help manage the land using the appropriate quantity of fertilizers at the correct place.

Figure 1 shows the application of GIS in smart farming. Food producers compete for land, water, and energy supplies and limit food production's detrimental environmental effects. The modification moved manufacturers from conventional farming (CA) to precision farming (PA). PA is introduced to adapt the tractors and machines with GPS sensors for knowledge management [38–40]. In the process of crop irrigation for UAV based on least squares, regression's loss function is the MSE. RMSE, the squared loss function, from which MSE is derived, penalizes larger errors more severely because it is formulated for monitoring crops. The PA extension is the major driving force in big data analysis (BDA) agriculture. The key priority of PA is to collect, handle, and use data for decision-making. PA requires a range of synchronizing technology to capture and interpret data. Although Figure 1 shows the Geographic Positioning System (GPS), remote sensor (RS), and geomapping sensors, the environmental geography division studies the geographical distribution of agriculture and its influences and laws. The geographical distribution of agriculture is subject to a set of laws indicative of its life support system. From the process of crop irrigation for UAV and crop monitoring, the measurement of erosion can be done in one of four ways: (1) modification in mass, (2) modification in the atmospheric boundary layer, (3) transformation in channel flow, and (4) depositional collection from corrosion

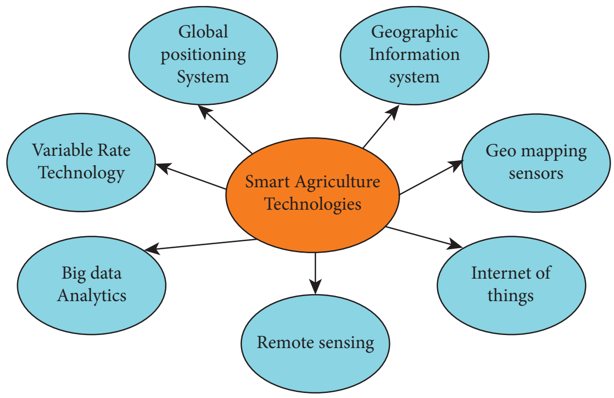


Figure 1: Application of GIS in smart farming.

plot lines and water sources. Agricultural geography is a field of physical science that focuses on the spatial interactions between agriculture and humans. In other words, the study of the phenomena and results in various areas contribute to creating the planet's top surface. Agricultural geographical maps reflect the distinction of the soil. They represent the ties between farming, nature, and economic conditions. Remote sensing provides soil humidity data, which helps measure soil moisture. Water resource mapping: remote sensing is important in mapping water supplies and can be used on a given piece of land for agriculture. The IoT is used in an agricultural environment to translate all elements and activities involved in agriculture into data through sensors, cameras, and other technologies. Big data provide farmers with granular data on precipitation levels, water cycles, and nutrient needs. This encourages them to make intelligent choices, such as cultivating and harvesting plants for better profitability. Ultimately the right choices raise agricultural yields. The production and execution of correct agriculture or site-specific agricultural practices have been enabled by integrating the Global Positioning System (GPS) and GIS. Millimeters (mm) per hour is the unit of measure for evapotranspiration. Water loss from a cropped area is measured using water depth units. Time can be measured in terms of an hour, a day, a decade, a month, or even an entire growing year in UAV for optimization and crop monitoring. In poor visibility field conditions, GPS helps farmers operate, for example, in mud, gravel, fog, and darkness. The VRT technology permits the application of fertilizers, pesticides, calcium, rinsing water, drainage, and other agricultural inputs at varying rates around the field without increasing the rate on machines manually or making multiple crossings [41, 42].

3.1. Case 1: Big Data-Based Smart Farming. Figure 2 shows the big data and GIS-based smart farming. Due to its unique capacity to visually reflect data, descriptive GIS analytics, tools, and applications can execute effective research with elaborated knowledge and transparency. Data filtering methods increase productivity. As of now, it is assisting in the analysis of decades' worth of climate and crop data, looking for trends that will allow farmers to forecast better crop yields and use UAV-enabled processes for irrigation and crop monitoring. Data extraction in farming operations can now benefit from the predictive capabilities generated by large datasets, as well as the proper operating decisions and process redesigns that these datasets allow, all thanks to the development of game-changing marketing strategies. Increased farm productivity, commercial viability, and stronger economic ties are part of agricultural development. There are a few key areas where agricultural change needs to be prioritized under transformation. The GIS analytics deals with internal device problems daily using spatial online analytical processing or surface-down approaches to assess soil and water consistency by implementing surface energy balance applications for soil and digital image processing. The economic and environmental efficiency of precision farming is assessed using pattern analysis to estimate the evapotranspiration rate needed for soil salination assessment. RS data are used for long-term acquisition, validation, and calculation of parameters to explain land cover change and measure soil erosion using unmanned area vehicles for optimization and crop monitoring. The topographic shuttle radar mission data serve as a baseline for testing the landscape characteristics. Agricultural greenhouse gas emissions are studied using economic and environmental models. GIS analytics uses hardware and programming to identify graphic and predictive trends within data and is primarily used to model future events. Various predictive analytics have been used, such as database mining, text mining, and forecasting. An adequate prediction approach is developed for the risks and uncertainty of supply chains for agriculture [43–45]. The crop monitoring uses spatial online analytical processing or surface-down approaches to assess soil and water consistency by implementing surface energy balance applications for soil and digital image processing in the GIS. UAV for optimizing irrigation is used to estimate the evapotranspiration rate required for soil salinity assessment in precision farming. The cost is minimized, and farmers and other interested parties are likely to obtain highly accurate knowledge of climate prediction and take advantage of favorable weather. In this analysis, we categorize GIS analyses' particular applications. The predictive GIS analytics applications is categorized into water/irrigation, soil, plant/agricultural, and fertilization systems.

Further experiments in water/irrigation and crops and agricultural systems have been carried out. Predictive GIS analytics are used when the data are forwarded to the spatially complex event processing engine after filtering and reprocessing.

Figure 3 shows the average error. The typical day of data gathered from different sensors is processed as part of the data preprocessing in the cloud network. The mean of the data is considered since it may include missing and noisy values. Since the data include multiple measuring units (categorical and numerical), standardization is carried out before using the proposed model. In addition to the aforementioned micromeasures and macromeasures, the resulting method defines the error as root mean squared error (RMSE) and mean squared error (MSE). GIS, or

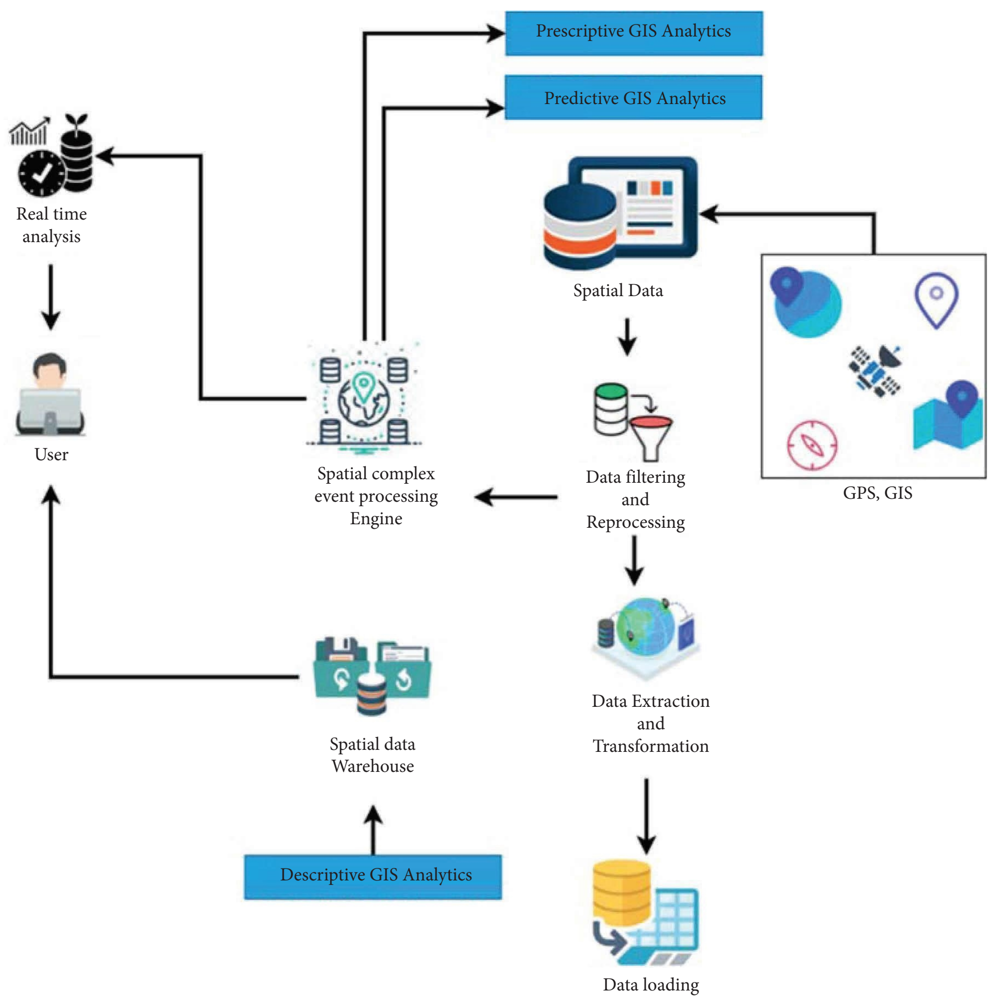


Figure 2: Big data and GIS-based smart farming.

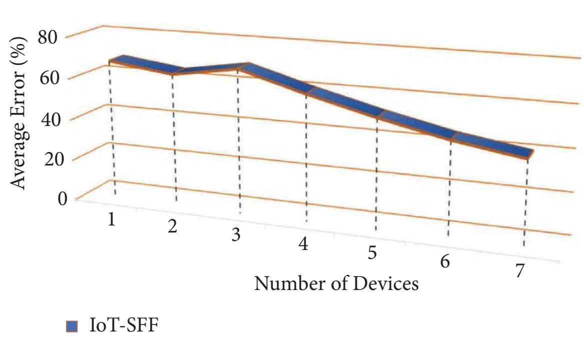


Figure 3: Average error.

geographic information systems, is a relatively new field of study in the information technology in unmanned area vehicles for optimization and crop monitoring. Natural resources used in food production, such as land, weather, hydrogeology, and a wide range of socioeconomic factors, can now be examined and analyzed with greater ease. To transmit data to GPS receivers on the ground, satellites in orbit around the Earth are used. Geographic information systems (GIS) are computer programs that make it possible to use data collected by GPS satellites.

Nevertheless, it should be noted that during testing, the stochastic descent of gradients does not require MSE or RMSE. Rather, the error term is expected between an altered sample and its prediction for big data nodes' weights. MSE and RMSE determine the average model absorption error as

$$\mathrm{MSE} = \frac{1}{n}\sum_{j=1}^{n}\left(X_j - X_j'\right)^2. \quad (1)$$

Here, n is the overall number of data samples, X_j is the target and j th instance, and X_j' is the output or product of the learning model's j th data instance.

Figure 4 shows the ratio for prediction. Decision-making calls for accurate information from sensor results. The big data from the sensor provide learning opportunities in a continuously evolving climate. Such decision-making can be short-term, medium-term, or long-term. When those requirements are met, automatic decisions from big data may be taken that require little to no human involvement. These automatic decisions could vary from temperature management to water supply control irrigation systems. Geospatial analysis and the agriculture stick are combined in this paper. It can be accessed electronically via a mobile phone and combined with various sensors and live data

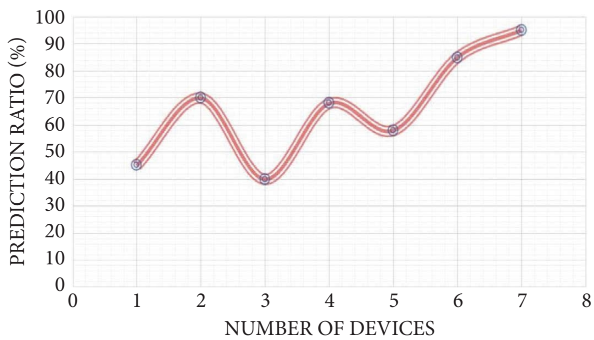


Figure 4: Ratio for prediction.

streams in unmanned area vehicles for optimization and crop monitoring. Testing on actual farmland ensures that the data feeds are accurate in various soil conditions. The use of big data in greenhouses will lead to the identification of ideal conditions for crop cultivation by observing data from the sensors on nutrients, yield, growth, perspiration, color, taste, transplantation, levels of light pest temperatures, and air quality. The precision, conciseness, completeness, and timeliness of data are critical. Several programs have been developed to enable farmers to decide on farms and animals in a cultivated way.

3.2. Case 2: IoT-Based Smart Farming. Figure 5 shows the IoT-based effective communication in agriculture. Smart agriculture based on IoT sensors monitors the environmental state of fields, soil, and crop development for professionals and growers. Sensors, drones, satellite systems, GPS systems, actuators, gateways, cloud servers, the Internet, and android cell phones are all part of the smart farming system. The actuator provides the central coordinator's response to an order, which powers the driving systems in smart farming in crop irrigation for UAV and crop monitoring. A central coordinator measures ground moisture and the actuators based on agricultural field sensor readings. The presented software and hardware led to the progress of these innovations in crop production. Big data, practical guidance, and recommendations from online expert guidance systems for farmers, pests, and disease management are described in [46, 47].

Figure 6 shows the crop yield level. The IoT-enabled precision farming technology ensures that farm efficiency increases and demand grows to satisfy the growing population's food requirements. Surface-down approaches to assess soil and water consistency by implementing surface energy balance applications for soil and digital image processing are two ways GIS analytics which are used daily to solve internal device issues. To prevent soil salinization, farmers use UAVs to monitor crops and analyze irrigation patterns to determine the evapotranspiration rate. Using IoT to boost weather consistency influences crop yields greatly, and one-way IoT has a positive effect on yield. The higher the precision it achieves, the less likely the crop will be damaged by unexpected circumstances, thus improving productivity. A connected farm IoT network has been conducted and found to increase yields by decreasing energy costs per acre for the average farm with IoT-enabled technology and water usage for irrigation. Food farmers drench their crops, limiting growth and yield and increasing the probability that fungal diseases emerge in the soil. Water can be processed, and overwatering challenges are avoided when the farmer has access to the data. It may indicate whether irrigation is inadequate and needs to be increased to optimize cultivation output. There are two types of formats stored: organized and curated. These are based on the smart farming analytics (SFA) data model. Here, the analytical system finds its reference point for reality in UAVs for optimization and crop monitoring. Files in the raw zones must be removed until new data is placed here to avoid undesirable outcomes. There are two types of zones: those containing raw data and those containing processed data.

Figure 7 shows the water management system based on big data and IoT. In several areas of agriculture, IoT is now a feasible database. A study has been conducted to use big data to tackle the large volume of data in many agriculture fields. The formats are stored based on the smart farming analytics (SFA) data model and include organized and curated data. This region becomes the analytical system's center of reality. The files placed under the raw zones should be removed until data are brought into this zone to avoid results. Based on unmanned aerial vehicles for irrigation and crop monitoring, the six steps of multicriteria decision-making include the following: (1) formulation of the problem, (2) identifying the necessities, (3) setting goals, (4) identifying various alternatives, and (5) developing criteria. Verifications of QC and "farming" laws are made at the tables. We provide a single source of truth/access to all main key performance indicators (KPI) for agricultural research. Data are saved in a format that data scientists and data visualization software can process. Promoting healthy water-based relations between biophysical and human processes and maintaining water control to minimize water leakage and recommend emergency measures. The water pressure is within acceptable bounds based on the analysis and simulation of water use patterns. They gather this knowledge, called historical usage data, and offer other data that can be used to predict the potential consumption of water. Multicriteria decision analyses are used in the prescriptive GIS analysis to gather knowledge about large and complex datasets in crop irrigation for UAV and crop monitoring. The MCDA method has been the method of choice for most researchers in their quest to identify the most important factors affecting agricultural productivity. Smart water dripping for farmers can help the automatic and productive use of soil-based irrigation methods based on soil temperature. The approach includes integrating smart farming big data technology into the next granularity stage, offering an infrastructure tailored to fulfill the SFA criteria for smart farm analytics.

GIS analytics are used daily to solve internal device issues using spatial online analytical processing or surface-down approaches to assess soil and water consistency by implementing surface energy balance applications for soil and digital image processing in Figure 8. Farming with precision is evaluated using UAVs in crop monitoring and irrigation

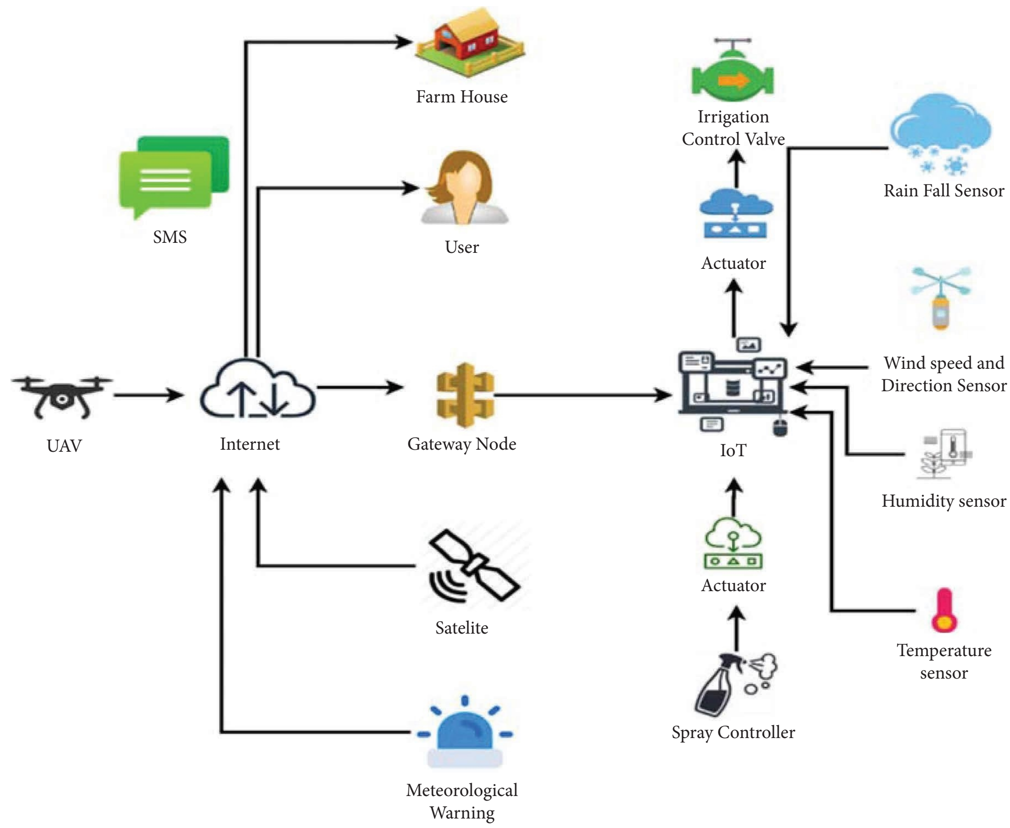


Figure 5: IoT-based effective communication in agriculture.

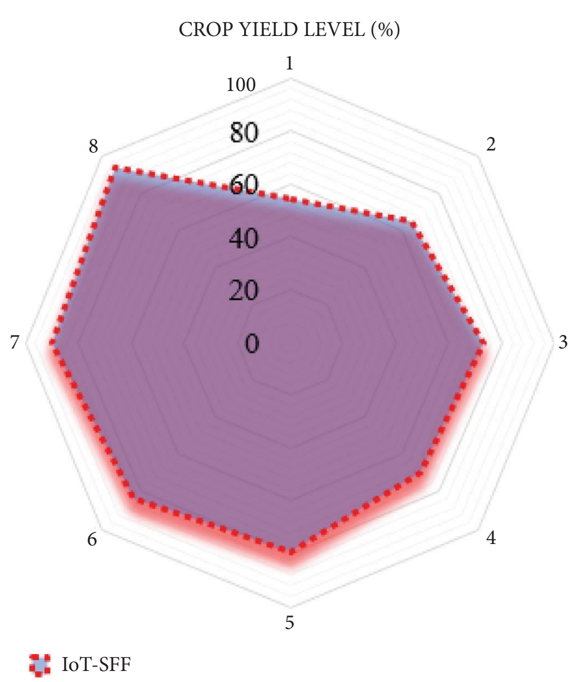


Figure 6: Crop yield level.

pattern analysis to determine the evapotranspiration rate required to prevent soil salinization. The suggested approach includes multiple data sources, data modeling, elements of applications, and technological limitations in UAVs. Our proposal for enhancing the efficiency of the work schedule, mastering the quality of the data from smart farms, and including the irrigation systems to promote agriculture is still being worked on.

Figure 9 shows the IoT-based smart farming. The earliest accuracy relied on satellites to pass seed knowledge to a central hub. Wi-Fi is available to link data directly to a farmer's smartphone from on-site instruments. Many farms that use precise farming use mesh networks that send Wi-Fi signals over several acres. Agriculture's key performance indicators (KPIs) stay updated on feed consumption for irrigation and monitoring, production, and costs in the UAV process. Agriculture and its output are impacted by making more money and being more productive. Time is money when it comes to farming programs based on IoT features. In the measurement of gain and loss, computers are used as records for the cost of manufacturing, shipping, farm processing, and details. The Internet allows farmers and traders to connect with experts in agriculture. In the cloud database, the storage of soil and water resources data and the network management of farm data are realized. In the agriculture knowledge system, multilevel decision-making information and climate growth are important. Developers may use this platform to visually determine how APIs function, the quantity and consistency of data, request processing speed, and resources' availability. The dashboard makes getting actual samples, which IoT offers via our agricultural applications, easier to enlighten rural areas, service water pumps, and run the computer system and telecommunications.

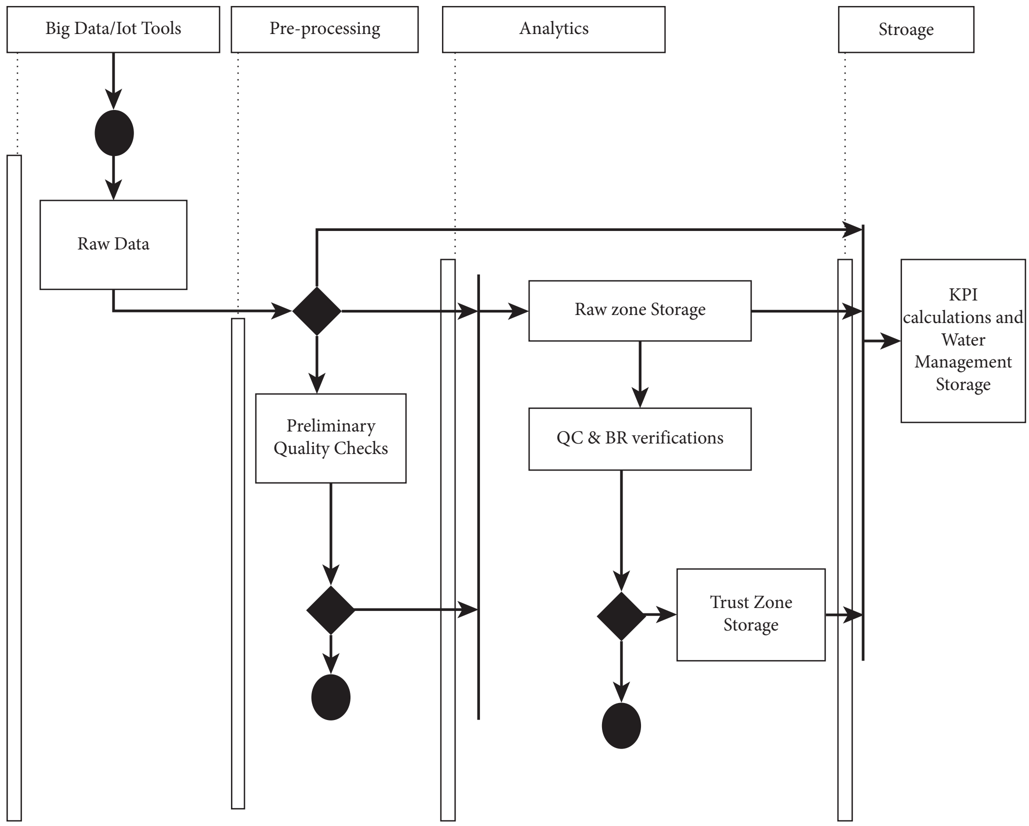


Figure 7: Water management system based on big data and IoT.

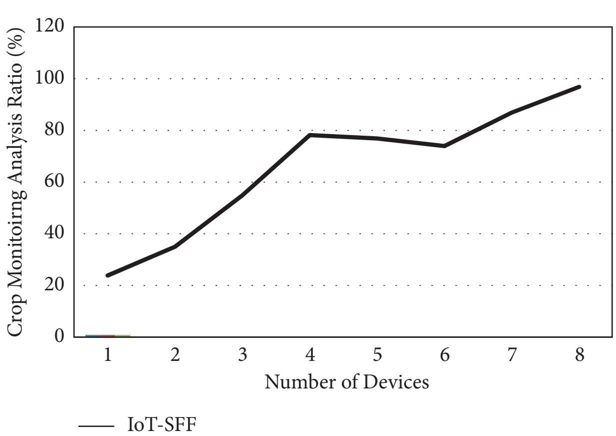


Figure 8: Crop monitoring ratio.

Figure 10 shows the accuracy ratio. A lack of awareness of the importance of climate in agriculture can significantly impact crop production and efficiency. However, when it comes to the Internet of things (IoT), it is possible to monitor the situation in real-time accuracy ratio irrigation and monitoring of crops using UAV. In and out of the agricultural sector, sensors have been installed to select the best cultivars for various climates, using environmental data. Environmental sensors, such as those measuring moisture, rainfall, temperature, and other variables in real-time, make up the IoT ecosystem. Many sensors are needed to monitor and optimize all of these parameters to serve the needs of smart farming. In adverse weather conditions, the need for human intervention increases productivity and yields greater returns on investment for farmers.

Figure 11 shows the cost ratio. The invention of numerous IoT-based devices for intelligent farming transforms every day, leading to crop production by enhancing it, reducing waste, and making it cost-effective. This paper is intended for farmers to generate live data on temperature and soil moisture. Other variables for accurate environmental monitoring are to improve their total production and the quality of products. In utilizing the UAV process for irrigation and monitoring, the robotics in farming GIS can provide accurate maps that include all the necessary information about crops in the field. Task maps and application maps are examples of task maps by precision methods that will use them to maintain the field. This paper combines the agriculture stick with geospatial analysis, and it can be accessed electronically by mobile telephone and combined with different sensors and live data stream. The proposed product is tested on live farming fields to ensure high precision in data feeds in various soil circumstances.

$$f(sa) = s_0 + \sum_{n=1}^{\infty} \left(s_n \cos \frac{n\pi x}{cd} s_n \sin \frac{n\pi x}{cd} \right), \tag{2}$$

$$(1 + x)^n = 1 + \frac{nx}{1!} + \frac{n(n-1)x^2}{2!} + \text{MSE}. \tag{3}$$

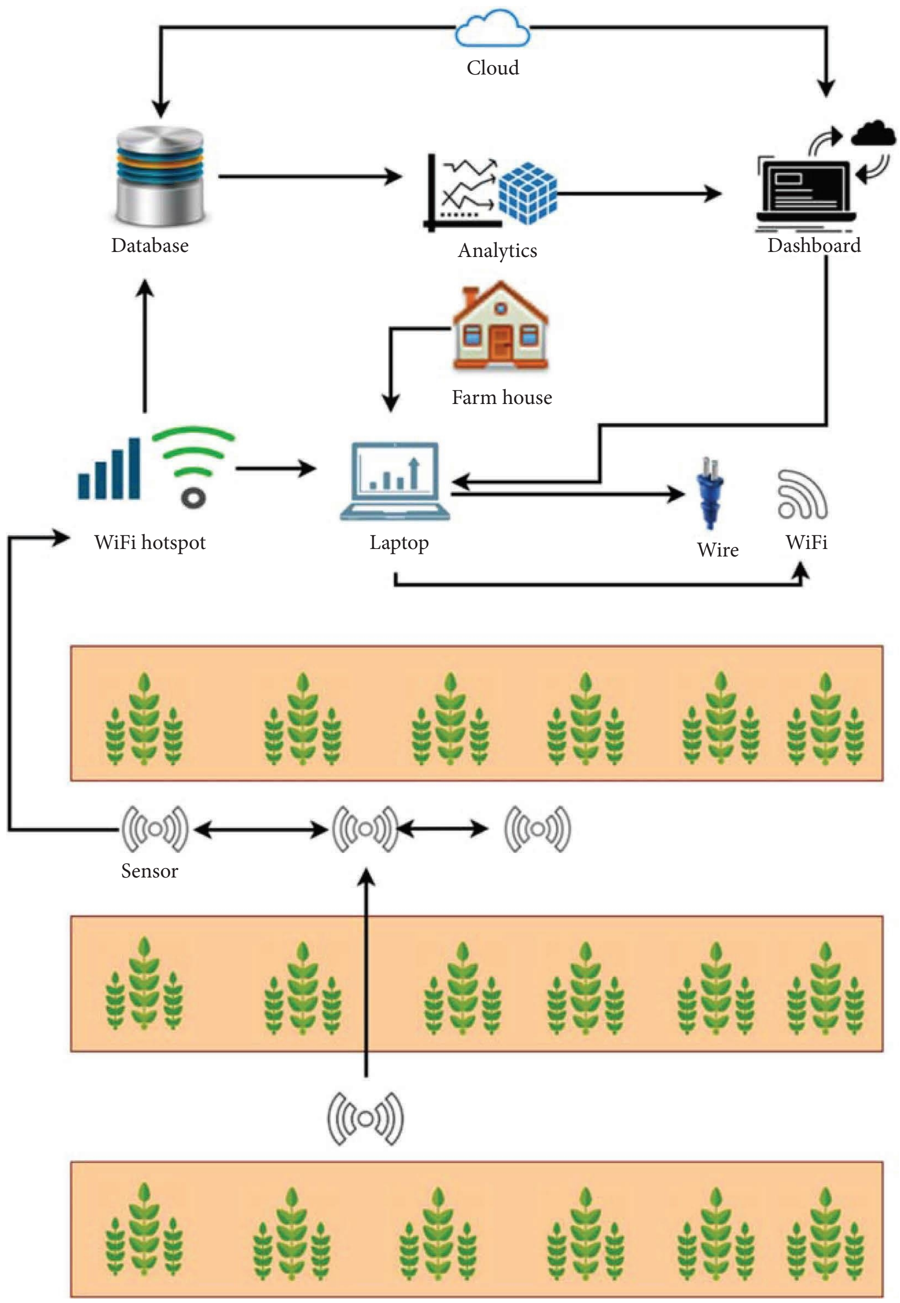


Figure 9: IoT-based smart farming.

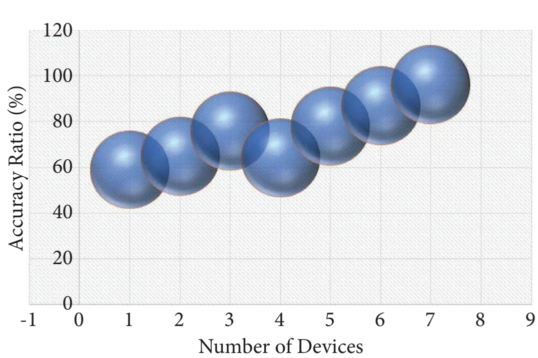


Figure 10: Accuracy ratio.

In the prescriptive GIS analysis, statistical algorithms $f(sa)$, simulations s_0, and multipronged decision analyses are used to collect knowledge of high-volume and complex data cd in crop irrigation for UAVs and crop monitoring with values obtained using binomial equations (2) and (3) with trigonometric values. For the most part, researchers have turned to the MCDA method when attempting to pinpoint critical variables that influence agricultural productivity.

The UAV process in irrigation and crop monitoring can collect operational data and impact operations more than manual practices shown in Figure 12. As a result, the use of robots in manufacturing can be further reduced, and the

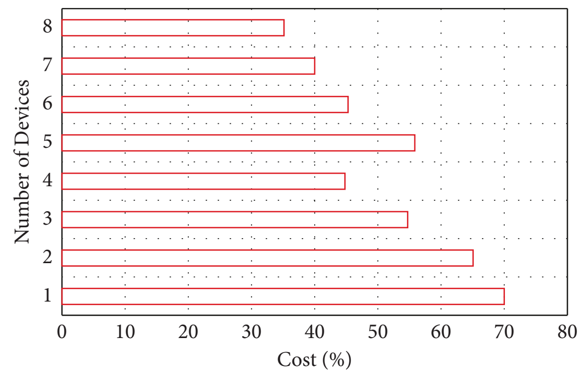


FIGURE 11: Cost ratio.

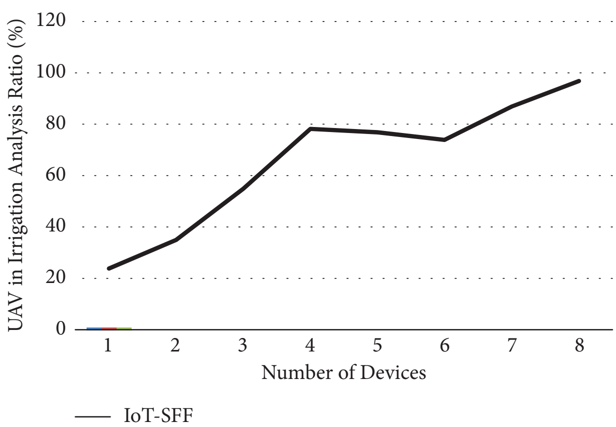


FIGURE 12: UAV in irrigation ratio.

accuracy and effectiveness of the inputs to the operation can increase. Agriculture is transforming with UAV is a term used to describe the ability of physical devices to communicate with each other over the Internet. With an IoT, devices all over a farm can collect data remotely and send it to the farmworker in real-time for crop irrigation and monitoring activities.

The proposed model is analyzed using sample smart farming implementation to measure various parameters and achieve high crop yield, accuracy ratio, prediction ratio, low cost, and average error; automation of greenhouses, crops, cattle, and livestock monitoring and management; farming with precision, smart farming with drones, and predictive analytics; systems for the complete management of a farm are examples of devices used with UAV-enabled processes for irrigation and crop monitoring.

4. Conclusion

This paper discussed the IoT and big data based smart farming technology to improve crop production. Hence, this paper proposed the IoT-SFF with GIS analysis to increase crop yield and fertilize inland smart agriculture. UAV-enabled irrigation and crop monitoring process can get information on soil moisture using remote sensing. This information is then used to determine the crops grown in the area. Farmers can use this information to determine how much water their soil needs by comparing its moisture content with other soils. The Internet of things (IoT) is illuminating agricultural management in a big way, which is why smart farming is becoming increasingly important. Sensors, power transmission, and feelings all contribute to the generation of enormous amounts of data. Big data tracking, evaluation, and value stream from such big data are crucial for smartly coordinating and managing farms. Even though this research, IoT-SFF, is focused on the existence in agricultural production of big data technology, IoT, and data management in the context of agriculture, these constraints are relevant to this research. This IoT-SFF recognizes large-scale analysis to play a major role in enhancing the efficiency of GIS implementation. In farms, the Internet of things enables the system across a farm to collect and transmit real-time data on a wide range of metrics to the farmer. Moisture content, contaminant application, dam thresholds, livestock wellbeing, irrigation, and monitoring can all be monitored by UAV devices, which can then be used to monitor barricades, automobiles, and snow conditions. They provide researchers, agriculture experts, and officials with recommendations for efficient management of large GIS data to increase farm productivity. The simulation findings show that the proposed IoT-SFF model improves crop yield ratio by 92.4%, prediction ratio by 97.7%, accuracy ratio by 94.5%, average error by 38.3%, and low-cost rate by 34.4%.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] A. Vangala, A. K. Das, N. Kumar, and M. Alazab, "Smart secure sensing for iot-based agriculture: blockchain perspective," *IEEE Sensors Journal*, vol. 21, no. 16, pp. 17591–17607, 2021.

[2] V. Higgins and M. Bryant, "Framing agri-digital governance: industry stakeholders, technological frames and smart farming implementation," *Sociologia Ruralis*, vol. 60, no. 2, pp. 438–457, 2020.

[3] P. K. R. Maddikunta, S. Hakak, M. Alazab et al., "Unmanned aerial vehicles in smart agriculture: applications, requirements, and challenges," *IEEE Sensors Journal*, vol. 21, no. 16, p. 1, 2020.

[4] N. Tantalaki, S. Souravlas, and M. Roumeliotis, "Data-Driven decision making in precision agriculture: the rise of big data in agricultural systems," *Journal of Agricultural & Food Information*, vol. 20, no. 4, pp. 344–380, 2019.

[5] K. Sekaran, M. N. Meqdad, P. Kumar, S. Rajan, and S. Kadry, "Smart agriculture management system using internet of things," *Telkomnika*, vol. 18, no. 3, pp. 1275–1284, 2020.

[6] D. S. Bullock, M. Boerngen, H. Tao et al., "The data-intensive farm management project: changing agronomic research through on-farm precision experimentation," *Agronomy Journal*, vol. 111, no. 6, pp. 2736–2746, 2019.

[7] T. C. Hsu, H. Yang, Y. C. Chung, and C. H. Hsu, "A Creative IoT agriculture platform for cloud fog computing," *Sustainable Computing: Informatics and Systems*, vol. 28, Article ID 100285, 2018.

[8] A. Anwer and G. Singh, "Geo-spatial technology for plant disease and insect pest management," *Bulletin of Environment, Pharmacology and Life Sciences*, vol. 8, no. 12, pp. 1–12, 2019.

[9] M. W. Convolbo, J. Chou, C. H. Hsu, and Y. C. Chung, "GEODIS: towards the optimization of data locality-aware job scheduling in geo-distributed data centers," *Computing*, vol. 100, no. 1, pp. 21–46, 2018.

[10] M. S. Farooq, S. Riaz, A. Abid, K. Abid, and M. A. Naeem, "A survey on the role of IoT in agriculture for the implementation of smart farming," *IEEE Access*, vol. 7, pp. 156237–156271, 2019.

[11] K. T. Liu, S. J. Chang, and S. Wu, "Fabrication and characterization of GaN ultraviolet photodetector prepared by growing on geometrical patterned sapphire substrate," in *Proceedings of the 2016 International Conference on Advanced Materials for Science and Engineering (ICAMSE)*, pp. 401–403, IEEE, Tainan, Taiwan, November 2016.

[12] A. T. Balafoutis, F. K. V. Evert, and S. Fountas, "Smart farming technology trends: economic and environmental effects, labor impact, and adoption readiness," *Agronomy*, vol. 10, no. 5, p. 743, 2020.

[13] F. Farivar, M. S. Haghighi, A. Jolfaei, and M. Alazab, "Artificial intelligence for detection, estimation, and compensation of malicious attacks in nonlinear cyber-physical systems and industrial IoT," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 4, pp. 2716–2725, 2020.

[14] P. Paul, P. S. Aithal, A. Bhuimali, and T. Kalishankar, "Environmental informatics vis-à-vis big data analytics: the geospatial & sustainable solutions," *International Journal of Applied Engineering and Management Letters (IJAEML)*, vol. 4, no. 2, pp. 31–40, 2020.

[15] A. Kumari, S. Tanwar, S. Tyagi, N. Kumar, M. Maasberg, and K. K. R. Choo, "Multimedia big data computing and Internet of Things applications: a taxonomy and process model," *Journal of Network and Computer Applications*, vol. 124, pp. 169–195, 2018.

[16] A. Wąs, P. Sulewski, E. Majewski, and P. Kobus, "Use of big data for assessment of environmental pressures from agricultural production," in *Management in the Era of Big Data*, J. Paliszkiewicz, Ed., pp. 137–152, Auerbach Publications, Boca Raton, Fla., USA, 2020.

[17] S. Garg, K. Kaur, N. Kumar, G. Kaddoum, A. Y. Zomaya, and R. Ranjan, "A hybrid deep learning-based model for anomaly detection in cloud datacenter networks," *IEEE Transactions on Network and Service Management*, vol. 16, no. 3, pp. 924–935, 2019.

[18] A. Kirkaya, "Smart farming-precision agriculture technologies and practices," *Journal of Scientific Perspectives*, vol. 4, no. 2, pp. 123–136, 2020.

[19] K. Kaur, S. Garg, G. Kaddoum, F. Gagnon, N. Kumar, and S. H. Ahmed, "An energy-driven network function virtualization for multi-domain software defined networks," in *Proceedings of the IEEE INFOCOM 2019-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 121–126, IEEE, Paris, France, April 2019.

[20] A. Farouk, J. Batle, M. Elhoseny et al., "Robust general N user authentication scheme in a centralized quantum communication network via generalized GHZ states," *Frontiers of Physics*, vol. 13, no. 2, Article ID 130306, 2018.

[21] K. G. Orjuela, P. A. Gaona-García, and C. E. M. Marin, "Towards an agriculture solution for product supply chain using blockchain: case study Agro-chain with BigchainDB," *Acta Agriculturae Scandinavica Section B Soil and Plant Science*, vol. 71, pp. 1–16, 2020.

[22] S. Sankar, P. Srinivasan, A. K. Luhach, R. Somula, and N. Chilamkurti, "Energy-awaregrid-based data aggregation scheme in routing protocol for agricultural internet of things," *Sustainable Computing: Informatics and Systems*, vol. 28, Article ID 100422, 2020.

[23] P. Hemalatha, K. Dhanalakshmi, S. Matilda, and M. BalaAnand, "Farmbot-a smart agriculture assistor using internet of things," *International Journal of Pure and Applied Mathematics, Special Issue*, vol. 119, no. 10, pp. 557–566, 2018.

[24] S. Balamurugan, B. A. Muthu, S. L. Peng, and M. H. A. Wahab, "Call for special issue papers: big data analytics for agricultural disaster management: deadline for manuscript submission: december 12, 2020," *Big Data*, vol. 8, no. 5, pp. 450-451, 2020.

[25] M. Saqib, T. A. Almohamad, and R. M. Mehmood, "A low-cost information monitoring system for smart farming applications," *Sensors*, vol. 20, no. 8, p. 2367, 2020.

[26] M. N. I. Sarker, M. S. Islam, M. A. Ali, M. S. Islam, M. A. Salam, and S. H. Mahmud, "Promoting digital agriculture through big data for sustainable farm management," *International Journal of Innovation and Applied Studies*, vol. 25, no. 4, pp. 1235–1240, 2019.

[27] M. D. Santos, L. L. Lacatan, and F. G. Balazon, "Cloudbased smart farming for crop production suitability using wireless sensor technology," *Test Engineering and Management*, vol. 81, no. 11-12, pp. 5043–5052, 2019.

[28] J. Munz, N. Gindele, and R. Doluschitz, "Exploring the characteristics and utilisation of farm management information systems (FMIS) in Germany," *Computers and Electronics in Agriculture*, vol. 170, Article ID 105246, 2020.

[29] S. Trilles, A. González-Pérez, and J. Huerta, "An IoT platform based on microservices and serverless paradigms for smart farming purposes," *Sensors*, vol. 20, no. 8, p. 2418, 2020.

[30] M. Maimaitijiang, V. Sagan, P. Sidike, A. M. Daloye, H. Erkbol, and F. B. Fritschi, "Crop monitoring using satellite/UAV data fusion and machine learning," *Remote Sensing*, vol. 12, no. 9, p. 1357, 2020.

[31] J. P. Sinha, "Aerial robot for smart farming and enhancing farmers' net benefit," *Indian Journal of Agricultural Sciences*, vol. 90, no. 2, pp. 258–267, 2020.

[32] T. Qureshi, M. Saeed, K. Ahsan, A. A. Malik, E. S. Muhammad, and N. Touheed, "Smart agriculture for sustainable food security using internet of things (IoT)," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 9608394, 10 pages, 2022.

[33] K. N.-E.-A. Siddiquee, Md. S. Islam, N. Singh et al., "Development of algorithms for an IoT-based smart agriculture monitoring system," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 7372053, 16 pages, 2022.

[34] L. H. Li, J. C. Hang, and Y. Gao, "Using an integrated group decision method based on SVM, TFN-RS-AHP, and TOPSIS-CD for cloud service supplier selection," *Mathematical Problems in Engineering*, vol. 2017, Article ID 3143502, 14 pages, 2017.

[35] Y. Li and L. H. Li, "Enhancing the optimization of the selection of a product service system scheme: a digital twin-driven Framework," *STROJNISKI VESTNIK-JOURNAL OF MECHANICAL ENGINEERING*, vol. 66, no. 9, pp. 534–543, 2020.

[36] L. H. Li and H. G. Wang, "A VVWBO-BVO-based GM (1,1) and its parameter optimization by GRA-IGSA integration algorithm for annual power load forecasting," *PLoS One*, vol. 13, no. 5, Article ID e0196816, May.

[37] J. Tian, D. Li, and X. Jia, "IoT smart agriculture and agricultural product income insurance participant behavior based on fuzzy neural network," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4778975, 12 pages, 2022.

[38] E. Ramirez-Asis, A. Bhanot, V. Jagota et al., "Smart logistic system for enhancing the farmer-customer corridor in smart agriculture sector using artificial intelligence," *Journal of Food Quality*, vol. 2022, Article ID 7486974, 8 pages, 2022.

[39] L. Li, C. Mao, H. Sun, Y. Yuan, and B. Lei, "Digital twin driven green performance evaluation methodology of intelligent manufacturing: hybrid model based on fuzzy rough-sets AHP, multistage weight synthesis, and PROMETHEE II," *Complexity*, vol. 2020, no. 6, Article ID 385392, 24 pages, 2020.

[40] L. Li, J. Hang, H. Sun, and L. Wang, "A conjunctive multiple-criteriadecision-making approach for cloud service supplier selection of manufacturing enterprise," *Advances in Mechanical Engineering*, vol. 9, no. 3, Article ID 168781401668626, 2017.

[41] K. Phasinam, T. Kassanuk, P. P. Shinde et al., "Application of IoT and cloud computing in automation of agriculture irrigation," *Journal of Food Quality*, vol. 2022, Article ID 8285969, 8 pages, 2022.

[42] A. H. Adow, M. K. Shrivas, H. F. Mahdi et al., "Analysis of agriculture and food supply chain through blockchain and IoT with light weight cluster head," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1296993, 11 pages, 2022.

[43] L. Li, B. Lei, and C. Mao, "Digital twin in smart manufacturing," *Journal of Industrial Information Integration*, vol. 26, no. 9, Article ID 100289, 2022.

[44] R. Sharma, S. Rani, and S. J. Nuagh, "RecIoT: a deep insight into IoT-based smart recommender systems," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 9218907, 15 pages, 2022.

[45] L. Li, T. Qu, Y. Liu et al., "Sustainability assessment of intelligent manufacturing supported by digital twin," *IEEE Access*, vol. 8, pp. 174988–175008, 2020.

[46] L. Li and C. Mao, "Big data supported PSS evaluation decision in service-oriented manufacturing," *IEEE Access*, vol. 8, no. 99, pp. 154663–154670, 2020.

[47] Y. Wang, H. Li, B. S.-X. Teo, A. A. Jaharadak, and A. Adam, "Image detection system based on smart sensor network and ecological economy in the context of fine agriculture," *Journal of Sensors*, vol. 2022, Article ID 8953914, 12 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 8380307, 15 pages
https://doi.org/10.1155/2022/8380307

Research Article

Machine Learning-Assisted Competency Modeling for Human Resource Management Jobs

Changfang Cao and Zhongying Zhang

School of Management Engineering, Anhui Institute of Information Engineering, Anhui, Wuhu 241000, China

Correspondence should be addressed to Zhongying Zhang; 2016039@aiit.edu.cn

Received 18 July 2022; Revised 4 September 2022; Accepted 14 September 2022; Published 13 October 2022

Academic Editor: R. Mo

The competency model has an important role in human resource management (HRM) as a scientific and objective talent assessment theory. With the widespread use of machine learning methods, machine learning assistance for HRM job competency assessment and determination provides a new way of thinking for talent assessment and management. The establishment of HR competency models based on machine learning methods can quickly and accurately help companies gain a competitive advantage. Based on this paper, we take Company *C* as the research object and establish the HR management job, competency model, based on a genetic algorithm optimized BP neural network (GA-BP). The results show that the model can be better applied to the analysis of the HR management job competency model of Company *C*. The prediction accuracy reaches 91.16% classification accuracy, which has a certain reference significance, and it has a certain reference and value for the HRM job industry.

1. Introduction

With the rapid development of computer technology, the trend of information diversification is increasingly obvious, and the competition between enterprises is more intense. In order to gain more competitive advantages, enterprises must grasp the strategic resources of human resources, play the core role of human resources management in the development of enterprises, and the recruitment management system must be constantly reformed to mature. Recruitment management is the basic work of an enterprise, which will have a significant impact on the management and operation of the enterprise [1–4]. How to solve the above problems requires HRM positions to focus on.

Facing cruel business competition, the survival of enterprises is closely related to the choice of talents. It can even be said that the healthy and sustainable growth and development of enterprises depend on screening potential talents and cultivating talents. Managers' competence is not static. Different times and different business environments have great differences in the requirements of enterprise managers' competence. The real environment has different requirements for managers' competence. As managers, they must dynamically adjust to meet the actual needs in order to remain invincible in the fierce competition. On this basis, how to measure whether employees are competent is an important part of talent search and training. It can be said that the relevant topics of talent competency are not only the subject of HRM but also a major issue to build the company core competitiveness, which is closely related to the company's short-term goals and long-term blueprint. Enterprises must properly measure the ability of employees, ensure that employees are delivered to different positions required by the company with high quality, and improve the adaptability of the company in the current unpredictable market competition environment. However, in fact, most companies do not really realize the importance of employee competency and do not attach importance to the assessment of employee competency. Job allocation is usually relatively reluctant, which may eventually lead to project failure or brain drain. Certainly, some companies have a clearer understanding of the importance of employees, but when evaluating and judging employees, they often focus only on obvious and objective aspects, such as their work experience, education, objective skills, and technology. Employees' life values, job motivation, personal will, their own characteristics, and the degree of job matching have not been in-depth exploration and understanding [5, 6]. In addition, some companies can only rely on years of human resources experience, interview experience, and managers' personal feelings when evaluating and judging employees, which will affect their work. It can be imagined that the results of human resources work have some effect, but not comprehensive. In general, employees are the spirit of the enterprise, evaluation, and judgment of employees is an important part of the healthy development of enterprises. The evaluation and judgment results directly determine the success or failure of the company's operation. The adverse circumstances listed above will lead to higher costs and greater employment risks, which to some extent affect and limit the normal development of the company.

From a realistic perspective, in the case of a sharp increase in the demand of employees in the company, the management position based on the competency model is of great significance. The society has devoted a wide range of attention to HRM positions, so that a variety of management models are derived and applied to the actual business management [7–9]. At present, the competency model is widely used internationally, and the competency model is more biased towards practical ability, which can help enterprises to improve recruitment efficiency, recruitment system more scientific and complete, and create a long-term strategic layout.

Competency model research has a long history. The concept of competency began in the United States in the early 1970s and has been widely used in Britain and other countries. After the concept of competency model is put forward, it has attracted great attention from the business community and academia and has become the focus of global management research. Feng Ming and Yin Mingxin systematically summarized the construction methods of competency model, including behavioral event interview method, functional analysis method, scenario method, performance method, and multidimensional method. Xu Feng analyzed the importance of competency model from the perspective of human resource performance management and proposed methods such as behavioral event interview and human resource index analysis. In the 1990s, the Hong Kong Management Development Center of China used the competency method to train and develop local managers [10, 11]. They used literature review and brainstorming to obtain 30 competency items of middle-level managers. Through the investigation and study of more than 2000 middle managers, Hong Kong Management Development Center has obtained several management competency groups: leadership, communication, team member spirit, team building ability, result orientation, individual driving force, planning ability, efficiency, business consciousness, decision-making ability, customer orientation, and management competency factors have corresponding behavior description. Wang Jicheng put forward several general models of competence in his master's thesis (1) professional and technical personnel model includes: achievement desire, influence, analytical thinking, initiative, self-confidence, interpersonal insight, information seeking, technical expertise, teamwork, and customer service awareness; (2) salesperson model includes: influence, desire for achievement, initiative, interpersonal insight, customer service awareness, self-confidence, public relations, analytical thinking, conceptual thinking, information seeking, authority awareness, related technology, or product expertise; (3) community service personnel model: influence, development subordinates, interpersonal insight, self-confidence, self-control, personality charm, organizational commitment, technical expertise, customer service awareness, teamwork, analytical thinking, initiative, development of others, self-confidence, command, information seeking, team leadership, conceptual thinking, county-wide ownership, public relations, and technical expertise; and (4) entrepreneur model: desire for achievement, initiative, opportunity capture, persistence, information seeking, quality and credibility awareness, systematic planning, analytical thinking, self-confidence, professional experience, self-education, influence, command, development of subordinates, and public relations.

In 1978, Gug divided competence into concept competence, interpersonal competence, and skill competence. In 1982, Richard Boyatzis believed that competency model is related to personal performance, which may be transformed into motivation, personality, technical ability, personal image, social role, and knowledge body. In 1984, Stuart E. Dreyfus and Hubert L. Dreyfus proposed that the competency level is divided into five levels, according to different objects using different development methods. In1993, Lyle M. Spencer and Sige M. Spencer believed that employees' ability-based selection methods should be based on the following assumptions: the more employees' ability meets the job requirements, the higher job performance, and satisfaction. A person's successful adaptation to work depends on (1) accurate evaluation of personal ability; (2) ability model for a given job; and (3) assessment methods to maintain a good balance between people and work. Peter Drucker put forward three hypotheses of business success, emphasizing that the necessary competence of employees to complete organizational mission is the most important one, and it is the fundamental premise to achieve business environment and organizational mission. In 1995, Lcdford believed that competence includes individual unique qualities, which can confirm clearly shown elements and the possibility of performance. In addition to focusing on the current performance, it also needs to focus on future performance. In 1996, Semark a British scholar argued that effective behaviour required by jobs was more important than potential factors. The general industry competency model is questioned and requires improvement in practice. In 2000, Nilan and Alldredge constructed the competency model of 3M company and extended it to the general competency model of administrative staff. In 2008, Peerasit proposed that in addition to knowledge and skills, core competence is also an indispensable factor for project managers [11–14].

In general, the establishment methods of competency model generally include induction and reasoning. Inductive method is based on the current position using the key sample method or behavioral event interview method for statistical analysis of the in-service staff, through the way of induction or brainstorming summarizes the quality characteristics and skill levels have a positive impact on job performance, based on specific behavior, develop competency model. The competency model established by the interview method is closer to the actual situation of the enterprise and is more suitable for mature and stable enterprises. However, the establishment process is time consuming and labor intensive, the operation is difficult, and the established model is not theoretically deduced, which often has certain limitations. Reasoning method is to deduce the competency model through the company's mission vision, core values, roles, and responsibilities of each position. The derivation process of the model can be completed by the company internal and canalso be completed by external professional consultants. This method makes the model logic clear by considering the company's strategy and subdividing job responsibilities globally [15–17]. However, due to the lack of specific behavior as the basis, the description of the model may be too vague. With the rapid development of machine learning and artificial intelligence technology, the extensive and in-depth study of competency provides a new direction for the future development of human resources management.

Based on this, this paper mainly combines machine learning algorithm to establish a new competency model of HRM posts. Through literature survey, questionnaire survey, and time interview analysis, a large number of basic research data are collected and sorted out. Combined with the internal HRM database of a Chinese enterprise management post, we further sort out and eliminate the invalid data that are not completely filled out or filled out with obvious problems, which lays the foundation data for the case study. Then genetic algorithm is used to optimize BP neural network, and finally the competency model of HRM based on GA-BP neural network is established to realize the intelligent evaluation of the overall competency of managers.

2. Concepts and the Features

2.1. Competency Concept. In the process of carrying out performance management and human resource management, enterprises must combine their own actual situation to build a scientific and perfect service system.

The concept of competency can be traced back to the1970s. In 1973, David C. McClelland, a professor of Harvard University, published "Testing for Competence Rather Than for Intelligence" in the *Journal of American Psychologists* [18–20]. He pointed out that the abuse of intelligence tests, sexual tests, academic tests, and grade scores to determine the irrationality of individual competence and proposed to replace the traditional intelligence measurement with competence, trying to find out the most obvious differences between the best performers and the average performers. The publication of this article marks the beginning of competency research. Since then, many scholars have devoted themselves to the study of competency and put forward the definition of competency from different perspectives, as shown in Table 1.

Competency refers to the deep-seated characteristics of individuals who can distinguish outstanding achievements from ordinary people in a job. In recent years, with the improvement of the theoretical level and practical ability of HRM in China, the concept of competency model has also been improved. Its specific meaning is a position in an organization or enterprise, according to its responsibilities and requirements, the centralized representation of the ability support elements needed to complete the work, generally includes the following aspects: (1) professional skills: professional skills are required by the incumbent, so that the staff can skillfully operate the business process of a position, play its unique advantages and creative potential, and can be competent for the job requirements of a position. Employee competence refers to a person's unique knowledge, skills, and personality. It is a unique skill and personality psychological characteristics compared with other personal characteristics on the basis of specific post competency elements. (2) Professional knowledge: it requires the incumbent to master a certain depth and breadth of knowledge related to their own fields, can accurately analyze the problems encountered in the research work, make the right choice, and have the necessary professional technical level or management experience. (3) Professional ethics: it requires people engaged in an industry to meet the basic requirements of the industry and society in terms of ideology, consciousness, and will. (4) Practical ability: it means that the incumbent can apply the learned operating skills to practical work.

In order to improve the competitiveness of enterprises, continuously improve the competency model of the enterprise. The basic evaluation index of competency model is the comprehensive quality of a position. Quality refers to the deep characteristics that distinguish people with outstanding achievements from those with general achievements in a certain work. In the quality dictionary, psychologists divide people's quality into six groups (achievement and action group, help and service group, impact and influence group, management group, cognition group, and personal efficacy group) with a total of 20 specific elements. The elements are divided into many levels, making a comprehensive summary of human knowledge, skills, social roles, self-concept, personality, and motivation, and forming a complete quality model for enterprise employees.

2.2. Characteristics and Classification of Competency Model. Quality refers to the deep characteristics that distinguish people with outstanding achievements from those with general achievements in a certain work. In the quality dictionary, psychologists divide people' s quality into six groups (achievement and action group, help and service group, impact and influence group, management group, cognition group, and personal efficacy group) with a total of 20 specific elements. The elements are divided into many

Table 1: Competency definition summary.

Scholars or institutional research	Definition of competency
Davld. C. MeClelland	Knowledge, skills, abilities, traits, or motivations directly similar to or related to work or work performance or other important achievements in life
Richard Boyatzis	A person has the potential characteristics that lead to good performance in a job (it may be motivation, trait, skill, self-image, or social role, or the knowledge entity he uses, etc.)
Lyle. M. Spencer	Potential personal characteristics related to effective or excellent job performance, including five dimensions: Knowledge, skills, self-concept, traits and motivation
Flei Shman, Wetrongen, Marshall-Mies	A mixture of knowledge, skills, abilities, motives, beliefs, values, and interests
Mirabile	Knowledge, skills, abilities, or characteristics associated with high performance in a position
Dubois	Competency is the necessary ability to achieve or exceed the expected quality level of work output: it is the potential characteristics of an employee, such as motivation, skills, self-image, social role, and knowledge. These factors will lead to effective or outstanding performance in the work
Spencer	Individuals have one or more potential traits, and these potential traits are related to their job or job performance, but also can be expected to reflect their behavior and performance.
Green	Written description of measurable working habits and personal skills used to achieve work objectives
Boyatzis	A person's potential characteristics, such as motivation, traits, skills, self-image or social role, or the knowledge entity he/she uses, will produce effective or excellent job performance
McLagan	Competency refers to knowledge, skills, and competencies that are sufficient to accomplish key work outcomes
Fletcher	Competency refers to some kind of behavior, which is concrete, observable, verifiable and can be classified logically

levels, making a comprehensive summary of human knowledge, skills, social roles, self-concept, personality, and motivation, and forming a complete quality model for enterprise employees. Figure 1.

There are three generally accepted competency models, namely, iceberg model, onion model and trapezoidal model. Figure 2. The details of these three general models are as follows. Figure 3.

(1) Iceberg model is proposed by Spencer competency iceberg model. As shown in Figure 1, the model divides competency into two parts. Specifically, the first part is the iceberg above the level of the revealed part, including knowledge and skills of two competency characteristics elements that are explicit competency characteristics elements. The second part is the iceberg hidden level below the part, including social roles, self-image, personal characteristics, and motivation, which are inherent, recessive, competency elements.

(2) Onion model is proposed by Boyatzis competency onion model. As shown in Figure 2, the principle of competency iceberg model is basically similar, and competency onion model from inside to outside illustrates that the various elements of competency can be gradually observed, measured characteristics. In the onion model, the explicit competency elements are placed at the outermost layer of the onion model, and the potential, hidden, and internal competency elements are placed inside the onion model. Thus, a layer-by-layer competency onion model is constructed from inside to outside. From the outside to the inside, the difficulty of observability, cultivation, and evaluation of competency elements gradually increases. The outermost competency elements are relatively easy to be observed and evaluated, and they are also most likely to be cultivated. However, the more the inner layer is, the more it can reflect the future work performance. Motivation and personal characteristics at the core layer are the most reliable and stable competency.

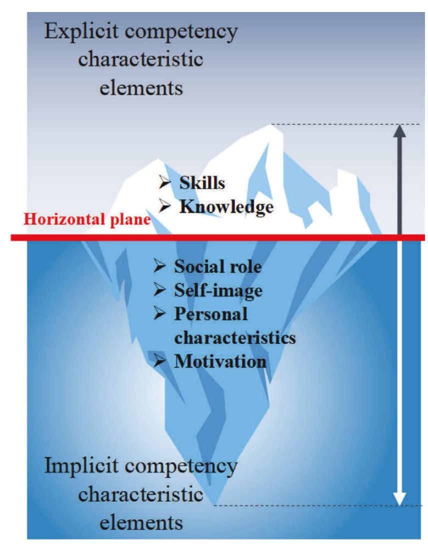


Figure 1: Iceberg model of competence.

(3) The trapezoidal model is proposed by the International Institute of Human Resources Management. As shown in Figure 3, the competency ladder model is divided into six levels from top to bottom, namely knowledge, skills, social roles, self-concept, personal traits, and motivations. Among them, self-concept

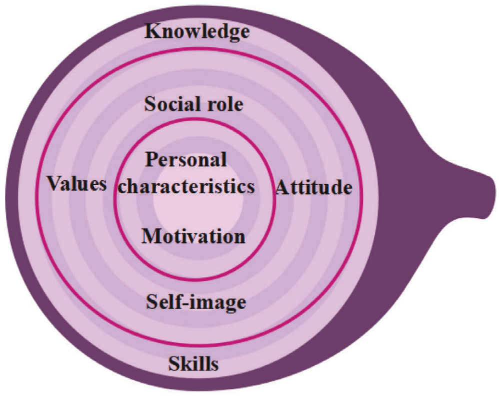


Figure 2: Competency model of the onion.

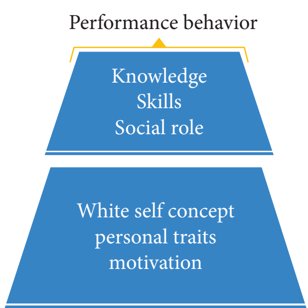


Figure 3: Competency trapezoidal model.

refers to the perception and understanding of individual identity. Specifically, the top of the ladder is personal performance behavior, and the six levels at the bottom of the ladder affect the reflection of individual' s job objectives to varying degrees, thus determining the top of the ladder's personal performance behavior.

With the rapid development of new computer technology, a series of new methods have emerged for competency model building.

3. Problems in HRM Positions

3.1. Analysis of Current Situation of Personnel in HRM. The work done by human resources management personnel is closely related to performance. If the management personnel want to complete the post work with high quality and efficiency, they must have the quality consistent with the post. HRM post personnel can successfully complete the work and reflect that the core quality of work value is the post competency model, including ability, characteristics, knowledge, skills, self-awareness, and values. At present, human resources management posts present the following phenomena:

(1) Multidimensionality: the competency of HRM posts includes various forms of quality, including both representational quality (knowledge, skills, etc.) and psychological quality (ability, characteristics, etc.). Multidimensional quality together constitutes the competency model of management posts.

(2) Dynamic: the environment of HRM position is constantly changing, including external macro environment, internal enterprise environment, and industry environment, which also causes the change of position environment and elements. In order to meet job requirements, job competency must be adjusted with the change of environment.

(3) Specificity: HRM post competency model will be affected by post environment, incentive mechanism, constraint mechanism, job responsibilities, and so on. It is closely related to specific jobs. The quality requirements of different jobs are also different, and different specific competency models are also different.

(4) Strategic: HRM post model includes a variety of core qualities, which can promote the efficient work of employees and help enterprises to achieve strategic objectives.

(5) Level: in the internal structure of human resources management posts, different grades have different levels, and the requirements are not the same, that is, each grade must have an independent post quality, that is, the post quality model is the comprehensive competency model of different grades of posts, and the level of post competency model is directly related to the level of grades.

3.2. Problems of HRM Personnel. In the process of enterprise development, the reasonable application of competency model, from the perspective of enterprise performance management to carry out dynamic management and service, ensure that the enterprise internal different business transactions and different business can achieve reasonable docking. But the current development has the following problems.

(1) The recruitment and selection of human resources management posts are backward. Enterprises are accustomed to the old system and norms. The selection and selection methods of talents in their management positions are relatively backward, and the selection criteria of talents are relatively simple. They pay attention to high academic qualifications, high professional titles, and rich working experience, but they are easy to ignore the specialty and potential of candidates. Many enterprises are accustomed to hiring familiar employees, and there are still related households in terms of promotion, which to some extent hinders the efficiency of human resources utilization in enterprises, and also affects the optimal utilization and composition of human resources within the organization.

(2) The performance management mechanism of human resources management posts needs to be improved. Although many enterprises have established

a simple performance management system, some performance management mechanisms are mere formality and do not really play their effectiveness. The main performance is not well-applied performance management system, and assessment implementation process is not standardized. In the process of performance evaluation, due to the large number of enterprises and complex relations, performance management often considers acquaintance relations, which brings some emotional colors to the implementation of performance evaluation and affects the fairness and rationality of HRM in the whole enterprise.

(3) The design of HRM training system needs to be improved. Enterprise internal training mode is mostly endogenous growth, training effect needs to be improved, and training system design needs to be improved. From all aspects such as entry and promotion, new recruits need prejob training to be employed. But in this process, there are also fewer participants training, training management is not standardized, many ginseng training frequency and frequency is low, training learning is not strict enough, training management is a mere formality, and it is difficult to achieve the training effect required by the enterprise.

(4) The personal development plan design of HRM post is not clear enough. Employees are social elites with high education and professional title and good interpersonal relationship. They have urgent needs for the future, but their career planning is not clear enough to meet the current situation. In state-owned enterprises, the phenomenon of seniors in the field of investment is relatively serious. Many old employees are used to passive arrangement of their own career paths and lack active planning. In addition, the Ministry of Human Resources Management of the staff professional channels lack of scientific and reasonable settings, business staff career channels single, many people have little hope for promotion, to a certain extent, affected the enthusiasm of professional and technical personnel.

(5) Talent inventory in human resources management positions is a mere formality. Traditional enterprises prone to error talent inventory purposes: in the process of business transformation to feel the pain of talent management to carry out talent inventory, and not the talent inventory as a routine management process. In addition, companies mistakenly think that the inventory result is to promote people, rather than more scientific classification of team personnel management, improve team effectiveness. Some enterprises' talent inventory is formalistic, with talent inventory every year, but no further follow-up and application of inventory results.

(6) Human resources planning needs to be improved. Some enterprises believe that the planning of HRM is not related to the strategic development and business objectives, so there is no idea of combining HRM with business objectives. The lack of reasonable planning and connection between human resource planning and enterprise management directly affects the sustainable development of enterprises. In the implementation process of human resource planning, some enterprises simply use past historical data as support to predict the current human resource data for the convenience of work. This measure is difficult to provide accurate human resource planning data for enterprises and cannot provide strong support for the talent training of enterprises.

3.3. Necessity of Building Manager Competency Model Based on Machine Learning. In order to ensure the sustainable development of enterprises and attract more excellent human resources, enterprises should build performance-based human resource management competency model to ensure that the talent market and the competitive market environment can achieve good coordination and matching, to ensure that the two can be fully integrated. It is believed that the application value of competency model in the field of HRM positions is huge, especially in recruitment, training, and employee development. The status of competency model system in the field of human resources management can be shown in Figure 4.

Step 1. As shown in the line 1 in Figure 4, when the strategy of the enterprise changes, the roles and responsibilities of various departments and employees have also changed accordingly, and the individual competency requirements of the enterprise will also be different from the past. At this time, the original competency model has been unable to accurately reflect the new enterprise development strategy needs of the staff ability and quality requirements. Therefore, enterprises need to re-adjust the competency model according to the new strategic objectives and business priorities.

Step 2. As shown in Route 2 in Figure 4, when the enterprise revises the original competency model according to the new strategic objectives in a timely manner, that is, after redefined the competencies and qualities necessary for the staff, it is not possible to evaluate the performance of the staff in accordance with the previous evaluation methods, but to evaluate the performance of the staff in daily work according to the competency elements defined in the new competency model, and there will be some competency assessment results.

Step 3. As shown in Route 3 in Figure 4, the existing competency level of the staff obtained through scientific evaluation in the second step is compared and analyzed with the competency level necessary to achieve the position, and the gap is concluded. In view of the existing gap, to design to improve staff competency

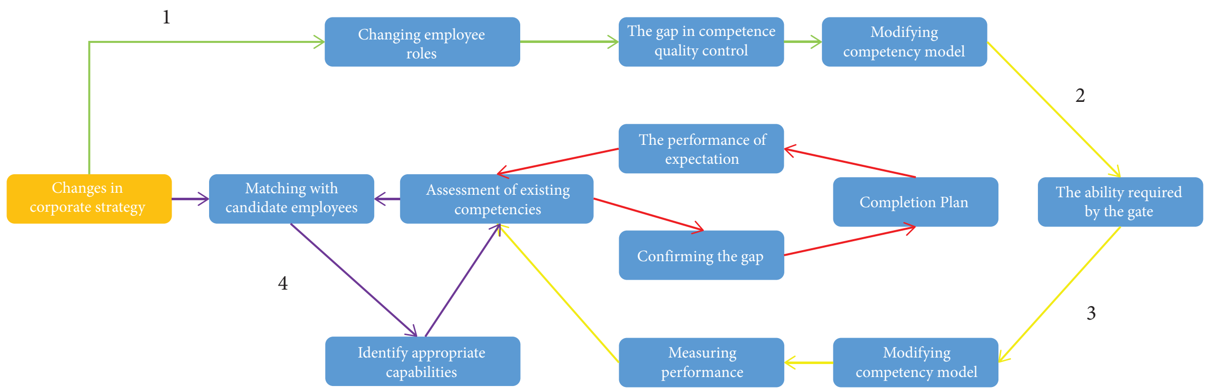


FIGURE 4: Competency model and HRM system.

plan, and targeted to develop the corresponding training plan.

Step 4. As shown in Figure 4 in line 4, if the second step is over, the existing competency level has reached the standard of the position and even reached the requirements of the upper level; or appropriate training makes the staff's competency has been improved, to meet the requirements of the position or the superior position, then the enterprise will be the existing staff competency level and competency model matching, as a basis, to determine whether the staff is to remain or promote the superior position.

The cycle of the above four steps, in which the work content and competency model are adjusted and adapted to the changes of corporate strategic objectives to ensure that HRM strategies and business activities are always consistent with strategic objectives. But in practical application, various factors are changing rapidly, making the competency model complicated. However, the combination of machine learning method in the competency model of HRM posts will greatly reduce the establishment time and result analysis of competency model.

4. Construction of Human Resource Competency Model Based on BP Neural Network

Through literature review, semi-structured interviews and open questionnaire survey, the competency modulus table of HRM positions in Company C is obtained, which provides basic data for the subsequent establishment of human resource competency model based on machine learning assistance [21–23]. Company C management post competency model design process is shown in Figure 5.

4.1. Literature Review. Through questionnaires, on-site interviews, literature review analysis, and other methods to obtain information on the competency and job performance of middle-level managers in enterprises, the theme of the survey interview is determined to be "list the key competencies you think you need to do a good job in middle-level management." Interviews with middle-level managers around the theme to obtain the ability characteristics that excellent middle-level managers should have. Five dimensions and 22 related competency elements are summarized. The index system of the general competency characteristics of management talents in Company C is shown in Figure 6.

4.2. Semistructured Interviews. A semi-open interview is conducted on the management staff of Company C, which mainly adopts one-to-one semi-structured interview. The semi-structured interview system is shown in Figure 7.

4.3. Open Form Questionnaire. The characteristics of the questionnaire in this paper are that participants can freely express their views and opinions according to their own experience and ideas and the way participants fill in the questionnaire is to answer freely. The structure chart of questionnaire competency index system is shown in Figure 8.

4.4. Company C Management Post Competency Index. The basic components of the enterprise competency model (knowledge, skills, social roles, etc.) were determined through a literature review. Semi-structured interviews refer to informal discussions conducted with a rough outline to obtain components of the enterprise competency model (values, attitudes, motivations, etc.). The open questionnaire survey gets the index size of each factor index among employees [24–26]. In this paper, the final determination of C Company management post competency model is shown in Figure 9.

A total of 110 samples of human resources management positions in Company C in 2021 are collected in Table 2. 110 samples can be divided into excellent performance and other performance samples, of which 64 are excellent performance samples and 46 are other samples.

5. Construction of Human Resource Competency Model Based on GA-BP Neural Network

5.1. Genetic Algorithm Model. Genetic algorithm is an evolutionary algorithm. Its basic idea is based on Darwim's

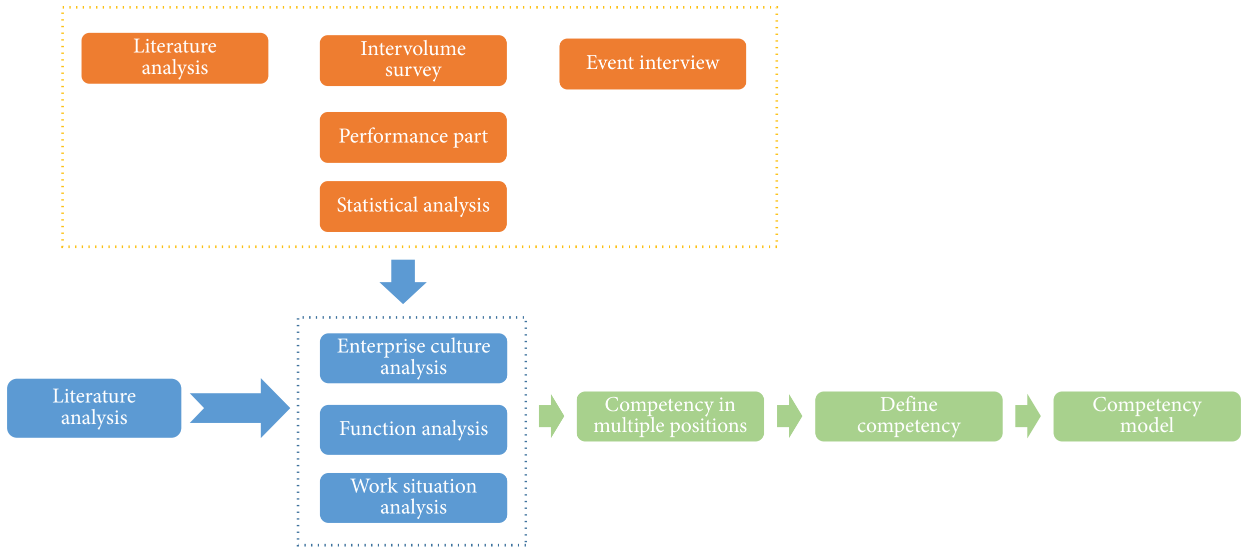


Figure 5: Competency model construction process of Company C.

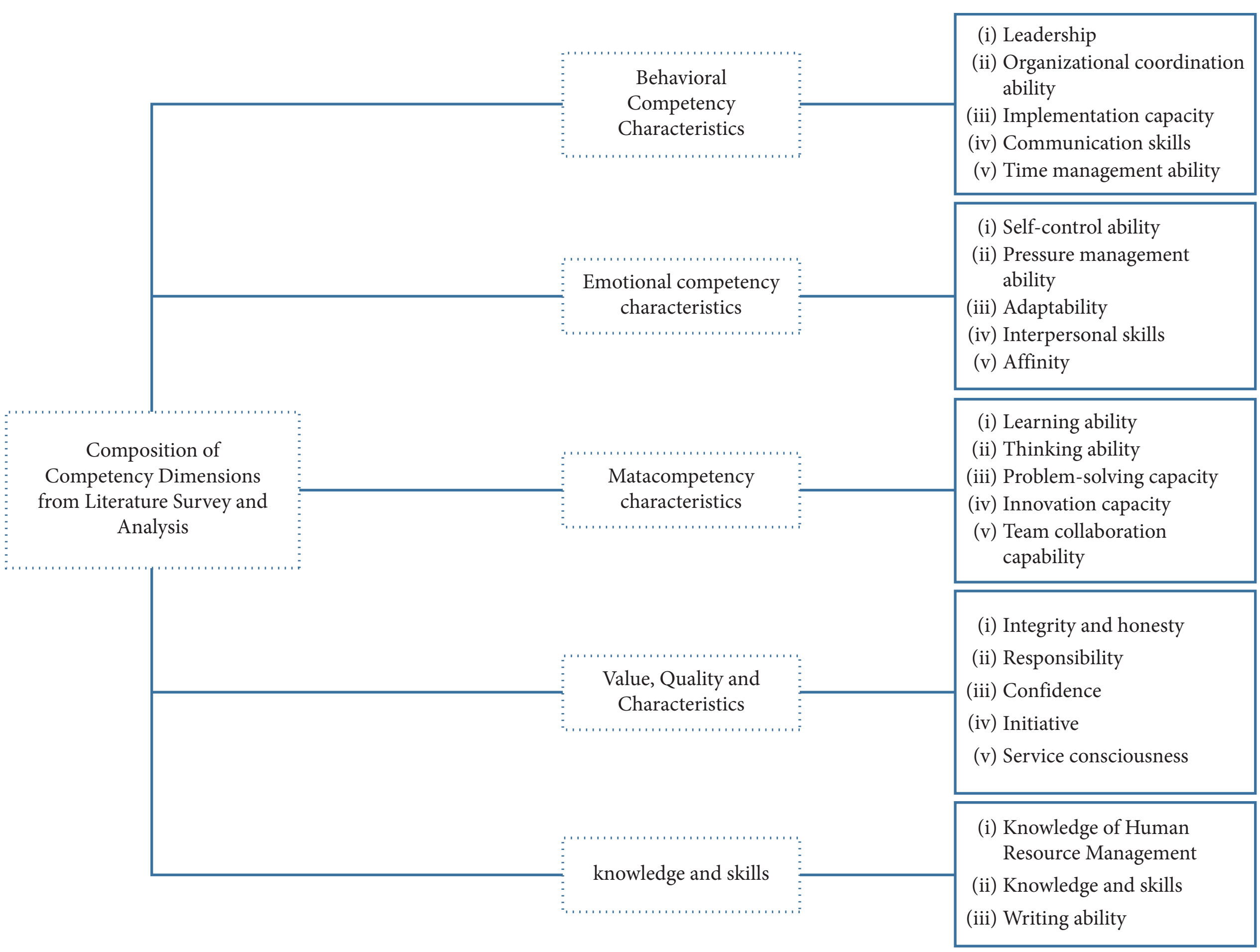


Figure 6: Competency index system from literature.

theory of evolution and Mendel's theory of genetics. It was first proposed by Professor J. Holland of Michigan University in 1962. It is a parallel random global optimization search algorithm based on natural genetic law and biological evolution. The algorithm is evolved from the genetic mechanism of natural competition and survival of the fittest. Through the combination of mathematical principles and computer simulation, a series of selection, crossover, and

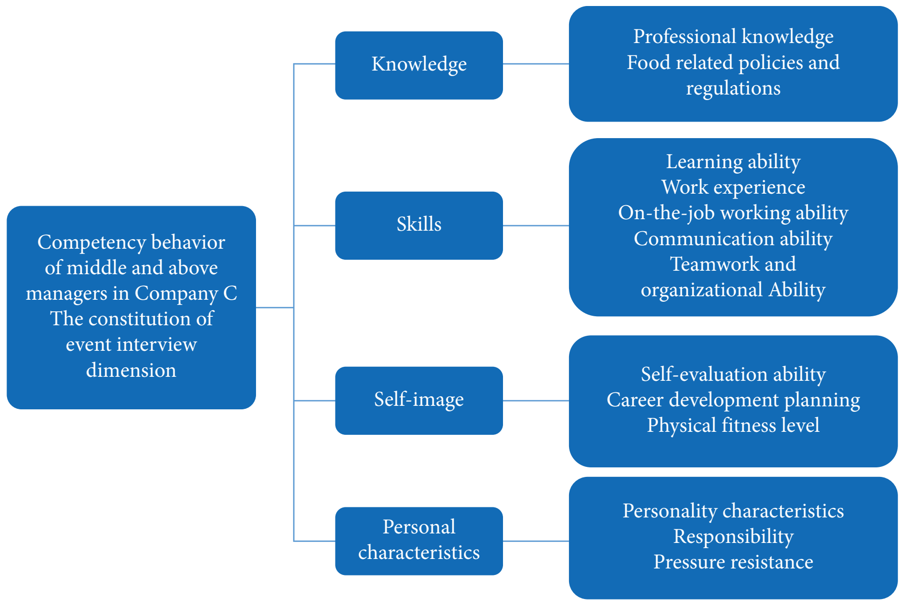


Figure 7: Semi-structured interview competency index system structure.

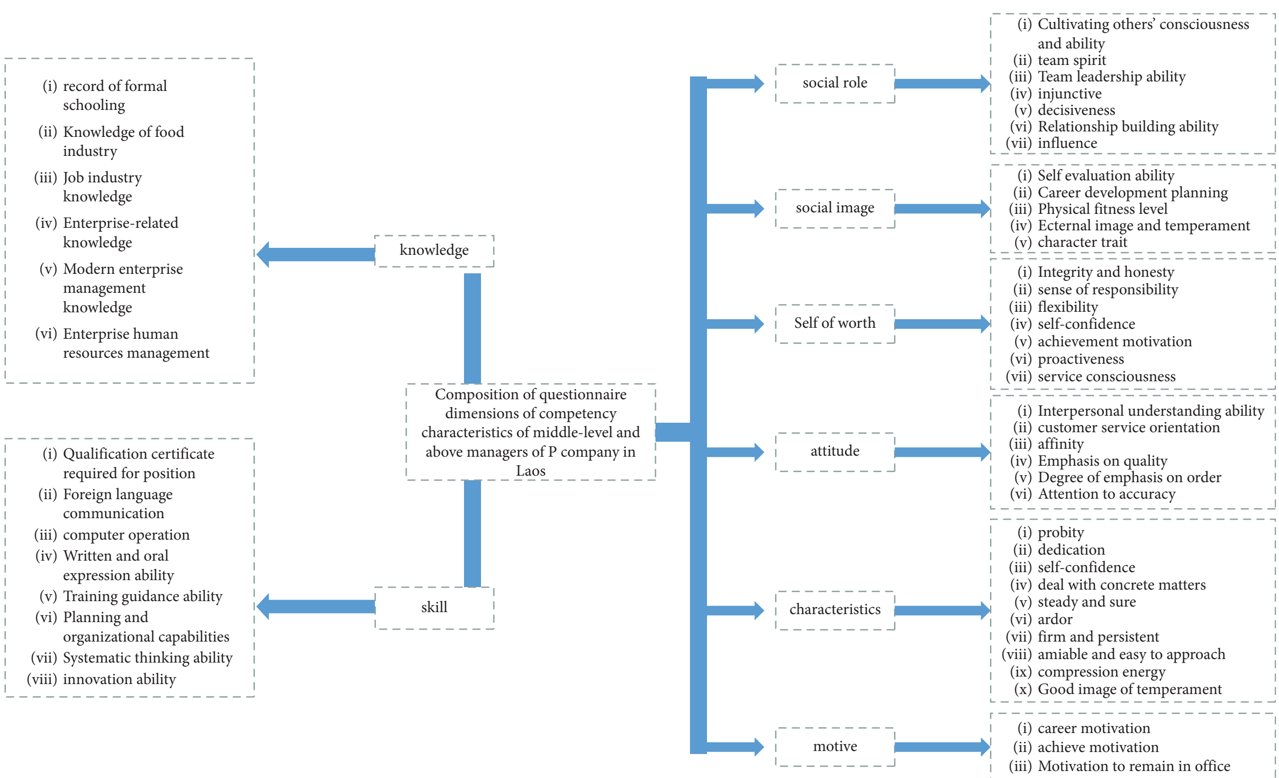


Figure 8: Questionnaire competency index system.

mutation operations are carried out on biological chromosome genes to imitate the biological evolution in nature, so as to obtain the best offspring, namely the optimal solution. Genetic algorithm has good global search ability. Compared with the traditional gradient descent algorithm, it is easy to fall into local optimal solution, and genetic algorithm has certain advantages. This method has the characteristics of efficient heuristic search and parallel

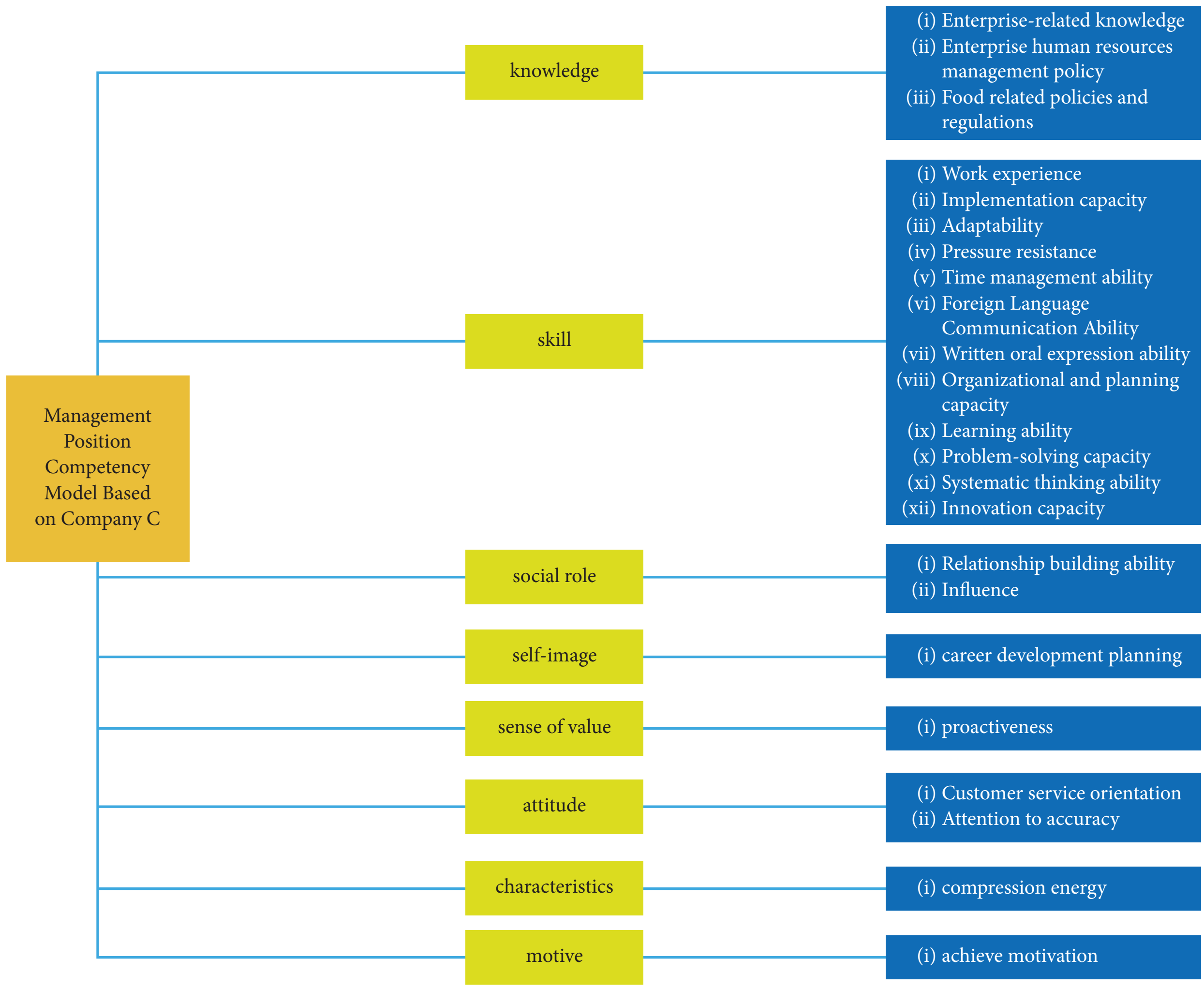


Figure 9: Company C management post competency model.

Table 2: Data set.

Sample number	Excellent performance samples	Other performance samples
110	64	46

computing and is usually used in functional optimization, combinatorial optimization. and production planning [27–30].

The main process of genetic algorithm is as follows:.

Step 1. Coding: coding the solution of the problem in the form suitable for genetic algorithm is called the representation of genetic algorithm, and the coding methods include binary coding, real number coding, integer coding, and tree coding.

Step 2. Generate the initial population, that is, select several of the solutions to form the initial population, which is generally generated by random method.

Step 3. Fitness evaluation: introduce fitness to evaluate the ability of each individual in the selected initial population to adapt to the surrounding environment. The specific method is to calculate the fitness value of each individual through the fitness function to compare the differences between individuals, mainly with or without constraints and constraints of two fitness functions.

Step 4. Selection: selection is to extract the individual with higher fitness value as the parent and enter the next stage of cross operation. The commonly used method is the roulette method. Since the probability of each individual being selected is proportional to the fitness value, the basic principle is to select according to the proportion of individual fitness value.

Step 5. Crossing: part of the genes in the coding chain of two parent individuals are exchanged to produce a new individual. Cross operation is the main means to generate new individuals. For real number coding, the cross method includes discrete recombination, linear

$$\begin{array}{l} \text{A: } 1100\ 0101\ 1111 \\ \text{B: } 1111\ 0101\ 0000 \end{array} \xrightarrow{\text{Crossing}} \begin{array}{l} \text{A: } 1100\ 0101\ 0000 \\ \text{B: } 1111\ 0101\ 1111 \end{array}$$

FIGURE 10: Crossover process.

$$\text{A: } 1100\ 0101\ 1111 \xrightarrow{\text{variation process}} \text{A: } 1100\ 0101\ 1101$$

FIGURE 11: Variation process.

recombination, and intermediate recombination. Cross operation is shown in Figure 10.

Step 6. Variation: mutation is to change the values of some loci of individuals randomly selected in the initial population with a certain probability (usually called mutation probability). The mutation operation is shown in Figure 11.

Step 7. Termination condition judgment: The optimal solution is obtained by the termination algorithm.

5.2. BP Neural Network. BP neural network is a neural network with three or more layers. Any network is composed of an input layer, an output layer, and one or more hidden layers in the middle layer. Each layer contains several neurons responsible for information processing. There is no connection between the neurons in the same layer, and the adjacent two neurons are fully connected by weights. However, each neuron only accepts the input of the previous layer and is responsible for exporting the information to the next layer without feedback to itself. The hidden layer can be one or more layers between the input layer and the output layer. The function of the internal representation of the hidden layer as the input mode is to extract the features contained in a class of input modes that are different from those contained in other types of input modes and to transfer them to the output layer. Due to the unique recognition function of the hidden layer, it is also called the feature extraction layer. The process of feature extraction in hidden layer is the process of self-organizing the weights between the input layer and the hidden layer. That is, the weights between the input layer, the hidden layer, and the output layer are gradually adjusted from the initial random value during the training process of the network, and finally the learning process that can characterize the characteristics of the input mode is achieved. The common three-layer neural network topology is shown in Figure 12.

The network of three layers, shown in Figure 12, consists of the input layer, the hidden layer, and the output layer. Full connection between the upper and lower layers and no connection between each layer of neurons. When a pair of learning samples are provided to the network, the activation value of neurons propagates from the input layer through the hidden layer to the output layer, and each neuron in the output layer obtains the input response of the network.

Data preprocessing is to convert various data and nondata into positive data indicators. Depending on the nature of the indicators, different conversion methods were used. For example, the data obtained by the competency model scale usually belongs to quantitative data such as the lowest score of 0 points, the highest score of 5 points. The data obtained after pretreatment should be the dimension matrix $n \times m$, where n is the number of samples, that is, the number of respondents; m is the number of evaluation index factors, namely the input of artificial neural network.

In the forward propagation mode, through the input vector $A_k = (a_1, a_2, \ldots, a_m)$, where $k = 1, 2, \ldots, n$ is the serial number m of the training sample, which is the number of neurons in the input layer. The input of the input layer is the evaluation result of each competency index. The function of the input unit is to accept the external input mode and transmit it to all the units of the hidden layer connected to it.

Calculate the hidden layer input $S_k = (S_1, S_2, \ldots, S_p)$, where p is the number of hidden layer neurons. The input value of the hidden layer is the function of the difference between the input weighted sum and the threshold:

$$s_j = \sum_{i=1}^{m} w_{ij} a_i - \theta_j \tag{1}$$

where a_i is the output value, w_{ij} is the link weight between the input layer and the hidden layer, $i = 1, 2, \ldots, m$ is the serial number of the input unit, $j = 1, 2, \ldots, p$ is the serial number of the hidden layer unit, and θ_j is the threshold of the hidden layer neurons.

Calculate the hidden layer output $B_k = (B_1, B_2, \ldots, B_p)$. The function of the hidden layer unit is to generate a set of output values by the over-type nonlinear function after weighted summation of all the input values and then transmit them to all the units connected to the output layer. The output value of the layer is S (Sigmoid) function.

$$b_j = f(s_j) = \frac{1}{1 + e^{-s_j}}. \tag{2}$$

Calculate the number of output layer input $L_k = (l_1, l_2, \ldots, l_q)$, q is the number of output layer neurons. The output layer input value is a function of the difference between the hidden layer output weighted sum and the threshold.

$$l_k = \sum_{j=1}^{p} v_{jt} b_j - r_t, \tag{3}$$

where v_{jt} is the link weight between the hidden layer and the input layer, q is the serial number of the output unit, and r_t is the threshold of the output layer neurons. In calculating the output layer output $C_k = (c_1, c_2, \ldots, c_q)$.

$$c_k = f(l_t) = \frac{1}{1 + e^{-l_t}}, \tag{4}$$

The inverse propagation mode is carried out, and the output vector $Y_k = (y_1, y_2, \ldots, y_q)$ is preset to calculate the correction error of the output layer.

$$d_t = (y_t - c_t) \cdot c_t (1 - c_t), \tag{5}$$

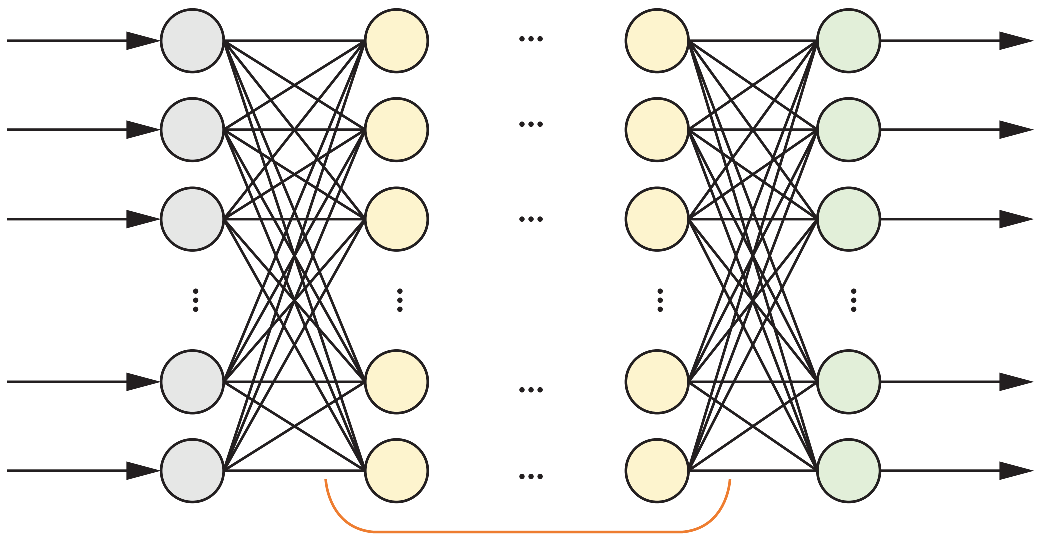


Figure 12: Three-layer BP network structure.

where d_t is the output value of the network, that is, the actual output of the t neuron, $(y_t - c_t)$ represents the absolute error between the expected output and the actual output. $c_t(1 - c_t)$ is the first-order differential of the output layer S function, which indicates that the deviation is adjusted according to the actual response of each element.

The connection weight correction between the output layer and the hidden cross is calculated as

$$v_{jt}(N+1) = v_{jt}(N) + a \cdot d_t \cdot b_j, \tag{6}$$

where a is the learning factor. The threshold correction between the output layer and the hidden layer is calculated as

$$r_t(\mathrm{N}+1) = r_t(\mathrm{N}) + a \cdot d_t. \tag{7}$$

The correction error of the hidden layer is calculated.

$$e_j = \left(\sum_{t=1}^{q} d_t \cdot v_{ij}\right) b_j\left(1 - b_j\right), \tag{8}$$

where $b_j(1 - b_j)$ is the first-order differential of the hidden layer S function. The meaning of e_j is like d_t. The correction error of each hidden layer is generated by the correction of q output units. Connection weights between the hidden layer and the output layer are calculated.

$$w_{it}(\mathrm{N}+1) = w_{it}(N) + \beta \cdot e_j \cdot a_i, \tag{9}$$

where β is the momentum factor. Then calculate the threshold correction between the hidden layer and the input layer.

$$\theta_{\mathrm{j}}(N+1) = \theta_{\mathrm{j}}(N) + \beta \cdot e_j. \tag{10}$$

Calculate the k^{th} sample error.

$$E_k = \frac{1}{2}\sum_{t=1}^{q}\left(y_t - c_t\right)^2 \tag{11}$$

Calculate the average error of all test samples.

$$E = \frac{1}{n}\sum_{k=1}^{n} E_{\mathrm{k}}. \tag{12}$$

After repeated iterations, when the error is less than the allowable value, the network training process is ended.

5.3. Genetic Algorithm Optimization Bp Neural Network. BP neural network is an algorithm based on the principle of gradient descent. The selection of initial weights and thresholds is random and lacks scientific basis, resulting in BP neural network is more likely to fall into local optimal solution and the global search ability is relatively poor, and the final prediction accuracy is affected. To solve this problem, genetic algorithm for distributed information exploration and collection can cover the whole space, with the characteristics of global search. Therefore, this paper will use genetic algorithm to optimize the original BP neural network. Aiming at the problem that the initial weights and thresholds of BP neural network are randomly generated, genetic algorithm is used to optimize the initial weights and thresholds. Combining the characteristics of the two, the GA-BP neural network is constructed to make up for the defects of the random initial weights and thresholds of the BP neural network. The GA-BP neural network has the advantages of the nonlinear mapping of the BP neural network and the global search of the genetic algorithm, which further improves the prediction accuracy of the BP network, and finally obtains the global optimal solution as shown in Figure 13.

5.4. Establishment of Human Resource Competency Model Based on GA-BP Neural Network. The process of genetic algorithm optimization mainly includes the determination of network basic topology, genetic algorithm optimization, and neural network prediction. First, the BP neural network topology is optimized by genetic algorithm, which mainly

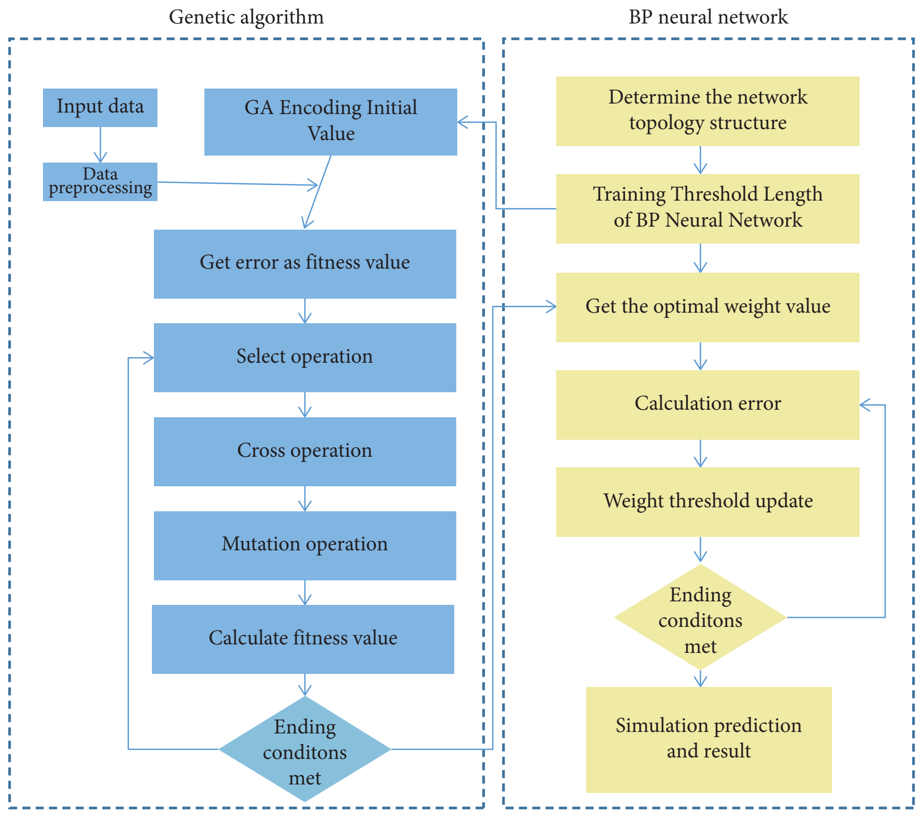


Figure 13: GA-BP neural network flow chart.

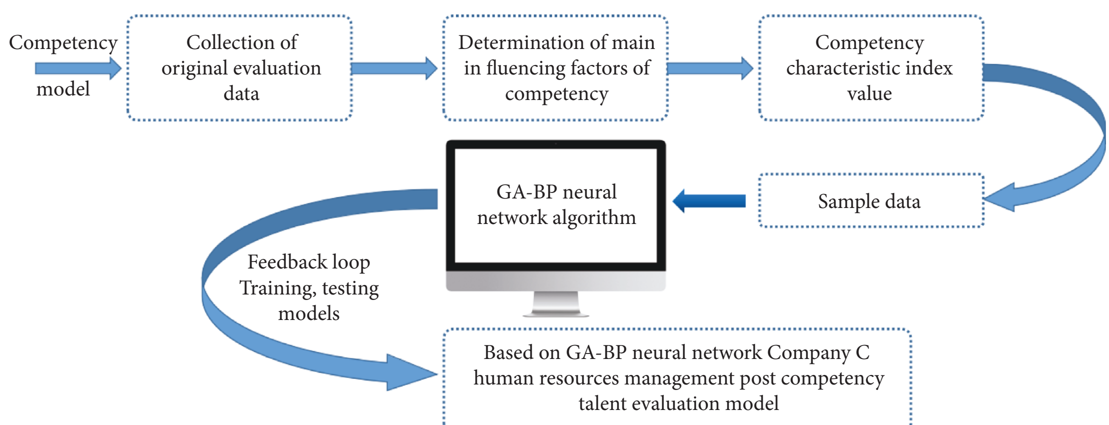


Figure 14: Based on GA-BP neural network Company *C* human resources management post competency model.

Table 3: Number of competency characteristic elements.

Source of characteristic elements	Total source	Number selected
Literature analysis	20	6
Questionnaire survey	50	10
Semi-structured interview	10	4
Total	80	20

includes the real number coding of the initial weights and thresholds of the network to obtain the initial population, and the fitness value of each individual is calculated by the fitness function. On this basis, the selection, crossover, and mutation operations are carried out. Finally, the corresponding individuals of the optimal fitness value, namely the new network weights and thresholds, are found in the continuous iterative evolution process. At this time, the optimized network starts to learn the training process and finally predicts the output value of the function. Establishment of human resource competency model based on BP optimized by genetic algorithm is shown in Figure 14.

There are 6 characteristic elements from literature analysis, 10 characteristic elements from questionnaire survey, and 5 characteristic elements from behavioral event interview. The total number of selected characteristic elements is 20 and is shown in Table 3.

Table 4: Division of training set and test set.

Data set	Sample number	Excellent performance	Other performance
Training sets	90	53	37
Testing set	20	13	7

Table 5: Confusion matrix of GA-BP neural network model.

Model		Excellent performance	Other performance	Prediction
GA-BP	Excellent performance	92.31	7.69	91.16
	Other performance	10.00	90.00	

The sample data set is divided into training set and test set, in which the proportion of training set is 80%, and the proportion of test set is 20%. The same principle is shown in Table 4.

The GA-BP neural network model is applied to the human resources management position of Company *C*. As shown in Table 5, it is found that the prediction accuracy of the model is very high, and the prediction accuracy reaches 91.16%. The competency model of HRM based on machine learning can scientifically, objectively, effectively, and quickly identify the employees with excellent performance from the candidate employees and can also effectively identify whether the employees have the potential to some extent.

6. Conclusion

With the rapid development of modern enterprises, middle managers, as the connecting hub between the grass roots and senior levels of enterprises, have been paid more and more attention to their competence and job performance. The research on the relationship between competency and job performance has attracted more and more attention. Managers' competency is not immutable. Different times and different business environments have great differences in the requirements of enterprise managers' ability. The actual environment has different requirements for managers' competency. As managers, they must dynamically adjust to meet the actual needs in order to remain invincible in the fierce competition.

This paper mainly studies the establishment of HRM post competency model based on machine learning and uses the GA-BP algorithm in machine learning to construct the HRM post competency model of Company *C*. The empirical research on the internal relationship between managers' competence and job performance shows that if job performance is different, their competence characteristics must be different. In this paper, 110 employees of management positions in Company *C* are completed and divided into experimental data sets by questionnaire survey and semi-structured interview. Taking the construction of management position competency model of Company *C* as the goal, the GA-BP neural network classification algorithm is used to construct the competency model of HRM positions based on GA-BP neural network. From the evaluation results, the HRM post competency method based on machine learning can minimize the impact of human factors and avoid the uncertainty of traditional methods. The quantitative analysis of GA-BP neural network model has obtained 91.16% classification accuracy, which verifies the application prospect of machine learning in the construction of auxiliary competency model.

Data Availability

The experimental data 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 to report regarding the present study.

Acknowledgments

The study was supported by the Key Teaching and Research Projects in Colleges and Universities in Anhui Province: Practical Case Driven Teaching Practice of Basic Courses of Business Administration—Taking "Principles Of Management" as an Example (2021jyxm0643) and Anhui University Humanities and Social Sciences Key Research Project: Digital Transformation Innovation and Practice of Enterprise Human Resources Management in the Digital Economy Background (SK2022A0756).

References

[1] V. A. Degterev and I. A. Larionova, *A COMPETENCY MODEL IN THE COMPANY IS A TOOL OF HUMAN RESOURCE MANAGEMENT*, 2015.

[2] G. M. McEvoy, J. C. Hayton, A. P. Warnick, T. V. Mumford, S. H. Hanks, and M. J. Blahna, "A competency-based model for developing human resource professionals," *Journal of Management Education*, vol. 29, no. 3, pp. 383–402, 2005.

[3] M. A. Yan, H. Wang, and S. O. Management, "A study of competency model of human resource managers in service industry:based on vision of human resource management in double perspective," *Commercial Research*, 2018.

[4] L. X. Zhang, L. L. Min, and Y. R. Fang, "A human resources management model based on competency," *Journal of Northeastern University*, 2006.

[5] S. M. Zhao, "A study of human resource management based on competency model," *Economic Management*, 2007.
[6] V. Loia, C. D. Maio, G. Fenza, F. Orciuoli, and S. Senatore, "An enhanced approach to improve enterprise competency management," in *Proceedings of the 2010, IEEE International Conference on Fuzzy Systems*, pp. 18–23, Barcelona, Spain, July, 2010.
[7] K. C. Wooten and M. Elden, "Cogenerating a competency based HRM degree: a model and some lessons from experience," *Journal of Management Education*, vol. 25, no. 2, pp. 231–257, 2001.
[8] M. A. Thomsenová, *Competency Model as an Effective Tool for the Human Resource Management*, Journal of Quality and Environmental Studies, 2017.
[9] S. H. Wang, *Foundation of Human Resource Management--Competency Model*, Journal of Capital University of Economics and Business, China, 2005.
[10] F. Trichet, M. Bourse, M. Leclere, and E. Morin, "Human resource management and semantic Web technologies," in *Proceedings.of the 2004 International Conference on Information and Communication Technologies: From Theory to Applications*, .
[11] L. I. Nai-Wen and P. Z. Zhi-Jiang, "Study on post safety competency model of special operation personnel in coal enterprise[J]," *China Safety Science Journal*, vol. 18, no. 1, pp. 14–19, 2008.
[12] R. M. Akbar and A. Akbar, *Human Resource Management Competency Model for Pakistani Banking Companies*, Academic Journals, 2012.
[13] S. V. Samarin, "Competency-based approach as a basis for controlling the development of innovations in a coal mining enterprise," *News of the Ural State Mining University*, vol. 1, no. 2, pp. 210–217, 2020.
[14] Y. Ma and H. Du, *Data Is Becoming a Core Competency of Enterprises[J]*, Springer Books, New york, USA, 2022.
[15] A. A. Mamun, S. A. Fazal, and N. Wan, "Entrepreneurial traits, competency, performance, and sustainability of microenterprises in kelantan, Malaysia[J]," *International Journal of Asian Business and Information Management*, vol. 12, 2021.
[16] C. S. Shen, *Human Resource Management Professionals Training Method Discuss—Based on Competency Model*, Science Technology and Industry, 2010.
[17] N. M. Salleh, A. K. B. H. Rosline, and D. K. A. Budin, "Human resource management roles & skills shortages in Malaysian organisations," *Open Journal of Social Sciences*, vol. 03, no. 03, pp. 219–226, 2015.
[18] G. Yifei, "Research on Competency Model of Enterprise Human Resource Management," *International Conference on Economics*, vol. 2018, 2018.
[19] F. Fenling, "Research on Enterprise Human Resource Management Competency Model," in *Proceedings of the 2018 International Workshop on Advances in Social Sciences (IWASS 2018)*, pp. 547–550, Fuzhou, China, April 2018.
[20] Xiaopeng, R. Gong, and Liu, "Research on the application of competency model in personnel management of private enterprises," *Conference Series*, vol. 1621, 2020.
[21] Yu Meilin and X. Duo, *Research on the Human Resource Management Mode based on Competency Model*, vip, vol. 2015, , pp. 39–41, 2015.
[22] Q. Xue, "Study on human resource management model based on competency model," *Value Engineering*, 2013.
[23] T. U. Huo-Ming and J. Zhao, "The research of public sector human resource competency model construction," in *Proceedings of the International Conference on Public Administration*, 2012.
[24] C. You and Y. Li, "The study on the human resource management of the logistics enterprises based on the competency model," in *Proceedings of the 2006 Interntional Conferecne on Management of Logistics and Supply Chain*, p. 2006, China, 2006.
[25] M. Muratbekova-Touron and Maral, "Why a multinational company introduces a competency-based leadership model:a two-theory approach," *International Journal of Human Resource Management*, vol. 20, no. 3, pp. 606–632, 2009.
[26] S. Zhao, X. Sun, S. O. Business, and N. University, "CEOs' transformational leadership behavior, strategic human resource management and firm performance in small and medium size enterprises:the moderating role of HRM competency," *Nankai Business Review*, 2016.
[27] M. Yuan and S. O. Business, *Employee Competency Development System (ECDS): A System to Improve Human Resource Based on Competency Model*, Human Resources Development of China, China, 2015.
[28] Y. Wang, W. D. Liu, and Z. W. Wang, *Exploration on Frame of Teaching Staff Human Resource Management Based on Competency in Higher Schools*, China Higher Medical Education, China, 2011.
[29] B. V. Le, *The Development of a Process Management Model for Herman IT's Competency-Based Human Resource Management and Development Contents*, 2015.
[30] S. Bolz, S. Kustermann, H. Hildebrandt, and K. L. Kohler, "To cultivate virtue: competency development to aim at excellent project manager(education of project management)," *Journal of the Society of Project Management*, vol. 14, no. 2, pp. 37–42, 2012.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 7318948, 11 pages
https://doi.org/10.1155/2022/7318948

Research Article

CCMbAS: A Provably Secure CCM-Based Authentication Scheme for Mobile Internet

Yu Zhang,[1,2] **Guangmin Sun**,[1] **and Peng Zhai**[1,2]

[1]*Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China*
[2]*Institute of Mathematics and Computer Science, Jining University, Qufu 273155, China*

Correspondence should be addressed to Yu Zhang; b201602005@jnxy.edu.cn

Received 14 July 2022; Revised 17 September 2022; Accepted 21 September 2022; Published 12 October 2022

Academic Editor: Dragan Pamučar

To improve the security of authentication system and strengthen privacy protection in mobile Internet environment, this paper proposes a provably secure Chebyshev chaotic map (CCM)-based authentication scheme (CCMbAS). The proposed scheme transformed the traditional public key of Chebyshev chaotic map into a private key and combined two private keys to compute a one-time key used to encrypt authentication information. The scheme is verified using security review of BAN logic and ProVerif simulation tool. The verification results confirm that the scheme is well secured against all existing security threats. Compared with similar schemes, the proposed scheme is more efficient and secure. The security analysis shows that the proposed scheme can fulfil secure demands and ensure the security of user's information in mobile Internet environment.

1. Introduction

Mobile Internet is the Internet and service that takes mobile network as access network. It has the characteristics of openness and complexity. With the rapid upgradation of mobile communication and the wide application of intelligent terminal, the application services provided by mobile Internet are more and more widespread. However, the network environment is more and more complex. Identity authentication is the first defender of information system, which can guarantee the security of system data and user information in complex network environment. It plays a key role in application system.

Aiming at solving the security threat of identity authentication system and protecting user's privacy information, Zhu et al. [1] proposed a biometrics-based multiserver key agreement scheme (BbKAS) on chaotic map cryptosystem. The encryption key of the scheme is not secure enough because the attacker can obtain encrypting key and crack encrypted information with dictionary attack of the intercepting information. Jiang et al. [2] proposed a new three-factor scheme. Ali et al. [3] proposed a three-factor identity authentication scheme based on RSA encryption algorithm.

To reduce the computational cost, Dong et al. [4] proposed a biometric verification-based authentication scheme (BVbAS) using Chebyshev chaotic mapping. The design of the scheme is unreasonable because the registry centre must provide all concerned information about all users and servers to each other before they request authentication. The design may result in a sharp increase in the communication cost of system. Otherwise, the authentication cannot be performed.

In general, the schemes can be classified into five groups in terms of the underlying intractability problem: based on discrete-logarithm problem [5–8], based on pairing [9, 10], based on chaotic map [11–13], based on integer-factorization problem [14], and based on hash function [15–18]. Among them, schemes based on elliptic curve bilinear pairings, such as a robust provable-secure privacy-preserving authentication protocol (PpAP) for Industrial Internet of Things [10], usually require large computation cost. Chaotic cryptography has become increasingly popular due to its lower computational complexity and higher asymmetric key security [19]. In view of the computing and security advantages of chaotic cryptography, CCMbAS is proposed to solve the problems of the above schemes.

2. Related Theoretical Knowledge

2.1. Fuzzy Extractor. In order to solve the contradiction between the variability of extracted biometric feature data and the input stability of traditional cryptography, Dodis proposed an algorithm of fuzzy extractors [20]. The algorithm could keep the numerical consistency of output results in the case of slight differences in the extracted biometric features.

Fuzzy extractor includes generation function $Gen(\cdot)$ and reproduction function $ep(\cdot)$, and $Gen(\cdot)$ is a probabilistic generating function. When the user inputs a biometric feature BIO_i, the function will generate a random string b_i limited to a fixed length ($b_i \in \{0,1\}^m$) and a public reproduction parameter P_i (as an auxiliary string), namely, $Gen(BIO_i) = (b_i, P_i)$, and $Rep(\cdot)$ is a deterministic reproduction function which can reproduce the biometric key according to the input biometric feature BIO_i' and corresponding public reproduction parameter P_i. If the Hamming distance between BIO_i and BIO_i' is within the preset fault tolerance threshold, $Rep(BIO_i', P_i) = b_i$. When $Gen(\cdot)$ and $Rep(\cdot)$ run in polynomial time, fuzzy extractor is very efficient. Without the aid of the original biometric feature, the biometric key cannot be reproduced with only the public reproduction parameter through calculation [4].

The application of fuzzy extractor can be effectively combined with cryptography in the field of authentication. In recent years, fuzzy extractor is used in many multi-factor authentication schemes [21–25].

2.2. Chebyshev Map

Definition 1. Chebyshev polynomial $\mathrm{Tn}(x)$ is the polynomial of n orders about x, where n is a natural number, $x \in [-1,1]$, and $T_n(x) = \cos(n * \arccos(x))$.

According to trigonometric transformation, Chebyshev polynomial iterative relation can be obtained as follows: $T_0(x) = 1$, $T_1(x) = x, \ldots, T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x)$.

Definition 2. The cryptosystem based on Chebyshev polynomial has the risk that session key is intercepted. In order to remedy the security defect, Zhang et al. [26] extended the domain of x from $x \in [-1,1]$ to $x \in (-\infty, +\infty)$ in 2008, that is, extended Chebyshev polynomial:

$$T_n(x) \equiv \left(2xT_{n-1}(x) - T_{n-2}(x)\right) \bmod p, \tag{1}$$

where ≥ 2, $x \in (-\infty, +\infty)$, and p is a big prime number. It still has the semigroup property:

$$T_r(T_s(x)) \equiv T_{rs}(x) \equiv T_s(T_r(x)) \pmod p. \tag{2}$$

Definition 3. It is a very hard problem of discrete logarithm to get r with the value x and y ($T_r(x) = y$). It is impossible in theory.

Definition 4. It is Diffie–Hellman problem to compute $T_{rs}(x)$ using $T_r(x)$ and $T_s(x)$. It is also impossible in theory.

Table 1: Symbol definitions.

Symbol	Definition
$(x, T_k(x))$	Public key of CA
k	Private key of CA
N	A large prime number
ID_i	The user's identity
PW_i	The user's password
ID_j	The server's identity
$Gen(\cdot)$	Generation function of fuzzy extractor
$Rep(\cdot)$	Reproduction function of fuzzy extractor
α_i	Public reproduction parameter
B_i	Biometric feature
b_i	Biometric key of the user U_i
$h(\cdot)$	Hash function
r_i, r_j	Random number
r_a, r_b, r_c	Random number
Δt	Valid time interval
$\oplus$	XOR
$\Vert$	Concatenation
t_i, t_j, t_1, t_2, t_3	Current time

3. Scheme Design

The authentication system consists of three parts: certificate authority (CA), user terminal, and server.

CA includes registration module, important data management module, and user authority management module.

User terminal includes registration module, biometric feature authentication module, password verification module, important data management module, and application interface module.

Server includes registration module, key agreement module, important data management module, and application platform interface module.

3.1. Symbol Definitions. The symbol definitions of the proposed scheme are shown in Table 1.

3.2. System Settings. CA first generates a private key k (assuming that the key is absolutely secure), then selects a random string x, and generates $T_k(x)$ through Chebyshev chaos map. The public key is published. The private key is hidden.

3.3. Registration Phase

3.3.1. Server Registration Phase. The registration process of the server is shown in Figure 1.

Step 1. The server S_j selects a unique identity ID_j and sends ID_j and the current time t_j to CA via secure channel.

Step 2. After receiving the registration request message $\{ID_j, t_j\}$ from S_j, CA first checks whether the time t_j exceeds the maximum valid time interval Δt or not. If the time interval meets the requirements, CA then checks whether the identity ID_j of the server is registered already or not. If the identity ID_j is registered

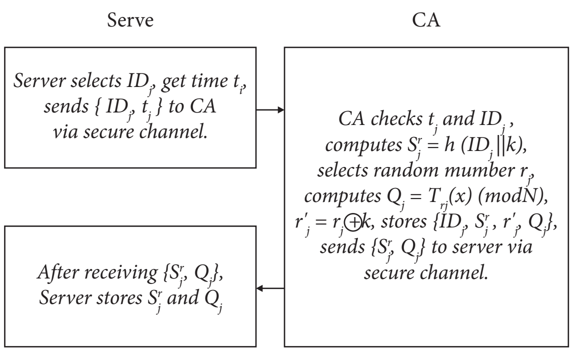


Figure 1: Server registration phase.

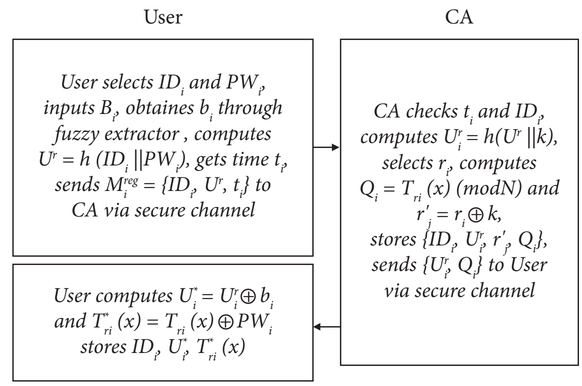


Figure 2: User registration phase.

already, CA rejects the registration request. Otherwise, CA computes $S_j^r = h(ID_j \| k)$, selects a random number r_j, and computes the key $Q_j = T_{r_j}(x)$ and $r_j' = r_j \oplus k$. The key Q_j cannot be published. Then, CA stores the data $\{ID_j, S_j^r, r_j', Q_j\}$ in the important data management module and sends the message $\{S_j^r, Q_j\}$ to server S_j via secure channel.

Step 3. After receiving $\{S_j^r, Q_j\}$, the server S_j stores them in the important data management module.

3.3.2. User Registration Phase. The process is shown in Figure 2.

Step 1. The user U_i selects a unique identity ID_i and W_i. Then, the biometric sample B_i is input through the sensor of biometric authentication module. The biometric key b_i is obtained by using the fuzzy extractor and its public reproduction parameter α_i. That is, $(B_i) = (b_i, \alpha_i)$.

Step 2. The user U_i computes $U^r = h(ID_i \| PW_i)$, gets current time t_i, and then sends the registration information $M_i^{reg} = \{ID_i, U^r, t_i\}$ to CA via secure channel.

Step 3. After receiving $M_i^{reg} = \{ID_i, U^r, t_i\}$, CA first checks whether the time t_i exceeds the maximum time interval Δt or not. If it exceeds the maximum time interval, CA rejects the user's request. If the result is eligible, CA checks whether the identity is registered already or not. CA forbids the user to register again. If the identity ID_i is not registered, CA calculates $U_i^r = h(h(ID_i \| PW_i) \| k)$, selects a random number r_i, and calculates the key $Q_i = T_{r_i}(x)$ (the public key transformed into private key) and $r_i' = r_i \oplus k$. Then, CA stores $\{ID_i, U_i^r, r_i', Q_i\}$ in the important data management module and sends $\{U_i^r, Q_i\}$ to user U_i via the secure channel.

Step 4. After receiving $\{U_i^r, Q_i\}$, user U_i calculates $U_i^* = U_i^r \oplus b_i$ and $T_{r_i}^*(x) = T_{r_i}(x) \oplus PW_i$ and then stores the information $ID_i, U_i^*, T_{r_i}^*(x)$ in the important data management module.

3.4. Login, Authentication, and Key Agreement Phase. If the user requests to login to the server, successfully authenticates his identity, and accesses resources, he/she must perform the steps shown in Figure 3.

Step 1. The user U_i inputs biometric feature through the sensor of biometric feature authentication module and uses fuzzy extractors and its public reproduction parameter α_i to obtain biometric key b_i' by calculating function $Rep(B_i', \alpha_i) = b_i'$. When the Hamming distance from B_i' and B_i is only less than the default tolerance threshold value, the equation $b_i' = b_i$ can be set up and the user U_i can pass biometric feature authentication. Then, the user U_i calculates $U_i^r = U_i^* \oplus b_i$.

Step 2. The user U_i inputs the correct password W_i and calculates the equation $T_{r_i}(x) = T_{r_i}^*(x) \oplus PW_i$ so as to pass password authentication.

Step 3. The user U_i selects a random number r_a as the temporary private key, calculates $k_1 = T_{r_a}(T_{r_i}(x))$, $M_i = h(ID_i \| ID_j \| U_i^r) \oplus k_1$, and $M_1 = \{M_i, ID_i, T_{r_a}(x)\}$, obtains the current time t_1, and sends the message $\{M_1, t_1\}$ to server S_j via public network. The key k_1 is the one-time key generated by calculation after the combination of the private key r_a and Q_i,

Step 4. After receiving $\{M_1, t_1\}$, server S_j first checks whether the time t_1 exceeds the maximum time interval Δt or not. If it exceeds the maximum time interval, the server rejects the user's request. If the result is eligible, server S_j selects a random number r_b as the temporary private key, calculates $k_2 = T_{r_b} T_{r_j}(x)$, $M_j = h(ID_i \| ID_j \| S_j^r) \oplus k_2$, and $M_2 = \{M_j, ID_j, T_{r_b}(x), M_1\}$, then obtains the current time t_2, and sends $\{M_2, t_2\}$ to CA via public network. The key k_2 is the one-time key generated by calculation after the combination of the private key r_b and Q_j.

Step 5. After receiving $\{M_2, t_2\}$, CA first checks whether the time t_2 exceeds the maximum time interval or not. If it exceeds the maximum time interval, CA rejects the request. If it is eligible, CA calculates $r_j = r_j' \oplus k$, $k_2' = T_{r_j}(T_{r_b}(x))$, $M_j' = M_j \oplus k_2'$, and $h(ID_i \| ID_j \| S_j^r)$ and verifies $h(ID_i \| ID_j \| S_j^r) = M_j'$. If the result is equal, CA authenticates the server S_j.

Step 6. Based on $M_1 = \{M_i, ID_i, T_{r_a}(x)\}$, CA calculates $r_i = r_i^{\oplus} \oplus k$, $k_1' = T_{r_i}(T_{r_a}(x))$, $M_i' = M_i \oplus k_1'$, and $h(ID_i \| ID_j \| U_i^r)$ and verifies $h(ID_i \| ID_j \| U_i^r) = M_i'$. If the result is not equal, CA stops authentication. If the result is equal, CA can authenticate the user U_i which applies for accessing the server S_j.

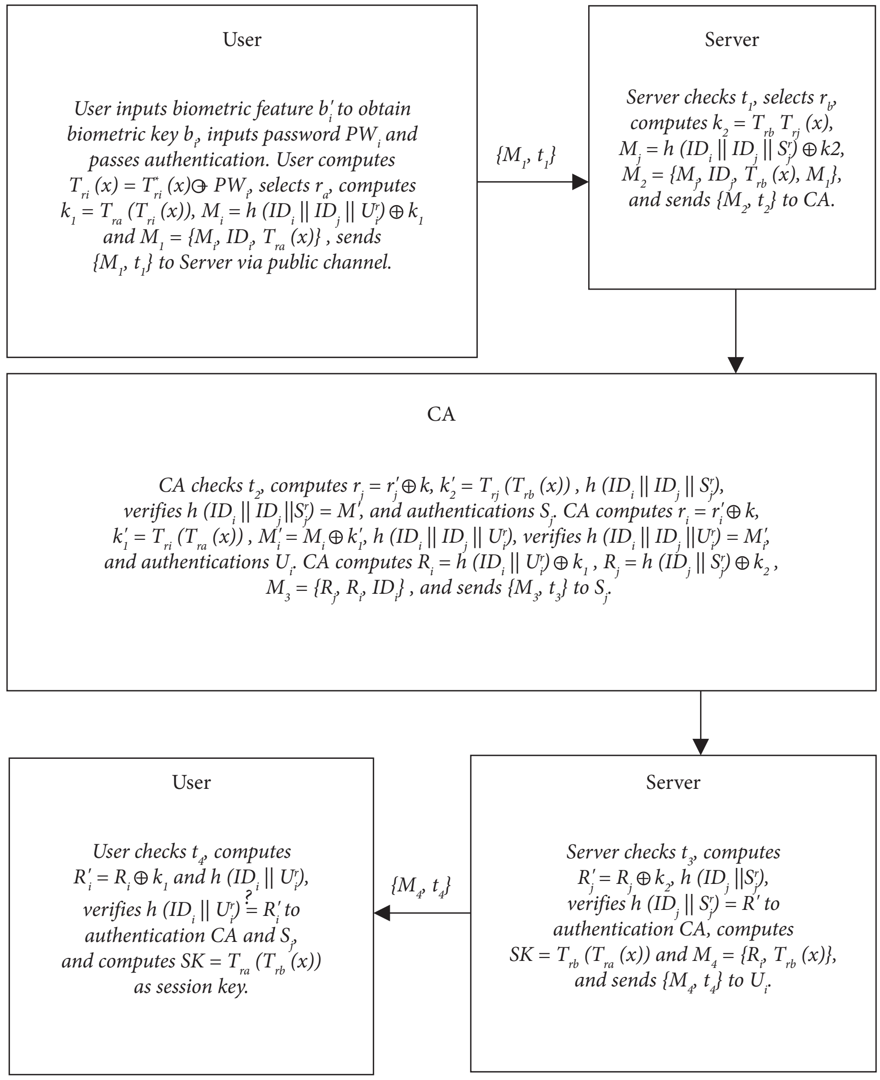


FIGURE 3: Login, authentication, and key agreement phase.

Step 7. CA calculates $R_i = h(ID_i \| U_i^r) \oplus k_1$, $R_j = h(ID_j \| S_j^r) \oplus k_2$, and $M_3 = \{R_j, R_i, ID_i\}$, obtains the current time t_3, and sends $\{M_3, t_3\}$ to the server S_j via public network.

Step 8. After receiving $\{M_3, t_3\}$, server S_j first checks whether the time t_3 exceeds the maximum time interval or not. If it exceeds the maximum time interval, the server will discard the received information. If the result is eligible, the server fetches R_j from M_3 and calculates $R_j' = R_j \oplus k_2$ and $h(ID_j \| S_j^r)$. Then, the server verifies $h(ID_j \| S_j^r) = R_j'$. If the result is not equal, the server stops authentication. If the result is equal, the server can authenticate CA. Then, the server calculates the session key $SK = T_{r_b}(T_{r_a}(x))$ which will be used with the user U_i, gets the current time t_4, and sends $\{M_4, t_4\} M_4 = \{R_i, T_{r_b}(x)\})$ to the user U_i via public network.

Step 9. After receiving $\{M_4, t_4\}$, user U_i first checks whether the time t_4 exceeds the maximum time interval. If it oversteps the maximum time interval, user U_i will discard the received information. If the result is eligible, the user calculates $R_i' = R_i \oplus k_1$ and $h(ID_i \| U_i^r)$. Then, the user verifies $h(ID_i \| U_i^r) \overset{?}{=} R_i'$. If the result is not equal, the user stops authentication. If the result is equal, the user can authenticate CA and the server S_j. Then, the user calculates the session key $SK = T_{r_a}(T_{r_b}(x))$ which will be used with the server S_j.

3.5. Password Change. If the user wants to change the password, the authentication must be completed of the user on the terminal first. Then, the user changes the password according to the steps of registration. The corresponding information stored in the user terminal and the CA can be updated.

3.6. Identity and Biometric Feature Change. If the user needs to change the identity, the identity can be changed by the similar steps of the password change. If the user needs to change the biometric feature, the biometric feature can be changed after the terminal authenticates the legitimate user.

4. Scheme Security

4.1. Security Analysis

4.1.1. Key Security. The user's biometric key b_i is generated by fuzzy extractor, so the attacker cannot get the user's biometric key through the fuzzy extractor without the user's biometric feature. In the proposed scheme, a double key combined encryption mechanism is designed. For example, the key k_1 is the one-time key generated by calculation after the combination of the private key r_a and Q_i. Because the one-time key k_1 is newly generated, the information encrypted with k_1 is difficult to crack. The user calculates $F_i^* = F_i \oplus b_i$, $U_i^* = U_i^r \oplus b_i$, and $T_{r_i}^*(x) = T_{r_i}(x) \oplus PW_i$ in order to hide U_i^r and $T_{r_i}(x)$ and then stores the information $ID_i, U_i^*, T_{r_i}^*(x)$ into the important data management module. Suppose that attacker can obtain the data stored in the user's terminal, and the encrypted information cannot be decrypted. Therefore, the information b_i, U_i^r, and $T_{r_i}(x)$ cannot be leaked or stolen.

4.1.2. Terminal Lost Attack. If the terminal device is lost, authentication requires not only the correct biometric feature information but also the correct password. The user's secret information stored in the terminal device is encrypted data. The attacker cannot provide the correct information and decrypt the stored secret information. Therefore, the system can ensure the security of the secret information in the case of terminal device loss.

4.1.3. Password Guessing Attack. In this scheme, user authentication includes two steps. If user wants to login successfully, the biometric feature and password must be correct. Without biometric feature of the legitimate user, the attacker cannot pass the initial biometric feature authentication. Therefore, the attacker cannot proceed the second step, password authentication. The shared session key generated temporarily is new and different each time. Attacker cannot guess the session key. Therefore, authentication system can effectively avoid password guessing attack.

4.1.4. Impersonation Attack. Because user authentication includes biometric feature and password, the attacker cannot pass through password authentication when he initiates impersonation attack in case of obtaining the user's biometric feature. If an attacker impersonates a legitimate user or server to transmit information, the user, server, or CA can identify the authenticity of the sender through calculation and the impersonation attack information.

4.1.5. Eavesdropping Attack. The scheme uses the randomness of hash function value to hide the authentication information transmitted in the public network and uses the one-off key randomly generated by Chebyshev chaos map to encrypt the authentication information. Under the premise of this double security, the attacker cannot get useful information by eavesdropping on the messages transmitted in the public network.

4.1.6. Denial-of-Service Attack. Within a certain time period, CA does not allow users using the same ID to apply for registration. Therefore, CA can avoid excessive consumption of server resources and effectively defend against denial-of-service attack.

4.1.7. Man-in-the-Middle Attack. Even if information of legitimate users or servers is intercepted and tampered by attacker, the attacker cannot pass the inspection and authentication of users or servers. Therefore, the attacker cannot steal the content from the information of user and server by attack.

4.1.8. Replay Attack. Time information is added to the transmitted information in the proposed scheme, which has the function of time stamp and can effectively avoid replay attack.

4.1.9. Privileged Insider Attack. In this scheme, CA uses its own private key to perform XOR operation to the key of user or server to hide the important information. The password of user is protected by one-way hash function when applying for registration and authentication, which also achieves the purpose of hiding important information. In this way, privilege attack can be effectively avoided.

4.1.10. Forward Security. The encryption key of authentication information is one-off in the process of certification. The sharing session key is also one-off after key agreement. The scheme has dual security by hiding and encryption. The attacker cannot crack the former session key.

4.1.11. Mutual Authentication. In the proposed scheme, the shared session key calculated only by the legitimate user and server will be the same. Therefore, the scheme can realize mutual authentication among CA, user, and server. Meanwhile, the scheme can ensure the communication security between legitimate user and server.

The comparison results in terms of security are shown in Table 2.

4.2. BAN Logical Proof

4.2.1. BAN Logic. Among the cryptographic protocol formal verification methods, BAN logic proposed by Burrows et al. in 1989 is the well-known one [27]. BAN logic is a kind of modal logic based on belief, which mainly includes the following three processing objects: subject, key, and formula. P, Q, and R represent the subject variable. K represents the key variable. X and Y represent the formula variable. A and B represent the two common subjects. S is the authentication

Table 2: Security comparison.

Security issues	BbKAS [1]	BVbAS [4]	PpAP [10]	Proposed scheme
Forward security	×	√	√	√
Key security	×	√	√	√
Mutual authentication	√	×	√	√
Key agreement	√	×	√	√
Privilege insider attack	√	√	√	√
Resist man-in-the-middle attack	√	√	√	√
Resist replay attack	√	√	×	√
Resist password guessing attack	×	√	√	√
Resist terminal loss attack	√	√	√	√

Table 3: The syntax and semantics of the BAN logical component.

Symbol	Definition
$P\| \equiv X$	Subject P believes X.
$P \lhd X$	Subject P receives the message X.
$P\| \sim X$	Subject P has sent out the message X.
$P\|\Rightarrow X$	Subject P has jurisdiction over X.
$\#(X)$	X is fresh.
$P \overset{K}{\leftrightarrow} Q$	K is the shared key of subjects A and B, which is unknown to other subjects.
$\longrightarrow^{K} P$	K is the public key of the subject P. The other subjects do not know the corresponding private key K^{-1}.
$\longrightarrow^{K} -1\ P$	K^{-1} is the private key of the subject P.
$P \overset{X}{\rightleftharpoons} XQ$	X is the shared secret between subjects P and Q, which is unknown to other subjects.
$\{X\}_K$	The ciphertext is obtained by encrypting X with the key K.
$\langle X\rangle_Y$	A cascade between message X and secret Y can prove that the message $\langle X\rangle_Y$ is sent by a certain subject.

Table 4: BAN logic inference rules.

Sequence number	Rules
R1	$P\| \equiv Q \overset{K}{\longleftrightarrow} P, P\lhd\{X\}_K \vdash P\| \equiv Q\| \sim X$
R2	$P\| \equiv \ \longrightarrow^{K} Q, P\lhd\{X\}_{K^{-1}} \vdash P\| \equiv Q\| \sim X$
R3	$P\| \equiv P \overset{Y}{\rightleftharpoons} YQ, P\lhd\{X\}_Y \vdash P\| \equiv Q\| \sim X$
R4	$P\| \equiv \#(X), P\| \equiv Q\| \sim X \vdash P\| \equiv Q\| \equiv X$
R5	$P\| \equiv Q \| \Rightarrow X, P\| \equiv Q\| \equiv X \vdash P\| \equiv X$
R6	$P\lhd(X,Y) \vdash P\lhd X$
R7	$P\lhd\langle X\rangle_Y \vdash P\lhd X$
R8	$P\| \equiv P \overset{K}{\leftrightarrow} Q, P\lhd\{X\}_K \vdash P\lhd X$
R9	$P\| \equiv \ \longrightarrow^{K} P, P\lhd\{X\}_K \vdash P\lhd X$
R10	$P\| \equiv P \longrightarrow^{K} Q, P\lhd\{X\}_{K^{-1}} \vdash P\lhd X$
R11	$P\| \equiv \#(X), \vdash P\| \equiv \#(X,Y)$
R12	$P\| \equiv X, P\| \equiv Y, \vdash P\| \equiv (X,Y)$
R13	$P\| \equiv (X,Y), \vdash P\| \equiv X$
R14	$P\| \equiv Q\| \equiv (X,Y), \vdash P\| \equiv Q\| \equiv X$
R15	$P\| \equiv Q\| \sim (X,Y), \vdash P\| \equiv Q\| \sim X$
R16	$P\| \equiv R \overset{K}{\longleftrightarrow} R', \vdash P\| \equiv R' \overset{K}{\leftrightarrow} R$
R17	$P\| \equiv Q\| \equiv R \overset{K}{\longleftrightarrow} R', \vdash P\| \equiv Q\| \equiv R' \overset{K}{\longleftrightarrow} R$
R18	$P \| \equiv R \overset{X}{\rightleftharpoons} XR', \vdash P \| \equiv R' \overset{X}{\rightleftharpoons} XR$
R19	$P \| \equiv Q \| \equiv R \overset{X}{\rightleftharpoons} XR', \vdash P \| \equiv Q \| \equiv R' \overset{X}{\rightleftharpoons} XR$

server. K_{ab}, K_{ac}, and K_{bc} represent the specific shared key. K_a, K_b, and K_c represent the specific public key. K_a^{-1}, K_b^{-1}, and K_c^{-1} represent the specific secret key. N_a, N_b, and N_c represent the temporary value. $h(X)$ represents the irreversible hash function of X.

(1) The syntax and semantics of the BAN logical component.

The syntax and semantics of the BAN logical component are shown in Table 3.

(2) BAN logic inference rules.

Message meaning rules: R1–R3. Nonce verification rule: R4. Jurisdiction rule: R5. Seeing rules: R6–R10. Freshness rules: R11. Belief rules: R12–R15. Key and secret rules: R16–R19. BAN logic inference rules are shown in Table 4.

4.3. Scheme Security Proof

(1) Initialization $S_j \lhd T_{r_a}(x)$, $U_i \lhd T_{r_b}(x)$

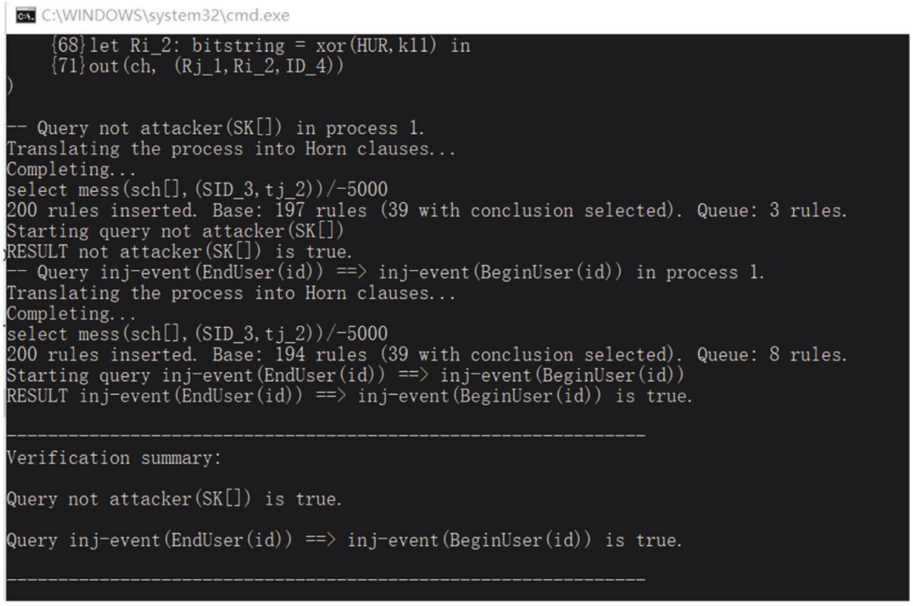


Figure 4: Performance result of ProVerif code.

(2) Establish security goals G1. $S_j| \equiv U_i| \equiv T_{r_a}(x)$, G2. $U_i| \equiv S_j| \equiv T_{r_b}(x)$, G3. $S_j| \equiv T_{r_a}(x)$, G4. $U_i| \equiv T_{r_b}(x)$, G5. $S_j| \equiv U_i| \equiv (T_{r_a}(x), T_{r_b}(x))$, G6. $U_i| \equiv S_j| \equiv (T_{r_a}(x), T_{r_b}(x))$

(3) Protocol formalization

F1. $U_i| \equiv T_{r_a}(x)$, $U_i|\Rightarrow T_{r_a}(x)$, F2. $S_j| \equiv T_{r_b}(x)$, $S_j|\Rightarrow T_{r_b}(x)$, F3. $S_j| \equiv U_i| \sim T_{r_a}(x)$, $S_j| \equiv \#(T_{r_a}(x))$, F4. $U_i| \equiv S_j| \sim T_{r_b}(x)$, $U_i| \equiv \#(T_{r_b}(x))$, F5. $S_j| \equiv U_i|\Rightarrow T_{r_a}(x)$, F6. $U_i| \equiv S_j|\Rightarrow T_{r_b}(x)$, F7.$S_j| \equiv U_i \overset{SK}{\leftrightarrow} S_j$, F8. $U_i| \equiv S_j \overset{SK}{\leftrightarrow} U_i$,

(4) Concrete proof process

V1. According to the rule R4 and formalization F3, $\mathrm{S_j} | \equiv \mathrm{U_i} | \sim \mathrm{T_{r_a}}(\mathrm{x}), \mathrm{S_j} | \equiv \#(\mathrm{T_{r_a}}(\mathrm{x})) \vdash \mathrm{S_j} | \equiv \mathrm{U_i} | \equiv \mathrm{T_{r_a}}(\mathrm{x})$ can be got. Therefore, the goal G1 is true.

V2. In the same way of V1 above, according to the rule R4 and formalization F4, the goal G2 is true.

V3. According to the rule R5 and formalization F5, $\mathrm{S_j} | \equiv \mathrm{U_i} | \equiv \mathrm{T_{r_a}}(\mathrm{x}), \mathrm{S_j} | \equiv \mathrm{U_i} | \Rightarrow \mathrm{T_{r_a}}(\mathrm{x}) \vdash \mathrm{S_j} | \equiv \mathrm{T_{r_a}}(\mathrm{x})$ can be obtained. Therefore, the goal G3 is true.

V4. In the same way of V3 above, according to the rule R5 and formalization F6, the goal G4 is true.

V5. According to goal G3, formalization F2, and rule R12, $S_j| \equiv T_{r_a}(x), S_j| \equiv T_{r_b}(x) \vdash S_j| \equiv (T_{r_a}(x), T_{r_b}(x))$ can be obtained. Therefore, the goals G5 and G6 are true.

Basing on the BAN logic proof, the proposed authentication scheme can achieve the predetermined security goal, which proves that the scheme is secure.

4.4. ProVerif Verification

4.4.1. ProVerif Code

```
( * ----------------channel-------------------- * )
free sch: channel [private]. ( * ---secure channel---- * )
free ch: channel. ( * ---unsecure channel---- * )
( * ----------------variable and constants-------------------- * )
free ID: bitstring. ( * ---User ID---- * )
const SID: bitstring. ( * ---Application server ID---- * )
const x: bitstring. ( * ---Seed for Chebyshev chaotic map ---- * )
const pw: bitstring [private]. ( * ---password of user---- * )
free treg: bitstring. ( * ---the time of registration---- * )
free s: bitstring [private]. ( * ---key of application server---- * )
free u: bitstring [private]. ( * ---key of user---- * )
free k: bitstring [private]. ( * ---key of CA---- * )
free B: bitstring [private]. ( * ---biometric of user---- * )
free w: bitstring [private]. ( * ---parameter of fuzzy extraction algorithm---- * )
```

```
free SK: bitstring [private]. (* ---the session key between user and application server-- *)
(* ----------------constructor-------------------- *)
fun H(bitstring): bitstring.
fun senc(bitstring, bitstring): bitstring.
fun T(bitstring, bitstring): bitstring. (* ---the Chebyshev chaotic map algorithm-- *)
fun xor(bitstring, bitstring): bitstring.
fun Concat(bitstring, bitstring): bitstring.
fun GEN(bitstring): bitstring. (* ---the GEN section of fuzzy extraction algorithm-- *)
fun REP(bitstring): bitstring. (* ---the REP section of fuzzy extraction algorithm-- *)
(* --------destructors&equations-------------------- *)
reduc forall m: bitstring, n: bitstring; sdec(senc(m,n),n) = m.
(* reduc forall a: bitstring, b: bitstring, x:bitstring; T(b,T(a,x)) = T(a,x) * T(b,x). ---the Chebyshev chaotic map algorithm-- *)
equation forall m: bitstring, n: bitstring; xor(xor(m,n),n) = m.
(* ----------------events --------------------- *)
event BeginUser(bitstring).
event EndUser(bitstring).
(* ----------------query --------------------- *)
query attacker(SK).
query id:bitstring; inj-event(EndUser(id)) = =>inj-event(BeginUser(id)).
(* ----------------process-------------------- *)
(* ----------------user process--------------------- *)
let user =
let (b) = GEN(B) in
let UR=H(Concat(pw, ID)) in
out(sch, (ID, UR, treg));
event BeginUser(ID);
in(sch,(URR:bitstring, Qi:bitstring)); (* -Input some data - *)
let URb = xor(URR, b) in
let Qip = xor(Qi, pw) in
new Bioaut: bitstring;
new pw: bitstring;
let b = REP(Bioaut) in
let Qi = xor(Qip, pw) in
new a: bitstring;
let tax = T(a,x) in
let k1 = T(u, tax) in
let HU=H(Concat(Concat(ID, SID), URR)) in
let MU = xor(HU, k1) in
out(ch,(MU, ID, tax));
in(ch,(Ri:bitstring, tbx:bitstring));
let Rii = xor(Ri, k1) in
let Hi = H(Concat(ID, URR)) in
if Hi = Rii then
let SK = T(a,tbx) in
0
).
(* ----------------Application Server AS process--------------------- *)
let AS =
!
(
new tj:bitstring;
out(sch,(SID,tj));
in(sch,(SR:bitstring, Qj:bitstring));
let SR=H(Concat(SID,k)) in
let Qj = T(s,x) in
in(ch,(MU: bitstring, ID: bitstring, tax: bitstring));
new b:bitstring;
let tbx = T(b,x) in
let k2 = T(s, tbx) in
let HS=H(Concat(Concat(ID, SID), SR)) in
let MS = xor(HS, k2) in
out(ch,(MS, SID, tbx, MU, ID, tax));
in(ch,(Rj: bitstring, Ri: bitstring, ID: bitstring));
let Rjj = xor(Rj, k2) in
let Hj = H(Concat(SID, SR)) in
if Hj = Rjj then
let SK = T(b,tax) in
out(ch,(Ri, tbx));
event EndUser(ID);
0
).
(* ----------------Certificate Authority CA process--------------------- *)
let CA =
in(sch, (ID:bitstring, UR:bitstring, treg:bitstring));

new u: bitstring;
let URR=H(Concat (UR,k)) in
let Qi = T(u,x) in
let ui = xor(u,k) in
out(sch,(URR, Qi));
in(sch, (SID:bitstring, tj: bitstring));
new s: bitstring;
let SR=H(Concat(SID,k)) in
let Qj = T(s,x) in
```

Table 5: Execution time comparison.

Scheme	Registration phase	Login, authentication, and key agreement phase	Password change phase
BbKAS [1]	$3T_H + T_{Gen} = 808T_H$	$12T_H + 6T_C + 8T_E + T_{Gen} = 845T_H$	$2T_H + T_{Gen} = 807T_H$
BVbAS [4]	$9T_H + T_{Gen} = 814T_H$	$21T_H + T_{Rep} + 4T_C = 216T_H$	$8T_H + T_{Rep} + T_{Gen} = 1000T_H$
PpAP [10]	$11T_H + 9T_P + 4T_A = 1201T_H$	$17T_H + 7T_P + 7T_A = 1277T_H$	$6T_H + 4T_P + 2T_A = 538T_H$
Proposed scheme	$3T_H + 2T_C + T_{Gen} = 814T_H$	$8T_H + T_{Rep} + 6T_C = 207T_H$	$2T_H + 2T_C + 4T_E + T_{Rep} = 201T_H$

```
let sj = xor(s,k) in
out(sch,(SR,Qj));

in(ch,(MS:bitstring, SID:bitstring, tbx:bitstring, MU: bitstring, ID: bitstring, tax: bitstring));
let s = xor(sj,k) in
let k22 = T(s, tbx) in
let MSS = xor(MS,k22) in
let HUS=H(Concat(Concat(ID,SID),SR)) in
if HUS = MSS then
let u = xor(ui,k) in
let k11 = T(u, tax) in
let MUU = xor(MU,k11) in
let HSU=H(Concat(Concat(ID,SID),URR)) in
if HSU = MUU then
let HUR=H(Concat(ID,URR)) in
let Ri = xor(HUR,k11) in
let HSR=H(Concat(SID,SR)) in
let Rj = xor(HSR,k22) in
out(ch,(Rj, Ri, ID));
0.
process (user| AS |CA)
```

4.4.2. Performance Result. The performance result is shown in Figure 4. From the result, we can see that our scheme is secure.

5. Performance

5.1. Computation Cost. According to literature [1, 10, 28–31] and the measured consumption time of the relative algorithms of the proposed scheme on our Intel Core i5-3470 platform, the details are shown as follows.

T_X: XOR. Because XOR operation time is very small, it can be ignored. T_H: hash operation. The hash operation time is 0.6 ms. T_C ($T_n(x)$mo dP): Chebyshev chaotic map. Its operation takes twice the time of hash operation. T_{GEN}: the time of obtaining public parameters and feature key from biometric feature by fuzzy extractor algorithm. The time is 805 times that of hash operation. T_{Rep}: the time of regenerating the biometric key from biometric feature and public parameter by fuzzy extractor algorithm. The time is 187 times that of hash operation. T_E and T_D: symmetric encryption operation and symmetric decryption operation. The operations of both of them take twice the time of hash operation. T_{ECC}: encryption or decryption of elliptic curve public key cryptography. The time is 968 times that of hash operation. T_P: the time of an elliptic curve point multiplication. The time is 126 times that of hash operation. T_A: the time of an elliptic curve point addition. The time is 14 times that of hash operation. The comparison results of execution time of the related schemes are shown in Table 5.

As can be seen from Table 5, the computation cost of two phases is the lowest respectively in our proposed scheme. The proposed scheme is superior to the similar scheme in [1, 4, 10].

Table 6: Communication cost comparison.

Scheme	Login, authentication, and key agreement phase (bits)
BbKAS [1]	$6L_M + 15L_{ID} + 8L_T = 2528$
BVbAS [4]	$7L_H + 3L_T = 1504$
PpAP [10]	$4L_H + 6L_P + 4L_T = 3072$
Proposed scheme	$6L_H + 5L_{ID} + 4L_M + 4L_T = 2144$

5.2. Communication Cost. Referring to [1, 4, 10], we set the length as follows. L_{ID}: the length of identity is 32bits; L_H: the length of hash function is 160bits; L_M: the output size of chaotic maps is 128bits; L_T: the length of time is 128bits because it can be considered as a random number; L_E: the length of symmetric encryption/decryption is 128bits; L_P: the output size of an elliptic curve point $P = (P_x, P_y)$ is 320bits; and L_r: the length of random nonce is 128bits.

Here only the often executed login and authentication phases are considered for cost calculations. The comparison results of communication cost for the protocols are presented in Table 6. It can be observed that our scheme is more efficient than the schemes [1, 10] in communication cost.

6. Conclusion

In order to improve security of authentication system and strengthen protection for sensitive information and privacy of users, a provably secure Chebyshev chaotic map (CCM)-based authentication scheme is proposed. The scheme uses hash function to hide user information and uses fuzzy extractor to authenticate user biometric feature. Especially, the proposed scheme transformed the traditional public key of Chebyshev chaotic map into a private key and combined two private keys to compute a one-time key used to encrypt authentication information. The results verified by BAN logic and ProVerif simulation tool confirm that the scheme is well secured against all existing security threats. Compared with similar schemes, the proposed scheme is more efficient and secure. Therefore, the proposed scheme has great application value in high security demands scenarios such as mobile payment and contactless access control. In the future, we will continue to further study authentication schemes for more complex network environment.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

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

Acknowledgments

This research was funded by the National Natural Science Foundation of China (grant no. 41706201).

References

[1] H. Zhu, X. Hao, and Y. Zhang, "A biometrics-based multi-server key agreement scheme on chaotic maps cryptosystem," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 6, no. 2, pp. 211–224, 2015.

[2] Q. Jiang, F. Wei, S. Fu, J. Ma, G. Li, and A. Alelaiwi, "Robust extended chaotic maps-based three-factor authentication scheme preserving biometric template privacy," *Nonlinear Dynamics*, vol. 83, no. 4, pp. 2085–2101, 2016.

[3] A. Rifaqat and P. A. Kumar, "A secure and robust three-factor based authentication scheme using RSA cryptosystem," *International Journal of Business Data Communications and Networking*, vol. 13, no. 1, pp. 74–84, 2017.

[4] X. Dong, M. Li, and Y. Du, "A biometric verification-based authentication scheme using Chebyshev chaotic mapping," *Journal of Beijing University of Aeronautics and Astronautics*, vol. 45, no. 05, pp. 1052–1058, 2019.

[5] S. Zhou, Q. Gan, and X. Wang, "Authentication scheme based on smart card in multi-server environment," *Wireless Networks*, vol. 26, no. 2, pp. 855–863, 2020.

[6] N. Lwamo, L. Zhu, and C. Xu, "Suaa: a secure user authentication scheme with anonymity for the single & multi-server environments," *Information Sciences*, vol. 477, pp. 369–385, 2019.

[7] N. Kumar, K. Kaur, S. C. Misra, and R. Iqbal, "An intelligent RFID-enabled authentication scheme for healthcare applications in vehicular mobile cloud," *Peer-to-Peer Networking and Applications*, vol. 9, no. 5, pp. 824–840, 2016.

[8] D. He, N. Kumar, and J. H. Lee, "Privacy-preserving data aggregation scheme against internal attackers in smart grids," *Wireless Networks*, vol. 22, no. 2, pp. 491–502, 2015.

[9] S. Roy, A. K. Das, S. Chatterjee, N. Kumar, S. Chattopadhyay, and J. J. P. C. Rodrigues, "Provably secure fine-grained data access control over multiple cloud servers in mobile cloud computing-based healthcare applications," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 1, pp. 457–468, 2019.

[10] D. Rangwani, D. Sadhukhan, S. Ray, M. K. Khan, and M. Dasgupta, "A robust provable-secure privacy-preserving authentication protocol for Industrial Internet of Things," *Peer-to-peer Networking and Applications*, vol. 14, no. 3, pp. 1548–1571, 2021.

[11] Y. Luo, Y. Liu, J. Liu, X. Ouyang, Y. Cao, and X. Ding, "ECM-IBS: a Chebyshev map-based broadcast authentication for wireless sensor networks," *International Journal of Bifurcation and Chaos*, vol. 29, no. 09, p. 1950118, 2019.

[12] X. Y. Guo, D. Z. Sun, and Y. Yang, "An improved three-factor session initiation protocol using Chebyshev chaotic map," *IEEE Access*, vol. 8, pp. 111265–111277, 2020.

[13] R. I. Abdelfatah, N. M. Abdal-Ghafour, and M. E. Nasr, "Secure VANET authentication protocol (SVAP) using Chebyshev chaotic maps for emergency conditions," *IEEE Access*, vol. 10, pp. 1096–1115, 2022.

[14] M. Fueyo and J. Herranz, "On the efficiency of revocation in RSA-based anonymous systems," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 8, pp. 1771–1779, 2016.

[15] S. Barman, H. P. H. Shum, S. Chattopadhyay, and D. Samanta, "A secure authentication protocol for multi-server-based E-healthcare using a fuzzy commitment scheme," *IEEE Access*, vol. 7, pp. 12557–12574, 2019.

[16] X. Zhang and J. Liu, "Multi-factor authentication protocol based on hardware fingerprint and biometrics," *Netinfo Security*, vol. 20, no. 8, pp. 9–15, 2020.

[17] A. Aug, "Security of cryptocurrencies in blockchain technology: state-of-art, challenges and future prospects," *Journal of Network and Computer Applications*, vol. 163, p. 765, 2020.

[18] S. Roy, S. Chatterjee, and A. K. Das, "On the design of provably secure lightweight remote user authentication scheme for mobile cloud computing services," *IEEE Access*, p. 1, 2017.

[19] S. Namasudra and P. Roy, "A new secure authentication scheme for cloud computing environment," *Concurrency and Computation: Practice and Experience*, vol. 29, pp. e3864–20, 2017.

[20] Y. Dodis, R. Ostrovsky, L. Reyzin, and A. Smith, "Fuzzy extractors: how to generate strong keys from biometrics and other noisy data," *SIAM Journal on Computing*, vol. 38, no. 1, pp. 97–139, 2008.

[21] D. A. Kumar, "A secure and robust temporal credential-based three-factor authentication scheme for wireless sensor networks," *Peer-to-peer Networking and Applications*, vol. 9, no. 1, pp. 223–244, 2016.

[22] M. Wazid, A. K. Das, S. Kumari, X. Li, and F. Wu, "Design of an efficient and provably secure anonymity preserving three-factor user authentication and key agreement scheme for TMIS," *Security and Communication Networks*, vol. 9, no. 13, pp. 1983–2001, 2016.

[23] R. Zhang, Y. Xiao, S. Sun, and H. Ma, "Efficient multi-factor authenticated key exchange scheme for mobile communications," *IEEE Transactions on Dependable and Secure Computing*, vol. 16, no. 4, pp. 625–634, 2019.

[24] C. Herder, L. Ren, M. Dijk, M. D. Yu, and S. Devadas, "Trapdoor computational fuzzy extractors and stateless cryptographically-secure physical unclonable functions," *IEEE Transactions on Dependable and Secure Computing*, vol. 14, no. 1, pp. 65–82, 2017.

[25] R. Amin, R. S. Sherratt, D. Giri, S. H. Islam, and M. K. Khan, "A software agent enabled biometric security algorithm for secure file access in consumer storage devices," *IEEE Transactions on Consumer Electronics*, vol. 63, no. 1, pp. 53–61, 2017.

[26] L. Zhang, "Cryptanalysis of the public key encryption based on multiple chaotic systems," *Chaos, Solitons & Fractals*, vol. 37, no. 3, pp. 669–674, 2008.
[27] M. Burrows, M. Abadi, and R. Needham, "A logic of authentication," *Proceedings of the Royal Society of London*, vol. 426, pp. 233–271, 1989.
[28] M. Zhang, J. Zhang, and Y. Zhang, "Remote three-factor authentication scheme based on Fuzzy extractors," *Security and Communication Networks*, vol. 8, no. 4, pp. 682–693, 2015.
[29] V. Odelu, A. K. Das, and A. Goswami, "A secure biometrics-based multi-server authentication protocol using smart cards," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 9, pp. 1953–1966, 2015.
[30] C. C. Chang and C. Y. Sun, "A secure and efficient authentication scheme for E-coupon systems," *Wireless Personal Communications*, vol. 77, no. 4, pp. 2981–2996, 2014.
[31] C. C. Lee, "A simple key agreement scheme based on chaotic maps for VSAT satellite communications," *International Journal of Satellite Communications and Networking*, vol. 31, no. 4, pp. 177–186, 2013.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9875103, 1 page
https://doi.org/10.1155/2023/9875103

Retraction

Retracted: The Modal Analysis of Multifactor Coupling of Regional Industrial Innovation

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] Y. Liu and Q. Guo, "The Modal Analysis of Multifactor Coupling of Regional Industrial Innovation," *Mobile Information Systems*, vol. 2022, Article ID 1261278, 12 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 1261278, 12 pages
https://doi.org/10.1155/2022/1261278

Research Article

The Modal Analysis of Multifactor Coupling of Regional Industrial Innovation

Yubo Liu and Qian Guo

School of Economics, Southwest Minzu University, No. 16 South 4th Section First Ring Road, Wuhou, Chengdu, China

Correspondence should be addressed to Yubo Liu; liuyubo@stu.swun.edu.cn

Received 25 August 2022; Revised 6 September 2022; Accepted 12 September 2022; Published 12 October 2022

Academic Editor: R. Mo

In order to maintain the sustainable development of resource-based industrial clusters, research on the coupling relationship between resource-based industrial clusters and regional innovation networks is proposed. Combined with the historical data of 11 typical resource-based cities from 2013 to 2021, the grey relational model is used to measure the correlation between resource-based clusters and regional innovation networks. And the concept of capacity coupling and the capacity coupling coefficient model in physics is used as a reference to obtain the coupling degree model of resource-based industrial clusters and regional innovation networks. The coupling degree model is used to measure the coupling degree of the two, and a reasonable analysis is carried out on the result of the measurement. The results show that there is a correlation between resource-based industrial clusters and regional innovation networks. And the correlation between the two is above 0.65, indicating a strong correlation. The coupling between resource-based industrial clusters and regional innovation networks has entered the subsequent high-level coupling stage from the run-in coupling stage that began in 2013 and continues to maintain a high-level coupling. It is pointed out that cluster governance plans should be formulated according to regional differences and local conditions, so as to guide the transformation and upgrading of resource-based industrial clusters and to avoid the negative impact on the regional economy due to resource exhaustion.

1. Introduction

Environmental problems are a global problem, and many countries around the world often use sustainable development strategies. Chinese Premier Zhu Rongji pointed out at the conference that China will unswervingly pursue the path of sustainable development. China began to implement the sustainable development strategy in 1994 [1]. The Seventeenth National Congress put forward the higher goal of sustainable development to build an ecological civilization.

Coordinated development is a prerequisite for sustainable development, and coordinated development is also a prerequisite for a country's stable, healthy, and efficient operation. Therefore, to achieve sustainable development, coordinated development must be achieved first. During the national "Eleventh Five-Year Plan" period, the grand drama of regional coordination and integrated development kicked off. The outline of the "Twelfth Five-Year Plan" put forward "promoting coordinated and interactive development of regions," while the "Thirteenth Five-Year Plan" clearly pointed out five new development concepts and put forward a coordinated development concept. In the 2016 government work report, Lin et al. proposed to promote coordinated regional development and ensure the coordinated operation of the economy and society [2].

The future 13th Five-Year Plan clearly proposes five new development concepts including the concept of coordination and the concept of green. At the same time of rapid development, it faces severe resource shortages and environmental pollution tests. At the same time, resource shortages also limit the speed of economic development, and the environmental carrying capacity also limits the healthy development of the economy. Rapid industrialization and urbanization lead to a shortage of mineral resources and increasingly serious regional and mobile environmental pollution, which has become a major constraint to the city with strong international competitiveness [3]. The issue of sustainable development is a hot topic of current research at home and abroad.

With the growth of the world population, the competition between resource depletion and environmental pollution has become more severe. Foreign researchers have studied the relationship between resources, the environment, and the economy for a long time. The conclusion can be divided into three stages. The first stage is the slow development stage of initial research, mainly before the Second World War. The famous German human geographer Ratzel proposed environmental determinism, proposing that "the environment will determine the development of a race." At this stage, scholars began to pay attention to the relationship between population and environment but did not pay much attention to the relationship between economy, resources, and environment. The second stage is the stage of rapid development, which mainly occurred from the late 1940s to the late 1980s. This period belongs to the post-World War II peace period. With the surge in population, the serious shortage of natural resources, and the serious damage to the ecosystem, the interrelationships between the economy, resources, and the environment have attracted more and more attention from scholars in various fields. The third stage is the comprehensive deepening stage since the late 1980s [4].

In the research of the ERE system, foreign scholars originated in the late 19th century, and domestic scholars began to study the coordinated development of the ERE system in the 1970s. From the concept of "garden city" in the United Kingdom to the establishment of a series of theoretical systems of regional economics, environmental economics, and other disciplines in the whole society, the research field of the ERE system has been continuously deepened, resulting in a research model with rich research indicators, extensive research content, diverse research methods, theoretical and empirical research, and quantitative and qualitative analysis.

2. Literature Review

Wang et al. proposed to explore the economic development model of resource-based regions and proposed that resource-based regions should develop a regional characteristic model (local modelling), which opened the prelude to the research of resource-based regions [5]. Zonzini et al. put forward the four-stage development theory of resource-based regions, arguing that the development of resource-based regions needs to go through infancy, growth, transition, and maturity. The four-stage development theory has laid the foundation for the staged research on resource-based regions by later scholars [6]. Xia et al. extended the stage theory of the life cycle of resource-based regions and proposed that the development of resource-based regions can be divided into six stages: construction, employment, transition, maturity, decline, and exhaust gas. These studies provide a theoretical basis for the academic exploration of resource-based regional development [7]. However, due to the depletion and nonrenewability of mineral resources, mineral resources in Western resource-based regions began to deplete in the late 1970s and early 1980s. At the same time, the resource-related extractive and manufacturing economies also began to decline. Therefore, Western scholars began to explore the transformation path of resource-based regions. Among them, Xu and others believed that the development of multilateral industries was an important path for the transformation of resource-based regions, but this was not conducive to the transformation of remote resource-based regions far from developed cities and towns [8]. Ahlemann et al. proposed that resource-based regions should get rid of economic dependence on mineral resources, not only to achieve industrial diversification but also to transform the labor market and product transactions [9]. In the 21st century, after a long period of theoretical exploration and scientific practice, many resource-based regions have gradually completed their economic transformation. For example, Li et al. analyzed the relationship between the control of mineral resources and the development of a resource-based regional economy, in order to achieve sustainable development in the context of the network era [10].

John and Edwards conducted research on industrial geography planning and urban construction planning in coal resource-based regions. Driven by the national macroeconomic policy, our country's industrialization has developed rapidly. Large-scale infrastructure construction and rapid urbanization have led to the prosperity and expansion of resource-based regions. Regarding the boosting effect of mineral energy on regional economic development, scholars undoubtedly hold a positive opinion [11]. Based on the new understanding of resource and environmental carrying capacity, Wang et al. proposed the calculation of ecological service capacity by characterizing the composition of resource and environmental carrying capacity (i.e., resource supply capacity, environmental pollution-absorbing capacity, ecological service capacity, and social support capacity) method to establish a more practical environmental carrying quantitative model [12]. Wu et al. sorted out and summarized the research results on the carrying capacity of resources and environment from the aspects of connotation, evaluation methods, and applications. They believed that the current academic research on the theoretical part of the carrying capacity of resources and environment was not deep enough and lacked unity. However, the scientific evaluation index system still lacked targeted research on different regions and cities in practical application [13]. Cao et al. reviewed the literature on tourism environmental carrying capacity at home and abroad in the past ten years and believed that the focus of future related research should be strengthened in five aspects: theory, method, norm, empirical, and practice [14]. Chen et al. took the western region of China as the research object, set four different waste emission level scenarios, and then used the resource and environmental carrying capacity as the standard to verify the potential GDP growth rate of Western provinces and cities. This is an in-depth application of resource and environmental carrying capacity [15].

The principle of connecting systems is that there is a good relationship between the two systems. Therefore, in the process of empirical analysis of the relationship between industrial clusters and regional innovation networks, it is necessary to first confirm the relationship between industrial clusters and regional innovation networks and then determine the degree of the relationship. In the research, the grey relational model is selected for subsequent related calculations to illustrate that there is a certain relationship between resource-based industrial clusters and regional innovation networks, and a coupling model based on efficacy function is selected to measure the degree of mutual influence between resource-based industrial clusters and regional innovation networks.

3. Research Methods

3.1. Coupling and Coordination. The level of resources and environment refers to the level of natural resources such as land, water, minerals, and energy and the level of ecological environments such as water environment, forest, atmosphere, and solid waste in the region. The comprehensive development level of expansion, construction land development, etc., is shown in Figure 1. The resource environment is the basic support for regional development and construction. With the increase of regional development intensity, the pressure and coercion on resources and the environment increases, which restricts the scale and intensity of regional development and construction. Regional development and construction change the development and construction mode according to the feedback information from resources and the environment, forming a dynamic interaction relationship [16].

In the physical sense, coupling refers to the phenomenon that two (or more than two) systems or patterns of movement affect each other through various interactions. Connections describe how systems or objects interact with each other. Through integration, the degree of connection and coordination determines the sequence and structure of the system through the critical areas; for example, the system tends to move from the dispute to the judgment. The key to the transformation of the body from disorder to order is in the intersection of nonorder in the body, which affects the level of change and the rules of the system, and the degree of connection is the measure that affects the integration [17].

The level of cooperation is used to measure the development of cooperation between systems or elements, which represents the will of the body from conflict to order and is the quantitative index of cooperation. It can analyze the cooperation between the systems and elements in a system, as well as the degree of connection between the two systems and the order level of the connection between the machines or products.

3.2. Analysis of the Interaction between Resource-Based Industrial Clusters and Regional Innovation Networks. Based on the analysis of the main topic, the main content, and characteristics of the production capital and the new market in the region, the production capital is related to regional innovation networks. Improving the development of the group will help to build a regional innovation network, find key points of innovation, improve regional products, improve regional business models, and improve the regional innovation network. At the same time, the development of new regional networks helps to change the structure of traditional products such as capacity building, improving the competitiveness of the group and thus helping to promote improved economic stability [18]. We consider the interaction between the two in terms of structure, function, and purpose. This study also explains the positive relationship between economic investments and new businesses as economic investments in the region. The connection process of regional innovation networks is shown in Figure 2.

3.3. Coupling Evaluation Index System of Resource-Based Industrial Clusters and Regional Innovation Networks. Based on the above points, evaluation indicators can be formulated to analyze the integration of regional economic resources. Figures 1 and 2 show the evaluation system of economic resource integration and the new regional economy, which consists of 14 specific indicators created at 5 levels. Table 1 shows the coupling evaluation index system of the resource-based industrial cluster and regional innovation network, and Table 2 shows the coupling evaluation index system of the resource-based industrial cluster and regional innovation network.

3.4. Selection of the Empirical Model

3.4.1. Grey Relational Model. The following are the specific analysis steps of the grey relational model:

(1) Define the reference and comparison sequence. The degree of relationship between the two systems of resource-based business clusters and regional innovation networks will be evaluated. Therefore, the definition of resource-based business group system index factors is continued as used in sequence and shown as follows $X_i\,(i = 1, 2, \ldots, n)$, where $n = 5$ in the research and in the follow-up. In the actual calculation process, X_i is implemented in the indicators, namely, X_1 is the number of industrial enterprises/built-up area, X_2 is the number of employees in the mining industry, X_3 is the total industrial output value, X_4 is the total profit, and X_5 is the growth rate of the regional GDP. Each index factor of the regional innovation network system is a comparison sequence, denoted as $X_j\,(j = 1, 2, \ldots, m)$. In the research, $m = 9$ is also implemented in the subsequent actual measurement and calculation process. X_j is implemented in the

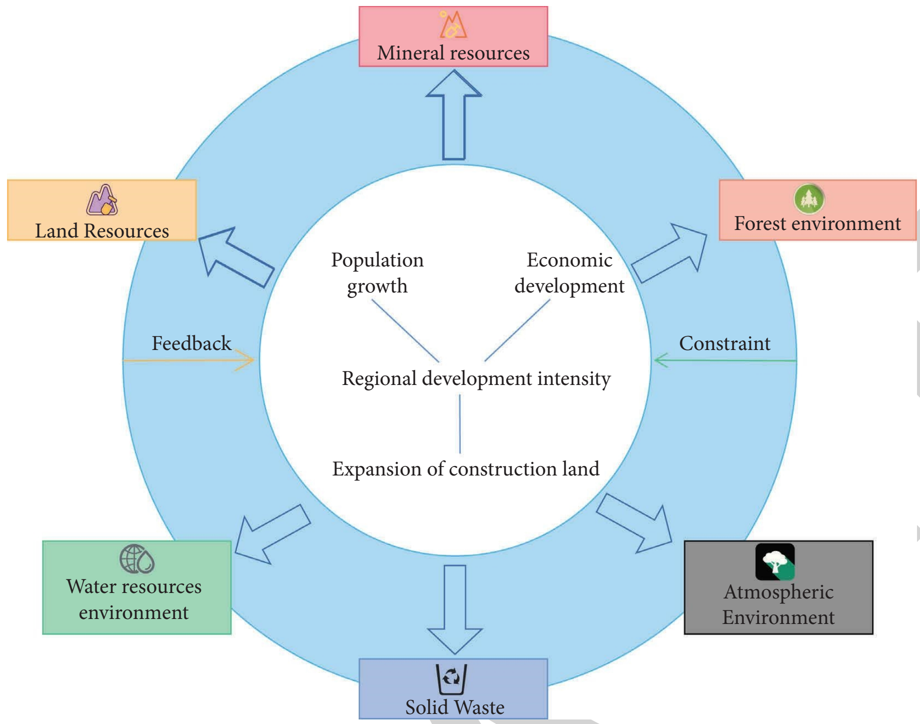


Figure 1: Schematic diagram of the interaction between the resource environment and regional development.

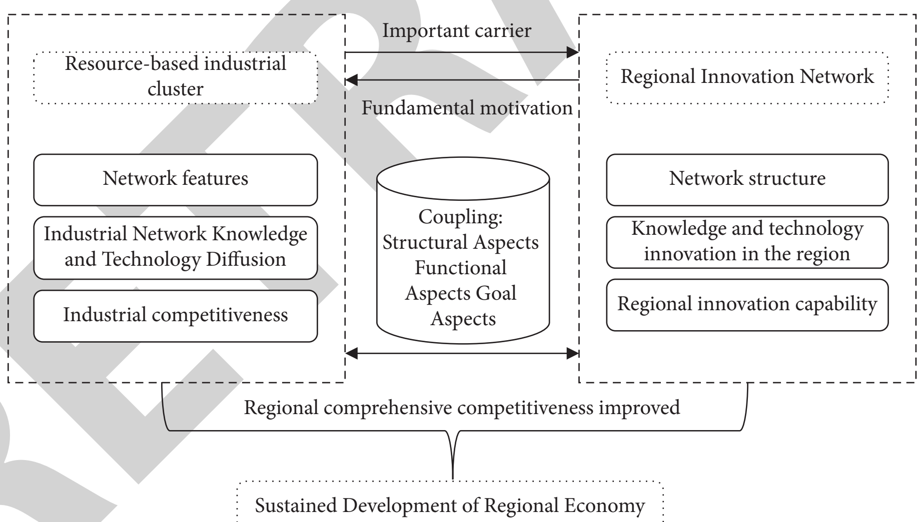


Figure 2: Coupling mechanism of resource-based industrial clusters and regional innovation networks.

index. Specifically, X_1 is the number of people for scientific research, technical services, and the geological exploration industry, X_2 is the number of employees in public management and social organizations, X_3 is the number of employees in the financial industry, X_4 is the number of patent applications, X_5 is the number of public library books, X_6 is the balance of various loans of financial institutions, X_7 is the amount of the infrastructure construction investment, X_8 is the number of Internet broadband access users, and X_9 is the amount of scientific and technological financial expenditure. Combined with the research, panel data will be selected. In general, the original data can be represented as $X_i = (x_i(1), x_i(2), \ldots, x_i(t))$ and $X_j = (x_j(1), x_j(2), \ldots, x_j(t))$, where the time and year are represented by t [19].

(2) Dimensionless processing of data. In order to minimize the deviation caused by the difference in the physical meaning and dimension of the index, the initial value processing transformation is adopted. At this time, x_i' and x_j' are the new data after

Table 1: Coupling evaluation index system of resource-based industrial clusters and regional innovation networks (target layer: resource-based industrial clusters).

Criterion layer	Indicator layer	Unit
Cluster size	Number of industrial enterprises/built-up area	Per square kilometer
	Number of employees in the mining industry	Ten thousand people
	Industrial output	Ten thousand yuan
Economic benefits	Total profit	Ten thousand yuan
	GDP growth rate	%

Table 2: Coupling evaluation index system of resource-based industrial clusters and regional innovation network (target layer: regional innovation network).

Criterion layer	Indicator layer	Unit
Network main body	Number of persons employed in scientific research, technical services, and geological prospecting	Person
	Number of employees in public administration and social organizations	Ten thousand people
	Number of people employed in the financial industry	Person
Network function	Number of patent applications	Pieces
	The number of books in public libraries	Thousands of copies
	Balance of various loans of financial institutions	Ten thousand yuan
Network environment	Investment in infrastructure construction	Ten thousand yuan
	Number of internet broadband access users	Thousands of families
	Government expenditure on science and technology	Ten thousand yuan

processing the original data, as shown in the following formula:

$$x_i = \left(1, \frac{x_i(2)}{x_i(1)}, \cdots, \frac{x_i(n)}{x_i(1)}\right),$$
$$x_j' = \left(1, \frac{x_j(2)}{x_j(1)}, \cdots, \frac{x_j(m)}{x_j(1)}\right). \tag{1}$$

(3) Here, $\rho \in (0, 1)$ is the resolution coefficient. In the research, the resolution coefficient $\rho = 0.5$ is used for subsequent calculations, as shown in the following formula:

$$\xi_i(j)(t) = \frac{\min_i \min_j \left|x_i'(t) - x_j'(t)\right| + \rho\max_i \max_i\left|x_i'(t) - x_j'(t)\right|}{\left|x_i'(t) - x_j'(t)\right| + \rho\max_i \max_j\left|x_i'(t) - x_j'(t)\right|}. \tag{2}$$

(4) The correlation degree is calculated to obtain the correlation degree matrix, where the correlation degree Υ_{ij} in the matrix is the following formula:

$$\Upsilon_{ij} = \frac{1}{k}\sum_{i=1}^{k} \xi_i(j)(t), \quad k = 1, 2, \ldots, n, k \text{ is the number of samples},$$
$$\Upsilon_i = \frac{1}{m}\sum_{j=1}^{m} \Upsilon_{ij}\,(i = 1, 2, \ldots, n). \tag{3}$$

3.4.2. Coupling Degree Model. The following are the specific implementation steps of the coupling model.

(1) Efficacy Function. Multiple measurement indicators can be used to comprehensively evaluate the system. Since each indicator has inconsistent dimensional units and values, in order to avoid interference with the measurement results, all measurement indicators in the system must be preprocessed, that is, the efficacy coefficient method is used in the research to further obtain the efficacy function values of the two systems [20]. The efficacy coefficient method can realize the dimensionless processing and the same direction processing of each index, and its calculation process is the following formula:

$$x_{ij}' = \begin{cases} \dfrac{x_{ij} - \min\left(x_j\right)}{\max\left(x_j\right) - \min\left(x_j\right)}, x_{ij}', \\ \\ \dfrac{\max\left(x_j\right) - x_{ij}}{\max\left(x_j\right) - \min\left(x_j\right)}, x_{ij}'. \end{cases} \tag{4}$$

In formula (4), i represents the system sequence parameter (the research refers to the year and there is $i = 1, 2, \cdots, 10$), j represents the index, and x_{ij} is the index value corresponding to each year, that is, the original data value. $\min(x_j)$ and $\max(x_j)$, respectively, represent the minimum and maximum value of the jth item index. $x_{ij}$is the index value after processing and is between 0 and 1, representing the contribution value of x_{ij} to the respective system. The larger the value, the greater the contribution of the value of the index to its system. In the research, the first formula is mainly adopted to realize the subsequent calculation.

Thus, the total efficacy function of each system is further obtained, which can be specifically expressed as the following formula:

$$U_{n-1,2} = \sum_{j-1}^{m} W_j x_{ij}'. \tag{5}$$

In formula (5), U_1 represents the total efficacy function value of the resource-based industrial cluster system, U_2 represents the total efficacy function value of the regional innovation network system, n represents the system, m is the number of indicators of a system (the research has $m = 5$ when $n = 1$ and $m = 9$ when $n = 2$), and W_j is the weight of each index of a system. The larger the U_1 value, the better the comprehensive development level of resource-based industrial clusters, and vice versa; the value of U_2 is larger [21].

(2) Coupling Model. The coupling degree can be used to measure the degree of mutual influence between two systems. Based on the power function, the coupling degree model of the interaction of multiple (n) systems is generalized, as follows:

$$C = n\left\{\frac{\left(U_1 \cdot U_2 \cdot \ldots \cdot U_n\right)}{\left[\prod\left(U_q + U_z\right)\right]}\right\}^{1/n}. \tag{6}$$

There are only two systems in the relationship between resource-based enterprises and regional innovation organizations. In order to simplify the research, the above multicommunication network model can be conveniently added to establish a model suitable for production capacity assessment. The ranking model and the model combined with the regional innovation network are shown in

$$C = 2\left\{\frac{\left(U_1 \cdot U_2\right)}{\left[\left(U_1 + U_2\right)\left(U_1 + U_2\right)\right]}\right\}^{1/2}. \tag{7}$$

3.5. Data Selection and Sources. The cluster referred to in the research is the chain structure of the industrial chain and its value chain with nonrenewable resources (mainly mineral resources) as the basic production factors. Therefore, in terms of data selection, the research will first remove the typical resource-based cities dominated by forest resources. Considering the availability of data, the research selects 11 provinces, cities, or autonomous regions in our country, and the resources cover coal, oil, 11 typical resource-based cities of four types of nonferrous metals and ferrous metals; we conduct empirical research from the data of nearly ten years from 2008 to 2017.

4. Results Analysis

4.1. Correlation Analysis between Resource-Based Industrial Clusters and Regional Innovation Networks. In order to analyze whether there is a correlation between resource-based industrial clusters and regional innovation networks, the following three points are discussed step by step in the research. First, according to the obtained correlation rank matrix, the average value of each column is obtained, and the correlation index value of each resource-based manufacturing group and regional innovation network system index is obtained. This is used to evaluate the existence of the relationship between indicators of capital-based business and regional innovation network systems. Secondly, the average value of the correlation measurement value Υ_i of each index of the regional innovation network and the resource-based industrial cluster system by row is obtained. Υ_i is used to judge whether there is a correlation between the indicators of the regional innovation network and the resource-based industrial cluster system. Finally, combined with the previous exploration process, the degree of relevance between the resource-based industrial cluster and the regional innovation network is analyzed as a whole. In this whole analysis process, it is necessary to combine the existing classification criteria of association degree in previous studies to judge the degree of association. The specific division criteria are shown in Table 3.

According to the correlation matrix calculated above, the calculation results are divided into 11 urban regions and the relationship between the business input index and the regional innovation network system is studied. We compare resource-based economic cluster and regional innovation index. The ranking table is shown in Table 4. The results corresponding to IR in this table are the five specific indicators of economic resources (number of industrial establishments/plants, workers in the mining sector, value of total output, total income, and GDP growth rate of the region) and regional innovations. The degree of freedom and the value of each degree of freedom vary little across cities, indicating that the relationship between the economy and regional capital has been poorly measured over the past decade. This gradual change suggests that the development of capital-based businesses will lead to new businesses in the region. In addition, resource-based commercial enterprises are the main source of new business opportunities in this

Table 3: Classification criteria of correlation degree.

Correlation degree	Evaluation description
$0 < \Upsilon_i \leq 0.35$	Weak correlation
$0.35 < \Upsilon_i \leq 0.65$	Moderate correlation
$0.65 < \Upsilon_i \leq 0.85$	Strong correlation
$0.85 < \Upsilon_i \leq 1$	High correlation

Table 4: Numerical table of correlation degree between each index of resource-based industry cluster and regional innovation network in 11 cities.

City	r_i				
	r_1	r_2	r_3	r_4	r_5
A city	0.807841	0.834576	0.772171	0.825325	0.7932
B city	0.904282	0.895912	0.906768	0.846104	0.884523
C city	0.761847	0.7643	0.645369	0.684501	0.717437
D city	0.854821	0.876952	0.875045	0.8402	0.827413
E city	0.854657	0.851421	0.877584	0.809168	0.850639
F city	0.829657	0.849554	0.871266	0.796046	0.828472
H city	0.827855	0.813827	0.809255	0.788645	0.807453
I city	0.756242	0.750166	0.715532	0.615913	0.7484
J city	0.800354	0.789741	0.825287	0.751	0.778468
K city	0.912926	0.827623	0.934925	0.910885	0.90965
L city	0.865352	0.875061	0.886214	0.851713	0.839547

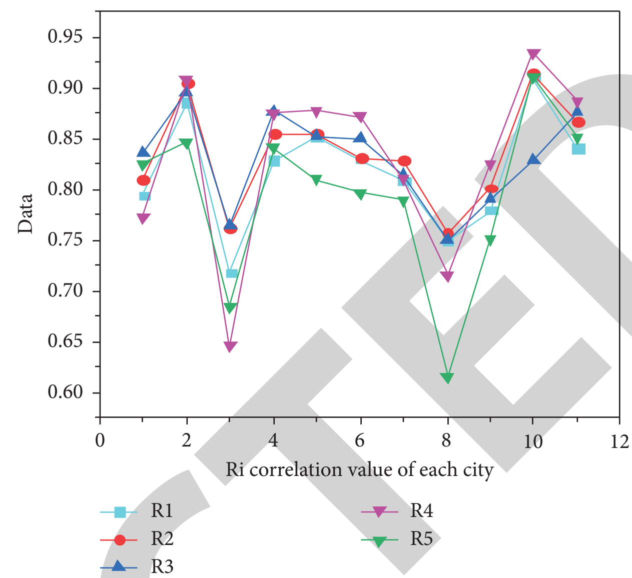


Figure 3: Numerical diagram of correlation degree between each index of resource-based industry cluster and regional innovation network in 11 cities.

region [22–26]. In order to clarify the relationship between economic resources and regional innovation indicators, this paper draws a quantitative report on the relationship between the resource-based economy of 11 cities and all regional innovation networks, as shown in Figure 3. The results shown in the figure show that the relationship between various indicators of capital market network marketing by the economy and regional economy in the last decade is almost more than 0.65, which illustrates the difference in corporate capital. There is a certain relationship between each index of the system and the regional innovation network system.

Then, to study the relationship between the regional innovation network and the resource-based economy [27, 28], the relationship between the regional innovation network and the resource-based economy in the business of 11 cities is shown in Table 5. The value corresponding to r_j in this table is the degree of correlation between the nine specific indicators of the regional innovation network and the resource-based industrial cluster system as a whole. It can be seen from this table that all the correlation degree values fluctuate within a range of only 0.2. This shows that there is a continuous relationship between regional innovation networks and resource-based business clusters, suggesting that the existence of regional innovation networks promotes the creation of resources. In addition, combined with this table, a numerical map of the correlation degree between each index of the regional innovation network and resource-based industrial clusters in 11 cities is drawn, as shown in Figure 4. Combined with the classification criteria of correlation degree, it can be seen that the correlation degree of each index in the regional innovation network system and the resource-based industry cluster system is mainly above 0.65, which also indicates that each index of the regional innovation network is at least strongly correlated with the resource-based industry cluster system.

Based on the above analysis process [29–31], it can be concluded that each index of the resource-based industry cluster system has a strong correlation with the regional innovation network system as a whole, and each index of the regional innovation network also has a strong correlation with the resource-based industry cluster system as a whole. Considering that the selected data objects themselves are representative and involve the east, west, and east regions of China, it can be inferred that, from the whole system, there is a strong correlation between the resource-based industry cluster and the regional innovation network.

4.2. *Coupling Analysis of Resource-Based Industrial Clusters and Regional Innovation Networks*

4.2.1. Time Evolution of the Coupling between Resource-Based Industrial Clusters and Regional Innovation Networks. The selected coupling degree model is used, combined with the coupling evaluation index system of resource-based industrial clusters and regional innovation networks constructed in the previous section and the corresponding index data, and the VBA language in EXCEL software [32–34] is used to calculate the model, so as to obtain the coupling degree measurement. Finally, the numerical result table of the coupling degree between resource-based industrial clusters and regional innovation networks as shown in Tables 6 and 7 is compiled.

According to the above-mentioned numerical table of coupling degrees, combined with the classification criteria of coupling degrees in the literature on the study of coupling relationships [35, 36], as shown in Table 8, this study will use

Table 5: Correlation degree of each index of innovation network in 11 cities and regions with resource-based industry clusters.

City	r_i								
	r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8	r_9
A city	0.84857	0.89352	0.86245	0.85773	0.77445	0.71898	0.68934	0.8162	0.79861
B city	0.92141	0.92687	0.92797	0.895	0.91858	0.8302	0.86825	0.90523	0.79765
C city	0.76292	0.7574	0.7564	0.67632	0.7275	0.68223	0.70732	0.66608	0.69657
D city	0.92494	0.94577	0.93843	0.70014	0.8706	0.85886	0.76626	0.78857	0.90066
E city	0.89297	0.90613	0.89868	0.76073	0.8957	0.8514	0.89737	0.8686	0.66672
F city	0.83386	0.90591	0.90669	0.6055	0.87575	0.84585	0.89933	0.80194	0.84027
H city	0.87771	0.88925	0.88887	0.66147	0.80773	0.8022	0.84972	0.78255	0.72503
I city	0.74222	0.77098	0.73469	0.69173	0.72155	0.73544	0.72871	0.6667	0.66343
J city	0.81583	0.88878	0.89385	0.6252	0.75806	0.81556	0.79196	0.74583	0.76764
K city	0.92294	0.92844	0.9293	0.82217	0.9267	0.90965	0.85883	0.92664	0.86763
L city	0.91898	0.92995	0.9468	0.8666	0.85418	0.68133	0.73141	0.88715	0.95561

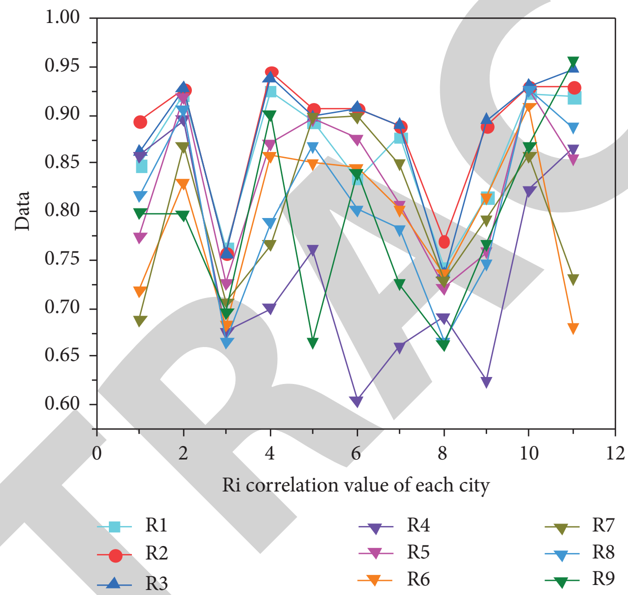


Figure 4: Numerical diagram of correlation degree between each index of an innovation network and resource-based industry cluster in 11 cities and regions.

Table 6: Coupling degree of resource-based industrial clusters and regional innovation networks (1).

	A city	B city	C city	D city	E city	F city
2008	0.596835	0.026749	0.883327	0.257738	0.325568	0.955275
2009	0.922196	0.808783	0.85817	0.666865	0.657443	0.998035
2010	0.9685	0.766024	0.960345	0.708246	0.770223	0.967442
2011	0.902481	0.815388	0.98278	0.872237	0.867297	0.998546
2012	0.986521	0.916816	0.988848	0.959186	0.947092	0.905854
2013	0.981608	0.917121	0.994489	0.990555	0.999042	0.774407
2014	0.985383	0.961182	0.982524	0.939589	0.99069	0.997452
2015	0.999612	0.993098	0.980505	0.913716	0.973032	0.876197
2016	0.95732	0.893216	0.974616	0.82412	0.931382	0.830295
2017	0.793052	0.914745	0.958772	0.865083	0.681858	0.836118

the resources based on business clusters and regional innovation linkages as the basis for its conclusions. The four levels of horizontal linkage, antagonistic linkage, run-on linkage, and high-level linkage are used to measure the linkage change process of assistance based on production groups and regional innovation networks [23]. In the low-level coupling stage, the mutual influence of resource-based industrial clusters and regional innovation networks is still slight, and their coupling relationship does not clearly exist. In the antagonistic coupling stage, the two systems of

Table 7: Coupling degree of resource-based industrial cluster and regional innovation network (2).

	G city	H city	I city	J city	K city
2008	0.668398	0.432233	0.824567	0.728933	0.608289
2009	0.930099	0.942502	0.898385	0.756595	0.687294
2010	0.781866	0.957656	0.965976	0.751045	0.629124
2011	0.842878	0.999956	0.911597	0.913888	0.85993
2012	0.9732	0.973742	0.89021	0.989817	0.964375
2013	0.998195	0.975976	0.996845	0.957277	0.996136
2014	0.998577	0.993567	0.994191	0.957176	0.977395
2015	0.896125	0.967348	0.945093	0.994486	0.788677
2016	0.929063	0.986957	0.957583	0.993952	0.591245
2017	0.918887	0.991957	0.942762	0.904041	0.842723

Table 8: Coupling level division.

Coupling	Level division
[0, 0.3]	Low-level coupling phase
(0.3, 0.5]	Antagonism phase
(0.5, 0.8]	Running-in stage
(0.8, 1]	High-level coupling stage

resource-based industrial clusters and regional innovation networks are internal or external. Driven by the driving force, the cross-catalytic mechanism of the two is gradually formed, and at this time, it is in a state of mutual restriction. In the running-in coupling stage, the interaction relationship between the two systems changes and the survival or development of one of them will affect the other. In the high-level coupling stage, resource-based industrial clusters and regional innovation network systems have entered healthy self-development. The exchange of matter, energy, and information between systems is in a reasonable and orderly state, and the two influence each other to the highest degree [24].

In order to more intuitively reflect the coupling of resource-based industrial clusters and regional innovation networks in these 11 typical resource-based cities from 2008 to 2017, the corresponding results are drawn, as shown in Figure 5. It can be seen from the figure that the coupling degree between resource-based industrial clusters and regional innovation networks in these 11 typical resource-based cities has shown a slow upward trend in the past ten years [37–39]. In addition to the large difference in degrees, the coupling degrees of these two systems in each city in the following years are both above 0.5 and fluctuate, which indicates that over time, the degree of mutual influence between resource-based industrial clusters and regional innovation networks has exceeded the low level. In the coupling and antagonistic coupling stages, the two systems have a strong coupling relationship, and the two systems will not cause constraints on the other due to the excessive advance and lag development of one of them, but both affect each other and develop in a positive direction [25]. At the same time, it is found that the coupling degree between resource-based industrial clusters and regional innovation networks in each city in 2008 is significantly different from

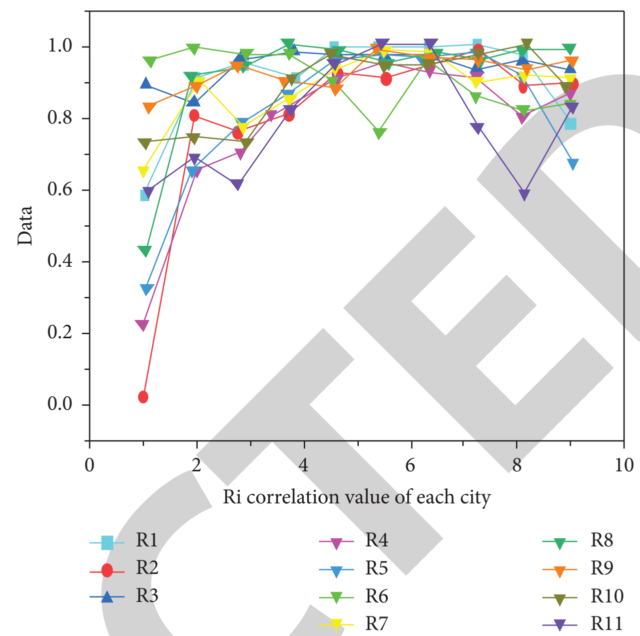


Figure 5: Coupling degree of resource-based industrial clusters and regional innovation networks.

other years. The year is generally low, indicating that the regional innovation network is initially formed at this time, and the resource-based industrial cluster is in the process of self-expanding development, relying on resource endowment to achieve regional economic development, and the current demand for innovation activities is less, which makes the two in this year. The degree of mutual influence is lower than in subsequent years. Then, in order to obtain the integration process of each city's production capacity and new development area in the past ten years, the average annual connection level of the two systems has been taken as a city model, and the model is drawn as shown in Figure 6. As this figure shows, the production capacity as a product and new industries in the region are intertwined. Both have entered the next phase of high connectivity from the level of connectivity that started in 2008 and continues to have high connectivity. This often shows that the links between the regional innovation networks of natural resources-based cities and natural resources-based industrial groups have directly entered into their state's development.

4.2.2. Spatial Analysis of Coupling between Resource-Based Industrial Clusters and Regional Innovation Networks. In 2017, according to the regional distribution of land parcels in the east, central, and west China, 11 traditional capital cities and newly connected capital business groups in the region were divided and classified by economic zone. As shown in Tables 9 and 10, the degree of connection between the two can be obtained. According to the table, spatial analysis is conducted on the connection between capacity and new markets in the region.

From Tables 9 and 10, it can be seen that the coupling degree of resource-based industrial clusters and regional innovation networks in these 11 typical resource-based cities

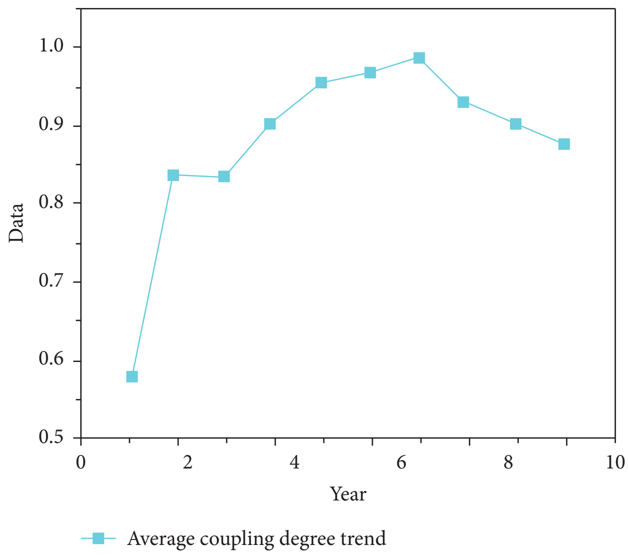


Figure 6: Trend chart of average coupling degree from 2008 to 2017.

Table 9: Spatial distribution of coupling degree between resource-based industrial clusters and regional innovation networks in 2017 (running-in coupling).

Eastern region	Central region
A city	E city

Table 10: Spatial distribution of coupling degree between resource-based industrial clusters and regional innovation networks in 2017 (high-level coupling).

Eastern region	Central region	Western region
B city	C city, D city	F city, G city, H city, I city, J city, K city

has obvious spatial characteristics in the geographical distribution. Overall, calls from the east to the central and western regions are on the rise. Most of the cities have high connectivity, and the western regions have high connectivity. Business development in the eastern and central regions has been successful, and urban planning and construction have become more rational and timely. The business-based business model established in the capital will change over time in response to new needs. The degree of interaction has gone from low to high [40–42]. The economic growth of cities in the western region is generally lower than that in the eastern and central regions. For urban resources based on natural resources, urban development is not sustainable in terms of resource utilization, and new employment opportunities must be created. The simplicity of the business chain is reduced. The integration of urban economic development based on urban capital and regional innovation leads to continuous fixed development, and there is always capital according to the business development, thus making the capital grow. In addition, the National Sustainable Development Plan 2013–2020 in 2017 will further connect resource-based industrial clusters and regional innovation networks in 11 natural resource-based cities. Tables 11 and 12 show location analysis and specific investment results by market and regional market, respectively. The interplay between capacity and new business in the region is important, the report notes. Focusing on developing and declining cities, this is based on the need for resource innovation as cities move into the next stage of development. Among them, resource-based industrial clusters need the promotion of regional innovation networks to achieve diversified development and avoid the weakening of clusters. In addition, except for growth-oriented cities, which do not have corresponding couplings, other types of cities show that resource-based industrial clusters and regional innovation networks have a high level of coupling. Industrial clusters are also in the budding development, relying on their operating mechanism to achieve the development of clusters, thereby bringing sustainable development to the regional economy.

Table 11: Spatial distribution of coupling degree by city type in 2017 (running-in coupling).

Mature city	Decay city
A city	E city

Table 12: Spatial distribution of coupling degree by city type in 2017 (high-level coupling).

Mature city	Decay city	Regeneration city
D city, F city, I city, K city	C city, G city, H city, J city	B city

5. Conclusion

There is a correlation between resource-based industrial clusters and regional innovation networks, and the correlation between the two is above 0.65, indicating a strong correlation. Whether from the perspective of the correlation degree between different indicators in the resource-based industrial cluster system and regional innovation network, or from the perspective of the correlation degree between different cities, the value of each correlation degree is almost the same. In recent 10 years, the relationship between regional innovation network and resource-based regional economy has not changed much, which indicates that resource-based industrial cluster is an important factor in regional development. Considering the whole set of resources as a business, the amount of difference between all relevant results is only 0.2, showing the relationship between the indicators of the new business area and the resource-based business. The presence of the regional innovation network is the catalyst for the development of the services sector of the industrial group, and the regional innovation network is the main driving force for the development of the services of the industrial group. From the point of view of the whole system, there is a strong relationship between the economic resources and the new economy in the region, and the degree of relationship between them is also strong.

The connection between the capital market and the regional economy has steadily increased from the level of connection that began in 2008 and continues. The relationship between the urban economy and the regional economy of 11 cities that can play an economic role shows slow growth in the last decade. The coupling degree of these two systems in each city is in the range of 0.5 or more, which means that with time, the degree of mutual influence between resource-based industrial clusters and regional innovation networks has exceeded the low-level coupling and antagonistic coupling stages. There is a strong coupling relationship between the systems, and the two systems will not restrict the other due to the excessively advanced and lagging development of one of the two systems. But they both affect each other and develop in a positive direction, and the two systems influence each other to a very high degree.

Data Availability

The dataset can be accessed upon request to the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] X. Liu and X. Zhang, "Noma-based resource allocation for cluster-based cognitive industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 8, pp. 5379–5388, 2020.

[2] F. Lin, W. Dai, W. Li, Z. Xu, and L. Yuan, "A framework of priority-aware packet transmission scheduling in cluster-based industrial wireless sensor networks," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 99, pp. 5596–5606, 2020.

[3] R. Umanodan, A. Shimazu, M. Minami, and N. Kawakam, "Corrigendum∗: effects of computer-based stress management training on psychological well-being and work performance in Japanese employees: a cluster randomized controlled trial," *Industrial Health*, vol. 58, no. 4, pp. 397-398, 2020.

[4] P. Jia, X. Wang, and K. Zheng, "Distributed clock synchronization based on intelligent clustering in local area industrial iot systems," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 6, pp. 3697–3707, 2020.

[5] K. Wang, Y. Zhou, Z. Liu, Z. Shao, and Y. Yang, "Online task scheduling and resource allocation for intelligent noma-based industrial internet of things," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 99, p. 1, 2020.

[6] F. Zonzini, A. Girolami, L. D. Marchi, A. Marzani, and D. Brunelli, "Cluster-based vibration analysis of structures with graph signal processing," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 99, p. 1, 2020.

[7] W. Xia, J. Zhang, T. Quek, S. Jin, and H. Zhu, "Mobile edge cloud-based industrial internet of things: improving edge intelligence with hierarchical sdn controllers," *IEEE Vehicular Technology Magazine*, vol. 15, no. 1, pp. 36–45, 2020.

[8] X. Xu, Y. Xu, M. H. Wang, J. Li, and Y. He, "Data-driven game-based pricing for sharing rooftop photovoltaic generation and energy storage in the residential building cluster under uncertainties," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 99, p. 1, 2020.

[9] F. Ahlemann, C. Legner, and J. Lux, "A resource-based perspective of value generation through enterprise architecture management," *Information & Management*, vol. 58, no. 1, pp. 1–17, 2021.

[10] X. Li, X. Yao, Z. Guo, and J. Li, "Employing the cge model to analyze the impact of carbon tax revenue recycling schemes on employment in coal resource-based areas: evidence from shanxi," *The Science of the Total Environment*, vol. 720, no. 10, pp. 137192.1–137192.14, 2020.

[11] John and Edwards, "Signal processing inspires network innovation: signal processing is the key as researchers work to boost network speed and capacity [special reports]," *IEEE Signal Processing Magazine*, vol. 37, no. 1, pp. 16–23, 2020.

[12] J. Wang and X. Cao, "Evolution mechanism of advanced equipment manufacturing innovation network structure from the perspective of complex system," *Complexity*, vol. 2021, no. 3, 12 pages, 2021.

[13] H. Wu and X. M. Gu, "Fuzzy principal component analysis model on evaluating innovation service capability," *Scientific Programming*, vol. 2020, no. 1, 9 pages, 2020, https://doi.org/10.1155/2020/8834901, Article ID 8834901.

[14] X. Cao, G. Zeng, L. Lin, and L. Zou, "Hierarchical characteristics and proximity mechanism of intercity innovation networks: a case of 290 cities in China," *Complexity*, vol. 2021, no. 1, 14 pages, Article ID 5538872, 2021.

[15] H. Chen, "Research on innovation and entrepreneurship based on artificial intelligence system and neural network algorithm," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, pp. 2517–2528, 2021.

[16] Y. Li, L. Bai, and B. Cheng, "Analysis of enterprise site selection and r&d innovation policy based on bp neural network and gis system," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 4, pp. 5609–5621, 2020.

[17] R. GiddeRanjitsinha, "Design optimization of micromixer with circular mixing chambers (m-cmc) using taguchi-based grey relational analysis," *International Journal of Chemical Reactor Engineering*, vol. 18, no. 9, pp. 182–192, 2020.

[18] Wenming, Y. Xie, S. Ma, S. Li, L. Zhang, and M. Ruan, "Optimizing soil dissolved organic matter extraction by grey relational analysis," *Pedosphere*, vol. 30, no. 5, pp. 15–22, 2020.

[19] W. Xie, Y. Ma, L. I. Shijun, S. Zhang, and L. Zhang, "Optimizing soil dissolved organic matter extraction by grey relational analysis," *Pedosphere*, vol. 30, no. 5, pp. 589–596, 2020.

[20] J. Tang, B. Xiong, Y. Li, C. Yuan, and Y. Qiu, "Faulted feeder identification based on active adjustment of arc suppression coil and similarity measure of zero-sequence currents," *IEEE Transactions on Power Delivery*, vol. 36, no. 99, p. 1, 2021.

[21] X. Ren, C. Li, X. Ma, F. Chen, H. Wang, and A. Sharma, "Design of multi-information fusion based intelligent electrical fire detection system for green buildings," *Sustainability*, vol. 13, no. 6, p. 3405, 2021.

[22] J. Jayakumar, 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, Article ID 5325116, 8 pages, 2021.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 7817309, 10 pages
https://doi.org/10.1155/2022/7817309

Research Article

A YOLOv3-Based Industrial Instrument Classification and Reading Recognition Method

Haifei Zhang, Qianqian Chen, and Liting Lei

School of Computer and Information Engineering, Nantong Institute of Technology, Yongxing Road 211, Nantong 226002, China

Correspondence should be addressed to Haifei Zhang; 46462490@qq.com

Received 8 August 2022; Revised 24 September 2022; Accepted 28 September 2022; Published 11 October 2022

Academic Editor: R. Mo

Aiming at the demand of industrial instrument reading, this study proposes a method of industrial instrument classification and reading recognition based on YOLOv3. Given that industrial meters can be divided into pointer meters and digital meters according to the dial type, this method conducts a reading study for each of the two types of meters. Firstly, the YOLOv3 model is trained to recognize and detect the meter types and classify the meters according to the values of the obtained classes. The pointer meter uses a Hough circle to detect the dial, extracts the scale and the pointer, calculates the angle between the 0 scale line and the pointer, and obtains the reading of the pointer meter. The digital meter extracts the digits by finding the contours of the dial and the digit area and then uses a support vector machine (SVM) to identify the extracted digits and output the readings of the digital meter. Through the test, the mean average precision (mAP) of the recognition model in this study is 93.73%. The absolute error of pointer meter reading is less than 0.1 in general, and the maximum relative error is 0.35%. The accuracy of the digital meter reading is 99.7%. The proposed method can accurately read the value of the instrument and meet the needs of industrial production.

1. Introduction

Nowadays, industrial instruments are becoming increasingly functional and play an increasingly important role in the industry, becoming an integral part of modern industrial production. They have become an integral part of modern industrial production. With the continuous integration and deepening of electronic information and modern industry, industrial production is gradually developing towards automation and intelligence. IoT-based instrument reading is widely used, but this method is only applicable to industrial instruments with communication interfaces for reading. In the current industrial life, there are still many traditional industrial meters without communication interfaces, which are mostly read manually. However, the manual reading of the meter is a large and inefficient workload, which can easily cause visual fatigue and thus lead to errors in the values read.

Industrial meters at this stage can be divided into pointer meters and digital meters according to the type of dial. Since there are advantages and disadvantages to both pointer and digital meters, both types of instruments are often used in combination in industrial production. This study analyzes the advantages and disadvantages of the instrument recognition and reading method used in Zhang's paper [1] and Lei's paper [2] and proposes an industrial instrument reading recognition system based on YOLOv3, which uses the YOLOv3 target detection algorithm to identify the instrument type. By calculating the included angle between the scale line and the pointer in the pointer instrument, the reading of the pointer instrument is recognized, and the support vector machine (SVM) method is used to identify the numbers in the digital instrument, and then, the reading result of the digital instrument is obtained. The test of the system realized by this method shows that this method has high accuracy in the type identification and reading of industrial instruments.

2. Related Work

With the rapid development of machine vision, instrument image processing technology has become a mainstream technology in the field of industrial instrumentation and is widely used throughout the production systems of industry.

In 2018, Kucheruk et al. [3] used template matching to precisely locate the center of the circle of the meter and the pointer and then calculated the rotation angle of the pointer and finally obtained the meter reading. In 2019, Chen et al. [4] proposed a fast and direct digital recognition algorithm for digital instruments based on feature extraction. In 2020, Guo et al. [5] used the GrayWorld algorithm to process digital meter images and then train a variable convolution character recognition algorithm to complete the recognition of digital meters. Based on machine vision, Zhang [1] firstly identified the type of instrument meters and then processed the pointer and scale, respectively, in pointer meters to read the pointer meters, while the digital meter adopted MNIST number set automatic recognition algorithm to recognize the reading. This method has an accuracy of 93.47% for meter type recognition and 98.5% for digital meter readings. Zhang et al. [6] read the pointer meter in HSV color space, determined the pointer centroid based on MBR, and was able to quickly locate the pointer position. In 2021, Li et al. [7] proposed an Otsu algorithm and an improved Hough algorithm, which can well avoid the influence of uneven illumination on pointer meter readings. Sowah et al. [8], using computer vision and machine learning technology, developed a new algorithm used for automatic meter reading, using a series of image contour filter cascading number classifier to extract the advanced features; this method does not rely on any information before the instrument is being read and can directly get the meter reading. Tang et al. [9] conducted automatic detection of pointer instrument through image recognition method, and the accuracy of reading recognition met the performance requirements, with efficient and intelligent operation characteristics. Wang et al. [10] combined target detection and computer vision, used fast-RCNN to detect the instrument panel, and then read the pointer indicator representation through Hough transform. Lei [2] designed a system capable of reading both pointer and digital meters. The system uses a target detection method to locate pointer meters and then combines image processing techniques to recognize the meter readings, while digital meters are read using a deep learning-based method. This system is very accurate for meter recognition, but can only identify and locate pointer meters with an accuracy of 81.4% using the minimum external matrix-based method, while the error using the PSPNet reading method is very high, with an average error value of 0.771.

3. YOLOv3-Based Recognition Model for Industrial Instruments

A model for industrial instrument type recognition is produced by collecting instrument images, labeling them with type, and then feeding them into the YOLOv3 neural network for training. The instrument images to be classified are fed into the trained model and the instrument types are identified by the model. Once the recognition is complete the instrument images are stored in a corresponding folder. The main process is shown in Figure 1.

3.1. YOLOv3 Principle. YOLOv3 borrows the ResNet idea by adding the residual module to the network and modifying Darknet-19 of YOLOv2 to Darknet-53 to form a deeper network layer. Fifty three layers of convolution are used in Darknet-53, where a large number of 3×3 and 1×1 convolution kernels are set up and $256 \times 256 \times 3$ is used as the output.

The structure of the YOLOv3 model [11–15] is shown in Figure 2. Each square in Figure 2 is a basic block consisting of Conv2d, BatchNorm2d, and LeakyReLU (except the last layer of each output) called Basic_Block. The resblock in Figure 2 is a residual block consisting of Basic_Block and a shortcut called Residual_Block. The rightmost three outputs in the diagram are called Header_Block and consist of a Basic_Block and a Conv2d. [+] in Figure 2 stands for splicing, i.e., splicing the channel dimensions of two tensors.

Figure 2 shows that the final outputs of the model are $y1$, $y2$, and $y3$. Assuming an output of $1 \times 3 \times 416 \times 416$ the output dimensions are $1 \times 255 \times 13 \times 13$, $1 \times 255 \times 26 \times 26$, and $1 \times 255 \times 52 \times 52$, three different scales are used to predict large, medium, and small targets, respectively.

Each piece of data used for training consists of a picture and a label, with the label content being the enclosing box and class of all targets of interest in the graph, with the enclosing box generally represented by the top-left vertex and bottom-right vertex. Intuitively, it is best if the output of the network is the same as the labels, i.e., $(x1, y1)$, $(x2, y2)$, and class. In practice, however, a direct regression of the above six values would not work well and the network would converge very slowly because the range of variation in the above six values is so large that the neural network would predict the data better and converge significantly faster after normalization. So, the YOLOv3 network chooses to predict a normalized data and encodes the six values mentioned above.

The relationship between the above output and the actual bounding box [16–18], confidence levels, and categories is as follows:

$$\begin{cases} b_x = c_x + \sigma(t_x), \\ b_y = c_y + \sigma(t_y), \\ b_w = p_w \cdot e^{t_w}, \\ b_h = p_h \cdot e^{t_h}, \\ b_p = \sigma(t_p). \end{cases} \tag{1}$$

It can be seen that the predictions t_x and t_y are the deviation of the center of the bounding box relative to the top-left corner of each grid, t_w and t_h are a factor of the width and height of the bounding box, t_p is the confidence level after sigmoid, and the category is the category corresponding to the subscript of the maximum value in the last 80 data.

3.2. Model Training. The images of pointer and digital meters were collected and labeled using LabelImg image labeling software to classify the images into two categories: pointer meters (labeled as "pointer") and digital meters (labeled as "digital"). The labeling effect is shown in Figure 3.

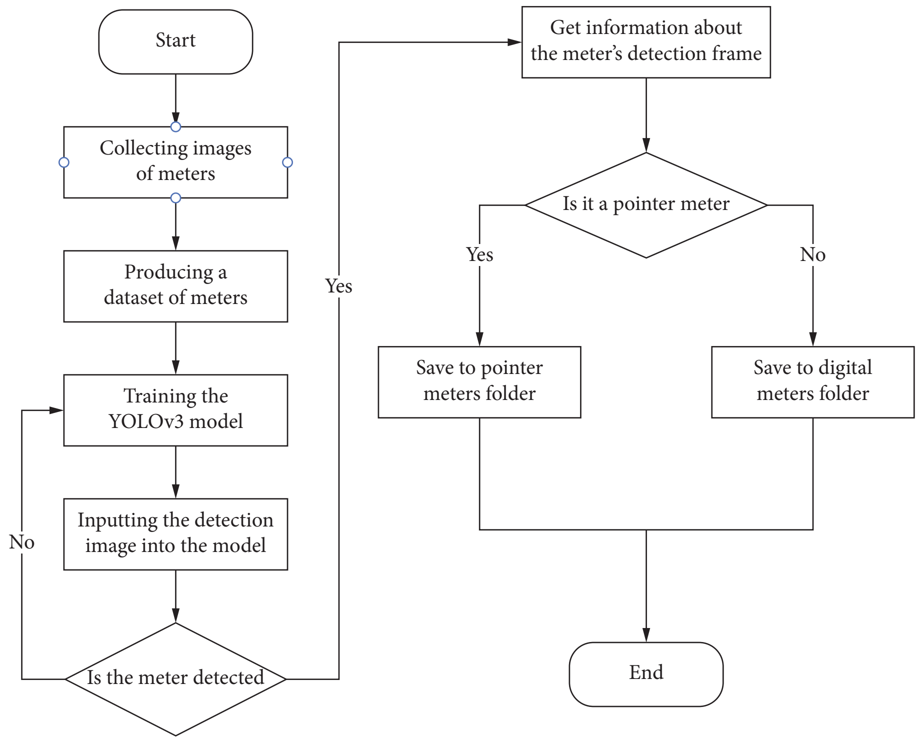


Figure 1: Algorithm flow diagram of the industrial instrument classification module.

After the images are annotated, the annotated file is divided into a training set and a test set, and the test set is fed into the YOLOv3 network [19–22] for training. The number of training rounds (epochs) is set to 200, the batch size of each iteration is 2, and the first 249 layers are unblocked so that they can be trained together.

3.3. Experimental Results and Analysis. Two hundred different meter images were selected for input into the system to test the recognition effect of the model, and some of the recognized images are shown in Figure 4.

The loss during training was recorded and plotted as a loss curve and compared with the model used in [1]. As can be seen from Figure 5, the YOLOv3 model [23–25] used in this study has a lower loss value and is better able to achieve the recognition of industrial meter types.

To test the recognition effect of the model more accurately, the Precision-Recall curve (i.e., P-R curve), with Recall as the horizontal coordinate and Precision as the vertical coordinate, is chosen to evaluate the results after recognition in this study; the results are shown in Figure 6.

As can be seen from Figure 6, the Average Precision (AP) of pointer instrument recognition in the test set is 93.62%, the AP of digital instrument recognition is 93.83%, and the mean Average Precision (mAP) of the industrial instrument recognition model based on YOLOv3 is 93.73%, which is 0.26% higher than [1] and has higher identification accuracy.

4. Reading Recognition of Pointer Meters

Reading recognition of pointer meters is described in the following sections.

4.1. Image Processing of Pointer Meters. The process of image processing of pointer meters is as follows.

4.1.1. Mean Value Filtering. Mean filtering means converting a pixel point (x, y) in an image to the mean of the pixel values around that pixel point, thus obtaining the grey scale $g(x, y)$ of the current point on the image, by accessing all the pixel points in the image, in turn, to achieve the mean filtering operation on the image. The basic principle [26–28] is as follows:

$$g(x, y) = \frac{1}{m}\sum f(x, y). \tag{2}$$

For pointer meters, the information on the dial contains a variety of information such as scale lines, scale values, and hands. The mean filtering of the pointer meter image can well remove the irrelevant information on the dial and facilitate the subsequent extraction of key information on the dial.

4.1.2. Greyscaling. Greyscaling is the conversion of a colourful image into a grey-only image. By means of greyscaling, the contrast of the image is increased and the

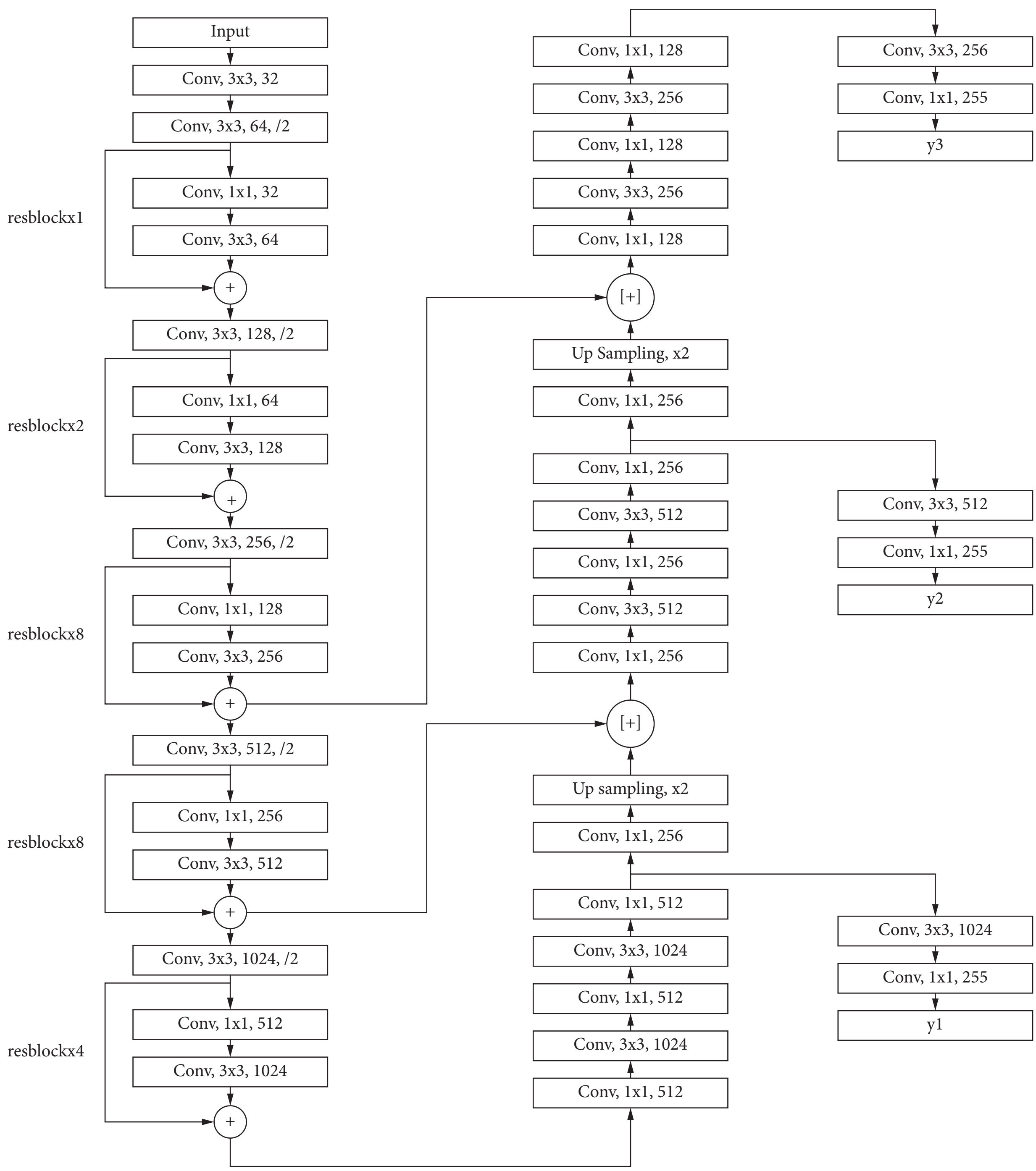


Figure 2: Structure of the YOLOv3 model.

dynamic range of the pixels is expanded, making the whole image finer, clearer, and easier to recognize for computer processing. The calculation process for greyscale is as follows:

$$\text{gray} = \sqrt[2.2]{\frac{R^{2.2} + (1.5G)^{2.2} + (0.6B)^{2.2}}{1 + 1.5^{2.2} + 0.6^{2.2}}}. \tag{3}$$

After greyscale processing, it is possible to obtain a greyscale image that is closer to human vision and contains only one color information per pixel in greyscale. The processed pointer meter image is shown in Figure 7.

4.2. Position the Dial of a Pointer Meter. Given that the dials of pointer meters are circular, detection of circular dials in meter images can be performed using the Hough circle

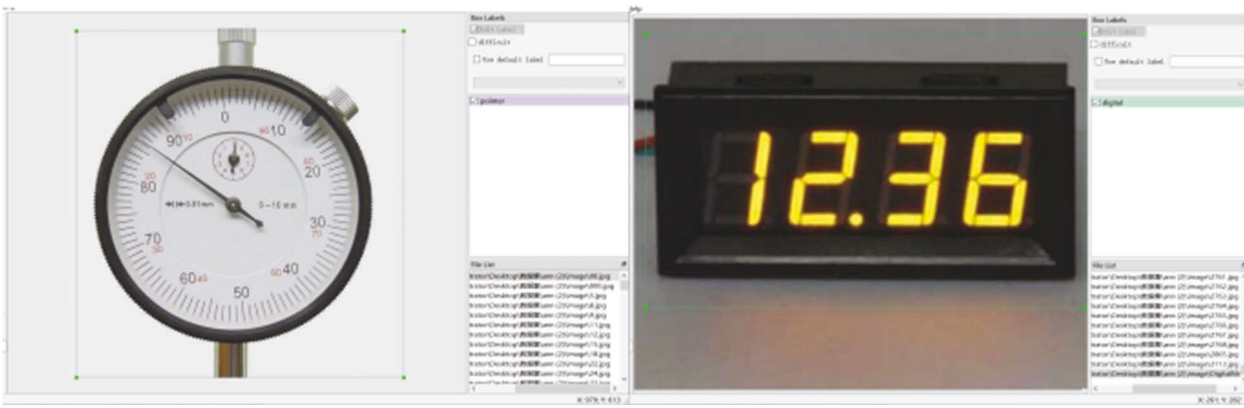


Figure 3: Image annotation effect.

Pointer Meters

Digital Meters

Figure 4: Recognition effect.

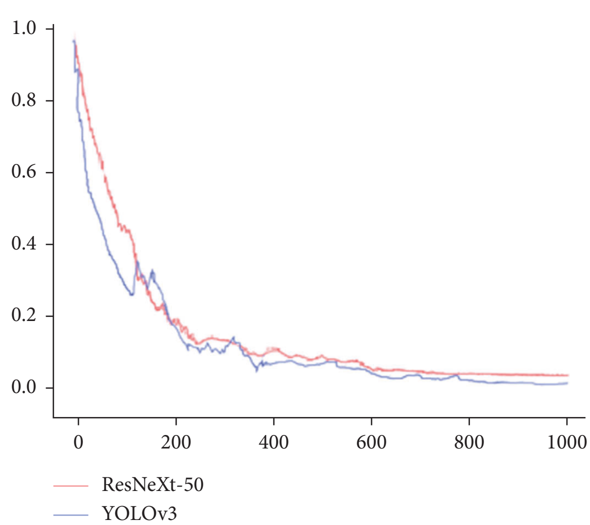
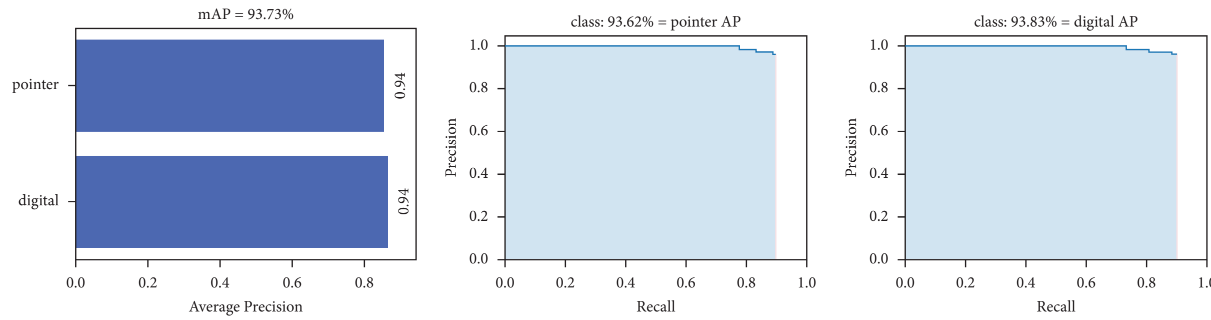
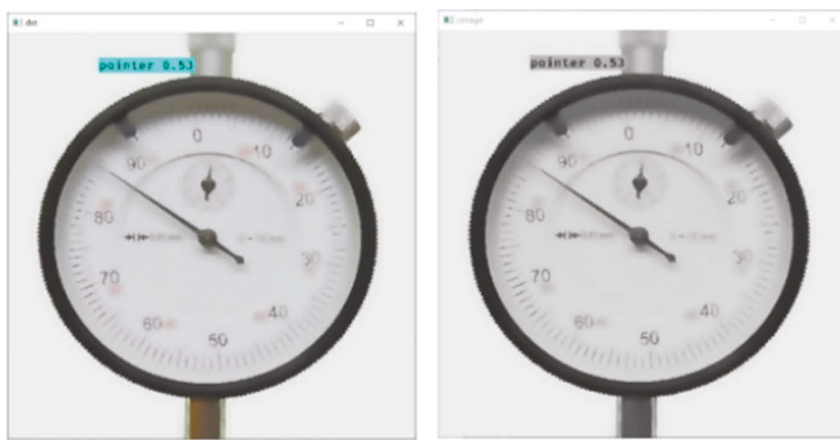


Figure 5: Loss curve comparison chart.



map pointer AP digital AP

Figure 6: Evaluation indicators in the training process.



Mean value filtering Greyscaling

Figure 7: Pointer meter image processing.

transformation, a feature-based digital image processing technique that extracts circles from the image. Candidate circles can be generated by voting to construct a cumulative coordinate plane in the parameter space and select the local maximum value of the accumulator matrix.

The equation of a circle in the Cartesian coordinate system is

$$(x-a)^2 + (y-b)^2 = r^2, \tag{4}$$

where the coordinates of the center of the circle are (a, b) and the radius is r, so equation (4) can also be expressed as

$$\begin{aligned} a &= x - r\cos\theta, \\ b &= y - r\sin\theta. \end{aligned} \tag{5}$$

So, a circle in a three-dimensional coordinate system a-b-r can be defined by a single point (a, b, r). If the circles in the image all pass through a certain pixel point, it is possible to transform these circles into curves in polar coordinate space. Multiple circles will intersect curves in polar space, and a threshold is set for these intersection points, above which the points are identified as circles.

4.3. Dial Information Extraction. The dial of a pointer meter contains a lot of important information, mainly the scale lines, the scale values, and the pointer. In order to achieve the reading of the meter, the information on the dial needs to be identified and extracted in steps, removing the distracting information on the dial and retaining the important information on the annunciator. By identifying the retained information, the reading function of the pointer meter is thus realized.

4.3.1. Extracting the Scale. The tick marks on the dial can be extracted by finding the tick outline, based on the characteristics of the tick marks. The principle is to find the black pixel and define it as the starting pixel, which is retraced each time it touches the black pixel and then goes clockwise around the black pixel visiting each pixel in its molar neighborhood until it touches the black pixel. The effect of extracting the tick marks is shown in Figure 8.

4.3.2. Locating the Center of the Circle. After selecting the scale lines, the least square method is used to fit the scale lines, and the fitting lines on the dial of the instrument can be obtained. As shown in Figure 9, the center of the dial can be obtained by taking the average of the intersection points of all the scale lines.

Figure 8: Extraction scale.

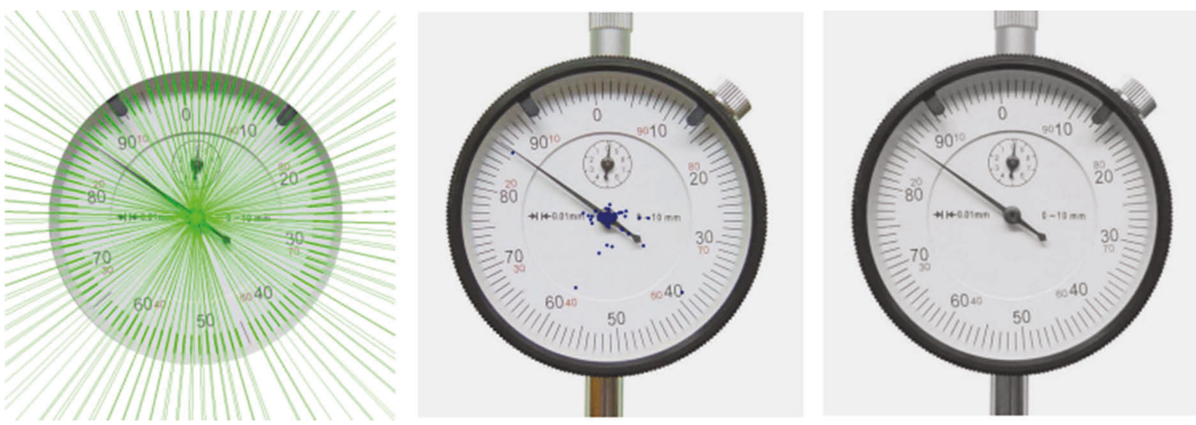


Figure 9: Positioning the center of the circle.

4.3.3. Extracting Pointers. Based on the analysis of the characteristics of pointer meters, the Hough linear transform was used to identify the pointer in the dashboard and to extract refinement of the pointer. The essence of the Hough linear is a mapping of the coordinates of the pixel points in a two-dimensional image, represented as a curve in polar coordinate space.

If the composition of pixels in an image intersects at a point, then all pixels in the image can be transformed into a curve in parametric coordinate space. After the transformation, each point (x, y) in the image space is mapped to a sine curve in (r) polar coordinate space. If the points in image space all intersect after the transformation into polar coordinate space, and the intersection points are all the same point, the parameter corresponding to that point is the parameter corresponding to the parametric equation of the line, i.e., the parameter of the line on which the pointer is located.

4.4. Calculation of Angles. The output of the meter readings is obtained by calculating the angle of rotation of the pointer and the angle of rotation of the total range and comparing them to obtain the final reading, where the position of the 0 scale is vector of the center of the circle coordinates $O(x, y)$ pointing to the start of the scale $A(x, y)$; the position of the pointer is vector of the center of the circle coordinates $O(x, y)$ pointing to the end of the pointer $B(x, y)$, and the position of the maximum range is vector of the center of the circle coordinates $O(x, y)$, pointing to the end of the pointer $C(x, y)$.The angle of rotation of the pointer is, the angle of rotation of the total range is; then, the formula is

$$
\begin{aligned}
\theta_1 &= \langle \vec{v}_0, \vec{v}_1 \rangle = \cos^{-1}\left(\frac{\vec{v}_0 \cdot \vec{v}_1}{\left\|\vec{v}_0\right\|_2 \cdot \left\|\vec{v}_1\right\|_2}\right), \\
\theta_2 &= \langle \vec{v}_0, \vec{v}_2 \rangle = \cos^{-1}\left(\frac{\vec{v}_0 \cdot \vec{v}_2}{\left\|\vec{v}_0\right\|_2 \cdot \left\|\vec{v}_2\right\|_2}\right).
\end{aligned}
\tag{6}
$$

4.5. Experimental Results and Analysis. Pictures of pointer meters with different values were entered into the system to test the pointer meter reading function, and some of the test results are shown in Table 1.

As can be seen from Table 1, the absolute error of the pointer meter reading method used in this paper is generally less than 0.1, and the maximum relative error is 0.35%. However, the maximum relative error of the mBR-based method in literature [6] is 0.67%, that of the method based on Otsu and improved Hough transform in [7] is 0.9%, and that of [2] is 1.5%. By comparison, the error of the pointer meter reading method adopted in this study is small and can meet the needs of pointer meter reading.

Table 1: Results of the pointer meter reading function test.

Sample number	Actual meter values	System reading results	Absolute error	Relative error (%)
1	33.2	33.14309709	0.05690291	0.17
2	15.3	15.35345867	0.05345867	0.35
3	5.3	5.311204024	0.011204024	0.21
4	26.8	26.84405347	0.04405347	0.16
5	28.3	28.39544074	0.09544074	0.34
6	0	0	0	0
7	7.5	7.511302457	0.01130246	0.15
8	26.3	26.24309709	0.05690291	0.22
9	14.7	14.64890989	0.05109011	0.35
10	33.4	33.36257512	0.03742488	0.11

Greyscaling Gaussian filtering Binarization

Figure 10: Digital meter image processing.

5. Digital Meter Reading Recognition

Digital meter reading recognition is described in the following sections.

5.1. Image Processing of Digital Meters. The process of image processing of digital meters is as follows.

5.1.1. Gaussian Filtering. Gaussian filtering is used to filter out normally distributed noise (Gaussian noise). The value of each pixel is weighted by the value of the same pixel and the values of neighboring pixels; the closer to the center, the higher the weight. The effect of Gaussian filtering is gentler compared to other filters and is able to maintain the invariance of image changes such as rotation and lighting changes. It is calculated as follows:

$$g(x, y) = \frac{1}{2\pi\sigma^2} e^{-x^2+y^2/2\sigma^2}. \quad (7)$$

The dial information on digital meters is relatively simple, with essentially only the digits and decimal point. The Gaussian filtered image of the meter still retains the digits and decimal point intact and has no effect on subsequent readings.

5.1.2. Binarisation. Binarisation is the setting of the pixel points in an image according to their grey scale values, converting the original greyscale map into a binary image containing only black and white information, i.e., only two values of 0 and 255. Binarisation is calculated as follows:

$$g = \omega_1 * \omega_2 * (\mu - \mu_2)^2. \quad (8)$$

Image binarisation removes distracting information from the digital dial, extracting information from the image more accurately, and retaining clearer numerical and decimal information.

The processed digital meter images are shown in Figure 10.

5.2. Digital Recognition. The SVM dataset was created from images collected on the Internet and contains 400 images from 0 to 9.

This study utilizes the histogram (HOG) method which has now been widely used for complex facial and gesture recognition. First, the Sobel derivative values of the horizontal and vertical sides of the image are calculated to obtain the gradient angle and gradient size of each pixel, and then, the angle of the gradient is converted to an integer between 0 and 16. The image is divided into four $10 * 10$ blocks and the histogram of the slope of each block is derived using the gradient value of that block as the weighted value. Each small cube is then represented by 16 vectors, and the whole image is then represented by the eigenvectors of the four blocks.

The experimental results show that the method can effectively eliminate background interference and extract 64 feature values. However, the extracted feature values are all too large for good recognition. So, the gradient histogram can be transformed by the Hellinger kernel function, which is calculated as follows:

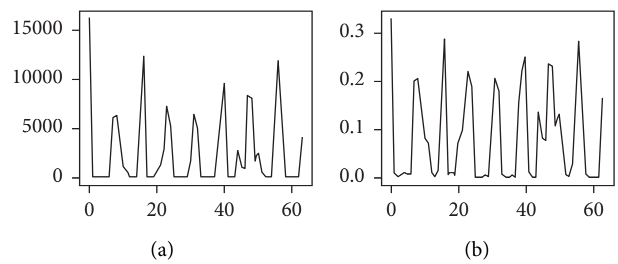


Figure 11: Hellinger kernel transformation. (a) Before the change. (b) After the change.

Table 2: Digital meter reading function test results.

Sample number	Actual meter values	System reading results	Absolute error
1	12.36	12.36	0
2	24.2	24.2	0
3	13.57	13.57	0
4	46.53	46.53	0
5	4.75	4.75	0
6	87.09	87.09	0
7	0	0	0
8	65.85	65.88	0.03
9	24.17	24.17	0
10	10.65	10.65	0

Table 3: Digital instrument recognition accuracy.

Methods	Accuracy (%)
MNIST dataset recognition method based on Keras framework [1]	99.13
Pixellink + CRNN [2]	99.3
GrayWorld + deform-conv [5]	99.45
The method used in this study	99.7

$$k_{HK}\left(x_i, x_j\right) = \sum_{q=1}^{d} \sqrt{x_{iq} x_{jq}}. \quad (9)$$

As shown in Figure 11, after the Hellinger kernel transformation, these data were transformed from the largest number to the smallest decimal and preserved their feature information well enough to be roughly viewed as a normalization process.

5.3. Experimental Results and Analysis. The digital meter reading function was tested by entering pictures of digital meters with different values into the system. Some of the test results are shown in Table 2.

The algorithm used in this study for digital instrument reading is compared with [1, 2, 5]. As can be seen from Table 3, the accuracy of the digital instrument reading method adopted in this study is 1.2% higher than [1], 0.3% higher than [2], and 0.25% higher than [5], with higher accuracy and more accurate reading results.

6. Conclusions

This study presents a YOLOv3-based classification and reading recognition method for Industrial Instrument, which achieves target detection of meters by training a YOLOv3 neural network model for the type recognition of meters. The Hough transform is used to locate the dials and hands of pointer meters, and then, the angle between the scale 0 mark and the readings of pointer meters are calculated. The dial of a digital meter is located by contour finding, the digital area is located, the digits are segmented, the segmented digits are identified using SVM, and the final output is identified and read.

The biggest advantage of this method is the ability to identify the type of industrial meter, classify the type of meter, and call different meter reading modules for the classified meter to read and identify it. The system can identify both pointer type meters and digital meters, where the mAP of the meter type identification model is 93.73%, the relative error of pointer meter reading is less than 3.5%, the error is small, and the accuracy of the digital meter reading is 99.7%, and the result is accurate.

Data Availability

The datasets analyzed during the current study are available in the paddle AI Studio repository (https://aistudio.baidu.com/aistudio/datasetdetail/157981) and public domain resources (https://aistudio.baidu.com/aistudio/datasetdetail/137555 and https://aistudio.baidu.com/aistudio/datasetdetail/124339).

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

Acknowledgments

This work was sponsored by the following: (1) the Training Project of Top Scientific Research Talents of Nantong Institute of Technology, under Grant no. XBJRC2021005, (2) the Science and Technology Planning Project of Nantong City, under Grant nos. JC2021132, JCZ21058, JCZ20172, JCZ20151, JCZ20148, and MS22021028, and (3) the Universities Natural Science Research Projects of Jiangsu Province, under Grant nos. 17KJB520031, 21KJD210004, 22KJB520032, 22KJD520007, 22KJD520008, and 22KJD520009.

References

[1] B. Zhang, *Research and Application of Automatic Recognition of Industrial Instrumentation Numbers Based on Machine Vision (MS Thesis)*, Taiyuan University of Science and Technology, Taiyuan, China, 2020.

[2] F. Lei, *Research on Automatic reading Method of Industrial Instruments Based on Machine Vision (MS Thesis)*, Taiyuan University of Science and Technology, Taiyuan, China, 2021.

[3] V. Kucheruk, I. Kurytnik, P. Kulakov, and J. Chen, “Definition of dynamic characteristics of pointer measuring devices on the basis of automatic indications determination,” *Archives of Control Sciences*, vol. 28, no. 3, pp. 401–418, 2018.

[4] G. Chen, Z. Hu, and C. Zheng, “Digital fast recognition algorithm for digital instrument based on feature detection,” *China measurement & testing*, vol. 45, no. 04, pp. 146–150, 2019.

[5] L. Guo, R. Han, and K. Xiao, “Digital instrument recognition based on deformable convolutional Neural Network,” *Computer science*, no. 10, pp. 187–193, 2020.

[6] Z. Zhang, F. Wang, E. Tian, and T. Chen, “Machine vision based reading recognition for a pointer gauge,” *Control Engineering China*, vol. 27, no. 3, pp. 581–586, 2020.

[7] Na Li, J. Wang, X. Wang, and S. P. Dong, “Pointer meter recognition method based on improved Hough transform,” *Chinese Journal of Liquid Crystals and Displays*, vol. 36, no. 8, pp. 1196–1203, 2021.

[8] R. R. Sowah, A. R. Ofoli, E. Mensah-Ananoo, G. A. Mills, and K. M. M. Koumadi, “An intelligent instrument reader: using computer vision and machine learning to automate meter reading,” *IEEE Industry Applications Magazine*, vol. 27, no. 4, pp. 45–56, 2021.

[9] J. Tang, Bo Tang, X. Fu, and A. Wang, “Automatic detection technology of instrument reading based on image recognition,” *Electronic Devices*, vol. 44, no. 3, pp. 725–730, 2021.

[10] L. Wang, P. Wang, L. Wu, L. Xu, P. Huang, and Z. Kang, “Computer vision based automatic recognition of pointer instruments: data set optimization and reading,” *Entropy*, vol. 23, no. 3, p. 272, 2021.

[11] Z. Wang, H. Zhu, X. Jia, Y. Bao, and C. Wang, “Surface defect detection with modified real-time detector YOLOv3,” *Journal of Sensors*, vol. 2022, Article ID 8668149, pp. 1-10 pages, 2022.

[12] L. h Li, J. c Hang, Y. Gao, and Cy Mu, “Using an integrated group decision method based on SVM, TFN-RS-AHP, and TOPSIS-CD for cloud service supplier selection,” *Mathematical Problems in Engineering*, vol. 2017, Article ID 3143502, pp. 1-14 pages, 2017.

[13] Y. Li and L. H. Li, “Enhancing the optimization of the selection of a product service system scheme: a digital twin-driven framework,” *STROJNISKI VESTNIK-JOURNAL OF MECHANICAL ENGINEERING*, vol. 66, no. 9, pp. 534–543, 2020.

[14] D. Shi and H. Tang, “A new multiface target detection algorithm for students in class based on bayesian optimized YOLOv3 model,” *Journal of Electrical and Computer Engineering*, vol. 2022, Article ID 4260543, pp. 1-12 pages, 2022.

[15] H. Zhai and Y. Zhang, “Target detection of low-altitude UAV based on improved YOLOv3 network,” *Journal of Robotics*, vol. 2022, Article ID 4065734, pp. 1-8 pages, 2022.

[16] L. Li, C. Mao, H. Sun, Y. Yuan, and B. Lei, “Digital twin driven green performance evaluation methodology of intelligent manufacturing: hybrid model based on fuzzy rough-sets AHP, multistage weight synthesis, and PROMETHEE II,” *Complexity*, vol. 2020, no. 6, Article ID 3853925, pp. 1-24 pages, 2020.

[17] H. Wang, L. Dong, H. Zhou et al., “YOLOv3-Litchi detection method of densely distributed litchi in large vision scenes,” *Mathematical Problems in Engineering*, vol. 2021, Article ID 8883015, 11 pages, 2021.

[18] L. Li, J. Hang, H. Sun, and L. Wang, “A conjunctive multiple-criteria decision-making approach for cloud service supplier selection of manufacturing enterprise,” *Advances in Mechanical Engineering*, vol. 9, no. 3, Article ID 168781401668626, 2017.

[19] W. Wang, B. He, and L. Zhang, “High-accuracy real-time fish detection based on self-build dataset and RIRD-YOLOv3,” *Complexity*, vol. 2021, Article ID 4761670, 8 pages, 2021.

[20] Q. Yang and L. Yu, “Recognition of taxi violations based on semantic segmentation of PSPNet and improved YOLOv3,” *Scientific Programming*, vol. 2021, Article ID 4520190, 2021.

[21] L. Li and C. Mao, “Big data supported PSS evaluation decision in service-oriented manufacturing,” *IEEE Access*, vol. 8, pp. 154663–154670, 2020.

[22] D. Chen, S. Sun, Z. Lei, H. Shao, and Y. Wang, “Ship target detection algorithm based on improved YOLOv3 for maritime image,” *Journal of Advanced Transportation*, vol. 2021, Article ID 9440212, 11 pages, 2021.

[23] B. Zhang, G. Wang, H. Wang, C. Xu, Y. Li, and L. Xu, “Detecting small Chinese traffic signs via improved YOLOv3 method,” *Mathematical Problems in Engineering*, vol. 2021, Article ID 8826593, 10 pages, 2021.

[24] L. Li, B. Lei, and C. Mao, “Digital twin in smart manufacturing,” *Journal of Industrial Information Integration*, vol. 26, no. 9, Article ID 100289, 2022.

[25] C. Liu and J. Li, “Self-correction ship tracking and counting with variable time window based on YOLOv3,” *Complexity*, vol. 2021, Article ID 2889115, 9 pages, 2021.

[26] S. Li, H. Zhao, and J. Ma, “An edge computing-enabled train obstacle detection method based on YOLOv3,” *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 7670724, 9 pages, 2021.

[27] L. Li, T. Qu, Y. Liu et al., “Sustainability assessment of intelligent manufacturing supported by digital twin,” *IEEE Access*, vol. 8, pp. 174988–175008, 2020.

[28] H. Dou and G. Wang, “A dynamic multitarget detection algorithm in front of vehicle based on embedded system and internet of things,” *Scientific Programming*, vol. 2022, Article ID 3585162, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 1604184, 11 pages
https://doi.org/10.1155/2022/1604184

Research Article

Study on Optimization of Marketing Communication Strategies in the Era of Artificial Intelligence

Lingying Wen,[1] Wen Lin,[2] and Mingde Guo[2]

[1]*Jiangxi University of Applied Science, Nanchang 330100, China*
[2]*Business College, Jiangxi Institute of Fashion Technology, Nanchang 330201, China*

Correspondence should be addressed to Mingde Guo; guomingde@jift.edu.cn

Received 25 July 2022; Revised 1 September 2022; Accepted 16 September 2022; Published 11 October 2022

Academic Editor: R. Mo

This paper starts from the research of computational advertising and scene theory, based on the background of artificial intelligence technology application, and studies the communication strategy of the computational advertising scene. It studies the three major strategies of computing advertising scene communication and deeply analyzes the principles and applications of scene insight strategy, content selection strategy, and community operation strategy. Results show the following: (1) under the background that the era of artificial intelligence has arrived, computing advertising uses artificial intelligence algorithms to complete breakthrough upgrades. Through the combination of algorithms and data, its intelligent upgrades are mainly manifested in three aspects: First, to achieve higher matching accuracy communication, the second is efficient customized communication, and the third is to realize the contextual interaction between advertisements and users; (2) the "intelligent" performance characteristics of intelligent scene dissemination of computing advertising are mainly in the application of scene intelligence technology, the intelligence of the data platform, and the intelligent construction of user portraits; (3) in the era of artificial intelligence, the communication strategies for computing advertising scenarios mainly include intelligent scenario insight strategies, content selection strategies, and community operation strategies. First, the intelligent scene insight strategy is mainly analyzed from two levels. On the one hand, it is a scene mining based on intelligent data and an in-depth analysis of the user tag system centered on intelligent algorithms. On the other hand, through the research on keyword search, the impact of intelligent upgrade on scenario insight is analyzed. Secondly, in terms of content selection strategy, the application of artificial intelligence technology has brought a new upgrade to the creation and recommendation mechanism of computational advertisements.

1. Introduction

With the development and popularization of Internet technology, computational advertising plays an increasingly important role in the advertising industry. Since the development of the Internet, it has experienced four evolutions: the era of fixed networks, the era of big data, the era of mobile Internet, and the era of artificial intelligence [1]. In this process, computational advertising has developed from the initial contract advertising, bidding advertising, search advertising, and other basic forms to the stage of data-driven programmatic advertising and the current stage of artificial intelligence. Based on the national strategic guidance plan and industry development trend, it is of importance to gain an in-depth understanding of the application of artificial intelligence technology in the Internet field, and explore the intelligent optimization path of the scene matching strategy of computing advertising, conduct a comprehensive analysis of real-time positioning scenarios, fragmented scenarios, and full-scenario matching modes, and summarize industry experience to provide a practical reference for the development of computational advertising. It provides a contemporary reference and reference value for the theoretical research of computational advertising [2].

Regarding the literature review of computational advertising, the author started from the concept of computational advertising that was first proposed by Andrei Broder, a senior researcher at Yahoo Research, at the 19th

ACM-SIAM Symposium in 2008 [3]. Scholars Wu and Hu [4] proposed at the IEEE conference to use a computational model to reconstruct the spread of advertisements in social networks and used this computational model to conduct a series of virtual experiments to obtain information on how to select an initial set of people to maximize observations and experimental implications of advertising communication effects. Since then, scholars' research on computational advertising has focused on the exploration of computational logic. The rapid update of Internet technology has continuously promoted the development of computing advertising in the direction of intelligence, especially the practical application of new theories and technologies such as cloud computing and artificial intelligence in the field of computing advertising [5]. It seems that computing advertising is about to usher in subversive changes. It can be seen that computational advertising is gradually becoming the most important part of the advertising industry in the Internet era. With the popularization of smart devices, traditional advertising has shown signs of fatigue, and computational advertising based on big data and artificial intelligence algorithm models has shown strong vitality [6]. The machine can quickly gain insight into the needs of consumers in real-time through accurate user portraits, and combine itself with the user's usage scenarios when computing advertisements are automatically pushed, effectively coping with the famous problem raised by advertising master John Wanamaker: "I Knowing that half of my advertising dollars are wasted, but the problem is I don't know which half."

Scene theory has produced many fruitful research results in the fields of culture, urban space, film art, and media [7]. In the literature review section, this paper mainly focuses on the literature review and elaboration of scene research in the fields of advertising, marketing, communication, and the Internet [8]. In the 1980s, when television media was popularized and covered in developed countries in Europe and America, McLuhan put forward the media scene theory based on his profound insight into the relationship between media and human behavior. In this theory, famous viewpoints such as "media is the extension of man" are included, which believe that the media changes the relationship of time and space so that people's perception ability is strengthened or extended. Merowitz's research on scenes is different from the material scenes that show space and believes that it is not the material site itself that determines the nature of people's interactions, but the pattern of information flow [9–11]. Extend the definition of the scene to the virtual space created by electronic media. In the physical space under the traditional concept and under the influence of the flow of media information, the virtual scenes created by the electronic media have gradually attracted the attention and research of scholars. The current society has entered the mobile Internet era, and the boundaries between real and virtual scenes have become more ambiguous.

In the context of the Internet and big data, a new advertising method driven by the industry, namely, computing advertising has come into vogue. According to public information on the Internet, Andrei Broder, who proposed the concept of computational advertising, pointed out the problems that computational advertising needs to solve from the perspective of the industry, and defined "computational advertising is to find a suitable advertisement for a user U under a given scenario C, in order to achieve" Optimal "match" [12]. Computational advertising has its unique characteristics in the selection of target groups, content production and delivery, advertising matching, and effective monitoring [13]. Different from traditional advertising, computational advertising is based on Internet technology to achieve process-based and efficient operation, and accurately convey it to target user groups [14]. The application of computational advertising to big data technology and artificial intelligence algorithms has formed an industrial scale. Internet giants at home and abroad have taken advantage of the huge traffic advantages of their platforms to obtain large-scale revenue from the advertising business every year. In China, Internet giants such as BAT have accumulated a large amount of user data after long-term development, and the application of artificial intelligence technology is constantly being strengthened [15].

Driven by artificial intelligence technology, big data, intelligent algorithms, and communication scenarios are deeply integrated through machine learning technology, and important breakthroughs have been made in product insights and consumer dynamics [16]. In the digital age, streaming media platforms focus more on data analysis to find the target audience that best meets product needs. Traditional algorithms draw user portraits by analyzing demographic attributes such as regional factors, age, and occupation, which traditional advertising relies on, as well as user behavior, but there are inevitably some unsolvable drawbacks [17]. For example, by optimizing the click-through rate of the target audience, media buyers can attract a lot of audiences that match the user's label to the traffic platform, but for the advertiser, these users are not necessarily the potential consumption of the product [18]. In contrast, digital advertising strategies that use artificial intelligence to identify and target potential customers without prejudice have a greater chance of finding real consumption, which is different from the segmented audience given by big data.

Artificial intelligence technology has brought disruptive changes to the creation of advertising content. For a long time, the creative thinking of advertisers is the only way to complete the creation of advertising content [19]. Due to the limitations of human thinking and energy, the creativity and creative content of advertisements are relatively limited [20]. At present, the demand for advertising creativity in China is increasing day by day, and the contradiction between supply and demand that advertisers create is low in quantity and quality. At present, the potential of artificial intelligence advertising content creation is far from being fully exploited and is currently only used to accomplish some simple goals [21]. For example, in terms of copywriting creation, the artificial intelligence copywriting recommendation system can automatically recommend the use of titles and description texts through a certain number of titles and description texts given by advertisers, which can not only help advertisers achieve the highest ROI but also save a lot of money, test time, and effort.

The traditional way of advertising can only be done by buying media and "casting a wide net," which is not only expensive but also difficult to grasp the accuracy of advertising [22]. Even in the Web 2.0 stage, programmatic advertising has greatly improved the accuracy of delivery, but computational advertising based on traditional algorithms, still has many shortcomings, such as the processing of true and false information from data sources, transparency of advertising, and false traffic, advertising effect evaluation,etc. [23]. The basis for intelligent recommendation of computational advertisements is to accurately divide the Internet population and create user portraits based on data management. The needs of Internet users are showing a trend of personalization and autonomy. Traditional algorithms are difficult to adapt to the development of the times, and they are increasingly unable to capture the purchasing habits of Internet users [24]. The number of Internet advertisements is extremely large, and the effect of advertisements needs to be monitored dynamically in real time. This link involves the technical application of various roles in the advertising chain, such as the real-time bidding technology related to DSP, the inquiry optimization technology related to the advertising exchange platform (ADX), and the network optimization technology related to the supply side platform (SSP). Online advertisements frequently require hundreds of millions of data requests every day, as well as high requirements for timeliness [25]. User portraits need to be updated in real time according to changes in each user's characteristics, and advertising creatives need to be adjusted. Artificial intelligence machine learning technology guarantees real-time updates of data, makes adjustments to user portraits, and formulates advertising strategies in real time to maintain optimal advertising effects.

With the advent of the mobile Internet era, the combination of people and scenes has become the development trend of computing advertising. The application of digital technology has always been the driving force for the development of computing advertising. The artificial intelligence algorithm intelligently upgrades the construction of user portraits, and promotes computing advertising into a scene-based communication mode. Based on the application of artificial intelligence technology, this paper studies the scenario communication strategy of computational advertising. First of all, based on the application of artificial intelligence technology, using "scenario five forces" and scenario theory to build an intelligent scenario communication strategy model. Then, it discusses the scenario communication strategy, and analyzes the three strategies of scenario insight, content selection, and community operation. At the same time, by analyzing relevant typical intelligent computing advertising communication cases, the practical application of scenario communication strategies in user usage scenarios is summarized and analyzed. Finally, the problems and deficiencies of the current computing advertising scene dissemination are sorted out and analyzed.

The research method of this paper mainly adopts the case analysis method, takes computational advertising as the research object, and focuses on the operation mechanism of the communication strategy of computational advertising in the era of artificial intelligence. Based on the field of computational advertising, this paper summarizes the four elements of scene communication, and studies the dimensions of scene time and space, communication content, scene connection medium, and user network behavior.

2. Research Ideas, Methods, and Technology Support

2.1. Research Ideas. The research of this paper begins with the author's personal experience of the rapid development of Internet advertising in recent years. After more than 20 years of development, domestic Internet advertising has rapidly become a new force in the advertising industry. The application practice of Internet advertising in programmatic marketing and big data marketing involves a lot of technical and theoretical knowledge of algorithms and machine learning. By collecting and sorting out relevant literature on the development of various Internet advertisements, the research object is preliminarily determined as computational advertising [26].

Computational advertising requires accurate personalized recommendations, and its core essence is to recommend the right advertisement to the right target group at the right time. The most effective way to realize this recommendation process is to combine people and scenes. The application of artificial intelligence technology allows the machine to place the most suitable advertisements through the user portrait model based on big data and real-time user mobile scene positioning. Therefore, the author determines the core of the research as the scene-based communication strategy of computational advertising [27].

Combining the previous research methods and research results, the research idea of this paper is roughly as follows: firstly, the domestic and foreign-related researches on computational advertising and scene theory are organized into a literature review; secondly, the basic concepts of computational advertising are sorted out and analyzed, and artificial intelligence technology is determined [28]. It clearly describes the role and impact of artificial intelligence on computational advertising; furthermore, it establishes a model for the principle of scene matching between computational advertising and audiences, and analyzes the elements of scene communication and the modeling mechanism. Then, combined with the theoretical basis of the previous article, the scene communication strategy of computing advertising is cut into three dimensions, namely, scene insight, content selection, and community operation, and the principle of the strategy is deeply analyzed [29]. At the same time, the practical application of the strategy is analyzed. Finally, it summarizes the existing problems of the current development of computational advertising and looks forward to its future.

2.2. Research Methods. Case study. This paper conducts a case study on the scenarios of computing advertising reach, focusing on the relevant cases of domestic Internet companies using scenario communication strategies for advertising communication. For different strategies, the author selects the cases of computing advertising communication of typical Internet companies such as BAT, as well as advertising communication cases based on artificial intelligence algorithms such as Toutiao Information Streaming Advertising and Google Advertising, and analyzes the specific strategy application and communication rules of computing advertising scene communication [30]. In addition, based on the "Five Forces of Scenarios" proposed by Robert Scober and Shell Israel and relevant domestic and foreign scenario theory research, the author constructs a model of computational advertising scenario communication strategies in the era of artificial intelligence. The calculated advertising communication data, communication paths, and industry-related scene communication strategies, etc., are used to further verify whether the selected cases conform to the model.

2.3. Technology Support. A user portrait is a labeled user model abstracted from information such as using social attributes, living habits, and consumption behavior. In the dissemination system of computational advertisements, user portraits are abstracted as the characteristic attributes of real users, which help the platform to realize the precise delivery and personalized recommendation of computational advertisements according to user characteristics and their preferences. This paper starts from the communication dimension of computational advertising, discusses the model construction of user portraits and its relationship with advertising communication, and deeply understands the important role that user portraits play in computational advertising. The concept of user portrait (persona) was first proposed by "Father of Interaction Design" Alan Cooper. Alan Cooper believes that user portraits are virtual representations of real users and target user models based on a series of attribute data [31].

With the development of the Internet, user portraits have been given new connotations based on information such as user demographics, online browsing content, online social activities, and consumption behaviors. The author sorts out and summarizes relevant literature and research, and summarizes the basic elements of user portraits into three: basic data, algorithm models, and user tags. The collection of basic data is the premise of depicting user portraits. In the current Internet environment, the establishment of a massive user database requires the use of big data technology to obtain user data. The algorithm model is the main tool to realize the full-link propagation of computational advertisements [32]. In fact, user portrait modeling is only one of the tasks that algorithm models need to complete, but process that describe user portraits, from the cleaning and filtering of massive raw data to data statistics and classification, and then to machine learning and text analysis, this process requires algorithmic models to complete. User tags are the core elements of user portrait construction. As a highly refined feature identifier specified by humans, tags are the most important output of user portraits, which make the user portrait model have practical significance.

In terms of computing advertising, the core work of constructing user portraits is mainly to use massive logs stored on the server and a large amount of data in the database for analysis and mining to create a "tag system," each tag can represent the ID of the dimension feature for the user. The process of user portrait modeling is mainly divided into three processes: basic data collection, behavior modeling, and portrait construction, as shown in Figure 1.

Through the application of artificial intelligence technology, the data value density of user portraits and the accuracy of personalized recommendations can be improved. AI systems perform cognitive functions, using machine learning techniques, to train how to act or react to an outcome and know to take the same action in the future.

The process of scene dissemination of computational advertisements needs to be supported by the underlying architecture such as technical foundation and data foundation [33]. In this process, multiple scene elements need to be coordinated to complete propagation at the same time. In the context of the application of artificial intelligence technology, the construction elements of the scene communication of computational advertising should include scene time and space, a scene connection medium, scene communication content, and user network behavior. On this basis, other factors will be introduced in the dissemination scenarios of computing advertisements, such as AR, VR glasses, smart wearable devices, etc., that enhance the scene experience. Based on the above analysis, we constructed a computational advertising scenario propagation model based on artificial intelligence technology (see Figure 2).

From this model, with the support of artificial intelligence technology, computing advertising is based on the current user network behavior, connects with users through smart devices, and finally disseminates content and user scenarios through scenarios such as scene insight, content selection, or community operations. Space-time to achieve the best match. From the user's point of view, the reach of the advertisement content is perfectly combined with the current user network behavior, which effectively reduces the user's rejection of advertisements [34]. At a deeper level, the delivery of advertising content is based on the labeling system of user portraits, and the depiction of user portraits is based on intelligent algorithms cleaning and filtering invalid raw data, and at the same time receiving current user network behavior data for analysis. As a result, the communication value of advertisements has been greatly improved, and the effect of triggering user resonance has been created to the greatest extent.

Disseminate content. This mode helps to advertise content creators to gain a deep insight into user needs and achieve precise matching with users from the content ontology.

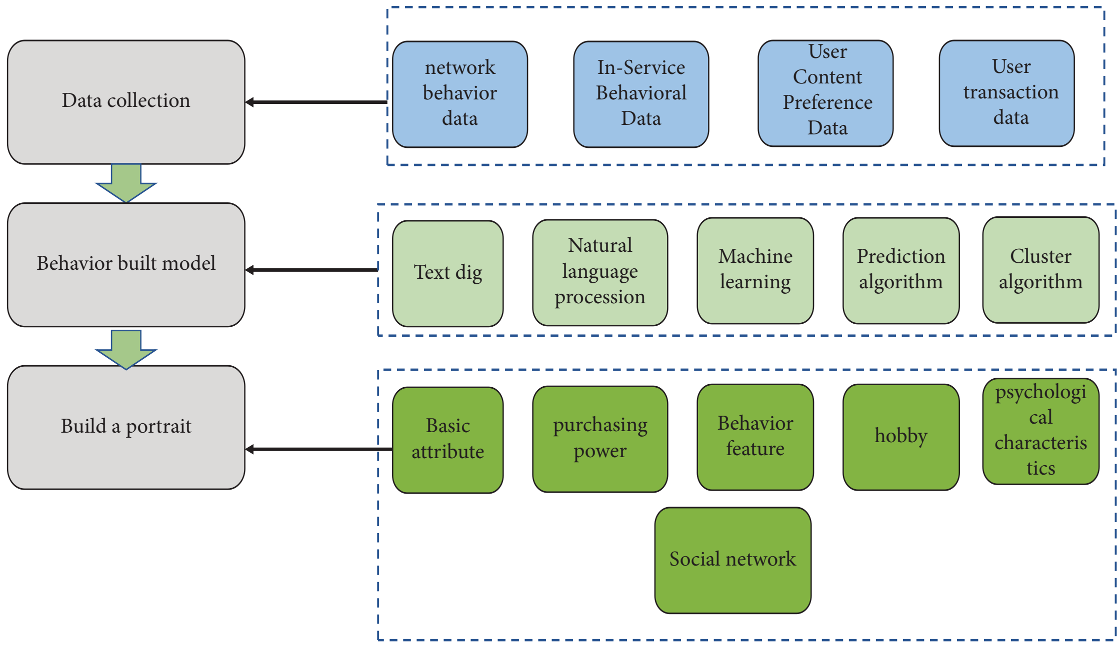


Figure 1: User portrait model.

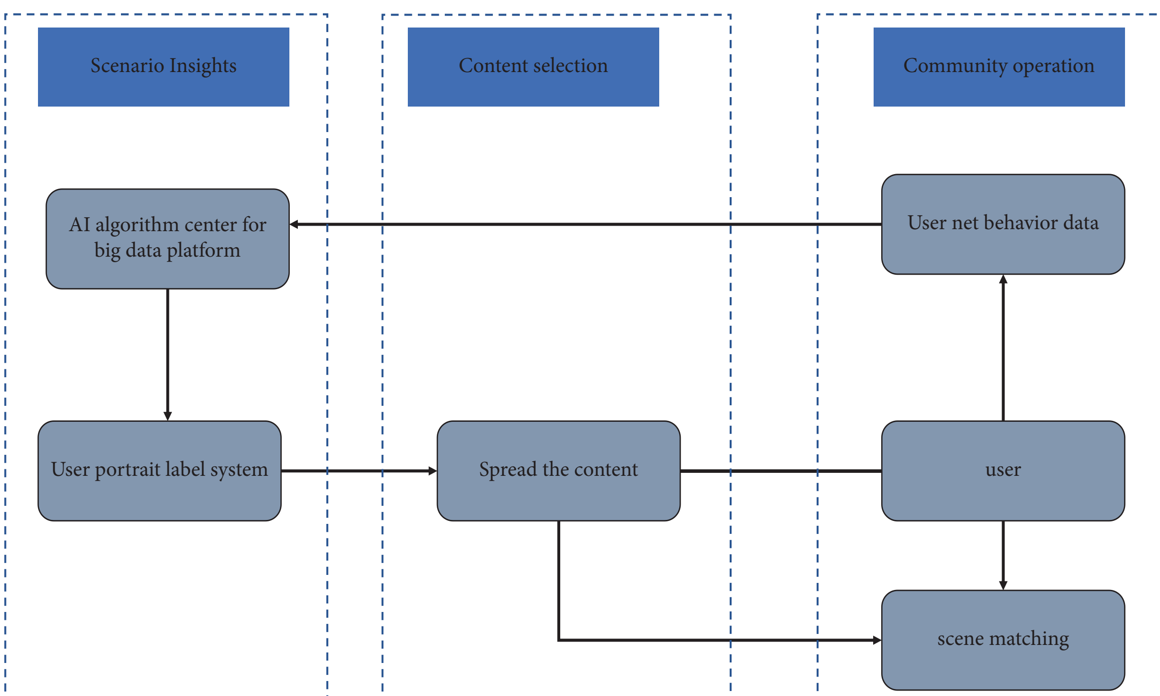


Figure 2: Propagation strategy model for computing advertising scenarios in the era of artificial intelligence.

From the perspective of scene connection media, with the help of the natural immediacy and interactivity of mobile devices, advertising content can quickly reach advertising users, break the limitations of time and space, and realize the combination of users and virtual scenes. When a user browses the current target content, it becomes a natural process for the user to accept the advertisement due to the seamless integration of the advertisement into the scene.

3. Results and Discussion

3.1. Intelligent Scene Insight Strategy. In the era of the mobile Internet, user scenario insight has become scientific and efficient. Through the user's authorization agreement on the smart terminal, the user's daily network behavior data are obtained and the user portrait is depicted. On this basis, users are analyzed according to their basic attributes, short-term behaviors, and long-term interests, and user interest tags are extracted. Take Twitter for example. According to public information, more than 15 billion articles and videos are displayed in the daily information flow of the Toutiao app, and the amount of data processed per day exceeds 8.2 PB. The Toutiao series of products generate 6 billion server requests per day. At the same time, Toutiao AI Lab continues to accumulate a large number of training samples and data based on rich and diverse application scenarios and massive users, train algorithm models, and establish a unified data warehouse to continuously train and improve artificial intelligence. When a user uses the Toutiao app, the algorithm center begins to interpret the content of the user's browsing and forms a preliminary user portrait in a few seconds. As the data generated by the reading behavior increases, the user portrait is continuously optimized. In this process, Toutiao tags users' interests and sets primary tags, secondary tags, and underlying interest tags, so as to optimize the selection of recommended content and information flow advertisements.

At present, the fusion of Twitter advertising and news has achieved an optimal balance. Toutiao adheres to the operating philosophy of "advertising is news, the news is advertising," and has made deep insights into the needs of consumers. On its official advertising website, it bluntly says, "No one likes advertising. At times, these contents happen to be advertisements." Therefore, Toutiao has launched three ways of advertising, namely, application startup fullscreen advertising, information flow large image advertisement, and information flow small image advertisement. We organize them as shown in Table 1.

Smart algorithms bring many fruitful advantages to Toutiao ads. According to official statistics from Toutiao, as of September 2020, Twitter had been installed by 450 million people, used by at least 90 million people every day, and users used it for an average of 100 minutes a day. Based on advanced machine hardware equipment and technology, Toutiao calculates user interests and intelligently recommends news based on interests within 5 seconds. At the same time, based on big data mining, multi-latitude targeted advertisements such as gender, age, interest, mobile phone, time, and region can be realized. According to the real-time data report, advertisers can check the data at any time, and click the user portrait of the advertisement, and directly formulate the advertisement delivery strategy on the advertisement delivery background. Using random sampling becomes infeasible because the total number of headline ads is not known. Therefore, this article adopts the method of convenience sampling for information flow advertisements in the "recommended" section of the Twitter App, and sorts out 18 advertisements, and records are shown in Table 2.

Before collecting and recording the advertisement push traces, the author registered the Toutiao account by binding the personal WeChat account. According to the default selection of the App system, I followed 19 tags in the personal channel, namely, Hotspot, Guangzhou, Video, New Era, Pictures, entertainment, Q&A, technology, car emperor, finance, military, sports, live broadcast, international, health, real estate, small video, NBA, and value points. Among the 18 information flow advertisements sorted out, except for one advertisement recommended by Toutiaohao and 2 random advertisements, the other 15 advertisements are related to the author's account information, browsing history, or festivals. On the one hand, the Toutiao algorithm pushes relevant information flow advertisements in fragmented scenarios according to the user's reading interest tags, reading history, and other data, and on the other hand, pushes according to the user's geographic location information. Based on the collection of user data, artificial intelligence algorithms update user portraits in real time, gain insight into users' reading scenarios, and allow advertisers to adjust advertising strategies promptly.

The core of computational advertising is to target users by keywords and terms, and people's online purchasing decisions depend on keywords and search results pages. It can be said that understanding keyword search is crucial to understanding user needs. From the perspective of scenario-based communication, keywords still have a strong guiding effect on advertising. For example, in terms of Google AdWords, keywords are the key to the success of AdWords campaigns, so PPC advertising (Pay Per Click, pay-per-click advertising) is also known as "keyword-driven" advertising. Keyword search is one of the important factors for calculating insights into advertising scenarios. Judging from industry practices such as the current Baidu series of products, Google Search, and Toutiao's recently launched "Toutiao Search" product, artificial intelligence has been applied to keyword search. In terms of search query words, it has broken through the results display of traditional algorithms, and the search algorithm has been upgraded intelligently.

The core of keyword search is to put users first, searchers convert their thoughts into words, search engines understand these words as keywords, and calculation advertisements will be displayed on the search results page. At present, keyword matching can be roughly expressed in three forms. In this paper, based on industry practice and related viewpoints, it is summarized as broad match, phrase match, and exact match. When broad matching is selected to be placed on a keyword, users searching for the keyword, or related variants of the keyword, such as singular and plural forms, synonyms, etc., may trigger the placement of the advertisement. In the phrase matching mode, the advertisement will be triggered when the user's search term contains the keywords preset by the advertiser. Precise matching requires that the user's search term and the keyword set by the advertiser are exactly the same before the ad can be triggered. The DNN model uses deep learning methods to train the model with 10 billion user click data and has more than 100 million parameters. In simple terms, artificial intelligence gives the system a large amount of training data and finds patterns from it. Finding patterns from massive data is what artificial intelligence is good at. It can quickly find possible ranking factors, adjust the weights of ranking factors, automatically iteratively calculate, and fit the calculation formula between ranking factors and user-satisfied search results.

Table 1: Formats and features of Twitter advertisement.

	Feature	Scope of use
Fullscreen ad	Huge user base, at least 10 million people see ads every day	Suitable for strengthening brand display; can be delivered by the province
Information flow big picture advertisement	Displays advertisements in news information flow and the advertisement picture is larger than the size of the news picture; attracts attention and is the first to be noticed by users	Gets high attention; precise targeted advertising; can be placed according to city districts\interests
Information flow thumbnail advertisement	A style consistent with the font size of the news picture, displaying advertisements in the news information stream; reduces the users' rejection of advertisements	Gets high attention; precise targeted advertising; can be placed according to city districts\interests

Table 2: In-feed ad records in the "Recommended" section of the Twitter app.

Serial number	Title of the ad.	Forms of the ad.	Source of the ad.
1	Guangzhou household registration quota development	Big picture of information flow	Geo-targeting
2	Medical advertisement	Big picture of information flow	Follow the label
3	Top oldies app downloads	Big picture of information flow	Random
4	Wework office space	Big picture of information flow	Age
5	Book shopping guide	Big picture of information flow	Random
6	Bazaar fair	Big picture of information flow	Time is approaching Children's day
7	Lighter ad	Big picture of information flow	Gender
8	New ad on Huawei P20	Big picture of information flow	Mobile phone brands
9	Kindergarten decoration	Big picture of information flow	Age
10	Men's shopping guide	Big picture of information flow	Gender
11	Classic arcade game app	Big picture of information flow	Age
12	Tea ads	Big picture of information flow	Papers
13	A school EMBA admissions advertisement	Small picture of information flow	Personal information
14	Lottery app	Big picture of information flow	Papers
15	Football mobile game	Big picture of information flow	Papers
16	Takeout	Small picture of information flow	Recommendation from APP
17	Mercedes Benz Guangzhou dealer	Small picture of information flow	Geographical location
18	Caivln Kioln bags	Big picture of information flow	Gender

3.2. Content Selection Strategy Based on Artificial Intelligence Technology. In the era of mobile Internet, it is easier for advertisers and brand owners to obtain big data portraits of current users or potential users than before, and to achieve accurate scene matching by using the tag body to select the most suitable advertising channels. It can be seen from the scenario communication model of computational advertising that the data platform can learn the user's behavior data before placing the advertisement and push the advertisement that is more in line with the user's preference. This article still uses Toutiao as an example. Toutiao's personalized recommendation mechanism has two basic recommendation logics in practical applications. First, when the user uses the app, the more types of actions the user takes, such as through human-computer interaction, the amount of reading content, the reading speed, the reading scene, etc., in action collection, the Toutiao algorithm center will begin to draw user portraits. When the detailed description is formed into a structured table, the Toutiao algorithm system determines that it "knows" everyone. Then, after "recognizing" each article and each keyword by the same technical means, the system algorithm center matches the user with the scene in real time. The higher the matching degree, the system determines that the information is the information that the user most wants to obtain, and then pushes it.

In addition, for users who do not have enough historical information in the database, the Toutiao system will interpret similar groups and push the information that users are most likely to be interested in. Therefore, everyone's headlines are different, that is, "what you care about is the headlines." According to the above two logics, the Toutiao system splits the matching and connection objects between articles and users in more detail. Its recommendation engine first divides articles into specific feature vectors according to keywords and categories and then locates users, and assigns them to a specific feature vector. Finally, the two are matched, and different information is pushed to each user according to the algorithm learned by the recommendation engine. Computational advertising identifies interest tags based on user portraits, uses Toutiao's personalized recommendation engine to match the most relevant creative materials, recommends multiple creatives at the same time, and completes targeted advertising.

The basis of Toutiao's personalized recommendation of news feed advertisements is the intelligent recommendation mechanism of its content. In order to achieve integration with information and reduce the psychological rejection of users, advertisements need to be consistent with the presentation of information. At present, based on the literature and product experience of the Toutiao App, we summarize Toutiao's article personalized recommendation mechanism as follows: recommendation of articles with similar themes, recommendation of news in the same city, a recommendation based on article keyword characteristics, recommendation of popular articles, and long-term selection of interests by users. Mechanisms such as tag recommendation and site source recommendation. In addition, recommendations may also be made based on basic user information, such as user interests, age, occupation, reading habits, and geographic location. The information flow advertisement is based on the tag system of the article and is allocated to the corresponding content area for display.

Scenario-based communication in the Internet era has the attributes of immediacy and dynamics, and artificial intelligence advertising creation is applied to the field of computational advertising along with the development of scenario-based thinking. Although artificial intelligence cannot currently perform advanced advertising creative work like professional advertising creators, in some special scenarios, artificial intelligence will come in handy promptly when a large number of advertising works cannot be provided by manpower alone. For example, since 2016, the intelligent design platform "Luban" developed by the Alibaba Cloud team has designed advertising posters with billions of dollars for the "Double Eleven" shopping festival and placed them in the specific scenes of users' shopping on that day. At present, artificial intelligence advertising creation mainly focuses on advertising design and copywriting design. In terms of advertising design, taking the "Luban" intelligent design platform as an example, the system can complete the work volume of 8,000 banner images per second through artificial intelligence algorithms and a large amount of data training, machine learning design, and output design capabilities. Customers provide four functions: one-click generation, intelligent typesetting, design expansion, and intelligent creation. In terms of AI copywriting design, the artificial intelligence advertising creation system has developed a "neuron" function similar to human memory. The system automatically "stores the memory" for the copy finally selected by the user during the use process. The copy selected by the user will be ranked at the top of the machine in the next similar search; the copy not selected by the user will be ranked in the next similar search by the machine. The latter position is no longer recommended; the system will also "remember" the annotation modification made by the user to improve the quality of the copy generated next time. In other words, the AI copywriting system can realize the automatic optimization algorithm of the machine according to the user behavior correction algorithm model.

3.3. Scenario Communication Strategy Based on Community Operation. Community connection is based on the user's social relationship chain. People with common interests and hobbies gather together spontaneously or under the subtle guidance of the media content platform to form a fixed community. Based on common values and aesthetic standards, it is easy to establish emotional relationships between users, thereby enhancing the user's community stickiness. With the development of the mobile Internet, mobile apps have become the entrance of scene dissemination, and various sub-scenarios such as beauty, fitness, taxi, travel, photography, etc., have rapidly formed a community form, providing users with a personalized scene service experience. Under the guidance of Internet thinking, the user-centered product concept is gradually recognized. When users join the community, they will perceive the community through behaviors such as observation, experience, and participation. The first step in retaining users is whether the community can bring the value expected by users. Secondly, if the community creates a contagious community social culture on the basis of the value it brings to users, and the community brings a sense of warmth to users, and it is possible to gain user loyalty and make it the spokesperson of the community.

Strong scene connections can bring better communication and conversion effects to computational advertising. In the e-commerce industry, content e-commerce has become a breakthrough in a new round of competition. Internet e-commerce companies such as Taobao, Netease Koala, and Xiaohongshu are already building a "scenario + community" e-commerce model. The headline product "Douyin" has opened the "commodity window" function on the video creator's personal homepage to try to get into the competition of e-commerce giants, and take the initiative in another way. It has become a trend in the industry that major platforms provide users with scene experience and emotional appeal through community connections, so as to use intelligent algorithms to seamlessly recommend advertisements or product information to users' usage scenarios.

With the change of users' online behavior habits, the focus of computing advertising has gradually shifted from the PC side to the mobile side, showing the characteristics of scene-based and community-based communication. Cai Qi believes that the changes in reading in online community communication are mainly manifested in four aspects, including the production revolution under cognitive surplus, the hybrid text in aggregation and differentiation, the shared reading driven by content and relationship, and the multi-terminal. Internet products need to deal with fierce market competition, compete with users, and begin to focus on network community operations. The core of the community operation strategy is to build an online community culture, provide users with high-quality content services, and meet the needs of users for individual value creation. Creating a community culture and establishing community-specific cultural symbols and community rituals are the basis for maintaining user participation and activity. High-quality content services include setting high-quality interactive themes for community members, providing differentiated content reading, and improving user retention time. In the interaction of topics of interest, users complete the expression and presentation of self-values, obtain value recognition from other users with the same interests, and find a spiritual sense of belonging in the virtual community.

The former artificial intelligence technology is still in the initial stage of application, and at the same time, limited by domestic economic and technological levels and other development factors, computational advertising also shows many problems and limitations. For example, at the data management and application level, the data on the data platforms of major companies is uneven, and there is a lack of effective communication between platforms due to factors such as competition and games. Transparency between platform data and advertisers and agencies still needs to be improved. At the technical level, the higher the level of artificial intelligence, the stricter the requirements for the complexity of the algorithm model.

The “Island” phenomenon of computational advertising monitoring data. With the advent of the information society, people realize that the value of data is becoming more and more important, and data are increasingly becoming an important part of the core assets and competitiveness of enterprises. At the same time, the multi-interest game in the data market has caused the phenomenon of data segmentation, which has caused the phenomenon of data silos to become more and more obvious. The interesting structure of all parties in the data transaction is unbalanced. In the data industry, the interesting relationship among users, owners, and users is unbalanced, especially the unequal rights and agreements between users and platforms, which has led to the aggravation of the phenomenon of isolated islands. Much of China’s data are collected in the databases of industry giants, telecommunications operators, and government agencies. On the other hand, without obtaining any data benefits, data producers also risk data misuse and damage to personal privacy and data security. In the long run, neither party in the data industry ecology will be the ultimate winner. Data silos aggravate the closure of data and the imbalance of all parties in the data chain. The collapse of the system will only damage the interests of all participants. The practice of data sharing lacks momentum. Data sharing has great practical difficulties, which is an important reason for the intensification of the phenomenon of data silos. China currently lacks laws and regulations that strictly regulate data sharing and openness. Relevant personnel are worried that the openness of government data sharing will cause information security problems, data leakage and loss of control, fear of data sharing and openness, and they dare not share their own data resources with others. Share openly. At the same time, the big data industry also has a low degree of resource openness and sharing, and the value of data is difficult to be effectively exploited and utilized.

The lack of interoperability and sharing of data is caused by various reasons. For example, in the concept of many enterprises or institutions, data resources are treated in the same way as other resources, and they believe that possession and private ownership are corporate wealth, resulting in the phenomenon of “data islands.” To break this phenomenon, this article provides several countermeasures. First of all, major companies or institutions need to establish and strengthen cooperative relationships and open up data sharing permissions. Here, agreements can be made through a third-party data platform, or data can be shared directly between enterprises in a mutually beneficial way, so that the most accurate data can be obtained for the calculation of advertising scenarios and accurate access can be achieved. Secondly, all parties involved in sharing data should establish a standardized data construction system and integrate big data resources. At the same time, the scope of data sharing can be extended through cross-business cooperation. Finally, establish an information protection system and improve the legal system. In the era of artificial intelligence, government agencies and enterprises work together to build a credit-sharing environment for big data.

4. Conclusions

This paper focuses on the three major strategies of computing advertising scene communication, and deeply analyzes the principles and applications of scene insight strategy, content selection strategy, and community operation strategy. Meanwhile, in view of the current problems in the communication of computational advertising scenarios, we analyzed the three aspects of the phenomenon of “data island,” traffic cheating and traffic hijacking, and brand safety in the propagation of computational advertising scenarios, and proposed some feasible improvements. This paper also discussed the intelligent upgrade of computing advertising, how to realize the scene-based communication of computing advertising and related communication strategies. The findings of this study can be roughly divided into the following points.

(1) The development of computational advertising enables smart upgrades. Under the background that the era of artificial intelligence has arrived, computing advertising uses artificial intelligence algorithms to complete breakthrough upgrades. Through the combination of algorithms and data, its intelligent upgrades are mainly manifested in three aspects. First, to achieve higher matching accuracy communication, the second is efficient customized communication, and the third is to realize the contextual interaction between advertisements and users. The key to the upgrade of the accurate dissemination of computing advertising lies in the combination of artificial intelligence algorithms and big data, intensive processing through data mining and analysis, and advertising based on the intelligent

user portrait tag system. Through intelligent user portraits, computing advertisements can achieve the intelligent push of "thousands of people and thousands of faces." The application of artificial intelligence technology enables the automatic push of computing advertisements to make intelligent decisions based on data analysis of users' real-time usage scenarios and to achieve contextual interaction between advertisements and users.

(2) The intelligent scene dissemination of computing advertising requires multiple conditions to be satisfied at the same time. The "intelligent" performance characteristics of intelligent scene dissemination of computing advertisements are mainly in the application of scene intelligence technology, the intelligence of the data platform, and the intelligent construction of user portraits. This paper summarizes and analyzes the role of three basic technologies on scene dissemination, including scene connection technology, data processing technology, and scene experience technology. With the advent of the era of mobile Internet and the gradual deepening of domestic research on scene theory, the connection between scenes, the connection between scenes and people, and the connection between people and people have become the focus of research on computational advertising scene dissemination. The realization of scene dissemination puts more stringent requirements on data processing technology. The key point is the degree to which artificial intelligence algorithms can process big data. Finally, the user's scene experience is realized through various types of intelligent terminal devices. In the entire scene dissemination of computing advertisements, the data management service provided by the data platform and the intelligent labeling system of user portraits plays a key role in the dissemination of the scene.

(3) Intelligent computing advertising scenario communication strategy. This paper believes that in the era of artificial intelligence, the communication strategy of the computing advertising scene mainly includes three aspects: intelligent scene insight strategy, content selection strategy, and community operation strategy. First, the intelligent scene insight strategy is mainly analyzed from two levels. On the one hand, it is a scene mining based on intelligent data and an in-depth analysis of the user tag system centered on intelligent algorithms. Secondly, in terms of content selection strategy, the application of artificial intelligence technology has brought a new upgrade to the creation and recommendation mechanism of computational advertisements. Smart user portraits are updated with real-time data to push advertisements that are more in line with user preferences. In addition, artificial intelligence content creation has been realized in poster design and advertising copywriting, which can meet the needs of large-scale advertising in specific scenarios, and provide new options for the content selection strategy of computing advertising scenarios. Finally, this paper argues that community is an important strategy for establishing strong connections in the scene. The strong connection of the scene can give a better communication effect to the calculation of the advertising code. The establishment of a community needs to establish an emotional relationship with each other based on the user's social relationship chain. At the same time, it is necessary to create a high-quality community culture based on intelligent data, integrate users with scenes, and effectively improve the reach of advertisements.

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] R. M. Quintana and S. R. Haley, "The persona party: using personas to design for learning at scale," *Chi Conference Extended*, vol. 1, no. 5, pp. 933–941, 2017.

[2] J.-J. Chang, J. H. Chen, and M. F. Tsai, "Corporate social responsibility, social optimum, and the environment-growth tradeoff," *Resource and Energy Economics*, vol. 69, Article ID 101311, 2022.

[3] W. Huang, "Research on advertising marketing strategies of Bilibili under the New media era," *International Journal of Educational Technology*, vol. 3, no. 3, 2022.

[4] J. Wu and B. Hu, "An agent-based simulation study of the dynamics of mobile viral advertising," in *Proceedings of the 2008 Winter Simulation Conference*, p. 2953, Miami, FL, USA, December, 2008.

[5] X. An, X. Li, L. Xiao, and Z. Lai, "A study on the content marketing strategy of pan-entertainment live broadcast platform," *International Journal of Educational Management*, vol. 7, no. 3, 2022.

[6] P. Sharma, J. S. Nirma, B. Jack, and R. Angeline, "Promoting SMEs in pacific island countries through effective marketing strategies: a systematic literature review and a future research agenda," *Journal of Electronic Commerce in Organizations*, vol. 20, no. 2, 2022.

[7] M. Zhang and G. Shi, "Consumers' impulsive buying behavior in online shopping based on the influence of social presence," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6794729, 2022.

[8] M. Su Han, H. Daniel Peter, Y. Wang, and H. Wang, "Consumer confidence and green purchase intention: an application of the stimulus-organism-response model," *Journal of Retailing and Consumer Services*, vol. 68, 2022.

[9] F. Shakouhi, R. Tavakkoli-Moghaddam, A. Baboli, and A. Bozorgi-Amiri, "Multi-objective programming and Six Sigma approaches for a competitive pharmaceutical supply chain with the value chain and product lifecycle," *Environmental Science and Pollution Research International*, vol. 21, 2022.

[10] J. J. Ferreira, P. M. Veiga, C. I. Fernandes, and S. Kraus, "B2B marketing strategies in healthcare management: intellectual structure and research trends," *Journal of Business & Industrial Marketing*, vol. 37, no. 8, pp. 1580–1593, 2022.
[11] M. Chen, F. Su, and F. Tai, "Major league baseball marketing strategies and industry promotion approaches," *Frontiers in Psychology*, vol. 13, 2022.
[12] C. Natalie, G. Frank, and W. Palmatier Robert, "A call for research at the public policy–marketing strategy interface," *Journal of Public Policy and Marketing*, vol. 41, no. 3, 2022.
[13] R. Ruggless, *Zaxby's Names Patrick Schwing to Head Marketing, strategy*, Nation's Restaurant News, Howick Place, London, 2022.
[14] H. Lv, "Smart product marketing strategy in a cloud service wireless network based on SWOT analysis," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 7539860, 11 pages, 2022.
[15] D. Richardson, "Attitude, integrity, relationships: rob falke's other definition of AIR," *Air Conditioning, Heating and Refrigeration News*, vol. 276, no. 3, 2022.
[16] Q. Qiu, M. Huang, and Z. Jing, "Research on the marketing strategy of agricultural products under the background of the epidemic-- taking xiaogang village as an example," *World Scientific Research Journal*, vol. 8, no. 6, 2022.
[17] P. K. Kannan, "Introduction: global marketing strategy," *International Journal of Research in Marketing*, vol. 39, no. 2, pp. 443-444, 2022.
[18] K. Chikhi and M. C. E. B. Djelloul, "The olive oil market in the mediterranean: what are marketing strategies for Algeria?" *Contemporary Agriculture*, vol. 71, no. 1-2, pp. 28–37, 2022.
[19] Z. Liu, "On the influence of hunger marketing strategy of luxury group in global market on company development," *Financial Engineering and Risk Management*, vol. 5, no. 2, 2022.
[20] G. Jonathan and J. Fast Nathanael, "The power to harm: AI assistants pave the way to unethical behavior," *Current Opinion in Psychology*, vol. 47, 2022.
[21] F. Lynn Thomas, M. Ottino Julio, M. Lueptow Richard, and B. Umbanhowar Paul, "Potentialities and limitations of machine learning to solve cut-and-shuffle mixing problems: a case study," *Chemical Engineering Science*, vol. 260, 2022.
[22] E. Rami, M. Tina, I. Ahmad, and A. Valdez Tulio, "Artificial intelligence and tele-otoscopy: a window into the future of pediatric otology," *International Journal of Pediatric Otorhinolaryngology*, vol. 160, 2022.
[23] P. J. Nerea and N. Valdecantos Vicente, "Review of the application of artificial neural networks in ocean engineering," *Ocean Engineering*, vol. 259, 2022.
[24] S. Bayomie Omar, F. L. de Cerqueira Rafael, N. Laura et al., "Detecting flooding state in extraction columns: convolutional neural networks vs. a white-box approach for image-based soft sensor development," *Computers & Chemical Engineering*, vol. 164, 2022.
[25] W. Cho Geoffrey, K. Ghanem Ahmed, G. Quesada Carlos et al., "Quantitative plaque analysis with A.I.-augmented CCTA in end-stage renal disease and complex CAD," *Clinical Imaging*, vol. 89, 2022.
[26] S. Coakley, R. Young, N. Moore et al., "Radiographers' knowledge, attitudes and expectations of artificial intelligence in medical imaging," *Radiography*, vol. 28, no. 4, pp. 943–948, 2022.
[27] T. Christoph, B. . Ann-Christin, M. . Felix, M. Yu, K. Arno, and S. Carsten, "Multi-modal framework to model wet milling through numerical simulations and artificial intelligence (part 2)," *Chemical Engineering Journal*, vol. 450, no. P1, 2022.
[28] S. Devagiri Jeevan, P. Sidike, N. Quamar, X. Yang, and S. Smith, "Augmented reality and artificial intelligence in industry: trends, tools, and future challenges," *Expert Systems with Applications*, vol. 207, 2022.
[29] M. R. M. Veeramanickam and P. Ramesh, "Analysis on quality of learning in e-Learning platforms," *Advances in Engineering Software*, vol. 172, 2022.
[30] M. Cao, Y. Hu, Ru Shi, and Z. Zhang, "REsearch on marketing strategy optimization of healthy food based on analysis of consumer behavior characteristics——take mint health APP for example," in *Proceedings of the 2021 2nd International Conference on World Economy and Project Management*, Barcelona Spain, January, 2022.
[31] L. Dong, H. Liu, and T. Yi, "Research on brand remodeling and marketing strategy optimization of Chinese domestic cosmetics in post-pandemic era," in *Proceedings of the 2021 3rd International Conference on Economic Management and Cultural Industry (ICEMCI 2021)*, Guangzhou, China, October, 2021.
[32] K. Xiao and X. Hu, "Study on maritime logistics warehousing center model and precision marketing strategy optimization based on fuzzy method and neural network model," *Polish Maritime Research*, vol. 24, no. s2, pp. 30–38, 2017.
[33] W. Liu, Y. Yang, and D. Xu, "Research on optimization of cross-border E-commerce marketing strategy of agricultural products under the background of environmental protection," *IOP Conference Series: Materials Science and Engineering*, vol. 612, no. 5, Article ID 052031, 2019.
[34] Y. Liu and W. Chen, "Optimization of brand marketing strategy of intelligent technology under the background of artificial intelligence," *Mobile Information Systems*, vol. 12, p. 2021, 2021.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 9266844, 13 pages
https://doi.org/10.1155/2022/9266844

Research Article

The Intelligent Selection Method of Distribution Sites Driven by the Intelligent Optimization Algorithm

Juanli Jin

School of Business, Wuhan Vocational College of Software and Engineering, Wuhan 430205, China

Correspondence should be addressed to Juanli Jin; 100219@yzpc.edu.cn

Received 31 August 2022; Revised 19 September 2022; Accepted 26 September 2022; Published 7 October 2022

Academic Editor: R. Mo

For the internal enterprise, the intelligent selection of logistics distribution sites can optimize the distribution route, which is conducive to reducing the total distribution cost and improving the enterprise revenue. This paper takes urban logistics distribution point selection optimization as the research content and compares and analyses the applicability of different intelligent algorithms to logistics data. The AFSA is selected as the optimization algorithm, and knowledge learning is introduced to optimize the algorithm. The optimized AFSA model is applied to the mathematical model of distribution point selection. An intelligent algorithm-driven logistics distribution point selection model is established, and the optimized AFSA is used to solve the problem. Based on the actual case of CSC Logistics Company, the route of the distribution point in a region of CSC Logistics Company is optimized and the model is validated and solved. The results show that GWO and AFSA search capabilities are significantly better than of other intelligent algorithms, but there is some instability in the GWO algorithm. The AFSA is most suitable for solving logistics-related problems. The optimized AFSA considering knowledge learning has high efficiency and good optimization results. CSC Company uses this intelligent algorithm to select distribution sites intelligently, which shortens the total logistics path and improves the distribution efficiency. The total mileage of the initial route is 115 km. After intelligent algorithm optimization, the total mileage of the distribution changes to 83 km, which reduces by 32 km and 28% compared with the original route. The whole distribution process saved about 1.5 hours, which fully optimized the efficiency of distribution.

1. Introduction

As the third profit source of enterprises, logistics plays an important role in the development of enterprises [1]. In June 2014, the State Council of China issued the medium- and long-term plan for the development of the logistics industry, which emphasized the importance of the logistics industry and its impact on people's livelihood. The logistics industry shows its strong vitality under the strong promotion of national policies. According to the data analysis of the official website of the National Bureau of Statistics, with the development of e-commerce, online sales have shown an "explosive" growth trend [2]. In the first half of 2018, the national online retail volume was about 4081 billion, an increase of more than 30% year-on-year. Among the 4081 billion retail sales in the whole network, the retail sales of physical goods was 3127.7 billion, an increase of more than 30%, accounting for about 17.4% of the total retail sales of consumer goods in the whole society [3]. Driven by expanding domestic demand and increasing consumer demand, the e-commerce logistics express market is also developing at a high speed. The popularity of e-commerce platforms such as Taobao, JD.com, and Pinduoduo has not only changed people's purchasing methods but also made the express logistics industry rise rapidly [4]. In 2010–2020 alone, nearly 3000 express logistics enterprises were added in China. According to the statistics of the State Post Office, in the past five years, the National Express service has developed rapidly. The scale of the business increased from 5.7 billion pieces of express delivery in 2010 to 10.3 billion pieces of express delivery in 2021. The growth rate of the express delivery business from 2013 to 2018 was about 40% per year, as shown in Figure 1. According to the statistical analysis of the State Post Office, in 2018, China's express business

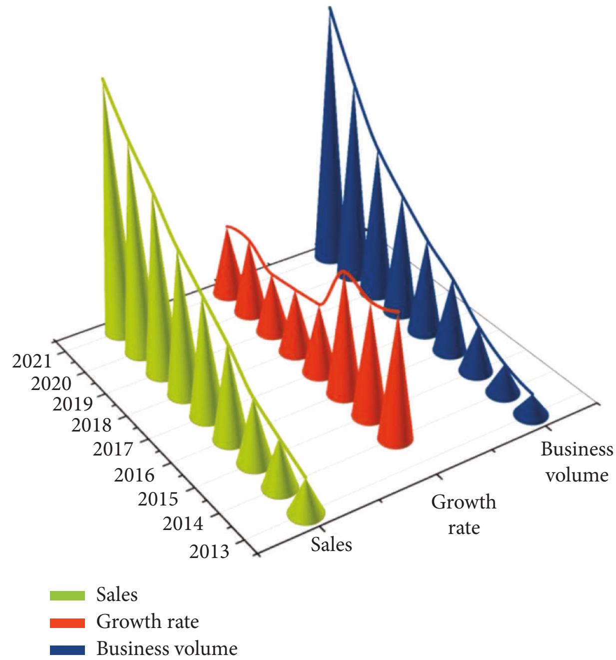


FIGURE 1: The express delivery business data.

volume exceeded 50 billion pieces for the first time [5]. China's express business volume continued to rank first in the world, accounting for nearly 50% of the global express business volume. The business volume has increased 19 times in the past 10 years, and it has ranked first in the world for 8 consecutive years, making China the first in the express industry among countries. At the same time, the logistics industry in other countries is also developing rapidly. For example, famous third-party logistics service giants have established integrated logistics service networks in various parts of India to provide seamless transfer services for outsourcing operators, importers, exporters and global trade operators in the Indian market. It can be seen that the logistics industry has a broad market prospect, which has a great impact on the development of enterprises and countries [6].

However, with the rapid development of the logistics express industry, many enterprises are also facing many problems [7]. The efficient and intelligent material distribution system and the scheme is a strong support for the development of the e-commerce industry. Modern logistics technology integrates Internet of things technology, path optimization, and decision-making control to improve the level of modern logistics [8]. Therefore, efficient material distribution optimization has become the breakthrough point of modern logistics transportation system reform. Logistics distribution mainly completes the delivery of goods from the distribution center to the customer's distribution point. At present, researchers usually regard the logistics distribution scheme as a mathematical model to solve the traveling salesman problem. However, the logistics distribution is affected by the commodity value, the types of commodities required by each distribution point, and the volume of distribution vehicles occupied by the commodities, so the distribution traveler model is more complex. Among them, the problem of "difficult selection of logistics distribution sites" is particularly prominent [9]. Logistics distribution refers to the logistics activities within a certain area, according to the customer's goods demand, through several logistics nodes, and through the operational processes such as loading and unloading, sorting, packaging, splitting, and assembly, finally, the designated goods are delivered to the customer at the designated place in a reasonable way [10]. Logistics distribution is a kind of operation mode in the modern circulation industry which integrates commodity flow, information flow, and transportation flow. The time of logistics distribution is highly variable, resulting in frequent "secondary delivery," which reduces the efficiency of distribution. However, if more distribution vehicles are added, it will not only increase the cost of logistics transportation but also increase the burden of urban traffic, making the traffic situation more crowded and complex. Therefore, the logistics industry is also constantly optimizing its own distribution mode to achieve cost and quality control while ensuring distribution efficiency [11].

After entering the 1990s, with the rapid development of electronic, communication, computer, and other technologies, researchers began to use computers as auxiliary tools to study the logistics scheduling problem from the perspective of artificial intelligence [12]. The intelligent optimization algorithm is a kind of meta-heuristic random search algorithm derived from biology. Because its search process is stochastic, can easily jump out of local search, and has the adaptive ability to the environment, some specific types of optimization algorithms have been initially applied in logistics distribution problems and have achieved initial results [13], for instance, genetic algorithm, simulated annealing algorithm, particle swarm optimization, hybrid optimization algorithm, immune optimization algorithm, ant colony optimization algorithm, and artificial fish swarm optimization algorithm. These intelligent optimization algorithms find the optimal solution in the search space through continuous iteration and overcome the dependence of quasi-Newton algorithm on the initial value of calculation to a certain extent. At the same time, they also solved the problem that the conventional algorithm is difficult to find the global optimum and ensured the convergence of the algorithm [14]. Most importantly, with the continuous improvement of the logistics information management system, various business data, traffic data, and express information can be collected in real time from the database platform of logistics enterprises [15]. These data record the timeliness, transportation mode, and distribution requirements of different types of express delivery and are comprehensive and detailed descriptions of the entire logistics distribution process. Through the statistical machine learning algorithm represented by the artificial neural network, the model is directly established based on a large amount of data. By fitting the distribution relationship of variables and targets, the impact mechanism of logistics transportation efficiency can be analyzed from multiple angles and all directions, avoiding the simplification of a large number of factors in the conventional modeling [16].

Therefore, the intelligent optimization algorithm used to solve complex optimization problems can often get better results and has potential application value for logistics distribution problems [17]. Based on this, the concept of intelligent distribution point selection came into being. This method is a modern comprehensive logistics distribution mode that can be perceived, analyzed, processed, and adjusted in time in the selection of logistics distribution sites [18].

The purpose of this paper is to study the application of intelligent optimization algorithms in the selection of logistics distribution sites. On the one hand, the applicability of different intelligent optimization algorithms in the optimization of logistics distribution is studied and analyzed, and an optimized algorithm suitable for large numbers of logistics is proposed, in order to further improve the efficiency and accuracy of the algorithm and provide an efficient, stable, and good performance solution to the path optimization problem. On the other hand, based on the proposed algorithm, a mathematical model driven by intelligent optimization algorithm is established, which provides data support for the decision-making scheme to improve the efficiency of logistics distribution. Therefore, this paper comprehensively uses logistics data and the intelligent optimization algorithm technology to explore a new development mode of express logistics and proposes an intelligent selection method of logistics distribution sites driven by the intelligent optimization algorithm.

2. Selection and Optimization of the Intelligent Optimization Algorithm

2.1. Comparison of Intelligent Algorithms. Although the traditional optimization algorithm has the advantages of fast calculation speed, less parameters, fast speed, simple model, and good robustness, it is easy to converge prematurely, the accuracy of seeking parameters is not high, and later evolution is easy to fall into local solution. The intelligent optimization algorithm has strong global search performance, can accurately find feasible solutions, has a wide range of parameter settings, and does not require high initial values of parameters.

In order to compare and analyze the advantages and disadvantages of different intelligent optimization algorithms, this paper compares quantum particle swarm optimization (QPSO), moth flame optimization (MFO), sine cosine optimization (SCA), simulated annealing (SA), grey wolf optimization (GWO), and the performance of the artificial fish swarm algorithm (AFSA) [19].

In this paper, three types of common business functions are selected from several logistics data for simulation experiments to evaluate the performance of the algorithm. Here, A is the cost test function, B is the traffic test function, and C is the distance calculation test function. In order to reduce statistical errors and produce convincing results, 30 independent experiments were performed for each algorithm. The operating system of the experimental environment is Windows 10, and the processor is Intel®Xeon®Gold 6154 CPU@3.00 GHz with 32 GB of memory. The integrated development environment is matlabr2017b, the population is set to 50, and the maximum number of iterations is 1000. The population was set to 50, and the maximum number of iterations was 1000.

For each test function, the performance of the algorithm is evaluated by comparing the convergence curves of the execution time to find the optimal value.

The solution results of six intelligent optimization algorithms on each test function are summarized in Table 1. The parameter AE in the table represents the average error between all the solved optimal values and the actual optimal values, ESTD represents the standard deviation of the error, mine represents the minimum error, and MaxE represents the maximum error.

For the logistics cost test function, GWO and AFSA show excellent solution accuracy. The average value of the solution is very close to the global optimal value. The standard deviation in 30 experiments is small, and the effect is the best. The QPSO effect is good, the average value and standard deviation are also within the acceptable range, the SCA and GWA effects are general, the average value of the solution obtained by SA is far from the real global optimal solution, and the optimal value obtained by multiple experiments is far less than other algorithms, with the worst effect.

For the traffic test function, GWO, AFSA, and MFO all find the global optimal value. The QPSO effect is general, and the SCA and SA effects are the worst. The test results are basically consistent with the results of the cost test function. It should be noted that there is a great difference between the average value of the solution obtained by GWO in the distance calculation test function and the real global optimal solution, and the test result is significantly different from that of AFSA.

On the whole, GWO and AFSA have better search ability than other algorithms, and AFSA has better effect than GWO. The results of QPSO and MFO are in the middle level, and the results of SCA and SA are poor.

To verify the stability and convergence of the six algorithms, we compare the convergence of each algorithm based on different functions, as shown in Figure 2. The GWO algorithm and the AFSA have premature convergence in the test function, but the accuracy of solution is better than of other algorithms. The two algorithms converge when the number of iterations is less than 400, and the solution accuracy is much better than other algorithms, which shows that the two algorithms can maintain a good balance between global search and local development. The accuracy of other algorithms improves with the increase of iteration times.

According to the horizontal analysis in Figure 2, the AFSA requires the least number of iterations when the same solution accuracy is reached. Although the GWO algorithm can also obtain the optimal value when performing global search, the convergence speed is slow. Although the search population is converging toward the solution center, many iterations are wasted on random target finding. Once the population diversity is reduced, there is no corresponding strategy to help jump out of the local optimal solution. Even

Table 1: The solution results of intelligent optimization algorithms.

Types	Statistical values	QPSO	MFO	SCA	SA	GWO	AFSA
A	AE	1.1	2.3	4.5	7.2	1.2	0.1
	EStd	3.4	5.3	2.9	3.1	0.8	0.2
	MinE	9.1	3.1	3.4	3.7	0	0
	MaxE	12.4	18.3	76.1	100.2	4.5	1.2
B	AE	2.2	3.3	2.5	8.2	3.4	0.9
	EStd	4.4	6.3	3.9	6.1	1.1	0.9
	MinE	6.7	0	12.4	13.1	0	0
	MaxE	9.4	4.3	66.1	88.2	5.5	0.3
C	AE	3.3	4.3	3.5	5.2	7.3	0.1
	EStd	7.4	7.3	3.9	5.8	3.2	0.2
	MinE	7.7	8.8	11.4	10.1	5.4	0
	MaxE	29.4	54.3	106.1	108.2	33.5	0.8

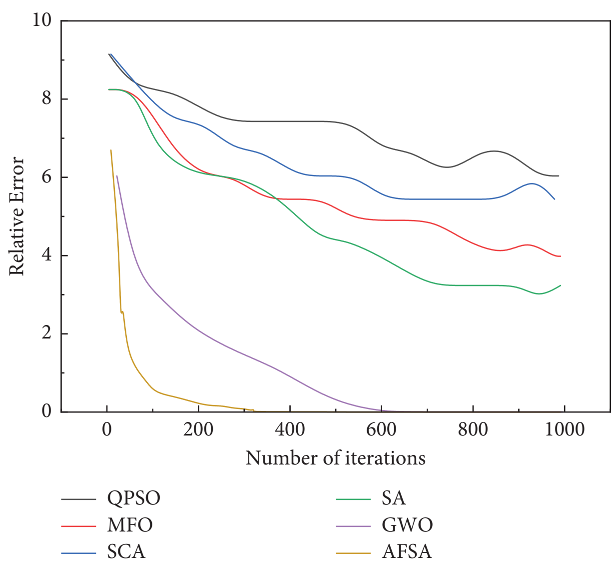


Figure 2: The convergence of each algorithm based on different functions.

if the maximum number of iterations is increased, the effect is not optimized, which indicates that the GWO algorithm has certain instability. Among all the algorithms, the SA algorithm has the worst convergence effect. In the process of solving, it falls into the local optimum many times earlier, resulting in huge errors and poor stability. Although the principle of the QPSO algorithm is simple, its global search ability is strong but its search accuracy is not enough, and the search range is concentrated in several regions. In conclusion, the artificial fish swarm algorithm is selected as the analysis algorithm to establish the model.

2.2. Artificial Fish Swarm Algorithm and Optimization. The artificial fish swarm algorithm is a swarm intelligence optimization algorithm that simulates the behavior of fish swarm [20]. The algorithm simulates four behaviors of fish, namely, random behavior, foraging behavior, swarm behavior, and tail chasing behavior. Each individual position in the fish school is to be solved. By comparing the food concentration in the field of vision of the four behavior modes, the objective function value is placed in the bulletin board to achieve the optimal solution of the problem. The calculation flow is shown in Figure 3.

The artificial fish swarm algorithm only refers to the information of the historical optimal position of the population in the process of optimization and lacks the interaction of information and knowledge between individuals and groups. Therefore, the algorithm can be optimized from the perspective of knowledge learning, making full use of the group knowledge information, and using the similarity clustering function to divide the fish population into different regions, so that the fish population can carry out targeted regional knowledge learning and dimension knowledge learning. The process of individual knowledge learning can also be regarded as a decision-making process. Decision-making is a high-level cognitive process, which usually refers to the decision-makers using scientific methods to mine the hidden useful knowledge of the given knowledge and integrate it to form an implementation plan in order to achieve a certain goal. The group is the collection of cognitive resources of each decision-making individual. The purpose of individual interaction is to make useful knowledge by the individual or the group, so as to fully and effectively transfer, share, and utilize the knowledge.

The basic equation of informatics is as follows:

$$Q[s + \Delta s] = Q[s] + \Delta s. \tag{1}$$

This formula shows that with the increase of learning time, the knowledge potential energy $Q[S]$ possessed by decision-makers also increases and the increased knowledge level Δs is related to their original knowledge and external knowledge. For individual fish, the knowledge potential energy can be expressed by the fitness value of its position.

Let t_{jx} be the position of individual j at time t; then, its corresponding knowledge level is F_{tjx}. That is, when an individual's fitness value is better, its knowledge potential energy is also larger; otherwise, the knowledge potential energy is smaller. The knowledge learning rate is given by the following formula:

$$\xi_i^t = \frac{e^{\mathrm{Score}_i^t}}{e^{\mathrm{Score}_1^t} + e^{\mathrm{Score}_2^t} + \cdots + e^{\mathrm{Score}_m^t}},$$
$$\mathrm{Score}_j^t = \begin{cases} 1, \; if\,(f_{\mathrm{worst}} = f_{\mathrm{best}}), \\ \dfrac{f_{\mathrm{worst}}^t - f_j^t}{f_{\mathrm{worst}}^t - f_{\mathrm{best}}^t}, \text{others}, \end{cases} \tag{2}$$

where f_{j}^t is the fitness value of individual J, f_{best}^t and $f_{\mathrm{worst}}^{\mathrm{t}}$ are the best and worst fitness values of the population at time t, and m is the number of fish stocks. It can be seen that the better the fitness value of an individual fish, the greater its knowledge potential energy score; on the contrary, the worse the fitness value of an individual fish, the smaller its knowledge potential energy score.

The positions determined by historical knowledge are as follows:

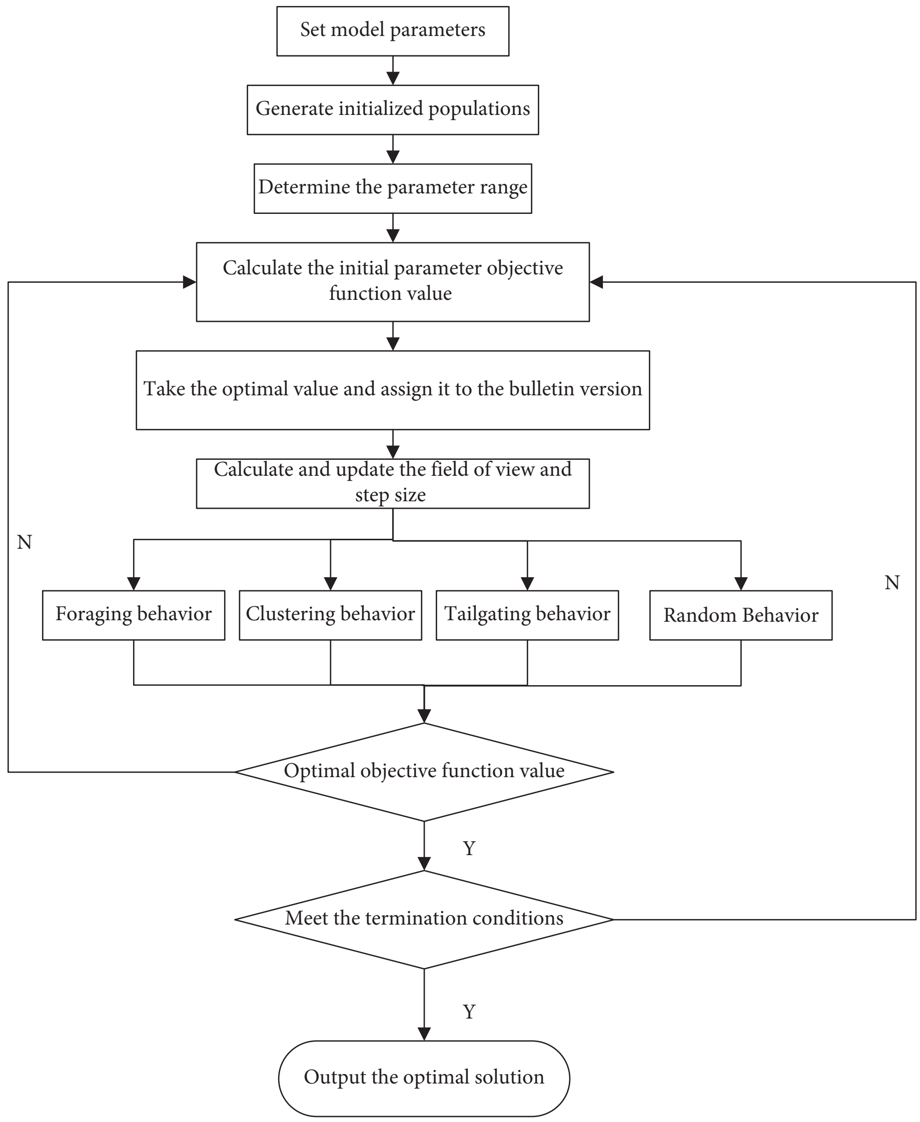


Figure 3: The calculation flow of AFSA.

$$p_i^t = \sum_{j=1}^{T} \xi_i^{t-j} x_i^{t-j}, \tag{3}$$

where t is the historical algebra and t_{ji} and x_{ij}^t are the learning rate and position of individual i at time t_j, respectively.

Compared with the standard artificial fish swarm algorithm, the evolution equation of the artificial fish swarm algorithm based on individual historical knowledge learning is as follows:

$$v_i^{t+1} = v_i^t + \left(x_{\mathrm{i}}^t - x^*\right) \cdot f_i \cdot \mathrm{rand1} + \left(x_{\mathrm{i}}^t - p_i^t\right) \cdot f_i \cdot \mathrm{rand2}. \tag{4}$$

2.3. Test Simulation. Two data sets A and B are selected for performance evaluation using benchmark test examples. Data set A consists of 12×6 to 9×9, composed of 11 groups of logistics data, with 1 instance in each group. Data set B consists of 20×5 to 75×20, with 3 instances in each group, and we repeated calculation for 30 times.

The algorithm is implemented by visual studio. In order to better evaluate the experimental results, the average percentage deviation (ARPD) is used as the performance parameter of the algorithm. The formula is defined as follows:

$$\mathrm{ARPD} = \frac{1}{N} \sum_{i=1}^{N} \frac{C_i(H) - C_i^*}{C_i^*} \times 100\%, \tag{5}$$

where n represents the number of times the instance runs, $C_i(H)$ is the $C_{\max}$ value obtained by the algorithm h to solve the instance I, and C_i^* is the best solution found so far. The smaller the ARPD, the higher the optimization efficiency of the algorithm. In the calculation process, the maximum number of iterations is 30; the population size is the product of the number of machines; the step factor is 0.1; the index parameter is 1.5; the discovery probability is 0.25.

In order to further verify the effectiveness of the optimized AFSA in solving the optimal value, the optimized AFSA is compared with OPSO, MFO, and SCA algorithms on data set A, and also with SA, GWO, and standard AFSAs on data set B. Figures 4 and 5, respectively, show the ARPD change curves of each algorithm on data sets A and B.

The solution efficiency of the optimized AFSA is 0.00 on the A data set, which is better than that of OPSO, MFO, and SCA. It shows that the algorithm can effectively solve small-scale data sets and has better optimization effect. At the same time, the discoverable AFSA is superior to other comparative algorithms in 17 of the 21 B data set cases. Compared with the average ARPD, QCS algorithm is the best, with the average ARPD of only 0.22, followed by the standard AFSA, which is 0.54, and the worst is SA, which is 0.88. Compared with other algorithms, the optimized AFSA improves by more than 20%. This shows that the optimized AFSA has some advantages over other algorithms.

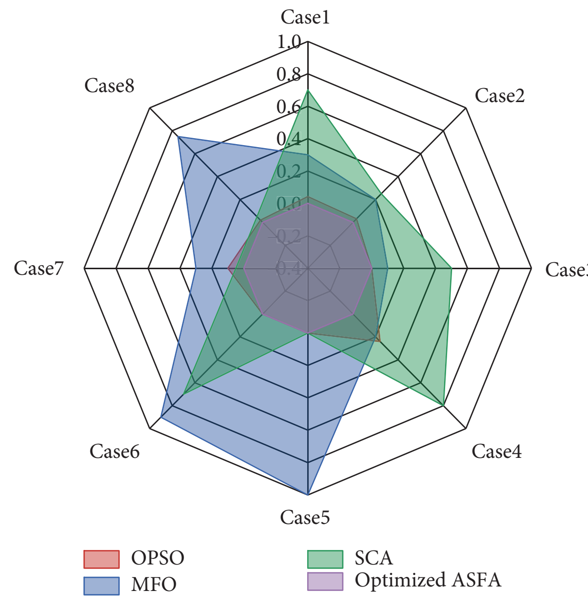


Figure 4: The ARPD based on data set A.

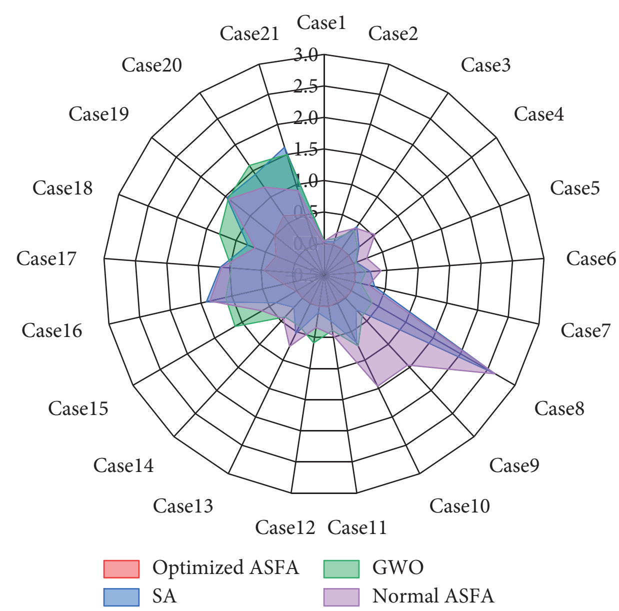


Figure 5: The ARPD based on data set B.

3. Design and Establishment of the Mathematical Model for Intelligent Selection of Logistics Distribution Sites

3.1. Model Establishment. Reasonable selection of distribution sites can provide convenience for couriers and customers, effectively improve distribution efficiency and reduce distribution costs, and have good social and economic benefits. The reasonable selection of logistics distribution sites can not only improve the quality of logistics service but also reduce the cost of logistics distribution [21]. The optimal selection model of logistics distribution sites needs to consider many factors such as environment, economy, customers, and geographical location [22]. In order to simplify the solution process, the following assumptions are made in the mathematical model:

(1) Multiple candidate distribution centers are responsible for delivering goods to customers (demand points), regardless of loading and unloading time

(2) In a fixed period of time given by a service period, each vehicle can carry out multijourney transportation of goods

(3) For each route, the vehicle departs from a candidate distribution center and returns to the candidate distribution center

(4) The capacity constraints of the candidate distribution centers are not considered

(5) The transportation cost per unit time of each distribution vehicle is a fixed value

(6) In each cycle (day), each demand point should be served by one vehicle only once

The operation process of the mathematical model for intelligent selection of distribution sites established in this paper is as follows:

(1) Determine the location and quantity of demand points.

(2) Determine the reasonable coverage of distribution sites.

(3) Select appropriate methods to determine the location of distribution sites.

(4) According to the actual geographical environment of the demand point, eliminate the nodes that are not convenient for the construction of the distribution point, and determine the final distribution point location according to the principle of the shortest “pick-up distance.”

The decision goal is how to determine the best location of the distribution point from the potential geographical location and optimize the vehicle route from the warehouse to the customer. On the premise of meeting the vehicle carrying capacity and the capacity of the distribution center, we reduce the operation and distribution cost of the entire logistics distribution system, so as to minimize the total cost of the entire logistics network and achieve the purpose of system optimization [23]. Based on this, a mathematical model is developed, which mainly studies the vehicle routing problem with multi-journey operation strategy and distribution center location based on dynamic demand. In this model, the customer location node and customer demand are uncertain factors. Other parameters and variables are defined as follows:

N: total number of demand points (customers)

i, j: subscript of demand point, $i, j = 1,2,3, \ldots, N$

W: candidate distribution centers

u, v: candidate distribution center

SC: construction cost of a single distribution center

R: the total number of vehicle k trips

q: the delivery quantity to the demand point J in one cycle

Q: the maximum loading capacity of the vehicle

VC: the vehicle start-up cost of each cycle

TC: vehicle transportation cost per unit time

t_{ij}: travel time from demand point I to demand point J

t_{ui}: travel time from the candidate distribution center u to the demand point I

T: maximum service time of a single delivery vehicle

M: an arbitrary large positive number

Decision variables are defined as follows:

x^r_{ijk}: if the vehicle k provides delivery service from the demand point i to the demand point j during the journey r, then $x^r_{ijk} = 1$; otherwise, $x^r_{ijk} = 0$

v^r_{kui}: if the vehicle k starts from the distribution center u and arrives at the demand point i, then $v^r_{kui} = 1$; otherwise, $v^r_{kui} = 0$

u^r_{kvj}: if vehicle k returns to v in distribution from demand point j, then $u^r_{kvj} = 1$; otherwise $u^r_{kvj} = 0$

Yu: if the candidate distribution center u is selected, $Yu = 1$; otherwise, $Yu = 0$

The model uses the following objective functions and constraints with the objective of minimizing the following values [24]:

$$Z = \mathrm{SC}\sum_{N=1}^{W} Y_N + VC\sum_{k=1}^{K}\sum_{N=1}^{W}\sum_{i=1}^{N}\sum_{r=1}^{1} V^r_{kui} + \mathrm{TC}\sum_{k=1}^{K}\sum_{r=1}^{R_i}\left(\sum_{N=1}^{W}\sum_{i=1}^{N} t_{Ni}\cdot V^r_{kui} + \sum_{i=1}^{N}\sum_{j=1}^{N} t_{ij}\cdot X^r_{ijk} + \sum_{V=1}^{W}\sum_{j=1}^{N} U^r_{kij}\right). \tag{6}$$

The restrictions are as follows:

$$\sum_{i=1}^{N}\sum_{k=1}^{K}\sum_{r=1}^{R_i} X^r_{ijk} = 1,\ j = 1, 2, \cdots, N, i \neq j, \tag{7}$$

$$\sum_{u=1}^{W}\sum_{i=1}^{N} V^r_{kui} = \sum_{u=1}^{W}\sum_{i=1}^{N} V^r_{kvj} r = 1, 2, \cdots, R_k, k = 1, 2, \cdots, K, \tag{8}$$

wherein equation (6) is the objective function, which represents the minimization of the total logistics system cost, specifically including the distribution center construction cost, vehicle start-up cost, and transportation cost. Equation (7) guarantees that only one vehicle serves each demand point. For any starting vehicle in any journey, the number of trips in and out of the distribution center should be the same, as shown in equation (8).

At the same time, the following constraints are set for the distribution process:

(1) Each distribution point can only be distributed by one distribution center
(2) The cargo volume of each distribution route shall not exceed the carrying capacity of the vehicle
(3) All vehicles meet the demand of the distribution point
(4) Ensure that the delivery vehicle is unique to the delivery point

3.2. Selection of Distribution Sites. In essence, the problem of selecting distribution sites is to transform the multi-distribution point problem into a single distribution point problem, which can be handled by the whole method. The whole method is to solve the problem from the perspective of the whole. At this time, the distribution centers and distribution points are regarded as a unified whole. Compared with the decomposition method, the whole method has strong global optimization ability. Figure 6 is a schematic diagram of the solution of the whole method [25]. The specific solution process of the whole method is as follows:

(1) First, a general distribution center shall be set in the system and directly connected with each actual distribution point. No distribution point is allowed between the virtual and real distribution centers. At this time, each actual distribution center is regarded as a special distribution point.
(2) It is assumed that all distribution vehicles in the distribution system start from the general

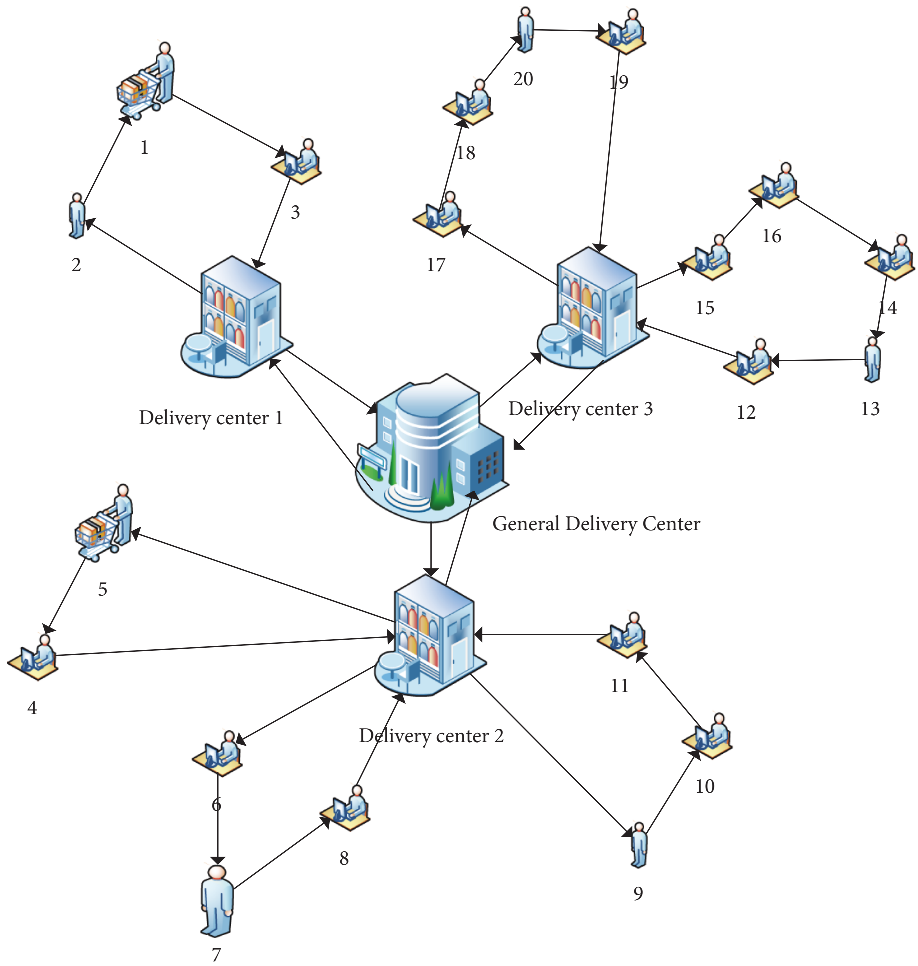


Figure 6: The schematic diagram of the solution of the whole method.

distribution center and must pass through the actual distribution center before driving to the distribution point for service. When the distribution vehicle completes the distribution task, it must first return to the original actual distribution center and then return to the general distribution center. If the distribution vehicle can continue to execute the task according to the model constraints after completing the distribution task and returning to the distribution center, the general distribution center can send the vehicle back to the distribution center in the previous distribution cycle to start a new round of distribution tasks.

(3) It should be pointed out that the general distribution center is only an ideal point, and the distance between the general distribution center and the actual distribution centers is zero. Therefore, there will be no travel distance, travel time, and no operating cost. It can be assumed as the only parking place of distribution vehicles in the distribution network, but each distribution vehicle is still parked in the actual distribution center.

When using the whole method to solve the multi-distribution center logistics vehicle routing problem, each distribution point to be served in this distribution network does not belong to a certain actual distribution center but it assigns each distribution point to be served to the corresponding actual distribution center according to the actual situation based on the entire distribution network. The introduction of the general distribution center makes it possible to directly model the entire distribution network, solves the problem that the time of the logistics vehicle routing problem cannot be directly connected, and effectively improves the global optimization ability of the logistics distribution network. Therefore, the overall method is adopted to analyze and solve the selection of logistics distribution sites. Based on this idea, the optimized AFSA intelligent algorithm in the previous section is introduced to optimize the selection of multiple distribution sites.

4. Application and Case Analysis

4.1. Current Situation Investigation. The time of this questionnaire survey is September 2021. A total of 1000 questionnaires were distributed to the operation area of CSC Logistics Company, of which 886 were recovered and 810 were valid. We carry out strict quality control on the data. The standards are as follows: (1) anonymous investigation; (2) random survey; (3) the baseline of dissatisfaction is dissatisfied + very dissatisfied more than 15%; (4) the baseline of satisfaction is very satisfied + more than 70% satisfied and less than 15% dissatisfied. The statistical results of the satisfaction rate questionnaire are shown in Figure 7. In the satisfaction survey of distribution efficiency, we mainly designed the timeliness of distribution and the management of distribution.

From the satisfaction survey on the distribution efficiency in this area, it can be seen that the proportion of users who are satisfied with the distribution speed is 32%, the proportion of users who choose to be very dissatisfied is 21%, and the proportion of users who choose to be dissatisfied is 22%. In terms of the proportion of satisfaction, the proportion of this area is not high. Users are generally satisfied with the distance from the distribution point and think it is convenient. In terms of timeliness of distribution, the proportion of very dissatisfied with the selection reached 31%, followed by the dissatisfied proportion of 28%, with very low satisfaction. To sum up, the distribution situation in this area needs to be optimized.

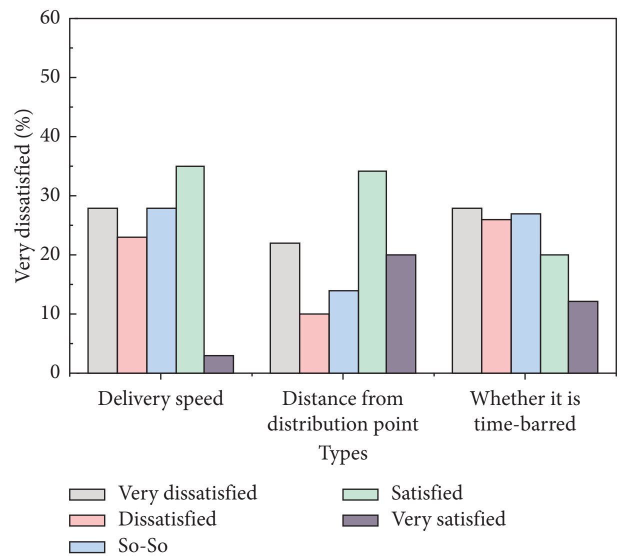


Figure 7: The satisfaction survey on the distribution efficiency.

4.2. Selection of Model Parameters. By studying the case of CSC Company's logistics distribution, combining the data of distribution center and distribution point in the case and the problem of CSC Company's distribution route optimization, the parameter selection is carried out. The logistics distribution project in the case is that a single distribution center provides distribution to multiple mutually dispersed distribution sites and needs to provide distribution within the time constraints of each distribution point. After the distribution service is completed, the logistics vehicles return to the distribution center. Therefore, the mathematical model established in this paper is applied to distribution planning, and the model parameters are collected and checked according to the actual data in the case. The optimization process is based on the optimized AFSA intelligent algorithm to solve the model, reducing the total cost of distribution and improving the distribution efficiency and logistics service level.

4.3. Distribution Site Data Preparation. The distance between each distribution point and the quantity of goods required by each distribution point are shown in Table 2 and Figure 8. The objective of the solution is to minimize the transportation cost on the premise of ensuring the distribution efficiency.

Table 2: The quantity of goods required by each distribution point.

Distribution site	X coordinate	Y coordinate	Business volume
1	4.6	3.6	220
2	2.2	0.8	520
3	1.6	4.8	760
4	0.8	2.4	520
5	2.2	2	310
6	3.8	1.6	370
7	1.4	4.6	160
8	4.2	1.2	490
9	4.8	1.4	340
10	4.2	3.4	430
11	2	2	310
12	2.4	0.4	220
13	0.6	4	340
14	3.4	3.4	610
15	1.8	1.8	190
16	4.4	0.8	130
17	1.2	4.4	190
18	4	2.2	160
19	4.8	5	310
20	3	3.6	520

4.4. Optimization of Route Cost and Distribution Efficiency. The optimized route is shown in Figure 9, and the specific cost indicators are shown in Table 3.

Comparing the results before and after optimization, the following can be seen:

(1) The total process is shortened, and the distribution efficiency is optimized. The total mileage of the initial distribution route is 115 km. After the optimization of the intelligent algorithm, the total mileage of the distribution route becomes 83 km, reducing 32 km and 28% compared with the original route. Under other conditions unchanged, the distribution vehicle travels at an average speed of 60 km/h, which saves about 1.5 hours in the whole distribution process and fully improves the distribution efficiency.

(2) The load rate is increased. After intelligent algorithm optimization, the distribution route only needs 4 vehicles to complete the distribution, reducing the use of distribution vehicles, reducing the company's fixed cost and improving the vehicle load rate. The average load of the distribution route optimized by the intelligent algorithm is 89.8%, and the optimized load rate is higher, which is more conducive to the development of the enterprise.

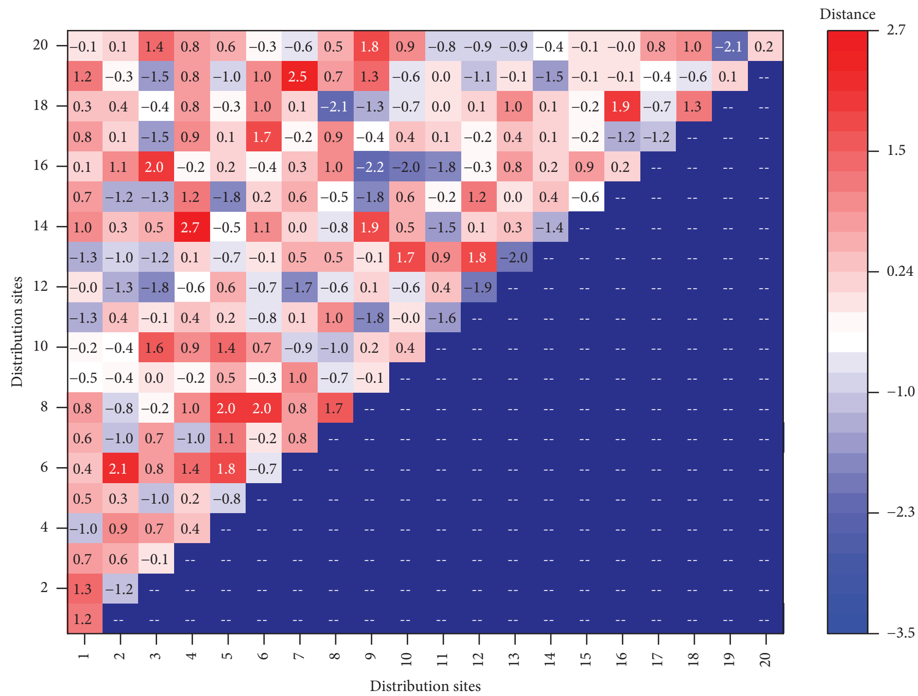


Figure 8: The distance between each distribution point.

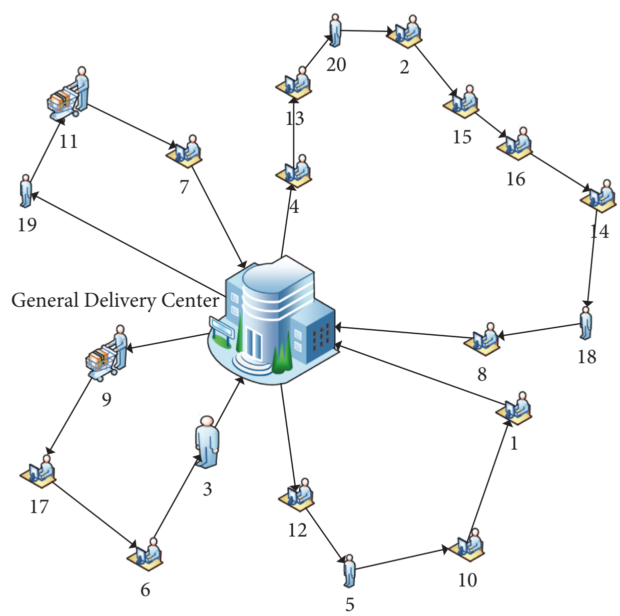


Figure 9: The optimized route.

Table 3: The specific cost indicators.

Number	Route	Total business volume	Load rating (%)	Mileage (km)
1	0–9–17–6–3–0	1630	92	15
2	0–12–5–10–1–0	1180	91	18
3	0–19–11–7–0	780	91	16
4	0–4–13–20–2–15–16–14–18–8–0	3480	85	34

(3) The total cost is reduced. After the intelligent algorithm optimization, the cost of the distribution route is lower than the original route. After optimization, the number of vehicles used becomes less and the fixed cost changes from 1967 yuan to 1599 yuan, a decrease of 15.8%. After optimization,

the total mileage becomes shorter, and the cost of goods damage in transit is greatly reduced. The cost of goods damage is changed from 801 yuan to 702 yuan, a decrease of 12.1%. At the same time, the total mileage of distribution after optimization is shorter, and the driver has enough time to deliver goods, which greatly improves customer satisfaction. The total cost is changed from 5099 yuan of the original route to 4038 yuan, a decrease of 17.2%.

(4) There is feasibility of intelligent algorithm. The optimized AFSA is superior to the original path in every index of the optimized distribution path, and the speed of finding the optimal solution is obviously faster than the genetic algorithm. This shows the feasibility of the optimized AFSA and the validity of the model established in this paper.

4.5. Suggestions. The construction of logistics is very important for fresh food and agricultural products enterprises. For the outside of the enterprise, it can guarantee the quality of products and quickly improve the enterprise image. For the internal part of the enterprise, the optimization of the distribution route is conducive to reducing the total distribution cost of the enterprise and improving the income of the enterprise. Based on the above research, the following suggestions are put forward for the future development of the company:

(1) Build the logistics distribution management system and optimize the route. In terms of the current development, CSC Company can consider establishing its own logistics distribution management system, which can not only simplify the logistics distribution information, strengthen the relationship between various logistics departments in the distribution center, and better cope with the changing market and suppliers. Moreover, by establishing the logistics distribution system, the rapid and accurate optimization of the distribution route can be realized. Select the appropriate distribution route, reduce the distribution cost, and improve the income of the enterprise.

(2) Control the number of distribution vehicles and improve the loading rate. The company can consider making improvements in other aspects while reducing costs by optimizing the distribution route. Regulate the number of outgoing vehicles of transport vehicles, and select the appropriate number of vehicles for distribution according to the demand of distribution sites. The distribution vehicle is different from the general vehicle, and the fixed cost caused by the vehicle loss and manpower and the transportation cost caused by the fuel consumption in the distribution process are much higher. To reduce the total distribution cost of the company, the method of increasing the vehicle weight rate of transportation can also be adopted. That is, when the transport vehicle is used correctly without speeding, overloading, fatigue driving, and other strict compliance with laws and regulations, the number of vehicles will be reduced and the load capacity will be increased. Considering the company's interests, the specific requirements of customers and the distribution sequence, the company can consider renting transport vehicles to reduce capital investment.

(3) "End joint distribution" can be implemented. Combined with the joint distribution from the secondary distribution center to the terminal distribution point, it will form a complete joint distribution from secondary distribution to the completion of distribution. If it can be implemented, it will be a further application of joint distribution. Joint distribution can be studied together with logistics express collection, and the two-way integration of express terminal collection and delivery can be implemented, so as to further improve logistics efficiency and reduce logistics costs.

(4) The number of distribution sites in the community can be further determined based on the optimal service radius of the end distribution sites and the time satisfaction of customers. At this point, we can start with the relationship between the service capacity and the service radius of logistics nodes, then find the cost of logistics nodes with different radii, introduce time satisfaction as a cost loss item, and finally find the partial derivative of the radius for the total cost to obtain the optimal radius.

(5) As for the model of distribution point selection, since the logistics data is in the hands of the express company and can only be obtained by investigation, the data volume is limited, which has a certain impact on the training of the model and the accuracy of the model. If the accuracy of the model is to be further improved, a large number of community data are required to train the model.

5. Conclusion

This paper focuses on the selection and optimization of urban logistics distribution sites and analyses the applicability of different intelligent algorithms to logistics data. The AFSA is selected as the optimization algorithm, and the optimization of the algorithm is realized by introducing knowledge learning. The optimized AFSA model is applied to the mathematical model of distribution point selection. An intelligent algorithm-driven logistics distribution point selection model is established and solved by using the optimized AFSA. Combined with the actual case of CSC Logistics Company, the route of logistics distribution point in a certain area is optimized and the model is verified and solved. The conclusions of this paper are as follows:

(1) On the whole, GWO and AFSA have better optimization ability than other intelligent algorithms and AFSA has better effect than GWO. Although the GWO algorithm can also get the optimal value when

performing global search, its convergence speed is slow and there is certain instability. Therefore, the AFSA is most suitable for solving logistics-related problems.

(2) Knowledge learning is introduced into the AFSA for improvement. The solution efficiency of the optimized AFSA is 0.00 on the A data set, which is better than OPSO, MFO, and SCA. It shows that the algorithm can effectively solve small-scale data sets and has better optimization effect. Compared with other algorithms, the optimized AFSA improves by more than 20%. This shows that the optimized AFSA has some advantages over other algorithms.

(3) After CSC Logistics Company uses this intelligent algorithm to select logistics distribution sites, the total process of logistics route becomes shorter and the distribution efficiency is optimized. The total mileage of the initial distribution route is 115 km. After the optimization of the intelligent algorithm, the total mileage of the distribution route becomes 83 km, reducing 32 km and 28% compared with the original route. About 1.5 hours were saved in the whole distribution process, which fully optimized the distribution efficiency.

(4) The construction of logistics is very important for fresh and agricultural products enterprises. The optimization of distribution route is conducive to reducing the total distribution cost of enterprises and improving the income of enterprises. For the future development of CSC Company, it is suggested to build a logistics distribution management system, optimize the route, control the number of distribution vehicles, and improve the loading rate.

Data Availability

The labeled data set used to support the findings of this work is available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] S. Winkelhaus and E. H. Grosse, "Logistics 4.0: a systematic review towards a new logistics system," *International Journal of Production Research*, vol. 58, no. 1, pp. 18–43, 2020.

[2] F. Sgarbossa, E. H. Grosse, W. P. Neumann, D. Battini, and C. H. Glock, "Human factors in production and logistics systems of the future," *Annual Reviews in Control*, vol. 49, pp. 295–305, 2020.

[3] E. Tijan, S. Aksentijević, K. Ivanić, and M. Jardas, "Blockchain technology implementation in logistics," *Sustainability*, vol. 11, no. 4, p. 1185, 2019.

[4] W. Liu, Y. Liang, X. Bao, J. Qin, and M. K. Lim, "China's logistics development trends in the post COVID-19 era," *International Journal of Logistics Research and Applications*, vol. 25, no. 6, pp. 965–976, 2022.

[5] V. Yavas and Y. D. Ozkan-Ozen, "Logistics centers in the new industrial era: a proposed framework for logistics center 4.0," *Transportation Research Part E: Logistics and Transportation Review*, vol. 135, Article ID 101864, 2020.

[6] L. Koh, A. Dolgui, and J. Sarkis, "Blockchain in transport and logistics – paradigms and transitions," *International Journal of Production Research*, vol. 58, no. 7, pp. 2054–2062, 2020.

[7] Y. Ding, M. Jin, S. Li, and D. Feng, "Smart logistics based on the internet of things technology: an overview," *International Journal of Logistics Research and Applications*, vol. 24, no. 4, pp. 323–345, 2021.

[8] X. Qin, Z. Liu, and L. Tian, "The optimal combination between selling mode and logistics service strategy in an e-commerce market," *European Journal of Operational Research*, vol. 289, no. 2, pp. 639–651, 2021.

[9] M. Pournader, Y. Shi, S. Seuring, and S. L. Koh, "Blockchain applications in supply chains, transport and logistics: a systematic review of the literature," *International Journal of Production Research*, vol. 58, no. 7, pp. 2063–2081, 2020.

[10] F. Facchini, J. Oleśków-Szłapka, L. Ranieri, and A. Urbinati, "A maturity model for logistics 4.0: an empirical analysis and a roadmap for future research," *Sustainability*, vol. 12, no. 1, p. 86, 2019.

[11] I. J. Orji, S. Kusi-Sarpong, S. Huang, and D. Vazquez-Brust, "Evaluating the factors that influence blockchain adoption in the freight logistics industry," *Transportation Research Part E: Logistics and Transportation Review*, vol. 141, Article ID 102025, 2020.

[12] L. Chen, M. Ma, and L. Sun, "Heuristic swarm intelligent optimization algorithm for path planning of agricultural product logistics distribution," *Journal of Intelligent and Fuzzy Systems*, vol. 37, no. 4, pp. 4697–4703, 2019.

[13] K. Kuru, D. Ansell, W. Khan, and H. Yetgin, "Analysis and optimization of unmanned aerial vehicle swarms in logistics: an intelligent delivery platform," *IEEE Access*, vol. 7, pp. 15804–15831, 2019.

[14] W. Hu, "An improved flower pollination algorithm for optimization of intelligent logistics distribution center," *Advances in Production Engineering & Management*, vol. 14, no. 2, pp. 177–188, 2019.

[15] B. Feng and Q. Ye, "Operations management of smart logistics: a literature review and future research," *Frontiers of Engineering Management*, vol. 8, no. 3, pp. 344–355, 2021.

[16] K. Leng and S. Li, "Distribution path optimization for intelligent logistics vehicles of urban rail transportation using VRP optimization model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 2, pp. 1661–1669, 2022.

[17] J. Wang, M. K. Lim, Y. Zhan, and X. Wang, "An intelligent logistics service system for enhancing dispatching operations in an IoT environment," *Transportation Research Part E: Logistics and Transportation Review*, vol. 135, Article ID 101886, 2020.

[18] Y. Li, M. K. Lim, and M. L. Tseng, "A green vehicle routing model based on modified particle swarm optimization for cold chain logistics," *Industrial Management & Data Systems*, vol. 119, 2018.

[19] L. Ma, J. Li, Z. Liu et al., "Intelligent algorithms: new avenues for designing nanophotonic devices [Invited]," *Chinese Optics Letters*, vol. 19, no. 1, Article ID 011301, 2021.

[20] E. B. Tirkolaee, A. Goli, and G. W. Weber, "Fuzzy mathematical programming and self-adaptive artificial fish swarm algorithm for just-in-time energy-aware flow shop scheduling problem with outsourcing option," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 11, pp. 2772–2783, 2020.

[21] M. Yu, G. Yue, Z. Lu, and X. Pang, "Logistics terminal distribution mode and path optimization based on ant colony algorithm," *Wireless Personal Communications*, vol. 102, no. 4, pp. 2969–2985, 2018.

[22] X. Xiong, F. Zhao, Y. Wang, and Y. Wang, "Research on the model and algorithm for multimodal distribution of emergency supplies after earthquake in the perspective of fairness," *Mathematical Problems in Engineering*, vol. 2019, Article ID 1629321, 12 pages, 2019.

[23] A. Javaheri, C. Brites, F. Pereira, and J. Ascenso, "Mahalanobis based point to distribution metric for point cloud geometry quality evaluation," *IEEE Signal Processing Letters*, vol. 27, pp. 1350–1354, 2020.

[24] J. Wang and X. Lu, "Sustainable and resilient distribution systems with networked microgrids [point of view]," *Proceedings of the IEEE*, vol. 108, no. 2, pp. 238–241, 2020.

[25] T. Wenzel, G. Santin, and B. Haasdonk, "A novel class of stabilized greedy kernel approximation algorithms: convergence, stability and uniform point distribution," *Journal of Approximation Theory*, vol. 262, Article ID 105508, 2021.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9824921, 1 page
https://doi.org/10.1155/2023/9824921

Retraction

Retracted: Construction of an Inquiry-Based Teaching Model for Ideological and Political Education in Colleges and Universities from the Perspective of Deep Learning

Mobile Information Systems

Received 13 September 2023; Accepted 13 September 2023; Published 14 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.

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] B. An and L. Gao, "Construction of an Inquiry-Based Teaching Model for Ideological and Political Education in Colleges and Universities from the Perspective of Deep Learning," *Mobile Information Systems*, vol. 2022, Article ID 9286979, 12 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 9286979, 12 pages
https://doi.org/10.1155/2022/9286979

Research Article

Construction of an Inquiry-Based Teaching Model for Ideological and Political Education in Colleges and Universities from the Perspective of Deep Learning

Bo An[1] and Lin Gao[2]

[1]*Hebei Chemical and Pharmaceutical Vocational and Technical College, Shijiazhuang, Hebei 050026, China*
[2]*Hebei Normal University, Shijiazhuang, Hebei 050000, China*

Correspondence should be addressed to Lin Gao; lingao@hebtu.edu.cn

Received 31 July 2022; Revised 7 September 2022; Accepted 13 September 2022; Published 7 October 2022

Academic Editor: R. Mo

Social development and people's daily life have undergone unprecedented great changes under the impact of the reconstruction of the knowledge economy and information age, and ideological and political education, as a social practice activity for cultivating people, has also been affected. The purpose of knowledge has shifted from retelling facts to discovering and applying knowledge, requiring learners to be able not only to have basic factual rules and operating procedures but also to experience irrational factors such as character, value, and spirit about knowledge and beyond it, so as to have a sense of the quality of deep learning. And, this deep learning ability has become a key ability necessary to support human survival and development in the future artificial intelligence era, and it is also the main basis for measuring the learning ability of learners. Because of this, deep learning research has set off a wave of teaching reforms around the world. This research is based on the analysis of the stage characteristics of smarter classrooms and the adaptation of inquiry-based teaching to the development stages of smarter classrooms. Inquiry-based teaching needs to be designed. The demand analysis of inquiry-based teaching design is carried out, and then the main line of inquiry-based teaching design is to focus on the development of students' wisdom to promote the development of students' wisdom. A detailed analysis and design are carried out, and the specific performance is given. Inquiry-based teaching is applied to practice and test its effectiveness. The interview method was used for the experimental class. The practical results show that the advanced design of inquiry-based teaching in smarter classrooms is very effective, providing a reference for smarter teaching in primary and secondary schools.

1. Introduction

The current society has developed rapidly, social life and the main contradictions in society have changed, and the development of informatization and intelligence has penetrated into all aspects of life, which determines that the function of ideological and political education and the role of value will inevitably change. On this basis, the urgent requirement for innovative development of ideological and political education is put forward, so that it can better guide and practice, do a good job in the educational transformation and quality improvement of people's thoughts, and finally promote the innovative development of ideological and political education. Looking back at the research on ideological and political education since the reform and opening up, it can be found that whether it is the subject of education research at all levels from the central to the local level, or the practice of ideological and political education reform in schools at all levels, the effectiveness of ideological and political education has been subject to the attention of theoretical researchers and scholars. However, the corresponding effect has been widely questioned by all walks of life for a long time. The ideological and political education in schools has become more and more powerless, and the effectiveness of ideological and political education is low, which seems to be unanimously agreed by everyone. There are many factors that affect the effectiveness of ideological and political education. Among them, the traditional, oldfashioned, and backward teaching of ideological and political education in schools, such as materialization of teaching objects, generalization of teaching content, and rigid teaching methods, is one of the key factors. In order to get rid of this dilemma of internal and external difficulties, school ideological and political education based on this must follow the pace of social development and carry out profound changes that meet the needs of current social development.

The great progress of technology has brought profound changes to ideological and political education. Modern classroom teaching is inseparable from the supporting role of the information environment. With the continuous application and development of the concept of smart education in subject teaching, it has become a traditional classroom teaching method. The internal driving force of the transformation of new classroom teaching methods is used to guide the development of teaching activities. The new classroom is the smart classroom, which is the realization form of the smart education concept focusing on the classroom. The smart classroom learning environment has changed from the traditional blackboard and podium-based environment to a smart learning environment. In addition, the way of interaction between teachers and students and the way of delivery of learning content have also changed. Education and social and economic development complement each other. In the industrial society before the information age, the teaching-based teaching mode can adapt to and meet the needs of education. However, with the advent and development of the information age, multimedia computers, interactive whiteboards, and more advanced tablet computers have entered the classroom one after another. The classroom environment is constantly changing, and the high-tech content is increasing. The smart classroom also has different stages of development. In different stages of the smart classroom, its teaching and model design is particularly important. At each stage, appropriate teaching design is carried out to meet the needs of the times, so that technology and teaching are perfectly combined, so as to better promote the cultivation of students' wisdom and the all-round development of education. At the same time, it is necessary to change the teaching method and further use advanced information technology to explore and establish a new teaching model of ideological and political education that is teacher-led and learner-centred. Cultivate students' learning ability in the information environment, cultivate students' habit of using information technology and information means to learn, encourage students to use information technology to learn actively, autonomously, and cooperatively in the information environment, and enhance students' learning in the network learning environment. The ability to ask questions, analyze problems, and solve problems improves the quality of student learning. Inquiry-based teaching is a modern teaching mode actively advocated by the new curriculum reform. It can mobilize students' enthusiasm for learning, guide students to conduct independent, cooperative, and inquiry-based learning, and focus on cultivating and developing students' intelligence. Students are the main body, changing the traditional teaching method and teaching mode in which students passively accept knowledge and repeat mechanical operations. At the same time, it is also a kind of teaching in which teachers are the advocates and facilitators of teaching, and students are active participants in learning methods and learning styles. Therefore, inquiry-based teaching can adapt to the needs of the continuous development of the information environment.

2. Related Work

The study of inquiry learning has a long history in foreign countries, which can be traced back to Socrates' theory of "midwifery," that is, through the teacher's continuous questioning of the students, the students can be stimulated to think in the questions. But the concept of "inquiry-based learning" was first put forward by Dewey. Before that, most scholars believed that the method of science education was to let students learn a lot of scientific knowledge, concepts, and principles through direct teaching. Based on this, Dewey proposed the use of inquiry methods in school science education in his speech at American Association for the Advancement of Science. American educator Schwab (Joseph J. schwab) first proposed the concept of "inquiry learning" in his speech "Science Teaching as Inquiry." He pointed out that "If students are to learn the methods of science, what better way to learn than by being actively involved in the process of inquiry?" This sentence has had a profound impact on inquiry learning in science education. Based on this, Schwab proposed an inquiry-based learning method based on reading literature rather than experiments, which he called "inquiry into inquiry." Such discussion allow students to understand how scientific knowledge arises and what are the basic elements of scientific knowledge? The research of foreign scholar FDochy pointed out two points: first, inquiry learning can promote students' mastery of skills, and students can keep knowledge longer through inquiry learning, but it affects the efficiency of knowledge learning. Second, students' original intelligence level is a very important moderating variable for students' skills and knowledge learning. There are huge differences in the skills and knowledge acquired by students with different intelligence levels in exploratory learning. That is to say, the effective development of inquiry-based learning requires the corresponding knowledge and ability of students. Moreover, compared with lecture style teaching, the individual differences in both the inquiry learning process and the learning results will be greater. Balci and Cakiroglu explored "inquiry-based learning teaching strategies." They believed that, in the process of learning science, there are many modes of inquiry. The 5Es learning cycle is one of the most common methods of inquiry, and it consists of the following steps: (1) participation: students participate in survey questions; (2) exploration: students plan, design, conduct experiments, and record experimental data; (3) Refinement: students extend and apply their findings in new contexts, especially in everyday life; (4) evaluate: in a variety of ways, evaluate students' performance during the experiment and their experimental results, such as an activity report, instructor-led observations during the activity, and student demonstrations of the experiment. Saksri Supasorn and Anchule Lordkam confirmed in their research that the use of five scientific inquiry learning activities (12 hours) to isolate natural substances is effective in improving students' academic performance. It can be seen that the implementation of daily life-based inquiry experiments is effective, and it can promote students' attitudes towards learning science, which in turn increases their conceptual understanding and academic performance. Therefore, it is best to use inquiry learning activities in the whole college learning process so that students can improve their scientific inquiry skills and the ability to build knowledge through the inquiry process [1–8].

3. Deep Learning Theory and Methods

3.1. Smart Classroom. The connotation of smart classroom is derived from the concept of smart education, and it is the result of the implementation of smart education in classroom teaching. Different researchers have different views on the concept of smart classroom. Therefore, there is no exact and unified concept. For example, Zhu Zhiting and others believe that smart education is to cultivate the learning wisdom of learners, based on information technology, make full use of the characteristics of information technology, create a smart learning space environment, and promote the comprehensive, coordinated, and sustainable development of learners. Tang Yewei and others believe that smart classroom is a new type of classroom, a personalized, digital, and intelligent classroom [9]. This paper believes that smart classroom is based on the theory of teaching and learning, guided by the concept of smart education, and takes the generation of students' wisdom as the fundamental purpose. With the application of information technology means such as learning analysis, it is helpful for teachers' teaching and students' personalized learning in an effective classroom, specifically as shown in Figure 1.

3.2. Inquiry-Based Teaching. Inquiry-based teaching, as its name implies, is an inquiry-based teaching. Among them, "inquiry" refers to the way of inquiry, usually referring to the methods, means and strategies used by natural and social phenomena. "Inquiry-based teaching" originated from Plato divides students into two different classes for different education: the first class teaches truth-telling education, and the second class allows students to take a critical approach to what they have learnt. Attitudes are questioned and can even be revised and improved. The teaching method used in the second class is "inquiry-based teaching." In China, it can be considered that Confucius is the first person to use inquiry-based teaching in teaching, and his heuristic teaching idea is to promote students' autonomy and inquiry-based learning. Inquiry-based teaching is to introduce the form of inquiry into teaching. Like scientific inquiry, let students learn in an inquiry-based way. This process is inquiry-based learning for students and inquiry-based teaching for teachers. Specifically, inquiry-based teaching refers to the creation of various teaching situations for students with the help of teachers' inspiration and guidance, using textbooks or digital teaching resources as the content and providing students with questions that can be fully questioned, explored, and discussed, so that students can take the initiative. It is a teaching form in which the knowledge learned can be applied in solving practical problems. Inquiry-based teaching focuses on cultivating and developing students' intelligence and creative thinking. It can mobilize students' enthusiasm and initiative, and cultivate students' ability to discover, ask, analyze, and solve problems. Inquiry-based teaching is a multilateral activity of teaching and learning in which students are the main body and teachers are the leading, and both teachers and students participate together. It is a special teaching method and teaching mode to cultivate students' scientific inquiry ability as shown in Figure 2 [10].

4. Design of the Teaching Mode of Inquiry-Based Ideological and Political Education in the Smart Classroom

4.1. Demand Analysis of Advanced Design of Inquiry-Based Ideological and Political Education Teaching. The ultimate goal of inquiry-based teaching in the integration stage of smart classroom is to focus on the generation of students' wisdom in the initial stage to promoting the development of students' wisdom ability. The inquiry-based teaching in the integration stage is designed in the environment of tablet computer intelligent mobile terminal. It takes the intelligent teaching goal formed by the core literacy of the subject + the three-dimensional teaching goal as the teaching goal. "Wisdom" teaches "wisdom" and forms a wisdom teaching system. Under the teaching concept of improving students' higher-order thinking ability, problem-solving ability, and core literacy of the subject, intelligent teaching is realized through the teaching links before, during, and after class. Therefore, inquiry-based teaching can be called "individual interactive" inquiry-based teaching, interspersed with the wisdom of combining online evaluation and offline evaluation, formative evaluation and summative evaluation, and knowledge skills and thinking ability. Teaching evaluation ultimately promotes the development of students' intellectual ability. The transition from the initial stage to the integration stage is reflected in various aspects such as theoretical basis, teaching objectives, technical conditions, program design of teaching activities, and teaching evaluation. The overall design of inquiry-based teaching from the initial stage to the integration stage is shown in Figure 3 [11].

4.2. Teaching Goals of Wisdom Ideological and Political Education. In the smart classroom environment, students' wisdom is constantly developing. The core of teaching design from the beginning to the integration stage is to promote the development of students' wisdom. Its teaching concept as a supporting condition has undergone great changes compared with the initial stage. Teaching goals have also changed. For the stage of the integration of technology and teaching concepts, only promoting the development of

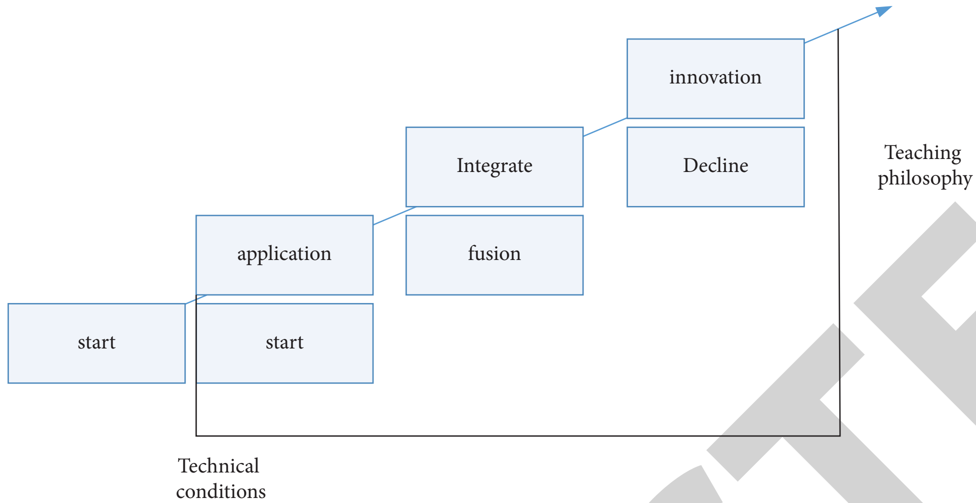


Figure 1: Technical conditions.

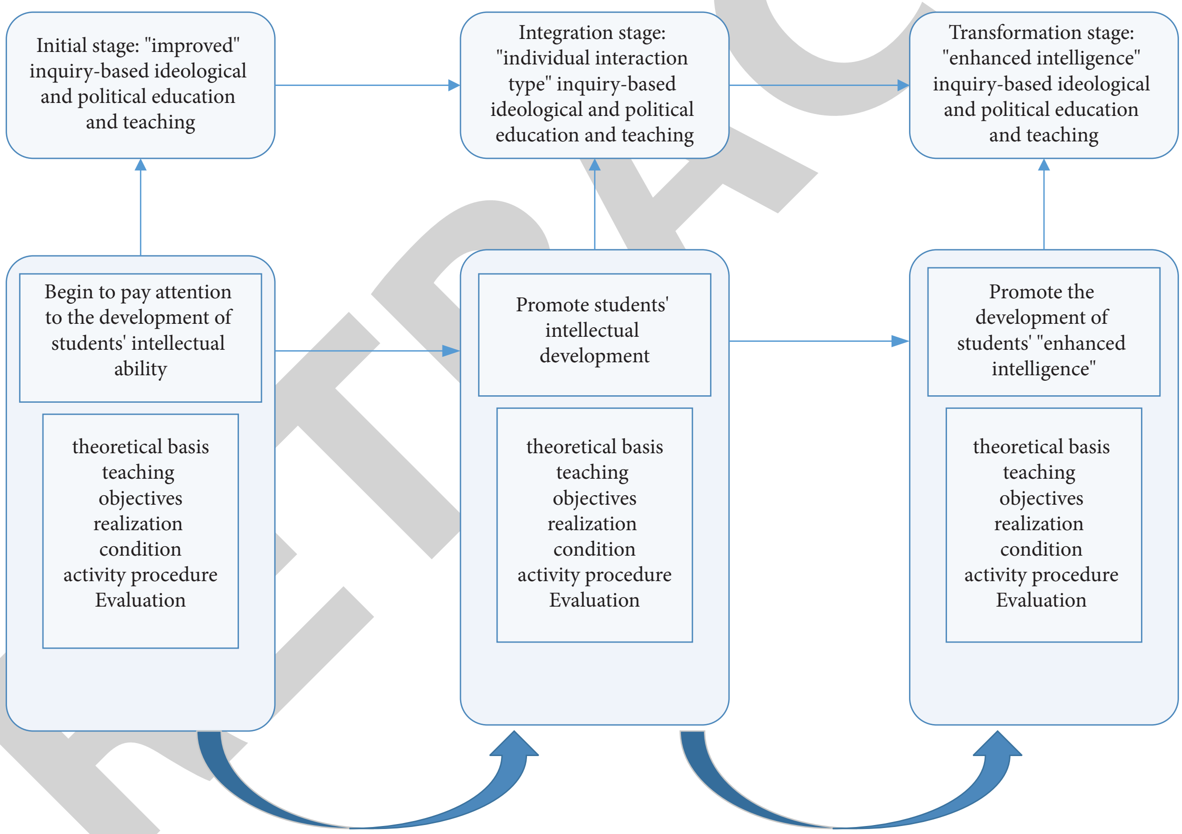


Figure 2: The advanced process of inquiry-based teaching in the smart classroom.

students' three-dimensional goals can no longer meet the ultimate goal of promoting the development of students' wisdom, and it cannot fully complete the purpose of "meaning construction" of knowledge. Therefore, in the context of the integration of smart classrooms, the teaching objectives of ideological and political education are transformed from three-dimensional objectives to intelligent teaching objectives of ideological and political education. The specific advanced performance is shown in Table 1 [12].

The "improved" inquiry-based ideological and political education and teaching in the multimedia network environment in the initial stage takes the three-dimensional goal as the teaching goal of ideological and political education. In the initial stage, students can get rid of the limitations of blackboards and chalks, and with the support of technology, they can actively discover problems, analyze problems, and try to solve problems with the cooperation of teachers and classmates. It should be able to initially take care of students'

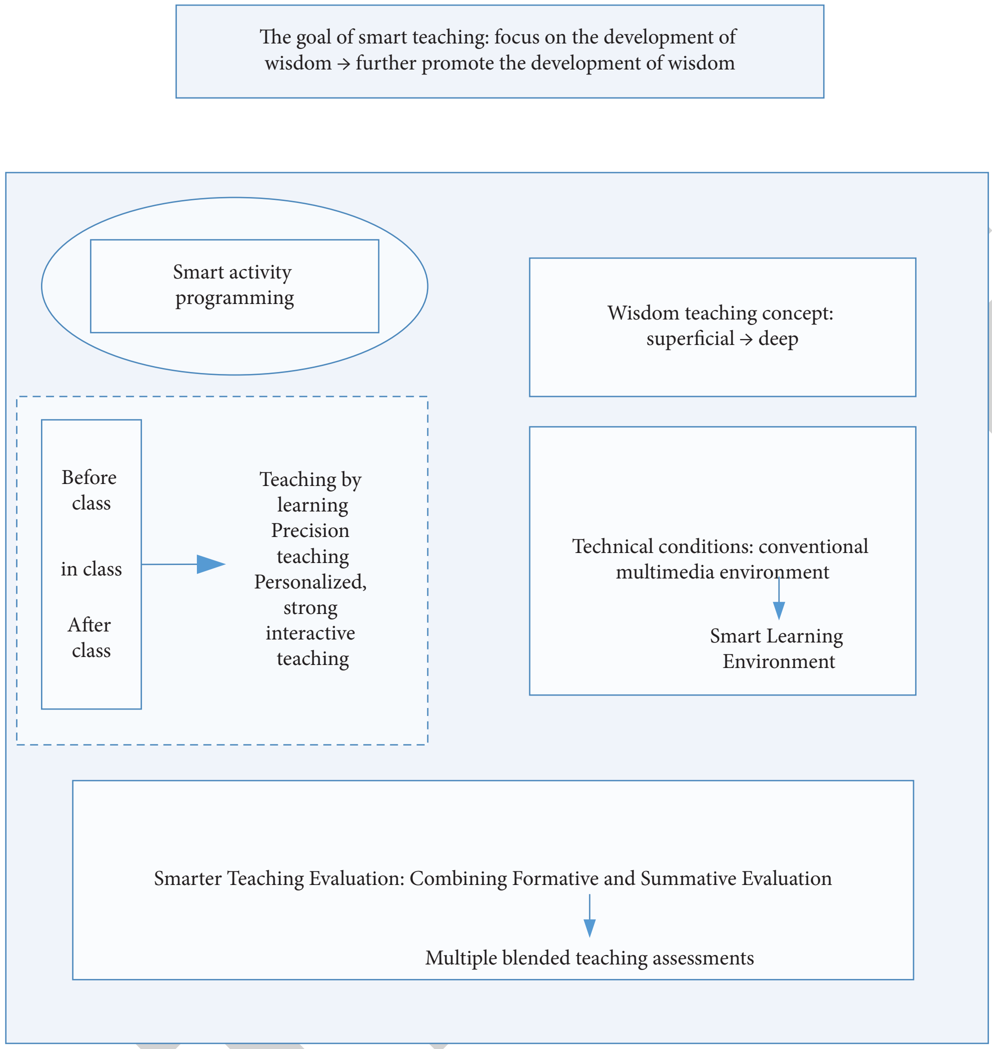


Figure 3: The overall design of inquiry-based teaching from the initial stage to the integration stage.

Table 1: Advanced performance of teaching objectives.

	"Improved" inquiry-based teaching	"Individual interactive" inquiry teaching
Teaching objectives	Relying on 3D goals to cultivate wisdom: knowledge and skills process and method emotional attitudes and values	Smart teaching goals: (1) Deep learning (2) Higher order thinking skills (3) The core competencies of students
Advanced performance: smarter teaching concept integrates three-dimensional goals and develops into smarter teaching goals		

higher-order thinking skills such as analysis, problem-solving, and innovative thinking. In addition, the technical conditions and teaching concepts in the initial stage have laid the foundation for the cultivation of students' innovative ability. However, in the integration stage, the technical conditions are more perfect, the teaching concept is more advanced, and the students' initiative is more prominent. After finding a problem, you can use tools to analyze and solve the problem by yourself, and under the drive of divergent thinking, you can innovate the problem-solving method to reflect your own personality. In this process, students' problem-solving ability is continuously improved, and high-level abilities such as innovative thinking are continuously improved. The "individual interactive" inquiry-based teaching in the integration stage is based on the three-dimensional teaching goal and integrates deep learning, higher-order thinking ability, and core literacy of disciplines to form a smart teaching goal of ideological and political education [13].

4.3. Technical Conditions. Compared with the initial stage, the technology in the fusion stage is more advanced, and its infrastructure, resources, and tools have advanced to a level that is compatible with the "personal interaction" inquiry-based teaching. In terms of the technical conditions of inquiry-based teaching design, the advanced performance from the initial stage to the integration stage is shown in Table 2.

As shown in the table above, with the development of informatization, its infrastructure has developed from a multimedia network environment in the initial stage to a mobile terminal device that promotes students' mobile learning. Ideological and political education teaching resources are more abundant, and the cloud, network, various rich media ideological and political education resources, digital ideological and political education resources, and generative ideological and political education resources provided by the smart classroom information environment composed of terminals, etc., promote students' personalized learning, digital learning, and interactive learning. In terms of tools, the integration stage supports the development of students' higher-order thinking and the solution of complex wisdom problems and can carry out personalized teaching and learning based on the students' whole-process learning data provided by ideological and political education learning analysis technology.

4.4. Smart Activity Programming. In terms of activity procedures, it is an inquiry-based teaching design, so teaching activities are carried out based on student autonomy, cooperation, and inquiry. Compared with traditional teaching, the two stages are "student-centered, with students being the main body and teachers being the facilitators and helpers" to carry out teaching activities. However, compared with the two stages, the "individual interaction" teaching in the integration stage has its own unique advantages, which can better promote the development of students' wisdom [14]. Before class, push personalized learning resources through the smart platform, which can realize "learning to teach"; in the classroom, due to the development of intelligent technology, students can better carry out exploration activities, and teachers can use various ideological and political education teaching tools to conduct teaching evaluation. After class, the design of inquiry-based teaching in the integration stage can effectively carry out personalized teaching and learning and improve the teaching of personalized teaching and differentiated teaching that cannot be well realized in the integration stage due to technology and other reasons. The advanced performance of inquiry-based ideological and political education teaching from the beginning to the integration stage is shown in Table 3.

Ideological and political education teaching activities from the beginning to the integration stage is to better realize the teaching concept of "determining teaching based on learning, precision teaching, interactive teaching, individualized teaching, and differentiated teaching" in order to realize the comprehensive development of cultivating students' wisdom. The smart activity program in the integration stage and the development from the start to the integration stage are mainly carried out from three parts: before class, during class, and after class [15].

4.4.1. Before Class. The preclass ideological and political education and teaching activities from the start to the integration stage mainly include teacher activities and student activities. However, due to the lack of support of advanced technical conditions and infrastructure, and the teaching concept cannot achieve the desired purpose, it cannot fully achieve accurate teaching. For the initial stage, teachers distribute study guide plans to students before class, and students conduct previews before class. Teachers analyze the completion of students' study guide plans, find problems, and improve teaching strategies for ideological and political education. Although teachers can analyze students according to their usual performances, they lack comprehensive understanding of students due to lack of systematic study analysis, the interaction between teachers and students before class is not strong, and the depth of communication is not enough. It cannot fully achieve personalized teaching and good teaching based on learning. With the emergence of the integration stage of smart classrooms, inquiry -based ideological and political education and teaching can be advanced, which effectively solves the problems existing in the initial stage. Due to the use of tablet computers, it is truly possible to teach by learning. It embodies the concept of personalized learning. The advanced development from the initial stage to the fusion stage is shown in Table 4 [16].

For the integration stage, the preclass ideological and political education and teaching activities mainly include learning situation analysis, resource push, independent learning, and exchange feedback. It fully reflects the independent learning of students. As shown in Figure 4, the teaching activities of teachers and students in the integration stage are shown in Figure 4.

4.4.2. Inclass. Ideological and political education in the initial stage of smart classroom is carried out in a multimedia network classroom environment, starting from the purpose of paying attention to the development of students' wisdom, aiming at cultivating innovative and creative thinking ability and at the same time emphasizing students' problem awareness and enthusiasm for learning. The cultivation of initiative is an inquiry-based ideological and political education and teaching model with the ultimate goal of promoting students' wisdom and development, with students' independent inquiry and cooperative learning as the main line. The main teaching activity process include creating a situation, eliciting a problem, discovering a problem, raising a problem, guessing a hypothesis, exploring a problem, solving a problem, communicating and interacting, evaluating feedback, and integrating innovation. Inquiry-based teaching in the initial stage can promote students' enthusiasm and initiative in learning more than traditional teaching and can also stimulate students' problem awareness. In the multimedia network teaching environment, teachers raise questions by creating an environment, students discover problems through teaching situations and establish connections between old and new knowledge, and teachers trigger students' conjectures and assumptions about problems through mind maps and cognitive tools. Then, they carry out independent

Table 2: Advanced performance of technical conditions.

	"Improved" inquiry-based teaching	"Individual interactive" inquiry teaching
Technical conditions	Infrastructure: multimedia network technology environment (electronic whiteboard). Resources: paper resources + digital network resources. Tools: external cognitive tools such as mind maps, concept maps, information tools, communication tools, and evaluation tools	Infrastructure: mobile terminal (tablet) smart learning environment. Resources: personalized, digitized web resources + generative resources. Tools: various cognitive tools
Advanced performance: effective application of smart learning environments such as smart mobile terminals and smart classrooms		

Table 3: Advanced performance of ideological and political education teaching activities program.

	"Improved" inquiry-based teaching	"Individual interactive" inquiry teaching
Teaching activities program	Before class: in the preview class: creating a situation, eliciting problems, discovering problems, raising questions, guessing assumptions, exploring problems, solving problems, communicating and interacting, evaluating feedback, integrating innovation after class: consolidation review	Before class (teaching based on learning): learning situation analysis, resource push, independent learning, communication and feedback during class (precision teaching, interactive teaching, and personalized teaching): creating situations, intelligent interactive teaching, independent cooperation and exploration, display and communication, real-time after-class testing (personalized teaching): personalized teaching and learning, communication and reflection, homework summary
Advanced performance: through the design of preclass, in-class, and after class in the integration stage, it can better realize teaching concepts such as teaching based on learning, precision teaching, strong interactive teaching, and personalized teaching and better promote the realization of smart teaching		

Table 4: Comparison of the advanced development of the preclass link from the initial stage to the integration stage.

	"Improved" inquiry-based teaching	"Individual interactive" inquiry teaching
Teacher activity	Distribute study guides Correction of the case study guide Analyze student preparations Improve teaching strategies	Study situation analysis Resource construction Resource release Instructional design
Student activities	Independent study guide	Self-study and exchange of experience

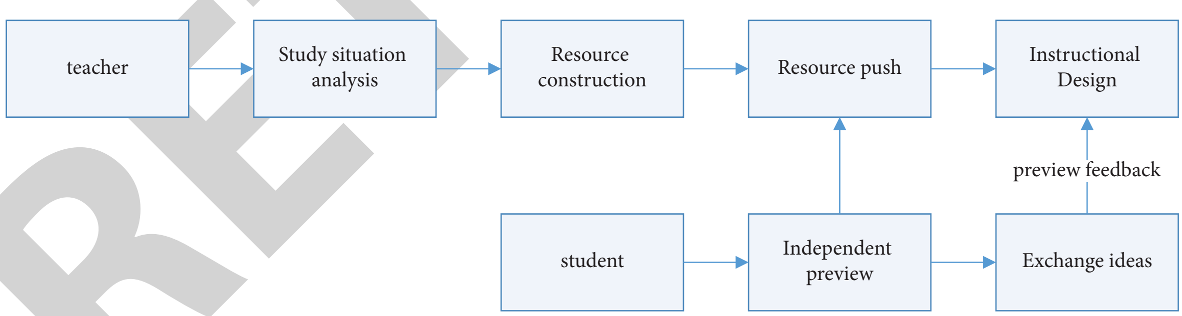


Figure 4: The process of preclass activities.

and cooperative exploration, and through multimedia computer upload function, demonstration function, and other functions, teachers and students can communicate and interact, evaluate, and give feedback, so that students can better reflect, improve innovative thinking ability, and improve teaching effect. . However, due to the limitations of teaching concepts and technical conditions, although the questions raised by students in the teaching process are explored through cooperation, most of them are answered with the help of teachers, and due to the lack of data provided by the smart learning platform in the students' learning process, a complete individualized teaching cannot carry out precise teaching, so the individualized development of students is insufficient. Therefore, in order to better promote the personalized development of students, promote strong interaction between teachers and students, and promote precision teaching, the inquiry-based teaching in the initial stage needs to develop to the integration stage, and due to the emergence of intelligent technologies such as big data and cloud computing, integration the inquiry-based teaching of ideological and political education in the initial stage can solve the shortcomings of inquiry-based teaching in the initial stage, and the inquiry-based ideological and political education teaching in the integration stage of smart classrooms solves the lack of interaction in the initial stage. Insufficient personalized development and other problems: we have achieved precise teaching, interactive teaching, and personalized teaching, which further promotes the development of students' wisdom from the initial stage of inquiry-based ideological and political education teaching to the advanced development of the integration stage of inquiry-based teaching [17]. The in-class links in the integration stage of the smart classroom are mainly based on the strong interaction between teachers and students. The teaching activities in the class mainly include creating a situation, intelligent interactive teaching, independent cooperation and exploration, display and communication, and real-time detection and other processes. The specific teacher activities and student activities are shown in Figure 5.

4.4.3. After Class. The after-school link in the smart classroom integration stage can make up for the lack of personalized training for students. Its main activities include teachers uploading courseware to students and assigning after-school homework. Students take the initiative to complete the tasks assigned by teachers according to their learning conditions. The initiative has been improved, but complete personalized learning has not been achieved, and the feedback of teachers is not immediate, which lacks the depth of the cultivation of students' intellectual ability. Therefore, teachers and students at the initial stage urgently need new teaching methods. In order to improve their abilities in all aspects, the design of inquiry-based ideological and political education and teaching in the integration stage can effectively carry out personalized teaching and learning and improve the personalized teaching and differentiated teaching that cannot be well realized in the integration stage due to technology and other reasons [18].

Therefore, the inquiry-based ideological and political education teaching in the initial stage will develop to the inquiry-based teaching in the integration stage as shown in Table 5.

After-school teaching is mainly based on individualized counseling, targeted, and differentiated teaching, its main activities include individualized teaching and learning, communication and reflection, and homework summary.

4.5. Teaching Evaluation of Wisdom Ideological and Political Education. The ultimate purpose of teaching activities is to achieve teaching goals, and teaching evaluation is also to better achieve teaching goals. The traditional classroom emphasizes the result and ignores the process, but the smart classroom pays more attention to the result and more attention to the process. Therefore, whether it is in the initial stage or the integration stage, it attaches great importance to the evaluation of students' input and the achievement of teaching goals, that is, formative and consequential. However, due to the development of technologies such as big data and learning analysis, the integration stage is more objective and accurate than the teaching evaluation in the initial stage. As well as teachers' evaluation of students, it is more conducive to the development of students' analytical thinking, critical thinking and other higher-order thinking, thereby contributing to the development of students' wisdom. Its advanced performance is shown in Table 6 [19].

5. Analysis of the Application of Inquiry-Based Ideological and Political Education and Teaching in Smart Classrooms

5.1. Statistical Analysis of Questionnaires. In order to facilitate the statistics and analysis of questionnaire data, this paper adopts the form of questionnaire star to distribute and fill in the questionnaire. At the same time, the questionnaire is filled out at the designated time in the ideological and political education class, which can ensure the recovery rate and efficiency. This research requires 55 students and 57 students from two experimental classes in two schools to participate in their own questionnaires. In this survey, 55 and 57 questionnaires were distributed, and 55 and 57 were recovered. The number of valid questionnaires was 54 and 56, respectively. The recovery rate of the two schools' questionnaires was 100%, the effective rate of W school was 98.18%, and the effective rate of Z school was 98.25% [20].

5.2. Analysis of the Process of Inquiry-Based Learning in Ideological and Political Education

5.2.1. Preclass. In the preclass aspect of ideological and political education, the investigation is mainly carried out from two aspects: the students' adaptability to the study guide and whether they can complete the preview independently. The survey results showed that 48.1% of the students agreed with the view that the teacher's preclass study guidance plan was suitable for them. Of course, a very small number of students felt that it was not suitable, accounting for 7.4%; Yes, the proportion reached 66.7%. It can be seen that the inquiry-based ideological and political education and teaching in the initial stage is very effective for students' learning before class. The specific data are shown in Figures 6 and 7 [21].

5.2.2. Inclass. In the initial stage of the smart ideological and political education classroom, the main links are the exchanges and learning between teachers and students and students. For the in-class links, this research mainly investigates the mastery of multimedia functions, the mastery of basic knowledge and skills, and the utilization of classroom teaching resources for ideological and political education. The survey results are shown in Table 7 [22].

It can be seen from Table 7 that most of the students "agree" with the learning in the ideological and political education class. Therefore, it can be seen that the teaching in the initial stage of the smart classroom is practical. Through learning in the multimedia network environment, 55.6% of the students can master the functions of multimedia computers or electronic whiteboards, 53.7% of the students can

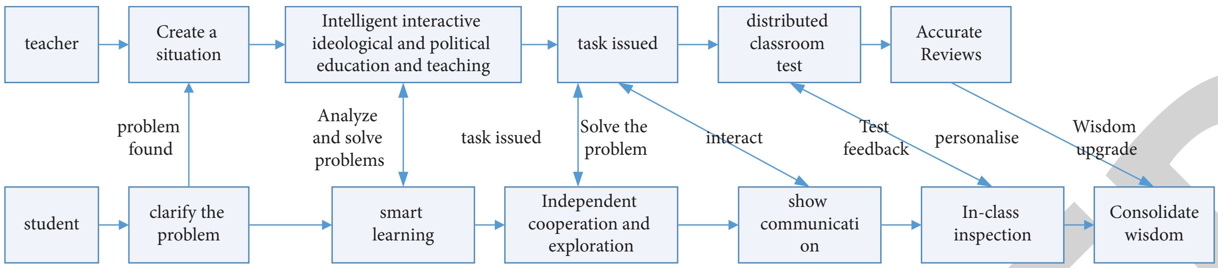


Figure 5: The activity process in the lesson.

Table 5: Comparison of the advanced development of the after-school link from the initial stage to the integration stage.

	"Improved" inquiry-based teaching	"Individual interactive" inquiry teaching
Teacher activity	Upload courseware homework	Personalized push. Personalized coaching. Teaching analysis and reflection
Student activities	Finish homework. Reflection and promotion	Finish homework. Summary reflection

Table 6: Advanced performance in teaching evaluation.

	"Improved" inquiry-based teaching	"Individual interactive" inquiry teaching
Teaching evaluation	Combining formative and summative assessments	Smart and precise teaching evaluation combining online and offline, formative and summative, knowledge skills and smart thinking
Advanced performance: The application of a variety of mixed teaching evaluation, from fuzzy evaluation to precise evaluation		

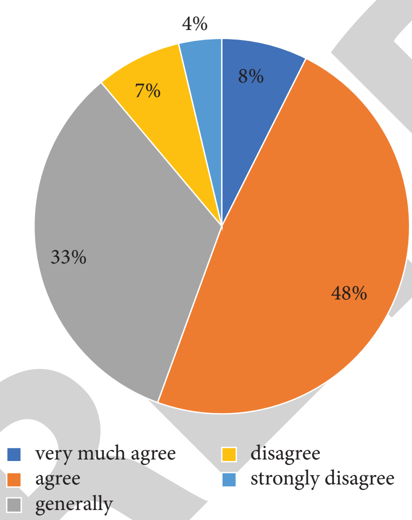


Figure 6: The adaptability of the lesson plan before the lesson.

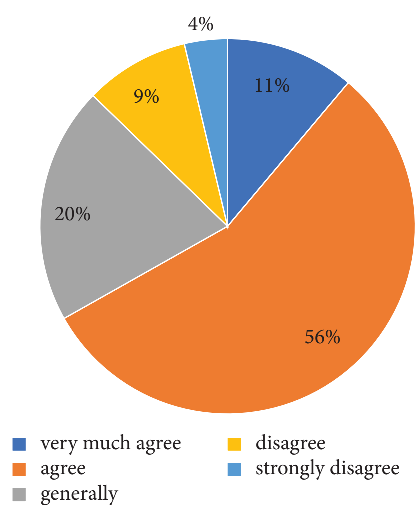


Figure 7: Independent preview.

master the basic knowledge and skills, and in the use of classroom resources, it is also more than the average students can use ideas Political education and teaching resources, but there are also some students who fail to make good use of ideological and political education and teaching resources. In the initial stage of smart classrooms in the multimedia network environment, most students can actively participate in ideological and political education and teaching activities, and in the classroom, most of the students can take the initiative to ask questions, but some students are not sure whether they can take the initiative to ask questions. In addition, in the interaction between teachers and students in the classroom, 14.9% of the students are not very good. However, most students can grasp what they have learned in time. Therefore, the inquiry -based ideological and political education and teaching carried out in the multimedia network environment is of positive significance [23].

5.2.3. After Class. The after-school link of ideological and political education mainly investigates the completion of tasks and the consolidation and improvement of students'

Table 7: Sections in the lesson.

	Very much agree	Agree	General	Disagree	Strongly disagree
Mastering the functions of multimedia computers	9 (16.7%)	30 (55.6%)	15 (27.8%)	0 (0%)	0 (0%)
Mastery of basic knowledge and skills	8 (14.8%)	29 (53.7%)	16 (29.6%)	1 (1.9%)	0 (0%)
Use of classroom resources	4 (7.4%)	27 (50%)	18 (33.3%)	3 (5.6%)	2 (3.7%)
Active participation in teaching activities	5 (9.3%)	28 (51.9%)	18 (33.3%)	2 (3.7%)	1 (1.9%)
Teacher-student interaction	4 (7.4%)	26 (48.1%)	16 (29.6%)	5 (9.3%)	3 (5.6%)
Proactive questioning	5 (9.3%)	27 (50%)	19 (35.2%)	2 (3.7%)	1 (1.9%)
Summary of the content of this lesson	4 (7.4%)	29 (53.7%)	17 (31.5%)	2 (3.7%)	2 (3.7%)

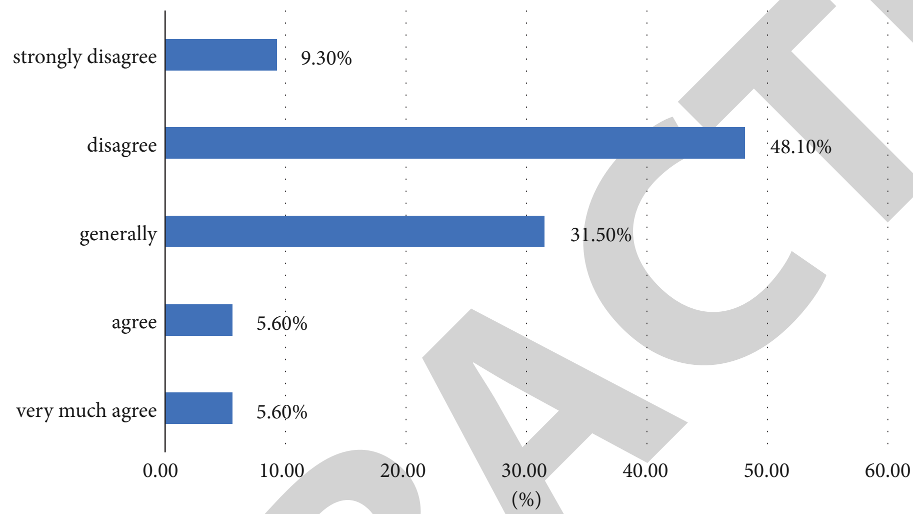


Figure 8: Active completion of tasks.

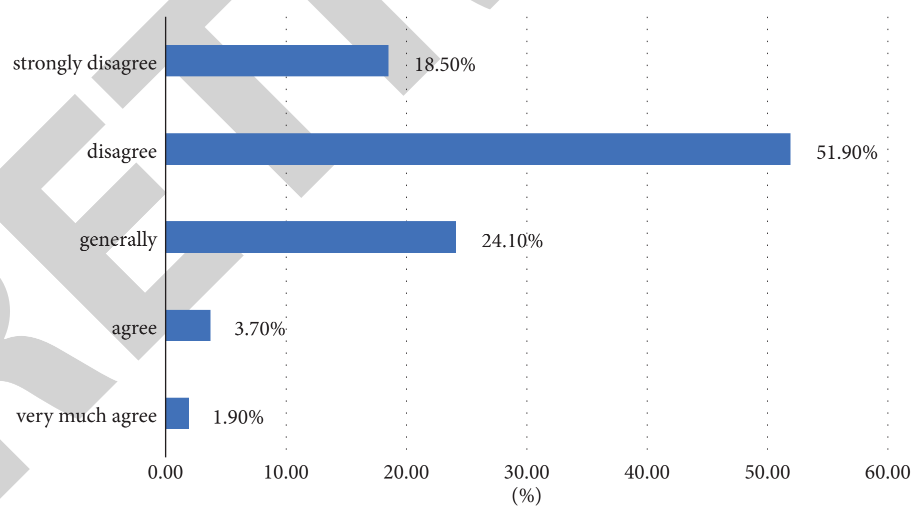


Figure 9: Self-consolidation and improvement.

autonomy. The specific situation is shown in Figures 8 and 9. 48.1% of students can complete the tasks assigned by teachers, and 9.3% of them are very good, but 11.2% of them cannot complete them well, but the overall degree of completion is very high; for students' ideological and political education courses, after the improvement of self-consolidation, more than 70% of students can achieve self-consolidation. It can improve yourself well. Therefore, the inquiry -based ideological and political education and teaching in the initial stage can well promote the improvement of students' knowledge and ability.

5.3. Investigation and Analysis on the Learning Effect of Ideological and Political Education. Taking W school as an example, the learning effect of W school students mainly

Table 8: The results of the survey on learning effects.

	Very much agree	Agree	General	Disagree	Strongly disagree
Mastery of basic knowledge	9 (16.7%)	23 (42.6%)	16 (29.6%)	6 (11.1%)	0 (0%)
Increased interest in learning	8 (14.8%)	22 (40.7%)	17 (31.5%)	7 (13.0%)	0 (0%)
Improve learning enthusiasm and initiative	9 (16.7%)	24 (44.4%)	15 (27.8%)	4 (7.4%)	2 (3.7%)
Increased awareness of the problem	4 (7.4%)	21 (38.9%)	19 (35.2%)	6 (11.1%)	4 (7.4%)
The improvement of imagination	6 (11.1%)	25 (46.3%)	18 (33.3%)	3 (5.6%)	2 (3.7%)
The improvement of independent cooperation and exploration ability	6 (11.1%)	23 (42.6%)	19 (35.2%)	4 (7.4%)	2 (3.7%)

includes the mastery of basic knowledge of ideological and political education, the improvement of learning interest, the improvement of learning enthusiasm, the improvement of initiative, the ability to ask questions, the improvement of imagination, and the improvement of independent cooperation and exploration. Several aspects of capacity improvement were investigated. Compared with the students' initial level, the students' abilities in all aspects after the experimental teaching tend to "agree." Among them, more than 59.3% of the students agree that the use of the inquiry -based ideological and political education teaching method has improved their mastery of basic knowledge. Similarly, in terms of enthusiasm and initiative in learning, more than 70% of the students believe that the enthusiasm and initiative in learning have been improved; in addition, in terms of the improvement of problem awareness and imagination, 35.2% and 33.3% believe that the effect is average, but large. Most of the students think that they have improved a lot, and they have improved their independent cooperation and exploration ability. Therefore, through this teaching method, students' learning effect has been significantly changed, their enthusiasm for learning has been improved, students' awareness of problems can be stimulated, and students' innovative thinking ability has been improved. The results are shown in Table 8.

6. Conclusion

As a hotspot of educational informatization research in China, smart classrooms in ideological and political education have received widespread attention. It implements the theory into the classroom teaching of ideological and political education, focusing on the teaching activities between teachers and students. Information technology and intelligent technology are important technical components of smart classrooms in ideological and political education. With their rapid development, smart classrooms show staged characteristics. Under these staged characteristics, each stage is brought about. Different inquiry-based ideological and political education teaching: this research aims to explore the development of inquiry-based teaching in smart classrooms of ideological and political education at different stages from the perspective of theory and practice and to conduct advanced design research on inquiry-based teaching in smart classrooms of ideological and political education, with a view to promote the development of students' intelligence. After reading a large number of relevant literature studies on smart classroom, smart classroom teaching, inquiry-based teaching, etc., this research conducted a necessity analysis of the advanced design of inquiry-based ideological and political education teaching in smarter classrooms and initially designed the inquiry-based ideological and political education. Based on this, the advanced design of inquiry-based teaching is carried out. Finally, in order to test the effectiveness of the advanced design of inquiry-based ideological and political education teaching in smart classrooms, students from W school were selected as subjects to conduct practical tests. The results show that the advanced design of inquiry-based ideological and political education teaching in smart classrooms is possible. This paper conducts in-depth research on the design of inquiry-based ideological and political education teaching in smart classrooms, which can provide some reference for schools, but the design needs to be further improved. Due to the limitations of various conditions, the inquiry-based teaching design of the smart classroom stage is only prospected in theory, and no further in-depth design is carried out. Therefore, it needs to be further improved in the future.

Data Availability

The dataset can be accessed upon request to the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] M. J. L. A. I. Timms, "Letting artificial intelligence in education out of the box: educational cobots and smart classrooms," *International Journal of Artificial Intelligence in Education*, vol. 26, no. 2, pp. 701–712, 2016.
[2] A. K. D. Jena, "Smart classroom an effective technology for teaching," *A ResearchAnalysis[J]Journal Of Educational Technology*, vol. 10, no. 1, pp. 55–64, 2016.
[3] W. G. C. W. Kumara, K. Wattanachote, B. Battulga, T. K. Shih, and W. Y. Hwang, "WA kinect-based assessment system for smart classroom," *International Journal of Distance Education Technologies*, vol. 13, no. 2, pp. 34–53, 2015.
[4] J. Lee, H. Lee, and Y. Park, "The smart classroom: combining smart technologies with advanced pedagogies," *Educational Technology*, vol. 53, no. 3, pp. 3–12, 2013.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 3537468, 10 pages
https://doi.org/10.1155/2022/3537468

Research Article

Challenges and Strategies of Student Management in Universities in the Context of Big Data

Yongzhi Zhang

School of Marketing Management, Liaoning Technology University, Huludao 125105, China

Correspondence should be addressed to Yongzhi Zhang; zhangyongzhi@lntu.edu.cn

Received 22 August 2022; Revised 5 September 2022; Accepted 12 September 2022; Published 28 September 2022

Academic Editor: R. Mo

As the most active and knowledge-intensive university, the forefront of the use of network information technology, the education, management, and service mode of the school, and the ideology, learning methods, and behavioural habits of teachers and students will be profoundly affected by the era of big data. The current form of student management in universities is too old and needs to be replaced by a new management system. With the promotion of the construction of smart campus, the rapid development of big data technology has realised the innovation of student management in colleges and universities. This paper takes students of Wuhan University as the research object. By collecting various application data in the campus information system, the K-means algorithm in cluster analysis is used to classify students' campus behavioural characteristics, and then the Apriori algorithm is used to correlate students' behavioural characteristics with their academic performance. The experimental results show that there is a close relationship between the consumption behaviour, work and rest behaviour, study behaviour characteristics of different student groups, and their academic performance. Using the results of these analyses, universities can adopt differentiated management measures for different categories of students, which can help improve students' academic performance as well as further enhance the efficiency of student management.

1. Introduction

With the abundance of Internet applications, every student and every teacher are producing a variety of data every day. When the amount of data reaches a certain level, it can be analysed and mined in a meaningful way, and the analysis of massive amounts of data will have a profound impact and influence on education information technology. At the same time, schools have more and more live data at their disposal, which provides new directions for research to find the reasons behind the patterns. For example, the campus network platform used by students to communicate generates a huge amount of data consisting of photos, videos, conversations, messages, emails, etc. every day, and these datasets reflect students' thinking, emotional direction, and behavioural dynamics and contain rich information and regular content. How to scientifically store, manage, and later effectively analyse and utilise huge amounts of data is of great importance in student management in universities.

Big data has become the microscope through which people can analyse their thoughts and observe their behaviour, as it goes beyond the description of the nature of things and can be mined and analysed to concretely quantify all aspects of people's thoughts and behaviour. Similarly, student management in universities needs not only experience to guide it but also science to lead it. In this respect, if student management in universities can use data more effectively to analyse problems and make decisions, the quality of work will be greatly improved. With the rapid development of the mobile Internet, students are already making greater use of information networks in their activities. For example, video filming and audio and photographic recordings are used extensively for webcasting while being published in conjunction with traditional formats. The links between these heterogeneous sources of information data, obtained in different forms, are so complex that they are difficult to count and describe in a traditional, simple way. Therefore, the scientific and rational analysis, collation and aggregation of information and data, and presentation of the results in an intuitive and decision-friendly form have become a major issue for student management in universities.

Student management often makes use of instant messaging devices such as mobile phones to gather information about current events on campus and to understand the dynamics of student thinking, which also generates a large amount of data flow, which reflects the data flowing through the system and shows the characteristics of dynamic data. The best way to understand students' needs and feedback is to tap into the information data generated by students and analyse their behaviour. The best way to understand students' needs and feedback is to use student data to analyse their behaviour, but one of the problems is to sort out the data and store it in a scientific way. At the same time, the security of information data, which is closely related to the physical and mental safety of students, is also an aspect that must be taken into account in order to ensure that information about students is not leaked and is not used maliciously by other institutions and people.

In the process of the continuous evolution of big data technology, the construction of university intelligent campus is in the midst of a profound change towards the Internet and big data, while the student management problems, also under the influence of Internet+, are increasingly highlighted. The ubiquitous mobile network makes students obtain various information through the Internet all the time and at the same time inevitably makes students' studies negatively affected to different degrees. For example, students do not develop good study habits and habits of life, using their mobile phones to watch TV series, play games, and brush microblogs and jitterbug during classes, leading to low efficiency in classroom learning, lack of interest in learning new technologies and knowledge, insufficient motivation to learn, and a preference for speculation and taking refuge in easy choices [1, 2].

There are three major problems with university education in the new information technology environment. First, students are not sufficiently motivated to learn subjectively. In the relaxed learning atmosphere of university, due to the lack of external pressure and effective supervision, many students often miss classes, leave early, and even desert class, a situation mostly seen in freshmen students, so it is also called new sickness. In the long run, the lack of energy input will lead to academic wastage. Second, students study inappropriately. Many students still maintain the study habits of their high school days after entering university and do not adapt to new teaching modes, such as MOOC and flipped classroom, resulting in half the effort and half the result, resulting in their academic performance being unsatisfactory. Third, students' learning disabilities have not been corrected. This is due to the teachers' teaching methods, as well as the students' personal subjective lack of effort. Of the three types of situation: not wanting to learn, not knowing how to learn, and wanting to learn but not being able to learn, the first two account for the vast majority of cases [3]. In response to these situations, universities urgently need to innovate the management methods of university students and promote refined and differentiated management methods. Big data technology provides new ideas for personalized management and intelligent decision making.

Chinese scholars have begun to study the application of big data technology in the education sector and the impact and changes it has brought [4], with research findings focusing on the use of big data technology to analyse student behaviour [5–7] and transform management thinking and models [8–10]. However, few research results have been published on the use of big data technology to fine-tune and differentiate the management of college students. In this paper, from the perspective of college student management, we use the data from the digital platform of education to explore the hidden information value and provide new ideas and methods for college student management.

At present, most of the student management workers in universities are able to adjust their working concepts, methods, and approaches in a timely manner and make full use of the convenient conditions brought by network information technology to improve the efficiency and quality of their work. However, in the era of big data, the existing technical means in universities can only analyse a small number of information data with similar categories and structured shapes and cannot yet collect, store, analyse, and visualise the results of big data in the modern sense, and data analysis techniques are not yet widely used and not familiar to most staff. Moreover, it is difficult to invest significant effort and financial resources in the current student management in universities to use such techniques to analyse the individual needs of students in terms of learning and development. At the same time, professional data mining, integration, and analysis skills are scarce. The value of big data is undeniable, but the limitations in the use of analytical talent and technology have made it difficult to fully realise the value of data for the time being.

The advent of the big data era is undeniable, and data as an important asset are changing the decision-making model of governments and enterprises. Student management in higher education must face the challenges of this change, change its thinking, and use all possible methods to fully exploit the value of big data and provide a variety of support for its work. The era of big data requires more complex talents with data technology skills. To fully utilise the role of big data technology in student management, teams with big data analysis skills need to be formed. Introduce more specific technical training in the team, master theories, methods, and tools for data analysis on big data platforms, build a cloud-based student management system, etc. Through real-time tracking and analysis of internal campus communication forums, it can keep abreast of and grasp the dynamics of students' thoughts, actively guide public opinion on hot issues and emergencies that students discuss and care about, and maintain campus security and stability.

From the perspective of the current team structure of student management in China's colleges and universities, it is difficult to enrich in a short period of time a teaching force with both a background of knowledge of big data technology and the laws of student management. Therefore, in order to realise the challenges and innovations of student management in universities under the background of big data, this paper

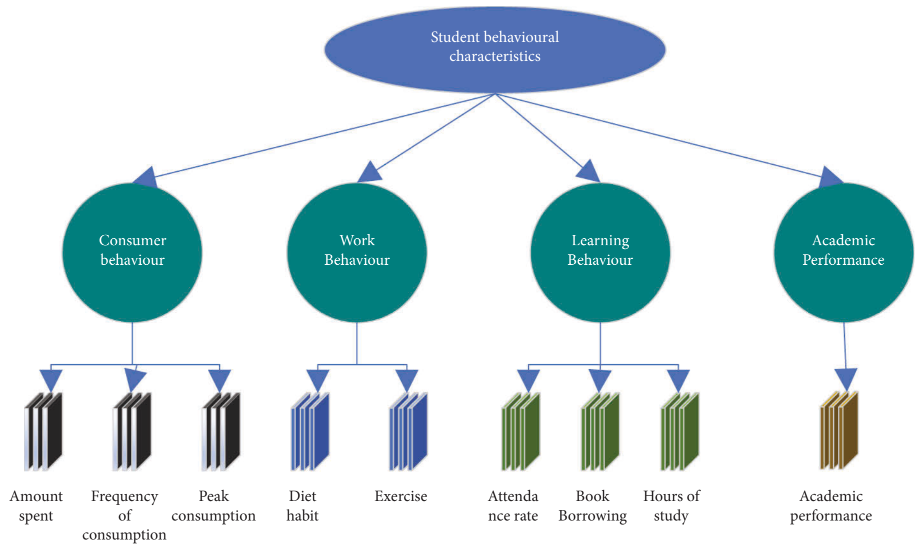


FIGURE 1: Student behavioural characteristics and student achievement research variables.

collects various application data in the campus information system, classifies students' campus behavioural characteristics by using the K-means algorithm in cluster analysis, and then uses the Apriori algorithm to correlate student behavioural characteristics with academic performance. The experimental results show that there is a close relationship between the consumption behaviour, work and rest behaviour, study behaviour characteristics of different student groups, and their academic performance. Using the results of these analyses, universities can adopt differentiated management measures for different categories of students, which can help to improve students' academic performance as well as further enhance the efficiency of student management.

2. The Application of Big Data in Student Management in Universities

In the management of students in colleges and universities, scientific and reasonable classification of students' behaviour in school, the formulation of corresponding management service methods for different types of students, and the provision of personalized support measures can improve the refinement of student management and teaching services and enhance the quality level of talent cultivation. In the era of big data on the Internet, students' one card consumption data, classroom card attendance data, book borrowing data, and Internet access data can dynamically and accurately map out students' behavioural characteristics and the behavioural habits hidden behind these data.

2.1. Experimental Data. In this paper, undergraduate students of Wuhan University Class 2021 were taken as the research object, and the behavioural data of four grades of undergraduate students for one academic year were selected, and the research variables of student behavioural characteristics and student performance were set, and the specific relationship is shown in Figure 1. Among them, the consumption behaviour data mainly came from the One Card system, and the work and rest behaviour data mainly came from the Sunshine Attendance System, the One Card system, and the Internet authentication system. The learning behaviour data mainly came from the teaching attendance system, the book lending system, and the Internet authentication system. The raw data extracted include 8426392 consumption flows from the One Card, 741904 attendance flows from the Sunshine Attendance System, and 929088 book borrowing flows. Due to the redundancy and structural inconsistency in the data, the student samples with serious missing information and some outliers were eliminated after data processing, and the final sample data of 12904 students were obtained.

Specific indicator codes for the student behavioural characteristics and student achievement research variables are given in Table 1.

TABLE 1: Specific information on student behavioural characteristics and student achievement research variables.

Category	Indicator name	Indicator codes
Consumer behaviour	Amount spent	x_1
	Frequency of consumption	x_2
	Peak consumption	x_3
Work behaviour	Diet habit	x_4
	Exercise	x_5
Learning behaviour	Attendance rate	x_6
	Book borrowing	x_7
	Hours of study	x_8
Academic performance	Academic performance	x_9

2.2. Clustering Analysis of Student Behaviour Using the K-Means Algorithm. The K-means algorithm was proposed by the academic Mac Queen in 1967, and it is one of the classical algorithms of clustering algorithms [11–13]. The core idea of the K-means algorithm is to find out K cluster centres $m_1, m_2, \cdots, m_k$, then make each data $x_i\,(i = 1, 2, \cdots, n)$, through continuous iteration, as close as possible to the cluster centre $m_j\,(j = 1, 2, \cdots, k)$ where it is located, and make the data not in the same cluster as far away from each other as possible. The calculation of cluster centres varies depending on the clustering algorithm, and the K-means algorithm determines the cluster centres based on the average of all data within a cluster.

Since the K-means algorithm always converges, K-means can always reach a steady state in a finite number of steps, i.e., the clustering centres will not change again [14–16]. Since changes in the clustering centres often occur in the course of the previous iterations, in order to optimise the time complexity of the algorithm, the iterative process can usually be stopped and the results can be output directly when only more than 99% of the data points belong to clusters that no longer change. This approach is effective in reducing the time complexity of K-means when dealing with larger data [17, 18]. In general, we can use the inverse of the spatial distance as the expression for calculating similarity, as shown in the following equation:

$$Sim\,(c, x) = \frac{1}{Dist_{L2}\,(c, x)}, \tag{1}$$

where c represents the cluster centre, x represents a particular data point, and $Dist_{L2}$ represents the Euclidean distance. The expression for calculating the Euclidean distance on an N-dimensional continuous space is given in (2), where x and y are two points in the N-dimensional space, and x_i represents the value of point x in the i-dimensional coordinates.

$$Dist_{L2}\,(x, y) = \sqrt{\sum_{N}^{1} (x_i - y_i)^2}. \tag{2}$$

In general, the cluster centre is the mean or median of all the data in the cluster, i.e., (3), where c_i is the cluster centre of cluster C_i and m_i is the number of data in cluster C_i.

$$c_i = \frac{1}{m_i} \sum_{x \in C_i} x. \tag{3}$$

For two different clustering results obtained by K-means on the same dataset, the superiority of the results can be judged by comparing their SSE, with a smaller SSE indicating a more desirable result. This is because a smaller SSE indicates a smaller sum of errors for all the data, which means that the clustering centre can better represent all the data in its clusters.

$$\text{SSE} = \sum_{K}^{i=1} \sum_{x \in C_i} di\,st\,(c_i, x)^2. \tag{4}$$

Clustering analysis is descriptive in the context of data mining tasks and is characterised by the fact that the output of the algorithm cannot be described in terms of correctness or incorrectness, i.e., there is no unique solution. Although there is no unique solution for this type of task, the evaluation of its results is important.

2.3. Analysis of the Association between Student Behaviour and Academic Performance at School Using the Apriori Algorithm. The Apriori algorithm is an algorithm for mining frequent item sets for Boolean association rules [19–21]. Apriori is a Boolean association rule frequent itemset mining algorithm that is part of unsupervised learning in machine learning [22, 23]. It is necessary to set a minimum confidence level to constrain the confidence level of association rules. The core of the Apriori algorithm is a two-stage recursive mining process, where the first stage is to find all combinations of factors with high frequency from the dataset, and the second stage is to find association rules that satisfy the requirements from all frequent items in the set. The formulae for support [24] and confidence [25] are as follows:

$$\begin{aligned} \text{support}\,(A \longrightarrow B) &= P\,(A \cup B), \\ \text{confi de nt}\,(A \longrightarrow B) &= \frac{P\,(A \cup B)}{P\,(A)}, \end{aligned} \tag{5}$$

where support indicates the probability of event A and event B occurring simultaneously and confidence indicates the probability of event B occurring simultaneously given the occurrence of event A. The results of the two confidence levels are shown in the following equation:

$$\begin{aligned} \text{Confident}\,(A \longrightarrow C) &= \frac{P\,(A \cup C)}{P\,(A)} = 1, \\ \text{Confident}\,(C \longrightarrow A) &= \frac{P\,(A \cup C)}{P\,(C)} = \frac{2}{3}. \end{aligned} \tag{6}$$

3. *K*-Means Clustering Analysis

Before the data are clustered and analysed, they are normalised and dimensionless. Cluster analysis was then performed to create a new data stream and the data were then processed using the K-means algorithm.

3.1. Analysis of Student Behaviour

3.1.1. Clustering Analysis of Student Consumption Behaviour. According to the evaluation criteria of the clustering algorithm, the best clustering effect was obtained when the number of clusters was set to 5. The clustered student classification results were analysed by comparing the mean of each cluster with the mean of the overall student indicator according to the actual situation of student consumption and noting H as the mean above the overall student indicator and L as the mean below the overall student indicator. The results of the cluster analysis of student consumption behaviour for each cluster are shown in Table 2.

The average monthly spending and the corresponding percentage of students in this category are shown in Figure 2.

Table 2: Cluster analysis results of student consumption behaviour.

Type of consumption	Number of students	Monthly consumption	Monthly consumption frequency	Monthly consumption peak	Comparison results
Type 1	1367	458.65	104.5	542.23	LHL
Type 2	2243	685.23	86.5	1125.32	LLH
Type 3	4363	752.35	65.2	856.27	HHL
Type 4	3168	845.24	95.8	1201.98	HHH
Type 5	1763	486.74	45.3	689.45	LLL

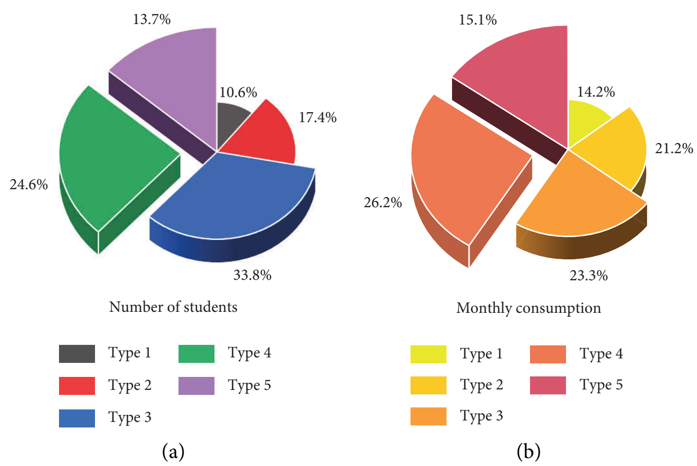


Figure 2: Average monthly consumption and corresponding number of students. (a) Number of students. (b) Average monthly consumption.

The data on the average monthly consumption frequency and the average monthly consumption peak are shown in Figure 3.

Figures 2 and 3 show that the characteristics of student consumption behaviour fall into five categories. The first group of students has the lowest monthly consumption level and the lowest monthly consumption peak but consumes frequently and belongs to the group with a low consumption level. The second group of students has a medium monthly consumption level but has a high monthly consumption peak and spends less frequently. The third group of students has a medium to high monthly consumption level, with a high monthly consumption frequency and a medium monthly consumption peak. The fourth group of students has the highest monthly consumption level, with frequent monthly consumption and the highest monthly consumption peak. The fifth group of students had a low monthly consumption level, the least frequent monthly consumption, and a low monthly consumption peak.

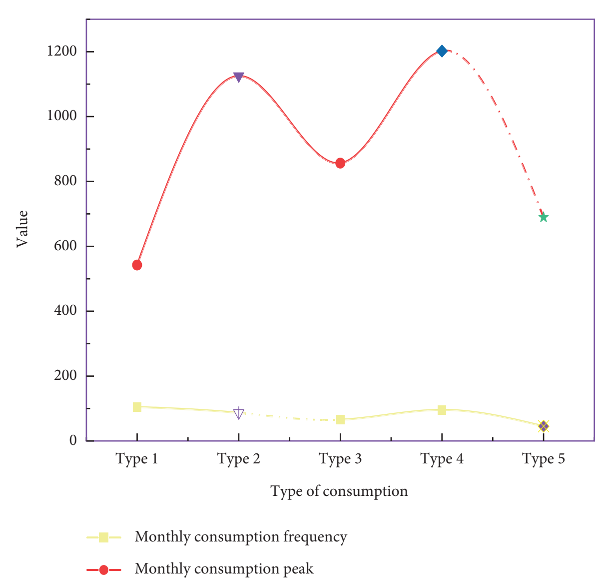


Figure 3: Average monthly consumption frequency and average monthly consumption peak.

3.1.2. Clustering Analysis of Students' Work and Rest Behaviour. The *K*-means algorithm cluster analysis was conducted on the eating and drinking habits, Internet habits, early waking habits, and physical exercise habits among the indicators of students' work and rest behaviour. According to the clustering average criterion, when the number of clusters is 3, the percentage of students in each type of cluster and the average value of the indicators are shown in Table 3.

Table 3 shows that the first group of students woke up early more often each month, ate more regularly in the school canteen, spent more time online, and participated in physical activity more often. The second group of students woke up more often each month, ate irregularly in the school canteen, spent the longest time online, and participated in very little physical exercise. The third group of students woke up early more often each month, but ate regularly in the

Table 3: Cluster analysis results of students' work and rest behaviour.

Type of work/ rest	Number of students	Number of regular meals	Number of early years	Hours on the Internet	Number of exercise sessions	Compare results
1	4720	23.6	20.3	175.24	53.2	HHLH
2	1326	9.3	6.9	210.31	35.2	LLHL
3	6858	13.5	19.1	206.79	28.6	LHHL

school canteen, spent more time online, and participated in physical activity less often.

A comparison of the data on the number of times students ate regularly and woke up early is shown in Figure 4.

A comparison of the data information on the number of hours spent online and the number of physical activity sessions for these three groups of students is shown in Figure 5.

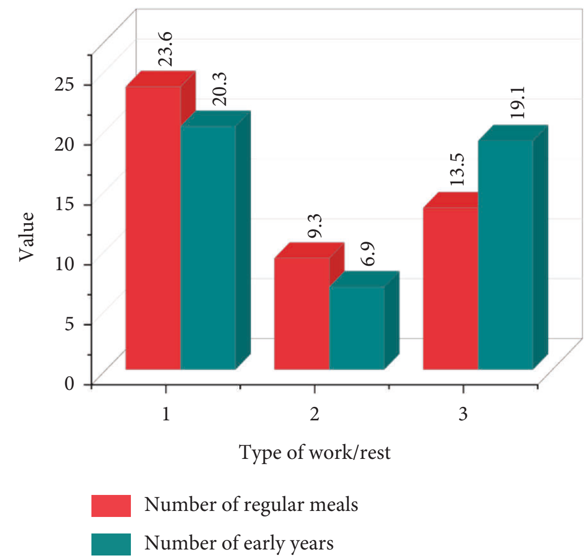


Figure 4: Comparison of data on the number of times students ate regularly and woke up early.

Figure 5: Comparison of the number of hours spent on the Internet and physical exercise among the three categories of students.

3.1.3. Clustering Analysis of Students' Learning Behaviour. *K*-means algorithm cluster analysis was conducted on four indicators of students' learning behaviour: class attendance, library borrowing, number of visits to the library, and length of study. The percentage of students in each category of clusters and the mean values of the indicators are shown in Table 4.

Table 4 shows that students in category 1 had the highest class attendance, borrowed fewer books, visited the library the most, and spent the most time studying. Category 2 students had the lowest class attendance, borrowed the fewest books, visited the library the least, and spent the shortest amount of time studying. Category 3 students had the highest attendance rate, borrowed the most books, visited the library more often, and studied for longer periods of time. The fourth group of students had an average attendance rate, borrowed fewer books, visited the library more often, and spent less time studying. The comparisons are shown in Figure 6.

3.2. Student Group Characteristics and Management Suggestions. Through the cluster analysis of the above three categories of students' different behavioural characteristics, we believe that the behavioural characteristics of different groups of students can be used as the basis for implementing personalized management for students in the management of university students and put forward suggestions for implementing personalized management, which are summarized in Tables 5–7. The following are some of the recommendations. Firstly, in view of the wide variation in student spending in the school canteen, it is important to pay particular attention to students who spend too much and too little and to develop a policy of poverty assistance for students in financial difficulty. Secondly, in view of the fact that students generally spend a lot of time on the Internet during the school year, administrators should have a clear understanding that, in the Internet era, although it is more convenient to obtain information from the Internet than from other means and students are becoming more and more dependent on the Internet, a reasonable code of conduct for students on the Internet should be formulated and the management and supervision of students on the Internet should be strengthened in view of students' health. Thirdly, in response to the situation that students do not invest enough in their studies, visit the library less often, and borrow fewer books, comprehensive measures to motivate university students to study should be formulated, such as setting up scholarships for excellence and launching reading day activities.

Table 4: Results of the cluster analysis of students' learning behaviour.

Type of study	Number of students	Attendance in class	Hours of study	Number of trips to the library	Book borrowing	Compare results
1	4202	95.63	153.6	29.3	7.6	HLHH
2	1534	86.26	84.3	11.5	3.4	LLLL
3	1588	92.37	132.4	26.1	15.8	HHHH
4	5580	90.12	116.5	16.7	8.3	LLLL

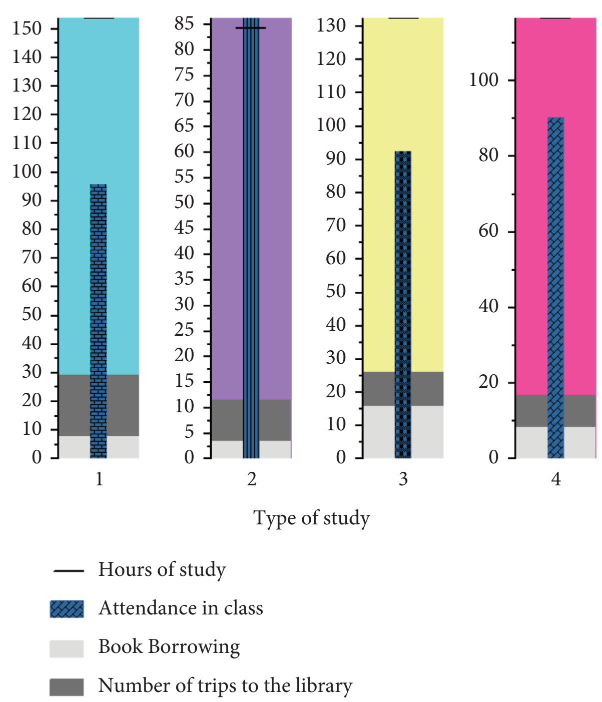


Figure 6: Comparison of data from the four categories of student learning behaviour cluster analysis.

4. Apriori Algorithm Analysis

4.1. Analysis of the Association between Students' School Behaviour and Academic Performance. The above clustering analysis has classified students' school behaviour into three categories: consumption behaviour, work and rest behaviour, and study behaviour. In order to further study the relationship between students' behavioural characteristics and their academic performance and to find out whether there is an inevitable connection between students' behavioural characteristics and their academic performance, the Apriori algorithm was chosen to conduct the correlation analysis and to explore the hidden correlations and patterns from the big data.

Before performing the Apriori algorithm correlation analysis, the student recruitment indicator data and the academic achievement indicator data were used as data sources. A new data stream is created in SPSS Modeler software, the Apriori algorithm model is constructed, and the relevant parameters are set. In the Apriori algorithm model, the five types of student fee behaviour, the three types of resting behaviour, the four types of learning behaviour, and the three types of student academic achievement were set as the antecedent and postecedent variables of the association rules. A support level of 10% and a confidence level of 80% were set for the association rule analysis, and a total of 24 association rules were obtained. In accordance with the objectives of this study, the association rules with an elevation greater than 1 were selected by eliminating and merging redundant rules, with the post-item being the association rule for student academic achievement, as shown in Table 8.

As can be seen from Table 8, the first rule reflects that 10.265% of the student population is characterised by not waking up early, eating irregularly, spending a lot of time online, and exercising less often, and 82.423% of these students are likely to be academically low achievers. The rule support was 8.461%, indicating that 8.461% of students in this type of behaviour were type 2 and had poor academic performance. The second rule reflects that 11.841% of the student population is characterised by low monthly consumption, low peak consumption, low consumption, and often waking up early, eating irregularly, spending more time online, and exercising less often, and that 83.246% of students in this category are likely to have average academic performance. The rule support of 9.857% indicates that 9.857% of students in this category have a consumption behaviour type of 5, a resting behaviour type of 3, and average academic performance. The third rule reflects that 53.156% of students are characterised as early risers, irregular eaters, spending more time on the Internet, and exercising less often, and 84.068% of students in this group are likely to be average achievers. The rule support of 44.687% indicates that 44.687% of students in this type of behaviour are of type 3 and have average academic performance.

The fourth rule reflects that 23.683% of students with average attendance, infrequent visits to the library, low book borrowing, short study time, early rising, regular eating, long Internet access, and regular exercise were 80.366% likely to be of average academic achievement. In contrast, the rule support of 19.033% indicates that 19.033% of students in this category have a study behaviour type of 4, a rest and relaxation behaviour type of 1, and have average academic performance. The fifth rule reflects that 10.308% of the student population is characterised by waking up early, eating more regularly, spending more time online, exercising regularly and being present in class more often, going to the library more often, borrowing more books, and studying for longer periods of time, and that 81.343% of students in this category are likely to be doing well academically. In contrast, the rule support of 8.385% indicates that 8.385% of students in this category have a

Table 5: Management characteristics of the consumption behaviour of student groups.

Types of consumer behaviour	Group characteristics
1	Low monthly consumption and frequent consumption
2	Medium monthly consumption and less frequent consumption
3	Medium monthly consumption and frequency of consumption
4	High monthly spending and frequent spending
5	Low monthly consumption and less frequent consumption

Table 6: Characteristics of student groups in managing their work and behaviour.

Types of work and rest behaviour	Group characteristics
1	Waking up early, eating more regularly, and exercising regularly
2	Not getting up early, not eating regularly, and not exercising as much
3	Frequent early risers, irregular diet, and low exercise frequency

Table 7: Characteristics of student groups for managing learning behaviour.

Types of learning behaviour	Group characteristics
1	High attendance, longer study time, and fewer books on loan
2	Low attendance, short study time, and few books on loan
3	Higher attendance, longer study time, and more books on loan
4	Average attendance, less study time, and fewer books on loan

Table 8: Results of association rule analysis.

Back items	Previous items	Rule support (%)	Degree of support (%)	Confidence (%)
Poor study performance	Type of work and rest behaviour = 2	8.461	10.265	82.423
Average study performance	Consumption behaviour type = 5	9.857	11.841	83.246
Average study performance	Type of work and rest behaviour = 3	44.67	53.156	84.068
Average study performance	Learning behaviour type = 4	19.033	23.683	80.366
Good study performance	Type of resting behaviour = 1	8.385	10.308	81.343

type 1 resting behaviour and a type 3 study behaviour and are doing well academically. A detailed comparison of the data from the analysis of the association rules is shown in Figure 7.

The use of big data technology to classify students into different behavioural characteristics can facilitate administrators to carry out targeted interventions and counselling for different types of student groups, which is conducive to personalized management of university students and further improves the efficiency of student management in universities; it can remind administrators to focus especially on those groups of students whose academic performance is poor due to high probability events and develop corresponding student management systems. The study also found that early morning exercise is a good way to improve academic performance. The study also found that regular early exercise, regular meals, and book borrowing contributed to improved academic performance. On the contrary, students who lacked good work and rest behaviour were more likely to have poor academic performance.

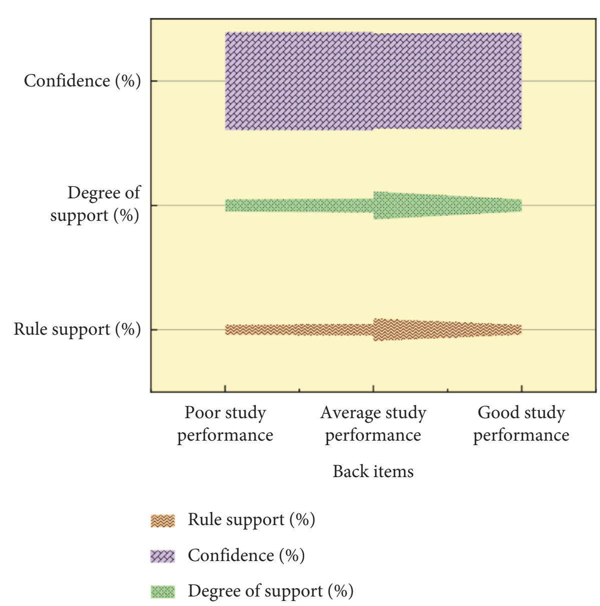


Figure 7: Comparison of association rule analysis results.

4.2. Innovation of Big Data Technology in the Management of University Students. Through the K-means clustering analysis and Apriori correlation analysis, the clustering results of students' school behaviours and the correlation rules between students' school behaviours and their academic performance were obtained. This not only correlates students' daily school behaviours with their academic performance but also distinguishes different groups of students with different characteristics. Based on the results of clustering analysis of student behaviour data, universities can develop differentiated management measures for different types of student groups to further improve student management efficiency and contribute to the innovation of student management in universities.

(1) The use of big data technology can transform the student management mode from unified management to personalized management.

In the context of the current high-quality development of higher education, the traditional management mode of college students can no longer meet the needs of the times. In the background of information technology, the smart campus construction of universities is changing day by day, the ways for students to obtain knowledge and information have become diversified, and they receive more and more information from the outside world every day. Through big data technology, universities can grasp the behavioural habits of students in daily consumption, work and rest, study, etc., understand the personalized needs of different student groups, and formulate corresponding management measures, so that the student management mode can change from uniform management to personalized management. This will change the management mode of students from unified management to personalized management.

(2) The use of big data technology can transform student management from passive management to active management.

As the cradle of talent training, universities need to keep pace with the development of society. The university's one card system, digital campus system, student teaching system, book lending system, sunshine attendance system, and other applications provide rich data resources for school management. By mining these data, we can obtain information on the learning and work habits of students across the university, thus changing the way student management works from trying to solve problems after they arise to taking the initiative to identify problems and solve potential problems immediately or set up preventions in advance. For example, in student management, individual students have abnormal behavioural characteristics but are not detected, often causing consequences before they are discovered, while the use of big data prediction type algorithm can predict and detect abnormal behaviour of students in time and arrange early intervention by tutors, thus transforming student management from passive management to active management.

(3) The use of big data technology can make the evaluation of student and student management performance shift from a single to a diversified approach.

At present, the evaluation of student and student management performance is mainly focused on academic performance ranking, which is one-sided in some ways. For one thing, the design of evaluation methods and indicators is not very reasonable; for another, there is a serious information mismatch between evaluators and students, and it is impossible to grasp students' daily performance and effort level in school systematically through effective technical means. The use of big data technology can collect data on the behaviour of university students in their school life and study and can analyse the behaviour of students in classes, grades, and even the whole university, so as to comprehensively and systematically judge the state of students' school life and study as well as the effectiveness of student management work, thus shifting the evaluation of students and student management work performance from singularity to diversity.

5. Conclusion

In view of the current team structure of student management in universities, it is difficult to fill up a faculty with the knowledge background of big data technology and the laws of student management in a short period of time, so it is necessary to give full play to the role of big data technology in student management. In this paper, we selected various types of behavioural data from the campus information system of four undergraduate students of Wuhan University Class 2021 for one academic year and used cluster analysis to classify students' behavioural characteristics in school and also to analyse the association between students' behavioural characteristics in school and their academic performance. The results of the experiment found that there is a close relationship between the consumption behaviour, work and rest behaviour, study behaviour, and academic performance of different student groups, which provides a basis for schools to adopt differentiated management measures for different types of students, and on this basis, suggestions are given for improving student management methods by using big data technology in the context of smart campus construction. Universities should make full use of the achievements of education informatization construction and process the large amount of data information obtained through big data technology and informatization means, so as to provide rich data support for school decision making and development.

Data Availability

The datasets used during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This study was supported by the Youth Project of Liaoning Provincial Department of Education (project no. lj2017qw001) in 2017.

References

[1] P. Hu, "Research on problem behavior management strategies for college students," *Chinese Journal of Education*, vol. 2, no. S1, pp. 236-237, 2015.

[2] G. Y. Ding and J. K. Cheng, "Research on students' academic performance in higher education: historical review, realistic problems and the way forward," *Heilongjiang Higher Education Research*, vol. 37, no. 10, pp. 52–56, 2019.

[3] H. Shang and D. Zhang, "Research on the status of academic guidance for college students in China--an empirical survey based on 56 colleges and universities across China," *China Higher Education Research*, vol. 3, no. 9, pp. 74–79, 2019.

[4] Y. Pei, S. Fu, and F. Wu, "Visual analysis of hot spots and problems in China's education big data research," *China Distance Education*, vol. 5, no. 12, pp. 46–53+80, 2017.

[5] F. Pan, "Analysis and research on user behavior of campus network," *Journal of Changji University*, vol. 10, no. 28, pp. 47–50, 2013.

[6] F. Deng and Z. Zhang, "Research on the construction of an early warning management platform for student campus behavior analysis based on big data," *China's Electro-Chemical Education*, vol. 7, no. 11, pp. 60–64, 2017.

[7] B. Liu and . Xin, "Empirical study on data mining analysis and application of smart classroom," *Electrochemical Education Research*, vol. 39, no. 6, pp. 41–47, 2018.

[8] J. Wang, "Analysis of the challenges and countermeasures of student management in universities in the era of big data," *Ideological and political education research*, vol. 30, no. 2, pp. 128–130, 2014.

[9] X. Wang, "The supporting role of big data thinking in the information management of students in colleges and universities," *Heilongjiang Higher Education Research*, vol. 8, no. 7, pp. 47–50, 2016.

[10] C. Wang, W. Li, and B. Sun, "Role transfer, big data and student education management model innovation," *China Adult Education*, vol. 8, no. 13, pp. 43–46, 2017.

[11] C. Wu, Y. Cheng, and Y. Zheng, "A review of K-means algorithm," *Modern Library and Intelligence Technology*, vol. 9, no. 5, pp. 28–35, 2011.

[12] S. Suganya, T. Meyyappan, and S. K. Santhosh, "Performance analysis of KMeans and KMediods algorithms in air pollution prediction," *International Journal of Recent Technology and Engineering*, vol. 8, no. 5, pp. 3573–3577, 2020.

[13] A. Nowak - Brzezi, "Outliers in rules - the comparision of LOF, COF and KMEANS algorithms," *Procedia Computer Science*, vol. 176, no. 2, pp. 1420–1429, 2020.

[14] P. S. Xiang, "A cloud detection algorithm for MODIS images combining kmeans clustering and otsu method," *IOP Conference Series: Materials Science and Engineering*, vol. 392, no. 6, pp. 125–129, 2018.

[15] Yi Yi Aung and M. M. Min, "An analysis of K-means algorithm based network intrusion detection system," *Advances in Science, Technology and Engineering Systems Journal*, vol. 3, no. 1, pp. 496–501, 2018.

[16] K. Peng, V. C. M. Leung, and Q. Huang, "Clustering approach based on mini batch kmeans for intrusion detection system over big data," *IEEE Access*, vol. 6, no. 8, Article ID 11906, 2018.

[17] E. L. A. L. L. A. O. U. I. Ahmad, "Clustering kmeans with evolutionary strategies," *International Journal of Imaging & Robotics*, vol. 17, no. 3, pp. 54–65, 2017.

[18] K. Omar and R. Faical, "A fast deterministic kmeans initialization," *International Journal of Applied Information Systems*, vol. 12, no. 2, pp. 6–11, 2017.

[19] W. D. Lai, D. M. Li, J. Yu et al., "An Apriori algorithm-based association analysis of analgesic drugs in Chinese medicine prescriptions recorded from patients with rheumatoid arthritis pain," *Frontiers in Pain Research*, vol. 3, no. 8, Article ID 937259, 2022.

[20] R. Dutta, "Apriori decay estimates for Hardy–Sobolev–Maz'ya equations," *Calculus of Variations and Partial Differential Equations*, vol. 61, no. 1, pp. 128–134, 2021.

[21] I. Daniel, "Near candidate-less Apriori with tidlists and other Apriori implementations," *International Journal of Applied Logistics*, vol. 12, no. 1, pp. 1–22, 2021.

[22] I. Aamo, "A forensic investigation of terrorism in Nigeria: an Apriori algorithm approach," *Journal of Information Security*, vol. 12, no. 04, pp. 270–280, 2021.

[23] K. Lutfiye, "Apriori association rule and K-means clustering algorithms for interpretation of pre-event landslide areas and landslide inventory mapping," *Open Geosciences*, vol. 13, no. 1, pp. 1226–1244, 2021.

[24] C. Yan, J. Zhang, and X. Lu, "Weighted Apriori algorithm with minimum support as interval value," *Journal of Liaoning University of Engineering and Technology*, vol. 35, no. 12, pp. 1512–1516, 2016.

[25] Z. Hong and F. Bian, "An improved Apriori algorithm based on support matrix," *Journal of Wuhan University (Natural Science Edition)*, vol. 6, no. 12, pp. 1246–1249, 2008.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 4950231, 9 pages
https://doi.org/10.1155/2022/4950231

Research Article

Analysis of Guzheng Music Style Transformation Based on Generative Adversarial Networks

Lei Zhang

College of Arts, Xiamen University, Xiamen 361005, China

Correspondence should be addressed to Lei Zhang; zl@xmu.edu.cn

Received 1 August 2022; Revised 25 August 2022; Accepted 6 September 2022; Published 22 September 2022

Academic Editor: R. Mo

With the emergence of more and more new music styles, Guzheng, a traditional national musical instrument that has been handed down for thousands of years, also faces new challenges. However, with the development of the times and aesthetics, composers have created various modern Guzheng music, thus injecting new vitality into the development of Guzheng. In order to study the mutual conversion between Guzheng music and other music, this paper proposes a Guzheng music conversion model based on Star generative adversarial networks (GAN), which can convert any length of music and realize the music conversion between various styles. Based on the idea of star GAN, this model uses Mel spectrum of music and random target style tags to train a general generator to generate a specified music style. It is found that the model proposed in this paper has better conversion effect between similar styles of music; that is, when Guzheng music is converted into pop and flute, the effect is the best, the average accuracy is 89.62% and 90.32%, respectively, and the cosine similarity is also 15.8 and 20.5.

1. Introduction

In recent years, the progress of Guzheng art is obvious to all. The creation and expression of Guzheng music compete with each other like a hundred flowers in full bloom. The traditional national instrument Guzheng, which has been inherited for thousands of years, has beautiful timbre and simple shape and is deeply loved by the public. The cultural heritage of the Guzheng has been developing over time. Since the 1950s, the Guzheng has entered the peak of innovation and creation of Guzheng music [1]. Especially after the 1990s, composers have compiled a variety of modern Guzheng music that conforms to the development of the times and aesthetics, breaking the limitations of the traditional Guzheng style, being all inclusive, integrating national characteristics, and boldly combining western music. At the same time, it has also made innovations in playing techniques, which has injected new vitality into the development of Guzheng. Up to now, while paying attention to preserving the characteristics of traditional Zheng music, Guzheng has been constantly absorbing, integrating, and innovating, constantly injecting new vitality, and bursting out new vitality [2].

In the contemporary social background, people's ability to adapt to the development of globalization is becoming stronger and stronger, and aesthetic changes are becoming faster and faster, so more and more new music forms emerge as the times require. Therefore, as contemporary musicians, we should treat this traditional and modern change with an inclusive heart, and we should adapt to and integrate into new music forms with our own strength. Keeping the traditional music form unchanged, we should better integrate the Guzheng with the current music form, instead of turning it into a "fast-food" performance form but should continue to develop better all the time [3]. This is the direction that every musician works together.

Reviewing the current academic research results, the research on Guzheng art focuses on performance skills, aesthetic characteristics, development process, etc. At the theoretical level, there is still insufficient effort to deconstruct the musical style of important Guzheng works and sort out and summarize the performance techniques, so there is a large research space [4]. Therefore, exploring how to change the style of Guzheng music will make Guzheng music absorb constantly in different cultures, regain its vitality, and emit irreplaceable unique charm.

The goal of music conversion task is to obtain new music with the overall melody or structure of the specified music and the musical characteristics of other specified tracks by learning and extracting the structure, melody, and other various information in the music of different tracks, and recombining them. Similar to image conversion tasks, music conversion can also be divided into two types: style conversion and domain conversion.

The transition from music to visual style is a challenging but important cross model learning problem in creative practice. The main difference between it and the traditional image style transfer problem is that the style information is provided by music rather than images. Assuming that the music features can be correctly mapped to the visual content through the semantic link between the two domains, we solve the music to visual style conversion problem through two steps: music visualization and style conversion. The music visualization network generates an image-based music representation from music data using an encoder generator architecture with a conditional generation countermeasure network. Music style transformation refers to the transformation of a piece of music into new music with the style characteristics of another piece of music while retaining its content characteristics. Among them, the content feature of music generally refers to the overall melody or structure of music, and the style feature generally refers to the genre of music, which is usually extracted from music by pretrained convolutional neural networks (CNN). The goal of music domain conversion is to learn the overall information of music in different domains under the same type, while retaining the overall content of any music in one domain, converting it into music that sounds to belong to another domain under the same type. Among them, type refers to the type of music features that are changed in the conversion process, and domain refers to the specific type of music features.

As a technology that converts music from source genre to target genre while keeping the overall content of source music unchanged, there are many possible application directions in today's huge digital music market. For example, when a musician reproduces another song of a different style, if the two music belong to roughly the same genre, they can usually create new songs by slightly changing their playing instruments and combining their own unique sounds. However, the reality is that the music works of the two musicians usually show different genre styles. In this case, the process of manually converting music belonging to one genre into music belonging to another genre will become very complex, and the workload will become very huge. Music genre conversion can assist musicians in accelerating this music creation process, and even completely realizing automation, so that people can enjoy music that cannot be created by humans alone.

In a word, as a new research field, music style transformation is expected to provide real, vivid, and creative music and become an auxiliary tool to accelerate the process of music creation. So far, generative adversarial networks (GAN) [5] have made many remarkable achievements in computer vision related fields such as random image generation [6, 7] and image-to-image conversion [8–10], but its research achievements in audio modelling and conversion are still relatively few. Choi et al. [11] proposed a model Star GAN that uses a single generator to realize unsupervised image to image conversion between unpaired image data in multiple domains. In view of the success of Star GAN in the image-to-image conversion task, Rizos et al. [12] used the same architecture to convert the Mel spectrum of speech between different emotions. Kameoka et al. [13] enhanced the additional classifier in the original Star GAN and used it to convert the Mel spectrum of speech between different speakers. So far, modelling with GAN in the audio field is still a challenge [14].

Through analyzing the existing research, this paper finds that the existing research mainly realizes the one-way conversion between the two genres, and the conversion between multiple genres of music is mainly achieved by repeatedly establishing the conversion model between the two genres. This method is easy to lead to problems such as a large number of models and complex practical applications. Based on this, this paper analyzes the style of Guzheng music and the style of other music genres, studies the conversion between Guzheng music and other music, establishes a Guzheng music style conversion model based on Star GAN, and realizes the mutual conversion of Guzheng music between different genres based on retaining the overall content of Guzheng music in a given variety of genres of music.

2. Generation Countermeasure Network and Related Technologies

With the continuous development of neural network model and deep learning technology and theory, deep learning models such as generative countermeasure network have been widely used in image, text, audio, and other fields. The basic idea of neural network, as a common machine learning model, is to use high-dimensional nonlinear functions to estimate or approximate model the data. Deep learning is an algorithm based on neural network to learn the hidden features in data [5]. With the continuous development of deep learning theory and technology, a large number of novel deep learning models have been proposed and applied to various scenarios.

2.1. Mel Spectrum. Mel scale is a nonlinear scale unit. It is defined and named according to Stevens, and Ge et al. [15] found that the human ear presents an approximate logarithmic relationship with the Hertz under the frequency scale when sensing the pitch intensity in sound in 1937.

$$\mathrm{m} = 2595 \log_{10}\left(1 + \frac{f}{700}\right). \tag{1}$$

For sound signals, when the frequency is greater than 500 Hz, the human ear's perception of pitch changes, and the

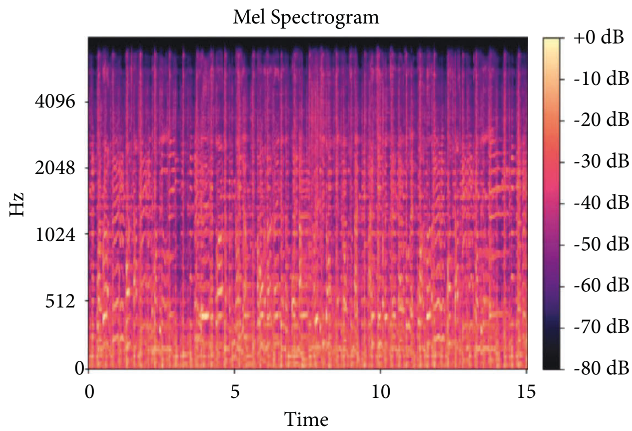


Figure 1: Mel spectrum.

frequency changes under the Mel scale are linear with the changes of Hertz under the frequency scale; that is, they are two times different from the changes of Hertz under the frequency scale. The approximate transformation between the frequency m under the Mel scale and the Hertz f under the frequency scale can be carried out through the above formula to obtain the time-frequency representation of the Mel scale spectrum as shown in Figure 1, also known as the Mel spectrum.

2.2. Griffin-Lim Algorithm. In order to effectively convert the time-frequency representation of low sampling rate obtained from time-frequency analysis of music back to time-domain waveform music, people have explored many different signal processing methods, among which Griffin-Lim algorithm [16] is the most widely used. This algorithm can effectively decode the time-frequency representation obtained from the commonly used STFT time-frequency analysis method back to the time-domain waveform through a large number of calculations [17].

Because the phase information will be lost in the time-frequency analysis of music, it is not easy to reconstruct the time-domain waveform music from the time-frequency representation without phase information. In order to estimate and reconstruct the phase information from the given time-frequency representation, Griffin-Lim algorithm uses an iterative method to minimize the mean square error between the time-frequency representation of the reconstructed waveform music and the given time-frequency representation in each iteration. The Griffin-Lim algorithm flow is shown in Algorithm 1.

2.3. Generative Adversarial Network. Generative adversarial network (GAN) was proposed by Goodfellow et al. [18] in 2014. Since it was proposed, it has been widely studied and is one of the hottest research topics in today's academic circles. Yann Le, the legend of deep learning, once said in the Q&A on quota that "Gan is the most interesting and exciting idea in the field of machine learning in the past decade."

GAN is mainly inspired by the zero-sum game idea in game theory. Its basic structure is shown in Figure 2, including two parts: generator (G) and discriminator (D). These two models are generally realized by neural networks, which can be constructed by deep neural networks with deep structure. In addition, they can also be realized by any form of differentiable system that can map data in space to other spaces [3]. The generator randomly samples from the latent space; that is, it takes the random noise vector as the input, and its goal is to capture and fit the distribution of the real sample and generate a data sample $G(z)$ as similar as the real sample as possible, so as to deceive the discriminator as much as possible. The discriminator is usually a binary classifier, which takes the real samples or the data samples generated by the generator as the input, and its goal is to distinguish the generated samples from the real samples as accurately as possible. The two models constantly compete and adjust parameters, and the ultimate goal is to make the discriminator unable to accurately judge the authenticity of the generated samples, that is, whether the generated samples are real samples. The loss function of GAN in the training process is shown in the following formula:

$$\min_{G}\max_{D} V(D,G) = E_{x\sim Pdata(x)}[\log D(x)] + E_{z\sim p_z(z)}[\log(1-D(G(z)))], \tag{2}$$

where $p_{\text{data}}(x)$ represents the data distribution of real samples, and $p_z(z)$ represents the data distribution of random noise vectors. GAN learns by playing games between generator and discriminator. The optimization of model parameters is a minimum maximum optimization problem. The confrontation game between the generator and the discriminator and the adjustment of parameters end at a saddle point, which is the minimum value for the generator; that is, minimize the function $V(D, G)$, and minimize the difference between the generated sample and the real sample. For the discriminator, which is the maximum value, that is, the maximization function $V(D, G)$, the discriminator's ability to distinguish between real samples and generated samples is improved as much as possible. In other words, the training and learning process of GAN is essentially a "binary minimax game" process, and its ultimate goal is to achieve Nash equilibrium [19, 20]. Therefore, it can be considered that the generator has learned the distribution of real samples and can generate enough "real" new data samples.

2.4. Convolutional Neural Network. Star GAN [11] is a model designed for unsupervised multidomain image-to-image conversion tasks. Its goal is to use a generator g to learn the star mapping relationship between images in multiple domains without any paired training data, in which the numbers 1 to 5 represent image data belonging to different domains.

Considering the above star topology, Star GAN uses a generator g and a discriminator d to form a confrontation network, and its model structure and training process are shown in Figure 3. Among them, D accepts the real image and the false image (generated by the generator) as input, respectively, and predicts the domain type of the input image

Required: time-frequency representation S, short-time Fourier transform f, inverse short-time Fourier transform $f-1$.
Output: waveform music signal xi reconstructed for the ith time.

(1) Randomly initialize the phase of time-frequency representation $P0$, and waveform music signal;
(2) Repeat;
(3) For i do;
(4) Time frequency analysis of $x0$, and get a new time-frequency representation $S0$ and phase $P0$;
(5) Discard the newly obtained time-frequency representation $S0$, and use the original time-frequency representation $S0$ and the newly obtained phase P0 to reconstruct the waveform music signal;
(6) Time frequency analysis of $x1$, and get a new time-frequency representation $S1$ and phase $P1$;
(7) Minimize the mean square error between the time-frequency representation of the reconstructed waveform music signal and the original time-frequency representation S;
(8) End for;
(9) Until error convergence.

ALGORITHM 1: Griffin-Lim algorithm flow

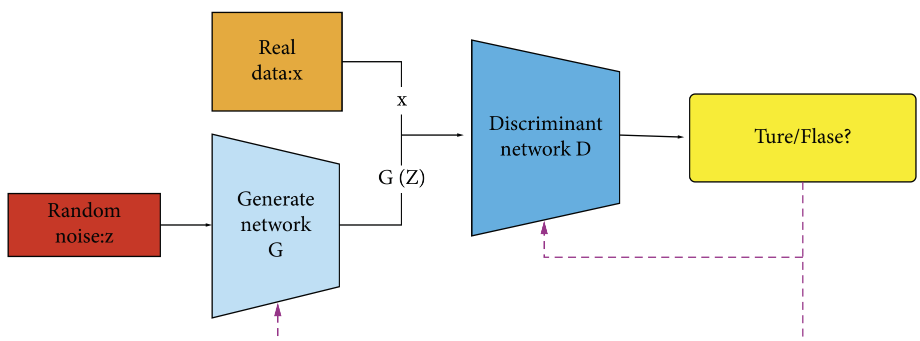


FIGURE 2: Generate the basic structure of confrontation network.

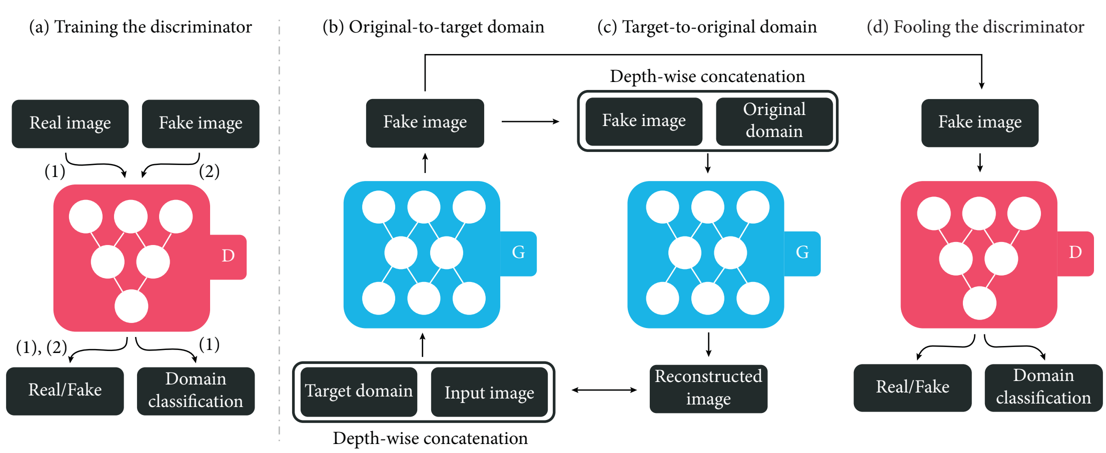


FIGURE 3: Star GAN model structure and training process [11].

while judging its authenticity. G simultaneously accepts the image and the target domain label as input and outputs the converted false image.

During training, Star GAN first uses the real image to train d to make it more capable of distinguishing whether the image is true or not and then trains g to convert the image from the source domain to the target domain and from the target domain back to the source domain, respectively, so that the converted false image can deceive D. The loss function used includes Wasserstein Gan objective with gradient penalty (WGAN-GP) [21], which can make the training process of Gan more stable.

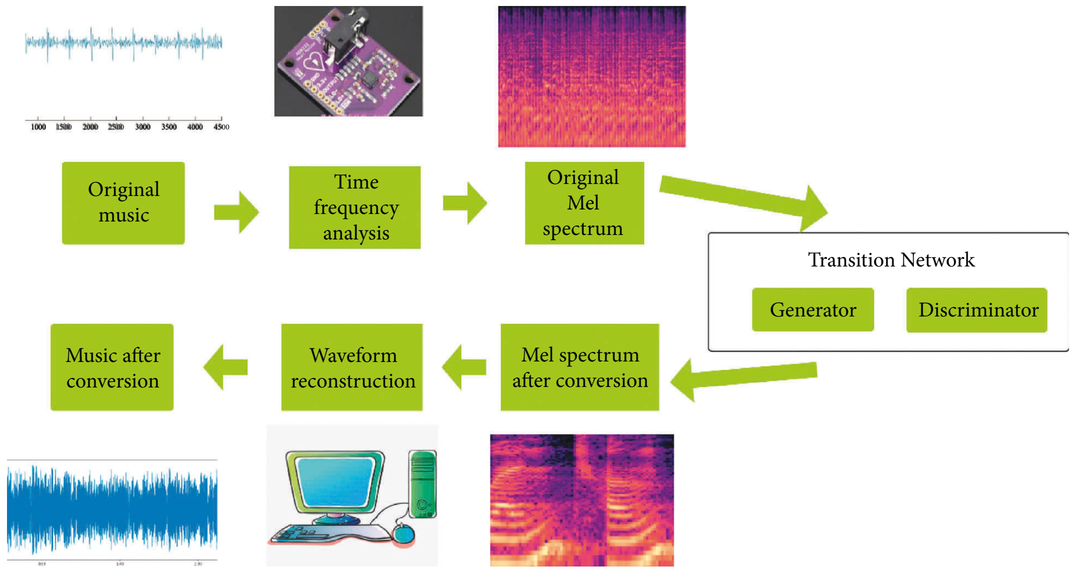


Figure 4: GZ-Star GAN model structure.

$$L_{adv} = E_x[D(x)] - E_{x,c}[D(G(x,c))] - \lambda_{gp} E_{\hat{x}}\left[\left(\left\|\nabla_{\hat{x}} D(\hat{x})\right\|_2 - 1\right)^2\right], \quad (3)$$

where x represents the input image, $\hat{x}$ represents the image evenly sampled between a pair of real and generated images, c represents the target domain label, and λ_{gp} is set to 10 in the experiment.

And two classification losses were used to predict the category of the real and converted images, respectively [11]:

$$L_{cls}^{r} = E_{x,c}[-\log D_{cls}(c \mid x)], \quad L_{cls}^{f} = E_{x,c}[-\log D_{cls}(c \mid G(x,c))], \quad (4)$$

where c' represents the source domain label. And a reconstruction loss based on cyclic consistency loss for constraining the model to retain the original input content during the conversion process [11]is used:

$$L_{rec} = E_{x,c,c}[\|x - G(G(x,c))\|_1]. \quad (5)$$

3. Guzheng Music Translation Based on Star Gan

There are two kinds of Guzheng music conversion, one-to-one music conversion and one-to-many music style conversions. One-to-one music conversion refers to the conversion of Guzheng music into another kind of music, while retaining the content of the conversion of Guzheng music. The goal of one-to-many music style conversion is to realize the mutual conversion between Guzheng music and mixed music, so that the content of the converted music is as close to Guzheng music as possible, and the style it belongs to should be as close to the music style it actually wants to convert to, also known as multimusic style conversion.

In order to solve the above problems as much as possible, this paper proposes a Guzheng music translation based on Star GAN (GZ-Star GAN) by learning the existing music style and instrument conversion models and generating a confrontation network model, which can convert any length of Guzheng music into a variety of target styles. Different from the original Star GAN network, this paper improves the network structure and loss function in GZ-Star GAN to improve the stability of the model during training, the retention effect of the overall content in Guzheng music, and the conversion effect between target music.

In GZ-Star GAN, first, based on the idea of Star GAN, this paper uses the Mel spectrum of music and random target style tags to train a general generator to generate Mel spectrum of the specified style and trains a discriminator to judge the authenticity and style type of Mel spectrum at the same time, so as to realize the style conversion. Secondly, the twin network is introduced to replace the cyclic consistency loss in the original Star GAN to constrain the eigenvectors of the Mel spectrum before and after the conversion, so as to make them equal as much as possible, so as to retain the overall content of the Guzheng music. Third, the differences between music styles are included in the loss function calculation as a weight, and the styles with small differences are given a larger weight value to improve the conversion effect between similar style music. At the same time, in order to improve the stability of the model during training, the hinge loss is introduced to replace the WGAN-GP counter loss in the original Star Gan, and the spectrum normalization operation is adopted for all convolution kernels in the generator and discriminator.

The structure of Guzheng music style conversion model GZ-Star GAN based on Star GAN is shown in Figure 4. It is composed of four parts: time-frequency analysis module, style conversion network, content retention network, and waveform reconstruction module. Among them, the time-frequency analysis module extracts the time-frequency representation of logarithmic amplitude Mel spectrum of time-domain waveform music through short-time Fourier transform and Mel filter. The music style conversion network includes a generator and a discriminator for performing style conversion on the time-frequency representation of the Mel spectrum of music. The content retention network is used to retain the overall content of Guzheng music in the process of style conversion. The waveform reconstruction module reconstructs the converted Mel spectrum to echo music through Griffin Lim algorithm.

3.1. Transformation Network. Based on the idea of the domain image to image conversion model star Gan, this paper constructs a style conversion network in GZ-Star Gan to realize the mutual conversion between Guzheng music and other music styles. The conversion network is composed of a generator and a discriminator. The generator's goal is to convert the Mel spectrum of Guzheng music into a specified target style. The discriminator's goal is to be able to distinguish the real Mel spectrum from the Mel spectrum converted by the generator and to identify the style of the Mel spectrum.

For the generator, GZ-Star Gan uses subpixel convolution network [22], instead of deconvolution network, and upsamples the Mel spectrum on the basis of the original Star Gan generator architecture, to prevent the deconvolution network used by the original Star Gan generator from generating chessboard artifacts in the converted Mel spectrum, thus causing serious noise in the reconstructed waveform music. For the discriminator, GZ-Star Gan refers to the network architecture of the discriminator in Star Gan and adds a network layer for classifying various styles on the basis of the standard discriminator structure to determine whether the Mel spectra of various styles converted by the generator belong to the correct target style.

For a given original Mel spectrum, in order to make the Mel spectrum converted by the generator real enough, the generator and the discriminator need to be trained at the same time. Because the WGAN-GP used by Star GAN uses Wasserstein distance to measure the difference between the two distributions, it will show the characteristics that the error is proportional to the convergence speed during training, and it is not suitable for tasks involving many different types of data and needs to maintain the relative balance of the overall effect.

Therefore, in order to make the confrontation between the generator and the discriminator, the learning process can be more stable. On the one hand, drawing on the recent research progress in the field of GAN, in GZ-Star GAN, spectral normalization (SN) [23] is used for all convolution cores in convolution neural networks used in generators and discriminators; that is, the network parameters of convolution neural networks at each layer are divided by the spectral norm of the network parameters at that layer, replacing the Wasserstein distance used in WGAN-GP networks as a method to realize Lipschitz constraint in the training process of the model. The calculation process of realizing Lipschitz constraint is shown as

$$\left\|\nabla_x (f(x))\right\|_2 = \left\|D_N \frac{W_N}{\sigma(W_N)} \cdots D_1 \frac{W_1}{\sigma(W_1)}\right\|_2 \leq \prod_{i=1}^{N} \frac{\sigma(w_i)}{\sigma(w_i)} = 1, \quad (6)$$

where $f(x)$ is the N-layer neural network with x as the input, $\|\nabla_x (f(x))\|_2$ represents the gradient of $f(x)$, D_N is the diagonal matrix used to represent the action of the nonlinear activation function RELU of the N-layer network, the positive and negative of its diagonal elements are consistent with the input, W_N is the parameter matrix of the N -layer network, and $\sigma(W_N)$ is the maximum singular value of the parameter matrix W_N.

3.2. Content Retention Network. Aiming at the problem that the original Star Gan uses the cyclic consistency loss to retain the original input content during the conversion process, it is easy to introduce additional noise outside the content, thus affecting the retention effect of the Guzheng music content. GZ-Star GAN introduced a content retention network outside the style conversion network to achieve the research goal of converting music between multiple styles and preserving the overall content of Guzheng music as much as possible. In order to realize the retention of the whole content of Guzheng music in the process of style conversion, this paper, based on the idea of double network, constrains the Mel spectrum feature vectors before and after the conversion to be equal as much as possible.

First, in GZ-Star GAN, the transformation vector between paired Mel spectrum sets (x_1, x_2) is defined as

$$t(x_1, x_2) = x_1 - x_2. \quad (7)$$

During training, the content preserving network S extracts and encodes the content feature information contained in the paired Mel spectrum set before conversion and the paired Mel spectrum set converted by the generator G. The training goal is to make the Mel spectrum set after conversion retain the overall music content in the Mel spectrum set before conversion as much as possible, even if the conversion vectors defined in formula (7) meet the relationship shown in formula (8) as much as possible, where $S(x)$ represents the output vector of the network with the combined x of the segmented Mel spectrum as the input, and the content is reserved.

$$t(S(\widetilde{M}_1, \widetilde{M}_2)) = t\Big(S\Big(\widehat{\widetilde{M}}{}_1'', \widehat{\widetilde{M}}{}_2''\Big)\Big). \quad (8)$$

In the training process, referring to the conversion loss originally proposed by Amodio et al. [24], the type conversion loss function is used to optimize the image-to-image conversion task. It should be noted that, in the training process, the loss of genre conversion is optimized by G and S networks together to achieve the goal of reducing iteration, which makes it easy for S to learn a simple identity map to meet the requirements. For example, for any input, s outputs a fixed value. Therefore, a standard twin boundary contrast

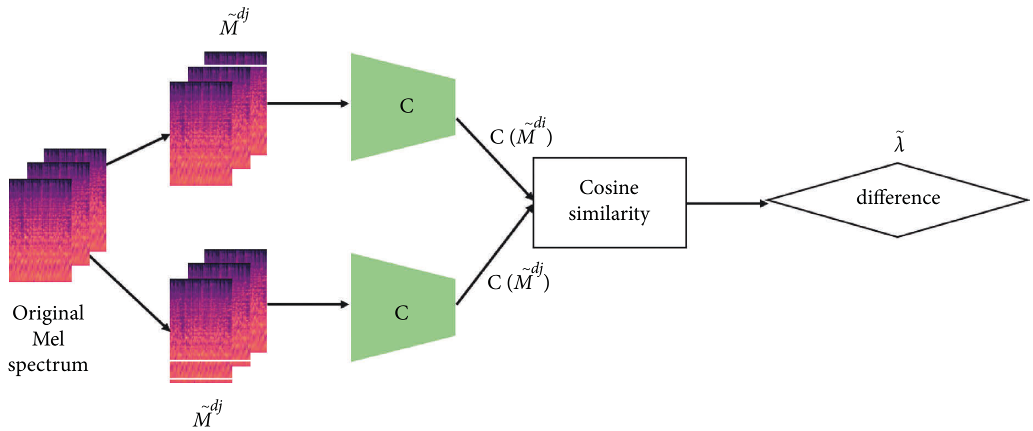


Figure 5: Loss item dynamic weighting scheme.

loss is added here for the content retention networks [25, 26] to avoid the above situation. Among them, δ is a fixed value, set to 2 in this article. By using this loss, the twin network can make the distance between any two vectors obtained by the input coding at least different and avoid the network learning becoming a simple identity mapping function.

3.3. Dynamic Weight of Loss Function. Aiming at the problem that the existing models tend to learn the conversion between style music with obvious differences and that the conversion effect between similar style music is poor, this paper proposes a dynamic weight scheme of loss function, which takes the differences between music styles as weights into the calculation of loss function and gives a larger weight value to the styles with small differences, so as to effectively increase the loss between similar style music and improve the conversion effect of the model [16]. The whole scheme includes two parts: calculating the difference between different styles of music and calculating the dynamic weight of loss function based on the difference between styles of music [27]. Its basic process is shown in Figure 5.

Among them, i and j represent the original Mel spectrum of Guzheng and other styles randomly selected from the original Mel spectrum collection. C represents the additional classifier, which represents the style feature vector extracted by the additional classifier respectively, and $C(i)$ and $C(j)$ represent the dynamic weight of the loss function calculated based on the difference between Guzheng and other music [28].

4. Music Conversion Experiment Based on GZ-Star GAN

4.1. Dataset. In order to evaluate the effectiveness of GZ-Star GAN model in the task of music style conversion and its flexibility in other types of music data, this paper uses GZ-Star GAN to carry out music style conversion experiments on GTZAN style music data set and music instrument conversion experiments on 415-Chinese traditional instrument (CTI) instrument music data set.

GTZAN data set is often used for waveform music style classification task, which contains 10 styles, each style is composed of 100 music samples, and the duration of each music sample is 30 seconds, a total of 1000 music samples. Each music sample in the dataset is a 22050 Hz mono 16 bit audio file in wav format. 415-CTI data set is collected and sorted out from various sources such as CD, microphone recording, and Internet according to the types of traditional Chinese musical instruments. It contains 12 kinds of traditional Chinese musical instruments, each of which is also composed of 100 music samples with a length of 30 s, a total of 1200 music samples [29, 30]. Each music sample in the dataset is a 44100 Hz mono 16 bit audio file in wav format.

The research goal of this paper is to realize the mutual conversion between various styles of music. Therefore, considering the calculation cost of the model and the representativeness of the selected styles and instruments, the music of blues, classical, pop, and jazz, as well as the music of suona, pipa, flute, and erhu, are selected from GTZAN and 415-CTI data sets for experiments, The data set is divided according to the ratio of training and testing 9 : 1. Finally, the number of music samples in the training set of the two data sets is 360, and the number of music samples in the test set is 40.

4.2. Experimental Results and Analysis. In order to evaluate whether the proposed GZ-Star GAN model really learned to convert music of different styles or instruments, this paper uses accuracy as the evaluation index. Accuracy refers to the probability that samples are recognized correctly, involving four concepts: true positive (TP) representing the number of positive samples recognized as positive cases, false positive (FP) representing the number of negative samples recognized as positive cases, true negative (TN) representing the number of positive samples recognized as negative samples, and false negative (FN) representing the number of negative samples recognized as negative cases. Its calculation formula is

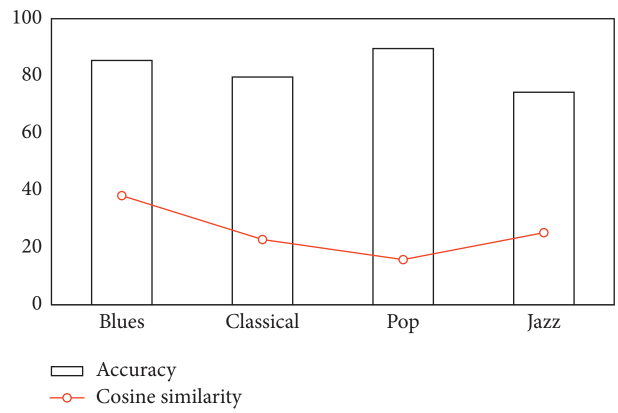


Figure 6: Test average accuracy and cosine similarity between different music.

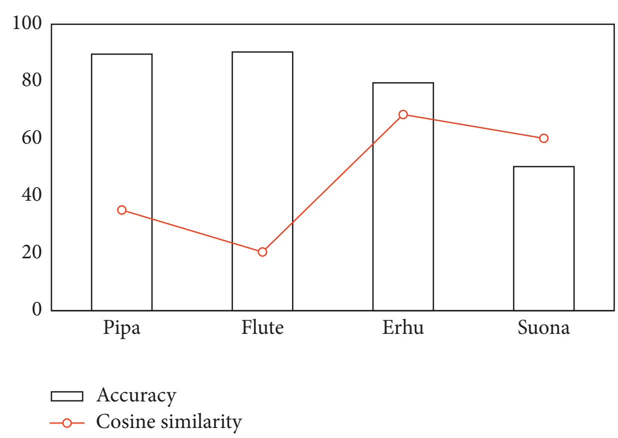


Figure 7: Test average accuracy and cosine similarity between different instruments.

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + TN + FP}. \quad (9)$$

In order to evaluate the conversion effect of musical instruments and styles of GZ-Star GAN model, it is applied to different music data sets to try to convert music between different instruments. Different music data sets are used to train the additional classifiers. The average accuracy and average cosine similarity of the trained additional classifiers on the corresponding test set are shown in Figure 6, where the value range of cosine similarity is [−1, 1]. The larger the value, the more similar the two genres of music. The smaller the value, the more dissimilar the two genres of music are. Figure 6 shows that the average accuracy between Guzheng and pop is 89.62%, indicating that the two kinds of music have higher similarities in some aspects. The average accuracy between Guzheng and jazz music is the lowest, 50.3%. At the same time, Figure 6 shows that, compared with other music, the cosine similarity between Guzheng and pop music is the smallest, which also shows that the two kinds of music have higher similarity.

Similarly, the instrument music data set is used to train the additional classifier, and its average accuracy and average cosine similarity on the corresponding test set are shown in Figure 7. As can be seen from Figure 7, the average accuracy between Guzheng and flute music is 90.32%, indicating that the two instruments have higher similarities in some aspects. The average accuracy between Guzheng and Suona music is the lowest, 50.3%. At the same time, Figure 7 shows that, compared with other instruments, the cosine similarity between Guzheng and flute music is the smallest, which also shows that the musical similarity of these two instruments is higher.

5. Discussion and Conclusion

Aiming at the conversion between Guzheng music and other music, this paper proposes a Guzheng music conversion model based on Star GAN, which can convert any length of music and realize the music conversion between various styles. The main work of this paper includes the following aspects:

(1) This paper presents a Guzheng music conversion model GZ-Star GAN. Based on the idea of Star GAN, this model uses the Mel spectrum of music and random target style tags to train a general generator to generate Mel spectrum of a specified style and train a discriminator to judge the authenticity and style type of Mel spectrum at the same time, which can convert Guzheng music of any length into other target music and realize the conversion of Guzheng music.

(2) In order to improve the retention effect of Guzheng music content in the process of music conversion, a content retention network is introduced to constrain the eigenvectors of Mel spectrum before and after conversion, making them equal as much as possible, which improves the retention effect of the overall content of Guzheng music.

(3) In order to improve the conversion effect between similar music in the process of music conversion, this paper adds the difference between music as a weight to a dynamic weighting scheme of loss function, and the music with small difference is given a larger weight value, which effectively increases the loss between similar music and improves the conversion effect of the model between similar music.

(4) In order to verify the music conversion effect of the model, experiments are carried out on the style and instrument music data sets, respectively. The experimental results show that the model proposed in this paper has a good conversion effect between similar styles of music and also has a good effect between other styles of music. When Guzheng music is converted into pop music and flute music, the effect is the best, the average accuracy reaches 89.62% and 90.32%, respectively, and the cosine similarity also reaches 15.8 and 20.5. It verifies the flexibility that this model can be applied to other types of music styles.

Data Availability

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

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] H. Gao and M. Yin, "On the basic musical style of the guzheng song everlasting longing for each other," *Northern Music*, vol. 40, no. 3, p. 25+38, 2020.

[2] S. Lu, "An analysis of the musical style and performance of Zheng song "western theme Fantasia"," *contemporary music*, vol. 18, no. 3, pp. 150–153, 2022.

[3] N. Meng and Z. Wang, "Investigation and Research on the application of Guzheng in contemporary rock music," *art evaluation*, vol. 35, no. 1, pp. 3–5, 2020.

[4] Y. Cheng, "Discussion on timbre influencing factors and skill improvement strategies in Guzheng performance," *Northern Music*, vol. 37, no. 19, pp. 76–78, 2020.

[5] Q. Wang, B. Cao, P. Zhu, N. Wang, Q. Hu, and X. Gao, "Face image translation based on self-discriminant cycle generation confrontation network," *Chinese Science: Information Science*, vol. 52, no. 8, pp. 1447–1462, 2022.

[6] I. Gulrajani, F. Ahmed, M. Arjovsky, D. Vincent, and C. Aaron, "Improved training of Wasserstein GANs," *Advances in Neural Information Processing Systems*, vol. 3, p. 30, 2017.

[7] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of GANs for improved quality, stability, and variation," 2017, https://arxiv.org/abs/1710.10196.

[8] J. Y. Zhu, T. Park, P. Isola, and A. Alexei, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, pp. 2223–2232, Venice, Italy, October 2017.

[9] P. Isola, J. Y. Zhu, T. Zhou, and A. Alexei, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1125–1134, Nashville, TN, USA, July 2017.

[10] T. C. Wang, M. Y. Liu, J. Y. Zhu, T. Andrew, K. Jan, and C. Bryan, "High-resolution Image Synthesis and Semantic Manipulation with Conditional GANs," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8798–8807, Salt Lake City, UT, USA, June 2018.

[11] Y. Choi, M. Choi, M. Kim, H. Jung-Woo, K. Sunghun, and C. Jaegul, "Star GAN: unified generative adversarial networks for multi-domain image-to-image translation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8789–8797, Salt Lake City, UT, USA, June 2018.

[12] G. Rizos, A. Baird, M. Elliott, and S. Björn, "Star GAN for Emotional Speech Conversion: Validated by Data Augmentation of End-To-End Emotion recognition," in *Proceedings of the ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3502–3506, IEEE, Barcelona, Spain, May 2020.

[13] H. Kameoka, T. Kaneko, K. Tanaka, and N. Hojo, "Non-parallel voice conversion with augmented classifier star generative adversarial networks," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 2982–2995, 2020.

[14] C. Donahue, J. McAuley, and M. Puckette, "Adversarial Audio synthesis," 2018, https://arxiv.org/abs/1802.04208.

[15] Y. Ge, X. sun, Q. Zhang, D. Wang, and X. Wang, "Face sketch synthesis based on cyclic generation confrontation network," *Journal of Jilin University (Science Edition)*, vol. 60, no. 4, pp. 897–905, 2022.

[16] D. Griffin and J. Lim, "Signal estimation from modified short-time Fourier transform," *IEEE Transactions on Acoustics, Speech, & Signal Processing*, vol. 32, no. 2, pp. 236–243, 1984.

[17] H. Liu, J. song, and S. Li, "Study on the discrimination between earthquake and ground pulsation based on the combination model of generating confrontation network and random forest," *Vibration and shock*, vol. 41, no. 15, pp. 312–318, 2022.

[18] I. Goodfellow, J. Pouget-Abadie, M. Mirza et al., "Generative adversarial nets," *Advances in Neural Information Processing Systems*, vol. 27, 2014.

[19] "Wang Haijun Minimax theorem of set-valued two functions," *Practice and Understanding of Mathematics*, vol. 5, no. 8, pp. 293–298, 2015.

[20] L. J. Ratliff, S. A. Burden, and S. S. Sastry, "Characterization and Computation of Local Nash Equilibria in Continuous games," in *Proceedings of the 2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 917–924, IEEE, Monticello, IL, USA, October 2013.

[21] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein Generative Adversarial networks," in *Proceedings of the International Conference on Machine Learning*, pp. 214–223, PMLR, Sydney, Australia, August 2017.

[22] W. Shi, J. Caballero, F. Huszár et al., "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1874–1883, Las Vegas, NV, USA, June 2016.

[23] T. Miyato, T. Kataoka, M. Koyama, and Y. Yuichi, "Spectral Normalization for Generative Adversarial networks," 2018, https://arxiv.org/abs/1802.05957.

[24] M. Amodio and S. Krishnaswamy, "Travelgan: image-to-image translation by transformation vector learning," in *Proceedings of the ieee/cvf Conference on Computer Vision and Pattern Recognition*, pp. 8983–8992, Long Beach, CA, USA, June 2019.

[25] I. Melekhov, J. Kannala, and E. Rahtu, "Siamese network features for image matching," *IEEE*, in *Proceedings of the 2016 23rd international conference on pattern recognition (ICPR)*, pp. 378–383, Cancun, Mexico, December 2016.

[26] E. J. Ong, S. Husain, and M. Bober, "Siamese network of deep Fisher-vector descriptors for image retrieval," 2017, https://arxiv.org/abs/1702.00338.

[27] Y. Wang, R. J. Skerry-Ryan, D. Stanton et al., "Tacotron: towards end-to-end speech synthesis," 2017, https://arxiv.org/abs/1703.10135.

[28] M. Zhao, J. Jiang, H. Yin, L. Li, and J. Yao, "Defocusing of single defocused image based on nonlocal generation countermeasure network," *Geographic information of Surveying and mapping*, vol. 47, no. S1, pp. 142–147, 2022.

[29] N. J. Conard, M. Malina, and S. C. Münzel, "New flutes document the earliest musical tradition in southwestern Germany," *Nature*, vol. 460, no. 7256, pp. 737–740, 2009.

[30] D. Eck and J. Schmidhuber, "A first look at music composition using lstm recurrent neural networks," *Istituto Dalle Molle Di Studi Sull Intelligenza Artificiale*, vol. 103, p. 48, 2002.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9782038, 1 page
https://doi.org/10.1155/2023/9782038

Retraction

Retracted: Artificial Intelligence Technology and Its Application in Improving Thought-Politics Education

Mobile Information Systems

Received 13 September 2023; Accepted 13 September 2023; Published 14 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] X. Bai and X. Wang, "Artificial Intelligence Technology and Its Application in Improving Thought-Politics Education," *Mobile Information Systems*, vol. 2022, Article ID 3150352, 11 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 3150352, 11 pages
https://doi.org/10.1155/2022/3150352

Research Article

Artificial Intelligence Technology and Its Application in Improving Thought-Politics Education

Xueyan Bai[1] and Xiajian Wang[2]

[1]*School of Marxism, Hohai University, Nanjing 211100, China*
[2]*School of Marxism, Nanjing University of Finance and Economics, Nanjing 210023, China*

Correspondence should be addressed to Xiajian Wang; 9119871089@nufe.edu.cn

Received 4 August 2022; Revised 29 August 2022; Accepted 5 September 2022; Published 20 September 2022

Academic Editor: R. Mo

As socialism with Chinese characteristics enters a new age, thought-politics education has also entered a new phase of development. Facing new opportunities and challenges, attaching importance to quality, pursuing quality, and improving quality have become the key directions for the development of thought-politics education. In the paper, a CNN-BiLSTM-based recommendation algorithm is proposed and combined with AI-related algorithms to be used in the task of improving thought-politics education. First, deep feature extraction in temporal and spatial dimensions by using LSTM and convolutional networks, respectively; then, by using multiscale attention fusion mechanisms to enhance the expressiveness of the features and make recommendations with the help of multilayer perceptron. Through extensive experiments, it is verified that the model has higher recommendation performance in this paper while maintaining higher real-time performance. It is tested on real data sets to verify that the model in the paper has better robustness.

1. Introduction

Nowadays, with the fast development, full integration, and widespread application of the new generation of network information technology represented by using mobile internet, internet of things, cloud computing, and artificial intelligence, humanity is gradually entering the "intelligent information era" of large-scale data mining, innovation [1], and application. The development and application of artificial intelligence and big data have become a new engine for economic transformation and growth and a new tool to reshape the country's competitive advantage, and it has also become a new approach to improving national governance capacity in countries around the world. In recent years, our governments have been actively promoting carrying out the national big data strategy, accelerating the pace of building digital China, the application of artificial intelligence in various areas of the economy and society continues to advance in-depth, and the scientific value and social benefits are gradually highlighted. As the frontier of contemporary information technology development, the wide range of information resources, advanced information processing technology, and new thinking paradigm of big data empower the innovation and development of daily thought-politics education of college students and promote the intelligence of daily thought-politics education of college students.

When socialism with Chinese characteristics enters a new age, A range of new changes, which have led to new opportunities for both daily thought-politics education for college students, are occurring in the context of daily thought-politics education of college students, their goals and tasks, and the target audience of their work. In terms of new opportunities, on the one hand, the overall strengthening of the state's leadership of thought-politics education in colleges and universities and the strengthening of the fundamental mission of "establishing moral education" in colleges and universities have brought a new historical opportunity for the daily thought-politics education of college students [2]. The national leaders emphasized that from the overall height of achieving the great rejuvenation of the Chinese nation and the strategic height of realizing the party's mission of governance and consolidating the party's ruling position, they have made a comprehensive, profound, and systematic exposition of a series of issues such as the strategic status, duty, and mission, objectives and tasks, guidelines and principles, methods and paths, reform and innovation of thought-politics education work in colleges and universities at the National Conference on thought-politics work in colleges and universities, the Symposium for Teachers of Thought-politics Theory Courses in Schools, the Symposium for teachers and students of Peking University, and the National Education Work Conference. Importantly, it provides the scientific guidance and fundamental followings for the thought-politics education work of colleges and universities in the new period to develop a new situation and achieve new results. In addition, some relevant policies and regulations have been issued one after another, providing programs and planning paths for strengthening and innovating thought-politics work in universities. Under this situation, the daily thought-politics education of college students, which is the main position of thought-politics education for college students, will usher in new development opportunities. On the other side, the fast development of the new round of information technology revolution with digitalization, networking, and intelligence as its core features and its extensive integration with the economy and society have provided new vehicles and tools for the daily thought-politics education of college students. Currently, the new generation of network information technology represented by using mobile internet, Internet of things, cloud computing, and artificial intelligence is developing rapidly; digitalization, networking, and intelligence are prominent features but also the core of the emerging information technology. The development and application of new network media and new technology of big data provide new vehicles and tools for the day-to-day thought-politics education of students in the college.

The effective application of artificial intelligence and big data technology in the improvement of thought-politics education is beneficial to enrich the theory of thought-politics education for college students. Especially in the last few years, to enhance the thought-politics education of students in the college to achieve affinity in the context of anti-epidemic is essentially the implementation of the professional ethics of college staff to care for students, which is the new requirement and development of the thought-politics education of students in the college in the new situation and new context. Secondly, it allows providing the diversified development needs of college students in the new age. College students entering the new era are more influenced and impacted by using the internet, have more ways and opportunities to contact new things, have more active thinking, more obvious personalized differences, more complex behavior patterns, and more diversified development needs for growth. With the assistance of big data technology and artificial intelligence technology, we are able to recommend positive cultural and educational columns to users in real time in the massive data, which can help meet the diversified development needs of college students in the new age [3]. It helps promote college students' identification with thought-politics education. The most direct and obvious effect of thought-politics work in colleges and universities is to gain the acceptance and recognition of university students, which is related to whether the overall goal of thought-politics education can be achieved. The research study on enhancing the affinity of thought-politics education for college students in the new age will provide colleges and universities with the choice of paths, to build an intimate relationship with students in the education process, and also improve, to some degree, the identity of college students for thought-politics education. It also helps to build a harmonious and friendly campus environment in the new age of universities. The education system of colleges and universities is a unified organic whole. To improve the affinity of thought-politics education of college students in the new age, there is an urgent need for the joint efforts of many departments of colleges and universities and the whole education system. The whole staff, whole process, and all-round education mode is not only beneficial to optimize the relationship between teachers and students but also to the friendly collaboration between various organizational departments of the school, thus promoting the formation of a harmonious and friendly campus atmosphere and the construction of a harmonious and friendly campus environment.

Improving the thought-politics education of university students with the help of artificial intelligence technology is beneficial to promote the implementation of the establishment of the basic tasks of ethical education in colleges and universities. "The foundation of colleges and universities lies in the establishment of moral education." To establish morality is to use moral education to guide, sensitize, and motivate college students, to urge them to form correct moral concepts and excellent moral qualities, and to practice a good moral reality. We shape the people that are under the concept of "people-oriented" and "student-oriented" education; the university students will be shaped by using various educational methods, guided to become ideal, competent, and responsible newcomers of the times, and promoted to the overall development of all human qualities. In particular, in today's rapidly growing network, how to accurately recommend positive cultural columns for colleges and universities to carry out the thought propaganda of the majority of young students, promote the core values of socialism, and help cultivate qualified socialist builders? As the main training ground for innovative talents, colleges and universities will become important participants, facilitators, and promoters under the wave of big data. As the most cutting-edge application of the new generation of information technology, the wide range of information resources, advanced information processing technology, and a new way of thinking contained in big data bring new impetus and space for the development and innovation of immediate, accurate, forward-looking and personalized daily thought-politics education for college students. Based on its resource advantages, technical advantages, and thinking advantages, big data has created a new space in cracking the lack of information acquisition ability, weakened subject-object relationship, and ineffective education methods in traditional thought-politics education in colleges and universities. Demonstrating new values in meeting the diversity of needs for thought-politics education resources, the plurality of practice development, and the complexity of thinking transformation in universities. It brings new opportunities in providing strategic assets for thought-politics education in universities, exploring the laws of education, and grasping opportunities. It is a real need for the daily thought-politics education of college students to conform to the development of the times by using big data resources, technologies, and methods to promote the innovation of thinking, optimization of supply, improvement of methods, and reconstruction of paradigms, constructing a scientific, digital and intelligent daily thought-politics education system for college students, and promoting the transformation and upgrading of daily thought-politics education of college students to “accurate thinking and politics” and “intelligent thinking and politics. It is also an important growth point and strong impetus for the daily thought-politics education of college students to further improve its quality and efficiency and give it a new chance.

2. Related Works

In terms of thought-politics education, there is a lack of empirical evidence that professional scholars in China have paid due attention to and systematically studied the quality improvement of thought-politics education for college students. There is no chapter dedicated to “quality of thought-politics education” in any domestic thought-politics education textbook. The academic literature on systematic research on quality improvement of thought-politics education of college students has not been searched in the core journal paper database; in general, journals, although articles related to the quality of thought-politics education are abundant, many of them have made profound theoretical discussions on the issue,improving the quality of thought-politics education from different aspects; most of the articles are not directly related to this field, but only mention it in passing when discussing other topics, and there are very few specialized studies on it.

2.1. Study on the Theoretical Basis of Daily Thought-Politics Education for University Students. Theoretical foundation research mainly focuses on the conceptual connotation, current characteristics, content system, main principles, and basic laws of daily thought-politics education of college students, which aims to systematically construct and professionally develop the daily thought-politics education of college students.

Definition of the connotation of daily thought-politics education for college students. There are three main tendencies in defining and interpreting the connotation of the concept of daily thought-politics education for college students: one is to focus on the “subject” [4]. This view emphasizes that the daily thought-politics education of college students refers to the educational activities organized and carried out mainly by student staff, such as counselors. For example, literature [5] believes that daily thought-politics education of college students refers to various educational activities organized and carried out by college student staff, such as counselors and class teachers, according to the Party’s educational policy and the requirements of carrying out thought-politics education work in colleges and universities, from the law of students’ cognitive development and in response to different students’ thought reality, with dormitories, class groups and grade work as the carrier. Literature [6] focuses on the “everyday” statement. This viewpoint defines daily thought-politics education for college students as a form of education other than classroom teaching in thought-politics theory courses, emphasizing the “daily” nature. Literature [7] points out that the daily thought-politics education of college students refers to the thought-politics education activities that are often carried out on weekdays, penetrates the daily study and life of college students, and plays a subtle role in the overall development and healthy growth of college students, with the characteristics of regularity, extensiveness, continuity, latent and permeability. Literature [8] focuses on the “carrier” saying, “This viewpoint focuses on the daily thought-politics education carried out by various carriers such as party and group organizations, social practice, campus culture, and network management”. For example, according to literature [9], the daily thought-politics education of college students is “the most basic and important way to carry out thought-politics education and daily management for students with the carrier of party and group organizations, association organizations and classes.

In addition, the issue of effectiveness has always been the central issue in the study of daily thought-politics education of college students. The researcher mainly explores the strategies and paths to enhance the effectiveness of daily thought-politics education of college students from the aspects of the education concept, education content, education method, team building, and education mode. (1) Update the concept of education. Literature [10] points out that “people-oriented” is the educational concept and methodological principle that must be adhered to improve the effectiveness of daily thought-politics education of college students. Literature [11] points out that respecting the equality of personalities in educational philosophy and realizing the transformation of subject-object dichotomy to the intersubjectivity relationship is the way out of the dilemma of daily thought-politics education of college students. (2) Enrich the education carrier. Literature [12] explored the use of various carriers such as party and caucus organizations, club activities, class collective construction, campus culture, social practice, and network management in the daily thought-politics education of college students to enhance their effectiveness. For example, in terms of cultural carriers, it is pointed out that the spirit of Jinggang Mountain and Qilu culture are used as carriers to carry out daily thought-politics education for college students. In terms of network carriers, it is pointed out that microblogs, new media, WeChat, self-media, public numbers, and APP software are used to enhance the effectiveness of daily thought-politics education of college students. In terms of activity carriers, literature [13] points out that classes, dormitories, and clubs are used to carry out daily thought-politics education of college students. (3) Improve education methods and approaches. Literature [14] points out the relationship between the daily thought-politics education methods of college students and the effectiveness of education in terms of promoting and leading roles and proposes ideas and strategies to enhance the effectiveness of education methods. Literature [15] explored the methods and approaches of daily thought-politics education for college students, such as whole process experiential, narrative therapy, and mainstream thought opinion guidance. (4) Innovative education model. In literature [16], Northeast Forestry University, for example, proposes to carry out “theme-driven education,” i.e., to determine the corresponding educational themes in different grades and different types of student groups and to provide students with highly targeted, time-concentrated, and thematic educational and practical activities in a phased, categorized, progressive, and hierarchical manner. Literature [17] proposes four approaches to daily thought education of college students: organizational construction, special lectures, program support, and individual counseling. Literature [18] points out the construction of a trinity of school, society, and family in the management of daily thought-politics education of college students. Literature [19] states that the effectiveness of work is enhanced through project-based and curriculum-based work of counselors. Literature [20] proposes to build a “five-in-one” system of daily thought-politics education for college students with institutionalized, refined, personalized, informationized, and branded management. (5) Improve the education mechanism. Literature [21] proposes to enhance the effectiveness by exploring the construction of demand mechanism, decision-making mechanism, implementation mechanism, evaluation mechanism, incentive mechanism, and guarantee mechanism for daily thought-politics education of college students. Literature [22] points out that strengthening collaborative mechanisms to enhance the effectiveness, including collaborative mechanisms between thought-politics theory courses and daily thought-politics education, and collaborative mechanisms between online thought-politics education, curriculum thinking and politics education, and daily thought-politics education. Literature [23] states that a long-term mechanism of daily thought-politics education for college students is built through classroom teaching, campus culture construction, school management, campus environment optimization, student club activities and social practice activities. Literature [24] points out that we should implement the “three close” requirements and establish a student evaluation mechanism based on the information network platform.

2.2. The Function of Deep Learning in Enhancing Thought-Politics Education. Recently, achievements of deep learning in natural language processing and computer vision have led the recommendation field to take notice of the robust tool, and scholars have begun to explore the methods of using deep learning so as to improve some of the insurmountable weaknesses of today’s recommender systems, such as data sparsity, cold starts, and poor interpretability. Particularly, the advent of CNN and RNN has been a great success in many natural language processing tasks. So, people started to experiment with deep learning methods, DeepCoNN, D-Attn, etc., to mine user preferences and make recommendations from the user’s perspective, which can be directly applied to predictive scoring [25]. DeepCoNN is composed of two parallel neural networks with CNN as the basic model, learning the representation of the learner and the thought-politics-cultural preferences of interest, respectively, and connecting the two sections at the top of the network to see the interaction, which demonstrates the validity of user-focused texts on thought-politics education for alleviating the sparsity problem.

The focus of the attention mechanism is to learn a weight to identify the degree of importance, and it has been widely used for natural language processing since its introduction, achieving state-of-the-art results in machine translation, reading comprehension, and speech recognition, as well as other areas. Thus, attention mechanisms receive attention in the field of recommendation and start being used for building recommendation algorithms that enhance thought-politics education. NARRE uses attention mechanisms to learn textual representations of different cultures of thought, to better model users and interests, predict interests, and generate explanations. Unlike the D-Attn word-level attention mechanism, NARRE uses a two-channel attention mechanism. Inspired by transformer, MPCN proposes a new pointer-based learning scheme without using RNNs and CNNs and relying entirely on attention mechanisms, which allows for deep textual interaction between the learner and the thought-politics culture of concern, with excellent results.

The development of NLP has contributed greatly to the application of text in the area of recommendation. Pretrained language models have evolved rapidly since they were proposed, resulting in lots of great approaches such as feature-based Elmo and fine-tuning-based OpenAI GPT. However, these language models are all unidirectional, limiting the representational power of pretraining. So, literature [26] proposed a bidirectional pretraining model, BERT, which reads the entire text at once using Transformer’s Encoder and this allows the model to learn based on both sides of the word and thus grasp more accurately the meaning expressed by the word in the sentence. Thus, BERT is naturally bidirectional, with a high generalization capability, and provides a good basis for downstream tasks.

In summary, this paper combines techniques such as CNN and LSTM in deep learning to construct recommendation models and apply them to the educational work of ideology and politics, which will greatly facilitate the rapid mining of columns such as education and culture of interest to learners in the massive data, quick reading to locate the knowledge of interest and real-time attention. Meanwhile, it is beneficial to enhance the cognitive ability of people’s

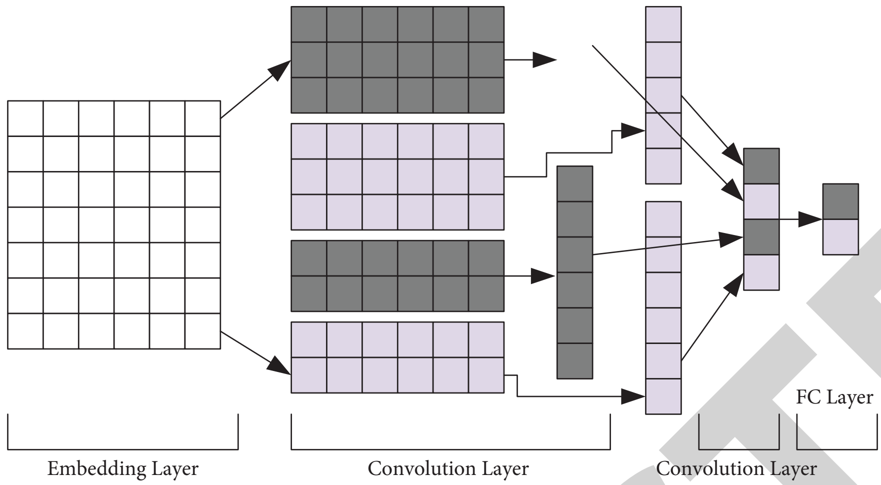


Figure 1: Convolution neural network model.

thinking and has a positive effect on improving the quality of all people.

3. Methodology

The recommended model of thought-politics education culture with artificial intelligence in this paper includes features extracted from the spatial dimension by using convolutional neural networks and temporal features extracted from the temporal dimension by using bidirectional long and short-term very easy neural networks, multiscale fusion based on the attributes of both features, and deep recommendation by using fused features. Section 3.1 gives the theory of convolutional neural networks; then, the process of feature extraction from the temporal dimension by long- and short-term memory neural networks is presented in Section 3.2; subsequently, a scale fusion attention is introduced in Section 3.3. Finally, Section 3.4 gives the strategy for computing the prediction scores and generating recommendations.

3.1. Convolutional Neural Network. The basic structure of a convolutional neural network [27] is composed of an input layer, a pooling layer, a fully connected layer, a convolutional layer, and an output layer. The convolution layer is an inner product of the embedding vector and the filter matrix. The pooling layer pools each feature map obtained from the convolution layer. Numerous studies by scholars have confirmed that the maximum pooling approach can extract better features and better results compared to the mean pooling, and current studies are using the maximum pooling approach. The fully connected layer takes the output features of the pooling layer as input and activates them by the activation function to obtain a feature vector of fixed dimension. CNNs are outstanding in classification problems, but there are few results in recommendation algorithms, mainly because the recommendation is a regression problem and the goals of the two are different. In the recommendation algorithm for text data, scholars fused CNN with a classical recommendation algorithm to build a hybrid recommendation model to achieve more accurate recommendation results. The structure of the convolutional network is given in Figure 1.

3.1.1. Input Layer. In the neural network, the embedded data have low dimensionality and can map discrete sequences into continuous vectors. Here, the input layer is used to convert the textual information of the learning resource into an embedding matrix, where each row in the matrix is a clause element, which can be expressed as equation (1).

$$D = \begin{bmatrix} w_{11} \cdots w_{1i} \cdots w_{1m} \\ w_{21} \cdots w_{2i} \cdots w_{2m} \\ \cdots\cdots\cdots \\ w_{n1} \cdots w_{ni} \cdots w_{nm} \end{bmatrix}, \tag{1}$$

where m indicates the dimensionality of the embedding, n indicates the number of words, and $[w_{i,1:m}]$ indicates the vector form of the i-th word.

3.1.2. Convolution Layer. Multiple convolution kernels of different sizes are used to make convolution operations on the embedding matrix, and the window size is the number of words covered by each convolution. In order to cover the whole word embedding vector, this paper sets the convolution size in the format of the number of words × vector dimension. The convolution operation flow is shown in Figure 2.

3.1.3. Pooling Layer. The pooling layer is mainly used after the convolution layer so as to reduce the feature map dimension and the number of network parameters by a downsampling operation. Common pooling operations include average pooling and maximum pooling. The pooling operation can ignore small variations in the feature maps and improve accuracy while effectively avoiding the overfitting phenomenon. Assuming that the feature map

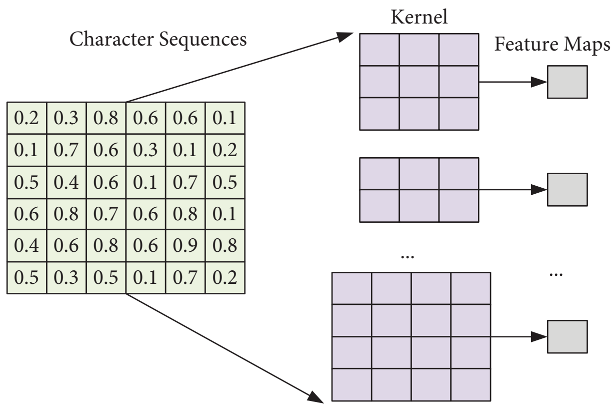


Figure 2: The process of the convolute operation.

obtained in the t-th convolutional layer is $M_t = \{m_1, m_2, ..., m_s\}$, the maximum pooling strategy is used to extract the maximum value for M_t, p_i denotes the pooling result of the t_i-th convolutional layer, and the pooling operation can be expressed as equation (2).

$$p_i = \max(M_t) = \max\{m_1, m_2, \ldots, m_s\}. \tag{2}$$

3.1.4. Fully Connected Layers. The main role of the fully connected layer is to synthesize the previously extracted feature values and output a fixed-size feature vector. Suppose there are m neurons in the fully connected layer, and after the ReLu activation function, a fixed size vector s is obtained, and it is the text feature vector of the learning resource. The calculation is shown in equation (3).

$$F_i = \sigma(w_i p_i + b_i), \tag{3}$$

where p_i denotes the output of the learning resource text information on the pooling layer, σ is the ReLu activation function, w_i denotes the weight, and b_i is the bias.

3.2. Bidirectional Long- and Short-Time Recurrent Neural Network. In the traditional recurrent neural network [28], the neurons between the layers are interconnected, which can preserve the short-range temporal features, but the gradient between the hidden layers is unstable, and there is the problem of gradient disappearance or gradient explosion. The long- and short-term temporal recurrent neural network structure allows the gradients to pass well through each hidden layer and can learn data with temporal characteristics such as text very well. The neurons of the LSTM structure learn only the information of the neurons in front of the layers, while the words before and after the words affect the semantic relations. Bidirectional long short-term recurrent neural network (BiLSTM) combines two sets of long short-term recurrent neural networks with opposite learning directions, which can better understand the contextual semantics compared to LSTM. The LSTM neural network consists of four main elements: forgetting gate f_t, input gate i_t, memory unit c_t, and output gate o_t. The forgetting gate determines the retention of the previous state information in the memory unit, the input gate controls the input of the current moment information in the memory unit, and the memory unit updates the memory state according to the current input information; then, the output gate determines the output result of the memory unit for the next state. The calculation process can be expressed as shown in equations (4) to (9).

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f), \tag{4}$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i), \tag{5}$$

$$\tilde{c}_t = \tan h(W_c[h_{t-1}, x_t] + b_c), \tag{6}$$

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t, \tag{7}$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o), \tag{8}$$

$$h_t = o_t * \tanh(c_t), \tag{9}$$

where W is the matrix multiplication operation, h_t denotes the state of the memory cell, x_t denotes the information input, b is the bias term of the function, $\sigma(\cdot)$ is the sigmoid function, and $*$ denotes the dot product operation.

The two-way long and short-term memory model is two sets of forward and backward LSTM models connected with sequential LSTM learning features h_t^{forword} , inverse LSTM learning features h_t^{backword}, and the two sets of LSTMs are connected into the final feature representation as shown in equation (10).

$$T_i = h_t^{\text{bilstm}} = h_t^{\text{forword}} \oplus h_t^{\text{backword}}, \tag{10}$$

where $\oplus$ is the concatenation operator. This two-layer structure allows the BiLSTM model to fully learn the contextual information of words in the input sequence data.

3.3. Multiscale Feature Fusion. The proposed attention mechanism is inspired by the human visual mechanism [29], and the basic idea is to weaken irrelevant information and increase the attention of focused information during the operation. In this paper, the time-dimensional features $T$and spatial-dimensional features F are fused by using the multiscale feature fusion attention mechanism as shown in Figure 3.

First, the matching matrices FA and FB between the temporal dimensional features F_i and the attribute features T_i of the spatial dimensional feature learning resource are calculated by using equation (11).

$$\begin{cases} FA = F_i \times T_i^T, \\ FB = T_i \times F_i^T. \end{cases} \tag{11}$$

Then, the attention distribution weights w_1 and w_2 of the matching matrices are calculated by using the SoftMax function, respectively, and the weights w are multiplied with the individual scale feature matrices to obtain the attention

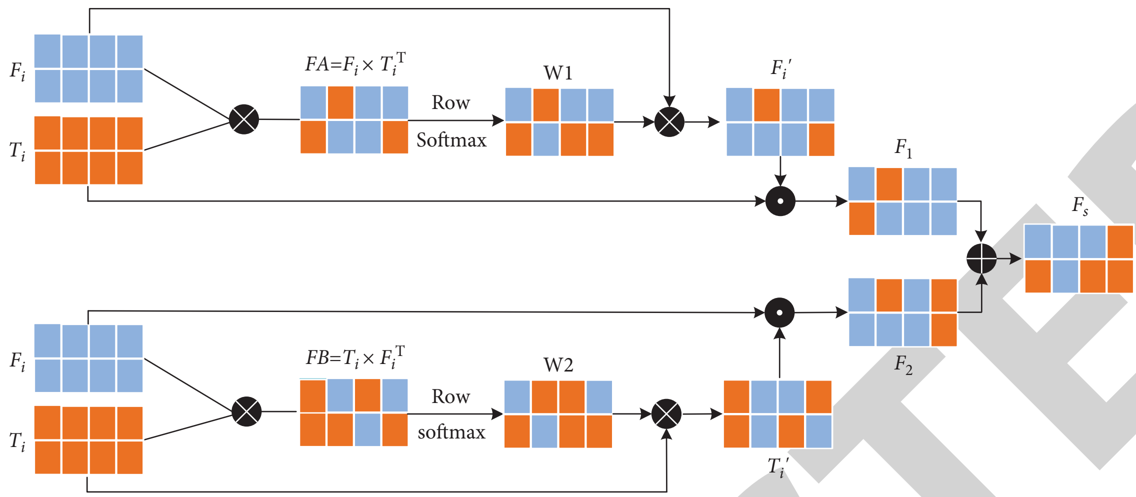


Figure 3: Multiscale feature fusion attention mechanism.

representation matrices F_i' and T_i', which is expressed as shown in equation (12).

$$\begin{cases} F_i' = F_i \times w_1, \\ T_i' = T_i \times w_2. \end{cases} \tag{12}$$

Finally, a multiplicative gating mechanism is used to multiply the attentional representation with another single-scale feature for the corresponding elements to obtain the interscale mutual attention matrices F_1 and F_2, which is expressed as shown in equation (13).

$$\begin{cases} F_1 = T_i \cdot F_i', \\ F_2 = F_i \cdot T_i'. \end{cases} \tag{13}$$

The F_1 and F_2 are jointly operated to obtain the final multiscale fusion features, represented as shown in equation (14).

$$Fs = F_1 \oplus F_2, \tag{14}$$

where $(\times)$ denotes matrix fork multiplication; $(\cdot)$ denotes matrix dot product; and $\oplus$ denotes Cat operation.

3.4. Predicting Scores and Generating Recommendations. The abovementioned multiscale fused features Fs are used as the input of the multilayer perceptron to predict the scores. Here, to achieve end-to-end optimization of the model in the paper, a cross-entropy loss function is used, and the weights of each layer are adjusted and determined by back-propagation. Finally, the prediction results are given quickly by mapping the activation function to the range [0, 1]. The calculations are shown in equations (15) to (17).

$$X_t = wh_t + b, \tag{15}$$

where h_t is the decoder output hidden vector. X_t is the fully connected result.

$$P(y \mid x) = \frac{e^{h(x, y_i)}}{\sum_{j=1}^{n} e^{h(x, y_i)}}, \tag{16}$$

where x is the fully connected result, y is the true description, and P is the SoftMax function.

$$L(\theta) = -\sum_{t=1}^{T} \log p\left(y^t \mid y_{1:t-1}^*\right), \tag{17}$$

where θ denotes the model cross-entropy loss balance parameter.

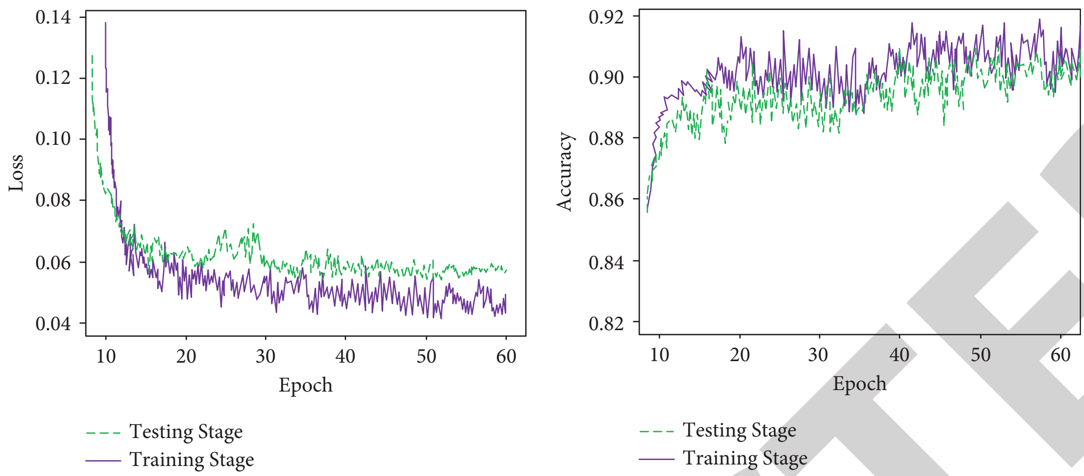
4. Experiments

The experimental running environment is ubuntu16 with 64G RAM and NVIDIA Tesla A100 GPU with 40G graphics memory; PyTorch deep learning framework supporting GPU acceleration is used, and the Cuda environment is NVIDIA CUDA 11.3 and the deep learning acceleration library of cuDNN.

In the experiments, the network was trained by using the stochastic gradient descent algorithm SGD with an initialized learning rate of 0.005 and a learning decay rate of 0.001, and the model training loss curve and accuracy curve are shown in Figure 4. The momentum factor is set to 0.8. In addition, to solve the model overfitting problem, Dropout is introduced to remove some neurons randomly, and Dropout takes the value of 0.5 in the paper. It can be seen from Figure 4 that when the number of model iterations Epoch is 40, the loss curves of both the training and test sets smooth out, and the loss values are below 0.06, indicating that the model has converged.

4.1. The Results and Analysis of Experiments

4.1.1. Comparison of Recommended Effects. To prove the validity of the model in the paper, comparative experiments are conducted under the same environment and evaluation index. The traditional collaborative filtering recommendation algorithm User-CF [30] based on user play records, the Gram-based personalized recommendation algorithm Gram-CF [31] and the collaborative filtering recommendation algorithm FCNN-CF [32] based on user preference statistics are selected as the comparison models. And



(a) (b)

FIGURE 4: Curves of the loss and accuracy during training and testing phases. (a) Loss curves. (b) Accuracy curves.

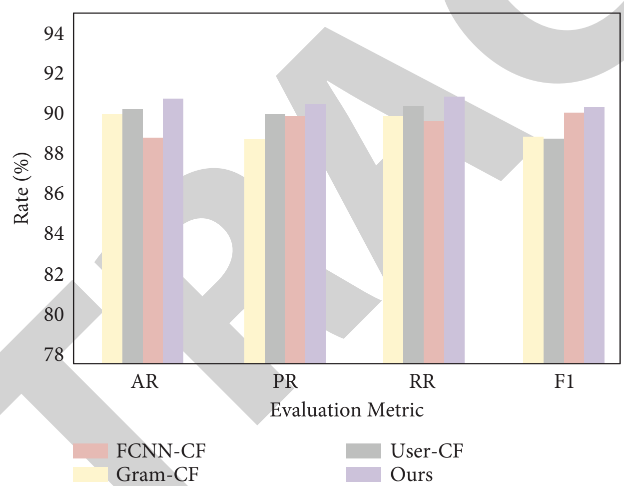


FIGURE 5: Comparison of the recommended performance of different models.

accuracy rate AR (accuracy rate), precision rate PR (precision rate, PR), recall rate RR (recall rate, RR), and $F1$ value are used as evaluation metrics.

From Figure 5, it can be seen that the model in this paper has improved 1.78% (89.7% vs. 91.3%), 1.22% (90.2% vs. 91.3%), and 3.05% (88.6% vs. 91.3%) in terms of accuracy compared to the comparison models FCNN-CF, User-CF, and Gram-CF recommendation models, respectively; in terms of precision, the models in this paper improved by 2.60% (88.5% vs. 90.8%), 0.78% (90.1% vs. 90.8%) and 1.79% (89.2% vs. 90.8%), respectively; in terms of recall, the models in this paper improved by 1.67% (90.1% vs. 91.6%), 1.10% (90.6% vs. 91.6%), respectively. 91.6% and 2.58% (89.3% vs. 91.6%), respectively. In terms of $F1$, the models in this paper improved by 1.58 (88.9% vs. 90.3%), 2.38% (88.2% vs. 90.3%), and 0.22% (90.1% vs. 90.3%), respectively. The abovementioned experimental results verify that the model in this paper has good recommended performance. The reason for this is that this paper performs feature extraction from two dimensions, such as temporal and spatial, and fuses them to strengthen the expressiveness of the features, effectively suppress edge information, and enhance the learning ability of the model for detailed information.

4.1.2. Robustness Testing. Besides, to verify the robustness of the model in this paper, robustness tests are conducted under the same data and experimental platform and compared with three current mainstream models and the experimental results are shown in Figure 6. It can be seen that the advantage of this paper gradually comes to the fore as the number of recommended projects increases. In particular, when the number of recommendations reaches 25, the model in this paper is higher than 81% for both AR, PR, RR, and $F1$ metrics. The abovementioned data further verify the robustness of the model in this paper.

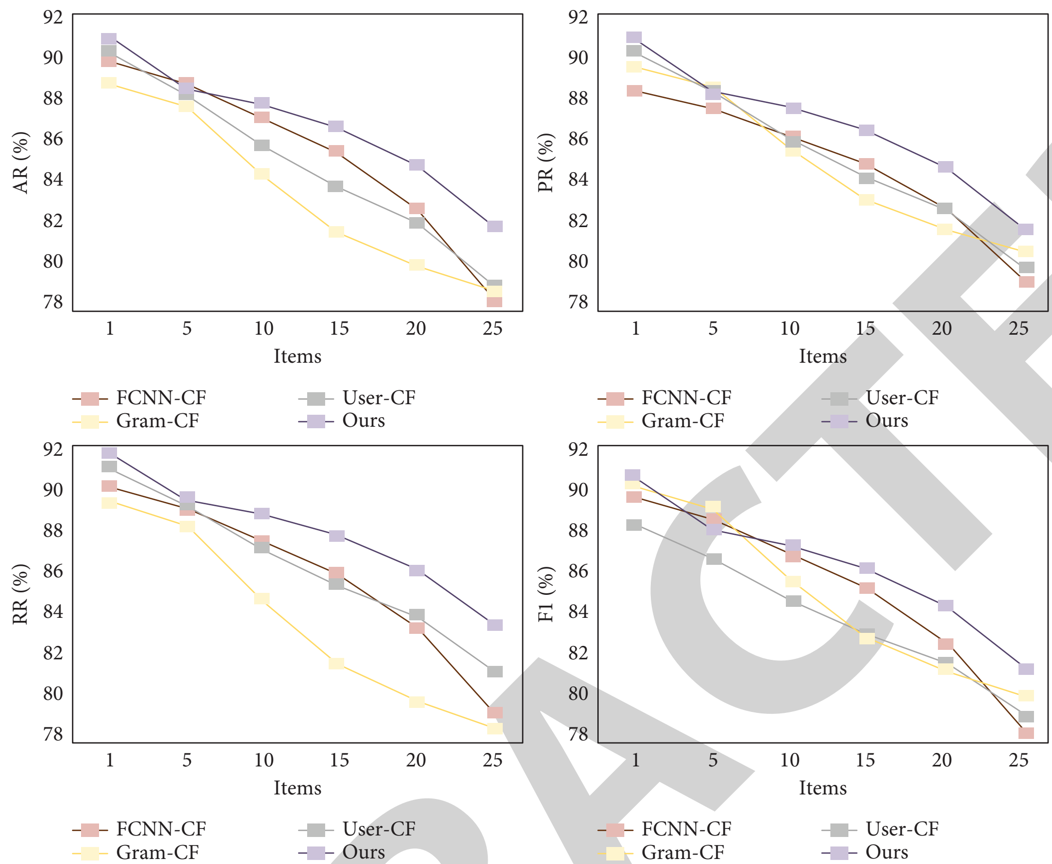


Figure 6: Effect of the number of recommendation items on the recommendation performance.

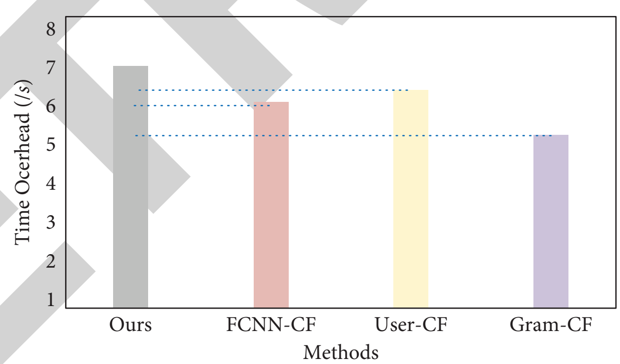


Figure 7: Comparison of the recommendation success rate of different models.

4.1.3. Recommended Real-Time Testing. Another requirement of the recommendation algorithm is that the model is real time. Here, to verify the real-time performance of the recommendation model in this paper, tests are conducted on the same data and environment. The results are shown in Figure 7. It can be seen that the model in this paper can achieve the recommended rate of 6.9video/s, the FCNN-CF model can achieve the recommended rate of 6.1video/s, and the User-CF model can achieve the recommended rate of 6.4video/s, and the Gram CF model can achieve the recommended rate of 5.3video/s. The abovementioned data show that the model in this paper can achieve a better recommendation success rate, mainly because this paper takes into account the personalized service of users and combines the attribute characteristics of learners and learning resources for a personalized recommendation.

4.2. Real-Life Example Analysis. The confusion matrix of this paper's method in six sets of real data is given in Figure 8, where the rows of the matrix denote the real labels and the columns denote the human models generated by the model. From the confusion matrix, it can be concluded that the successful accuracy of generating human morphological models for the six testers in the six sets of experiments was 92.48%, 93.28%, 93.26%, and 93.57%, respectively. The

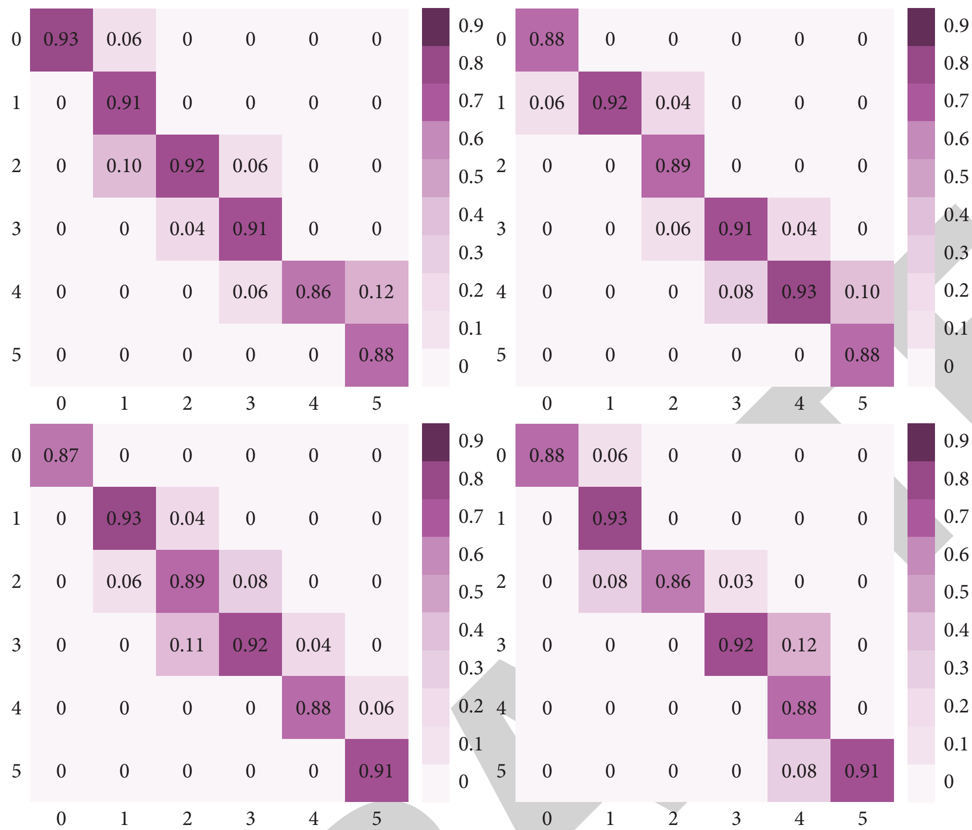


Figure 8: Confusion matrix.

abovementioned data show that the model in this paper tends to perform stably on multiple sets of experimental results and also has good real-time performance, which verifies that the model in this paper has good robustness.

5. Conclusion

In this paper, a recommendation algorithm is proposed in combination with artificial intelligence-related algorithms and applied to the task of enhancing the cognitive ability of thought-politics education, which will greatly facilitate the rapid excavation of the civic culture section of interest to learners in the massive data, quick reading to locate the knowledge of interest and real-time attention. Meanwhile, it helps to improve the cognitive ability of people's minds and has a positive effect on improving the quality of all people. Through testing in a variety of civics-related cultural columns, it is verified that the model in this paper has better real-time performance while maintaining higher detection accuracy and outperforms the mainstream comparison model in several indexes, which has a certain application value.

Data Availability

The datasets used during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The research was supported by: general project of Science Fund National Social, Research on the Construction Logic of the Ruling Discourse System of the Communist Party of China in the New Era (No: 21BDJ055).

References

[1] H. Kamalova and H. Tuychieva, "Improving the spiritual immunological education of academic lyceum students specific issues," *Asian Journal of Multidimensional Research*, vol. 10, no. 4, pp. 616–620, 2021.
[2] S. A. R. Khan, A. Sharif, H. Golpîra, and A Kumar, "A green ideology in Asian emerging economies: from environmental policy and sustainable development," *Sustainable Development*, vol. 27, no. 6, pp. 1063–1075, 2019.
[3] Q. J. Wang, G. F. Feng, Y. E. Chen, J. Wen, and C. P. Chang, "The impacts of government ideology on innovation: what are the main implications?" *Research Policy*, vol. 48, no. 5, pp. 1232–1247, 2019.
[4] S. Canagarajah and S. Dovchin, "The everyday politics of translingualism as a resistant practice," *International Journal of Multilingualism*, vol. 16, no. 2, pp. 127–144, 2019.
[5] J. D. García-Castro, G. B. Willis, and R. Rodríguez-Bailón, "I know people who can and who cannot: a measure of the

Hindawi
Mobile Information Systems
Volume 2022, Article ID 3899472, 9 pages
https://doi.org/10.1155/2022/3899472

Research Article

Study and Design of Scalable Multimedia Service System

Ziyu Zhao

Yeungnam University, Gyeongsan, Gyeongbu 38541, Republic of Korea

Correspondence should be addressed to Ziyu Zhao; h7276@hdc.edu.cn

Received 22 August 2022; Revised 6 September 2022; Accepted 13 September 2022; Published 20 September 2022

Academic Editor: R. Mo

In recent years, with the continuous advancement of multimedia technology and network technology, network-based multimedia applications are increasing day by day, both in scale and in the number of users. However, the randomness, time-varying, and complexity of user behavior and network make it impossible to guarantee the service quality (bandwidth, delay, etc.) of multimedia services, thus making it difficult to meet the demands of growing users for multimedia services. While meeting the application requirements, how to design and develop a scalable multimedia service system has become one of the key issues to support good multimedia services and better develop multimedia service services. In addition, the surge in user demand comes with various challenges. High-definition and smooth multimedia services require stable bandwidth. However, the randomness, time-varying, and complexity of user behaviors and network states usually make it difficult to guarantee the quality of services. It is of practical significance to study how to overcome these problems, and it is also a hotspot of current research. This paper mainly discusses the extensibility elements of the multimedia service system and puts forward the relationship between the extensibility of various elements and some design principles of the extensibility. In order to overcome the influence of random variation of the wireless channel on video playback, this paper adopts scalable coding technology. That is, through the analysis of the playback process, the underflow probability is proposed and defined, thereby establishing a constrained optimization problem. Finally, this paper also proposes the multimedia adaptive transmission in the network to do the optimization.

1. Introduction

Today, everyone's material living standards are improving day by day, and people are not only satisfied with the enjoyment brought by material things but more and more people have begun to pay attention to improving the quality of spiritual life. Education and entertainment, as two main ways in people's spiritual life, have played an important role in daily life [1, 2]. No matter whether it is educational service or entertainment service, the main service form is to provide multimedia service. The emergence of multimedia services, especially the emergence of video and audio services, is not like the previous e-mail, simple WEB and FTP services. Its service form begins to shift from reliable interaction to real-time interaction, which puts forward higher requirements for network performance [3, 4].

With the rapid development of multimedia technology, network technology, and communication technology, the network multimedia service technology based on streaming media transmission has been widely used [5, 6]. With the ever-increasing demands of network media users, the existing media service technologies are challenged. For example, the media resource management system is only aimed at a single media resource owner, video-on-demand (VOD) users need to register multiple times to get corresponding services, and media resource owners all want to serve more VOD users with their media resources. Obviously, the existing media resource service technology can no longer meet the requirements [7, 8]. A scalable system should be able to provide better performance and more functions. The overall computing power of the system should scale proportionally with the increase in resources. Ideally, the rate of growth is linear.

With China's emphasis on education and increased investment in school software and hardware resources, the continuous improvement of school teaching quality and the continuous increase in the number of students have placed high requirements on school teaching equipment. Therefore, many colleges and universities have added multimedia equipment for teaching. With the increase of equipment, how to manage equipment has become a big problem, especially the unified remote management of equipment has put forward higher requirements for schools. How to standardize the school's multimedia equipment, informatization, and automatic management of equipment management is very important. Hence, the multimedia teaching equipment can play a greater role; for example, remote control of multimedia teaching equipment is to centrally manage all classroom multimedia equipment. It can also reduce the manual workload, improve work efficiency, improve the standard of school equipment management, effectively improve the school's management of multimedia teaching equipment, and effectively improve the school's multimedia equipment [9, 10]. Management can reduce the possible mistakes of manual management of equipment, causing delays in class or failing to turn on equipment on time, etc. The use of remote centralized control can effectively reduce manual errors and improve equipment utilization. It plays an important role in the improvement of equipment management in colleges and universities [11, 12]. According to the characteristics of traditional media publishing, multimedia publishing has some inherent properties, such as point-to-face publishing, a large amount of published data, and little feedback. These characteristics determine some network technologies for media services.

At present, many multimedia teaching equipment in some universities relies on manual management because equipment management is not standardized, and equipment management information is low. With the increase of multimedia teaching equipment, the manual management method is becoming more and more inappropriate to better manage the equipment. In fact, there are many nonstandard management, which need to be improved. Statistical data are also prone to errors, and it is also very difficult to share the statistical data and information for the second time [13, 14]. Without a unified equipment management system, the information sharing of multimedia teaching equipment is very cumbersome, and even cannot be shared, resulting in information redundancy and junk information. It is urgent to have a unified platform to manage equipment information [15, 16].

In the past ten years, with the maturity of multimedia technology and information technology in teaching, the increase of national support, and investment in multimedia teaching, multimedia classrooms have been popularized in major universities. Multimedia teaching can vividly display sounds and animations, which greatly improves the enthusiasm of students in class and improves teaching efficiency. In the early multimedia classrooms, various teaching devices were simply connected, and each device had its own switch, which was inconvenient to use and not conducive to safe use. The solution for many colleges and universities is to install a central control system to manage the equipment of the entire school in a unified manner, efficiently manage various multimedia equipment, achieve the effect of centralized management, and realize the unified management of multimedia classrooms [17, 18]. The multimedia central controller solves the problem of using the multimedia classroom but still cannot solve the management of the multimedia classroom. Nowadays, the multimedia classrooms in major universities are distributed in different teaching buildings, and there are geographical distances [19, 20]. There are many floors up and down in the same teaching building, which is inconvenient to manage. Nowadays, many schools use manpower to manage and open the central control cabinet and turn on the equipment on time before class. However, one person can manage up to 20 multimedia classrooms. For example, in different teaching buildings, many multimedia classroom managers are required. In view of the above difficulties in management, a multimedia central management system based on a campus network can be built to solve the above problems. Cross-segment management can also solve the unified management of different campuses, and even managers can manage through the network at home [21, 22].

From the analysis of the existing media service site architecture, it is found that the media resource management systems of these media service sites have the following shortcomings. Firstly, the limitations of the media resource management system services, the media service system, and the site are bound together. Different media service sites use their own media resource management system to manage and publish their media resources. Secondly, the scalability of the media resource management system is poor. That is, it is difficult to add media servers and media resource owners to the existing system, and it is necessary for the existing system architecture to be greatly changed. Thirdly, it is inconvenient for users to order on demand, and the resources and services provided by the media resource management systems of different sites are independent of each other. When the media service site where the VOD user is located cannot provide the required service, he has to go to another media service site to register to obtain the required service [23, 24].

The existing media resource management system is oriented to the service of a single media resource owner owned by the system and cannot achieve unified management of the resources of multiple media resource owners. Extending to more types of resources is more difficult. The requirement of the scalable media service technology is to establish a service-expandable media resource management system, which allows to dynamically add media resource owners to the system, and the media resource owners can also add the required media servers at will, so as to realize the Internet environment. To achieve the goal of managing the resources of multiple different media resource owners and being able to manage their own media resources in the network environment, it is necessary to effectively distinguish and utilize the relevant attributes of the media resource owners. From the perspective of media resource owners, each media resource station is equivalent to a traditional resource publishing point with independent service capabilities, with its own media resources, management system, and service mechanism. The media resource owner can use the media resource station to define the scope of their own media resource management. All media resource stations form a huge capacity media resource library, which brings more resource choices to VOD users [25, 26].

The scalable system has three expansion aspects: resources, applications, and technology replacement. From a broad perspective, resources include hardware resources and software resources, and hardware resources include the number of CPUs, storage devices, and IO devices. Software resources include, for example, operating systems and application software. In the era of increasingly complex Internet applications, the concept of resources extends to the level of user needs. Not only resources have to be extensible but applications must also be extensible. That is, when the same program runs on a scalable system, its performance improves proportionally with the scale. Two important considerations are scalability of machine size and scalability of problem size. Technology generation scalability is a local concept rather than a global concept. It mainly refers to that the computer system can improve the performance or function of the system by means of upgrading components on the basis of maintaining the original usability with the change and replacement of technology. The upgrade here may be to replace the CPU with a higher frequency, or to increase the amount of memory in the system, or to upgrade the operating system to a newer version [27].

2. Multimedia Adaptive Transmission in Network

2.1. Background of Multimedia Adaptive Transmission. The quality of video services is affected by various subjective factors, which vary from person to person, time to place, and place. Therefore, subjective evaluation is too complicated and difficult to implement. In practical applications, an easy-to-implement objective quality assessment method is required. At the same time, in order to ensure the accuracy of the results, it is required that the results of subjective evaluation and the results of the objective evaluation have good correlation or consistency under the same test environment. In this paper, we mainly consider the impact of different adaptive algorithms on service quality. Therefore, usually under the same premise of other factors, that is, except for the differences of adaptive algorithms, other influencing factors at the service level, user level, and environment level are kept unchanged to compare the algorithm performance. This allows us to define some objective quality of service metrics for adaptive multimedia applications. Intuitively, users usually want to watch videos with good image quality and smooth playback. Ignoring the influence of users and the environment, when the video sources used are the same, the service quality mainly depends on factors such as video quality, smoothness of playback, and smoothness of the video.

For the problem of scalable video transmission in wireless networks, the core is how to design an efficient and reliable code stream adaptive algorithm. In practical applications, from the user's point of view, it is generally hoped that when there is enough video data in the buffer, the allowable received video bit rate is greater than the currently available bandwidth, so as to obtain better video quality. Similarly, when the buffer data are insufficient, it is hoped that the received video bit rate is smaller than the currently available bandwidth to buffer more video data, thereby avoiding playback interruption caused by sudden network deterioration. Therefore, to design an effective code stream adaptive algorithm, it is necessary to comprehensively consider the changes in the wireless channel and the state of the buffer. However, channel variations in practical systems are usually unpredictable. To achieve adaptive video transmission, we design a scalable video transmission system. During the video transmission process, the mobile terminal dynamically selects the number of layers of the scalable video according to the bit rate adaptive algorithm and sends a bit rate adjustment request to the video server. The video server dynamically extracts the corresponding code stream from the scalable video file according to the request of the terminal. To overcome the stochastic nature of wireless channels (such as random interference, user movement, and multipath), we model the problem using a probability-based approach.

2.2. Multimedia Adaptive Transmission. Suppose the encoded video contains a base layer and L-1 enhancement layers. In the nth time slot, the number of layers of the video received by the user is expressed as

$$l_n = \{1, 2, 3, \ldots, L\}, \tag{1}$$

where L is the number of layers.

The current buffer size is Q_n, and the number of video frames arriving is expressed as

$$A_n = \{0, 1, 2, 3, \ldots, A\}. \tag{2}$$

Among them, A is the number of video frames.

It is assumed that the frame arrival process is an independent and identically distributed random process. The change process of the receiving buffer queue can be described as

$$Q_{n+1} = \max\{Q_n - 1\} + A_n. \tag{3}$$

Among them, Q is the cache. In order to measure the influence of various random factors, the underflow probability at time n is defined; that is, the playback interruption probability is

$$p = P(Q_n < Q_L). \tag{4}$$

Using the underflow probability, the requirement of no playback interruption can be transformed into a constraint that the underflow probability must be lower than the requirement. Thus, a constrained optimization problem can be constructed. That is, in order to maximize the video bit rate while ensuring uninterrupted playback, the equation is as follows:

$$\begin{aligned}&\max, l_n,\\&\text{s.t. } p = P\left(Q_n < Q_L\right) < q,\\&l_n \le L_m,\end{aligned} \tag{5}$$

where q represents the tolerable underflow probability.

For a given time slot k, according to (3), the variation of the play queue can be obtained as

$$I_k = 1 - A_k. \tag{6}$$

Among them, I_k measures the mismatch between the current video bit rate and the channel throughput. Predict the channel change in N time slots in the future; that is, the prediction interval length is the change from time n to $n+N$ play queue:

$$I^{n+N} = \sum_{k=1}^{N} I_{n+k}. \tag{7}$$

Therefore, the length of the playback queue at time $n+N$ is

$$Q_{n+N} = Q_n - \sum_{k=1}^{N} I_{n+k}. \tag{8}$$

The probability of underflow at time $n+N$ is

$$p = P\left(Q_n - \sum_{k=1}^{N} I_{n+k} < Q_L\right). \tag{9}$$

The expectation that defines the average change of the play queue for N time slots in the future is

$$E\left[\frac{\sum_{k=1}^{N} I_{n+k}}{N}\right]. \tag{10}$$

In practical applications, the average frame rate can be used to characterize fluency:

$$f_T = \frac{S(T)}{T}. \tag{11}$$

Among them, $S(T)$ represents the total number of frames when a video with a length of L minutes is played, and T is the actual playback time. Statistics on the interruption frequency or average interruption time during the playback process can also represent the smoothness of the playback. That is, the interrupt frequency is as follows:

$$f_C = \frac{C(T)}{T}. \tag{12}$$

Among them, $C(T)$ represents the total number of interruptions. Figure 1 shows the random variation of the bandwidth of the 3G downlink channel with time. Therefore, it cannot provide a reliable bandwidth guarantee for video transmission in the wireless network.

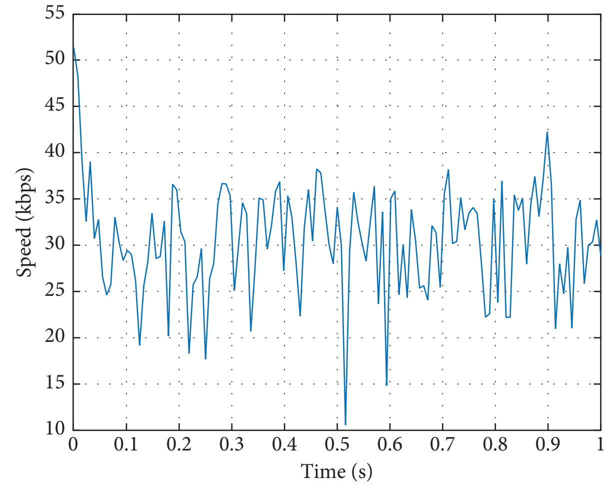


Figure 1: 3G downlink channel throughput over time.

3. Research and Design of Scalable Multimedia System

3.1. Multimedia System. If resources can be increased to meet increasing performance and functional requirements, or resources can be reduced to reduce costs, from the perspective of hardware and software, we can call a computer system that meets this characteristic scalable. The multimedia service system refers to a service system that provides multimedia content such as video, audio, animation, text, 4D. At present, the main multimedia service systems include Huawei's distance education system, Lanyu Video Network Multimedia Acquisition, Editing, Broadcasting and Publishing System, Shihan Technology Distance Education System, and Jintong Network Multimedia Teaching System. Among these systems, the distance education system provides a complete set of solutions for server, network, and client. The distance education system of Shihan Technology mainly provides the server system, and the Lanyu Video Network Multimedia Collection, Editing, Broadcasting, and Publishing System provide high-quality video services. The current multimedia system design mainly focuses on network access, multimedia information production management, and user access control. For example, the core layer of the system in Huawei's solution supports the access of real-time interactive video terminals such as IP, El, ISDN, and V35 and completes the exchange of images, voice, and data in distance education services. The user access layer is used to set up professional electronic classrooms and provide powerful real-time network video teaching functions. Another example is the Lanyu Video Network Multimedia Collection, Editing, Broadcasting, and Publishing System, which mainly focuses on the production and management of streaming media sources, access to various access devices (such as set-top boxes, Cable Modem access) and other charging strategies. In the design system, the current multimedia service system is mostly based on the layering of solutions. For example, Huawei's distance education system is divided into the management and control layer, system core layer, and teaching user access layer. The last two layers are network transmission and access. Huawei divides the multimedia service system into a business support layer, a media exchange control layer, and a user access layer. These layers focus on the overall network solution and do not involve the hierarchical division of the multimedia service management system. In terms of expansibility, these multimedia service systems focus more on the expansion of multimedia content formats and the expansion of access devices or the expansion of a certain aspect. For example, the Jintong network multimedia teaching system supports customizable user interface expansion, supports PC, STB, and POA terminal access type expansion, and supports secondary development (including teaching management and courseware production, playback). Another example is the Lanyu Video Network Multimedia Collection, Editing, Broadcasting, and Publishing System which supports the expansion of the video server but does not support the expansion of the text service. From the above several current typical multimedia service systems, we can see that the design of the current multimedia system does not clearly propose the management elements and scalable elements in the multimedia system and focuses more on the network extension and services of the multimedia system. The overall construction elements of the current multimedia service system have been very clear, but the extensible elements of the multimedia service system have not yet been systematically proposed. But clarifying the extensible elements has great guiding significance for designing an extensible multimedia service system.

The multimedia information service system supports four roles: end user, enterprise/merchant, carrier administrator, and call center agent. Terminal users can log in to the service platform by mobile phone and browse the information they need through comprehensive search, categorical search, and navigation functions. For registered users, they can further choose to bookmark and subscribe to the information of interested enterprises/merchants and publish relevant comments for the enterprises/merchants. In addition, terminal users can further open the function of personal trading. After a successful opening, personal trading users can publish their own information on the video navigation platform, which can be classified as real estate trading, rent seeking, video job hunting, personal goods flea market, etc. Enterprises/merchants can apply for the multimedia information release service according to the relevant regulations of telecom operators. After the application is successful, enterprises/merchants can release their multimedia information on the system. In addition, according to their own characteristics, enterprises/merchants can also cooperate with Unicom to provide users with in-depth value-added services, such as booking hotels and booking air tickets. According to the needs of developing multimedia information services, operators have administrators with different permissions. Different administrators use different accounts to log in to the system and display different management portal pages based on the rights of their accounts to perform different management functions. The call center agent can log in to the multimedia information service system using the assigned account and password and send the enterprise/merchant information to the terminal user in the form of a message based on the user's query requirements.

The multimedia system in this paper refers to the sending center system that mainly provides video and audio services and supplements text services. The system should not only provide reliable and high-quality media services but also provide a flexible and extensible system framework, so that the system can be easily expanded according to business needs. Each system has different scalable resources, and multimedia systems also have corresponding resources. Analyzing and clearly defining these resources is the basis of the scalable system design, and the analysis of the expansion motive force of the multimedia system is the basic basis of the system's scalability. Finally, the scalability of the business management software design is analyzed. The scalable system has the following characteristics, including the scalable element is the resource, extensible aspect package hardware and software, the drive for scalability comes from performance, functionality, and cost requirements, and the scalable way is to grow and shrink.

The multimedia service system is a system. Although it can provide a single super-capable server and a super-fast internal bus to support this system, this is not the final form of system design, so the multimedia service system must be interconnected with multiple servers, and services are provided in the form of network data, so a network layer must be required. In addition to providing services, the multimedia service system must provide some auxiliary services, such as authorization services, which should be accurately called service support, so there should be a service support layer and a service layer. Provide many services such as on-demand and live broadcast, organize and manage services reasonably, and then form an operable business, and there must be a business layer on the service layer. According to the principle of layered design, we divide the entire multimedia system into four layers, as shown in Figure 2. The main function of the network layer is to transmit various multimedia services under a certain network bandwidth and network equipment. Network bandwidth and network equipment become the two major resources of the network layer. The functions of the multimedia information service system for end-users include self-browsing function, manual service function, SMS receiving function, and welcome word function. The services for the call center include obtaining user's personal information, fashion shopping information query, wonderful tourism information query, and flavor food information query. The administrator provides information management, log management, and statistics management services.

The main function of the service bearer layer is to provide network resource allocation and scheduling, and a safe and reliable mechanism for various services. This layer converts network resources into resources of this layer for upward provision and shields the visibility of lower-layer resources from the upper layer.

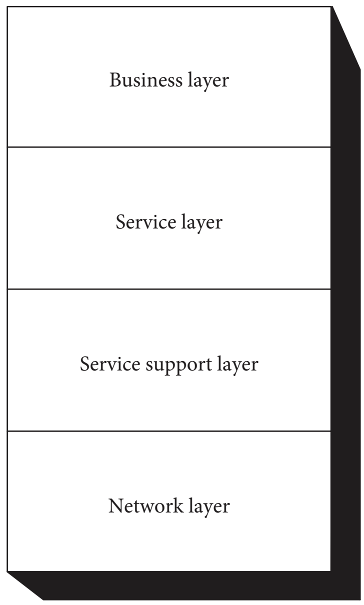


Figure 2: Multimedia service system hierarchy.

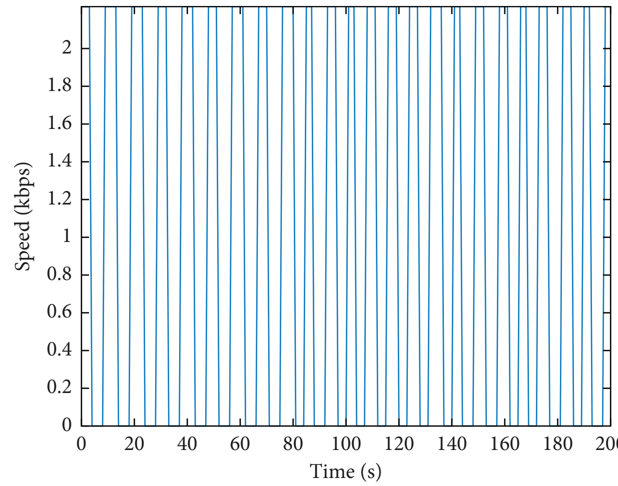


Figure 3: Transmission speed based on the probability of underflow.

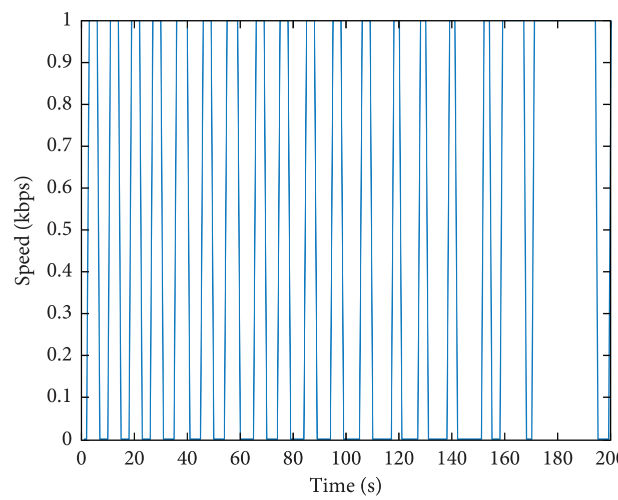


Figure 4: Transmission speed of the conventional method.

3.2. Discussion. The network layer corresponds to the transport layer and all the following layers of the OSI seven-layer protocol model, including server equipment and network connection equipment. The service support layer is to provide functional support for multimedia services and provide network layer resource scheduling functions. The service layer mainly refers to services corresponding to multimedia systems, such as on-demand, live broadcast, and file download. The business layer is corresponding to multimedia applications, such as distance education and entertainment services. Each layer has corresponding operation elements, and these elements are all resources corresponding to the upper layer. However, in order to strictly distinguish resources, we do not call all elements resources. Figure 3 is the transmission rate using the algorithm based on the probability of underflow. In contrast, the transmission rate with the conventional method is also shown in Figure 4, where it can be seen that the proposed method has a faster speed.

Multimedia services mainly have four transmission forms: reliable publishing, real-time publishing, reliable interaction, and real-time interaction. The design of the functional unit also needs to support these transport forms. In addition to the support for the four transmission forms, the business needs to be managed. Therefore, according to the different business support functions, it can be divided into three types of units: business unit, business support unit, and operation management unit. The business unit is the unit that directly supports the four transmission modes of multimedia services. The basic unit includes a real-time stream sending unit, a non-real-time stream sending unit, a reliable data sending unit, and an interactive access unit. The real-time streaming unit provides real-time-one-way service delivery mode services. The non-real-time streaming unit provides services in a non-real-time-one-way service transmission mode, which is unreliable. The reliable data transmission unit provides services in a reliable-one-way service transmission mode. The interactive access unit provides services in the reliable-interactive service transmission mode, and together with the real-time streaming unit, it provides services in the real-time-interactive service mode. Figure 5 shows the video quality at different channel qualities. As the channel quality improves, the video quality gets better. It shows that the algorithm can realize the adaptation of bit rate and video quality under different channel qualities. Figure 6 shows the interruption frequency of video on mutated channels. As can be seen, the interrupt varies with different conditions, and its maximum is nearly 1.6, which also can be explained in Figure 7.

The business support unit does not directly support the four transmission modes of multimedia services but provides a unit with auxiliary functions for service access. The basic unit has a program navigation sending unit and an authorized sending unit. The program guide sending unit sends the receiving information describing the service entity and the program. The authorization sending unit sends the

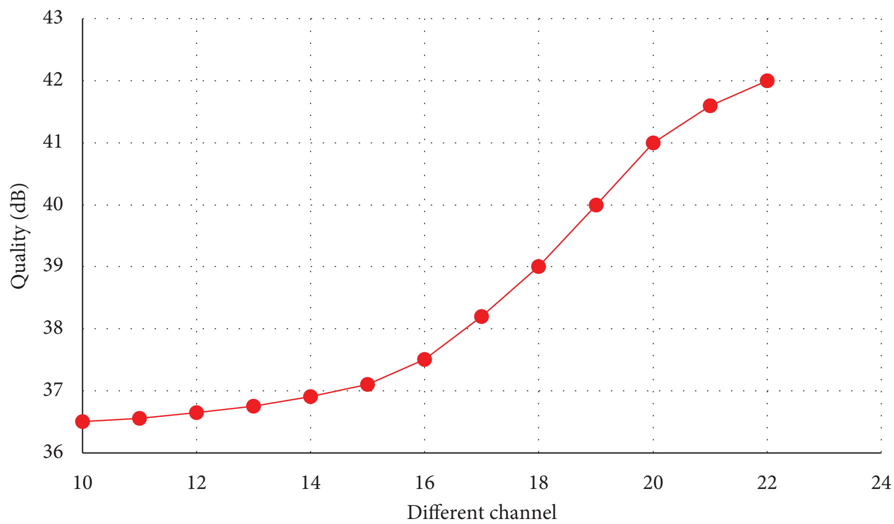


Figure 5: Video quality at different channel qualities.

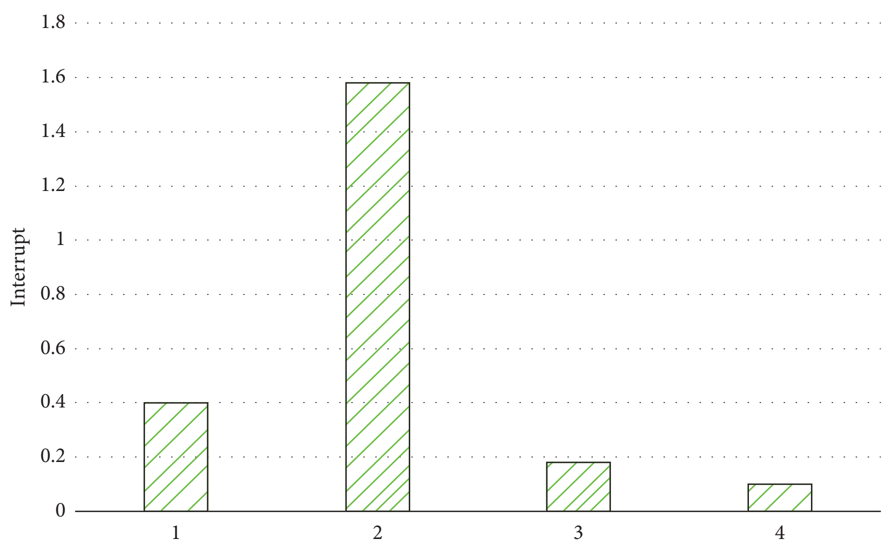


Figure 6: Interruption frequency of video on mutated channels.

service access control information. The operation management unit performs hardware management and broadcast monitoring of the entire system. It does not directly participate in service transmission but plays a vital role in the normal development of services and the management of business units and business support units. The basic unit includes a broadcast control unit and a flow monitoring unit. The broadcast control unit allocates, dispatches, and manages system resources, and manages and commands the status of each business unit and business support unit. The flow monitoring unit monitors the network flow of the entire system in real time. Strictly distinguishing these units can reflect two expansion advantages, including the scalability of incremental business development and the scalability of service capabilities by simply superimposing the same functional unit. Business development always starts with low risk and low investment, and there should be multiple options for starting a business. In the early stage of business development, for services that only require a reliable-one-way service transmission mode, only reliable data transmission units, business support units, and operation management units are required; for real-time-one-way service transmission mode, only a real-time stream sending unit, business support unit, and operation unit are required. For the later stage of subservice development-integrated services, all business units, including business support units and operation units, are needed. The evaluated data are plotted in Figure 8.

If a business needs to add streaming media services to the pure data business, it is only necessary to add a real-time stream sending unit and a non-real-time stream sending unit to the business unit. If a service is to be authorized and controlled, an authorization sending unit can be added to the service support unit. If you need to support interactive services, adding an interactive access unit can meet the requirements.

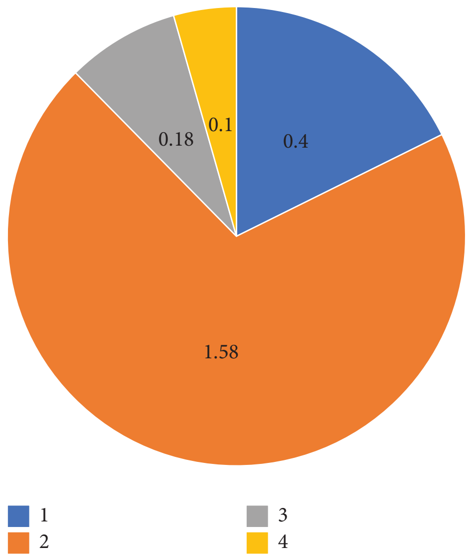


Figure 7: Percentage figure of the interruption frequency.

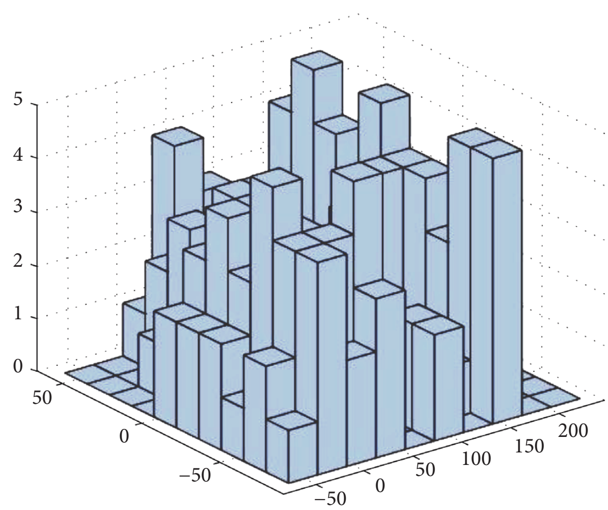


Figure 8: Evaluated data.

The expansion of business service types can not only expand different units (as described above) to achieve the expansion of service units but also expand the same business units according to the capabilities of each business unit to expand services. For example, it is necessary to expand the reliable data transmission service. Under the condition of the limited capacity of one reliable data transmission unit, another reliable data transmission unit can be added to expand the service capacity. Scalability not only refers to the scalability of software functions but also refers to the scalability of physical hardware and connections, and functional units also have the scalability of their physical connection methods. Considering the expansion of service capacity, the network traffic characteristics of different functional units, the sharable characteristics of functional units, and the physical connection method of functional classification placement are adopted for the connection of the entire functional unit. That is, units with similar network traffic characteristics and service capacity scalability constitute a physical connection unit, while the combination of multiple physical connection units constitutes the system connection structure of the entire functional unit. However, the current load balancing scheduling is also considered based on the simplicity of the algorithm and the quickness of the scheduling, and the goal of full utilization of resources is still avoided.

4. Conclusion

Under the current situation of network service turning to multimedia service, this paper analyzes the actual use of multimedia service system. Aiming at the lack of clear definition and distinction of scalable elements in current multimedia systems, this paper proposes a design framework for scalable multimedia systems. This paper analyzes the extensible elements of the multimedia service system, proposes a hierarchical system structure, and proposes extensible resources for each level. Among them, the main extension elements are concentrated in the service support layer. Aiming at the expansion analysis of the software and hardware system of the functional units and channels of the service support layer, the embodiment of the scalability under such a design is proposed, and the expansion of business services is the driving force of the system scalability.

However, the scalability introduced in this paper is still lacking in consideration of the scalability of the cost; that is, the expansion cost needs to be controlled and optimized at the same time of expansion.

Data Availability

The datasets supporting the conclusions of this article are available from the corresponding author upon reasonable request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] X. Y. Zhong, X. Tian, K. D. Thilak, and M. Anbarasan, “Machine learning-based multimedia services for business model evaluation,” *Computers & Electrical Engineering*, vol. 97, Article ID 107605, 9 pages, 2022.

[2] R. Palash, S. Sarker, M. Abdur Razzaque et al., “AI-enabled mobile multimedia service instance placement scheme in mobile edge computing,” *Computer Networks*, vol. 182, pp. 232–241, 2020.

[3] R. C. Park and E. J. Hong, “Urban traffic accident risk prediction for knowledge-based mobile multimedia service,” *Personal and Ubiquitous Computing*, vol. 26, pp. 1–11, 2020.

[4] Y. Jeon, “RF channel management scheme for seamless multimedia service in 802.11 wireless networks,” *Multimedia Tools and Applications*, vol. 79, no. 37-38, pp. 1–18, 2020.

[5] J. X. Zhou, “Customization of on-site assembly services by integrating the internet of things and BIM technologies in modular integrated construction,” *Automation in Construction*, vol. 126, pp. 321–332, 2021.

[6] F. J. Siddiqui, H. Ashraf, and A. Ullah, "Dual server based security system for multimedia Services in Next Generation Networks," *Multimedia Tools and Applications*, vol. 79, no. 11-12, pp. 7299–7318, 2020.
[7] T. Cao, "Reliable and efficient multimedia service optimization for edge computing-based 5G networks: game theoretic approaches," *IEEE Transactions on Network and Service Management*, vol. 99, p. 1, 2020.
[8] Y. Woo and I. Kim, "Framework for multimedia service using multicast in CVCN network," *Journal of Advanced Information Technology And Convergence*, vol. 9, no. 2, pp. 55–63, 2019.
[9] T. Cao, "Credible and economic multimedia service optimization based on game theoretic in hybrid cloud networks," *Transactions on Emerging Telecommunications Technologies*, vol. 33, pp. 1–8, 2019.
[10] H. Zhang, "Trade openness and green total factor productivity in China: the role of ICT-based digital trade," *Frontiers in Environmental Science*, vol. 1, pp. 98–102, 2021.
[11] B. Zhang, Y. Chen, J. L. Yu, and Z. Han, "An indirect-reciprocity-based incentive framework for multimedia service through device-to-device multicast," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 11, pp. 10969–10980, 2019.
[12] H. Bayrak and G. Nur Yilmaz, "A depth perception evaluation metric for immersive user experience towards 3D multimedia services," *Multimedia Systems*, vol. 25, no. 3, pp. 253–261, 2019.
[13] A. Ominike, "A quality of experience hexagram model for mobile network operators' multimedia services and applications," *International Journal of Computer*, vol. 34, no. 1, pp. 1–8, 2019.
[14] J. Kharel and S. Y. Shin, "Multimedia service utilizing hierarchical fog computing for vehicular networks," *Multimedia Tools and Applications*, vol. 78, no. 7, pp. 9405–9428, 2019.
[15] K. Yaser, "Dynamic framework to mining Internet of Things for multimedia services," *Expert Systems*, vol. 1, Article ID e12404, 2019.
[16] H. Kim, H. J. Kwon, and K. K. Kim, "Modified cyber kill chain model for multimedia service environments," *Multimedia Tools and Applications*, vol. 78, no. 3, pp. 3153–3170, 2019.
[17] J. Du and Z. Y. Yu, "Building a cross-border E-commerce ecosystem model based on block chain + internet of things," *Security and Communication Networks*, vol. 2021, Article ID 6451721, 45 pages, 2021.
[18] C. Singhal and S. De, "eM-SON: efficient multimedia service over self-organising Wi-Fi network," *IET Networks*, vol. 7, no. 4, pp. 181–189, 2018.
[19] S. Kumawat and M. Ravi Kumar, "Design of Variable Weight code for multimedia service in SAC–OCDMA systems," *IET Optoelectronics*, vol. 12, no. 2, pp. 56–64, 2018.
[20] H. Dang, "Design of long-distance multimedia physical education intelligent system," *Basic and Clinical Pharmacology and Toxicology*, vol. 127, p. 140, 2020.
[21] J. Lu and T. Valaec, "College multimedia teaching system devise and application evaluation," *International Journal of Electrical Engineering Education*, vol. 1, Article ID 002072092092847, 2020.
[22] Y. Liu, L. Zhu, and F. Liu, "Design of multimedia education network security and intrusion detection system," *Multimedia Tools and Applications*, vol. 79, no. 25-26, pp. 18801–18814, 2020.
[23] H. Zheng and Z. Perez, "Design of multimedia engineering teaching system based on internet of things technology," *International Journal of Continuing Engineering Education and Life Long Learning*, vol. 29, no. 4, pp. 1–7, 2019.
[24] W. Liang, "Scene art design based on human-computer interaction and multimedia information system: an interactive perspective," *Multimedia Tools and Applications*, vol. 78, no. 4, pp. 4767–4785, 2019.
[25] P. Li, "Design of meyer's theory-based high quality piano multi-media system," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 12, no. 01, p. 95, 2017.
[26] S. Harris Heidi and M. Greer, "Using multimedia for instructor presence in purposeful pedagogy-driven online technical writing courses," *Journal of Technical Writing and Communication*, vol. 52, no. 1, pp. 110–131, 2022.
[27] K. C. Moen, "The impact of multi-media presentation format: student perceptions and learning outcomes," *Scholarship of Teaching and Learning in Psychology*, vol. 7, no. 4, pp. 278–287, 2021.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 6131572, 12 pages
https://doi.org/10.1155/2022/6131572

Research Article

A Machine Learning Assessment System for Spoken English Based on Linear Predictive Coding

Lu Wang

School of Liberal Education, Chengdu Jincheng College, Sichuan, Chengdu 611731, China

Correspondence should be addressed to Lu Wang; wanglu@scujcc.edu.cn

Received 3 August 2022; Revised 18 August 2022; Accepted 30 August 2022; Published 20 September 2022

Academic Editor: R. Mo

In the teaching of English, there is an increasing focus on practical communication skills. As a result, the speaking test component has received more and more attention from education experts. With the rapid development of modern computer technology and network technology, the use of computers to assess the quality of spoken English has become a hot topic of research in related fields at present. A machine learning assessment system based on linear predictive coding is proposed in order to achieve automatic scoring of spoken English tests. First, the principle of linear predictive coding and decoding is analyzed, and the traditional linear predictive coding and decoding algorithm is improved by using hybrid excitation instead of the traditional binary excitation. Second, the overall structure of the machine learning assessment system is designed, which mainly includes division into four modules: acoustic model acquisition module, speech recognition module, standard pronunciation transcription module, and decision module. Then, the speech recognition module is implemented by an improved linear predictive speech coding method to acquire the feature parameters of the speech signal and generate the speech feature vector. Finally, the convolutional neural network algorithm is used to train the speech features so as to implement the acoustic model acquisition module. The experimental results show that the improved linear predictive speech coding method yields more natural and higher intelligibility speech signals. The designed machine learning evaluation system is able to accurately detect information about the quality of the learner's pronunciation.

1. Introduction

The focus of modern English language teaching is on the development of students' general application skills, including listening and reading skills. Among these, speaking training and speaking assessment have received increasing attention. There are generally two types of assessment for speaking tests: an automated assessment and a manual assessment by experts. With the continuous development of random computer technology, automated assessment of speaking tests is beginning to be used in a variety of industries [1–6]. For example, speaking assessment systems can be used during telephone interviews to automatically score the English proficiency of interviewees. In addition, online teaching application scenarios in the education industry can use speaking assessment systems to automate the scoring of students' speaking quality. Automated speaking assessment systems can give objective scores based on the test taker's performance in a timely manner and are not subjectively influenced by personal factors [7, 8].

As competition in business continues to intensify, there is an increasing demand for complex talents. Companies require these people to have not only solid professional knowledge, but also to be able to express themselves proficiently in English, so speaking skills are quite important. Unlike traditional written English teaching, oral teaching focuses on standard pronunciation. Although the forms of teaching have diversified, spoken English teaching is still at an artificial stage at this stage. In the traditional language teaching process, teachers provide comprehensive training such as listening, reading, and writing to students through a face-to-face approach, so as to achieve the purpose of developing students' language communication skills [9–11]. Among them, the learning and training of standard spoken language is the foundation and focus of English learning. Due to the constraints of teachers' resources, learning costs, and learning locations, the effect of traditional speaking learning and training is not satisfactory. Teachers need to spend a lot of time and effort conducting various subjective tests on students, resulting in ineffective work efficiency, especially in large-scale speaking test scenarios.

Currently, researchers are beginning to experiment with computer-assisted pronunciation training systems to address these problems [12–14]. The core issue of computer-aided pronunciation training systems is pronunciation bias testing, i.e., pronunciation bias assessment. Pronunciation bias assessment is the assessment of the standard of the learner's pronunciation and the assignment of a corresponding score or grade, which is the core function of a computer-aided pronunciation training system. Pronunciation bias assessment is mostly a confidence-based method. The phoneme sequence is first standardized and sliced to obtain more accurate phoneme boundary information. Then, the confidence of the phonemes in each speech segment is calculated, and the pronunciation bias is measured by the confidence score. Common confidence calculation methods include log-likelihood, log-likelihood ratio, log-posterior probability, and Goodness of Pronunciation (GOP) [15–17]. In addition, some methods combine confidence calculations with pronunciation features, which yield better joint score results. In order to assess pronunciation bias with high accuracy, more and more researchers are focusing on the detection of pronunciation bias at the phoneme level.

There are two ideas for the study of automatic detection of pronunciation bias at the phoneme level [18]. One is an automatic method for the detection of pronunciation bias based on acoustic phonetics. Such methods are based on a statistical analysis of speech. The other is an automatic method of pronunciation bias detection based on automatic speech recognition technology.

1.1. Pronunciation Bias Detection Based on Acoustic Phonetics. Pronunciation bias detection based on acoustic phonetics finds a specific combination of features by extracting structural, acoustic, and perceptual features of the speech to be tested. Then, pronunciation bias detection is achieved by statistically examining. A similarity calculation or a classifier is usually chosen for the differentiation of pronunciation bias types.

Morlett Paredes et al. [19] proposed a hybrid method based on time-domain features and phoneme boundary information for pronunciation bias detection of basic English pronunciation units, with remarkable results. This hybrid method used a multilayer perceptron as a classifier. Nakamura et al. [20] extracted several resonance peaks from different frames after pre-processing the speech to be measured. Then, a Gaussian Mixture Model (GMM) was used for classification and vowel articulation bias detection was achieved. Dashti and Razjmoo [21] defined a resonance peak that reduces ambient noise. This resonance peak is able to simulate the vocal tract shape properties. Articulatory bias detection is then performed by calculating the degree of structural distortion (Bhattacharyya distance) between the speech to be measured and the standard speech.

1.2. Pronunciation Bias Detection Based on Automatic Speech Recognition Technology. Automatic speech recognition is essentially a classification matching problem, while pronunciation bias detection is a classification regression problem, so pronunciation bias detection can be solved using speech recognition technology. Pronunciation bias detection based on automatic speech recognition is simpler than pronunciation bias detection based on acoustic phonetics. This is because automatic speech recognition can use a language model to counteract the effects of imprecise acoustics and thus output a legitimate sequence of characters. Therefore, this study chose to use automatic speech recognition to implement a spoken English assessment system. The key elements of automatic speech recognition technology include the extraction of speech feature parameters and the selection of acoustic models, both of which are also the focus of this study.

First of all, the extraction of speech feature parameters is a key step in the process of dynamic speech recognition, and the selection of parameters directly affects the overall performance of the system. After the speech signal has been preprocessed, it needs to be extracted and analyzed for the feature parameters. The most typical method of extraction is the use of vocoders.

The vocoder was born in the 1920s at Bell Labs in the USA. Since then, the vocoder has seen a period of rapid development. A large number of researchers have been working on speech coding and speech synthesis, and have achieved considerable results. The basis of the vocoder is Linear Predictive Coding (LPC). In the early 1980s, the US Department of Defense published LPC-10. Liu et al. [22] used LPC to build a parametric pronunciation bias database and combined it with a Gaussian Hidden Markov Model to achieve classification detection of pronunciation bias. Hiroya and Mochida [23] used LPC to extract speech feature parameters and then used the linear discriminant analysis or decision trees to train classification models to achieve pronunciation bias detection.

Second, for pronunciation detection, the constraint provided by the language model is not helpful as it leads to missed detection of incorrect pronunciations. Therefore, robust acoustic models are important to distinguish between those with standard pronunciation and those with abnormal pronunciation. In traditional speech recognition, the Gaussian Hidden Markov Model (GMM-HMM) has been the dominant acoustic model [24]. However, with the continuous development of deep learning techniques, deep learning models are gradually being used more often in speech recognition tasks. A convolutional neural network (CNN) is a multilayer perceptron that incorporates convolutional computation. CNN is one of the representative algorithms of deep learning [25] and is commonly used to analyze visual images. The CNN consists of an input layer, a convolutional layer, a ReLU activation layer, a pooling layer, and a fully connected layer. The CNN is also known as a "translation-invariant artificial neural network."

When applied to automatic speech recognition applications, in terms of input, CNN-based automatic speech recognition techniques are broadly divided into two types: one is to use traditional acoustic feature parameters as input, such as Mel Frequency Cepstrum Coefficient (MFCC) [26], LPC [27], and Fbank [28]. The other is to use original time-frequency spectrum as input, that is, to treat the time-frequency diagram as an image. Er et al. [29] analyzed the research on deep learning techniques in speech recognition and the key problems to be solved. Nakashika et al. [30] used recurrent neural networks for speech recognition and the recognition accuracy was high.

From a pronunciation bias detection perspective, we want to retain as much of the original information as possible in the features received at the input. This is because the original information is the most realistic representation of the quality of the learner's spoken language. However, time-frequency maps can cause information loss in the frequency domain, which is detrimental to pronunciation bias detection. Therefore, the automatic speech recognition technology in this paper uses acoustic feature parameters as input information. Due to the short-time smoothness of spoken English, the feature parameters of the acoustic model in pronunciation bias detection are updated less frequently, which effectively reduces the coding bit rate (below 2.4 kb/s or even below). The simple LPC vocoder is able to achieve a range of 0.8 to 2.4 kb/s in terms of coding efficiency, which just meets the coding bit rate requirements [31–33]. Therefore, LPC is used for speech signal feature extraction, and the features are trained by convolutional neural network algorithm to complete speech recognition. The aim of this study is to adopt LPC to extract acoustic feature parameters and use CNN as an acoustic model for pronunciation bias detection to automate the detection of English pronunciation bias.

In order to achieve automatic scoring of spoken English tests, a machine learning assessment system based on linear predictive coding is proposed, which mainly consists of being divided into four modules: acoustic model acquisition module, speech recognition module, standard pronunciation transcription module, and decision module. The improved stimulated linear predictive speech coding method is used to obtain the feature parameters of the speech signal and generate the speech feature vector to implement the speech recognition module. Finally, the CNN model is used to train the speech features so as to implement the acoustic model acquisition module. The experimental results show that the improved LPC + CNN-based evaluation system can accurately detect pronunciation bias information.

The main innovations and contributions of this paper include.

(1) How accurately unvoice/voiced tones judgments are made is important for spoken English assessment systems. Therefore, the traditional LPC algorithm is improved by using hybrid excitation instead of simple binary excitation. In the acoustic feature parameter extraction process, the sub-band sound intensity of the speech signal is extracted using a split-band hybrid excitation technique in addition to the extraction of the fundamental tone period required by the traditional LPC model.

(2) An English spoken pronunciation evaluation system based on improved LPC and CNN is constructed. The improved LPC algorithm is used to obtain the feature parameters of the speech signal and generate the speech feature vector, thus realizing the speech recognition module. A CNN is used to train the speech features, thus realizing the acoustic model acquisition module.

The rest of the paper is organized as follows: In Section 2, the representative spoken pronunciation assessment system was studied in detail, while Section 3 provides the improved LPC algorithm. In Section 4, the machine learning evaluation system based on ILPC + CNN was studied in detail, while Section 5 provides experimental results and analysis. Finally, the paper is concluded in Section 6.

2. Representative Spoken Pronunciation Assessment System

Since the 1990s, many technology companies and research institutes have conducted in-depth research in the field of pronunciation bias testing and have achieved remarkable results, and launched various application systems, as shown in Table 1. These systems have been widely used in areas such as computer-aided pronunciation training, computer-aided language learning, and computer-based speaking proficiency testing. For example, the DISCO (Development and Integration of Speech technology into Courseware for language learning) project at the University of Nijmegen (Netherlands) [34]. The DISCO system automatically detects pronunciation deviations and grammatical errors in the speech to be tested and generates detailed feedback on the errors checked. The HUGO system, developed by Kyoto University in Japan for Japanese learners of English, uses a decision tree technique based on linguistics and a phonological database to check pronunciation bias.

3. Improvements to the LPC Algorithm

3.1. Principle of LPC. The most basic low-rate speech coding method is linear predictive coding. In speech signal analysis linear prediction not only enables predictive functions but also provides a very good estimation of the vocal channel model parameters. Linear prediction analysis can provide a set of speech signal model parameters that accurately represent the spectral amplitude of the speech signal. The basic idea of linear predictive analysis is to use the p sample point values of the previous set of data to predict the sample point values of the current or next set. LPC can simulate the human articulatory system very well and therefore has some advantages in the extraction of English speech feature parameters [35]. After waveform interception and noise filtering of the speech signal, multiple frames of speech signal

Table 1: Representative spoken pronunciation assessment systems.

System name	Research and development institutions
EduSpeak SDK7	Stanford research institute, USA
DISCO	Lanemegen University, The Netherlands
SCILL	University of Cambridge, UK, Massachusetts Institute of Technology, USA
TBALL	University of California, Los Angeles, USA
HUGO	Kyodo University, Japan
LISTEN	Carnegie Mellon University, USA
ISLE	University of Leeds, UK, University of Hamburg, Germany
EyeSpesk	EyeSpesk inc
Enunciate	The Chinese university of Hong Kong
PLASER	Hong Kong University of Science and Technology
National general language testing system	KDDI Corporation
Versant	Pearson Corporation

in a certain time period can be obtained by frame sampling and combined with a linear time domain model to achieve feature parameter extraction.

Let $s(n)$ represent the speech signal. According to the LPC principle, $s(n)$ can be represented by the previous p sample points.

$$s(n) = a_1 s(n-1) + a_2 s(n-2) + +a_p s(n-p) \tag{1}$$

where a_1, a_2, a_p denote linear prediction coefficients.

Let $\hat{s}(n)$ be the predicted speech signal, then its representation is shown as follow:

$$\hat{s}(n) = \sum_{k=1}^{p} a_k s(n-k) \tag{2}$$

The prediction error is calculated as shown as follow:

$$\begin{aligned} e(n) &= s(n) - \hat{s}(n) = s(n) - \sum_{k=1}^{p} a_k s(n-k) \\ E &= \sum_n e^2(n) = \sum_n \left[s(n) - \sum_{k=1}^{p} a_k s(n-k) \right]^2 \end{aligned} \tag{3}$$

Let $\partial E/\partial a_k = 0\,(1 \le k \le p)$, then all coefficients can be solved and a stable speech feature signal can be obtained.

The basic principle of the linear predictive vocoder is that the model parameters are encoded with the excitation parameters using linear predictive analysis in an all-pole vocal channel model, resulting in the transmission of high-quality speech at low bit rates (below 2.4 kb/s). The principle of the linear predictive vocoder is shown in Figure 1. At the receiver end of the linear predictive vocoder, the prediction coefficients obtained from the linear predictive analysis can be used to synthesize the transmitted speech directly [36]. Figure 2 shows the coding principle of the LPC-10 vocoder.

First, after a low-pass analog filter, the LPC-10 vocoder performs an A/D conversion at a sampling rate of 8 kHz to obtain the digitized information of the speech. The digitized speech is then processed simultaneously in two steps. (1) The excitation information is processed. After the speech has been framed, the characteristic parameters of each frame are extracted and encoded for transmission. After encoding, the fundamental tone period (Pitch) and the voiced/unvoice sign (V/UV) of each frame are obtained. The fundamental tone period is calculated using the average amplitude difference function (AMDF) method. (2) The extraction of the vocal channel parameters is processed.

Because most of the energy of the speech signal is concentrated in the low-frequency range and the power spectrum decays with frequency, the LPC needs to preprocess the speech signal first so that the power spectrum on the high frequencies can be increased, thus improving the accuracy of speech channel parameter extraction.

$$H_{pw}(z) = 1 - 0.9375z^{-1} \tag{4}$$

where $H_{pw}(z)$ denotes the transfer function of the preprocessing filter.

3.2. Improvements to Incentive Sources. Conventional LPC algorithms use simple binary excitation sources (voiced/unvoice) to excite the synthesizer. Due to the low robustness, the quality of the speech synthesized by the binary excitation source is poor in the presence of high speech noise. Real-life English speech often has both voiced/unvoice tones, especially in noisy speech segments. Therefore, the result of the voiced/unvoice tones judgment can directly affect the quality of speech recognition. Therefore, improvement of the excitation source is important for spoken English evaluation systems. In this paper, a hybrid excitation is used instead of the traditional binary excitation, thus proposing an improved LPC algorithm (ILPC). In terms of parameter extraction, in addition to extracting the fundamental tone period required for conventional LPC, the hybrid excitation technique is also used to extract the sub-band sound intensity in the speech signal.

The steps for extracting the fundamental tone period in ILPC arithmetic are shown as follows:

Step 1: after passing the speech signal $x(n)$ through a low-pass filter at 900 Hz, the first 20 output values are removed to obtain $x'(n)$.

Step 2: find the maximum amplitude value of the first 100 samples and the maximum amplitude value of the last 100 samples of $x'(n)$ respectively. Select the smallest value as the threshold level L.

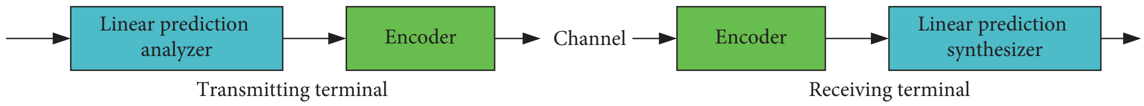


Figure 1: Linear predictive vocoder.

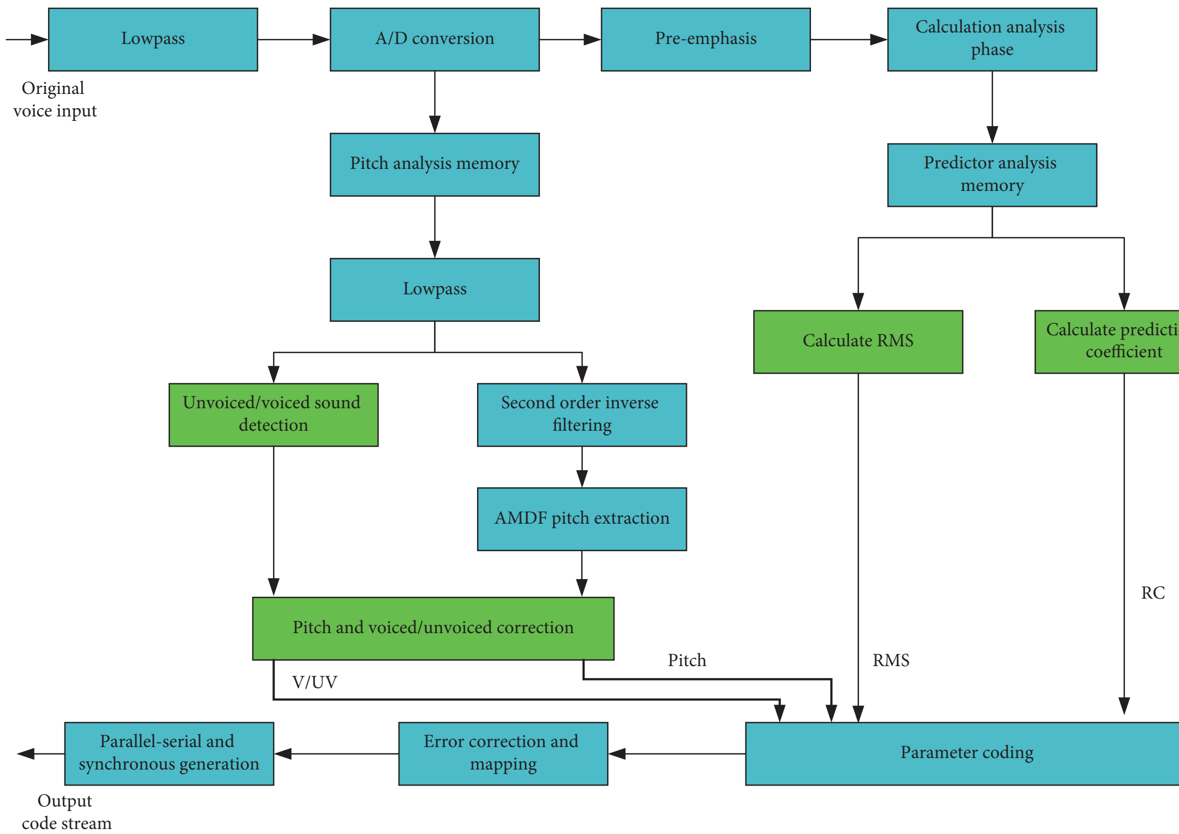


Figure 2: Encoding principle of the LPC-10 vocoder.

Step 3: make center-decimation and three-level decimation of$x'(n)$ to obtain $y(n)$ and $y'(n)$ respectively.

Step 4: find the correlation $R(k)$ between the signals $y(n)$ and $y'(n)$;

$$R(k) = \sum_{n=21}^{300} y \times y'(n+k) \quad (5)$$

where k ranges from 20 to 150 and $R(0)$ is the short-time energy.

Step 5: use the peak detector to find the maximum value of the correlation value $R_{\max}$. If $R_{\max}$ is less than $0.25R(0)$, this frame is considered as voiced tones and the fundamental tone period is set to $P = 0$. Otherwise, this frame is considered unvoiced tones, and the fundamental tone period is set to $P = k$.

The process of extracting the sub-band sound intensity in the ILPC calculation is shown in Figure 3.

After passing through the bandpass filter, the speech signal is extracted to obtain the fundamental tone period. The result of passing a frame of speech signal through each of the five sub-band filters is shown in Figure 4. The sound intensity of the five sub-bands is calculated as follows: 0.2452 for the first sub-band; 0.4478 for the second sub-band; 0.1893 for the third sub-band; 0.3707 for the fourth sub-band; and 0.3874 for the fifth sub-band.

For each unvoiced tone frame or dithered turbulent frame, the sound intensity of the speech signal in each sub-band is calculated separately. In forming the excitation signal, the sound intensity will determine the weighting of the pulse and noise sources in each sub-band, resulting in an excitation signal for the entire frequency band.

4. ILPC + CNN Based Machine Learning Evaluation System Design

4.1. General System Architecture. The automatic detection of pronunciation bias is a simulation of the human subjective detection process. Through machine learning of the manual detection results, automatic detection can even outperform human experts. The machine learning evaluation system for spoken English designed in this paper is shown in Figure 5. The system is divided into four modules: an acoustic model acquisition module, a speech recognition module, a standard pronunciation transcription module, and a decision module.

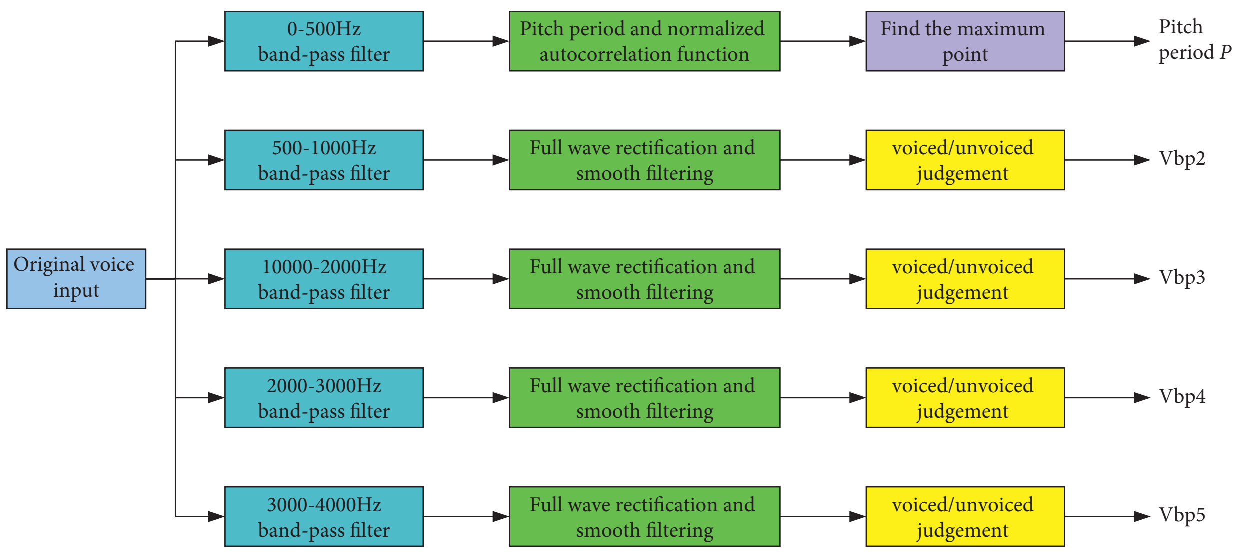


FIGURE 3: Extraction process of sub-band sound intensity.

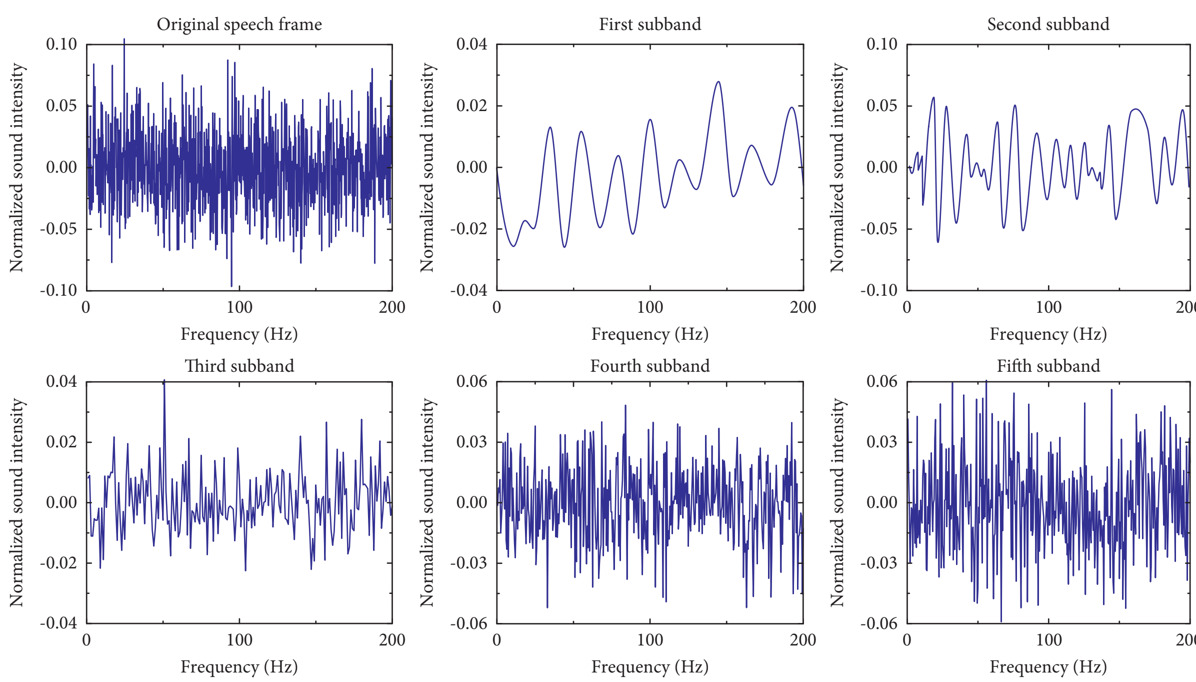


FIGURE 4: 5 subbands of a frame of speech.

4.2. ILPC-Based Speech Recognition Module. In this paper, the ILPC algorithm is used to implement a speech recognition module so that the learner's basic pronunciation units (phonemes), including legal and illegal pronunciation unit sequences, can be accurately identified. Automatic speech recognition aims to detect the content of the learner's pronounced text and to output legitimate character sequences by using acoustic models that can counteract the effects of undesirable acoustics. In the acoustic feature parameter extraction process, we use a split-band hybrid excitation technique to extract the sub-band sound intensity of the speech signal in addition to the parametric fundamental tone period, resulting in an accurate voiced/unvoiced tones judgment.

4.3. CNN-Based Acoustic Model Acquisition Module. The main function of the acoustic model acquisition module is to train an acoustic model. The trained acoustic model will be used in the speech recognition module. In traditional speech recognition, the Gaussian Hidden Markov Model (GMM-HMM) has been the dominant acoustic model. However, with the continuous development of deep learning techniques, deep learning models are gradually being used more often in speech recognition tasks. A convolutional neural network (CNN) is a multilayer perceptron that incorporates convolutional computation [37]. CNN is one of the representative algorithms of deep learning and is commonly used to analyze visual images. Therefore, in this paper, CNNs are used to implement the acoustic model acquisition module.

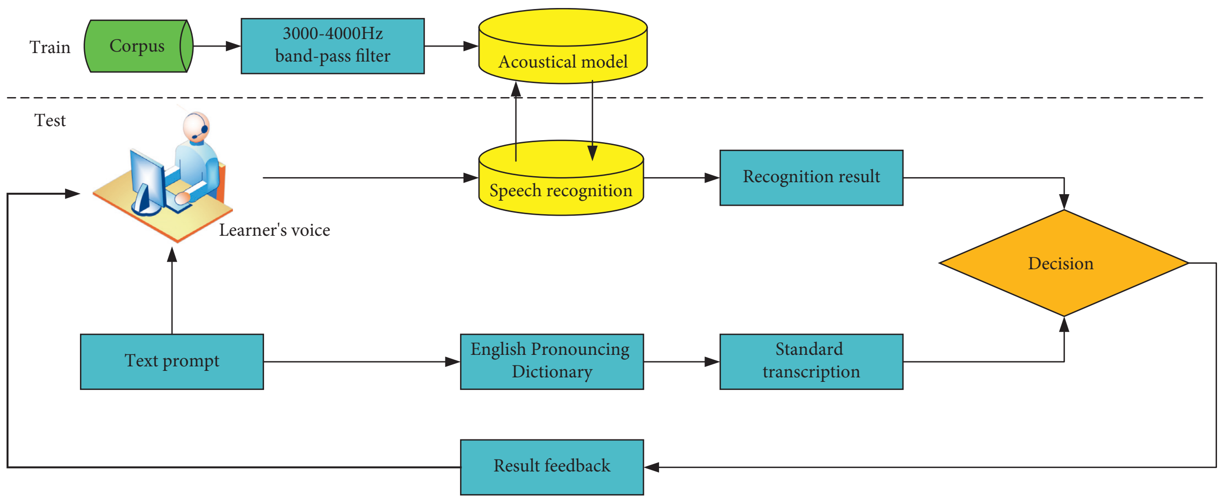


FIGURE 5: Overall system architecture.

The input to the CNN is the acoustic feature parameters obtained by the ILPC algorithm. the structure of the CNN is shown in Figure 6. Let the sample set of speech features be $= (x_1, x_2, x_N)$. First, the m speech features are convolved in the layer l of the CNN [38–40].

$$x_{l,j} = f\left(\sum_{j\in m} x_{l-1} * k_{lj} + b_{l,j}\right), \tag{6}$$

where k_{lj} and $b_{l,j}$ represent the weights and biases of the features j in the$l$ layer respectively, and $*$ represents the convolution operation.

$$f(z) = \frac{1}{1+e^{-z}}. \tag{7}$$

Then, the convolution operation is performed on the m features of the N samples. Let the size of the convolution kernel be $h \times w$.

$$g(x) = \max_{1\le k\le h\times w} (x_k). \tag{8}$$

A new sample is obtained again after the convolution operation and a transformation operation is performed on it.

$$x_j^l = f\left(\sum_{i=1}^{M} a_{ij}\left(x_i^{l-1} * k_i^l\right) + b_j^l\right). \tag{9}$$

The restrictions are shown as follows:

$$\sum a_{ij} = 1, 0 \le a_{ij} \le 1. \tag{10}$$

After obtaining the fully connected layer of the convolutional neural network, the classifier is selected to predict the sample class.

In traditional acoustic model training, the label corresponding to each frame of data needs to be known in order to train effectively. Therefore, the speech signal needs to be forcibly aligned prior to training the model. Although there are some relatively mature open source alignment tools available, there are significant constraints on the performance of speech recognition techniques with forced alignment. In CNN-based acoustic models, we want to leave more tasks to the neural network to perform, such as learning how to align autonomously. Therefore, predictive alignment techniques are used to solve this problem. The loss function for predictive alignment is defined as shown below [41–43].

$$\mathrm{L}(S) = -\ln \prod (\wedge)_{(Y,\hat{y})\in S} = -\sum_{(Y,\hat{y})\in S} \ln p(\hat{y} \mid Y), \tag{11}$$

where $p(\hat{y} \mid Y)$ denotes the probability when the input sequence is $\hat{y}$ and the output sequence is Y, and S denotes the training set. It can be seen that prediction alignment can directly output the predicted probability of a sequence without the need for external post-processing. With the help of predictive alignment, a large amount of manual resources can be saved, thus increasing the efficiency. In this paper, the acoustic model acquisition module is built by combining prediction alignment and CNN, as shown in Figure 7.

5. Experimental Results and Analysis

In order to verify the performance of ILPC + CNN in the quality assessment of spoken English, various experiments were conducted using separate speech samples with different accents. The experimental speech data were obtained from the open-source website VoxForge (https://www.voxforge.org/zh). The parameters of the experimental dataset are shown in Table 2, with a ratio of 3 : 1 between training and test samples. The parameters of the CNN model are set as shown in Table 3. The sampling rate of the audio data is 16000 Hz and the sample size was 16 bit. The number of texts (number of pronounced sentences) is 2268 and the total number of phonemes is 44359. The total number of speakers is 10 including 5 males and 5 females. First, the effect of different frame rates on ILPC performance was tested. Second, the effect of different convolutional kernel sizes on the recognition performance was tested. Finally, the designed system was compared with other spoken pronunciation evaluation systems.

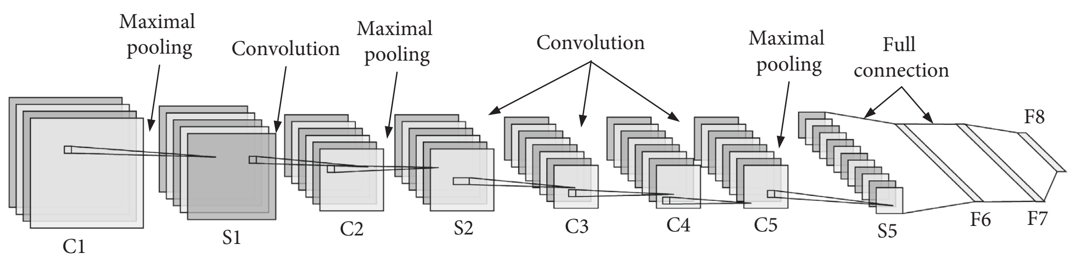


Figure 6: Structure of CNN.

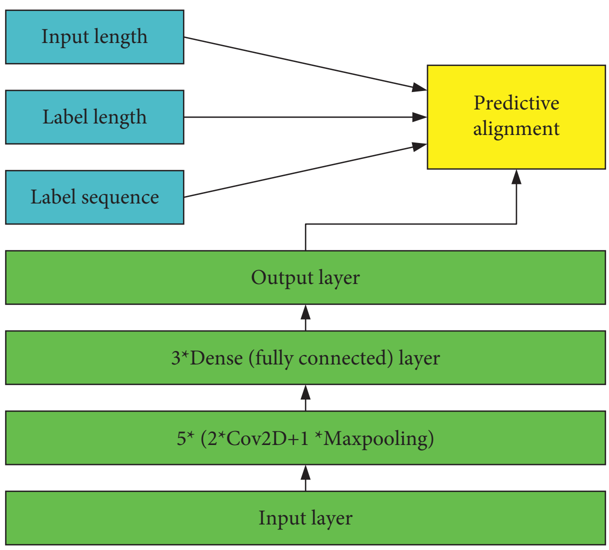


Figure 7: CNN-based acoustic model acquisition module.

Table 2: Parameters of the experimental data set.

Sample number	Type of sample	Sample size
1	American English	8276
2	British English	8144
3	European English	7768
4	Canadian English	3411
5	Australian English	2247
6	Indian English	2412

Table 3: Parameters of the CNN model.

Parameter settings	Numerical values
Learning rate	0.008
Batch_size	16
Convolution kernel	32/3 ∗ 3
Window size	2 ∗ 2
Dropout	0.3
Epoch	300
Optimizer	Adam

5.1. Effect of Different Frame Rates on ILPC Performance. In order to obtain the best frame rate setting, the speech recognition accuracy of the six datasets at different frame rates was verified, as shown in Table 4.

It can be seen that as the number of frames extracted increases, the recognition accuracy keeps improving. As the frame rate increases to 180 Hz, the ILPC algorithm shows a high recognition accuracy. As the frame rates increases to 200 Hz, datasets 1, 2, 3, and 6 show a decrease in recognition accuracy, while datasets 4 and 5 show a very small and almost negligible increase in recognition accuracy. When ILPC was used to extract features of speech signals with different sample types, too high a frequency would increase the computational effort of speech recognition, while too low a frequency would drop important features of the speech signal. Therefore, the frame rate used in the subsequent ILPC algorithms was 180 Hz.

Table 4: Speech recognition accuracy of different frame rates.

Data sets	Frame rate (Hz)				
	60	100	150	180	200
1	0.6133	0.7491	0.8479	0.9291	0.9287
2	0.6578	0.7674	0.8667	0.9307	0.9306
3	0.6344	0.7232	0.8227	0.9267	0.9265
4	0.6473	0.7512	0.8461	0.9244	0.9246
5	0.6022	0.7334	0.8635	0.9074	0.9077
6	0.5613	0.7219	0.8218	0.9083	0.9080

5.2. Effect of Convolutional Kernel Size on Speech Recognition. To further verify the effect of convolutional kernel size on speech recognition performance, the English speech recognition accuracy under different convolutional kernel conditions was tested, as shown in Table 5.

It can be seen that the recognition accuracy of English speech decreases when the size of the convolutional kernel increases. This may be because fewer speech features are involved in the operation when the convolutional kernel size is too large, resulting in a decrease in speech recognition accuracy. The comparison shows that the recognition accuracy of CNN is higher when the convolutional kernel size is $2 * 2$ and $3 * 3$. However, $2 * 2$ is more time-consuming than $3 * 3$ in CNN operations, so to improve real-time performance, the convolutional kernel size was $3 * 3$ in subsequent experiments.

5.3. Performance Analysis of ILPC. The excitation signal of the ILPC vocoder is compared with that of the LPC vocoder, as shown in Figure 8.

It can be seen that the excitation signal obtained by ILPC is a mixed excitation signal. Each frame of speech is no longer pure unvoice tone or voice tone but contains a distinct periodic pulse string and a little noise. As a result,

Table 5: English speech recognition accuracy under different convolution kernel conditions.

Data sets	Maximum accuracy	Average accuracy rate	Standard deviation
Convolution kernel size 2 ∗ 2			
1	0.9491	0.9291	1.75e – 003
2	0.9663	0.9307	1.92e – 003
3	0.9318	0.9267	1.66e – 003
4	0.9491	0.9244	1.73e – 003
5	0.9265	0.9074	1.81e – 003
6	0.9214	0.9083	1.65e – 003
Convolution kernel size 3 ∗ 3			
1	0.9482	0.9286	1.77e – 003
2	0.9658	0.9301	1.96e – 003
3	0.9316	0.9259	1.73e – 003
4	0.9492	0.9238	1.72e – 003
5	0.9263	0.9066	1.82e – 003
6	0.9209	0.9078	1.66e – 003
Convolution kernel size 4 ∗ 4			
1	0.8471	0.8312	4.13e – 003
2	0.8642	0.8476	3.37e – 003
3	0.8225	0.8164	5.68e – 003
4	0.8366	0.8132	4.05e – 003
5	0.8389	0.8246	3.49e – 003
6	0.8217	0.8059	4.22e – 003
Convolution kernel size 5 ∗ 5			
1	0.7323	0.7191	8.24e – 003
2	0.6824	0.6162	7.93e – 003
3	0.6917	0.6694	7.27e – 003
4	0.7318	0.7161	7.13e – 003
5	0.6429	0.6225	6.83e – 003
6	0.6835	0.6634	6.65e – 003

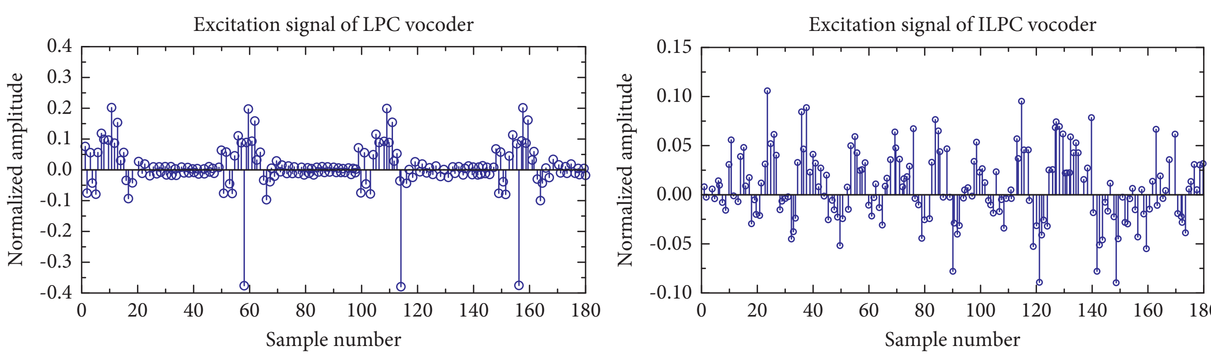


Figure 8: Comparison of excitation signals.

the speech signal obtained by ILPC is more natural and better defined. Conventional LPC uses a simple binary excitation signal to process the input sequence. Compared to the conventional LPC algorithm, ILPC based on hybrid excitation gives a waveform that more closely resembles that of the original speech signal. ILPC algorithm can get speech signals with high naturalness, and its waveform is almost consistent with the original waveform.

To verify the performance of the ILPC-based speech recognition module, 1000 samples were taken from each of the six datasets to form a speech hybrid dataset containing 6000 samples. The spoken pronunciation bias of this hybrid dataset was examined using LPC + CNN and ILPC + CNN respectively. The frame rate was 180 Hz and the convolutional kernel was 3 ∗ 3. The detection results are shown in Table 6 and Figure 9.

It can be seen that after ILPC feature extraction, the detection accuracy of CNN is significantly improved. Due to the lower robustness, the speech quality of LPC is poor in the case of very noisy speech, which is due to the fact that real-life English speech usually has both voice tones and unvoice tones, especially in transition segments and very noisy speech segments. When using ILPC for feature extraction of the captured speech signal, each frame of speech is no longer pure voice tones and unvoice tone, thus retaining as much of the original information as possible. ILPC + CNN converges at about 140 iterations, whereas LPC + CNN takes about 180 iterations to stabilize. In

Table 6: Detection accuracy of CNN and LPC + CNN.

Algorithms	Highest recognition accuracy	Average recognition accuracy	Minimum recognition accuracy
LPC + CNN	0.8573	0.8362	0.7935
ILPC + CNN	0.9625	0.9197	0.9046

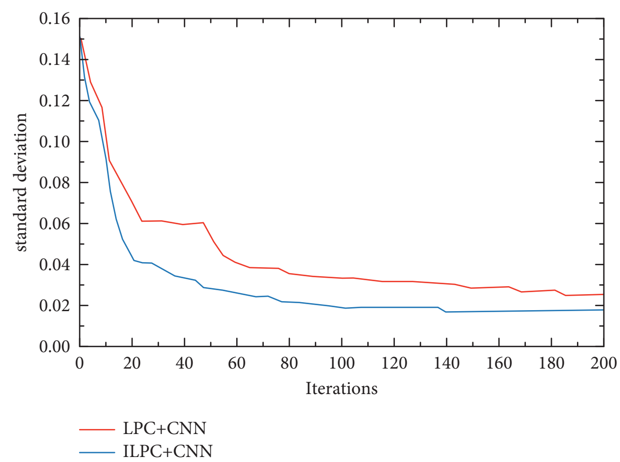


Figure 9: Standard deviation of detection for CNN and LPC + CNN.

Table 7: Articulatory phoneme detection for different systems.

System name	Number of correctly pronounced phonemes	Number of mispronounced phonemes
Manual (labelled) inspection	38326	6033
SCILL	32952	10407
TBALL	34465	9894
HUGO	32559	9568
LISTEN	34061	9714
ISLE	34954	9842
EyeSpesk	35058	8745
Enunciate	33857	9983
PLASER	36447	8243
ILPC + CNN based system	37826	7856

addition, the standard deviation of ILPC + CNN is smaller compared to LPC + CNN.

5.4. Performance Comparison of Different Spoken Language Assessment Systems. In contrast to traditional spoken pronunciation assessment methods, the training data for this experiment did not require manual annotation. Using the above speech mixture dataset containing 6000 samples, the designed system was compared with other spoken pronunciation assessment systems, the results are shown in Table 7.

A total of 44359 phonemes (initials, finals, and tones) were obtained from the speech mixture dataset. The manual detection results showed that 6033 phonemes were mispronounced in this speech data, 10407 mispronounced phonemes were detected by the SCILL system and 9894 mispronounced phonemes were detected by the TBALL system. The system designed in this paper (ILPC + CNN) detected 7856 mispronounced phonemes, which is the closest to the manual (labeled) detection result. The experimental results show that ILPC + CNN algorithm can indeed reduce the misjudgment rate of pronunciation deviation. This indicates that the feature parameters obtained by ILPC using hybrid excitation reflect well the characteristics of the original speech signal and therefore the decoded speech quality is better and the speech is clearer.

Finally, the experiment classified the 64 pronunciation errors into three types, namely initial errors, final errors, and tone errors. These three types of pronunciation errors were counted and the results are shown in Figure 10.

As you can see, of the 3 types of pronunciation errors, intonation is the most likely to occur. Therefore, learners of English need to focus on intonation. The next problem is rhyme errors. Compared to the other two types of pronunciation errors, vowel errors are easier to solve. The phenomenon that tone errors are much higher than the other two types of errors is in line with linguistic laws and therefore the experimental results are reliable.

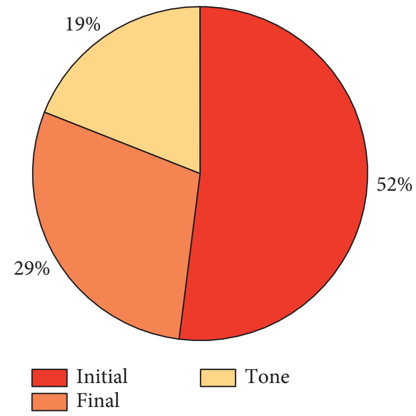


Figure 10: Percentage of 3 types of pronunciation errors.

6. Conclusion

In this paper, a machine learning evaluation system for spoken English based on ILPC + CNN algorithm is constructed so as to automate the detection of learners' pronunciation errors. The designed system consists of four main modules: acoustic model acquisition module, speech recognition module, standard pronunciation transcription module, and decision module. The speech recognition module uses the ILPC algorithm to obtain the feature parameters of the speech signal and generate the speech feature vector. The acoustic model acquisition module uses a CNN model to train the speech features and the input to the CNN is the acoustic feature parameters obtained by the ILPC algorithm. The experimental results show that the feature parameters obtained by ILPC using hybrid excitation reflect the characteristics of the original speech signal very well, and therefore the decoded speech quality is better and the speech is clearer. Compared with other spoken English evaluation systems, the ILPC + CNN-based machine learning evaluation system can reduce the misjudgment rate of pronunciation bias.

Data Availability

The experimental 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 to report regarding the present study.

References

[1] W. Y. Kim, G. S. Lee, Y. H. Kim et al., "Immunolocalization of pki α, βI, βII and γ in adult and developing rat kidney," *Electrolyte & Blood Pressure: E & BP*, vol. 5, no. 2, pp. 75–88, 2007.

[2] Z. Gang, "Quality evaluation of English pronunciation based on artificial emotion recognition and Gaussian mixture model," *Journal of Intelligent & Fuzzy Systems*, vol. 40, no. 4, pp. 7085–7095, 2021.

[3] J. Kim and D. A. Craig, "Validation of a videoconferenced speaking test," *Computer Assisted Language Learning*, vol. 25, no. 3, pp. 257–275, 2012.

[4] E. D. Quaid, "Reviewing the IELTS speaking test in East Asia: theoretical and practice-based insights," *Language Testing in Asia*, vol. 8, no. 1, pp. 2–9, 2018.

[5] H. T. D. Huang and S. T. A. Hung, "Comparing the effects of test anxiety on independent and integrated speaking test performance," *Tesol Quarterly*, vol. 47, no. 2, pp. 244–269, 2013.

[6] S. Karim and N. Haq, "An assessment of IELTS speaking test," *International Journal of Evaluation and Research in Education*, vol. 3, no. 3, pp. 152–157, 2014.

[7] B. Bridgeman, D. Powers, E. Stone, and P. Mollaun, "TOEFL iBT speaking test scores as indicators of oral communicative language proficiency," *Language Testing*, vol. 29, no. 1, pp. 91–108, 2012.

[8] T. P. Stones, "Transcription and the IELTS speaking test: facilitating development," *ELT Journal*, vol. 67, no. 1, pp. 20–30, 2013.

[9] H. T. D. Huang, S. T. A. Hung, and L. Plakans, "Topical knowledge in L2 speaking assessment: c," *Language Testing*, vol. 35, no. 1, pp. 27–49, 2018.

[10] F. Nakatsuhara, C. Inoue, and L. Taylor, "Comparing rating modes: analysing live, audio, and video ratings of IELTS speaking test performances," *Language Assessment Quarterly*, vol. 18, no. 2, pp. 83–106, 2021.

[11] P. Seedhouse, "The dual personality of 'topic' in the IELTS Speaking Test," *ELT Journal*, vol. 73, no. 3, pp. 247–256, 2019.

[12] J. Li, "An evaluation of IELTS speaking test," *OALib*, vol. 06, no. 12, pp. 1–17, 2019.

[13] K. Evanini, M. Heilman, X. Wang, and D. Blanchard, "Automated scoring for the*TOEFL Junior*®Comprehensive writing and speaking test," *ETS Research Report Series*, vol. 2015, no. 1, pp. 1–11, 2015.

[14] S. Roshan, "A critical review of the revised IELTS speaking test," *International Journal of English Language Education*, vol. 2, no. 1, p. 120, 2013.

[15] A. Latifa, A. Rahman, A. Hamra, B. Jabu, and R. Nur, "Developing a practical rating rubric of speaking test for university students of English in parepare, Indonesia," *English Language Teaching*, vol. 8, no. 6, pp. 166–177, 2015.

[16] Y. Zhan and Z. H. Wan, "Test takers' beliefs and experiences of a high-stakes computer-based English listening and speaking test," *RELC Journal*, vol. 47, no. 3, pp. 363–376, 2016.

[17] S. O'Grady, "The impact of pre-task planning on speaking test performance for English-medium university admission," *Language Testing*, vol. 36, no. 4, pp. 505–526, 2019.

[18] R. Koizumi and A. Hirai, "Comparing the story retelling speaking test with other speaking tests," *JALT Journal*, vol. 34, no. 1, p. 35, 2012.

[19] A. Morlett Paredes, A. Gooding, L. Artiola I Fortuny et al., "The state of neuropsychological test norms for Spanish-speaking adults in the United States," *The Clinical Neuropsychologist*, vol. 35, no. 2, pp. 236–252, 2021.

[20] K. Nakamura, T. Toda, H. Saruwatari, and K. Shikano, "Speaking-aid systems using GMM-based voice conversion for electrolaryngeal speech," *Speech Communication*, vol. 54, no. 1, pp. 134–146, 2012.

[21] L. Dashti and S. A. Razmjoo, "An examination of IELTS candidates' performances at different band scores of the speaking test: a quantitative and qualitative analysis," *Cogent Education*, vol. 7, no. 1, Article ID 1770936, 2020.

[22] P. Liu, S. Li, and H. Wang, "Steganography integrated into linear predictive coding for low bit-rate speech codec," *Multimedia Tools and Applications*, vol. 76, no. 2, pp. 2837–2859, 2017.

[23] S. Hiroya and T. Mochida, "Speech sound naturalness alters compensation in response to transformed auditory feedback," *The Journal of the Acoustical Society of America*, vol. 140, no. 4, p. 3228, 2016.

[24] L. Wang, N. Kitaoka, and S. Nakagawa, "Robust distant speaker recognition based on position-dependent CMN by combining speaker-specific GMM with speaker-adapted HMM," *Speech Communication*, vol. 49, no. 6, pp. 501–513, 2007.

[25] F. Demir, M. Turkoglu, M. Aslan, and A. Sengur, "A new pyramidal concatenated CNN approach for environmental sound classification," *Applied Acoustics*, vol. 170, no. 6, Article ID 107520, 2020.

[26] Y. C. Zeng, Y. Y. Chen, and Y. H. Mao, "Mel frequency Cepstrum coefficient extraction method based on empirical mode decomposition and combined spectrum of fourier transform and wigner distribution," *Natural Science Journal of Xiangtan University*, vol. 132, no. 5, pp. 563–573, 2015.

[27] Y. Hiwasaki, K. Mano, and T. Kaneko, "An LPC vocoder based on phase-equalized pitch waveform," *Speech Communication*, vol. 40, no. 3, pp. 277–290, 2003.

[28] J. Cao, M. Cao, J. Wang, C. Yin, D. Wang, and P. P. Vidal, "Urban noise recognition with convolutional neural network," *Multimedia Tools and Applications*, vol. 78, no. 20, pp. 29021–29041, 2019.

[29] M. B. Er, E. Isik, and I. Isik, "Parkinson's detection based on combined CNN and LSTM using enhanced speech signals with Variational mode decomposition," *Biomedical Signal Processing and Control*, vol. 70, Article ID 103006, 2021.

[30] T. Nakashika, T. Takiguchi, and Y. Ariki, "Voice conversion using RNN pre-trained by recurrent temporal restricted Boltzmann machines," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 3, pp. 580–587, 2015.

[31] A. Harma and U. K. Laine, "A comparison of warped and conventional linear predictive coding," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 5, pp. 579–588, 2001.

[32] R. J. Javier and Y. Kim, "Application of linear predictive coding for human activity classification based on micro-Doppler signatures," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 10, pp. 1831–1834, 2014.

[33] M. W. Spratling, "A review of predictive coding algorithms," *Brain and Cognition*, vol. 112, pp. 92–97, 2017.

[34] D. Bosbach and L. Z. Evins, "The European DISCO project: deep geological disposal of modern spent nuclear fuel," *Safety of Nuclear Waste Disposal*, vol. 1, pp. 233-234, 2021.

[35] M. F. Anjum, S. Dasgupta, R. Mudumbai, A. Singh, J. F. Cavanagh, and N. S. Narayanan, "Linear predictive coding distinguishes spectral EEG features of Parkinson's disease," *Parkinsonism & Related Disorders*, vol. 79, pp. 79–85, 2020.

[36] S. K. Selvaperumal, C. Nataraj, and V. Thiruchelvam, "Speech to text synthesis from video automated subtitling using Levinson Durbin method of linear predictive coding," *International Journal of Applied Engineering Research*, vol. 11, no. 4, pp. 2388–2395, 2016.

[37] A. Pandey and D. L. Wang, "A new framework for CNN-based speech enhancement in the time domain," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 7, pp. 1179–1188, 2019.

[38] S. Kwon and S. Kwon, "A CNN-assisted enhanced audio signal processing for speech emotion recognition," *Sensors*, vol. 20, no. 1, pp. 183–201, 2019.

[39] V. Passricha and R. K. Aggarwal, "A hybrid of deep CNN and bidirectional LSTM for automatic speech recognition," *Journal of Intelligent Systems*, vol. 29, no. 1, pp. 1261–1274, 2019.

[40] Z. Zhao, Q. Li, Z. Zhang, N. Cummins, H. Wang, and J.. Tao, "Combining a parallel 2D CNN with a self-attention Dilated Residual Network for CTC-based discrete speech emotion recognition," *Neural Networks*, vol. 141, pp. 52–60, 2021.

[41] J. Zhao, X. Mao, and L. Chen, "Learning deep features to recognise speech emotion using merged deep CNN," *IET Signal Processing*, vol. 12, no. 6, pp. 713–721, 2018.

[42] L. Comanducci, P. Bestagini, M. Tagliasacchi, A. Sarti, and S Tubaro, "Reconstructing speech from CNN e," *IEEE Signal Processing Letters*, vol. 28, pp. 952–956, 2021.

[43] X. Chen, "Simulation of English speech emotion recognition based on transfer learning and CNN neural network," *Journal of Intelligent & Fuzzy Systems*, vol. 40, no. 2, pp. 2349–2360, 2021.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 7877032, 11 pages
https://doi.org/10.1155/2022/7877032

Research Article

Intelligent Recognition of Traffic Signs Based on Improved YOLO v3 Algorithm

Zhonglai Yang

Shanxi Police College, Shanxi, Taiyuan 030000, China

Correspondence should be addressed to Zhonglai Yang; yangzhlai@chd.edu.cn

Received 15 July 2022; Revised 10 August 2022; Accepted 25 August 2022; Published 20 September 2022

Academic Editor: R. Mo

In recent years, assisted driving and autonomous driving technology have been paid more attention to by the public. Road sign recognition is of great practical significance for the realization of auto-driving technology. In the actual traffic environment, the traffic signs have the problems of small detectable volume, low resolution, unclear characteristics, and easy to be disturbed by the environment. In order to better realize road traffic sign recognition, this paper improves and optimizes the YOLO v3 network derived from YOLO v3 structure algorithm, enhances the data of the traffic signs by using color enhancement and other technologies, and improves the original FPN structure of the YOLO v3 network algorithm to 52×52. Then, the secondary sampling output characteristic diagram 108 in the YOLO v3 network is used $\times$ 108 solutions to solve these difficulties of picture size and image distortion. Use 5, 9, and 13 fixed-size pools in front of the surface of the control architecture, then the output characteristics are associated with the original characteristics of the picture so that inputs of different sizes can obtain the same output. Finally, we use the intermediate class K algorithm to group the TT100K landmark data set, reconsider the original network parameters, and compare the TT100K data set with the small target determination algorithm, such as YOLO v3 network model and improved YOLO v3 network model. The results show that compared with the traditional YOLO v3 algorithm, the optimized YOLO v3 road sign recognition algorithm has a significant improvement in sign recognition accuracy, sign recognition speed, and learning cost. When the change of FPS is very small, the recall rate and accuracy will be greatly improved. At the same time, compared with other small target detection algorithms, the improved YOLO v3 algorithm has more accurate and faster detection accuracy.

1. Introduction

Road traffic signs play an irreplaceable role in traffic order and safety. They gather road information, warnings, prohibitions, and other information into a simple sign to guide and restrict drivers to drive safely. The setting of traffic signs maintains the safety and smoothness of road traffic to a great extent. And as a sign to assist road safety, traffic signs also provide a simple and clear breakthrough for the development of intelligent transportation. Road traffic signs are usually composed of some simple words or symbols and have color characteristics that form a sharp contrast with the surrounding environment so that road traffic signs can better attract the attention of drivers. Traffic signs with different symbols and colors represent different traffic information [1].

For now, driverless technology has been widely concerned. Hope to improve the safety factor of driverless vehicles on the road, it needs to improve the car's perception of surrounding things, real-time and accurate detection of all targets on the road surface is an important part of environmental perception [2]. Nowadays, the traffic sign recognition system mainly samples the road traffic signs, further detects and recognizes the collected sample information, outputs the recognized results, compares the original image with the traffic sign database to give the results, and finally, sends out warnings and other information through the control center. Because the traffic sign recognition system is usually used in high-speed motor vehicles, the input signal needs to be processed in the embedded equipment of motor vehicles [3]. Facing such complicated steps, how to make the traffic sign recognition system in the embedded equipment, is higher real-time accuracy of the common difficulty, we face. At the same time, driverless driving has attracted more and more global attention. It is very important to ensure the safety of driverless vehicles in the situation, it is necessary to perceive the surrounding environmental information. Among them, real-time and accurate detection of all targets on the road surface is an important part of environmental perception. By identifying long-distance targets, more time can be provided for vehicle decision-making and control. Usually, long-distance targets (such as traffic signs) are smaller in size, occupy few pixels, and have no obvious features in the image, which makes it difficult to detect and recognize them in real-time [4]. Therefore, how to accurately recognize traffic signs, while ensuring real-time performance, is the key problem to be solved. Up to now, the recognition of traffic signs is mainly divided into two directions: conventional feature extraction and deep learning.

1.1. Basics Feature Extraction Method for Traffic Standard Recognition. There are generally three traditional traffic sign recognition schemes: color-based traffic sign recognition method, shape-based traffic sign recognition method, and machine learning-based traffic sign recognition method. See Figure 1 for the traditional traffic sign recognition process.

1.1.1. Color-Based Traffic Sign Recognition Method. Road signs generally contain red, yellow, and blue colors. This bright color leads to strong separability of feature information in the image, and color space threshold segmentation is relatively easy. Up to now, many scholars have reached a color-based traffic sign recognition method. The recognition method based on the color of road signs adopts the method of dividing the color spatial distribution of road signs to realize the detection and recognition of road signs, then extracts the feature information of the segmented image, and finally, classifies the extracted feature information through SVM classifier. RGB color space model proposes an image segmentation algorithm, which improves the operation speed of the algorithm [5]. Yang and Wu [6] proposed a two-stage algorithm for road traffic sign detection. The algorithm first calculates the color probability and then converts the image into a probability model for feature extraction. The extracted feature information is passed through the integral channel to reduce the error. Yuan et al. [7] used edge information to detect color changes in local areas of traffic signs.

1.1.2. Detection Scheme Using Road Sign Shape Recognition. Because the shape of road signs with different meanings varies greatly, we can recognize traffic signs by recognizing their shapes. We call this recognition algorithm based on the shape of traffic signs. This kind of algorithm first extracts the feature information of traffic sign shape and then classifies the extracted feature information through different shapes. Moreno et al. [8] and others detect traffic signs by limiting the hough transform of geometry in a certain area, which

Figure 1: Schematic diagram of traditional detection algorithm.

improves the robustness of the detection system. Boumediene et al. [9]proposed a coding gradient detection scheme for road sign damage and occlusion, which improved the poor detection effect of traffic signs that detect damage and occlusion. Pei et al. [10] proposed a low-rank matrix recovery architecture with a detection model to solve the problem that the relativity of characteristic information in traffic signs is easy to be ignored, which can better use the relativity of traffic signs to identify road signs.

1.1.3. Recognition Scheme Based on the Shape of Road Signs. The road sign detection scheme based on machine learning usually uses the moving window method to detect the given traffic sign images in turn, and the researchers manually select and extract the image feature information. In the research of target detection based on machine learning, Dalal [11] proposed the HOG algorithm in 2005. The working principle of the algorithm is to use the gradient direction distribution histogram in the image to describe the location-specific data of the feature information in the image and normalize it. This algorithm can effectively detect the local data of target features in the image, and then the HOG + SVM [12] structure has been continued, which also has a great adverse effect on road sign recognition. Because traffic signs have distinct color information, Huan et al. [13] added color information to the HOG algorithm to expand and achieved good results in traffic sign distinction. According to Lecun et al. [14], research findings a variant gradient direction histogram feature based on HOG algorithm, and trained a single classifier to detect traffic signs through a limit learning machine, which improved the detection efficiency without reducing the detection accuracy.

1.2. Road Sign Recognition Algorithm Based on Deep Learning. The computer technology research of road sign detection schemes using deep learning methods for recognition is also gradually maturing. The rise of convolutional neural networks makes the deep learning method using deep neural networks combined with different training methods shine in the field of computer vision. Since Geoffrey Hinton [15] proposed the research of artificial intelligence in 2006, deep learning has rapidly swept all research fields of computer technology, among which the most representative algorithm is a convolutional neural network (CNN). In computer vision, the convolution neural network solves the problems of difference recognition accuracy and slow recognition speed at the current stage and can extract the feature information in the image more efficiently and accurately. With the development of CNN, two-stage network structures such as RCNN [16], VGG [17], and AlexNet [18] for image classification and one-stage network structures based on

SSD [19] and YOLO [20] series algorithms for target detection have been successively extended. The algorithm first extracts the features of the target image, then generates candidate regions through the extracted feature information, and finally, uses convolution neural network to classify. In contrast with the traditional target detection algorithm, two-stage algorithm solves the shortcomings of more feature information, a large amount of data, slow detection rate, poor generalization ability, and so on. Single-phase architecture is also known as the identification framework of application regression. It mainly uses the idea of regression theory to give the area, information directly through the backbone network, and discards the candidate areas and RPN network, respectively. Compared with two-stage algorithms, this algorithm can recognize faster, but the recognition accuracy is not as good as two-stage algorithm. With the deepening of the research on deep learning algorithms, more and more scholars study the use of deep learning algorithms to identify road signs. Zuo et al. [21] proposed cascaded RCNN algorithm, which has a detection accuracy of 99% on CCTSDB data set, but the detection rate is relatively slow. Jianming et al. [22] used faster CNN to detect traffic signs and optimized the detection performance. Jianming et al. [22] reduced the amount of calculation and parameters of the algorithm by clipping the network on the basis of YOLO v2 [23] and enhanced the detection performance of small target traffic signs by meshing the input characteristic image.

2. Main Problems of Traditional Traffic Sign Detection Algorithm

2.1. Main Problems of Color-Based Traffic Compilation Detection Algorithm. Different types of traffic signs have different colors. For example, red traffic signs generally indicate prohibited behaviors. Different color combinations of traffic signs also convey different information. The identification of traffic signs can effectively read the meaning of traffic signs. With the deepening of the research on the color of road signs, the color-based road sign detection architecture has greatly improved the detection speed and accuracy. However, traffic signs are often on open and exposed roads, sometimes facing the influence of illumination, fading, occlusion, and bad weather, which makes the results obtained by the color-based detection algorithm unstable, resulting in wrong detection results and missed detection.

2.2. Main Problems of Road Sign Shape Recognition Architecture. The shape of traffic signs is an important feature of traffic sign information. For example, triangles often indicate reminders, and circles indicate prohibition or release of prohibition. Effective identification of traffic sign shapes can solve the initial reading of traffic sign information. For the detection algorithm of road signs shape, although the recognition accuracy of traffic signs has been greatly improved after continuous improvement research, due to the complexity of the road environment, the detection results when the traffic signs face occlusion, deformation and other situations are unsatisfactory [24]. In addition, the amount of calculation required to extract the shape feature information of traffic signs is large, it increases the calculation time of the model and requires higher computing power of the machine. Although many scholars are also studying the detection algorithm of unifying the color and shape of road signs, it also models size reduction and improves the real-time performance of the algorithm, but the reliability and real-time performance of this traditional traffic sign detection algorithm are still difficult to meet people's requirements for safe driving [25].

2.3. There Are Main Problems in Traffic Sign Recognition Algorithm Using Machine Learning. Although the traffic sign detection framework based on machine learning has a great improvement in the detection accuracy compared with the traffic sign detection algorithm based on color and shape, this kind of detection algorithm has higher requirements for feature extraction. In addition, the detection algorithm based on machine learning usually needs to manually select the region of feature information, which makes this kind of algorithm have a high workload and poor real-time performance. For traffic sign detection, the target detection algorithm based on machine learning still has some limitations.

3. Basic Principle of YOLO v3 Algorithm

YOLO v3 is a target detector. Its backbone architecture uses Darknet-53 instead of Darknet-19. There are 53 convolution layers in total. The network structure is shown in Figure 2.

Darknet-53 uses RESNET's residual idea for reference to form a residual structure, which can well control the spread of gradients, avoid situations that are not conducive to training, such as gradient disappearance or explosion, and greatly reduce the difficulty of training deep networks. The main part of the network is composed of five other debris. Multiple residual units form a residual block and each residual unit is constituted of two DBL modules and quick links. The deep separable convolution model is shown in Figure 3.

Darknet-53 minimum weight DBL module is composed of convolution, packet standardization, and leakage recovery firing. YOLO v3 divides the forecast into 13×13, 26×26, and 52×52. These three parameters push the three performance graphs to the test level. In particular, the features of low-level mapping have a small sensitive field and strong small target detection ability, while the features of depth mapping have a large sensing range and improving the performance of detecting large targets [26]. Therefore, YOLO v3 has obvious advantages in determining the size of detection targets. Because YOLO v3 network has high learning efficiency and strong adaptability to different task scales in complex traffic scenes, TT100K [18] signaling data set is used to improve, train and test YOLO v3 network.

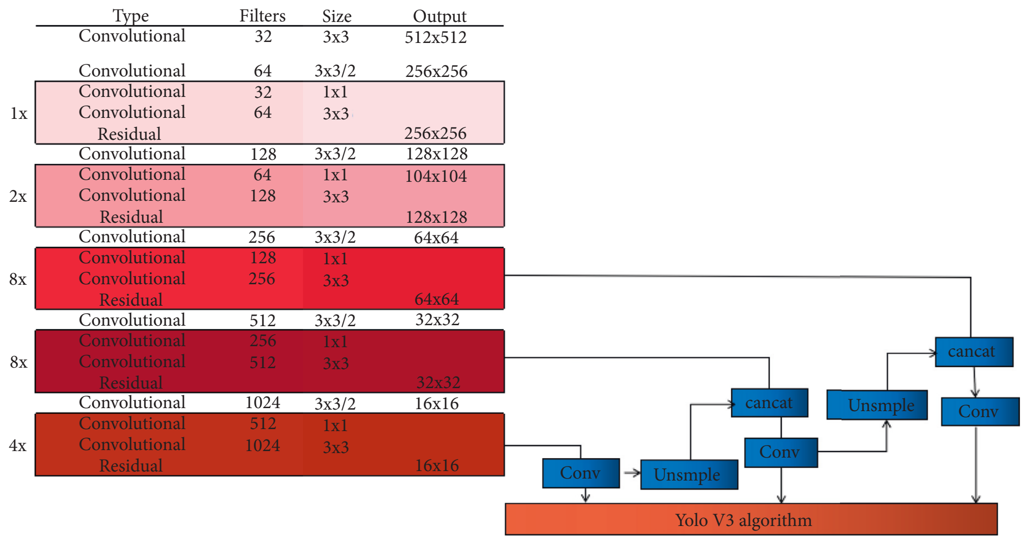


FIGURE 2: Basic model of YOLO v3 framework.

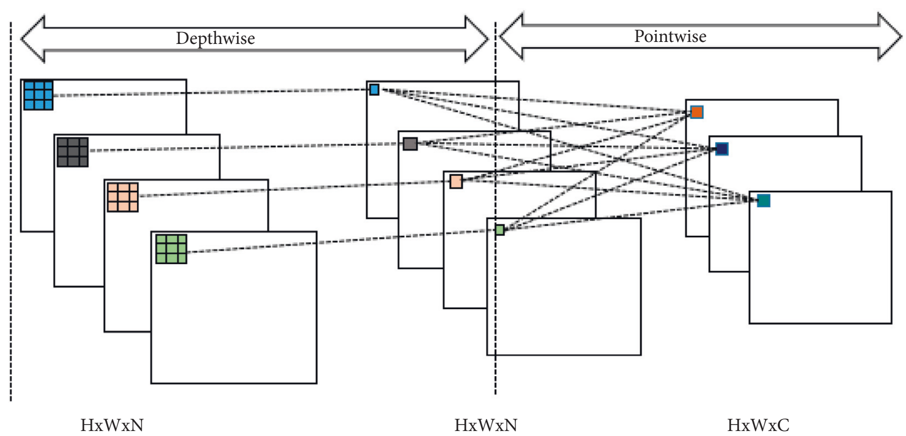


FIGURE 3: Deep separable convolution model.

4. Improvement of 4 YOLO v3 Algorithm

Aiming at the low accuracy of identification of the original YOLO v3 neural network for long-distance lower target objects, this paper improves the algorithm composition, *K*-means network structure, and loss function.

4.1. Improvement of Network Structure. Since the deep network of the original deep structure of YOLO v3 is conducive to the detection of large targets, and the shallow structure is convenient for the detection of small targets because the shallow algorithm structure passes through small convolution layers [27], it lacks deep semantic features, contains less semantic information, and has weak feature representation ability, these features affect the detection of small targets, which depends on the shallow algorithm structure. In order to improve the feature extraction ability of the detection algorithm structure, this paper uses Inception architecture that can enrich the features of the shallow network for reference.

As shown in Figure 4, inception the neural network operation and pool operation are performed on the identified image, and the output results are spliced into the deep marking feature image of different convolution kernel sizes such as 1×1, 3×3, 3×3, or 5×5. The information of different perception domains can be obtained from the input picture data, these operations can be combined, and all the structure can be combined to improve the image quality representation. Inspired by the concept architecture, a concept redefinition module structure is proposed and applied to the shallow layer network of YOLO v3. Compared with the traditional YOLO v3 network, the recognition algorithm of the shallow layer network has a stronger ability to extract the specific representation of the picture, and the information extraction abundance of the recognition system is improved. The improved YOLO v3 algorithm is also more closely combined with the feature points of the neural network, the recognition and perception efficiency of the image is higher, and the recognition ability of small traffic signs is improved.

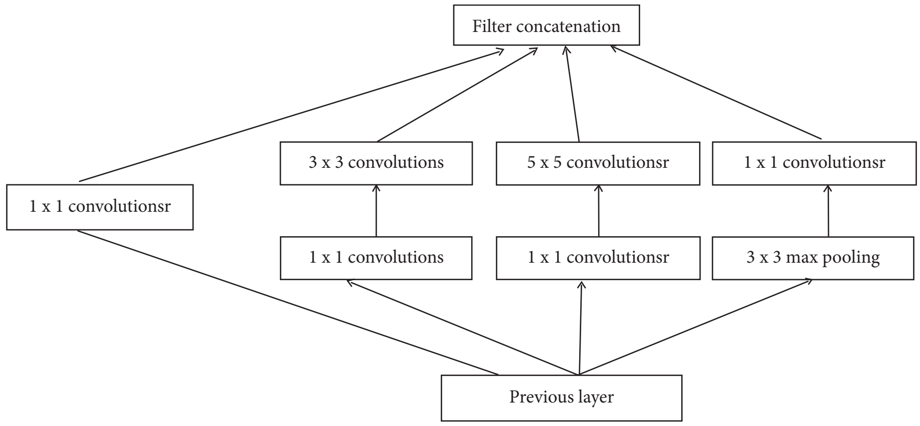


Figure 4: Structure diagram of inception.

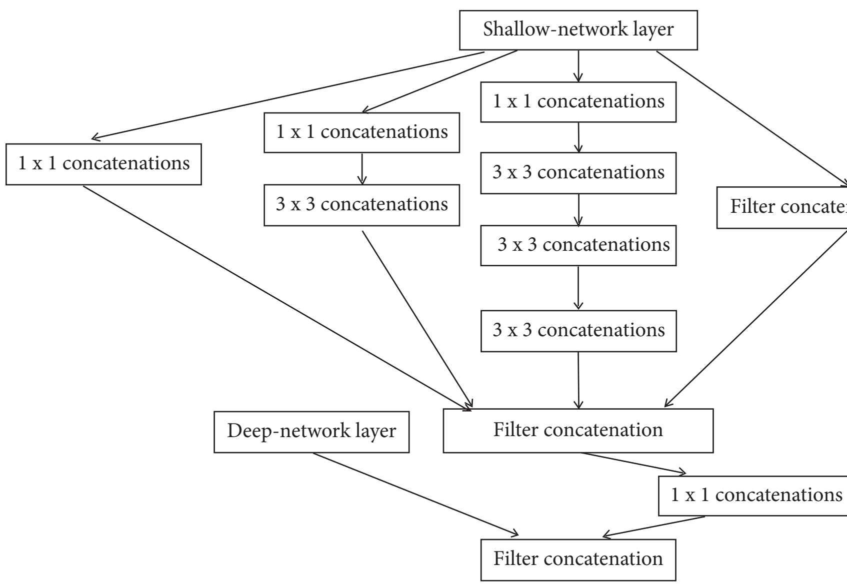


Figure 5: Structure diagram of inception-redefined module.

Figure 5 shows the structure of the initial redefinition module. The two ends of the structure are shallow network layer and deep network layer, respectively, which are used to connect to the network of YOLO v3. The structure consists of four substructures: the first substructure is 1×1 volume accumulation; the second branch is composed of a convolution of 1×1 and then a convolution of 3×3; the third branch consists of the convolution of one 1×1 followed by the convolution of three 3×3; and the fourth branch consists of the maximum pooling layer. A 7×7 convolution effectively extracts basic information from various small pictures. In this paper, using three 3×3 convolutions instead of one 7×7 convolution can save $7 \times 7(3 \times 3 \times 3) = 1.81$ times the calculation amount, which can improve the calculation speed. Front 3×3 convolution is 1×1. The convolution layer can reduce the number of input channels, effectively reduce the number of input parameters, and increase the parallel ability of the architecture. Benefit from 1×1 convolution passes through the ReLU activation function [28], so the generalization performance of neural network can be improved through the introduction of nonlinearity data pool for extraction image features. The four branches extract features of different scales that increase the adaptability of the network to different scales and obtains the information from multiple scales, respectively, [29]. Then the feature maps under the four branches are fused, and finally, the number of output access is reduced through 1×1

Table 1: Test results of different picture sizes.

Picture size	Test model	mAP	FPS
416×416	YOLO v3	0.431	54.4
	Improved YOLO v3	0.514	50.7
608×608	YOLO v3	0.691	34.2
	Improved YOLO v3	0.752	31.3
1024×1024	YOLO v3	0.791	18.7
	Improved YOLO v3	0.834	16.4

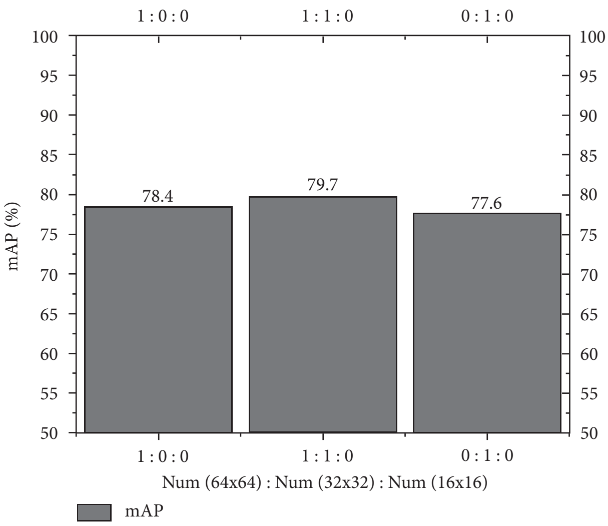


Figure 6: Comparison of map values of inception redefined module under different distribution proportions 4.2k-means algorithm improvement.

convolution layer. The channel ratio of the characteristic image output from the left 1×1 convolution to the right 3×3 convolution will affect terminal identification accuracy. In this paper, different proportions were tested, and the quantity ratio of 1 : 1 was finally selected with the highest accuracy. Table 1 shows the mAP values for different proportions.

Introduce the concept-redefined module structure into the output 64 in Figure 1, 64×64 (and 32×32) between the characteristic diagram and concatenation to form the YOLO v3 improvement network. The introduction method is shown in Figure 4. By connecting the shallow-network layer to the deep-network layer, the combination of deep information and shallow information is more conducive to the prediction of small targets. For output 64×64 in terms of a characteristic graph, the channel flux capacity and size of deep network characteristic graph are $64 \times 64 \times 128$. The size and surface channels network characteristic map are $64 \times 64 \times 256$, then the fused feature map are $64 \times 64 \times 384$.

As shown in Figure 6, this paper attempts to apply several distributions of the inception redefined module structure. Finally, this paper selects the distribution with the best result. Compared with shallow information, deep information can provide more image features. The multidimensional and multilevel convolution kernel in the improved YOLO v3 algorithm also provides convenience for the perception of visual field information in different ranges, and the sentence information abundance improves the perception ability of small targets.

In order to improve the accuracy of the traffic sign recognition algorithm for image proportion recognition, this paper improves the disadvantages of the original Yolo v3 and K-means clustering algorithm that lack filtering function and proposes an improved k-means clustering algorithm. Based on this, this paper proposes an improved K-means algorithm. First, the invalid data in the data set are eliminated by calculating the width-height ratio of the object coordinates, and the valid data are retained. Next, use k-means architecture operation to classify the obtained data. The mother is to obtain the size and proportion of the anchor. Finally, the classification results are added to the YOLO layer for training and recognition. The execution order of the improved k-means algorithm is as follows: Input: dimension file in the data set. Output: width, height, and proportion of anchor box. Where I is the number that marks the target.

(1) Eliminate significance of data annotation in data set.

   (1) for $i = 1$ to total do
   (2) Write coordinate data from the dimension file of the data set.
   (3) Mark according to the following rules. In the upper left corner of the annotation box, the standard x-axis is $X_{\min}$, the coordinate of the lower left corner of the annotation box on the x-axis is $X_{\min}$, the coordinate of the small left corner on the y-axis is $Y_{\min}$, the x-axis of the lower right corner of the coordinate axis is marked as $X_{\max}$, and the upper right corner of the y-axis is marked as $Y_{\max}$.
   (4) $\mathrm{D}x = X_{\max} - X_{\min}$, $\mathrm{d}y = Y_{\max} - Y_{\min}$, if $\mathrm{D}x = 0$ or $\mathrm{d}y = 0$, the mark data correspondence to Dx and Dy is meaningless.
   (5) $Q = \mathrm{D}x - \mathrm{D}y$, if $0.3 < Q < 1$, next, make data annotation by comparison to DX and Dy is valid, and else annotation data are invalid.
   (6) Filter all meaningful comments in the information set.
   (7) end for

(2) effective annotation data are clustered

   (1) choose k clusters intentionally, and choose the central initial cluster of k aimlessly.
   (2) do
   (3) Calculate the IOU value and cluster center of all valid annotation data.
   (4) Data points with large IOU values will be divided into clusters location of cluster center.
   (5) The new family center point is generated from the central set of each cluster data point selected.
   (6) While (the cluster center moves).

The optimized K-means clustering algorithm can effectively ignore the adverse impact of invalid annotations on the clustering center, significantly improve the fit between the anchor box and traffic signs, and significantly improve the recognition accuracy of YOLO v3 network model.

4.2. Optimization Loss Algorithm. The data lost in YOLO v3 algorithm are mainly divided into coordinate regression loss, confidence loss, and clustering loss. For the loss function of coordinate regression calculated by mean square error, the size of the target can directly resulting in decreased accuracy of coordinate regression, using IOU as the target scale in YOLO v3 will bring two problems: first, when IOU (A, B) is equal to 0 (A, B are the forecast boundary box and the real boundary box, respectively), that is, when A and B do not overlap, it is impossible to know whether a and B are adjacent to each other or far away, that is, it cannot reflect the distance between them, and its gradient will be zero, so it cannot be optimized; the second is that IOU (A, B) is not 0, that is, when a and B overlap, the specific overlap of the two cannot be reflected, and in the case of different distances, different proportions, and different aspect ratios, using IOU as a loss, the regression situation is usually incomplete. Compared with IOU, it only focuses on the areas where clusters overlap, unlike IOU, which focus only on overlapping areas, GIOU not only pays attention to the specific overlap of superposition areas but also at the same time, there is enough computing power to match in other nonoverlapping areas, which can better feedback the matching degree between objects. As shown in Figure 5, the IOU values are all 0.33, but there are three different overlaps, that is, the GIOU values are 0.33, 0.24, and 0.1 from left to right, respectively. GIOU is defined as follows:

$$\mathrm{GIOU} = \mathrm{IOU} - \left(C - S_{\mathrm{union}}\right)/C. \tag{1}$$

In this formula, C is the minimum superposition area of a and B. The value range of IOU is [0, 1], and the value range of GIOU is [−1, 1]. For GIOU, when the predicted value completely coincides with the actual value, the value is 1. When the two do not overlap and approach infinity, GIOU takes the minimum value 1. Then, GIOU is the preparation of expressing the measurement accuracy, which can accurately reflect the difference between the predicted value and the true value. Therefore, this paper selects GIOU to replace the coordinate regression loss, and the formula is given as follows:

$$L_{\mathrm{Got}} = 1 - \mathrm{GIOU}. \tag{2}$$

The confidence loss function is given as follows:

$$\begin{aligned} L_{\mathrm{coaf}} = {} & -\sum_{i=0}^{s^2}\sum_{j=0}^{B} I_{ij}^{\mathrm{oljj}}\left[\widehat{C}_i \log\left(C_i\right) + \left(1-\widehat{C}_i\right)\log\left(1-C_i\right)\right] \\ & - \lambda_{\mathrm{nooblij}}\sum_{i=0}^{s^2}\sum_{j=0}^{B} I_{ij}^{\mathrm{objj}}\left[\widehat{C}_i \log\left(C_i\right) + \left(1-\widehat{C}_i\right)\log\left(1-C_i\right)\right]. \end{aligned} \tag{3}$$

The first term in this formula is the confidence error of the prediction frame containing the target. The second item is the confidence error of the prediction frame without targets. S^2 is the number of grids of markers in the input image; I_{ij}^{oljj} indicates whether the jth anchor box of the ith grid catches the target, which is 1 or 0; C_i is the confidence score of the real box; $\widehat{C}_i$ is the confidence score of the prediction box.

$$L_{\mathrm{cls}} = -\sum_{i=0}^{s^2} I_{ij}^{\mathrm{obj}} \sum_{C\subset \mathrm{Class}} \left[\widehat{p}_i(C)\log\left(p_i(C)\right) + \left(1-\widehat{p}_i(C)\right)\log\left(1-p_i(C)\right)\right]. \tag{4}$$

When the first anchor box of the ith grid captures the target, the bounding box generated by the anchor box will calculate the classification loss function. In Equation (4), C represents the detected target category, and $p_i(C), \widehat{p}_i(C)$ express the probability of the real box and the prediction box belonging to category C in grid i, respectively. The final improved loss function is given as follows:

$$\mathrm{Loss} = L_{\mathrm{Glov}} + L_{\mathrm{coof}} + L_{\mathrm{cls}}. \tag{5}$$

5. Experimental Results and Analysis of Recognition Algorithm

In order to test the accuracy of the optimized YOLO v3 algorithm designed in this paper to recognize traffic signs, this paper carries out two inspections, which are the comparative experiment of different improved YOLO v3 algorithms and the use of three image input sizes (416×416, 608×608, and 1024×1024), and verify these two tests from the aspects Average detection accuracy (map), number of detected pictures per second (FPS) and accuracy recall (P-R) curve.

5.1. Road Sign Sample Collection and Experimental Platform. In this experiment, traffic sign data set tt100 is used, which provides 100000 2048×2048 images, of which 30000 traffic sign instances are small targets. There are 45 types of objects in the data set, representing the corresponding traffic signs, respectively: *i*2, *i*4, *i*5, *il*100, *il*60, *il*80, *io*, *ip*, *p*10, *p*11, *p*12, *p*19, *p*23, *p*26, *p*27, *p*3, *p*5, *p*6, *pg*, *ph*4, *ph*4.5, *ph*5, *pl*100, *pl*120, *pL*20, *pl*30, *pl*40, *pl*5, *pl*50, *pl*60, *pl*70, *pl*80, *pm*20, *pm*30, *pm*55, *pn*, pne, po, pr40, w13, w32, w55, w57, w59, and wo. First of all, we should put forward the pictures without landmark files in the traffic sign data set, select 6105 pictures for algorithm practicing, and 3070 pictures for accuracy and speed detection.

Experimental hardware: software system ubuntu22.04, *d* Neural network learning framework pytorch1.4, CPU AMD-r5 5600g, 32 GB memory, NvidiageForceRTX3080Ti ∗ 2 GPU, 24 GB video memory.

5.2. Algorithm Training and Detection. In this experiment, we trained the YOLO v3 architecture and the improved YOLO v3 architecture and used the conventional parameter optimization method of YOLO v3 to optimize the parameters. The initial learning rate is 0.001, and the maximum iterations are 300 cycles. The training rate is set to decay 10 times when the number of iterations is 75 epochs, 150 epochs, and 250 epochs, respectively. The data set is enhanced by flipping, translation transformation, and other methods. At the same time, multiscale training is adopted to make the scale float up and down in the set range, so as to achieve a better training effect. First, test the two models after training and calculate their precision and recall. The formulas for calculating precision and recall are given as follows:

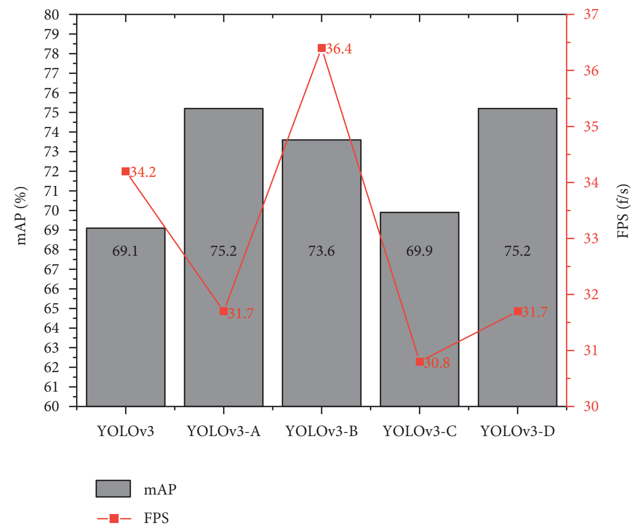


Figure 7: Improved detection results of YOLO v3 and YOLO v3.

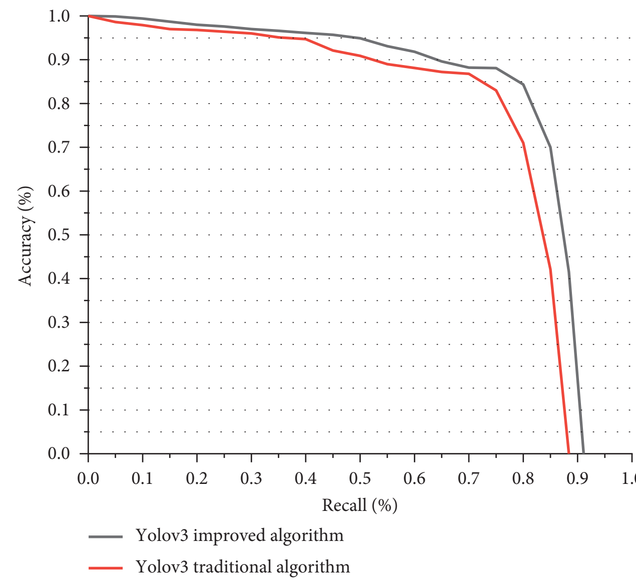


Figure 9: *P*-*R* curve (picture size 608 × 608).

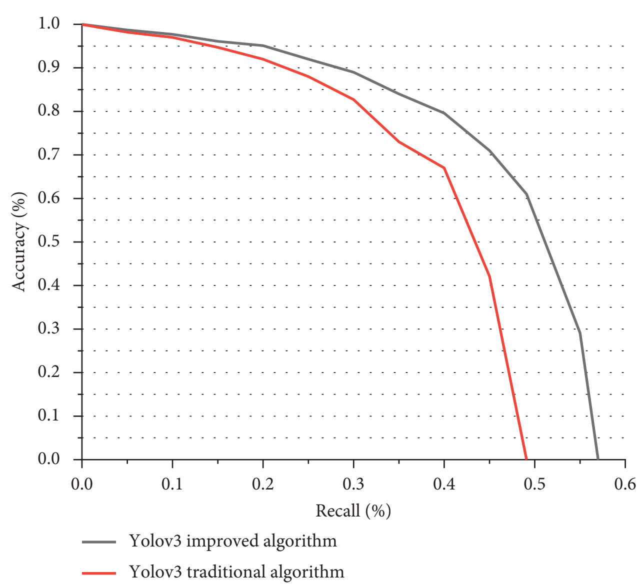


Figure 8: *P*-*R* curve (picture size 416 × 416).

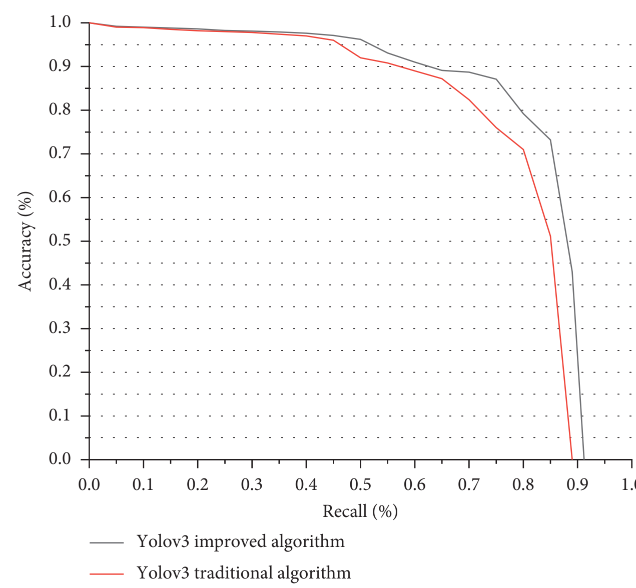


Figure 10: *P*-*R* curve (picture size 1024 × 1024).

$$p = \frac{MP}{MP + OP},$$
$$R = \frac{MP}{MP + ON}. \tag{6}$$

In the formula, MP is the number of positive classes predicted by positive classes, OP is the number of negative classes predicted by positive classes, and ON is the number of positive classes predicted by negative classes. By setting a fixed threshold, the prediction results of the detector are arranged in descending order of confidence, and positive prediction samples are generated respective, the P and *R* values can be calculated and the P-R curve can be drawn.

5.3. Comparison Results and Analysis

5.3.1. The Contrast Experiment of Improved. YOLO v3 is based on traffic signs (named YOLO v3-A), improved YOLO v3 network detection layer and FPN structure (YOLO v3-B), added spatial pyramid module (YOLO v3-C), and added to the above three improved YOLO v3-D networks. The input image size is 608 × 608. These four optimized models have been tested on TT100K photo acquisition and conventional YOLO v3 network. The experimental results are shown in Figure 7.

Figure 7 shows that the detection rate and accuracy of the optimized YOLO v3 architecture on the TT100K traffic sign data set are higher than that of the conventional YOLO v3

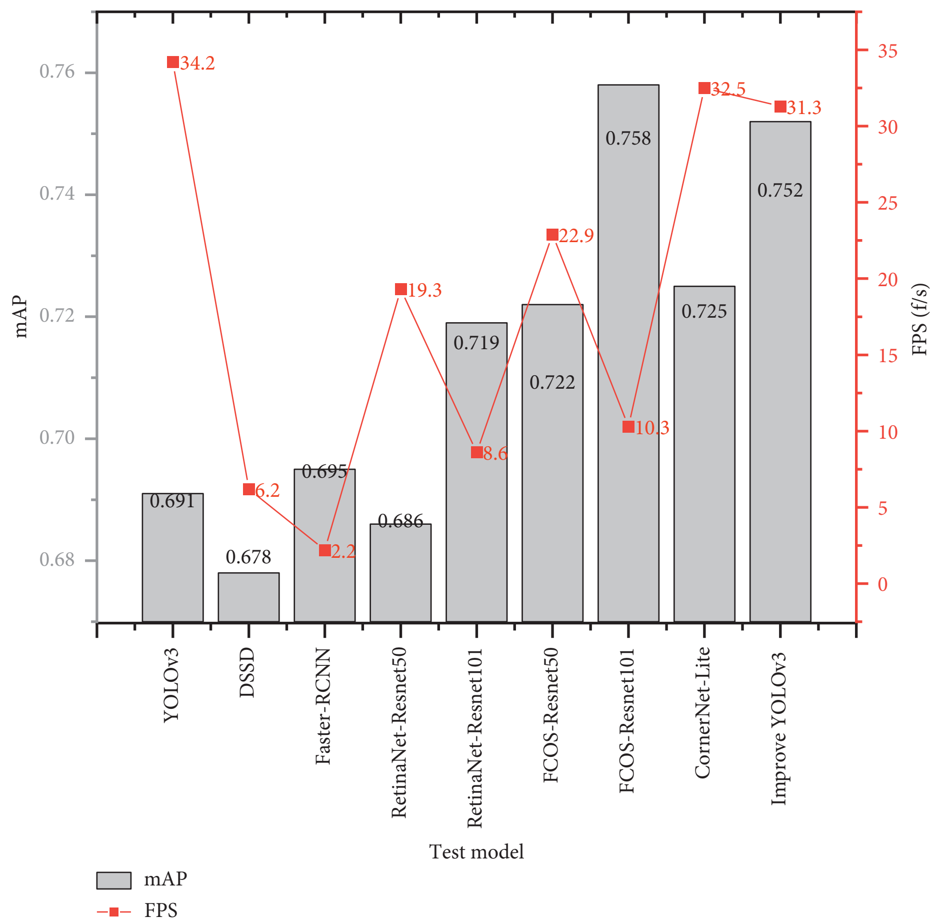


Figure 11: Comparison of recognition results of different recognition architectures.

architecture. Among the four optimized YOLO v3 models, the mapping of YOLO v3-D model reaches 75.2%, and the effect is the best among the four networks. Although its FPS has been reduced to 31.3 f/s, it can still meet the needs of implementing traffic sign detection and recognition.

5.3.2. Improved YOLO v3 Experiment with Different Picture Sizes. Furthermore, test the effectiveness of the optimized YOLO v3 algorithm, and carry out comparative analysis experiments on the optimized YOLO v3 neural network algorithm (YOLO v3-D) and the original YOLO v3 algorithm when the image input sizes were 416, 608, and 1024, respectively. When the input sizes are 416, 608, and 1024, respectively, the P-R curve comparison diagrams of the YOLO v3 network and the improved YOLO v3 network model are shown in Figures 8–10. Accuracy and recall rate of optimized YOLO v3 algorithm model are higher than that of the YOLO v3 network to a certain extent.

There are 3,070 pictures and 7,700 targets in the test set. Two network models are tested, respectively. Their map and FPs are shown in Table 1.

The input dimensions are 416×416, 608×608, and 1024×1024, the map of optimized YOLO v3 algorithm model increased by about 8.3%, 6.1%, and 4.3%, respectively, while FPS did not decrease significantly; at size 416×416 and 608×608, it has the characteristics of fast detection and can meet the needs of road standard identification in reality.

5.3.3. Comparative Analysis of the Optimized Algorithm and the Original Algorithm. In order to further test the recognition effect of the optimization algorithm, the input image size is 608×608. The improved YOLO v3 algorithm is compared with RetinaNet, FCOS, CornerNet, and other advanced small target detection algorithms. The final comparative analysis is shown in Figure 11.

It can be seen that (Figure 10)) only FCOS algorithm is higher and is better than the optimized YOLO v3 framework in recognition accuracy. However, its FPS is far lower than the optimized YOLO v3 architecture; the detection speed of CornerNet algorithm is similar to that of the improved YOLO v3 algorithm, but its mAP is 2.7% lower than that of the improved YOLO v3 algorithm. Experiments show that the optimized YOLO v3 algorithm proposed in this paper can achieve good results in traffic sign recognition. When recognizing small traffic signs in TT100K traffic sign big data set and traffic signs with small range occlusion and long distance, the improved YOLO v3 algorithm has also significantly improved its recognition efficiency and accuracy in the experimental results.

6. Conclusion

This article introduces an optimization model utilized on YOLO v3, which aims to solve the *e* problem of low accuracy of road traffic sign recognition, in the task of road sign recognition, the detection mode needs to deal with many parameters and slow speed. Aiming at the shortcomings of YOLO v3, the algorithm architecture, *K*-means clustering algorithm, and loss function are optimized, which greatly improves the accuracy and speed of the detection framework. The simulation results show that the optimized YOLO v3 framework has more advantages for small traffic standard recognition. The detection experiments on three different resolution photos show that compared with conventional YOLO v3 algorithm, recognition accuracy improvement of optimization algorithm 8.1%, 5.9%, and 4.6%, respectively. Under the premise of ensuring that the gap between FPS is small, the recall rate and accuracy rate have been significantly improved. In general, the main advantage of the optimized YOLO v3 algorithm in road sign detection and recognition is that the recognition efficiency is improved and the recognition accuracy is higher. In particular, the recognition rate is higher for small and distant traffic signs and traffic signs that are covered by foreign matters in a small range. It can be seen that the improved YOLO v3 has higher usability in actual road traffic.

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 they have no conflicts of interest.

References

[1] R. Saleh and H. Fleyeh, "Factors affecting night-time visibility of retroreflective road traffic signs: a review," *International Journal of Traffic and Transportation Engineering*, vol. 11, no. 1, pp. 115–128, 2021.

[2] H. Liu, *Train Unmanned Driving Algorithm Based on Reasoning and Learning strategy*Elsevier, Netherlands, 2021.

[3] P. P. . Thakur, "Detection and classification of traffic signs for driverless cars," *Bioscience Biotechnology Research Communications*, vol. 13, no. 14, pp. 517–521, 2020.

[4] F. Franzen, C. Yuan, and Z. Li, "Traffic sign recognition with neural networks in the frequency domain," *Journal of Physics: Conference Series*, vol. 1576, no. 1, Article ID 012015, 2020.

[5] A. W. Utoyo, H. D. Aprilia, R. A. D. R. I. Kuntjoro, and A. Kurniawan-Jakti, "Visual communication analysis the effect of signs and colors on traffic safety in Jakarta," *IOP Conference Series: Earth and Environmental Science*, vol. 729, no. 1, p. 6, Article ID 012087, 2021.

[6] Y. Yang and F. Wu, *Real-Time Traffic Sign Detection via Color Probability Model and Integral Channel Features*, Springer, Berlin Heidelberg, 2014.

[7] X. Yuan, L. F. Liu, C. H. Li, and Y. Y. Qu, "Unifying visual saliency with hog feature learning for traffic sign detection," in *Proceedings of the Intelligent Vehicles Symposium*, IEEE, Xi'an, 2009.

[8] D. . Moreno, F. Roberto Pichler, and A. Quesada Arencibia, "Fast road sign detection using hough transform for assisted driving of road vehicles," in *Proceedings of the 10th International Conference on Computer Aided Systems Theory*, Springer, Berlin Heidelberg, 2005.

[9] M. Boumediene, C. Cudel, M. Basset, and A. Ouamri, "Triangular traffic signs detection based on RSLD algorithm," *Machine Vision and Applications*, vol. 24, no. 8, pp. 1721–1732, 2013.

[10] D. Pei, F. Sun, and H. Liu, "Supervised low-rank matrix recovery for traffic sign recognition in image sequences," *IEEE Signal Processing Letters*, vol. 20, no. 3, pp. 241–244, 2013.

[11] N. Dalal, "Histograms of oriented gradients for human detection," in *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Diego, CA, USA, 2005.

[12] I. M. Creusen, R. G. J. Wijnhoven, and E. Herbschleb, "Use of artificial neural network in pattern recognition," *Color Exploitation in Hog-Based Traffic Sign detection[C]//IEEE International Conference on Image Processing*, Springer, Berlin Heidelberg, 2020.

[13] Z. Huang, Y. Yu, J. Gu, and H. Liu, "An efficient method for traffic sign recognition based on extreme learning machine," *IEEE Transactions on Cybernetics*, vol. 47, no. 4, pp. 920–933, 2017.

[14] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–44, 2015.

[15] R. Girshick, J. Donahue, T. Darrell, and M. Jitendra, "Rich feature hierarchies for accurate object detection and semantic segmentation," *IEEE Computer Society*, vol. 580–587, 2013.

[16] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *Computer Science*, vol. 1409, 2014.

[17] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, no. 2, 2012.

[18] W. Liu, D. Anguelov, and D. Erhan, *SSD: Single Shot MultiBox Detector*, pp. 21–37, European conference on computer vision, Springer, Cham, 2016.

[19] J. Redmon, S. Divvala, and R. Girshick, "You only look once: unified, real-time object detection," in *Proceedings of the Computer Vision & Pattern Recognition*, IEEE, NV, USA, 2016.

[20] J. Zhang, Z. Xie, J. Sun, Z. Xin, and W. Jin, "A cascaded R-CNN with multiscale Attention and imbalanced samples for traffic sign detection," *IEEE Access*, vol. 8, no. 99, p. 1, 2020.

[21] Z. Zuo, Y. Kai, Z. Qiao, W. Xu, and L. Ting, "Traffic signs detection based on faster R-CNN," in *Proceedings of the 2017 IEEE 37th International Conference on Distributed Computing Systems Workshops (ICDCSW)*, IEEE, Atlanta, GA, USA, 2017.

[22] Z. Jianming, H. Manting, J. Xiaokang, and L. Xudong, "A real-time Chinese traffic sign detection algorithm based on modified YOLOv2," *Algorithms*, vol. 10, no. 4, p. 127, 2017.

[23] S. U. Kim and J. W. Lee, "Traffic sign recognition, and tracking using RANSAC-based motion estimation for autonomous

vehicles," *Journal of Institute of Control, Robotics and Systems*, vol. 22, no. 2, pp. 110–116, 2016.

[24] X. Yang, W. Liu, S. Zhang, L. Wei, and T. Dacheng, "Targeted Attention Attack on Deep Learning Models in Road Sign Recognition," *IEEE*, vol. 8, 2020.

[25] Y. Sun and L. Chen, "Traffic sign recognition based on CNN and twin support vector machine hybrid model," *Journal of Applied Mathematics and Physics*, vol. 9, no. 12, p. 21, 2021.

[26] Z. Zhu, S. Zhang, H. Xiaolei, L. Baoli, H. Shimin, and L. Dun, "Traffic-signdetectionandclassificationinthewild," in *Proceedings of the 2016IEEEConferenceonComputerVisionand PatternRecognition(CVPR)*.

[27] A. Hechri and A. Mtibaa, "Two-stage traffic sign detection and recognition based on SVM and convolutional neural networks," *IET Image Processing*, vol. 14, no. 5, pp. 939–946, 2020.

[28] R. Ayachi, M. Afif, Y. Said, and A. B. Abdelali, "Real-time implementation of traffic signs detection and identification application on graphics processing units," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 35, 2021.

[29] W. Sun, Y. Du, X. Zhang, and G. Zhang, "Detection and recognition of text traffic signs above the road," *International Journal of Sensor Networks*, vol. 35, no. 2, p. 69, 2021.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 7289046, 9 pages
https://doi.org/10.1155/2022/7289046

Research Article

The Mechanism of Unsafe Behavior of Employees in High-Risk Positions under the Adjustment Effect of Hardy Personality

Yan Li, Yi Zhang, Jingfeng Wang, and Zhen Liu

School of Management, Xi'an University of Science and Technology, Xi'an, Shaanxi 710054, China

Correspondence should be addressed to Yan Li; guanliliyan@xust.edu.cn

Received 5 July 2022; Revised 12 August 2022; Accepted 17 August 2022; Published 19 September 2022

Academic Editor: R. Mo

In order to study the impact of human and environmental matching on the unsafe behavior of employees in high-risk positions, explore the intermediary role of work satisfaction and the adjustment role of hardy personality, and improve the safety management level of employees in high-risk positions. The data collected from 501 employees in high-risk positions were empirically analyzed from an individual personality perspective with the help of AMOS 26.0 and SPSS 25.0. The research results showed that staff matching and human-organization matching have significant negative effects on unsafe behavior; job satisfaction plays a complete intermediary role in the negative relationship between human-organization matching and unsafe behavior; hardy personality can adjust the relationship between job satisfaction and unsafe behavior.

1. Introduction

Employees in high-risk positions have been in high pressure for a long time, and the "three violations" behaviors and unsafe accidents occur one after another, but most of the accidents are rooted in human factors [1]. However, existing safety management systems and policies in China are generally effective in preventing unsafe actions and behaviors of workers in high-risk positions. How to reduce the incidence of safety accidents from the perspective of matching individual workers in high-risk positions with their environment has become an urgent problem to be studied and solved [2].

Individual-environment matching refers to a stable state between the individual's own characteristics and all environmental factors that can interact with it [3]. Human-organization matching and human-post matching, as two important indicators in the matching of human body and environment, are the direct factors that drive individuals to exert their own creativity and improve safety performance [4]. The degree of matching between the individual and the environment is reflected by the employee's job performance, job behavior, job satisfaction, and other behaviors. Zhao had empirically tested that the value matching and demand matching of enterprise employees have a significant impact on job satisfaction [5]; Zhang et al. showed that factors such as emotional experience, material exchange, value matching, and fair distribution can predict the teachers' job satisfaction [6]; Huang took hotel employees as the object and found that organizational matching would have a positive impact on employees' job satisfaction [7]. Therefore, when the person-organization matching and the matching degree between people and positions are high, the job satisfaction of employees will be improved [8]. The values of individuals and the organization will tend to be consistent, and it will be more conducive to good communication and communication among employees. This way the employees can be corrected for unsafe behavior.

As a subjective controllable factor, job satisfaction is affected by individual factors such as personal values [9], individual characteristics [10], and personality traits [11]. Judge explored the relationship between the Big Five and job satisfaction through meta-analysis, and also found that high-level agreeable personality is more likely to obtain satisfaction and face the stress in life [12], in addition to the direct effect of personal characteristics, but also through the form of moderator variables on job performance or outcome [13], on the relationship between job satisfaction and turnover intention [1]. To sum up, personality, as an individual characteristic variable, is likely to have an effect on job satisfaction in the form of a moderator variable rather than a main effect [14]. Therefore, personality as a variable to explore the relationship between employee satisfaction and behavior has important research value. It can be seen that most of the current research on employee behavior is from the perspective of job requirements-ability matching and other individual levels, but few scholars have combined personality traits and job satisfaction to study the impact of employee behavior in high-risk positions such as coal and construction.

To sum up, based on the individual-environment fit theory and cognitive toughness theory, this paper established a model of factors influencing unsafe behaviors of employees in high-risk positions with job satisfaction as the mediating variable and hardiness personality as the moderating variable. The effect mechanism of job matching on the unsafe behavior of employees in high-risk positions is researched, and the moderating effect of individual hardiness personality on job satisfaction and unsafe behavior is further studied.

2. Theoretical Analysis and Research Hypothesis

2.1. Person-Post Matching and Person-Organization Matching. Person-post matching and person-organization matching were first used in corporate HR operations, and both have received much attention from scholars since Winston proposed them as indicators of job and organizational characteristics and individual fit. Social cognitive theory suggests that there is a certain functional relationship between human behavior and the individual and the environment, and thus human behavior is influenced by individual and environmental differences [15]. Person-position matching reflects the consistency of individual's own experience, knowledge, skills, and job requirements, etc.; person-organization matching reflects the consistency and compatibility of goals and values between individuals and organizations. Person-organization matching is a balanced state between individuals and organizations, and person-position matching is a harmonious state between individuals and positions, and matching organizations and positions help improve employees' job satisfaction and help employees regulate their own behavior.

Numerous scholars had shown through empirical studies that matching has significant effects on job performance and job satisfaction, as shown in Table 1, but most of the literature directly considers the role of matching on the influenced factors. However, very few scholars consider the role of moderating and mediating variables.

2.2. Job Satisfaction and Its Influencing Factors. Job satisfaction is a subjective attitude of workers towards their work environment and work experience, and is a "detector" for enterprise or organization management. When individual job satisfaction is low, the tendency of unsafe behavioral intentions such as unfocused and perfunctory work is easy to occur. Domestic and international scholars developed job satisfaction survey scales from multiple dimensions, such as the Minnesota Satisfaction Scale MSQ, and combine different assessment methods to quantitatively measure employee job satisfaction and prevent turnover and the occurrence of behaviors inconsistent with the organizational goals. From the perspective of value matching, different scholars extracted the factors influencing job satisfaction varying as shown in Table 2.

Safety material culture and safety institutional culture can all exert a significant influence on employees' safety behavior through job satisfaction [27]. From organizational identity theory, job satisfaction plays a mediating role between individual and organizational value matching and work engagement [28]. When the job satisfaction of employees in high-risk positions is improved, the individual's ability to work can be recognized by the organization, which is conducive to the individual's sense of responsibility for work and makes the unsafe actions and behaviors such as irregular action operation and inattention improved and corrected.

2.3. Hardy Personality and Its Regulatory Role. Personality traits, as one of the indicators that distinguish individuals from each other, are widely used in studies to regulate work attitudes, job satisfaction, and individual behavior, as shown in Table 3. Hardy, as one of the personality traits, refers to an individual's attitude towards dealing with daily unexpected or stressful events, and is a measure of the size of an individual's ability to regulate himself after a setback [35, 36]. Individuals in the same scenario have different attitudes toward difficulties, frustrations, and stress. Individuals with higher hardy view setbacks as a prerequisite for success and a necessary path to progress, even in low job satisfaction environments, by adhering to strict rules and regulations. Hardy personality is mostly used in human resource management to study its moderating effect on employees' intention to leave. An increase in the level of hardiness personality can buffer the effect of lower job satisfaction on employees' intention to leave [37–40].

2.4. Unsafe Behavior. The study of unsafe behavior, a hot topic of current research in the field of management, does not have a clear and uniform definition. This article refers to unsafe behavior that is caused by the human dynamic factors of violations and accidents. In high-risk industries, employees with low job satisfaction are prone to actions that violate the work safety regulations and lead to unsafe behaviors, but improving employee job satisfaction can have a corrective effect on the unsafe behaviors of construction workers [41].

In addition to the influence of employees' own emotions, welfare benefits, organizational safety, and various indicators of performance appraisal are significantly associated with employees' unsafe behaviors [42]. According to Kristof's classification, the impact of person-organization match on safety behavior was explored in three dimensions: value

Table 1: Matching as an influencing factor.

Author	Influenced factors
Kim et al. [15]	Job satisfaction, organizational commitment, job performance, and citizenship behavior
Chen and Wang [16]	Work attitude
Jia et al. [17]	Employee mobility, turnover rate
Long and Zhao [18]	Job satisfaction, willingness to leave
Liu and Wu [19]	Farmworker job performance
Tan [20]	Job satisfaction, job engagement, and organizational support
Xi et al. [21]	Job satisfaction, propensity to leave
Zhao [22]	Employee emotional commitment
Yang [8]	Organizational commitment, employee relationship performance, work engagement, and motivation to learn
Sun and Zhou [23]	Job satisfaction, sense of organizational support, and burnout

Table 2: Factors influencing job satisfaction.

Author	Influencing factors
Locke [9]	Job, salary, recognition, promotion, benefits, working conditions, personal values, job expectations, etc.
Arnold and Feldman [24]	Job supervisor, financial compensation, promotion, work environment, work group
Robbins [11]	Challenging work, supportive work environment, cordial co-worker relationships, fair compensation, personality and job match
Zhao [5]	Matching values, matching needs, and matching abilities
Wang and Yang. [25]	Compensation, benefits, career development, performance recognition and work environment, requirements-ability match
Chen et al. [26]	Ability matching
Zhang et al. [6]	Emotional experience, material exchange, value matching, fair distribution and workload
Huang [7]	Organizational value matching, value matching and matching of competencies and job requirements

Table 3: Personality traits as moderating variables.

Author	Moderated variables
Liu [29]	Corporate culture and job satisfaction
Huang and Zhou [30]	Employee tendency to leave
Mao and Sun [31]	Uncivilized behavior and attitudes
Yu and Zhou [14]	Work-family balance and job satisfaction
Jinlei and Jiang [32]	Personal characteristics and entrepreneurial intentions
Chen and Xiong [33]	Workplace rejection and employee adaptive performance
Liu et al. [34]	Psychological factors and information sharing behavior

match, demand-supply match, and requirement-competency match, and there was a negative relationship between competency, job conditions, job rewards, and interpersonal relationships in the person-post match dimension and the occurrence of employee insecure behavior [1]. Therefore, the positive effects of person-organization matching and person-post matching on safety compliance behavior and safety participation behavior are significant; and the positive effects played by person-post matching and person-organization matching are different in different contexts [43].

Based on the above analysis, this paper proposes the following hypothesis.

H1: Person-post matching positively affects the job satisfaction of employees in high-risk positions;

H2: Person-organization matching positively affects the job satisfaction of employees in high-risk positions.

H3: Person-organization matching negatively affects the unsafe behavior of employees in high-risk positions;

H4: Person-post matching negatively affects unsafe behaviors of employees in high-risk positions.

H5: Job satisfaction negatively affects the unsafe behavior of employees in high-risk positions.

H6: Job satisfaction plays an intermediary role between person-post matching, person-organization matching, and unsafe behaviors.

H7: Hardy personality plays a moderating role between job satisfaction and unsafe behavior.

2.5. Theoretical Model. Based on the above-mentioned theoretical analysis and hypothesis, a theoretical research model of influencing factors of unsafe behaviors of employees in high-risk positions with job satisfaction as the mediating variable and hardy personality as the moderating variable is constructed, as shown in Figure 1.

3. Research Methods

3.1. Data Collection. Taking high-risk workers such as construction workers and coal mine workers in Shaanxi Province, Sichuan Province, Shanxi Province, and other provinces as the survey subjects, the questionnaires were distributed and collected through online methods such as WeChat based on the principle of random sampling.

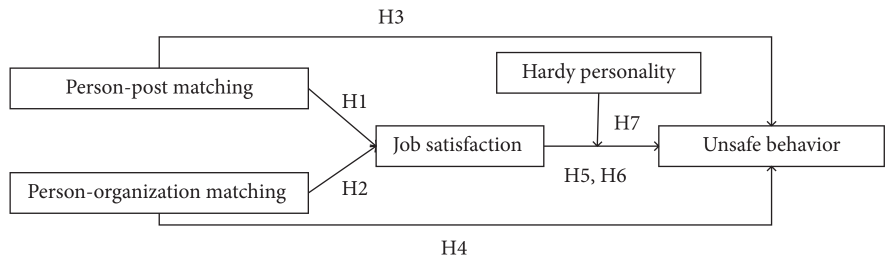


Figure 1: Theoretical hypothesis model.

Table 4: Basic information statistics of the survey sample.

Projects	Variable	Proportion (%)
Age	Under 25 years old	15.2
	26–36 years old	32.9
	36–45 years old	33.3
	Over 46 years old	18.6
Education level	Junior high school or below	26.1
	High school or technical secondary school	35.7
	Junior college	19.2
	Undergraduate	16.4
	Postgraduate and above	2.6
Length of service	5 years and below	35.9
	6–10 years	33.1
	11–20 years	20.6
	More than 20 years	10.4

Questionnaires were issued from March 6 to June 30, 2021. A total of 580 questionnaires were distributed during the survey, and 530 questionnaires were returned. After excluding invalid questionnaires, 501 valid questionnaires were finally obtained, with an effective rate of 86%. See the basic information in Table 4.

3.2. Variable Measurement. There are 39 items in this questionnaire, including 3 items for basic information, and the other variables are measured using the maturity scale of previous research. Using SPSS25.0 to test the reliability and validity of the scale, the overall Cronbach reliability coefficient of this questionnaire is 0.857, and the combined reliability coefficient is 0.952, indicating that the questionnaire is suitable for exploratory analysis. The reliability and validity test results of each dimension are as follows:

(1) Person-post matching. Using Wu [44] to develop the scale of the human-post matching measurement table, it contains 10 items, such as "I think I adapt to the labor load of my position.". The Cronbach reliability coefficient is 0.937, and the combined reliability coefficient is 0.966.

(2) People-organization matching. Using the Cable and DeRue scale [45], it contains 9 items, such as "My values are very similar to the organization's values.". The Cronbach reliability coefficient is 0.930, and the combined reliability coefficient is 0.956.

(3) Job Satisfaction. Using the scale of Chinese scholar Zhang Huifeng, it contains 6 items, such as "Are you satisfied with your salary?" The Cronbach reliability coefficient is 0.873, and the combined reliability coefficient is 0.887.

(4) Hardy personality. Using the scale compiled by Lu Guohua and Liang Baoyong in 2008, it contains 4 items, such as "work and study will bring me fun.". The Cronbach reliability coefficient is 0.850, and the combined reliability coefficient is 0.808.

(5) Unsafe behavior. According to the occupational characteristics of high-risk positions in China, the scale was adjusted to include 7 items, such as "I occasionally dismantle machinery and equipment and other safety devices." The Cronbach reliability coefficient is 0.888, and the combined reliability coefficient is 0.910.

4. Research Result

4.1. Descriptive Statistics and Related Analysis. In this study, SPSS25.0 was used for descriptive statistical analysis, and the results of the mean, standard deviation, and correlation coefficient of each variable are shown in Table 5. Person-post matching and job satisfaction ($r = 0.522$, $P < 0.01$), person-organization matching and job satisfaction ($r = 0.567$, $P < 0.01$) are all significantly positively correlated; person-post matching and unsafe behavior ($r = -0.355$, $P < 0.01$), person-organization matching and unsafe behavior ($r = -0.4777$, $P < 0.01$), job satisfaction and unsafe behavior ($r = -0.701$, $P < 0.01$) are significantly negatively correlated, and the relevant assumptions are obtained after initial verification.

4.2. Confirmatory Factor Analysis. In this study, AMOS26.0 software was used to carry out a confirmatory factor analysis and test of related variables to examine the discriminative validity of five constructs: hardy personality, person-post match, person-organization match, job satisfaction, and unsafe behavior. The analysis results are shown in Table 6. According to the results of confirmatory factor analysis, it is found that the five-factor model fits the actual data best ($\chi^2/\mathrm{df} = 1.295$, $\mathrm{CFI} = 0.986$, $\mathrm{AGFI} = 0.922$, $\mathrm{IFI} = 0.986$, $\mathrm{NFI} = 0.942$, $\mathrm{RMSEA} = 0.024$). The five study variables of the theoretical hypothesis distinguished good validity.

Table 5: Variable description statistical analysis.

Variables	M	SD	1	2	3	4	5	6	7	8
(1) Age	2.55	0.961	1							
(2) Education level	2.34	1.11	0	1						
(3) Length of service	2.05	0.989	0	-0.417^{**}	1					
(4) Hardy personality	3.4576	0.6624	0.008	0.017	0.051	1				
(5) Person-post matching	4.0816	0.86301	−0.067	−0.04	−0.011	0.385^{**}	1			
(6) Person organization match	4.2861	0.76685	0.049	-0.093^{*}	0.072	0.426^{**}	0.439^{**}	1		
(7) Job satisfaction	4.2139	0.82717	0.19	0.003	0.047	0.577^{**}	0.522^{**}	0.567^{**}	1	
(8) Unsafe behavior	4.0853	0.82498	-0.133^{**}	−0.021	−0.083	-0.569^{**}	-0.355^{**}	-0.477^{**}	-0.710^{**}	1

Note: *, $P < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

Table 6: Model fit test.

Models	χ^2	χ^2/df	CFI	AGFI	RMSEA	NFI	IFI
Judgment criteria	The smaller, the better	<3.0	>0.9	>0.8	<0.08	>0.9	>0.9
Five-factor model	593.272	1.295	0.986	0.922	0.024	0.942	0.986
Four-factor model	1585.827	2.697	0.907	0.810	0.058	0.860	0.907
Three-factor model	3658.808	6.191	0.713	0.480	0.102	0.676	0.714
Two-factor model	4513.207	7.611	0.633	0.422	0.115	0.601	0.634
One-factor model	5569.040	9.375	0.534	0.364	0.129	0.507	0.536

Note: Five-factor model: hardy personality, person-post match, person-organization match, job satisfaction, and insecure behavior; four-factor model: hardy personality and person-organization match combined into one factor, person-post match, job satisfaction, and insecure behavior; three-factor model: hardy personality, person-post match, and person-organization match combined into one factor, job satisfaction, and insecure behavior; two-factor model: hardy personality, person-post match, person-organization match, and job satisfaction combined into one factor and insecure behavior; one-factor model: all variables aggregated into one factor.

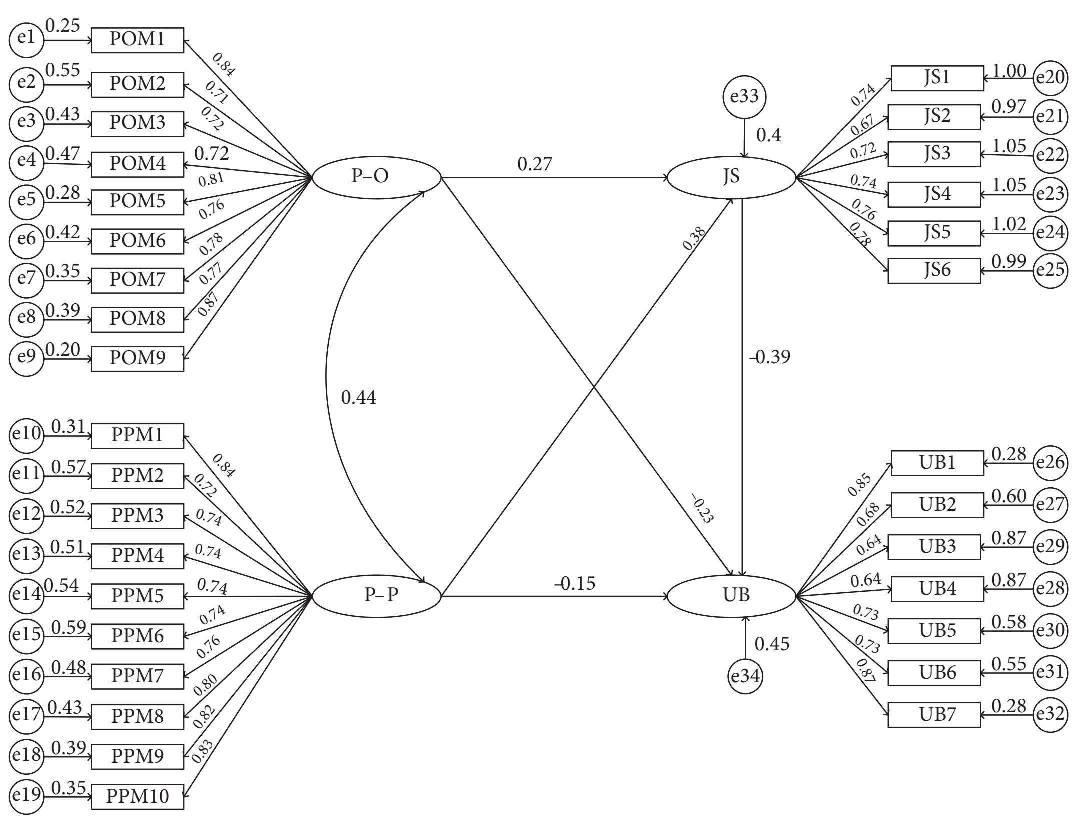


Figure 2: Structural equation model and standardized path diagrams.

4.3. *Hypothetical Test.* After confirming that the measurement model has a good degree of fit, this paper further verifies all the research hypotheses. The analysis result of the structural equation model is shown in Figure 2 (P-P indicates person-job match; P-O indicates person-organization match; JS indicates job satisfaction; and UB indicates unsafe behavior.). Among them, POM1–POM9 are the observed variables matched by person-organization; PPM1–PPM10 are the observed variables matched by person and post; JS1–JS6 are the observed variables of job satisfaction; UB1–UB7 are unsafe observed variables of behavior; e_1–e_{34} are the residuals of each observed variable. It can be seen from Figure 2 that person-post matching has a significant positive effect on job satisfaction ($\beta = 0.38$, $p < 0.05$), hypothesis 1 has been verified; person-organization matching has a significant positive effect on job satisfaction ($\beta = 0.27$, $p < 0.05$), hypothesis 2 is verified; person-post matching has a significant negative impact on unsafe behaviors ($\beta = -0.15$, $p < 0.01$), hypothesis 3 is verified; person-organization matching has significant effects on unsafe behaviors Negative influence ($\beta = -0.23$, $p < 0.05$), hypothesis 4 is verified; job satisfaction has a significant negative influence on unsafe behavior ($\beta = -0.39$, $p < 0.05$), hypothesis 5 is verified.

4.4. Test of the Moderating Effect of Hardy Personality. This paper uses SPSS25.0 to test the moderating effect of hardy personality. The test results are shown in Table 7. It shows that the interaction item of hardy personality and job satisfaction have significant negative regulation of unsafe behaviors ($r = -0.113$, $p < 0.001$). Hypothesis 7 has been verified.

To further confirm the moderating role of hardy personality between job satisfaction and unsafe behaviors, the hardy personality level was divided into high and low groups, and subjects with scores above 1 standard deviation were regarded as the high hardy personality group, and vice versa as the low hardy personality group. And then do a regression analysis of job satisfaction on unsafe behaviors. The results are shown in Figure 3. High hardy personality level and high job satisfaction level have the lowest impact on unsafe behaviors, that is, high hardy personality can weaken the impact of unsafe behaviors on high-risk employees when job satisfaction is low.

Table 7: Test results of main effect and the moderating effect.

Variables	Without adjustment		With adjustment
	Job satisfaction	Unsafe behavior	Unsafe behavior
Age	0.034	-0.179^{***}	-0163^{**}
Education level	0.103^{*}	-0.016	0.025
Length of service	0.034	0.059	0.057
Person-post matching	0.344^{***}	-0.19^{***}	
Person-organization matching	0.258^{***}	-0.209^{***}	
Job satisfaction		-0.343^{***}	-0.491^{***}
Hardy personality			-0.152^{***}
Hardy personality ∗ job satisfaction			-0.113^{***}
R^2	0.263	0.365	0.299
$\triangle R^2$	0.255	0.343	0.018
F	35.249^{***}	47.336^{***}	42.302^{***}

Note: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

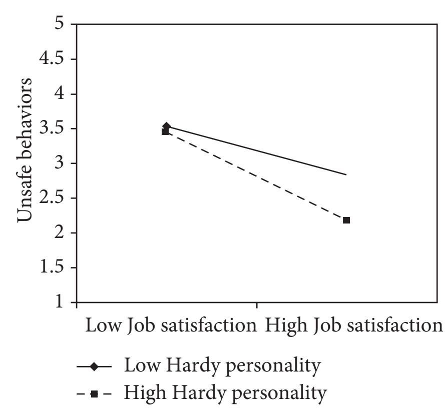


Figure 3: The moderating effect of hardy personality on job satisfaction and unsafe behavior.

4.5. Mediation Test and Adjusted Mediation Test. This paper used the Bootstrap method on the structural model to test the total effect and the mediating effect of job satisfaction between person-post matching and unsafe behavior, and person-organization matching and unsafe behavior. The Bootstrap sampling frequency is set to 5000, the confidence level is set to 95%, and the confidence interval estimation method of deviation correction is used to estimate the interval. The test results of the mediation effect of job satisfaction are shown in Table 8.

It can be seen from the table that job satisfaction plays a complete mediating effect in the relationship between person-post matching affecting unsafe behaviors and person-organization matching affecting unsafe behaviors. The lower limit and upper limit of the confidence interval do not include a value of 0, indicating a complete mediation of job satisfaction. The effect is significant, and Hypothesis 6 is verified. Therefore, the improvement of the person-post matching degree and the person-organization matching degree of high-risk employees can improve the individual job satisfaction of employees, thereby reducing the occurrence of unsafe behaviors and reducing the incidence of accidents.

This paper again used the bootstrap method to verify the significance of the moderated mediation effect model. The test results showed that hardy personality can adjust job satisfaction and has a significant indirect effect on the unsafe behavior of employees in high-risk positions (indirect effect value is −0.253, 95% confidence interval level is [−0.293, −0.016]). In addition, the interaction item (cognitive hardiness ∗ job satisfaction) has a path coefficient of −0.253 for unsafe behaviors of employees in high-risk positions, $p = 0.005$, indicating that the higher the level of hardy personality, the person-post match and the person-organization match respond through job satisfaction. The weaker the effect of unsafe behavior, the stronger the opposite.

5. Management Enlightenment

Based on the matching perspective and individual personality perspective, this paper constructed the role model of human-job matching and human-organization matching on

Table 8: Mediating effect analysis results.

Effect	Path relationship	Effect size	95% confidence interval
Total effect	Person-post matching ⟶ unsafe behavior	−0.289	[−0.338, −0.243]
	Person organization match ⟶ unsafe behavior	−0.365	[−0.370, −0.285]
Direct effect	Person-post matching ⟶ job satisfaction ⟶ unsafe behavior	−0.145	[−0.191, −0.099]
	Person organization match ⟶ job satisfaction ⟶ unsafe behavior	−0.252	[−0.272, −0.175]
Indirect effect	Person-post matching ⟶ job satisfaction ⟶ unsafe behavior	−0.144	[−0.177, −0.118]
	Person organization match ⟶ job satisfaction ⟶ unsafe behavior	−0.113	[−0.132, −0.075]

the influence mechanism of unsafe behavior of employees in high-risk positions, and found that employee satisfaction and hardiness personality of employees in high-risk positions have a significant negative influence on unsafe behavior. Therefore, this paper provides management suggestions on how to improve the job satisfaction of employees in high-risk positions and enhance the psychological hardiness of individual employees for the safety management of employees in high-risk positions.

(1) Improving the match between employees in high-risk positions and their jobs. Managers of high-risk positions should not only train employees with job-related knowledge and skills, but also focus on providing employees with positions that match their abilities, fully mobilize employees' initiative and enthusiasm, and thus enable employees to effectively improve their level of safety behavior skills.

(2) Improving the match between individual employees in high-risk positions and the organization. Managers should communicate and communicate with employees promptly on time to provide emotional care and spiritual support and encouragement for employees in high-risk positions, so that the employees can quickly integrate with the organization and establish safety values and goals consistent with the organization, which can eliminate employees' psychological isolation, cultivate employees' initiative to participate in safety activities and awareness of safety management, and improve employees' ability to independently solve hidden safety problems.

(3) Focus on employee job satisfaction and fulfillment. Due to the special nature of the industry of high-risk jobs, managers should pay great attention to the job satisfaction of employees, and can create a harmonious working atmosphere by improving wages, vocational training, and welfare.

(4) Strengthen the level of resilience of trained employees. A high level of hardiness can weaken the negative impact of job satisfaction on unsafe behavior, so managers should take reasonable measures to continuously strengthen the training of the hardiness level of employees in high-risk positions and improve their ability to withstand and resist stress in the face of work difficulties. When employees are faced with an unsafe accident, they can face the danger and effectively mitigate the adverse effects of negative events on individual employees and reduce the incidence of unsafe behavior accidents.

6. Conclusion

In this paper, the empirical test found that person-post matching and person-organization matching have a significant negative influence on employee insecurity in high-risk positions; job satisfaction plays a fully mediating role between man-position matching and insecurity in high-risk positions, and the fully mediating role of job satisfaction is regulated by the hardy personality, the higher the level of the hardy personality of employees whose job satisfaction has a negative influence on insecurity will be weakened. The above findings provide a reference value for the management of employees in high-risk positions in order to weaken the damage of safety accidents [46–49].

Data Availability

The dataset 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 National Natural Science Foundation of China (51604216 and 51874237), Shaanxi Provincial Social Science Foundation Project (2020R010), The Ministry of Education of Humanities and Social Science Project (21YJA630050), and Xi'an Social Science Planning Fund Project (22GL38).

References

[1] S. Li, *Research on the Influence of Person-Job Fit Degree on Unsafe Behavior of Coal miners[D]. Xi'an*, Xi'an University of Science & Technology, 2018.

[2] N.W. Li, "Analysis and countermeasures of safety drive loss of employees in high-risk positions," *Safety Now*, vol. 41, no. 12, pp. 1–6, 2020.

[3] Li Ru and S. Zhao, "New Changes and Coping Strategies of the person-environment fit in Digital Economy[J]," *Nanjing Journal of Social Sciences*, vol. 11, no. 09, pp. 37–44, 2021.

[4] 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[J]," *Personnel Psychology*, vol. 58, no. 2, pp. 281–342, 2005.
[5] H. Zhao, "Long Lirong. Person-organization fit and job satisfaction: Comparative research on value Congruence, Need-supply fit and demand-ability fit[J]," *Industrial Engineering & Management*, vol. 14, no. 04, pp. 113–117+131, 2009.
[6] M. Zhang, Y. Chang, and C. Long, "The Dominance analysis on the job satisfaction of Kindergarten teachers: Feeling exchange, value Fitness, material exchange, distributive Justice and Workload[J]," *Journal Of Southwest University (Natural Science Edition)*, vol. 38, no. 08, pp. 162–166, 2016.
[7] L. Huang, "Research on the relationship between value matching and employee satisfaction[J]," *Modern Business*, vol. 8, no. 04, pp. 55–57, 2021.
[8] J. Yang, "An empirical study on the influence factors of the staff working Efficiency of the cultural industry person-organization values fit[J]," *Special Zone Economy*, vol. 20, no. 04, pp. 65–69, 2016.
[9] E. A. Locke, *The Nature and Causes of Job Satisfaction: The Handbook of Industrial and Organizational Psychology [R]*, R and McNally, Chicago, IL, 1976.
[10] C. Wang, "Analysis of factors affecting job satisfaction[J]," *Marketing Research*, vol. 22, no. 02, pp. 29–32, 2005.
[11] S. P. Robbins, *Organizational behavior[M]*, pp. 7–752, Pearson Education, UK, 2008.
[12] T. A. Judge, D. Heller, and M. K. Mount, "Five-factor model of personality and job satisfaction: a meta-analysis," *Journal of Applied Psychology*, vol. 87, no. 3, pp. 530–541, 2002.
[13] T. A. Judge and R. J. Larsen, "Dispositional affect and job satisfaction: a Review and theoretical Extension," *Organizational Behavior and Human Decision Processes*, vol. 86, no. 1, pp. 67–98, 2001.
[14] Yu Yue and M. Zhou, "The effect of work-Family balance on work engagement and job satisfaction: the moderating of personality[J]," *Chinese Journal of Clinical Psychology*, vol. 24, no. 03, pp. 504–508+513, 2016.
[15] T. T. Kim, O. Karatepe, and G. Lee, "Psychological contract breach and service innovation behavior: psychological capital as a mediator[J]," *Service Business*, vol. 12, no. 2, pp. 1–25, 2017.
[16] W. Chen and C. Wang, "The mechanism of the effect of person-job matching and person-organization matching on work attitudes[J]," *Journal of Psychological Science*, vol. 7, no. 04, pp. 979–981, 2007.
[17] J. Jia, Y. An, and M. Sun, "Research on the control mechanism of knowledge employee turnover based on value matching [J]," *Science and Technology Management Research*, vol. 29, no. 07, pp. 341–343+340, 2009.
[18] L. Long and H. Zhao, "Individual-organization Fit: the dominant effect of performance and social responsibility[J]," *Chinese Journal of Management*, vol. 6, no. 06, pp. 767–775, 2009.
[19] P. Liu and Y. Wu, "An analysis of the impact of individual-organization matching on the job performance of migrant workers—with employee relationship as a mediating variable [J]," *Chinese Rural Economy*, vol. 23, no. 06, pp. 63–71, 2011.
[20] X. Tan, "Effect of person-organization value fit on employees' job involvement and perceived organizational support[J]," *Journal of Psychological Science*, vol. 35, no. 04, pp. 973–977, 2012.
[21] Y. Xi, C. Dai, and H. Yang, "Person-organization value fit, job satisfaction and turnover intention[J]," *Journal Of Nanjing Normal University (Social Science Edition)*, vol. 25, no. 01, pp. 38–47, 2014.
[22] H. Zhao, "The impact mechanism of person-organization value fit on affective organizational commitment[J]," *Business And Management Journal*, vol. 14, no. 11, pp. 165–175, 2015.
[23] X. Sun and C. Zhou, "The Influence of person-job fit on preschool teachers' job burnout: mediating of job satisfaction and moderating of perceived organizational support[J]," *Studies in Early Childhood Education*, vol. 21, no. 01, pp. 42–53, 2020.
[24] H. J. Arnold and D. C. Feldman, "A multivariate analysis of the determinants of job turnover," *Journal of Applied Psychology*, no. 3, pp. 350–360, 1982.
[25] H. Wang and J. Yang, "The effect of employee total rewards and demands abilities fit on job satisfaction: taking non-state-owned enterprises for example[J]," *On Economic Problems*, vol. 7, no. 05, pp. 73–78, 2015.
[26] J. Cheng, Y. Su, and Y. Wang, "A job fit, employability and turnover intention: the mediating effect of job satisfaction[J]," *Contemporary Finance & Economics*, no. 02, pp. 66–75, 2015.
[27] Y. Fang, L. Qin, and L. Zhang, "Influence of safety culture and the sense of job satisfaction on the safety behaviors of the employees—an empirical research on the structural equation model[J]," *Journal of Safety and Environment*, vol. 19, no. 06, pp. 2022–2032, 2019.
[28] Q. Zhang and Y. Yang, "The impact of value matching on work engagement: Double intermediary effect based on organizational Identification and job satisfaction[J]," *Journal of Liaoning University (Philosophy and Social Sciences)*, vol. 48, no. 02, pp. 59–68, 2020.
[29] S. Liu, "Organizational culture and job satisfaction: the moderating effect of personality[J]," *Acta Scientiarum Nauralium Universitaties Pekinensis*, vol. 9, no. 02, pp. 251–256, 2007.
[30] Y. Huang and Q. Zhou, "Research on the moderating effects of the Locus of control in employees' turnover intention [J]," *Journal of Beijing Institute of Technology (Social Sciences Edition)*, vol. 12, no. 04, pp. 32–37, 2010.
[31] C. Mao and J. Sun, "Impact of workplace incivility: the moderating effect of proactive personality[J]," *Chinese Journal of Management*, vol. 10, no. 05, pp. 708–714, 2013.
[32] H. Jinlei and S. Jiang, "Research on the influence of individual characteristics on Farmers' Entrepreneurial intention——based on the moderating effect of proactive personality[J]," *Social Sciences in Guangxi*, vol. 1, no. 06, pp. 84–88, 2016.
[33] X. Chen and J. Xiong, "The effect of workplace exclusion on the adaptive performance of knowledge workers: the moderating role of personality traits[J]," *Leadership Science*, vol. 21, no. 32, pp. 44–46, 2017.
[34] R. Liu, H. Cao, and L. Liu, "An empirical study on psychological factors of information Sharing in social Commerce Websites[J]," *Soft Science*, vol. 31, no. 12, pp. 101–107, 2017.
[35] E. R. Dogaheh, M. Khaledian, and A. R. Mohammadiarya, "The relationship of psychological hardiness with emotional Intelligence and Workaholism[J]," *International Letters of Social and Humanistic Sciences*, vol. 5, no. 5, pp. 1–9, 2013.
[36] L. Li, B. Lei, and C. Mao, "Digital Twin in Smart Manufacturing," *Journal of Industrial Information Integration*, vol. 26, no. 9, Article ID 100289, 2022.
[37] X. Zhang, S. Deng, and Z. Zou, "Cognitive hardiness: a cognitive style buffering between Nonfit organization and turnover intention[J]," *Chinese Journal of Applied Psychology*, vol. 26, no. 03, pp. 250–258, 2020.

[38] L. Li, T. Qu, Y. Liu et al., "Sustainability assessment of Intelligent Manufacturing supported by Digital Twin," *IEEE Access*, vol. 8, pp. 174988–175008, 2020a.

[39] L. Li and C. Mao, "Big data supported PSS Evaluation Decision in service-Oriented Manufacturing," *IEEE Access*, vol. 8, pp. 154663–154670, 2020.

[40] L. Li, C. Mao, H. Sun, Y. Yuan, and B. Lei, "Digital Twin driven green performance Evaluation methodology of Intelligent Manufacturing: Hybrid model based on Fuzzy Rough-Sets AHP, Multistage Weight Synthesis, and PROMETHEE II," *Complexity*, vol. 2020, no. 6, pp. 1–24, 2020b.

[41] G. Ni, L. Gao, and X. Wang, "Research on correction mechanism of job satisfaction on unsafe behavior of construction workers[J]," *Journal of Safety Science and Technology*, vol. 17, no. 09, pp. 156–162, 2021.

[42] J. Wang, X. Wang, and C. Wang, "Research on coal miners' unsafe behavior based on job satisfaction[J]," *China Safety Science Journal*, vol. 28, no. 11, pp. 14–20, 2018.

[43] A. Chuang, C. T. Shen, and T. A. Judge, "Development of a Multidimensional Instrument of person-environment fit: the perceived person-environment fit scale (PPEFS)," *Applied Psychology*, vol. 65, no. 1, pp. 66–98, 2016.

[44] T. Wu, *Study of the Relationship between Person-Job Fit and Job Performance[D]*, Hangzhou University of Electronic Science and Technology, 2015.

[45] D. M. Cable and D. S. DeRue, "The convergent and discriminant validity of subjective fit perceptions," *Journal of Applied Psychology*, vol. 87, no. 5, pp. 875–884, 2002.

[46] A. H. Busari, Y. H. Mughal, S. N. Khan, S. Rasool, and A. A. Kiyani, "Analytical cognitive style moderation on promotion and turnover intention," *The Journal of Management Development*, vol. 36, no. 3, pp. 438–464, 2017.

[47] Y. Peng and C. Mao, "The impact of person–job fit on job satisfaction: the mediator role of Self Efficacy," *Social Indicators Research*, vol. 121, no. 3, pp. 805–813, 2015.

[48] T. Straatmann, S. Königschulte, K. Hattrup, and K. C. Hamborg, "Analysing mediating effects underlying the relationships between P–O fit, P–J fit, and organisational commitment," *International Journal of Human Resource Management*, vol. 31, no. 12, pp. 1533–1559, 2020.

[49] Li Yan, Z. Zhao, and S. Tian, "Review of research on unsafe behavior of miners [J]," *Journal of Safety Science and Technology*, vol. 12, no. 08, pp. 47–54, 2016.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 8274618, 10 pages
https://doi.org/10.1155/2022/8274618

Research Article

Modeling and Simulation of Restorative Indoor Environment Based on Neural Network

Xianyao Ding,[1] **Haocheng Feng**,[2] **and Weilun Shan**[1]

[1]*Dong-eui University, Busan 47340, Republic of Korea*
[2]*Japan Advanced Institute of Science and Technology, Nomi 9231211, Japan*

Correspondence should be addressed to Weilun Shan; sjrrjs@dlut.edu.cn

Received 10 July 2022; Revised 4 August 2022; Accepted 13 August 2022; Published 14 September 2022

Academic Editor: R. Mo

In the busy modern society, after experiencing fast-paced work, what people most desire is to have a relatively relaxed and comfortable indoor rest scene. At the same time, with the improvement of material living standards, people have higher and higher requirements for the office and indoor rest environment, and the indoor temperature environment is very important to create a sense of comfort. A good indoor somatosensory temperature environment can not only convey the degree of somatosensory comfort but also relax people's bodies, in which the air conditioning control system can restore the indoor temperature environment. Based on the abovementioned problems, this paper uses the neural network modeling method, based on the indoor comfort index SET* value and other factors that can affect the environmental temperature, aiming at the human body temperature comfort, constructs a BP neural network model with indoor environmental parameters as the input index, and uses the improved particle swarm optimization algorithm to optimize the model, so as to realize the real-time control of SET* value, and then analyzes the relationship between indoor environmental thermal factors and SET* value, The model is optimized through simulation experiments to improve the optimization degree of the model. The practice shows that the restorative indoor environment model based on improved BP neural network constructed in this paper can optimize the room parameters well, so as to improve indoor comfort and give people a warm accommodation environment.

1. Introduction

WTO points out that health refers to people having a relatively healthy physical and mental state, and being able to better adapt to society, rather than simply referring to physical diseases. Only when you are physically and mentally healthy can you call it health [1]. Health is not only about balancing yourself but also about balancing the relationship between yourself and your surroundings.

In the view of environmental psychology, the connection between people and the environment exists naturally. Some external environments not only provide us with daily needs but also inject new spiritual sources into us on the premise of ensuring the material basis. This is the origin of "restorative environment." Running around the bustling city 24 hours a day, people seem to be firmly shackled. Different types of environmental pollution such as haze and noise are all around, and people are gradually far away from the simplest natural living state. In the long run, anxiety and anxiety will breed [2]. Nowadays, more and more subhealth people exist, and there are also more and more people with psychological problems [3–5]. Based on this, the vast majority of people want to be able to concentrate on their work tasks without external interference. According to the theory of self-control force model, if you want to maintain concentration in a noisy environment, you must consume self-control resources, resulting in fatigue [6]. Fortunately, this situation can compensate for self-control resources by eating, resting, and other different ways [7]. Baumeister and others made a very vivid metaphor, comparing self-control to muscles, which will be tired after use and recover to a certain extent after rest [8].

The research on the restorative environment has a relatively long history and has formed a quite deep theoretical foundation. In the 19th century, Olmsted presented the word "rehabilitation" to the world. In his view, people can get relaxation and calm from the natural environment, and self-heating can also help urban residents release pressure. This discovery caused a shock in the field of environmental psychology. Then Kaplan and Talbot formally put forward the term "restorative environment," believing that the restorative effect of the natural environment is quite good, which is conducive to people's physical and mental health. With the continuous development of environmental psychology, until the middle and late last century, the research on restorative environment has achieved a great breakthrough and formed a lot of mature theories and views, such as environmental load theory, wake-up theory, and so on. Among them, Ulrich et al. have repeatedly demonstrated that people can effectively relieve pressure and burden by being close to nature and contacting nature, and this recovery effect has encountered great resistance in the urban environment [9]. Based on the abovementioned theory, the subsequent research on this theory mainly focuses on the application and practice of environmental space, considers the effect of rehabilitation through various methods such as experiments and on-site investigations, and further deepens and expands this theory by using data. Korpela et al. found that restorative experience is related to people's preferred places. In fact, people can recover their emotions to a certain extent in familiar scenes and places [10]. et al. analyzed the environmental restoration of the monastery by means of a questionnaire and believed that the restoration function of the monastery did exist [11]. By investigating the restoration of urban parks and forests in Zurich, et al. found that sports in green spaces can relieve physical and mental stress, improve his own happiness index, and support the theory of restorative environment [12]. Gutierrez et al. deeply studied the application of restorative landscape in the campus environment, so as to reduce the load and pressure of teachers, students, and employees, which further expanded the restorative environment theory in the practice of space [13]. In addition to the abovementioned researchers and their research results, there are also a large number of research books, documents, and materials on restorative environmental theory abroad. They have made an in-depth analysis of the theory and its application effect in specific environmental sites through visits, investigations, field visits, logical derivation, and other methods, and also provided a good theoretical basis and scientific guidance for this study.

With the rapid development of science and technology, people are more inclined to spend most of their time indoors. Therefore, whether the indoor environment meets the thermal comfort of human body not only has a significant impact on people's work efficiency but also affects people's physical and mental health. Nowadays, almost all traditional air-conditioning control systems only use single temperature control to adjust the indoor thermal environment, not directly based on people's thermal comfort. Such a control method is not only of general effect but also great energy consumption. Therefore, the author believes that we should first choose an appropriate evaluation index of thermal comfort, evaluate whether the human body is in the thermal comfort feeling according to this index, and control it to make it tend to the range of human comfort. Only in this way can the indoor environment truly meet the thermal comfort of the human body.

Nevertheless, there is relatively little research on the restorative theory of the indoor environment. However, people spend more than 80% of their time indoors. A good indoor environment is conducive to improving work efficiency. At present, air conditioning systems only use temperature to measure indoor environmental comfort, ignoring other variables related to human thermal comfort [14–16]. Therefore, in many cases, it cannot meet people's needs for thermal comfort, and it will also lead to an increase in the operating cost and energy consumption of the air conditioning system. Now we use the standard effective temperature SET* index recommended by ASHRAE standard and widely used to measure the thermal environment. SET* is directly related to people's thermal feeling rather than the air temperature. It was proposed by Gagge and comprehensively considered the effects of temperature, humidity, average radiation temperature, wind speed, etc., under the influence of different activity levels and clothing thermal resistance. SET* can more effectively reflect the thermal comfort of the human body than simple air temperature [17–20]. The subjective evaluation criteria of heat sensation are divided into seven levels: cold, cool, slightly cool, moderate, warm, slightly warm, and hot. When the indoor environment is to be comfortable, the range of SET* index is [21, 22].

The SET* index is based on the nonlinear and time-varying characteristics of the thermal comfort index of the physiological response model. The traditional calculation method needs to calculate the influence of air temperature, humidity, wind speed, average radiation temperature, and other parameters on the temperature and humidity of human skin through repeated iterations. The calculation formula is complex and cannot be determined in real-time. Therefore, it is difficult to meet the requirements of real-time control of the air conditioning system [23–25]. Many studies assume that the sample data is taken within a certain range, but the measured sample data are more conducive to the training of SET* index model.

Therefore, in this paper, four environmental factors are obtained through actual measurement as sample data, and the particle swarm optimization algorithm is used to optimize BP neural network. The method can calculate the value of SET* in real-time, solve the complex iterative operation, and improve the convergence speed and prediction accuracy of the neural network. The corresponding neural network model is established.

2. Neural Network Model and Optimization Algorithm

2.1. BP Neural Network Model. BP (back propagation) structure is shown in Figure 1. Neurons are arranged in layers, independent within layers, and fully connected between layers. Information is transmitted and processed forward in one direction in the network, and there is no feedback loop. Where, $x = (x_1, x_2, \ldots, x_n)$ represents an n-dimensional input sequence composed of n feature

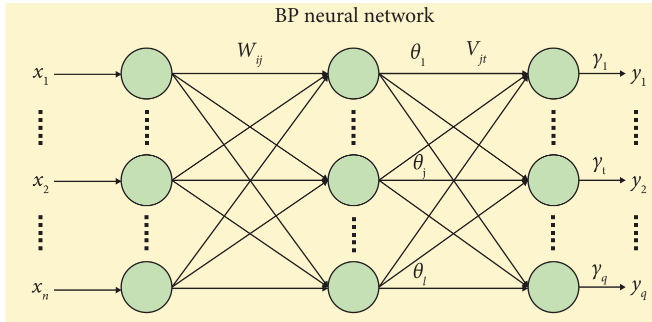


FIGURE 1: BP neural network structure.

information; W_{ij} is the connection weight between the input layer and the hidden layer; V_{jt} is the connection weight between hidden layer and output layer neurons, and $\theta_j\,(j = 1, 2, \ldots, l)$ is the threshold of hidden layer; $\gamma_t\,(t = 1, 2, \ldots, q)$ is the threshold of the output layer; $y = (y_1, y_2, \ldots, y_q)$ is the model output sequence; $\varphi(\cdot), \psi(\cdot)$ represents the activation function of the network middle layer and output layer, respectively, [26–28].

2.2. BP Neural Network Learning Process. The learning process of BP neural network is as follows:

(1) First, we initialize the network learning parameters of the new BP neural network model, update the weight threshold of the network, and specify the training error E, rated accuracy, and learning rate according to experience or applicable trial and error method η, the learning rate is [0, 1];
(2) After modifying some basic parameters of the network, we input the training set data into the model, and then calculate the forward output of each node of the network;
(3) After completing the second step, it is also necessary to calculate the deviation between the expected output and the network output;
(4) After the third step, calculate the parameter correction information of each layer;
(5) Repeatedly adjust the weights of each neuron node. If $E < \eta$ or the maximum number of training times is reached, the algorithm ends.

2.2.1. Signal Forward Propagation Process. The input signal u_j of the j-th node of the hidden layer is as follows:

$$u_j = \sum_{i=1}^{n} W_{ij}x_i + \theta_j. \tag{1}$$

The output signal s_j of the j-th node after the hidden layer passes through the function $\varphi(\cdot)$ and is activated as:

$$\begin{aligned} s_j &= \varphi(u_j) \\ &= \varphi\left(\sum_{i=1}^{n} W_{ij}x_i + \theta_j\right). \end{aligned} \tag{2}$$

The input signal c_t of the t-th node of the output layer is as follows:

$$\begin{aligned} c_t &= \sum_{j=1}^{l}\left(V_{jt}s_j + \gamma_t\right) \\ &= \sum_{j=1}^{l}\left(V_{jt}\varphi\left(\sum_{i=1}^{n} W_{ij}x_i + \theta_j\right) + \gamma_t\right). \end{aligned} \tag{3}$$

After the output layer is activated by function $\psi(c_t)$, the output signal y_t of the t-th node is as follows:

$$\begin{aligned} y_t &= \psi(c_t) \\ &= \psi\left(\sum_{j=1}^{l}\left(V_{jt}\varphi\left(\sum_{i=1}^{n} W_{ij}x_i + \theta_j\right) + \gamma_t\right)\right). \end{aligned} \tag{4}$$

2.2.2. Signal Back Propagation Process. After the forward signal is processed by the output layer, the deviation between the network output and the actual output is calculated layer by layer, and the weight threshold of each layer of the network is inversely modified according to the gradient descent algorithm to optimize the structural parameters [29].

Suppose that the training sample contains P data. For a single sample data point $p\,(p = 1, 2, \ldots, P)$, the error is as follows:

$$E^{(p)} = \frac{1}{2}\sum_{t=1}^{q}\left(\hat{y}_t^{(p)} - y_t^{(p)}\right)^2. \tag{5}$$

The error of the whole training sample set is as follows:

$$E^{(P)} = \frac{1}{2}\sum_{p=1}^{P}\sum_{t=1}^{q}\left(\hat{y}_t^{(p)} - y_t^{(p)}\right)^2. \tag{6}$$

Where $\hat{y}_t^{(p)}$ and $y_t^{(p)}$ represent the actual output and network output of the t-th node of the output layer when the input sample is p.

$$\begin{aligned} \Delta V_{jt} &= -\eta\frac{\partial E}{\partial V_{jt}}, \\ \Delta\gamma_t &= -\eta\frac{\partial E}{\partial\gamma_t}, \\ \Delta W_{ij} &= -\eta\frac{\partial E}{\partial W_{ij}}, \\ \Delta\theta_j &= -\eta\frac{\partial E}{\partial\theta_j}. \end{aligned} \tag{7}$$

It can be obtained after correction as:

$$\begin{aligned}
\Delta V_{jt} &= -\eta \sum_{p=1}^{P} \sum_{t=1}^{q} \left(\widehat{y}_t^{(p)} - y_t^{(p)}\right)\psi'(c_t),\\
s_j &= \eta\delta_t s_j,\\
\Delta\gamma_t &= -\eta \sum_{p=1}^{P} \sum_{t=1}^{q} \left(\widehat{y}_t^{(p)} - y_t^{(p)}\right)\psi',\\
(c_t) &= \eta\delta_t,\\
\Delta W_{ij} &= -\eta \sum_{p=1}^{P} \sum_{t=1}^{q} V_{jt}\left(\widehat{y}_t^{(p)} - y_t^{(p)}\right)\psi'(c_t)\varphi'\left(u_j\right),\\
x_i &= \eta\delta_j x_i,\\
\Delta\theta_j &= -\eta \sum_{p=1}^{P} \sum_{t=1}^{q} V_{jt}\left(\widehat{y}_t^{(p)} - y_t^{(p)}\right)\psi'(c_t)\varphi'.\left(u_j\right) = \eta\delta_j,
\end{aligned} \tag{8}$$

where δ_t and δ_j represent the error signals of the output layer and the hidden layer, respectively.

2.3. Particle Swarm Optimization Algorithm. Particle swarm optimization (PSO) algorithm is a mobile search process in which Dr. Kennedy and Dr. Eberhart, after an in-depth analysis of the natural bird predation process, idealize the individual of the birds swarm into a particle, draw lessons from the individuality and sociality of each bird in the group behavior, and simulate and simplify it into a particle to find the individual optimal solution in the feasible solution space and the global optimal solution at the same time [30–33]. In the process of updating each position, the particle should not only refer to the historical optimal value recorded by itself but also consider the optimal value of another individual search, and adjust its search direction under the comprehensive guidance of this information to make the group approach the optimal extreme value.

The position of a single particle in PSO algorithm represents a feasible solution to the optimization problem, which is substituted into the objective function to evaluate the fitness, and the particle position is constantly updated according to the fitness comparison results to solve the problem to be optimized. PSO algorithm first randomly generates an initial population with m particles in the D-dimensional feasible solution space and is randomly equipped with a certain flight speed. For the i-th particle, its particle position is recorded as $\overrightarrow{x}_i = (x_{i1}, x_{i2}, \ldots, x_{iD})$, its flight speed is recorded as $\overrightarrow{v}_i = (v_{i1}, v_{i2}, \ldots, v_{iD})$, the individual historical optimal value is recorded as $\overrightarrow{p}_i = (p_{i1}, p_{i2}, \ldots, p_{iD})$, and the group historical optimal value is recorded as $\overrightarrow{p}_g = (p_{g1}, p_{g2}, \ldots, p_{gD})$ in the iteration process, the particle updates its speed and position through (9) and (10):

$$v_{id}^{k+1} = \omega_{v_{id}}^{k} + c_1 r_1\left(p_{id}^{k} - x_{id}^{k}\right) + c_2 r_2\left(p_{gd}^{k} - x_{id}^{k}\right), \tag{9}$$

$$x_{\mathrm{id}}^{k+1} = x_{\mathrm{id}}^{k} + v_{\mathrm{id}}^{k+1}, \tag{10}$$

where ω is the inertia weight, and the range is within the range of $[0, 1]$; $i = 1, 2, \ldots, m$; $d = 1, 2, \ldots, D$; k Is the current search algebra; c_1 and c_2 are learning factors, which generally take equal nonnegative values; The random numbers r_1 and r_2 are independent of each other and range between [0, 1]. Usually, the position and speed of particles should be prevented from crossing the boundary, which is limited to a specific space, that is, $v_{\mathrm{id}} \in [-v_{\max}, v_{\max}]$, $x_{\mathrm{id}} \in [-x_{\max}, x_{\max}]$, to avoid invalid search and speed up the optimization process.

The standard PSO algorithm flow is as follows:

(1) Initialize the group size, inertia weight, learning factor, initial particle speed and position, and specify the particle speed limit $[-v_{\max}, v_{\max}]$ and position limit $[-x_{\max}, x_{\max}]$;

(2) Calculate the fitness value of each particle according to the objective evaluation function of the problem to be optimized;

(3) Compare the current fitness value of each particle with its own historical optimal function value. If the former evaluation value is better, the individual optimal position will be updated to the current position coordinates, otherwise, it will remain unchanged;

(4) Compare the historical local optimal value of all particles with the global optimal value of the population, and use the particle position with a better fitness function value to eliminate the position of the global optimal value, otherwise, it will remain unchanged;

(5) Update the velocity and position of particles according to equations (9) and (10);

(6) If the maximum number of iterations has been reached or the preset fitness minimum error is met, the algorithm ends, otherwise, return to program (2) to enter the next iteration.

2.4. Restorative Environment Theory. "Restorative environment" is based on the dimension of environment setting, trying to update and restore the physical and mental resources and abilities that have been continuously consumed by human beings. It can effectively reduce people's pressure, reduce people's bad emotions, reduce inner fatigue, and ensure the healthy development of the body and mind. At this stage, the recovery process visible to the naked eye, such as improving the task of directional attention, actively changing emotions all depend on the behavior of individuals using resources [7]. In addition, the healing effect of the environment refers to that the environment can continuously restore and update the physical and mental resources and abilities consumed by human beings. At the same time, at the level of environmental psychology, it can scientifically evaluate the social and psychological attributes of the environment, and can also be effectively applied to physical and mental healing, landscape design, urban planning, and other aspects [9].

We can summarize the positive role of environment in promoting human beings from three aspects: comprehensively improving human long-term happiness and health; recover human mental fatigue in a short time; let human beings quickly get rid of the pain of disease and recover their health as soon as possible. If the abovementioned requirements are met at the same time or in part, it can be called a restorative environment. The indoor ambient temperature in this study, that is, the short-term regulation of temperature, is the expected goal. The design of indoor temperature constructed by neural network modeling technology has a restorative effect and is discussed in the field of industrial design.

Since the 1980s, Kaplan and others have studied the restorative environment in depth by combining theory and empirical research, and successively put forward the theory of attention restoration and decompression. In the process of studying the restorative environment, both of them can be called two core theories. Although they have their own emphasis on evaluation criteria, concept definition, action methods, etc., from the perspective of research and development, the possibility of their final integration is greater, and the evaluation of environmental quality will become increasingly rich and systematic.

2.5. Standard Effective Temperature SET.* The SET* index comprehensively considers different activity levels and clothing thermal resistance. It is based on the human physiological response model and is obtained from the analysis of the physical process of human heat transfer, so the calculation is very complex. The SET* index is based on the two-node model theory of human body temperature regulation. That is, the model is regarded as two layers, the core layer, and the skin layer. Its model can be expressed by two heat balance equations as:

$$
\begin{aligned}
(1-\alpha)mc_{p.b}\frac{(dt_{cr}/dt)}{A_D} &= M-W-(0.023M(44-p_a)\\
&\quad +0.0014M(34-t_a))\\
&\quad -(t_{cr}-t_{sk})\times(5.28+1.163\times skbf),\\
\alpha mc_{Rb}\frac{(dt_{cr}/dt)}{A_D} &= (t_{cr}-t_s)\times(5.28+1.163\times skbf)\\
&\quad -\left(f_{cl}(h_c+h_r)\left(t_k-\frac{h_c t_a+h_r t_{mrt}}{h_c+h_r}\right)\right.\\
&\quad \left.+(0.06+0.94w)16.7h_c(p_{sk}^{*}-p_a)F_{pcl}\right),
\end{aligned} \tag{11}
$$

where t_{cr} and t_{sk} is the temperature of the core layer and skin layer; SkBF is the peripheral blood flow rate (L/hm^2); M is the metabolic rate of human body, and W is the external work of the human body; α Is the proportion of human skin layer; m is the body weight (kg); $c_{p.b}$ is the specific heat capacity of the human body (kJ/kg); A_D is the total skin surface area obtained by Dubois formula; t_a is the air temperature; t_{mrt} is the average radiation temperature; h_r and h_c are convective and radiant heat transfer coefficients; f_{cl} is the area coefficient of clothing; F_{pcl} is the permeability coefficient of clothing; w is skin moisture; p_a and p_{sk}^{*} refers to the water pressure under the air temperature and the saturated water pressure under the skin temperature.

SET* an index that considers the comprehensive effect of air temperature, humidity, wind speed, average radiation temperature, clothing thermal resistance, and different activity levels on human thermal sensation. The calculation of SET* is complex, and it is necessary to repeatedly iterate through the computer. First, the physiological parameters of the human body must be calculated by using the two-node model, and the heat exchange equilibrium equation between humans and the environment can be obtained. If H_{sk} is the heat loss of skin, it is expressed by the following equation as:

$$H_{sk} = h_s(t_k - \mathrm{SET}^*) + wh_{se} \times (p_{sk} - 0.5p_{\mathrm{SET}^*}), \tag{12}$$

where h_s is the standard convective heat transfer coefficient $(W/m^{2°}\cdot C)$; $h_{s,e}$ is the standard evaporation heat transfer coefficient $(W/m^2\cdot kPa)$; p_{SET^*} saturation pressure of water vapor in air at skin temperature (kPa); p_{SET^*} is the saturated water pressure (kPa) at SET*.

3. Algorithm Improvement and Performance Test

3.1. Improvement of Particle Swarm Optimization Algorithm. The standard particle swarm algorithm processes one set of solutions in parallel to update another set of solutions in the process of optimization. The particles move randomly in the complex field, and the balance between local search and global search is achieved through the evolution of particle cooperation and competition. Global convergence and antijamming capability [34, 35]. However, the local search ability of the algorithm is poor, and the optimization result is easy to fall into the local minimum value of the objective function so the algorithm converges slowly in the later stage, and a premature phenomenon occurs.

In view of the defects in the standard PSO algorithm, after further exploring the ways of bird swarm communication, Hu Wang et al. Verified that the search performance of BPSO algorithm has a low correlation with speed, and then abandoned the speed parameter in the iterative update of BPSO, and proposed an improved simplified particle swarm optimization algorithm, referred to as SPSO. SPSO only completes particle optimization through position information, effectively reducing the human interference factors added by the initialization of speed term, the accuracy and stability of the algorithm are improved. The improved particle position update formula is as:

$$x_{\mathrm{id}}^{k+1} = \omega_{x_{\mathrm{id}}}^{k} + c_1 r_1\left(p_{\mathrm{id}}^{k} - x_{\mathrm{id}}^{k}\right) + c_2 r_2\left(p_{\mathrm{gd}}^{k} - x_{\mathrm{id}}^{k}\right). \tag{13}$$

3.2. Performance Test of Improved Particle Swarm Optimization Algorithm. In order to verify that PSO algorithm has better optimization and convergence ability, this paper compares PSO and SPSO through three standard test functions with different characteristics. The specific information of each test function is shown in Table 1.

Table 1: Three classic test functions.

Function	Dimension	Value range	Optimal value
$f_1(x) = \sum_{i=1}^{d}(x_i^2/4000) - \prod_{i=1}^{d}\cos(x_i/\sqrt{i}) + 1$	10	[−600, 600]	0
$f_2(x) = \sum_{i=1}^{d} x_i^2$	2	[−5.12, 5.12]	0
$f_3(x) = 10\,d + \sum_{i=1}^{d}[x_i^2 - 10\cos(2\pi x_i)]$	10	[−5.12, 5.12]	0

Table 2: Test function optimization simulation results.

Function	Algorithm	Optimal solution	Average value	Success rate/%
f_1	bPSO	97.6325	21.0257	0
	sPSO	0	0.0007	100
f_2	bPSO	21.7645	39.8742	0
	sPSO	0.0031	0.0029	100
f_3	bPSO	81.4574	102.4315	0
	sPSO	0	0	100

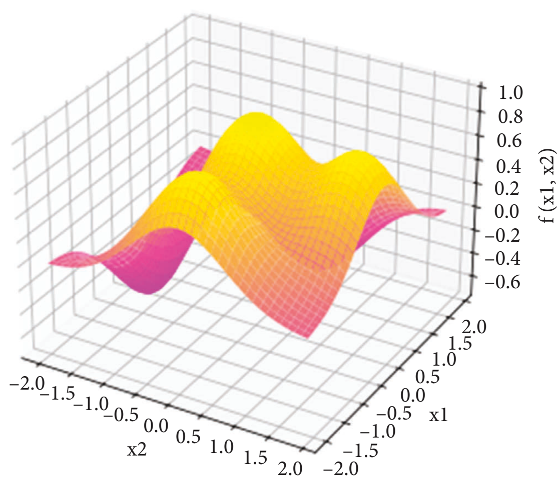


Figure 2: Function f_1 test function image.

The specific test contents are as follows: set the PSO and SPSO algorithm optimization test to be carried out in the same programming environment, specify that the number of particles per generation is 30, the maximum number of optimization is 200, and the two algorithms run 50 times each; inertia weight is set as: $\omega_{\max} = 1.0$, $\omega_{\min} = 0.2$; the acceleration factor of PSO and SPSO algorithm $c_1 = c_2 = 2.0$; during the test, the function to be optimized is the test function; and the value of the function dimension is different, including 10 and 2. The performance of the algorithm is investigated by three indicators: the optimal value, the average value, and the success rate. The simulation results of test function optimization are shown in Table 2.

From the simulation results in Table 2, it can be seen that under the same operating environment and parameter settings, the global optimization effect of the improved SPSO algorithm is ideal, in which f_1, f_2 and f_3 can search for the theoretical optimal solution, reflecting the strong optimization ability of the algorithm for different dimensional functions, while BPSO algorithm only finds the theoretical target value in the solution process of low dimensional f_2

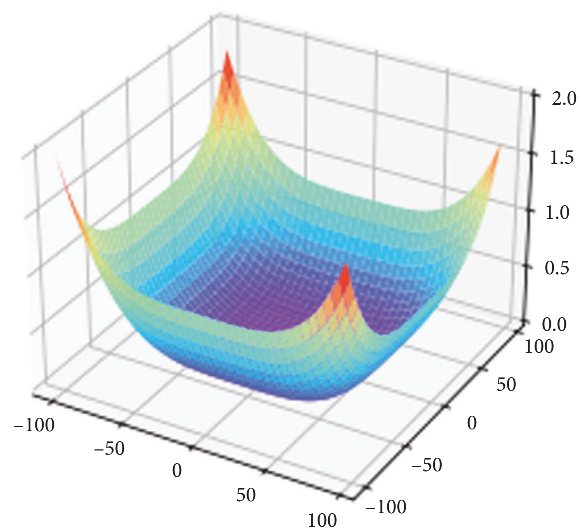


Figure 3: Function f_2 test function image.

function, and its optimal performance is far inferior to that of SPSO algorithm. For the index of success rate, the probability of SPSO algorithm meeting the rated accuracy is 100%. Relatively speaking, the success rate of BPSO is low, even 0. Figures 2–4 shows the images of each test function. It can be seen from Figures 2–4 that the optimization paths between different test functions are different. When f_1 function is optimized, the paths are diverse, but it is not conducive to the rapid generation of optimal results. The path optimization strategies of f_2 and f_3 functions are basically the same, but compared with f_2 function, it can be more selective and faster in terms of optimization paths.

Figures 5–7 is the curve diagram of the optimization process of each test function algorithm. It can be seen from Figures 5–7 that the optimization process curves for different optimization functions are also different. Among them, for function 1, the overall changes tend to be consistent, but the number of iterations required is longer, and it takes more practice; for function 2, the overall performance is better, which is an ideal state; for function 3, although the number of iterations of the improved PSO model is reduced, the function is changing. The fluctuations in the process are large, which is not conducive to the optimization of the model, and the results also have large uncertainties. Therefore, by synthesizing the optimization curves of three different optimization functions, it can be found that the SPSO algorithm is obviously due to the ordinary PSO algorithm, and its outstanding performance is that the convergence speed is faster and the accuracy is higher. Especially in the process of solving complex functions, the optimization ability is more prominent.

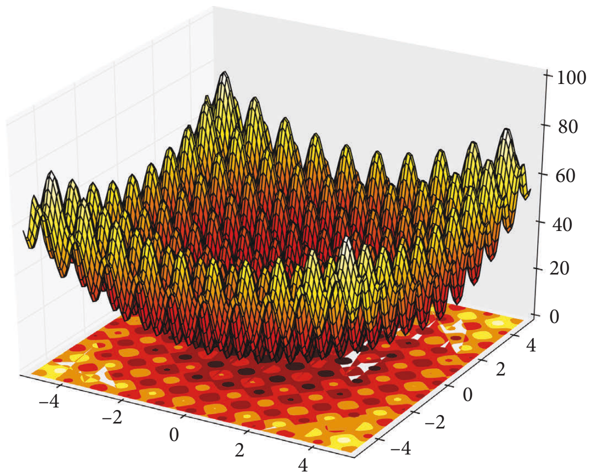


FIGURE 4: Function f_3 test function image.

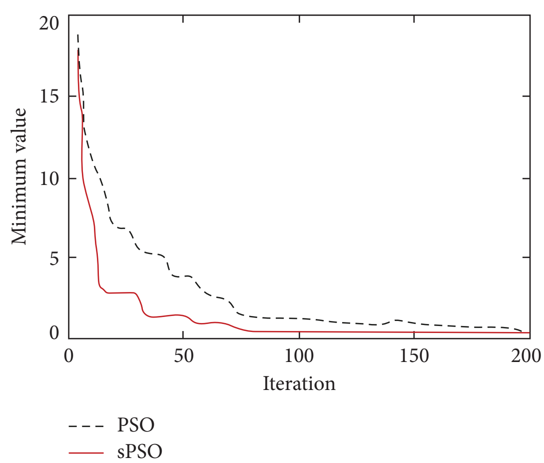


FIGURE 5: Graph of f_1 function optimization process.

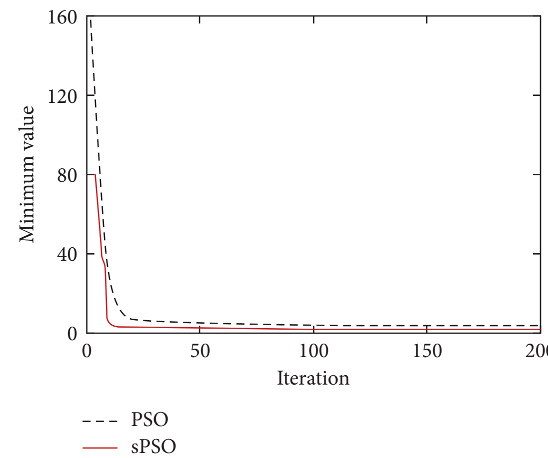


FIGURE 6: Graph of f_2 function optimization process.

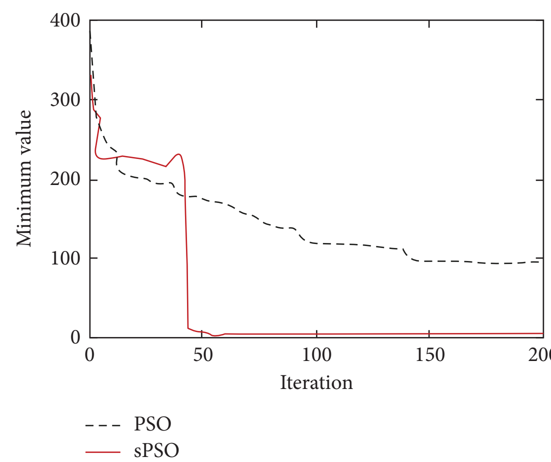


FIGURE 7: Graph of f_3 function optimization process.

To sum up, it can be seen from the optimization process curves of different functions that the improved particle swarm optimization algorithm (SPSO algorithm) after the improvement of the optimization algorithm in this paper is obviously due to the common PSO algorithm, which is characterized by faster convergence speed and higher accuracy, especially in the solution process of complex functions, the optimization ability is more prominent. In the actual combat process, it can also show better processing ability.

3.3. Construction of Restorative Indoor Environment Model Based on Neural Network. The standard BP neural network adopts the gradient descent method to constantly adjust the weight and threshold of the network to minimize the sum of squares of the network error. However, finding the minimum value in the error plane has its own limitations and shortcomings, mainly manifested in the fact that the learning rate is too small, resulting in too long training time, local extremum, slow convergence speed of the algorithm, poor numerical stability, etc. Therefore, this paper proposes a method to optimize the BP neural network model by using the improved particle swarm optimization algorithm, namely the SPSO BP neural network model. The flow chart of the improved neural network-based restorative indoor environment model algorithm is shown in Figure 8.

4. Simulation Experiment

4.1. Data Acquisition. After the model structure of comfort index SET* is established through the collection of experimental data, we should train according to the actual data. The experimental data were measured in a room on the second floor of the State Key Laboratory of Subtropical Building Science, South China University of technology. The room includes the measurement and sensing of four environmental factors: temperature and humidity, average radiation temperature, and wind speed, as well as comfort control equipment such as fans and air conditioners. Maintain ventilation during the experiment, other comfort control devices do not work.

A sht71 digital temperature and humidity sensor are installed in the room, it is used to measure the temperature

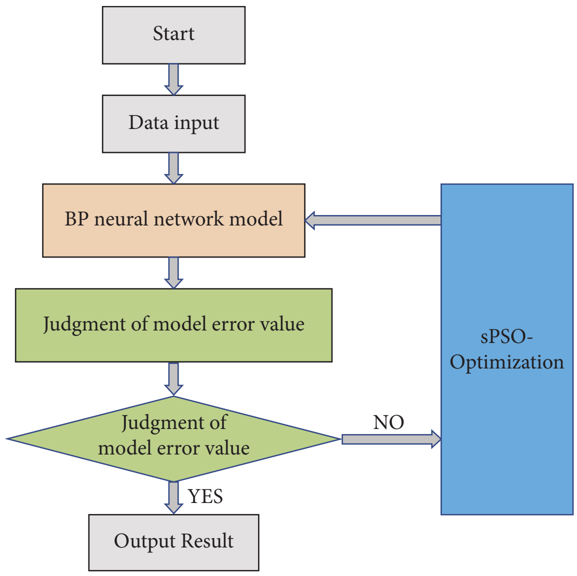


Figure 8: Model algorithm flow chart.

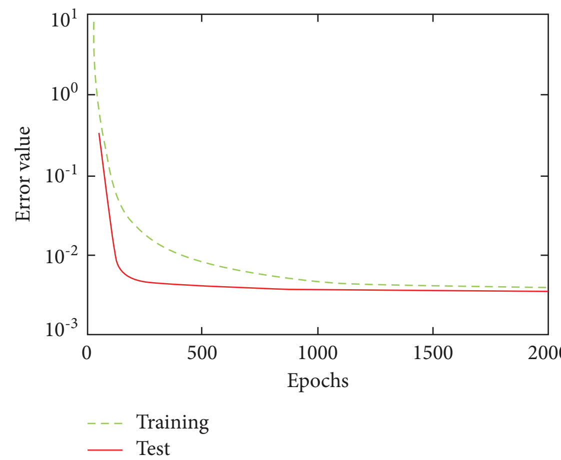


Figure 9: Iteration and convergence curve of standard BP neural network model.

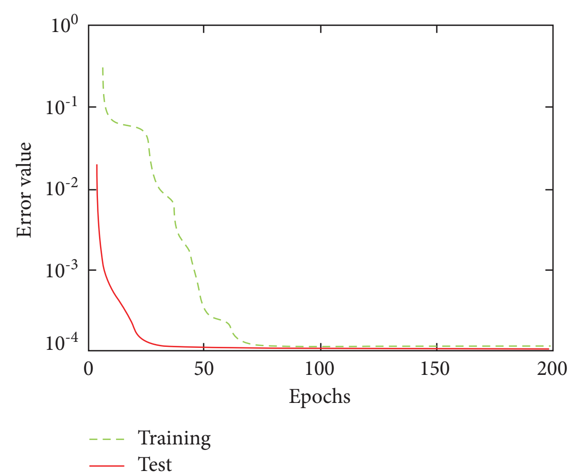


Figure 10: Iteration and convergence curve of BP neural network model optimized by improved particle swarm optimization algorithm.

and humidity of the indoor environment. The average radiation temperature was measured by Agilent 34970 A data acquisition instrument. Using SDP1000/SDP2000 micro differential pressure sensor to measure indoor wind speed, install the wind speed sensor in the center of the indoor roof. Since the thermal resistance of human clothing and the metabolic rate of the human body are variables related to human body, these two variables are uncontrollable factors in comfort control, for South China, only summer comfort is considered, for South China, only summer comfort is considered, when the thermal resistance ICOL of human clothing is 0.5 and the metabolic rate M is 58.2 W/m^2, the indoor comfort index SET* value is obtained by iterative calculation.

The experiment is carried out in the experimental room for measurement, and the experimental data collected by the sensor passes through the AD conversion module. After being converted into a digital signal, it communicates with PC through RS232, processes the received experimental data, and displays and stores it. From 1080 groups of environmental factors measured at different times collected by the sensor and SET* values obtained by iterative calculation, 880 groups of data are selected as training samples, and another 200 groups of data are selected as BP neural network model test sample data.

4.2. Model Training. Using the BP neural network model based on Improved Particle Swarm Optimization constructed above is used to train the experimental data, which is programmed with Python language of tensorflow2.0 software, and the network error target is set at 10^{-4}. Figures 9 and 10 show the iteration and convergence curves of the standard BP algorithm and the BP neural network algorithm model based on the improved particle swarm optimization.

Based on the Figures 9 and 10, it can be found that, first of all, from the iteration times of the above two figures, the standard BP neural network model has too long training time and local extremum due to too small learning rate, slow convergence speed of the algorithm, poor numerical stability, and is difficult to converge. It does not converge until the iteration times reach 2000. It can be seen that if the standard BP neural network model is used for modeling the restorative environment, it will spend a lot of time on training time; however, compared with the standard BP neural network model, the BP neural network model optimized based on the improved particle swarm optimization algorithm not only has less iterations but also has a high learning rate. The model can converge quickly. When the iterations are about 50 times, it will converge regionally, which can save a lot of time for model training. Secondly, based on the initial value of model error, the initial value of standard BP neural network model error is much larger than the initial value of error in the training of BP neural network model optimized based on improved particle swarm optimization algorithm, and its maximum value is close to 10, while the maximum initial value of error in the training of BP neural network model optimized based on improved particle swarm optimization algorithm is not more than $10^{0.5}$; At the same time, the final error difference between the two is also large, in which the final error of the standard BP neural network model is greater than 10^{-3}, while the final error of the BP neural network model optimized based on the improved particle swarm optimization algorithm is less than 10^{-4}.

To sum up, it can be found that BP neural network is used to establish the prediction model of indoor environmental comfort index SET*, and the relationship between environmental factors and SET* is obtained, which overcomes the complex iterative operation of thermal environment parameters on human thermal sensation in the traditional model. In addition, the BP neural network model optimized by the improved particle swarm optimization algorithm significantly improves the convergence speed of model training, The results show that there is little error between the prediction value of the improved neural network model and the expected value of the traditional model, which ensures the effectiveness of the prediction model, so it can meet the requirements of real-time in the control process of environmental control equipment such as air conditioning system.

5. Conclusion

With the improvement of material living standards, people have higher and higher requirements for the office and indoor rest environment, and the indoor temperature environment is very important to create a sense of comfort. A good indoor somatosensory temperature environment can not only convey the degree of somatosensory comfort but also relax people's bodies. At present, there is little research on this aspect at home and abroad, which leads to the stagnation of this theory. Therefore, in order to promote research in this field, this paper uses the neural network modeling method, based on the indoor comfort index SET * value and other factors that can affect the ambient temperature, aiming at the human body temperature comfort, adopts the improved particle swarm optimization algorithm to optimize the BP neural network method, constructs a neural network-based restorative indoor environment model, realizes the real-time control of SET* value, and then analyzes the relationship between indoor environmental thermal factors and SET* value, The model is optimized based on simulation experiments. The simulation results show that the error between the prediction value of the improved neural network model and the expected value of the traditional model is very small, which ensures the effectiveness of the prediction model. Therefore, it can meet the real-time requirements in the control process of environmental control equipment such as air conditioning systems. Relevant experiments not only verify the effectiveness of the improved algorithm in this paper but also use this algorithm model to optimize the indoor environment and ensure indoor comfort.

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] S. O. Ro, "A Study on the Physical, Mental and Social Factors Influencing the Health Status of Aged Women in Korea," *Korean Journal of Women Health Nursing*, vol. 2, no. 1, pp. 53–67, 1996.

[2] C. Caballero-Arce, A. Insausti, and J. B. Marco, "Lighting of Space Habitats: Influence of Color Temperature on a Crew's Physical and Mental health," in *Proceedings of the 42nd International Conference on Environmental Systems*, p. 3615, CA, USA, July 2012.

[3] R. Slabe-Erker and S. Ličen, "Dejavniki Gibalne Aktivnosti in Z Zdravjem Povezane Kakovosti Življenja (Factors Influencing Leisure-Time Physical Activity and Health-Related Quality of life)," vol. 48, no. 2, 2014.

[4] B. Oladunni, P. A. Lyoka, and D. T. Goon, "Perceived motivational factors influencing students with disabilities towards sports participation in Amathole district, Eastern Cape Province, South Africa," *African Journal for Physical Health Education Recreation & Dan*, vol. 21, no. 4.2, pp. 1389–1401, 2015.

[5] A. Ghahramani, "The Effect of the Relaxation Training on the General Health and Selected Physical Fitness Factors Affecting Seniors balance," vol. 43, no. 8, pp. 608–614, 2009.

[6] R. F. Baumeister, E. Bratslavsky, M. Muraven, and D. M. Tice, "Ego depletion: is the active self a limited resource?" *Journal of Personality and Social Psychology*, vol. 74, no. 5, pp. 1252–1265, 1998.

[7] L. W. Zhang, "Control of Ego-Depletion:Research and Application in Competitive Sports," *China Sport Science*, vol. 36, no. 5, pp. 506–515, 2013.

[8] R. F. Baumeister, K. D. Vohs, and D. M Tice, "The strength model of self-control," *Current Directions in Psychological Science*, vol. 16, no. 6, pp. 351–355, 2007.

[9] R. S. Ulrich, R. F. Simons, B. D. Losito, E. Fiorito, M. A. Miles, and M. Zelson, "Stress recovery during exposure to natural and urban environments," *Journal of Environmental Psychology*, vol. 11, 1991.

[10] K. M. Korpela, T. Hartig, F. G. Kaiser, and U Fuhrer, "Restorative experience and self-regulation in favorite places," *Environment and Behavior*, vol. 33, no. 4, pp. 572–589, 2001.

[11] P. Ouellette, R. Kaplan, and S. Kaplan, "The monastery as a restorative environment," *Journal of Environmental Psychology*, vol. 25, no. 2, pp. 175–188, 2005.

[12] R. Hansmann, S. M. Hug, and K. Seeland, "Restoration and stress relief through physical activities in forests and parks," *Urban Forestry and Urban Greening*, vol. 6, no. 4, pp. 213–225, 2007.

[13] J. Gutierrez, "Restorative Campus Landscapes: Fostering Education through Restoration," Kansas State University, Manhattan, 2013.

[14] M. Daly, "Association of ambient indoor temperature with body mass index in england," *Obesity*, vol. 22, no. 3, pp. 626–629, 2014.

[15] P. Bustelo, B. Cristina, G. Fernandez, M. Llorián, and C. Lovelle, "IoFClime: The Fuzzy Logic and the Internet of Things to Control Indoor Temperature Regarding the Outdoor Ambient conditions," *Future generations computer systems*, vol. 76, pp. 275–284, 2017.

[16] S. Vellingiri, P. Dutta, S. Singh, L. M. Sathish, S. Pingle, and B. Brahmbhatt, "Combating climate change-induced heat stress: assessing cool roofs and its impact on the indoor ambient temperature of the households in the urban slums of ahmedabad," *Indian Journal of Occupational and Environmental Medicine*, vol. 24, no. 1, 2020.

[17] D. G. Markov, P. S. Yordanov, I. S. Simova, M. Ivanov, N. Kehayova, and E. Georgiev, "On the Influence of Indoor Temperature on Occupant's Performance," in *Proceedings of the International Scientific Conference*, Ruse University, Bulgaria, October 2014.

[18] D. Meana-Llorián, C. González García, B. C. Pelayo G-Bustelo, J. M. Cueva Lovelle, and N Garcia-Fernandez, "IoFClime: the fuzzy logic and the Internet of Things to control indoor temperature regarding the outdoor ambient conditions," *Future Generation Computer Systems*, vol. 76, pp. 275–284, 2017.

[19] J. Gwak, M. Shino, K. Ueda, and M. Kamata, "An investigation of the effects of changes in the indoor ambient temperature on arousal level, thermal comfort, and physiological indices," *Applied Sciences*, vol. 9, no. 5, p. 899, 2019.

[20] B. Mabuya and M. Scholes, "The three little houses: a comparative study of indoor and ambient temperatures in three low-cost housing types in gauteng and mpumalanga, South Africa," *International Journal of Environmental Research and Public Health*, vol. 17, no. 10, p. 3524, 2020.

[21] A. Jo, H. Sang, A. Mjl et al., "Machine Learning-Based Discrimination of Indoor Pollutants Using an Oxide Gas Sensor Array: High Endurance against Ambient Humidity and Temperature," *Sensors and Actuators B*, Article ID 131894, 2022.

[22] L. Ma, N. Shao, J. Zhang, and T. Zhao, "The influence of doors and windows on the indoor temperature in rural house," *Procedia Engineering*, vol. 121, pp. 621–627, 2015.

[23] H. J. Kang, Y. P. Min, I. K. Lee et al., "Effects of Ambient Temperature and Dietary Glycerol Addition on Growth Performance, Blood Parameters and Immune Cell Populations of Korean Cattle steers," *Asian-Australasian Journal of Animal Sciences*, vol. 30, no. 4, 2017.

[24] M. K. Sharp, "Indoor comfort achieved exclusively from ambient sources across US climates," *Journal of Solar Energy Engineering*, vol. 143, no. 6, pp. 1–28, 2021.

[25] S. Yan and X. Li, "Analytical expression of indoor temperature distribution in generally ventilated room with arbitrary boundary conditions," *Energy and Buildings*, vol. 208, Article ID 109640, 2020.

[26] B. Sadeghi, "A BP-neural network predictor model for plastic injection molding process," *Journal of Materials Processing Technology*, vol. 103, no. 3, pp. 411–416, 2000.

[27] B. Lu and Y. Wang, "Overview of Handwritten Numeral Recognition Based on BP Neural network," in *Proceedings of the 2011 International Conference on Computer Science and Network Technology*, IEEE, Harbin, December 2011.

[28] M. Y. Chen and D. F. Chen, "Early cost estimation of strip-steel coiler using BP neural network Machine Learning and Cybernetics," in *Proceedings of the. 2002 International Conference on*, IEEE, Beijing China, November 2002.

[29] B. Zhao, C. X. Guo, B. R. Bai, and Y Cao, "An improved particle swarm optimization algorithm for unit commitment," *International Journal of Electrical Power & Energy Systems*, vol. 28, no. 7, pp. 482–490, 2006.

[30] G. C. Chen and Y. U. Jin-Shou, "Particle swarm optimization algorithm," *Information and Control*, vol. 186, no. 3, pp. 454–458, 2005.

[31] J. Kennedy, "Particle swarm optimization," *Proc. of 1995 IEEE Int. Conf. Neural Networks, (Perth, Australia), Nov. 27*, vol. 4, no. 8, pp. 1942–1948, 2011.

[32] C. J. Liao, C. T. Tseng, and P. Luarn, "A discrete version of particle swarm optimization for flowshop scheduling problems," *Computers & Operations Research*, vol. 34, no. 10, pp. 3099–3111, 2007.

[33] R. C. Eberhart, "Comparing Inertia Weights and Constriction Factors in Particle Swarm optimization," in *Proceedings of the 2000 IEEE Congress on Evolutionary Computation*, IEEE, La Jolla CA USA, July 2002.

[34] J. K. Zhang, S. Y. Liu, and X. Q. Zhang, "Improved Particle Swarm optimization," *Computer Engineering & Design*, vol. 193, no. 1, pp. 231–239, 2007.

[35] B. Liu, L. Wang, Y. H. Jin, F. Tang, and D. X Huang, "Improved particle swarm optimization combined with chaos," *Chaos, Solitons & Fractals*, vol. 25, no. 5, pp. 1261–1271, 2005.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 8188571, 11 pages
https://doi.org/10.1155/2022/8188571

Research Article

Analysis and Improvement of Tennis Motion Recognition Algorithm Based on Human Body Sensor Network

Neng Pan,[1] Yang Wang,[2] and Jun Tong [3]

[1]*School of Physical Education, Yunnan Normal University, Kunming 650500, Yunnan, China*
[2]*School of Rehabilitation, Kunming Medical University, Kunming 650500, Yunnan, China*
[3]*Dept. of Physical Education, Kunming Medical University, Kunming 650500, Yunnan, China*

Correspondence should be addressed to Jun Tong; tongjun@kmmu.edu.cn

Received 9 July 2022; Revised 6 August 2022; Accepted 17 August 2022; Published 14 September 2022

Academic Editor: R. Mo

With the development of digital image technology and human body sensor network, the error recognition model of tennis serving motion constructed by human beings based on this technology has also attracted more and more attention of the tennis lovers. Among the traditional methods, the tennis serve error recognition method mainly includes the tennis serve error recognition method based on feature resolution reconstruction, but it is difficult to realize the tennis serve error recognition. The traditional method of tennis serve action error recognition has poor feature recognition ability and detection and recognition ability. In this paper, a tennis serve error identification method based on human body sensor network (body area network) is proposed, and the simulation test analysis is carried out. The experimental results show that this method has superior performance in improving the ability of tennis serve error recognition. Relevant research can not only improve the accuracy of tennis movement error recognition, but also effectively promote the technical application and theoretical improvement of this technology in related fields.

1. Introduction

With the development of digital image technology, combined with image information processing and information recognition technology, a tennis serve action error recognition model is established. The image information fusion method is used to process the image information of tennis serve action, and a tennis serve action error feature analysis model is established [1, 2]. The tennis serve action error information is analyzed by combining the joint feature analysis and image edge contour detection methods. To improve the analysis ability of tennis movement characteristics, the research on relevant tennis serve movement error recognition methods has attracted great attention [3–6].

Relevant research needs to face the difficulties of technical diagnosis. The action of tennis is based on various motion systems, which is a mechanical system. Every human behavior has a basic mechanical structure. Finding out the key mechanical characteristics of each action is the key to evaluate the quality of the action. However, the motion system is composed of multiple human links. Whether it is to analyze the relatively simple gait or to optimize the complex tennis technology, it involves the movement and stability of multiple human links. The resulting large amount of biomechanical data needs to be comprehensively analyzed by experts from multiple disciplines. In view of the complexity of data processing, in the 1960s, foreign experts in the field of human body began to try to use computers to collect and process the biomechanical data of human gait, infer the data, and draw medical response suggestions for diseases in motion control, so as to assist the clinical doctors in surgery and treatment. At present, there have been several major breakthroughs in the construction of computer programs based on human body sensor networks, which are mainly divided into two different methods: expert system and neural network technology.

In recent years, with the change of technology, WBAN (wireless body area networks) has been widely concerned by academia and industry [7–9]. WBAN integrates many cutting-edge technologies such as wireless communication, sensor data acquisition and processing, intelligent wearing and human health, which will bring great changes and impacts on human production and life in the future. It can be widely used in medical [10, 11], physical rehabilitation [12–14], military [15, 16], sports [17–19], and other fields. Especially in sports competition, WBAN is a network composed of sensor nodes with different functions that are reasonably distributed on sports facilities according to their respective needs and characteristics. The position of the sensor can be located on the body, on the body surface, or in the body. It can carry out real-time and dynamic monitoring of various motion characteristic parameters of the sphere, and expand various applications based on this, so as to better serve the competitive life of human sports and escort the cause of human health [20–22].

Tennis, as a very popular sport mode, has strong antagonism and interest, and requires high technical requirements. Therefore, it requires professional guidance and training. However, this part of the cost exceeds the affordability of some people, and it also requires special appointment time, resulting in the decline of the feasibility of citizens engaging in this sport [23]. In order to solve this problem, many sports software have emerged in the market. People can watch teaching videos with this software, which is popular with many athletes and amateurs. However, this software can only mechanically display human movement behavior, poor interactivity, and cannot correct tennis errors in time, resulting in a decline in learning effect. So far, some progress has been made in the field of error action recognition. For example, linshuiqiang et al. [24] proposed the pose sequence finite state method to improve the universality and efficiency of users' recognition of natural behavior, so as to establish the limb node coordinate system. Taking the user as the starting point, when describing the number of limb nodes of users' behavior, it is no longer the device space, but the user space. Experiments show that this method can be well extended and has strong applicability. It can meet the needs of somatosensory interactive applications, but the recognition accuracy is very low; Shixiangbin et al. [25] adopted K-means clustering algorithm in human motion video sequence clustering, so as to improve the practicability and accuracy of motion recognition and reduce the complexity of the computer in the calculation process. By extracting the key frames of human motion video sequence from the clustering data, they have two features: the position of joint points in the key frames and the skeleton angle between various parts of the human body, The SVM classifier [26–28] is used to classify the action sequences. The experimental results show that the recognition accuracy is not very high.

Related research needs to face the difficulty of technical diagnosis. Human action is based on the musculoskeletal system, which is a mechanical system. Every human action has a basic mechanical structure. Finding out the key mechanical characteristics of each movement is the key to evaluate the quality of the movement. However, the musculoskeletal system is composed of multiple human links. Whether analyzing the relatively simple gait or optimizing the complex tennis technology, it involves the movement and stability of multiple human links. The resulting large amount of biomechanical data needs to be comprehensively analyzed by experts from multiple disciplines. In view of the complexity of data processing, in the 1960s, foreign experts in the field of human body began to try to use computers to collect and process biomechanical data of human gait, infer the data, and come up with medical response suggestions for diseases in motion control, so as to assist clinical doctors in surgery and treatment [29]. At present, there have been several major breakthroughs in the construction of computer programs based on artificial intelligence methods, which are mainly divided into two different methods: expert system and neural network technology. It is worth noting that various sports activities, including tennis, have become popular since the twentieth century. Although they have brought many benefits to people's physical and mental health, the injuries of the musculoskeletal system have also increased. This kind of injury is common not only among professional athletes, but also among the general public in their participation in sports leisure activities. Therefore, more and more attention has been paid to the movement quality of human movement. Among them, the more mature is the FMS movement screening system established by American physiotherapist gray cook [30]. At present, various intelligent products for physiological monitoring of professional athletes have been relatively mature, but artificial intelligence products for human movement quality monitoring have not appeared in China.

Based on the above research background, in view of the shortcomings of the above research, this paper combines the existing technology, fully analyzes the application prospect of the existing technology, applies the machine vision technology to the sports tennis wrong technical action recognition, and constructs the related wrong action recognition model, in order to improve the sports tennis wrong technical action recognition ability.

2. Image Acquisition and Feature Analysis of Tennis Serve

2.1. Image Acquisition of Tennis Serve. In order to realize the error recognition of tennis serve action based on image segmentation, it is necessary to first build the image acquisition model of tennis serve action. Combined with the feature analysis of tennis serve action video acquisition image, the output template feature matching model of tennis serve action image acquisition is established by using template matching and optimized feature detection methods, combined with image template feature matching and information fusion methods, Through the matching results of the action database, edge contour feature detection is carried out, and the joint image segmentation technology is used to match the position information and action information of the tennis serve action, so as to realize the error detection. The overall structure is shown in Figure 1.

According to the overall structure model shown in Figure 1, we can use the texture detection and fuzzy feature matching technology to detect the corners in the image acquisition process of tennis serve action. According to the

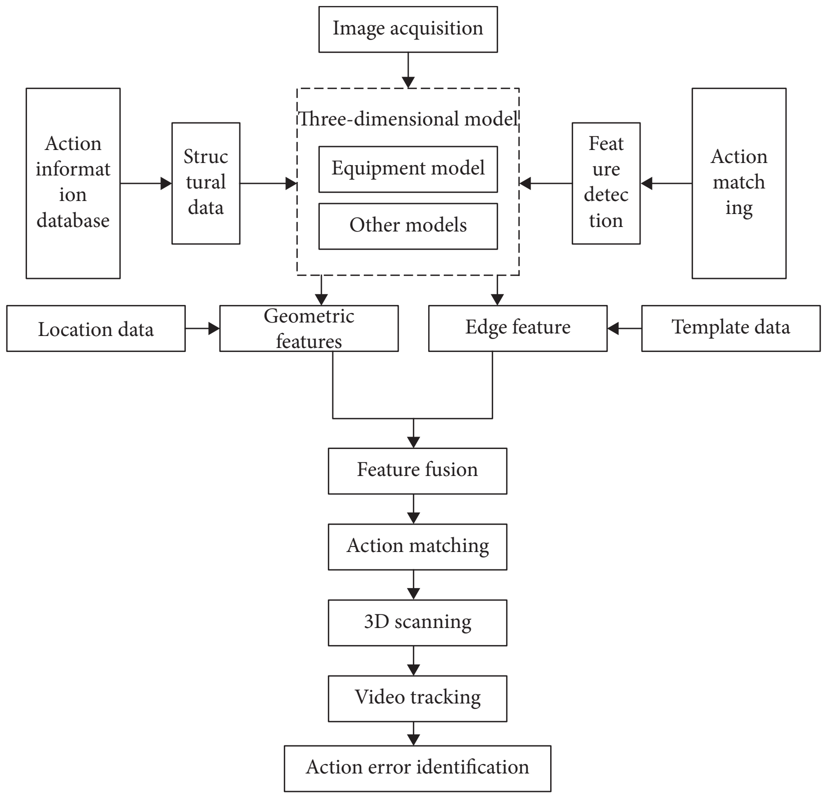


Figure 1: Overall structure model of error identification for tennis serve.

joint information feature point scanning method, we can realize the texture matching of the image of tennis serve action, and the image pixel distribution is as follows:

$$\overline{m}_i = \frac{1}{N} L(x, y, \sigma) \sum_{k=1}^{N} y_k I(x, y) \frac{s(n) m_i}{w}, \tag{1}$$

where: $I(x, y)$ represents the image-related parameters of the tennis serve action, $s(n)$ represents the image-related pixel value of the tennis serve action, and $L(x, y, \sigma)$ represents the feature recognition coefficient of the image of the tennis serve action.

Calculate the map prediction result of the image of tennis serve, which is expressed as:

$$\mathrm{L}(k, i) = \exp[G(x, y, \sigma)\varphi(k) \times \alpha \cdot D(i)], \tag{2}$$

where: $G(x, y, \sigma)$ is the non-significant feature of the image, and $\varphi(k)$ is the correlation factor. Using the correlation detection method, combined with the feature noise reduction technology of the tennis serve video capture image, the gray pixel value of the tennis serve video capture image is $E(d(x, y))$. Through template information matching and filter analysis, the tennis serve video feature analysis and corner information detection model are established, and the optimized collection results of the tennis serve video capture image are obtained.

2.2. Image Acquisition of Tennis Serve. Combined with the remote information recognition method, the feature detection model of tennis serve image is established. Combined with the scale transformation method, the feature collection of tennis serve image is carried out, and the sparsity fusion model of tennis serve image is obtained. Then, the fuzzy fusion control function of tennis serve video acquisition image is

$$F_d = \sum_{p=1}^{n} E_p A_i^2 \left[r_p(k) - y_p(k) \right], \tag{3}$$

where: A_i^2 represents the edge scale component of the tennis serve action image.

Using the edge scale feature segmentation method, the order mixed cumulant of the tennis serve action image is expressed as:

$$G_{new} = h(j) h^3 (j + \tau) \left(T^2 / \Delta^2 \right), \tag{4}$$

where: $h(j)$ is the feature set. The fourth-order cumulative mixed feature quantity of tennis serve action image is expressed as:

$$H_s = m f_s \sum_{k=1}^{n} (g_k - g_i), \tag{5}$$

where: f_s is the trend function and g_k is the directivity of image features. According to the contour information feature point acquisition results of the tennis serve video acquisition image, the boundary feature quantity of the tennis serve video acquisition image is reconstructed, which is expressed as follows:

$$R_i = 3A_i^4|\gamma|\sum_{i=0}^{s} h_i^4(j), \tag{6}$$

where: γ refers to the action frequency. Through the multimode state detection results, the multimodal high-frequency components and low-frequency components of the tennis serve video capture image are obtained. Combined with the multiscale detection results of the tennis serve video capture image, the edge scale information components of the tennis serve video capture image are

$$SNR_i = Kr + \sum_{i=0}^{s} |\gamma|A_i^4, \tag{7}$$

where: Kr is the image feature distribution point set. Carry out multimode feature calibration on the texture information feature distribution points of the tennis serve video capture image to obtain the noise kurtosis ratio of the tennis serve video capture image:

$$K(a,b) = \sum_{i=0}^{s} h_b^2(j)\sum_{V_m}^{m} K_r V_a, \tag{8}$$

where: V_a refers to the action track of tennis serve. The method of bilateral index random distributed detection is adopted to realize the feature extraction of tennis serve video capture image.

3. Recognition and Optimization of Tennis Serve Movement Error

3.1. Analysis on Error Characteristics of Tennis Serve. Combined with the scale transformation method, the multiscale feature segmentation of tennis serve video capture image is carried out [31], and the edge contour feature detection model of tennis serve video capture image is constructed. The average pixel set of tennis serve video capture is

$$\overline{x}_T = \frac{1}{T}\sum_{i=1}^{t} x_i\gamma K(a,b), \tag{9}$$

where: $x_i = (x_1, x_2, \cdots, x_T)$ is the error distribution template matching set of the tennis serve video capture image, and T is the pixel distribution density of the tennis serve video capture image.

The error information detection model of the video capture image of multi tennis serve is established, and the distributed pixel set of the error vector of the video capture image of tennis serve is obtained as follows:

$$F = \sum_{i=1}^{c} \mu_{ik} \times \eta J(\omega)\overline{x}_T. \tag{10}$$

Calculate the pixel intensity of the error frequency response modulus at $z = e^{\pm j\omega_0}$ of the tennis serve video capture image. According to the difference of pixel intensity, the dynamic image adaptive detection matrix of the tennis serve video capture is described as:

$$C = \frac{1}{n}\sum_{i=1}^{n} (X_i - \overline{X})\omega_{Br} I_x I_y. \tag{11}$$

According to the map feature matching, the tennis serve video capture image is collected and fused, and the pixel feature component of the tennis serve video capture image is obtained as follows:

$$E(T_n) = \frac{\tau \cdot \Phi(T_n)}{n}\sum_{i=0}^{N} CT_n\left(g_i^2\right), \tag{12}$$

where: $\Phi(T_n)$ represents the pixel value of the input tennis serve video capture image. According to the edge pixel distribution of the tennis serve video capture image, the edge information feature weighting function $f(g_i)$ of the tennis serve video capture image is obtained as follows:

$$f(g_i) = c_1 R_m(n)E(T_n)\sum_{i=0}^{n} \frac{\rho_j}{|v_i|}. \tag{13}$$

The error characteristic analysis model of tennis serve motion video capture image is obtained, and the error recognition is carried out according to the difference level of motion error [31].

3.2. Modelling of Individual Goals for Older People. Through the adaptive learning and scale transformation methods, the joint feature point location and fuzzy motion feature detection of the tennis serve video capture image are carried out. The segmentation expression of the tennis serve video image is recorded as:

$$G = W_i^T x_{ir} H_1 - f(g_i)kS^*(\omega)e^{-j\omega t_0}, \tag{14}$$

where: $H_1 = tr(x_{ir} - x'_{irp})A_{irp}$,

$$\begin{cases} GT^{\{3\}}(t) = T^{\{3\}}(t) - G^{\{3\}}(t) \\ AJ^{\{3\}}(t) = J^{\{3\}}(t) - A^{\{3\}}(t) \\ BK^{\{3\}}(t) = K^{\{3\}}(t) - B^{\{3\}}(t) \\ EP^{\{3\}}(t) = P^{\{3\}}(t) - E^{\{3\}}(t) \\ FQ^{\{3\}}(t) = Q^{\{3\}}(t) - F^{\{3\}}(t) \end{cases}$$

represents the statistical binary degree of the image, and tr represents the pixel tracking distribution track of the image.

Set the filter input as the power spectral density of $G_n(\omega) = N_0/2$, and combine the image filtering and optimized feature detection technology to obtain the dynamic distribution function of tennis serve video acquisition as follows:

$$HR(W_i) = Gs_d(t)^T \sum_{i=1}^{s} W_i^T x_d, \tag{15}$$

where: $s_d(t)$ is the $M \times d$ direction matrix; x_d is a $d \times 1$ image signal vector. According to the image segmentation results, the ambiguity boundary error function of the tennis serve video capture image is obtained as follows:

$$U(x,y) = HR(W_i)\sum_{c=1} V_c(x)p(y_s \mid x_s), \tag{16}$$

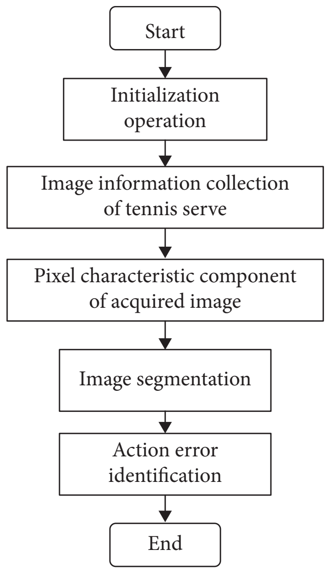


Figure 2: Implementation process of improved algorithm.

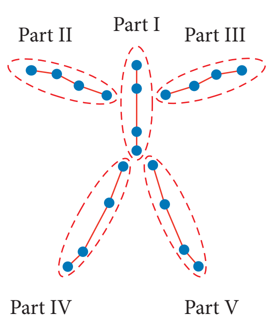


Figure 3: Diagram of division of five major parts of human body.

where: s represents the moment invariant feature of tennis serve video distribution, and $V_c(x)$ is the meta vector set of tennis serve video images. Thus, the action error identification output of tennis serve is

$$f_R(z) = U(x, y) \sum_{m=1}^{d} \lambda_m e_m * f(z), \tag{17}$$

where: $f(z)$ is the prior information of the tennis serve motion error, and $*$ is the convolution operation. Therefore, according to the image segmentation results, the tennis serve motion error recognition optimization is realized. The implementation process is shown in Figure 2.

4. Design of Recognition Method for Sports Tennis Wrong Technical Action

4.1. Extracting the Characteristics of Wrong Technical Movements in Sports Tennis. According to the understanding of the human body structure, the human body can be divided into five parts, namely, the trunk, left arm, right arm, left leg, and right leg. The trunk is an important part to support the human body. Some joints of the waist of the human body can reflect the information of their motion characteristics, while the motion information characteristics of the hands and feet are represented by the joints of the limbs of this part. The division results of the five major parts of the human body are shown in Figure 3.

Through the division of human body structure, the combination of these five parts can be used to represent some basic movements of the human body, so the hierarchical strategy is adopted in the classification. The first level: the actions of the five related combination modes should be summarized into a large category. For example, only two arm movements are the combination of the second part and the third part, which is the result of rough motion classification. Level 2: reclassify the actions of the same combination mode to determine the detailed classification of actions. The joint angle feature vector formed by the projection on the two-dimensional plane is verified from the 17 joint angles of the human body, and it is used as the first rough classification feature of human motion. According to the principle of kinematics, the features of the same combination of human body are extracted [8]. According to the principle of kinematics, the features of the same combination of human body are extracted. The complete movements of tennis players can be divided into active and auxiliary movements. The main action reflects the overall state of the motion mode, while the auxiliary action reflects the local state of the motion mode. Only by combining the characteristics of the main and auxiliary actions can we express this action more accurately. For the five major parts of the human body, such as the trunk, left arm, right arm, left leg, and right leg, the limb vectors in the three-dimensional space coordinates are established, respectively, which are expressed as the following formula:

$$\begin{cases} GT^{\{3\}}(t) = T^{\{3\}}(t) - G^{\{3\}}(t), \\ AJ^{\{3\}}(t) = J^{\{3\}}(t) - A^{\{3\}}(t), \\ BK^{\{3\}}(t) = K^{\{3\}}(t) - B^{\{3\}}(t), \\ EP^{\{3\}}(t) = P^{\{3\}}(t) - E^{\{3\}}(t), \\ FQ^{\{3\}}(t) = Q^{\{3\}}(t) - F^{\{3\}}(t), \end{cases} \tag{18}$$

where, {3} represents three-dimensional space; Indicates the moment when the limb moves; $GT^{\{3\}}$, $AJ^{\{3\}}$, $BK^{\{3\}}$, $EP^{\{3\}}$ and $FQ^{\{3\}}$, respectively, represent the limb vectors of the human body's trunk, left arm, right arm, left leg, and right leg in the three-dimensional space. In sports tennis, the contribution of human motion expression is different. Two joint angles are selected from each part as active joint angles. The following formula can be used to calculate the size of each joint angle of the human body in the three-dimensional space, so the angular velocity of the human joint angle is $\omega(t)$:

$$\omega(t) = \theta(t+1) - \theta(t). \tag{19}$$

The motion sequence of human body is continuous and changes with time. When the pitch angle changes, the angular velocity value will be generated. The limb vector and angular velocity of the active joint angle are the performance of the overall coordinated movement of the human trunk and limbs, and the change of the distance between joint points is represented by the bending of the human limbs and trunk. The human body also projects the YOZ side plane from the left view direction. The distance from the five parts of the tennis player to the joint point is

$$\begin{cases} d_{Gr}(y,z) = \sqrt{\left(y_G(t) - y_r(t)\right)^2 + \left(z_G(t) - z_r(t)\right)^2}, \\ d_{CL}(y,z) = \sqrt{\left(y_C(t) - y_L(t)\right)^2 + \left(z_C(t) - z_L(t)\right)^2}, \\ d_{CM}(y,z) = \sqrt{\left(y_C(t) - y_M(t)\right)^2 + \left(z_C(t) - z_M(t)\right)^2}, \\ d_{GR}(y,z) = \sqrt{\left(y_G(t) - y_R(t)\right)^2 + \left(z_G(t) - z_R(t)\right)^2}, \\ d_{GS}(y,z) = \sqrt{\left(y_G(t) - y_S(t)\right)^2 + \left(z_G(t) - z_S(t)\right)^2}, \end{cases} \tag{20}$$

where, $d(y,z)$ represents the Euclidean distance between two joints of the sports tennis player. In order to eliminate the differences between different individuals in sports tennis, each item in (18) and (20), as well as the width of the human shoulder and the mean value of the Euclidean distance between joints are standardized to obtain:

$$\begin{aligned} d_{AB}(x,y,z) &= \sqrt{\left(x_A - x_B\right)^2 + \left(y_A - y_B\right)^2 + \left(z_A - z_B\right)^2}, \\ \overline{d}(y,z) &= \frac{\sum_{t=1}^{n} d(t)}{n}, \end{aligned} \tag{21}$$

where: d_{AB} refers to the width of human shoulder; $\overline{d}$ represents the mean value of the Euclidean distance between joints in the five major parts of the human body.

According to the division results of the five parts of the human body, the limb vectors of the human body's trunk, left arm, right arm, left leg, and right leg in the three-dimensional space coordinates are established, and then the wrong technical action characteristics of sports tennis are extracted by using the distance between the five parts of the human body's bones and joints.

4.2. Motion Tracking and Adjustment Based on Body Area Network Technology. In the sports tennis teaching video or the live video of the game, the position of the camera is dynamic. Therefore, there are many non-static pictures in the sports tennis match video. Because the camera is always moving, it cannot well reflect the athlete's information. Therefore, it is necessary to use machine vision technology to track the moving target to obtain the athlete's activity field, and adjust the athlete's image within this range to offset the influence of external factors.

In the expanded tracking target area, symmetrical vertical and horizontal tracking is carried out to generate and adjust the image of the moving target. Calculate the "centroid" coordinate of the target area through (22), move the centroid coordinate (m_x, m_y) according to the central area, and then generate the tennis player target tracking image.

$$\begin{cases} m_x = \dfrac{\sum_{xeR}\sum_{yeR} x \cdot f(x,y)}{\sum_{xeR}\sum_{yeR} f(x,y)}, \\ \\ m_y = \dfrac{\sum_{xek}\sum_{yeR} y \cdot f(x,y)}{\sum_{xeR}\sum_{yeR} f(x,y)}. \end{cases} \tag{22}$$

After such processing, the target tracking image sequence can be adjusted. The adjusted image sequence only includes the athlete's limb motion and the motion generated by the racket, and cannot reflect the camera motion in the original image.

By tracking the sports tennis goal, the tennis player's activity area is obtained. By calculating the centroid coordinates of the sports tennis wrong technical action recognition area, the sports tennis wrong technical action tracking image is generated to realize the tracking and adjustment of sports tennis wrong technical action.

4.3. Design the Identification Process of Sports Tennis Wrong Technical Action. On the basis of sports tennis wrong technical action, the identification process of sports tennis wrong technical action is designed by calculating the sports tennis wrong technical action descriptor.

Machine vision technology is based on the characteristics of optical flow. The optical flow information set is established by the time displacement of each pixel, which requires high accuracy of optical flow. In order to transform the optical flow vector of moving video into a vector field, and then form a movable spatial distribution relationship, especially the optical flow field needs to be analyzed. Because the optical flow can only reflect the motion information in the tennis player's foreground image when tracking the image, the background in the tracking image will have an impact on the calculation of optical flow. So, we must clear up the background first. Not only the uniformity of background color should be considered, but also the global foreground image of tennis players should be obtained by processing the region growth algorithm based on the Gaussian mixture model. The athlete-centered background color elimination is shown in Figure 4.

According to the definition of the camera, the user can manually adjust the sports tennis error technical action image sequence, and then estimate the length of the light field. First of all, according to the change degree of flash and camera intensity, the brightness of the action image of sports tennis error technology is tracked in real time, which will lead to the error of optical flow calculation results. Therefore, image difference should be used to distinguish brightness and eliminate the influence caused by brightness change. Secondly, by analyzing the theory of biological vision system, it can be seen that machine vision cells are very sensitive to the edge movement of objects. In terms of direction and speed, different optical flows are formed due to different images to reflect the impact of human visual system on Logistics movement. Based on the difference image, the Horn–Schunck algorithm is used to estimate how tennis players track the Horn–Schunck optical flow field, as follows:

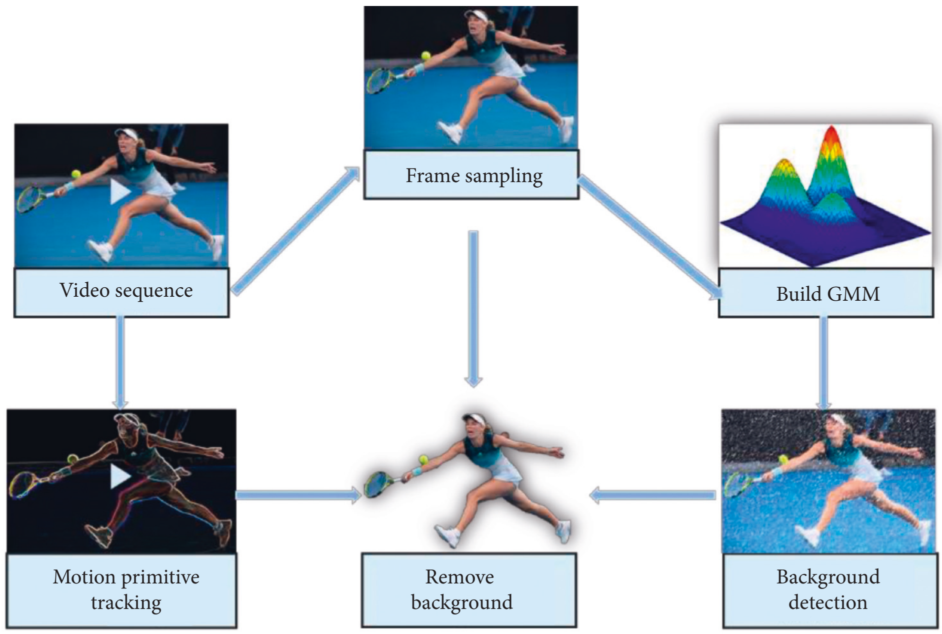


FIGURE 4: Athlete-centered background color elimination.

$$\begin{cases} DI_i = HC_i - HC_{i-1}, \\ OFF_i = HS(DI_i), i = 2, \cdots, N, \end{cases} \quad (23)$$

where: DI_i refers to the difference image of tracking sports tennis error technical action images HC_i and HC_{i-1}; HS refers to the estimation expression of horn Schunck algorithm; OFF_i represents optical flow field; N indicates the number of image sequence alignment of sports tennis wrong technical action.

There is a correlation between the athlete's position and the relative displacement of the body in the image of adjusting the wrong technical action of sports tennis, and this displacement exists in the corresponding image area. For different postures of tennis players, the spatial distribution of optical flow field is different.

According to the kernel density estimation and grid histogram of sports tennis error technical action image, the optical flow histogram is used as the motion descriptor of sports tennis players in the swing process. For the optical vector of a given optical flow field coordinate P, the components in the horizontal and vertical directions are $G_x(P)$ and $G_y(P)$, respectively. The amplitude $M(P)$ and direction angle $\theta(P)$ of sports tennis players' wrong technical action can be defined by using the following formula, which is expressed as follows:

$$\begin{cases} M(P) = \sqrt{G_x^2(P) + G_y^2(P),} \\ \\ \theta(P) = \arctan \dfrac{G_x(P)}{G_Y(P)}. \end{cases} \quad (24)$$

To sum up, on the basis of machine vision technology, the characteristics of sports tennis wrong technical action are extracted. By tracking and adjusting sports tennis wrong technical action, the sports tennis wrong technical action descriptor is calculated, and the recognition of sports tennis wrong technical action is realized.

5. Simulation Test Analysis

5.1. Error Technical Action Identification Test. In order to verify the performance of the sports tennis wrong technical action recognition method based on the human body area sensor network (body area network) technology, the athlete wrong technical action data collected in the tennis competitions of the 2016 Olympic Games, the 2017 China Masters, the Sudirman cup, and the China open were used. Sports tennis error technical action video is stored in MPEG-1 compression format, and the size of each video frame is 352*288. The types of tennis wrong technical actions in the above four matches are marked in the manual mode. The experimental data are shown in Table 1.

In order to quantitatively evaluate the recognition method of sports tennis wrong technical action based on machine vision technology, recall rate index and accuracy rate index are introduced to determine the recognition ability of each swing in sports tennis wrong technical action. The calculation method of recall rate index and accuracy rate index is as follows:

$$\begin{cases} R = \dfrac{n_c}{n_c + n_m} \times 100\%, \\ \\ p = \dfrac{n_c}{n_c + n_f} \times 100\%, \end{cases} \quad (25)$$

where: n_c refers to the number of sports tennis wrong technical actions correctly identified; n_m refers to the number of unrecognized sports tennis wrong technical

Table 1: Experimental data.

Sports tennis video clip	Left swing times	Up swing times	Right swing times
Tennis competition of 2016 olympic games	112	38	124
2017 China masters	706	234	291
Sudirman cup 2017	297	202	712
2017 China open	124	264	153
Total	1239	738	1280

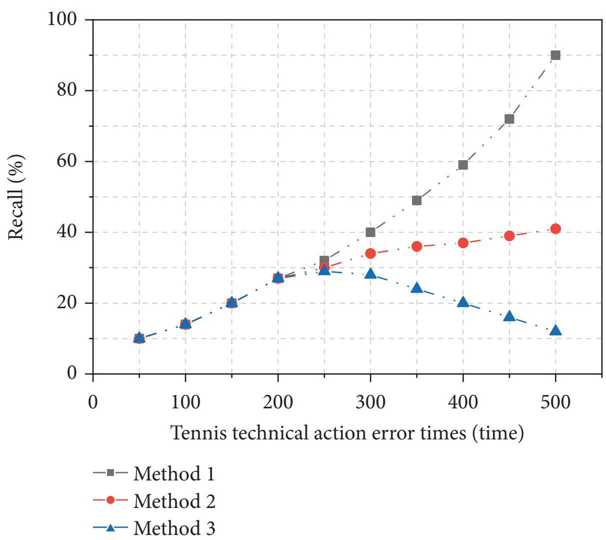


Figure 5: Recall test results under different methods.

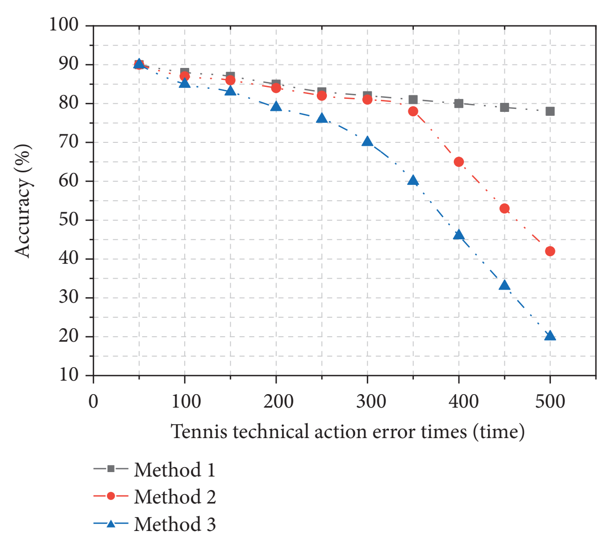


Figure 6: Accuracy test results under different methods.

actions; n_f indicates the number of sports tennis wrong technical actions incorrectly identified.

The recall rate test results of tennis error technical action recognition based on body area network are shown in Figure 5, and the accuracy rate test results are shown in Figure 6.

According to the experimental data in Table 1, three cross-validation strategies were used to form the training set and test set of the experiment. Taking the number of sports tennis wrong technical actions as the independent variable, the methods in literature [24] and literature [25] were introduced as comparison. After three iterations, the recall rates of the three action recognition methods were tested. The results are shown in Figure 5. It can be seen from the results in Figure 5 that when the number of wrong technical actions of sports tennis is less than 250, the action recognition recall rate of the three action recognition methods is basically the same. When the number of actions exceeds 250, when this method is used to identify sports tennis wrong technical actions, the recall rate of sports tennis wrong technical action recognition is getting higher and higher. According to the method of literature [24], due to the low quality of the identified sports tennis wrong technical action video, the recall rate appeared a turning point when the number of actions was 300 and began to decrease gradually. However, due to the influence of camera movement, the recall rate in reference [25] shows an upward trend, but the upward trend is relatively slow. Therefore, it can be concluded that the sports tennis wrong technical action recognition method based on machine vision technology has good performance in the recall rate test of sports tennis wrong technical action recognition.

It can be seen from the results in Figure 6 that with the increase of the number of sports tennis wrong technical actions, the methods in literature [24] and literature [25] cannot effectively distinguish when the movement direction of tennis players changes, resulting in a gradual decrease in the accuracy of sports tennis wrong technical action recognition. However, when this method is used to identify sports tennis wrong technical actions, the trend of the accuracy of identification is slow. When the number of sports tennis wrong technical actions reaches 500, the accuracy of action identification is as high as 77%.

5.2. Error Test of Service Action. In order to test the implementation of the method in this paper, set the number of pixels sampled for tennis serve video information is 250, the number of frames of the image is 120, and the image segmentation coefficient is 0 18. The regularization fusion parameter of image detection is 0 14. According to the above parameter settings, the tennis serve action error identification is carried out, and the original collected tennis serve action image acquisition results are shown in Figure 7.

Taking the image in Figure 7 as the research object, the tennis serve action error recognition is realized, and the image error recognition results are shown in Figure 8.

Figure 7: Image acquisition results of tennis serve.

Figure 8: Image processing and recognition of tennis serve.

Table 2: Error rate of tennis serve action recognition.

Iterations	Paper method	Method 1	Method 2
100	0.019	0.154	0.314
200	0.016	0.087	0.297
300	0.012	0.026	0.188
400	0.001	0.017	0.125

According to the analysis of Figure 8, the image resolution of tennis serve error recognition by this method is high, and the recognition ability is good. The accuracy of tennis serve error recognition by different methods is tested, and the comparison results are shown in Table 2.

According to Table 2, in the above three different tennis serve action recognition methods, the error rate shows a negative correlation trend with the increase of network iteration times, but for different methods, the error rate is quite different. When this method is used to identify the tennis service action error, its recognition accuracy is higher than that of literature [24] and literature [25], and the maximum error is close to zero. When using literature 24 and literature 25 to identify the tennis service action error, the error is much greater than that of the method used in this paper. This experiment verifies that the tennis movement recognition model based on human body area sensor network (body area network) has good recognition ability and certain use advantages.

6. Conclusion

Build a tennis serve error recognition model, and use the image information fusion method to process the image information of tennis serve. This paper proposes a tennis serve error recognition method based on image segmentation. Through the matching results of the action database, the edge contour feature detection is carried out, and the joint image segmentation technology is used to match the position information and action information of the tennis serve action. Through the adaptive learning and scale transformation methods, the joint feature point location and fuzzy action feature detection of the tennis serve action video capture image are carried out to realize the optimization of tennis serve action error recognition. The research shows that this method has better error and recognition performance in tennis serve error recognition.

In order to correct the wrong technical action of sports tennis in time, this paper proposes a tennis wrong technical action recognition method based on human body area sensor network (body area network). The human body area sensor network technology is used to extract the specific characteristics of sports tennis wrong technical action. Through tracking and adjusting Sports Tennis wrong technical action, the sports tennis wrong technical action recognition process is designed. The recognition of sports tennis wrong technical action is realized. The experimental results show that compared with the other methods, this method has better recognition performance, higher accuracy, the lowest error rate, and strong applicability. This proves that this research has high reliability and strong applicability, and the related technology and theory can also provide reliable theoretical support for the development of related research.

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.

References

[1] M. Ponakala, "Complexity Analysis of Wireless Body Area Networks Using Soft Computing Technique," *Information and Communication Technology for Competitive Strategies*, vol. 401, 2023.

[2] D. M. Perera, G. M. D. Menaka, W. Surasinghe, D. K. Madusanka, and T. D. Lalitharathne, "Development of a Vision Aided Automated Ball Retrieving Robot for Tennis Training Sessions," in *Proceedings of the 14th IEEE International Conference on Industrial and Information Systems*, IEEE Kandy, Sri Lanka, December 2019.

[3] S. S. Kim, E. H. Kim, and E. J. Kim, "A study of lower extremities joint moment pattern by stance types in tennis serve," *Korean Journal of Sport Biomechanics*, vol. 18, no. 2, pp. 41–48, 2008.

[4] Y. Yang and D. Hu, "Technical analysis of adaptive neuron fuzzy intelligent system in tennis serve," *Journal of Intelligent and Fuzzy Systems*, no. 1, pp. 1–10, 2021.

[5] W. Chiang, Y. T. Wang, and L. C. Liang, "The Relationships between Tennis Serve and Neuromuscular Control in Young Athletes," 2015.

[6] K.-W. Lee and K. Young, "The effect of repetition training type on improvement of skill movement of serve, stroke and volley in tennis-lovers," *Korean Journal of Sports Science*, vol. 21, no. 6, pp. 1173–1183, 2012.

[7] A. Hooda, S. Sharma, and S. Anima, "Wireless body area network[j]," *Journal of Electronics, Communication and Instrumentation Engineering Research*, vol. 6, no. 3, pp. 29–36, 2016.

[8] H. B. Li, K. Y. Yazdandoost, and B. Zhen, "Wireless body area network[J]," *Proc International Symposium on Medical Information & Communication Technology*, vol. 3, no. 3, pp. 203–210, 2010.

[9] E. Jovanov, A. Milenkovic, C. Otto, and P. C. de Groen, "A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation," *Journal of NeuroEngineering and Rehabilitation*, vol. 2, pp. 1–6, 2005.

[10] J. J. Jijesh and H. Shivashankar, "Implementation of health monitoring in sensor platform for wireless body area network," in *Proceedings of the 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, IEEE, Bangalore, India, May 2018.

[11] K. P. Wagh and S. O. Rajankar, "Wireless Body Area Network : Real-Time Health Monitoring System," in *Proceedings of the IJCA Proceedings on National Conference on Advancement in Electronics and Telecommunication Engineering*, Pune, 2012.

[12] K. Karunanithy and B. Velusamy, "Edge device based efficient data collection in smart health monitoring system using wireless body area network," *Biomedical Signal Processing and Control*, vol. 72, Article ID 103280, 2022.

[13] J. Ahmad and F. Zafar, "Review of Body Area Network Technology & Wireless Medical Monitoring," *International Journal of Information and Communication Technology*, vol. 2, no. 2, 2012.

[14] F. Touati, O. Erdene-Ochir, W. Mehmood et al., "An experimental performance evaluation and compatibility study of the bluetooth low energy based platform for ECG monitoring in WBANs," *International Journal of Distributed Sensor Networks*, vol. 2015, no. -9-2, p. 191, 2015.

[15] E. E. Egbogah and A. O. Fapojuwo, "Achieving Energy Efficient Transmission in Wireless Body Area Networks for the Physiological Monitoring of Military Soldiers," in *Proceedings of the Military Communications Conference*, IEEE San Diego, CA, USA, November 2014.

[16] A. M. Mile, "Modeling on body delay tolerant network sink locality of wireless body area networks for different body postures," *Global Journal of Computer Science and Technology*, vol. 20, no. 1, 2020.

[17] F. Ullah, M. Z. Khan, M. Faisal, H. U. Rehman, S. Abbas, and F. S. Mubarek, "An energy efficient and reliable routing scheme to enhance the stability period in wireless body area networks - ScienceDirect," *Computer Communications*, vol. 165, pp. 20–32, 2021.

[18] S. Raghuraman, K. Venkatraman, Z. Wang et al., "Immersive Multiplayer Tennis with Microsoft Kinect and Body Sensor networks," in *Proceedings of the Acm International Conference on Multimedia*, ACM, Nara, Japan, October 2012.

[19] F. D. Franco, C. Tachtatzis, B. Graham et al., "The Effect of Body Shape and Gender on Wireless Body Area Network On-Body channels," in *Proceedings of the Antennas & Propagation*, IEEE, Cairo, Egypt, October 2010.
[20] C. Otto and A. Milenkovic, "System architecture of a wireless body area sensor network for ubiquitous health monitoring," *Journal of Mobile Multimedia*, vol. 1, no. 4, pp. 307–326, 2006.
[21] C. Poon, Y. T. Zhang, and S. D. Bao, "A novel biometrics method to secure wireless body area sensor networks for telemedicine and m-health," *IEEE Communications Magazine*, vol. 44, no. 4, pp. 73–81, 2006.
[22] C. C. Y. Poon, Y. T. Zhan, and S. D. Bao, "A Novel Biometrics Method to Secure Wireless Body Area Sensor Networks for Telemedicine and M-Health," *IEEE Communications Magazine*, vol. 44, 2006.
[23] X'. Hongxiang and Dongping T., "Analysis of factors affecting the development of tennis fitness consumption of urban residents in Guangzhou," *Journal of Guangzhou Institute of Physical Education*, no. 06, pp. 54–57, 2007.
[24] S. Lin, Y. Wu, and Y. Wu, "Posture sequence finite-state machine method for motion recognition," *Journal of Computer-Aided Design & Computer Graphics*, vol. 26, no. 9, pp. 1403–1411, 2014.
[25] liushuanpeng Shixiangbin, "Zhang Deyuan Human motion recognition method based on key frame," *Journal of System Simulation*, vol. 27, no. 10, pp. 2401–2408, 2015.
[26] H. Meng, N. Pears, and C. Bailey, "Human action classification using SVM_2K classifier on motion features," in *Proceedings of the multimedia content representation, classification and security, international workshop*, Istanbul, Turkey, September 2006.
[27] K. C. Lin, S. H. Hsu, and J. C. Hung, "Adaptive SVM-Based Classification Systems Based on the Improved Endocrine-Based PSO Algorithm," in *Proceedings of the International Conference on Active Media Technology*, Springer-Verlag, Berlin, Heidelberg, 2012.
[28] K. M. Chit and K. Z. Lin, "Audio-Based Action Scene Classification Using HMM-SVM Algorithm," *International journal of advanced research in computer engineering & technology*, vol. 2, 2013.
[29] J. Cancela, M. Pastorino, A. Tzallas et al., "Wearability assessment of a wearable system for Parkinson's disease remote monitoring based on a body area network of sensors," *Sensors*, vol. 14, no. 9, pp. 17235–17255, 2014.
[30] F. Philp, D. Blana, C. Stewart, E. Chadwick, and A. Pandyan, "Does the Total Functional Movement Screen (FMS) Score Have Predictive Validity in Football," 2018.
[31] H. Tang, "Detection algorithm of tennis serve mistakes based on feature point trajectory," *Advances in Meteorology*, vol. 2022, Article ID 6584827, 8 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 6515974, 10 pages
https://doi.org/10.1155/2022/6515974

Research Article

Analysis on the Classification and Evaluation System of Talents in Colleges and Universities from the Perspective of AHP

Dong Wei,[1] Ling Guo ,[2] and WenRan Dong [3]

[1]*Office of Human Resources, ShanDong Jianzhu University, Jinan 250101, Shandong, China*
[2]*School of Art, ShanDong Jianzhu University, Jinan 250101, Shandong, China*
[3]*Personnel Department, Shandong University of Finance and Economics, Jinan 250014, Shandong, China*

Correspondence should be addressed to WenRan Dong; rcc_wd@sdjzu.edu.cn

Received 15 July 2022; Revised 12 August 2022; Accepted 23 August 2022; Published 9 September 2022

Academic Editor: R. Mo

With the continuous advancement of higher education reform, talent classification evaluation has become a new trend in the management and development of teachers in colleges and universities. As an important foundation for the development of teachers in colleges and universities, talent evaluation not only plays the role of the baton but also promotes the process of educational reform in colleges and universities. As the difficulty and key point of the reform of the personnel system in colleges and universities, the implementation of the classification and evaluation of talents in colleges and universities is conducive to mobilizing their enthusiasm and creativity and is an inevitable trend in the reform and development of higher education. However, at present, teaching, scientific research, and social services are mainly used as the important basis for teacher evaluation in college talent evaluation. The evaluation mechanism is not perfect, the evaluation standard is single, the evaluation methods are convergent, and evaluation mechanism of the academic circles on evaluation, etc. The research on the problem is still very weak, and the development of college teachers is increasingly affected. It is urgent to build a set of accurate classification and evaluation systems for college talents. In view of the many problems existing in the above-mentioned talent classification and evaluation mechanism in colleges and universities, this paper adopts AHP and fuzzy evaluation method, adopts this method to evaluate talents in the stage of talent introduction in colleges and universities, classifies talents according to the evaluation results, and finally uses the classification results as it is tested in practice to verify the effectiveness of the method constructed in this paper.

1. Introduction

With the rapid development of economy, the social demand for talents is also increasing, and the quality of higher education has become the focus of social attention. In order to better improve the teaching level of college teachers, we must strengthen the research of classification evaluation systems [1–3]. Because the traditional teaching quality research adopts qualitative analysis, the results are highly one-sided, which cannot accurately and comprehensively reflect the objective reality, and due to the limitations of technology and other reasons, the evaluation system is too cumbersome, the process is too complex, and the results lag is serious, which is not very helpful for the improvement of teaching quality [4–7]. The evaluation of teaching quality is mainly aimed at the comprehensive evaluation of teaching ability, teaching process, and teaching effect, which not only provides a certain data reference for teachers to improve the teaching content but also provides more basis for formulating a scientific and reasonable classification evaluation system [8, 9].

At present, China still has problems in talent evaluation methods and mechanism construction, such as weak classification evaluation management, convergence of evaluation standards for all kinds of talents, backward evaluation methods, and low integration of evaluation indicators with social needs. The existence of these problems is crucial to further deepen the reform of talent evaluation mechanisms [10–12]. As a large country of human resources, although China has unique advantages in the development and training of human resources, colleges and universities are gathered, scientific research institutes are numerous, and the industrial economy is booming, there are still various drawbacks in the evaluation of all kinds of talents; In terms of talent resource evaluation and use, the concepts and methods are not advanced enough, and there are many areas that need to be improved [13–16]. In the new era, China's demand for independent innovation is more urgent than ever before. Therefore, more high-level, high-tech and interdisciplinary talents are needed to help the process of social development. However, under the influence of the contradiction between the continuous progress of demand and the stagnation of technology, China's talent evaluation system is always in the development stage. It can be seen that a series of work such as talent training, evaluation, and management in China are facing great challenges, and it is also a key stage of comprehensive reform [17, 18]. Therefore, taking talent evaluation as the starting point and connecting point of reform, research, and explore the corresponding talent evaluation index system and evaluation methods and corresponding ways to improve, so as to improve the employment environment, enhance the attraction to talents, and promote the best use of talents, which is also one of the basic work in talent management [19].

In order to solve the problems of incomplete talent evaluation system and improper talent classification in our country, based on previous studies, this paper adopts AHP and fuzzy evaluation method to evaluate in the talent introduction stage of colleges and universities, and classifies talents according to the evaluation results. Finally, the classification results are used for practical testing, so as to promote the improvement of talent evaluation system in our country.

2. Problems in the Classification and Evaluation of Talents in Colleges and Universities at the Present Stage

2.1. Problems in the Evaluation Subject

(1) Alienation and dislocation of evaluation subject. Talent evaluation should be oriented by innovation, quality, contribution, and other indicators. However, in the actual implementation process, administrative power is often greater than academic power in China. The evaluation process is controlled by the leadership intention of the evaluation subject unit, and the correct and scientific evaluation method is alienated into a subjective evaluation method involving complex "relationships" [14, 20].

(2) The professional quality of the evaluation subject is low. Compared with foreign professional evaluators who have received systematic training, although China also has some excellent professional evaluators of talents in different industries, the number of such professionals is insufficient, and the quality level is uneven, which seriously restricts the development and progress of talent evaluation in China.

2.2. Problems in the Evaluation Object

(1) The definition of scientific and technological talents is unclear. At present, academia, industry, and government departments have different understandings of the definition of talents, and lack an accurate definition of the deep-seated connotation and characteristics of talents. Ji x[12] pointed out that the definition of talent and related concepts in China is vague, and there is no unified and scientific definition. Hua et al. [14] pointed out that although many domestic scholars have made various interpretations of the connotation of innovative talents, they have not reached a consensus, which makes the definition of talents a form, which is not conducive to the accurate evaluation of all kinds of talents.

(2) Insufficient talent classification and evaluation: at present, the domestic talent evaluation system is general, lacking the classification mechanism and corresponding evaluation indicators for various types of talents. Wang et al. [21] pointed out that the lack of classified evaluation is one of the difficulties faced by the evaluation of scientific research talents in China. Wang [22] pointed out that the classification and evaluation of talents is the requirement of talent science and technology concept, but at present, the classification standards of talents are different, which leads to the evaluation results cannot truly reflect the value and contribution of talents, and the guiding role is not obvious.

2.3. Problems in Evaluation Standards

(1) The pertinence of talent evaluation standards is poor. At present, the evaluation criteria of talents in China are often determined by the evaluation subject, and will change with the change of evaluation objectives. Therefore, the existing evaluation criteria are not reasonable, not only lack scientificity but also have strong fuzziness. Cui [23] pointed out that the existing talent evaluation standards do not emphasize systematic innovation enough, which is not conducive to promoting talent output and transformation. Powell et al. [18] evaluated high-level scientific and technological talents from multiple comprehensive standards such as talent quality level, academic performance, and influence, avoiding the untrue evaluation results caused by a single evaluation standard.

(2) The evaluation index is lack of operability. At present, the selection of talent evaluation indicators in China has not overcome the shortcomings of traditional evaluation indicators, and still focuses on "morality, ability, diligence, performance" and other aspects. On the one hand, static indicators lack adjustability, and cannot well meet the needs of actual evaluation; On the other hand, the selection of evaluation indicators is divorced from the principle of seeking truth from facts, resulting in the

measurement of some indicators cannot be operated, and the subjectivity is large.

2.4. Problems in Evaluation Methods

(1) There are few original theories, which copy the theoretical framework of foreign countries. The weakness of theoretical construction makes the scientific foundation of talent evaluation in our country insufficient, especially the talent evaluation theory with Chinese characteristics that is suitable for the actual situation of our country has not been fully formed. If we copy the theoretical evaluation method that is not suitable for the development of colleges and universities in our country, it will cause the risk of "going astray" to the domestic talent evaluation work.

(2) The evaluation method is single and subjective. At present, China mainly adopts the weighted talent evaluation method, evaluates talents through the evaluation index system formulated in advance, and obtains the total score of talents according to the weight of each evaluation index. Although this method pays attention to the comprehensive evaluation of scientific and technological talents, it also ignores the correlation between various indicators, and does not consider the systematic effect of the talent subject formed by various indicators in the talent evaluation process, so that the evaluation results cannot accurately reflect the quality, ability, and skills of talents.

(3) The evaluation system is unreasonable, such as a single evaluation standard, no differential treatment of scientific and technological personnel at different levels and scientific research posts of different nature, the complexity and variability of teachers' work are not effectively reflected, the phenomenon of seniority and thesis orientation is serious in the evaluation process, and the phenomenon of ignoring or avoiding qualitative indicators such as the morality and scientific research quality of the evaluated personnel occurs from time to time.

(4) It is easy for some teachers to ignore the research of teaching art and the improvement of teaching quality.

(5) Emphasize quantity and neglecting quality, that is, focusing on quantity in the evaluation of scientific research achievements such as papers, works, fund projects, and patents. This evaluation without quality as the core is easy to lead to the emergence of a large number of low-level achievements.

(6) The classification of scientific researchers is general, and even the differences among the most basic scientific research teachers, teaching teachers, and teaching and scientific research teachers are not reflected, and the differences between disciplines and majors are impossible to talk about.

2.5. Problems in Evaluation Procedures

(1) The evaluation process lacks impartiality. First of all, government departments and organizations will have an impact on talent evaluation, which destroys the objectivity and impartiality of the talent evaluation process. Secondly, the evaluation process is easily affected by economic factors, such as bribery, which not only affects the authenticity of talent evaluation but also leads to a bad atmosphere in the process of talent evaluation. Finally, talent evaluation is vulnerable to interpersonal relationships in the process, and "colleague" and "leadership" evaluation procedures are inevitable, making the evaluation results unable to truly reflect the actual situation of talents.

(2) The evaluation process lacks effective supervision. The unreasonable talent evaluation process directly leads to the distortion of the evaluation results, which is far from the actual situation. However, at present, there is no effective supervision mechanism or relevant regulatory agencies in China, and the function of external supervision is often a mere formality, which leads to the fact that the supervision procedure does not improve the objectivity and impartiality of the talent evaluation process.

2.6. Problems in Evaluation Results

(1) The application of evaluation results is insufficient. The application field of domestic talent evaluation results is relatively narrow, and the value of evaluation results cannot be effectively transformed into market value. This drawback is mainly manifested in two cases: one is formalism, which regards the evaluation process as a process or procedure, and the subsequent evaluation results are shelved; The other is to overemphasize the role of talent evaluation, resulting in the resistance of the evaluated to talent evaluation.

(2) The evaluation results are lacking follow-up evaluation and supervision feedback. From the actual operation of the domestic innovative talent selection project, it can be seen that the government departments strictly control the whole process of the selection of various talent projects according to the system, which means that excellent individuals or entrepreneurial teams will be able to obtain high social funds, strong support, and good social reputation, but there are obvious deficiencies in the follow-up implementation of performance appraisal and market supervision of the construction of various talent selection projects.

3. Construction of Talent Classification and Evaluation System in Colleges and Universities from the Perspective of AHP

3.1. AHP Theory. AHP is a qualitative and quantitative evaluation and analysis method proposed by Professor satti, an American operational research scientist. Its principle is relatively simple, that is, the evaluator decomposes the problem into several levels and elements by analyzing the governing relationship, internal essence or influence factors of a complex problem, and compares, judges, and calculates among the elements at the same level, so as to obtain different ranking schemes and provide decision-making basis for selecting the optimal scheme. The advantage of this analysis method is that it makes use of less quantitative information and mathematicizes the decision-making thinking process, which has high reliability and small error. The talent classification using this method mainly includes four steps as shown in Figure 1.

3.1.1. Establish Hierarchical Structure Model. When using analytic hierarchy process to make decisions, we should first clarify the purpose, scope, and requirements of the evaluation object of the decision-making goal, as well as the original information we have, and analyze the relationship between various elements; Secondly, the system problem is divided into levels and layers, and a hierarchical structure model is constructed. In this model, the problem is decomposed into several elements, and these elements are grouped according to attributes to form a mutually disjoint hierarchy. In the hierarchy, the upper layer dominates the elements of the lower layer, forming a layer by layer domination relationship from top to bottom according to the hierarchy:

(1) The highest level: the target level, which usually has only one element, is the desired result of the decision or the goal to be achieved;
(2) Intermediate layer: in order to achieve the intermediate link covered by the goal, it can be composed of several levels such as criteria and sub criteria, so the intermediate layer is also called the criteria layer;
(3) The bottom layer: also known as the scheme layer, mainly includes the solutions and measures to be taken to achieve each goal.

Generally speaking, the number of levels of a hierarchical structure model is related to the complexity of system problems and the level of detail that needs to be analyzed. In addition, in order to avoid the difficulty of comparing and judging because of too many dominant elements, the next level of elements dominated by each element in each level is required to be no more than nine.

3.1.2. Build the Judgment Matrix of Pairwise Comparison. With the establishment of hierarchical structure, the membership relationship between the upper and lower elements is also basically determined. Next, we need to build a pairwise comparison judgment matrix. That is, the elements at the upper level are the evaluation criteria, and the elements at this level are compared and weighted according to the relative importance of the evaluation criteria. The form of judgment matrix is shown in Table 1.

In order to quantify and compare the results, the 9-level scaling method is introduced, that is, the natural number 1–9 and its reciprocal are used to compare the relative importance of the two elements, so the evaluator's thinking decision process can be quantified more accurately. The 9-level scaling method and its significance are shown in Table 2.

3.1.3. Hierarchical Single Sorting and Consistency Verification. The maximum eigenvalue and its corresponding eigenvector W are calculated for the judgment matrix A, so the weight vector corresponding to the evaluation element is obtained.

$$AW = \lambda_{\max} W. \tag{1}$$

The steps of using the square root method to calculate $\lambda_{\max}$ and W are as follows.

First, multiply the weight value obtained when each element in the judgment matrix A is compared with other elements, and open the final result to the power of n to obtain a new set of vectors:

$$\overline{W} = \left\{ \prod_{j=1}^{n} \alpha i j \right\}^{1/n}. \tag{2}$$

Then, normalize the new vector to obtain the feature vector:

$$W_i = \frac{\overline{Wi}}{\sum_{i=1}^{n} \overline{Wi}}. \tag{3}$$

Calculate the maximum eigenvalue $\lambda_{\max}$ of the judgment matrix:

$$\lambda_{\max} = \sum_{i=1}^{n} \frac{(\mathrm{AW})\mathrm{i}}{nWi}. \tag{4}$$

Because the judgment matrix is obtained by comparing the elements in pairs, the weights of the elements are estimated. If the deviation of valuation is too large, and there is a serious inconsistency of thinking and judgment, the matrix must be corrected. The purpose of consistency verification is to verify whether such a deviation exists in the matrix. The matrix is valid only if it passes the consistency verification. The steps are as follows:

(1) Calculate the consistency index CI:

$$CI = \frac{(\lambda_{\max} - n)}{(n-1)}, \tag{5}$$

where $CI = (\lambda_{\max} - n)/(n-1) = (4.12 - 4)/(4-1) = 0.04 < 1$, according to Table 3, since the dimension of matrix A is $n = 4$ and $RI = 0.9$.

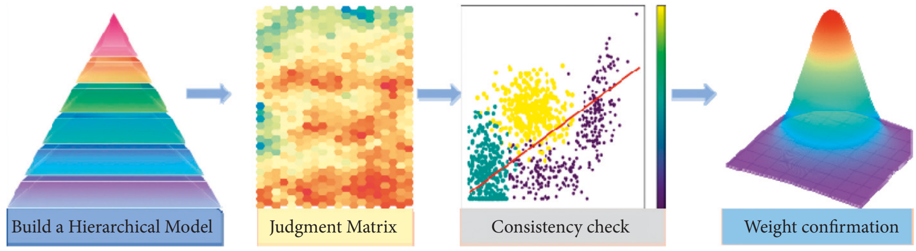


Figure 1: AHP talent decision process.

Table 1: Basic form of judgment matrix.

A	A_1	A_2	…	A_{n-1}	A_n
A_1	$a_{1,1}$	$a_{1,2}$	…	$a_{1,n-1}$	$a_{1,n}$
A_2	$a_{2,1}$	$a_{2,2}$	…	$a_{2,n-1}$	$a_{2,n}$
⋮	⋮	⋮	…	⋮	⋮
A_{n-1}	$a_{n-1,1}$	$a_{n-1,2}$	…	$a_{n-1,n-1}$	$a_{n-1,n}$
A_n	$a_{n,1}$	$a_{n,2}$	…	$a_{n,n-1}$	$a_{n-1,n}$

Table 2: AHP 9 scale method and its significance.

Scaling	Meaning
1	Elements Ai and A_j are compared, both have equal importance
3	Elements Ai and A_j are compared, A_i is slightly more important than A_j
5	Elements Ai and A_j are compared, A_i is obviously more important than A_j
7	Elements Ai and A_j are compared, A_i is more strongly important than A_j
9	Elements Ai and A_j are compared, A_i is more extremely important than A_j
2, 4, 6, 8	Represents the median value of the above adjacent judgments

Table 3: Average random consistency test index RI.

Order	1	2	3	4	5	6	7	8	9	10
RI	0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49

Table 3 shows the RI values of 1–10 dimensional matrices.

(2) Calculate the random consistency ratio CR:

$$CR = \frac{CI}{RI}. \tag{6}$$

The average random consistency index RI is obtained based on the arithmetic mean of the eigenvalues of the random judgment matrix after repeated calculations. When $CR < 0.1$, this judgment matrix is determined to meet the consistency requirements.

According to the above, $CR = CI/RI = 0.04/0.9 = 0.044 < 0.1$, the CR of each judgment matrix is less than 0.1, so the consistency test is passed.

Similarly, according to the above steps, calculate the weight of layer C indicators relative to layer B, and the weight of layer D indicators relative to layer C in turn, and conduct consistency test. After calculation, all indicators also pass the consistency test.

3.1.4. Combination Weight Calculation and Total Hierarchy Sorting

(1) In order to choose a better place, after calculating the weight vector of a group of elements relative to a certain criterion element of the upper level, we also need to synthesize the weights of each single level element to obtain the combined weight of the lowest level relative to the highest level target.

If the lowest layer is set as K layer, the weight of its index I relative to the index of the previous layer is $w^{(k)}$, and the weight of the index of this layer relative to the index of the next higher layer is $w^{(k-1)}$, and so on, then the weight $w_i^{(k)}$ of the index of the lower layer relative to the highest layer can be obtained:

$$w_i^{(k)} = w^{(k)} w^{(k-1)} w^{(k-2)} \ldots w^{(2)}. \tag{7}$$

(2) Consistency verification shall be carried out after the total ranking of the levels. When $CR < 0.1$ inch, the overall ranking consistency verification of the judgment level passes:

$$\mathrm{CR} = \frac{\left(\mathrm{a}_1 \mathrm{CI}_1 + \mathrm{a}_2 \mathrm{CI}_2 + \ldots + \mathrm{a}_m \mathrm{CI}_m\right)}{\left(\mathrm{a}_1 \mathrm{RI}_1 + \mathrm{a}_2 \mathrm{RI}_2 + \ldots + \mathrm{a}_m \mathrm{RI}_m\right)}. \tag{8}$$

3.2. Fuzzy Comprehensive Evaluation Theory and Steps

(1) Determine evaluation factors: set $U = \{u_1, u_2, \cdots, u_m\}$ as m factors (i.e., evaluation index) to describe the evaluated object; M is the number of evaluation factors, which is determined by the specific index system.

(2) Determine the evaluation level: let $V = \{v_1, v_2, \cdots, v_m\}$ be n kinds of decisions (i.e., evaluation grade) that characterize the evaluated object; N is the number of comments, and each comment can correspond to a model subset. The specific level can be described in appropriate language according to the evaluation content. For example, the evaluation of talents in this paper v = {Excellent, Good, General, Poor, Bad}.

(3) Construct evaluation matrix: first of all, for the single factor u_i (i = 1, 2, ⋯ , m) in the key factor set, from the u_i, determine the affiliation of the thing to the decision level v_j ($j = 1, 2, \cdots, n$), so as to obtain the single factor evaluation set of the ith factor U_i: $r_i\,(r_{i1}, r_{i2}, \cdots, r_{im})$.

In this way, the evaluation set of m such factors constitutes a general evaluation matrix R, that is, each evaluated object determines the fuzzy relationship R from U to V, which is a matrix, as shown in equation (9).

$$R = \left(r_{\mathrm{ij}}\right)_{\mathrm{mxn}} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{bmatrix}, \tag{9}$$

where r_{ij} means that starting from the factor u_i, the evaluation object can be rated as the membership of v_j, ($i = 1, 2, \cdots, m$; $j = 1, 2, \cdots, n$).

Specifically, r_{ij} represents the frequency distribution of the ith factor U_i on the j-th comment v_j, which is generally normalized to meet $\sum r_{\mathrm{ij}} = 1$.

(4) Determine the fuzzy weight a of evaluation factors: after obtaining the fuzzy relation matrix, it is still unable to evaluate things. The status and function of each factor in the evaluation factor set in the "evaluation objective" are different, that is, each evaluation factor occupies a different proportion in the comprehensive evaluation. A fuzzy subset a on u is to be introduced, which is called the weight or weight assignment set, $A = (a_1, a_2, \cdots, a_m)$. Where $a_1 > 0$, and $\sum a_i = 1$, it reflects a tradeoff of various factors.

(5) A is synthesized in the R of each evaluated object by using a suitable synthesis operator to obtain the fuzzy comprehensive evaluation result vector B of each evaluated object.

(6) Analyze the result vector of fuzzy comprehensive evaluation.

In order to make the evaluation results more realistic, we can make full use of the information brought by B and consider the rating parameters of various grades and the evaluation result B as a whole. In this way, we can set the parameter column vector relative to each level as

$$C = \left(c_1, c_2, \cdots, c_n\right)^T. \tag{10}$$

Then, the evaluation result of grade parameters is

$$(B \times C) = P. \tag{11}$$

P is a real number, which reflects the inductive information brought by hierarchical fuzzy subset B and hierarchical parameter vector C. In many daily applications, it is a very effective comprehensive parameter.

3.3. Construction of AHP Fuzzy Comprehensive Evaluation Model. The AHP fuzzy comprehensive evaluation system structure comprehensively constructed by the above steps is shown in Figure 2. The comprehensive classification evaluation model constructed in this paper has the following advantages:

(1) Because colleges and universities have the characteristics of complex personnel, great uncertainty and many complex factors when introducing talents, one research method is no longer suitable for the classification and evaluation of talents, so it needs the combination of multiple analysis methods.

(2) In the specific work of talent classification and evaluation, the process of talent introduction itself is very uncertain. Although analytic hierarchy process can solve the problem of qualitative and quantitative multi-objective decision making, it has limitations and needs to be combined with the comprehensive evaluation results of fuzzy comprehensive evaluation methods.

(3) Increase the convenience of operation and calculation of the evaluation method, and make the results more simple and clear.

(4) The evaluation index is more comprehensive, and the index evaluation that cannot be quantified is transformed into quantitative evaluation.

4. Case Analysis

4.1. Questionnaire Survey and Data Statistics

4.1.1. Investigation Purpose. In order to highlight the applicability of the AHP fuzzy comprehensive evaluation model constructed in this paper to the talent evaluation of ordinary colleges and universities, the author conducted a sample survey of teachers and managers from four higher vocational and undergraduate colleges in China. A total of 400 relevant talents were selected for this sampling.

4.1.2. Investigation Form. This paper adopts the method of interview and questionnaire. Interviews and surveys were conducted with some teachers and administrators of

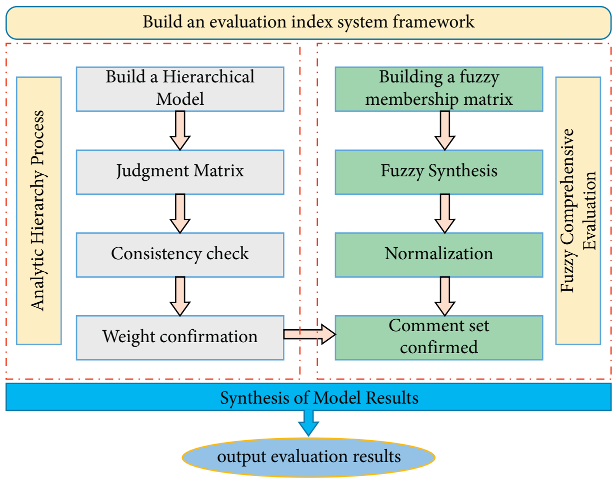


Figure 2: AHP fuzzy comprehensive evaluation model.

national and autonomous regional backbone schools and universities with obvious professional characteristics. The main issues involved include: the current situation of talent training in Colleges and universities, the construction of internship and training projects inside and outside schools, the construction of "double qualified" teachers, the enthusiasm and investment of enterprises in the process of talent training, the role of enterprises in the training process, the role of enterprises in the training of talents, the employment status of students, and the satisfaction of students with teachers and colleges.

The main evaluation indicators included in the questionnaire are shown in Table 4.

4.1.3. Investigation Implementation. It took half a year to complete the distribution, recovery and data statistics of the three questionnaires. A total of 400 questionnaires were distributed, 367 of which were recovered, with a recovery rate of 91.75%, 3 invalid questionnaires, 364 valid questionnaires, and an effective recovery rate of 91%. See Figure 3 for details.

4.2. Evaluation Effect of AHP Fuzzy Comprehensive Evaluation Model. It is known that the weight set of talent evaluation indicators of four colleges and universities can get $W = \{0.154, 0.234, 0.186, 0.157, 0.269\}$ through calculation in Chapter 3, and the P calculated by Excel formula MMULT is a 364×1, and the scores and ranking of all survey objects can be observed from the results. On this basis, according to Table5, the talent level of the above four universities is shown in Table 6. $V = \{\text{Excellent, Good, Average, Poor, Bad}\}$

This result shows that the talent level of the above four universities is in a normal distribution. The "excellent" and "poor" students account for 5.7% and 3.3% of the total sample, respectively, and the students at the intermediate level account for 91% of the total sample. The score difference is small, and the distribution is relatively average. It can be seen from this that the talent level of the above four universities is high, the overall quality of talents is good, and they generally have the basic conditions to become high-quality university talents; A few students have outstanding talent characteristics; There are also fewer teachers whose overall level is poor, and individual teachers do not reflect the development characteristics of talents.

Table 4: Test sample data.

First-level indicator	Second-level indicators
Research level	Paper Patent
Professional skills	Book Experience Job title
Ideological and moral	Political performance Morality Learning attitude Sense of responsibility
Learning ability	Self-study ability Amateur specialty Creativity

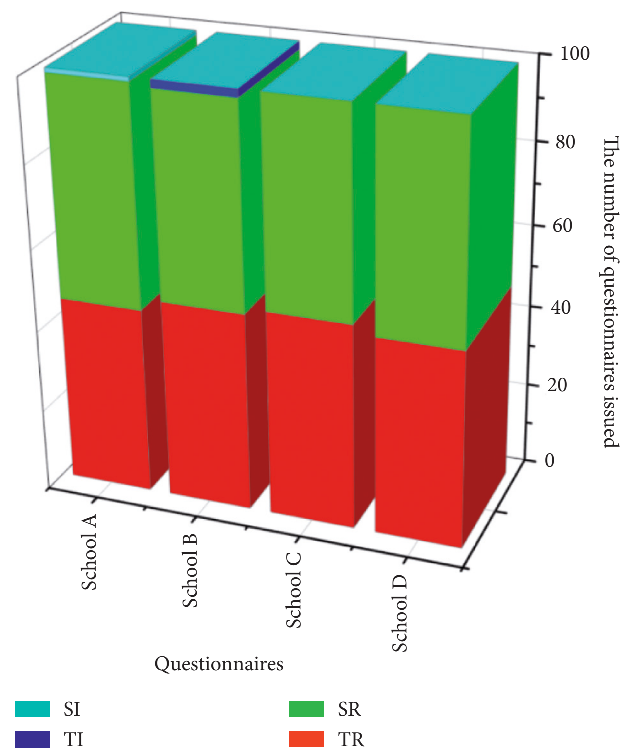


Figure 3: Statistical chart of talent evaluation questionnaire.

Table 5: Evaluation results calculated by principal component analysis.

Overall rating	Number of samples	Evaluation
85–100	19.0	Excellent
75–85	108.0	Good
65–75	110.0	Generally
55–65	109.0	Poor
0–55	18.0	Bad

Table 6: Evaluation results of AHP fuzzy comprehensive evaluation model on talents in four colleges and universities.

Overall rating	Number of samples	Evaluation
85–100	21.0	Excellent
75–85	114.0	Good
65–75	128.0	Generally
55–65	89.0	Poor
0–55	12.0	Bad

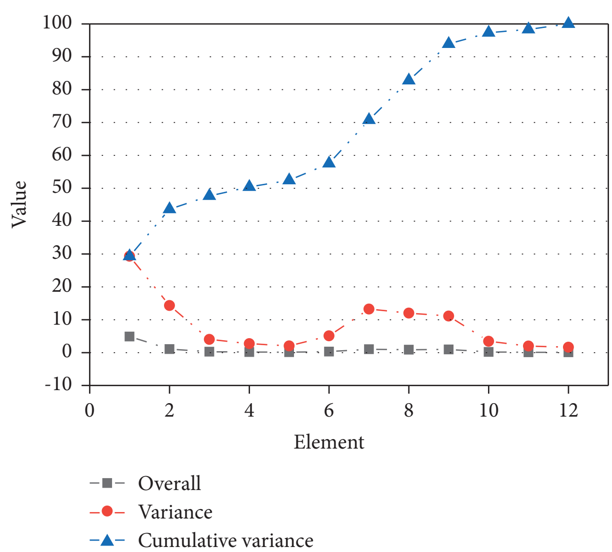


Figure 4: All explanatory variables of principal component analysis.

Figure 5: Explain the correlation between variables.

4.3. Research Based on Principal Component Analysis. In addition, this paper uses 364 effective samples for principal component analysis. By dimensionality reduction, a few principal components are extracted, and the linear function of the original variables is established. Calculate the comprehensive score of each survey factor according to the variance contribution rate of each variable. On the basis of scoring, the overall distribution was observed hierarchically to verify the evaluation results. Overall, 364 groups of survey data were imported into SPSS software to build the database, and the results were obtained by principal component analysis. See Figures 4 and 5 to show the correlation between the components.

As can be seen from Figure 4, the cumulative variance rate of some principal components is more than 80%. The eigenvalue of the first principal component is 4.872, which can explain 29.322% of the data samples, and the latter eigenvalue and variance rate gradually decrease. Therefore, we select 8 principal components for analysis, and calculate the scores of 8 principal components, respectively, by obtaining the principal component regression equation, as shown in Table 5.

As can be seen from Table 6, according to the principal component score, the sample sizes of “excellent” and “poor” are 19 and 18, respectively, accounting for 5.22% and 4.95% of the total sample, which are at a relatively low level; The “medium,” “poor,” and “poor” of the middle level are 78, 78, and 66, respectively, and the overall middle level accounts for more than 90% of the total. It can be seen from this that the results of principal component analysis are consistent with those of AHP fuzzy analysis and evaluation method, that is, the talent level of the above four universities is at the medium level, and the distribution is mainly in the “good” and “middle” range. The number of teachers with basic quality is large, and the foundation for talent introduction is good.

In addition, the factor load of the principal component can also be observed from the coefficient of the principal component equation, that is, the interpretation degree of the principal component to each explanatory variable. The first principal component can better explain these four variables, and has a better ability to explain the indicators of various dimensions such as vocational skills, innovation ability, and diligence. Among them, the third, fifth, and sixth principal components have a better ability to explain work attitude, responsibility, etc., combined with the other principal components, the extracted eight principal components can better explain all variables. Therefore, the first principal component better reflects the scientific research level of the introduced talents; the second principal component can identify the education level of the introduced talents; the third principal component can identify the application of the learned theory transformation in discipline competitions and academic papers; and the fourth principal component can measure their work, respectively.

In the AHP expert scoring and weighting, the scientific research level ability and vocational skill ability are given a large weight, which are, respectively, in the first and second place of the total weight ranking, representing the full attention of colleges and universities to teachers’ scientific research level and teaching and educating ability. This is the basis for higher education to achieve quality improvement and the most fundamental task for colleges and universities to cultivate professional talents for various industries. Therefore, the first and second principal components can be used to measure the overall sample.

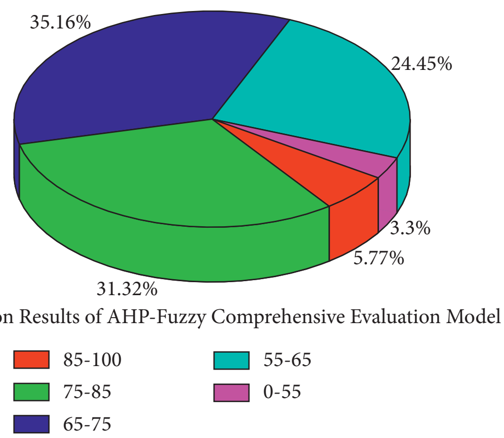


Figure 6: Evaluation results of AHP fuzzy comprehensive evaluation model.

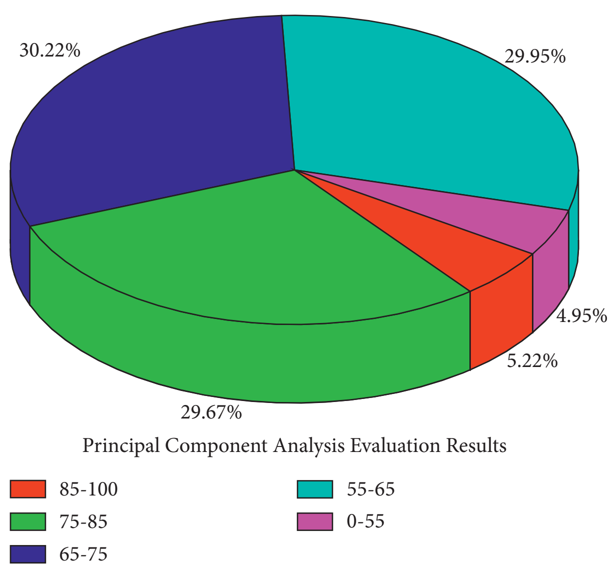


Figure 7: Evaluation results of principal component analysis.

In order to observe the results of the two methods more clearly, we use the histogram sub table to represent the international talent level distribution obtained by the AHP fuzzy comprehensive evaluation method and the principal component analysis method, as shown in Figures 6 and 7.

From the pie chart, the distribution similarity of the comprehensive scores of the two methods is very high, the number of samples of the maximum and minimum values is very small, and the samples are mainly distributed in the representative middle level interval; The sample size of each adjacent interval distribution has little difference, and the distribution is relatively uniform; The maximum sample size calculated by the two methods is within the range of “good” and “general,” and the estimation conclusions of the samples are consistent, indicating that the evaluation results are relatively robust.

It is worth noting that in the results of principal component analysis, the sample size in the “excellent” and “poor” range is more differentiated than the same kind of AHP fuzzy evaluation method, and compared with the principal component analysis method, the screening of excellent and poor talents is more obvious, and it is more in line with the talent introduction strategy of colleges and universities. Therefore, in practical application, the AHP fuzzy evaluation method constructed in this paper has more practical application value. This verifies the accuracy of the model constructed in this paper again.

5. Conclusion

In order to establish a correct employment orientation, encourage and guide the professional development of talents, mobilize the enthusiasm of talents for innovation and entrepreneurship, and speed up the construction of a strong country with talents, this study investigates and analyzes the situation and current situation of the existing talent evaluation system in China in terms of evaluation objectives, evaluation indicators, evaluation methods, evaluation cycle, evaluation procedures, and the use of evaluation results. It also summarizes the problems in the current talent evaluation system of colleges and universities and the various constraints behind the problems, and then clarifies the key elements for different talent evaluation. Based on the different scientific research stages of scientific and technological talents, the domestic talent evaluation indicators are divided into four first level indicators and 12 second level indicators; This paper uses AHP fuzzy comprehensive evaluation method to establish a comprehensive classification evaluation model of college talents, and analyzes and evaluates the results of the two models based on the talent classification index. The example shows that the talent evaluation system based on AHP fuzzy comprehensive evaluation method is more suitable for the talent introduction strategy of colleges and universities at the present stage. This study can not only promote the progress of related model building theory but also provide theoretical guidance and practical reference for the construction of similar systems.

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 they have no conflicts of interest.

References

[1] Y. Stukalina, “Addressing service quality issues in higher education: the educational environment evaluation from the students’ perspective,” *Technological and Economic Development of Economy*, vol. 18, no. 1, pp. 84–98, 2012.

[2] O. Dysthe, K. S. Engelsen, and I. Lima, “Variations in portfolio assessment in higher education: d,” *Assessing Writing*, vol. 12, no. 2, pp. 129–148, 2007.

[3] S. Saxena, “From oil to non-oil: how are private higher education institutions confronting with quality issues?” *Journal of Social & Administrative Sciences*, vol. 3, 2016.
[4] N. D. Hanh, “Review some issues of quality assurance, quality accreditation for higher education institutions and reality of vietnam,” *Science and Education Publishing Co. Ltd*, vol. 7, 2019.
[5] J. Chen, “Research on teaching quality of teachers in colleges based on factor Analysis and degree of grey incidence,” *Journal of Guiyang University (Natural Sciences)*, 2016.
[6] R. Jia-Gang, “Research on teaching quality of young teachers in private undergraduate colleges based on factor Analysis and cluster Analysis——taking shangqiu university as an example,” *Mathematics in Practice and Theory*, 2015.
[7] Y. Guzmán and J. Rafael, “Complex epistemology and educational quality,” *Systematization of experience in higher education institutions*.
[8] H. Liu, “Research on the improvement of young teachers’ teaching ability in applied undergraduate colleges based on teaching scholarship,” *Journal of Hubei Correspondence University*, 2018.
[9] V. Lipska, “Prospects of implementing educational audit in Ukrainian higher education institutions,” *Education - Technology - Computer Science*, no. 1, p. 7, 2012.
[10] A. Mcdonnell and D. G. Collings, *The Identification and Evaluation of talent in MNEs*, Taylor & Francis, Oxfordshire, UK, 2011.
[11] L. C. Huang, P. Wu, B. S. Jaw, and Y. H. Wu, “A study of applying fuzzy analytic hierarchy process on management talent evaluation model,” in *Proceedings of the IFSA World Congress and 20th NAFIPS International Conference, 2001. Joint 9th*, British Columbia, Canada, July 2001.
[12] X. Ji, L. Wang, and H. Xue, “Comprehensive evaluation of talent growth factor index system based on hesitant fuzzy language,” *Journal of Mathematics*, 2021.
[13] P. C. Petrantonakis and I. Kompatsiaris, “On the talent vs. Luck-based evaluation of the classification process,” *IEEE Access*, no. 99, p. 1, 2019.
[14] H. Hua, J. Ye, and C. Chai, “A talent classification method based on SVM,” in *Proceedings of the International Symposium on Intelligent Ubiquitous Computing & Education*, Washington, DC, USA, May 2009.
[15] Y. He, S. Gao, and N. Liao, “An intelligent computing approach to evaluating the contribution rate of talent on economic growth,” *Computational Economics*, vol. 48, no. 3, pp. 1–25, 2016.
[16] J. Han, “System optimization of talent Life cycle management platform based on decision tree model,” *Journal of Mathematics*, vol. 2022, Article ID 2231112, , 2022.
[17] T. Xia, “Based on bd physical education teaching evaluation system research,” *Journal of Physics: Conference Series*, vol. 1744, no. 3, Article ID 032010, 2021.
[18] J. R. Powell, C. J. Kambites, M. Reed, P. Gaskell, N. Lewis, and N. R. Curry, “The Campaign for the Farmed Environment: evaluation of partnership, local and wider impacts. Final Report to Defra,” *Zh.vychisl.mat.mat.fiz*, vol. 5, no. 1, pp. 1473–1483, 2012.
[19] G. Ding, Y. Deng, and S. Lin, “A study on the classification of China’s provincial carbon emissions trading policy instruments: taking Fujian province as an example,” *Energy Reports*, vol. 5, pp. 1543–1550, 2019.
[20] H. Pang, C. Wang, Z. Hu, Library, and S. University, “A theoretical study on the index pool and indicator system for the evaluation of talent introduction in the construction of “World-class” universities,” *Journal of Intelligence*, 2019.
[21] Y. P. Wang, “Competition management system on cultivating competitive sports talent in Chinese schools,” *Journal of Shandong Institute of Physical Education and Sports*, 2013.
[22] K. W. Wang, F. M. Yang, X. Chen, A. Y. Liang, H. C. Cao, and Y. Hou, “Classification evaluation on scientific literacy of TCM practitioners,” *Education of Chinese Medicine*, 2012.
[23] C. Cui and Q. Li, “Software talent evaluation based on dependent degree of e in the case of MIS project management,” *Acta Analysis Functionalis Applicata*, vol. 15, no. 3, pp. 225–233, 2013.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9861498, 1 page
https://doi.org/10.1155/2023/9861498

Retraction

Retracted: Deep Learning-Driven Financial Management Innovation Upgrade for Universities

Mobile Information Systems

Received 13 September 2023; Accepted 13 September 2023; Published 14 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] Y. Guan and C. Chen, "Deep Learning-Driven Financial Management Innovation Upgrade for Universities," *Mobile Information Systems*, vol. 2022, Article ID 9328712, 8 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 9328712, 8 pages
https://doi.org/10.1155/2022/9328712

Research Article

Deep Learning-Driven Financial Management Innovation Upgrade for Universities

Yu Guan[1] and Chen Chen[2]

[1]*College of Information Engineering, Fuyang Normal University, Fuyang 236041, Anhui, China*
[2]*Finance Office of Fuyang Normal University, Fuyang 236041, Anhui, China*

Correspondence should be addressed to Chen Chen; 200807022@fynu.edu.cn

Received 11 July 2022; Revised 3 August 2022; Accepted 11 August 2022; Published 7 September 2022

Academic Editor: R. Mo

This study aims to improve the quality of college financial management and reduce the risk of college financial management, and a college financial system based on multiscale deep learning is designed in this paper. This paper designs a university financial system based on multiscale deep learning. In the hardware design, the system adds multiple sensors and scans all the information in the financial database using a coordinator. In the software design, the weights that can connect the financial information of the same attribute are set by establishing a database form; according to the multilayer perceptual network topology, a full interconnection model based on multiscale deep learning is designed to realize the system's deep extraction of data. The experimental results show that the financial risk is based on the risk warning capability for university finance, and compared with the system under the traditional design, the university finance system designed at this time has the most categories of financial information parameters extracted.

1. Introduction

Financial management of colleges and universities is an important part of public financial management; therefore, the quality of financial management work of colleges and universities will be directly related to the public financial security [1]. The application of modern science and universities and the continuous deepening of the reform supported by the Internet technology are specifically manifested in the following aspects [2–4].

Firstly, it strengthens the construction of financial management information standard system of colleges and universities under the background of the Internet [5]. Given the complexity of financial management work of colleges and universities at the present stage, the financial management department of colleges and universities should systematically sort out its work content and establish the information standard system of college financial management, so as to significantly improve the quality of college financial management work [6–8].

While optimizing the financial management path, we will also implement and improve the corresponding financial management mechanism [9, 10]. At present, the network service platform becomes an important support for the financial management work of colleges and universities. Based on the many advantages of the Internet, the financial management work of colleges and universities should be adjusted with reference to the information service platform, and it should optimize the management mechanism, simplify the management process, and improve the efficiency of the financial management information service platform [11].

Taking advantage of Internet information transmission can effectively improve the quality of financial management information service [12]. For example, the management of the scientific research budget in colleges and universities needs to consider the use of fixed assets, labor cost expenditure, and other related funds, and in order to strengthen the management of the scientific research budget, it is necessary to rely on the information service platform, and the scientific researchers should report their demands through the financial management information service platform and financial management The financial data will be changed after approval by the financial management department, thus making the financial management mechanism more flexible and efficient [13].

The financial management concept cannot keep up with the development of the current situation. (1) The financial management of colleges and universities lacks the necessary modern management concept, financial risk prevention consciousness, and cost management consciousness. At present, compared with enterprises, the financial management of colleges and universities in China not only has great shortcomings in the management content, means and methods, but also lacks the awareness of financial risk prevention and cost management in the management concept. Under the condition of the market economy, the connotation of college financial management is increasingly rich, the function is expanding, and college financial management personnel, if only stay in simple bookkeeping, accounting and other affairs, will not be able to cope with the complex economic situation and the rapid development of college needs. (2) The financial management system of universities is not smooth enough. Although in the early 1990s, China began to explore the establishment of a chief accountant system in colleges and universities, there are not many colleges and universities that established the position of chief accountants. At present, it is more and more difficult to adapt to the requirements of financial management in schools where "the president of the school (college) in charge of financial work acts as the chief accountant." The head of the school (college) in charge of financial work should be very familiar with all aspects of teaching management in the school. However, due to the limitations of the profession, they are not familiar with the information on the fund operation of the whole university, the legal compliance of economic activities, and the efficiency effect and so on. Therefore, it is urgent for colleges and universities to set up proficient chief accountants in economic business to assist the president of colleges and universities to comprehensively lead the financial management of the university, improve the management mechanism of colleges and universities, strengthen the economic management of colleges and universities, and improve the efficiency and benefit of economic operation.

Optimizing the financial risk management mechanism of colleges and universities and building financial risk models are important [14, 15]. The innovation of financial management of colleges and universities in the background of the Internet can use the network data for reasonable control of financial risks and establish a financial risk model, and with the increase of financial data of colleges and universities on the Internet, the prediction accuracy of the financial risk model will be higher, so as to ensure the orderly development of financial management of colleges and universities [16].

2. Related Work

Many scholars have made breakthroughs in the research of financial intelligence in recent years, and intelligent financial technology has been applied to some extent in large and medium-sized universities. From system construction to management changes, promising results have been achieved. Literature [17] proposed the realization path of the model, literature [18] proposed the realization and application of financial sharing based on a cloud platform and made a more in-depth study on tax management optimization, and literature [19] proposed corresponding countermeasures to the difficulties of financial data sharing. Tan Qing proposed the construction model of a dynamic accounting information platform. Literature [12] studies the financial analysis system of large universities and points out that business, financial, and management information systems within universities need to be organically integrated. Realizing the unification of bill processing is conducive to the centralization of work and personnel in colleges and universities. Literature [13] based on realizing the business-financial integration between marketing and finance departments through financial intelligence proposes an integrated control scheme to improve the efficiency of financial control and to improve the experience of business departments. Literature [14] explains the application of API technology to realize information interaction and settlement internally, with partners, and with open unspecified objects and demonstrates the great convenience of this technological advancement for the financial management of universities. Literature [15] examines the implementation of financial intelligence in the manufacturing industry from procurement, manufacturing, warehousing, and cost management and proposes a cost management platform based on target costs, using scenario simulation for cost adjustment until the target needs of cost management are met. It also makes flexible budgeting more feasible.

In [16], more technical exploration of the evolution of financial technology upgrade from information technology to automation, intelligence, and digitalization is conducted. It is considered that financial management intelligence is a financial change driven by both information technology development and quantitative multiplication of university scale. Technological advancement is the external driving force, and the scale multiplication of universities is the internal motivation. Literature [17] analyzed the application scenarios of the more mature automation and intelligent technologies at present. It is pointed out that a good IT foundation is a necessary basis for AI implementation. It is considered that the current intelligent technology has been able to do more in the automatic processing of internal processes, automatic data exchange between universities and partners, automatic processing of accounts, intelligent audit, intelligent audit, supplier management, and risk prevention and control. Several intelligent scenarios and application processes are designed. The issues to be considered by universities in choosing RPA and the reference path for implementing RPA are proposed. It is also the fundamental criterion to evaluate the success of intelligence.

Literature [20] conducted more research on the change of financial management under the intelligent scenario and believed that the focus of finance people will gradually shift to management, and control will become the core of financial management [18] and put forward the theory of control mechanism [19]. Combining the wisdom of finance people with artificial intelligence to achieve the integration of human-computer intelligence improves the unstructured data processing capability of AI. With the powerful computing ability of AI, it can simulate a variety of scenarios and input emotional parameters for scientific calculation, thus adjusting the influence of financiers and college decision makers due to financial psychological accounts and helping rational decisions made by colleges and universities. The financier assumes the role of a basic knowledge inputter of colleges and universities and promotes the evolution of financial intelligence, which is more in line with the personalized needs of colleges and universities. At the same time, the focus of finance people shifts from information collection to management control.

Literature [19] studied the case of establishing the financial information system by combining financial outsourcing and a self-built financial center. It was pointed out that direct promotion from the top, adequate research and preparation in advance, and gradual implementation in phases according to the actual situation of each ministry are the guarantees of successful implementation. The establishment of the professional financial management subcenter, which is dedicated to promote, is the key to achieving the change in each business. In terms of talent, team building needs to be cultivated and introduced at the same time. Emphasis on the exchange of relevant departments is as important as providing professional training for continuous superior talent allocation and for establishing a dynamic financial organization system.

This paper introduces the composition of the financial management risk system of colleges and universities in detail and proposes a scientific construction plan for the financial risk early warning system of colleges and universities according to the special characteristics of financial management work of colleges and universities. For example, the risk of fixed assets management of colleges and universities is mainly directly related to the frequency of use, curriculum arrangement setting, and scientific research project declaration. Taking relevant data as the influencing factor of the risk model of fixed assets management of colleges and universities, it can improve the quality of fixed assets management and financial risks caused by improper management of fixed assets. 3 Hardware design of the university financial system is based on multiscale deep learning. The main problems of the current system are as follows.

The internal control system is not sound and becomes a mere formality. Budget management is not perfect. The standard of asset allocation is unreasonable. The overall quality of financial management personnel is not high, which restricts the improvement of financial management levels in colleges and universities, and there is lack of an effective cost accounting system. To meet the multiscale deep learning requirements of the university financial system, the system hardware needs to be redesigned so that the hardware meets the fully interconnected mode of the multilayer perceptron network under multiscale deep learning. It is known that the most critical hardware for the financial system to achieve the fully interconnected mode is the sensor and the coordinator. We can use the sensor to improve the transmission, processing, storage, display, recording, and control of electrical signals, and we can use the coordinator to scan the financial data in the system and select the appropriate parameters to build a multilayer perceptual network that can be used for multiscale deep learning. The block diagram of the hardware design of the financial system of the university is shown in Figure 1.

According to Figure 1, the coordinator ZigBee is connected to each node in the acquisition unit, processing unit, and control unit in order to obtain the parameters suitable for the whole network and to establish a network with complete coverage of the system. The newly selected sensor model is MGS-2483-H24, and the hardware is connected to the data collection unit for searching the financial data in the university financial system. According to the block diagram, one sensor is connected to each collection node to ensure that all collected data can be uploaded to the data processing center of the system. According to the hardware design block diagram shown in Figure 1, the hardware of the financial system is reconnected to realize the hardware design of the university financial system that satisfies multiscale deep learning.

3. Software Design of the University Financial System Based on Multiscale Deep Learning

3.1. Creating Database Form to Set Connection Weights. In the financial system of higher education, the financial information covered is of various types, complex contents and refined subjects may appear at the third or fourth levels, and the system may carry out multiscale deep learning by designing a database form with a sharp, clear objective and distinct type division and by setting a weight value to connect the same attribute information. The basic information about college users and college workers is shown in Table 1 and Table 2.

With Table 1 and Table 2 as the two major categories for financial information retrieval, the financial information of universities in terms of customers and employees is recorded and updated in real time. Set up information such as customer payment items, drug categories, and charge numbers is shown in Table 1; and set up information such as basic salary, job salary, seniority salary, employee benefits, incentive salary, and social insurance is displayed in Table 2. According to all the financial information of the university, we set the weights of the connected financial information, and the equation of the change of the weights is shown in the following equation:

$$\Delta q_{ij} = -\mu \frac{\partial D}{\partial q_{ij}} = -\mu \frac{\partial}{\partial q_{ij}} \left(\sum_{m=1}^{m} D_m \right) = \sum_{m=1}^{m} \left(-\mu \frac{\partial D_m}{\partial q_{ij}} \right), \quad (1)$$

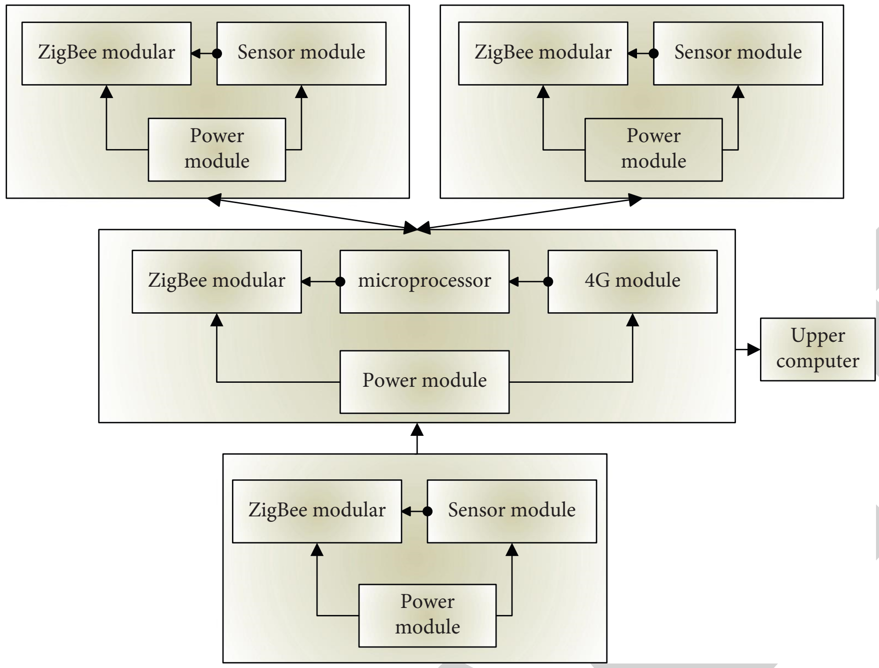


Figure 1: Hardware design framework of the financial system in higher education.

Table 1: Basic information of users.

Yuduan name	Data type	Length	Remarks	Crux
EID	Varchar	10	Customer number	Yes
Name	Varchar	10	Full name	No
Tel	Varchar	10	Contact information	Yes
DID	Varchar	20	Affiliated unit	No
Area	Varchar	50	Place of residence	Yes
DName	Varchar	10	Unit type	No

Table 2: Basic information of employees.

Yuduan name	Data type	Length	Remarks	Crux
EID	Varchar	10	Employee number	Yes
Name	Varchar	10	Employee name	Yes
Sex	Varchar	10	Gender	No
Birthday	Date	10	Date of birth	No
Hometown	Varchar	10	Native place	No
Address	Varchar	50	Current residence	Yes
Tel	Varchar	10	Contact number	Yes
ID_number	Varchar	10	ID number	Yes
Department	Varchar	20	Department	Yes

where q_{ij} denotes the connection weights between financial information; Δq_{ij} denotes the change of weights; i and j denote two random financial information; μ denotes the learning rate of the system; m denotes the same attribute of financial information; D_m denotes the global error value. According to the information of the established database form, the weights that can connect the financial information of universities are set.

3.2. Designing a Fully Interconnected Model Based on Multiscale Deep Learning. It is known that the perceptron has a single-layer computing capability and belongs to a kind of feed-forward network, and according to the set weights, bottom-up information transmission can be performed for each layer of the network. Therefore, according to this function of the perceptron, a multilayer perceptual network is constructed, and in this way, the full interconnection pattern of the system information is set; the connections between neurons in different layers are used to mine the financial information of the university.

The perceptron is used as each node in the neural network, and a dynamic connection weight is set according to the result of equation (1), and then, the perceptron is used to learn this weight; the topology of the multilayer perceptual network is schematically shown in Figure 2.

In Figure 2, x denotes the random financial information; H denotes the input layer unit; hk denotes the hidden layer unit; K denotes the output layer unit; Y denotes the final result. According to Figure 2, neurons in the same layer are not connected to each other and neurons in two adjacent layers are fully connected to each other, and the data transmitted from the input to the output are calculated layer by layer through directional information transfer. The multilayer perceptual network setup contains not only input and output layers but also one or more hidden layers, allowing the system to extract financial data features with associated properties within the system during deep learning, which is the forward propagation system interconnection. In contrast, the reverse system interconnection of the perception network is designed using the back-propagation algorithm, which is a secondary assignment of weights, feeding the input pattern into the system from the input side and adjusting the weights of the previous layer by using the error between the output value and the target output value, thus realizing the full interconnection pattern of the multilayer perception network.

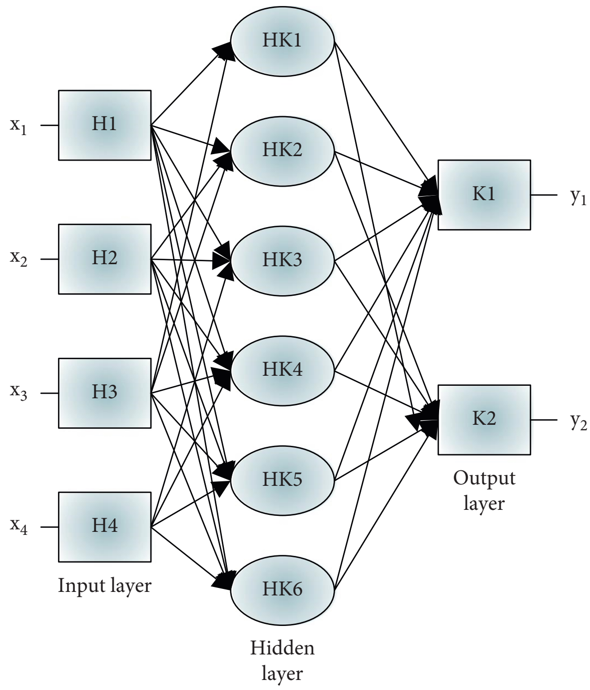


Figure 2: Topology of the multilayer sensing network.

It is known that the input layer of the perceptual network has a input neurons, the output layer has b output neurons, the hidden layer has e hidden neurons, q_{ij} plus formula (1) sought is the connection weights between the output hidden layer and the output layer, and p_{ij} is the connection weights, and then, the output results of the hidden neurons and the output layer neurons under forward propagation are shown in the following equation:

$$\begin{cases} K = f_1\left(\sum_{i=1}^{a} p_{ij}X_n\right), \\ Y = f_2\left(\sum_{i=0}^{e} p_{ij} - K\right). \end{cases} \tag{2}$$

Using the above equation, the multilayer perception network completes the forward propagation and establishes the spatial mapping from a dimension to b dimension. Under the definition of the backpropagation algorithm, let the input data be c, denoted by $x_1, x_2 \ldots x_c$, and the output value is calculated after the cth data are input into the network. Assuming that the error function is squared, the error between the cth output value and the target output value is calculated as shown in the equation.

$$\Delta Y = \frac{1}{2}\sum_{c=1}^{n}(X_c - Y_c), \tag{3}$$

where ΔY is the error and X_c denotes the original input and also the desired output. Adjusting the connection weights p_{ij} between the input layer and the output hidden layer according to this result, the equation of change of p_{ij} is calculated as follows:

$$\Delta p_{ij} = \sum_{i=1}^{c}\sum_{j=1}^{n}\mu(X_c - Y_c)f_2'(S_{HK})q_{ij}f_1'(S_K)q_{ij}, \tag{4}$$

where $f_2'(S_{HK})$ is the partial differential of the hidden layer transfer function and $f_1'(S_K)$ is the partial differential of the output layer transfer function. By adding p_{ij} to the result of equation (4) and similarly q_{ij} to the result of equation (1), we can use the connection weights of the two directions to realize the data analysis of positive and negative directions and complete the setting of the full interconnection mode of the multilayer perception network so that the design of the university financial system based on multiscale deep learning is completed [21–23].

4. Empirical Results

4.1. Model Evaluation. Constructing a financial reporting fraud identification model with good out-of-sample prediction ability is crucial to the research in this paper. For both the deep learning model and the benchmark model, three types of evaluation metrics, precision, recall, and F1-score, are used to measure the classification performance of the model on the test set.

The precision and recall rates can be expressed as

$$\text{Precision} = \frac{TP}{TP + FP},$$
$$\text{Recall} = \frac{TP}{TP + FN}. \tag{5}$$

In particular, TP (true positive (TP)) represents the MD&A text that is predicted by the model to be a fraud sample and is itself a fraud sample; FP (false positive (FP)) represents the MD&A text that is predicted by the model to be a fraud sample and is itself a nonfraud sample; FN (false negative) represents the MD&A text that is predicted by the model to be a nonfraud sample and is itself a fraud sample [24–27]. The F1-score is a commonly used statistical measure of model dichotomization performance and can represent a harmonic mean of the model's precision and recall rates, and it is expressed as

$$F1 - \text{score} = 2 \cdot \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{6}$$

The values of all three metrics are between 0 and 1, and the value closer to 1 indicates a better classification performance of the model. However, all three categories of indicators are calculated by assuming a positive sample for fraudulent texts and a negative sample for nonfraudulent texts. In order to investigate the classification performance of the model in different categories, the macroaverage method was introduced, which is the arithmetic average of the evaluation index values obtained when the fraud and nonfraud samples are positive, respectively.

4.2. Empirical Results and Analysis. Tables 3 and 4 summarize the prediction results of the deep learning model and the benchmark model on the out-of-sample dataset (evaluation metrics are macroaveraged) and the prediction performance of the deep learning model in different categories, using MD&A texts from the university periodic reports, respectively. The architecture of the deep learning model is based on a word embedding model and a character-level convolutional nerve network, while two types of statistical models (logistic regression and plain Bayesian models) and three types of shallow models (support vector machines, random forests, and gradient boosting decision trees) are selected for the benchmark model. Based on the evaluation metrics presented in Tables 3 and 4, the empirical results of the study can be summarized in the following four points: first, the classification performance of the models implemented in the study, both the deep learning model and the other benchmark models, is greater than 0.7, indicating that the models can effectively use the textual information in MD&A for financial reporting fraud identification; second, the classification performance of deep learning is significantly higher than that of the other benchmark models. The framework constructed in this paper can better identify fraudulent financial reports than the models used in traditional intelligent financial reporting fraud detection studies. The evaluation index values of the deep learning models on both types of MD&A text sets are greater than 0.82, indicating that the models show better fraud identification ability on different types of datasets [28–30].

4.3. Financial System Performance Testing. Based on the experimental data obtained from above, the system functions and operation are black box tested through the system's dynamic income statement function module and financial trend analysis, etc. After logging into the system, we enter the dynamic income statement interface. We then analyze the correctness of the detailed data in the financial multidimensional analysis table according to the test cases, check the income and expense reports, and check whether the comparison of the same period of the previous year has been shown graphically. We enter the financial trend analysis page, and first, we check the correctness of the parameter setting on the left side, set the year of salary withdrawal as a random year from January 1, 2014, to December 31, 2018, select the section department, submit the query, and check the section financial trend analysis chart and related financial index information as well as equipment yield tracking, future growth trend, and cost management trend analysis in the returned results. The efficiency of the financial system operation is tested as shown in Figure 3.

We compare the proposed system with the other three groups of methods. As can be seen from Figure 3, the operating efficiency of the system designed in this paper is significantly higher than that of the system used in the control experiment conducted by the three groups, and the highest operating efficiency of the system designed in this paper is close to 96%, indicating that the performance of the university financial system designed in this paper is superior. Although system B outperforms systems A and C, its performance is still lower than that of our method.

Table 3: Classification performance of deep learning models and benchmark models on MD&A text datasets.

Model	Macro_ Precision	Macro_ Recall	Macro_ F1-score
Word embedding + character-level CNN	0.89	0.850	0.851
Logistic regression	0.80	0.81	0.801
Naive Bayes	0.795	0.795	0.795
Support vector machine	0.811	0.82	0.812
Random forest	0.767	0.768	0.768
XGBoost	0.734	0.734	0.734
LightGBM	0.769	0.769	0.769

Table 4: Classification performance of deep learning models on two types of MD&A text datasets.

Sample category	Precision	Recall	F1-score
Fraud	0.87	0.83	0.84
Nonfraud	0.84	0.86	0.85

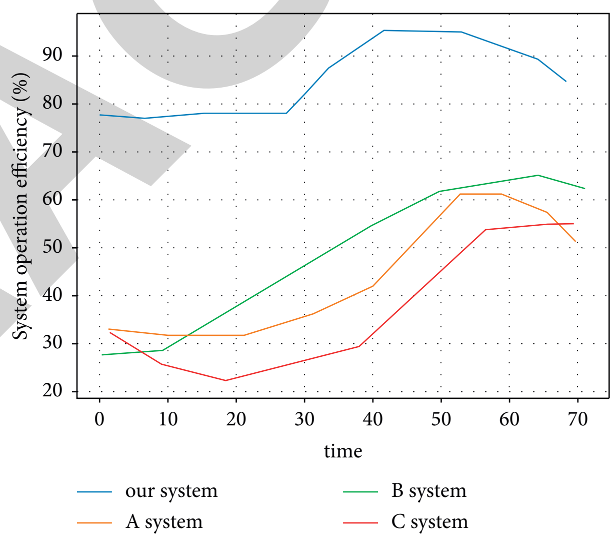


Figure 3: Operation efficiency under different system conditions.

5. Conclusions

Financial reports, as publicly disclosed information, directly reflect the operation of universities and therefore become an important medium for universities to commit fraudulent acts. Effective identification of financial reporting fraud has become one of the important tools to regulate the operation order of the financial market. This paper constructs a character-level CNN model for identifying fraudulent financial reports of colleges and universities by taking advantage of the breakthroughs in NLP in deep learning technology and uses the MD&A text in financial reports as the analysis sample. The results show that the character-level CNN model can still show better classification performance on a small sample dataset where the shallow model has obvious advantages, without complex textual feature directed extraction, and this result also provides direct evidence for using deep learning techniques to improve existing financial reporting fraud identification methods. In addition, the study not only demonstrates the advantages of deep learning models in identifying financial reporting fraud but also the models constructed show better prediction performance on MD&A texts, which reflects the utility and information value of public texts disclosed by universities. This reflects the usefulness and information value of the public text disclosed by universities. At the same time, the textual disclosure of financial reports, as a reliable and easily available data source, can provide good data support for related studies.

Data Availability

The experimental data 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

The study was supported by the Quality Engineering Project of Department of Education, Anhui Province: Research on the Teaching Model of Bisection Class in the Intelligent Classroom Environment take financial management section for example (2021jyxm1120); Quality Engineering Project of Department of Education, Anhui Province: Research and Practice of Financial Management in the context of new liberal arts (2021sx118); Quality Engineering Project of Department of Education, Anhui Province: Exploration and Practice of Software Engineering Practice Teaching System under the background of new engineering (2021sx119); Research Projects in the Humanities and Social sciences of College of Information Engineering, Fuyang Normal University: A study on the relationship between ESOP plan and pay stickiness (FXG2021SZ01); Key Research Project of Talent Fund of College of Information Engineering, Fuyang Normal University: Research on the Intermediary Effect between Cash Dividend and Financial Performance (2022xgrcxm03); Key Research Projects of Humanities and Social Sciences in Universities of Anhui Province: Research on Carbon Information Disclosure and Enterprise Performance under the background of carbon neutrality (SK2021ZD0060).

References

[1] B. Wei, R. A. Hamad, L. Yang et al., “A deep-learning-driven light-weight phishing detection sensor,” *Sensors*, vol. 19, no. 19, p. 4258, 2019.

[2] K. B. Hansen and C. Borch, “Alternative data and sentiment analysis: prospecting non-standard data in machine learning-driven finance,” *Big Data & Society*, vol. 9, no. 1, Article ID 205395172110707, 2022.

[3] I. Lee and Y. J. Shin, “Machine learning for enterprises: applications, algorithm selection, and challenges,” *Business Horizons*, vol. 63, no. 2, pp. 157–170, 2020.

[4] X. Xie, W. Zhang, H. Wang et al., “Dynamic adaptive residual network for liver CT image segmentation[J],” *Computers & Electrical Engineering*, vol. 91, 2021.

[5] A. Samacoits, P. Nimsamer, O. Mayuramart et al., “Machine learning-driven and smartphone-based fluorescence detection for CRISPR diagnostic of SARS-CoV-2,” *ACS Omega*, vol. 6, no. 4, pp. 2727–2733, 2021.

[6] D. Appelt, C. D. Nguyen, A. Panichella, and L. C. Briand, “A machine-learning-driven evolutionary approach for testing web application firewalls,” *IEEE Transactions on Reliability*, vol. 67, no. 3, pp. 733–757, 2018.

[7] J. Zhang and R. X. Gao, “Deep learning-driven data curation and model interpretation for smart manufacturing,” *Chinese Journal of Mechanical Engineering*, vol. 34, no. 1, p. 71, 2021.

[8] M. C. Elish, “The stakes of uncertainty: developing and integrating machine learning in clinical care,” *Ethnographic Praxis in Industry Conference Proceedings*, vol. 1, pp. 364–380, 2018.

[9] P. Ding, M. Jia, and X. Zhao, “Meta deep learning based rotating machinery health prognostics toward few-shot prognostics,” *Applied Soft Computing*, vol. 104, Article ID 107211, 2021.

[10] W. Ma, Z. Liu, Z. A. Kudyshev, A. Boltasseva, W. Cai, and Y. Liu, “Deep learning for the design of photonic structures,” *Nature Photonics*, vol. 15, no. 2, pp. 77–90, 2021.

[11] A. Boutet, J. Germann, D. Gwun et al., “Sign-specific stimulation ‘hot’and ‘cold’spots in Parkinson’s disease validated with machine learning,” *Brain communications*, vol. 3, no. 2, Article ID fcab027, 2021.

[12] D. Krishnamurthy, H. Weiland, A. Barati Farimani, E. Antono, J. Green, and V. Viswanathan, “Machine learning based approaches to accelerate energy materials discovery and optimization,” *ACS Energy Letters*, vol. 4, no. 1, pp. 187–191, 2018.

[13] J. Samuel, G. G. M. N. Ali, M. M. Rahman, E. Esawi, and Y. Samuel, “Covid-19 public sentiment insights and machine learning for tweets classification,” *Information*, vol. 11, no. 6, p. 314, 2020.

[14] X. Xie, X. Pan, W. Zhang, and J. An, “A context hierarchical integrated network for medical image segmentation,” *Computers & Electrical Engineering*, vol. 101, Article ID 108029, 2022.

[15] M. N. Alonso, G. Batres-Estrada, and A. Moulin, “Deep learning in finance: prediction of stock returns with long short-term memory networks,” *Big Data and Machine Learning in Quantitative Investment*, vol. 1, pp. 251–277, 2018.

[16] G. Herrigel, V. Wittke, and U. Voskamp, “The process of C hinese manufacturing upgrading: transitioning from unilateral to recursive mutual learning relations,” *Global Strategy Journal*, vol. 3, no. 1, pp. 109–125, 2013.

[17] A. Pramod, H. S. Naicker, and A. K. Tyagi, “Machine Learning and Deep Learning: Open Issues and Future Research Directions for the Next 10 Years,” *Computational analysis and deep learning for medical care: Principles, methods, and applications*, vol. 3, pp. 463–490, 2021.

[18] Q. Tong, P. Gao, H. Liu et al., “Combining machine learning potential and structure prediction for accelerated materials design and discovery,” *Journal of Physical Chemistry Letters*, vol. 11, no. 20, pp. 8710–8720, 2020.

[19] W. Zhu, C. Ma, X. Zhao et al., “Evaluation of sino foreign cooperative education project using orthogonal sine cosine optimized kernel extreme learning machine,” *IEEE Access*, vol. 8, pp. 61107–61123, 2020.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 5907051, 10 pages
https://doi.org/10.1155/2022/5907051

Research Article

Deformation Prediction and Analysis of Soft Rock Roadway with High Altitude and Large Buried Depth Based on Particle Swarm Optimization LSTM Model

Bin Du, Huahui Yi, and Fan Yang

Xi'an Technology University, Xi'an, Shaanxi 710062, China

Correspondence should be addressed to Bin Du; dubin@xatu.edu.cn

Received 19 July 2022; Revised 16 August 2022; Accepted 23 August 2022; Published 6 September 2022

Academic Editor: R. Mo

Deformation prediction is an important basis for roadway information construction, especially for soft rock roadway at high altitude and large buried depth, whether the deformation of roadway surrounding rock can be effectively and accurately predicted is an important basis for judging the stability of roadway surrounding rock. However, at present, the research on the informatization construction of the roadway is not in-depth, and the intelligent prediction technology for the deformation of the surrounding rock of the roadway is still in its infancy, and the accuracy of deformation prediction is also low. Therefore, based on the research of domestic and foreign researchers, in order to solve the breakthrough of related technology, this paper puts forward the deformation prediction and control technology of high altitude and deep buried soft rock roadway based on a neural network model. This method is based on the traditional prediction model and is replaced by the neural network, so as to improve the problems of low accuracy and large prediction deviation in the related deformation prediction of the traditional prediction model. At the same time, aiming at the problem of poor local weight and network search ability, an improved method using particle swarm optimization algorithm is proposed, which effectively considers the influence of local and global factors on the combined weight. Finally, the improved deformation prediction model of high altitude and deep buried soft rock roadway based on particle swarm optimization LSTM model is applied to an engineering example and compared with the traditional model to explore its feasibility and effectiveness. The results show that the prediction model has higher prediction accuracy than the traditional prediction model, and the relative deviation of the prediction results is controlled within 2%. At the same time, compared with other models (BP neural network model), it has relatively higher accuracy and stability. The research results can provide a new idea for the deformation prediction of soft rock roadway with high altitude and deep burial.

1. Introduction

With the rapid development of China's economy, the demand for energy is increasing. Energy mining, especially coal mining, has entered deep mining from shallow mining. In order to ensure the safety of mine production, how to accurately predict the deformation of surrounding rock of deep roadway and give a reasonable support scheme is particularly important.

At present, with the increase of mining depth of coal mines in various geological environments, affected by geological conditions, it is inevitable to face more and more soft rock problems. Many scholars have done a lot of research on the deformation prediction of deep soft rock roadway, and have also achieved many important research results. The deformation prediction methods of surrounding rock of roadway commonly used in engineering circles [1–3] mainly include theoretical formula calculation method, experimental method, engineering experience method, numerical simulation method, neural network deformation prediction method, and so on.

Although the theoretical calculation method can get a more accurate analytical solution, it must assume harsh preconditions in the calculation process, and can only solve relatively simple mechanical models. For underground engineering, the geological environment of rock mass is complex, and the layout and section size of roadway are complex, which is generally difficult to be solved by theoretical calculation method [4–6]. Experimental methods generally include indoor test and in-situ test. When carrying out the indoor test of samples, due to the damage caused by the sampling process to the influence of stress disturbance and size effect, the obtained rock mass mechanical parameters are much different from the actual rock mass mechanical parameters [7]. On the other hand, because the rock mass is a heterogeneous structure, the rock mass is divided by different joints and fissures, and the joints and fissures are filled with different substances, resulting in the extremely uneven stress of the rock mass and the complex geological environment structure of the rock mass, which makes the rock mass in different parts have great discreteness. It is difficult to comprehensively reflect the mechanical parameters of rock mass through field tests, and the mathematical model thus established is more difficult to reflect the real deformation, and the deformation prediction thus obtained is more difficult to achieve high prediction accuracy [8–11]. Because the rock masses of different strata and different sections are very different, the role of the engineering experience method in different engineering environments can be ignored. For underground engineering, the numerical calculation method [12–15] has obvious advantages over other methods. It can calculate most underground engineering problems, with wide applicability, fast calculation speed, and low cost. However, during the modeling of numerical simulation, due to the difficulty of obtaining parameters, the numerical model thus constructed is also difficult to reflect the real situation. The natural accuracy of roadway prediction is also difficult to meet people's expectations.

As a new deformation prediction technology developed in recent years, the neural network prediction method, due to its strong nonlinear processing ability [16–19], when predicting the deformation of roadway surrounding rock, does not need to assume harsh preconditions like the theoretical method, nor many parameters like the numerical simulation method, but only some basic parameters and data to realize the accurate prediction of surrounding rock, and gradually liked by people [4, 20–23]. At present, the artificial neural network model has been widely used in various fields of engineering, and the prediction accuracy of surrounding rock deformation of relevant tunnels has also been improved year by year, but there is still a large research space [3, 24, 25].

In view of the above problems, in order to explore the deformation law of soft surrounding rock under the geological conditions of high altitude and large buried depth and accurately predict it, this paper, based on the large buried deep soft rock roadway in a high altitude area in China, uses neural network and field monitoring methods to carry out the deformation prediction research of large buried deep soft surrounding rock roadway in high altitude area, so as to provide a reference for the design of similar high altitude and large buried deep soft surrounding rock roadway provide a reference for construction and deformation prediction.

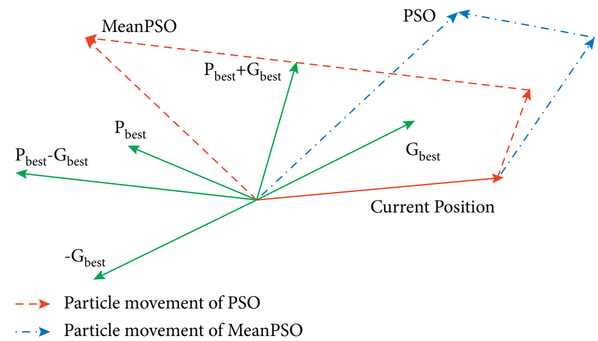


Figure 1: MeanPSO algorithm and particle evolution process of PSO algorithm.

2. Overview of Particle Swarm Optimization Algorithm and Human Long-Term and Short-Term Memory Neural Network

2.1. General PSO. PSO algorithm is a search method based on group cooperation, which is generated by simulating birds seeking the optimal route when foraging [26–28]. In PSO algorithm, there is a particle space, in which each particle has its own initial position and initial velocity, forming a d-dimensional space with m particles. Through continuous iteration, the speed and position of particles are constantly updated to form local optimal solutions and global optimal solutions.

In ordinary PSO algorithm, particles search for the optimal particle by learning their own historical experience (P_{best}) and group experience (G_{best}). For the optimization problem with variable $X = (x_1, x_2, \cdots, x_D)$ and objective function $\min\{f(x)\}$, the particle update formula of the standard PSO algorithm is formulas (1) and (2)

$$v_{id}(t+1) = wv_{id}(t) + c_1 r_1\left(P_{best_{id}} - x_{id}(t)\right) + c_2 r_2\left(G_{best_{id}} - x_{id}(t)\right), \tag{1}$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1), \tag{2}$$

where $v_{id}(t+1)$ stands for speed and $x_{id}(t+1)$ are stands for position, w stands for the weight, and w in ordinary PSO decreases with the number of iterations.

2.2. Improved PSO Algorithm. Through the above analysis, it can be found that for the ordinary PSO algorithm, its speed and position update of particles are relatively common in the path exploration of particles, which is often difficult to meet the needs of practical projects. In the face of complex data, it has great limitations. Therefore, aiming at the above problems, this section focuses on the improvement of particle speed and position update. The specific improvement process includes the following two points.

2.2.1. Improvement of Particle Velocity Update Formula. Reference [27] proposes a mean particle swarm optimization (MeanPSO) algorithm, which uses the linear combination $P_{best_{id}} + G_{best_{id}}/2$ and $P_{best_{id}} - G_{best_{id}}/2$ of individual optimization and group optimization to replace $P_{best_{id}}$ and $G_{best_{id}}$ in formula (1), respectively, and obtains a new particle velocity update formula, which is formula (3)

$$v_{id}(t+1) = wv_{id}(t) + c_1 r_1\left(\frac{P_{best_{id}} + G_{best_{id}}}{2} - x_{id}(t)\right) + c_2 r_2\left(\frac{P_{best_{id}} - G_{best_{id}}}{2} - x_{id}(t)\right). \quad (3)$$

The particle motion process in MeanPSO algorithm and PSO algorithm is shown in Picture. From Figure 1, it can be seen that the particle search interval in meanpso algorithm is wider, which makes the algorithm more likely to search for the global optimal solution in the early stage of evolution.

2.2.2. Improvement of Particle Position Update Formula. Reference [29] proposes a hierarchical simplified PSO algorithm with average dimension information. Phspo algorithm discards the particle velocity update item in PSO algorithm and introduces the concept of average dimension information, that is, the average value of all dimensional information of each particle, and the calculation formula is formula (4). At the same time, PHPSO algorithm decomposes the particle position update formula into three modes, They are equations (5)–(7), respectively.

$$P_{ad}(t) = \frac{1}{D}\sum_{i=1}^{D} x_{id}(t), \quad (4)$$

$$x_{id}(t+1) = wx_{id}(t) + c_1 r_1\left(P_{best_{id}} - x_{id}(t)\right), \quad (5)$$

$$x_{id}(t+1) = wx_{id}(t) + c_2 r_2\left(G_{best_{id}} - x_{id}(t)\right), \quad (6)$$

$$x_{id}(t+1) = wx_{id}(t) + c_3 r_3\left(P_{ad} - x_{id}(t)\right). \quad (7)$$

Among them, equation (5) is conducive to the global development ability of the algorithm; equation (6) is conducive to the local exploration ability of the algorithm, to helps the algorithm jump out of the local optimization; equation (7) helps to improve the convergence speed of the algorithm. In the iterative process, the algorithm selects different modes based on probability to update the particle position.

Reference [26] points out that using "$X = X + V$" to update the particle position helps to improve the local exploration ability of the algorithm, and "$X = wX + (1-w)V$" helps to Polish up the global development ability of them. This paper proposes an adaptive particle position update mechanism, as shown in equations (8) and (9)

$$P_i = \frac{\exp\left(fitt\left(x_i(t)\right)\right)}{\exp\left(1/N\sum_{i=1}^{N} fit\left(x_i(t)\right)\right)}, \quad (8)$$

$$x_{id}(t+1) = \begin{cases} wx_{id}(t) + (1-w)v_{id}(t+1), & p_i > rand \\ x_{id}(t) + v_{id}(t+1), & else \end{cases}, \quad (9)$$

where $fit(\cdot)$ stands for the fitness value of particles, and N stands for the number of particles in the population. In equation (8), p_i stands for the ratio of the current particle fitness value to the average fitness value of all particles in the population.

Combined with the improved optimization strategy of particles for particle speed and position, a new adaptive particle speed and position update strategy can be obtained, and its update method and process are shown in equations (10) and (11)

$$\begin{cases} v_{id}(t+1) = wv_{id}(t) + c_1 r_1\left(\dfrac{P_{best_{id}} + G_{best_{id}}}{2} - x_{id}(t)\right), \\ +c_2 r_2\left(\dfrac{P_{best_{id}} - G_{best_{id}}}{2} - x_{id}(t)\right), \\ x_{id}(t+1) = wx_{id}(t) + (1-w)v_{id}(t+1), \end{cases} \quad (10)$$

$$\begin{cases} v_{id}(t+1) = wv_{id}(t) + c_1 r_1\left(p_{ad} - x_{id}(t)\right), \\ \qquad +c_2 r_2\left(G_{best_{id}} - x_{id}(t)\right), \\ x_{id}(t+1) = x_{id}(t) + v_{id}(t+1). \end{cases} \quad (11)$$

The above adaptive strategy refers to p_i in literature [27–30] as the adaptive judgment condition. When $p_i > \delta$, the fitness value of the current particle is much larger than the average fitness value of all particles in the population, indicating that the algorithm is in the initial stage of search or the current particle distribution is relatively scattered. At this time, equation (10) should update the particle speed and to update the position. Equation (10) introduces a linear combination of individual optimization and population optimization into the speed update term, and the location update adopts "$X = wX + (1-w)V$" to improve the global search ability of the algorithm; When $p_i < \delta$, there is little difference between the fitness value of the current particle and the average fitness value of all particles in the population, indicating that the algorithm is in the middle and late stage of search or the current particle distribution is relatively concentrated. At this time, equation (11) should be used to update the particle speed and position. In equation (11), position update uses "$X = X + V$" to ensure the local exploration ability of the algorithm and prevent the algorithm from falling into local optimization when solving complex multimodal functions.

In PSO algorithm, w is particularly important. It is of great help to the global and local development of this algorithm. In an ordinary PSO, the linear decreasing method is generally used to change w. Within a certain range, this method has a positive effect on it, but when the data becomes very nonlinear and more complex, this method has little effect. Therefore, we need to find a better way to update the weight.

Heavy nonlinear adjustment strategy. This paper adopts the nonlinear change inertia weight w based on the logistic chaotic map proposed in reference [28, 31]. As a nonlinear map, chaotic map is widely used in evolutionary computation because its chaotic sequence has good randomness and spatial ergodicity. Among them, logistic map is widely used, which can produce [0, 1). Equation (12) gives the definition of logistic mapping, and the definition of w is as shown in equation (13).

$$r(t+1) = 4r(t)(1-r(t)), r(0) = rand, \tag{12}$$

$$w(t) = r(t)w_{\min} + \frac{(w_{\max} - w_{\min})t}{T_{\max}}, \tag{13}$$

where $r(t)$ is the random number generated by the iteration of equation (12). The change of inertia weight is shown in Figure 2.

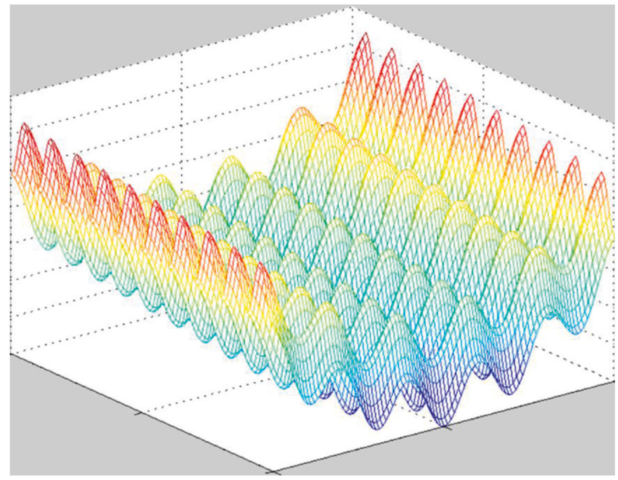


Figure 2: Variation diagram of inertia weight with iteration times.

2.3. LSTM. LSTM neural network was proposed by Hochreiter and Schmidhuber in 1997. It can measure itself and effectively make up for the shortcomings of recurrent neural network (RNN). Compared with RNN, LSTM adds memory units, including forgetting gate, updating gate, and output gate, which can use historical information. LSTM memory unit has long-term and short-term memory mechanism, which is suitable for processing data sequences with certain time intervals. The unit structure of LSTM neural network is shown in Figure 3. It mainly includes three "gate" structures and memory cells. The function of the forgetting gate is to decide which information should be discarded or retained, which determines the cell state c at the last time_ How much $t-1$ is reserved to the current time c_t; the input gate is used to update the status, which determines the input x of the network at the current time_t how much is saved to the unit state c_t; the output gate is used to determine the value of the next hidden state, control unit state c_t how many outputs are there to the current output value of LSTM h_t. The key to understanding short-term and long-term memory neural networks is the rectangular box below, which is called memory block.

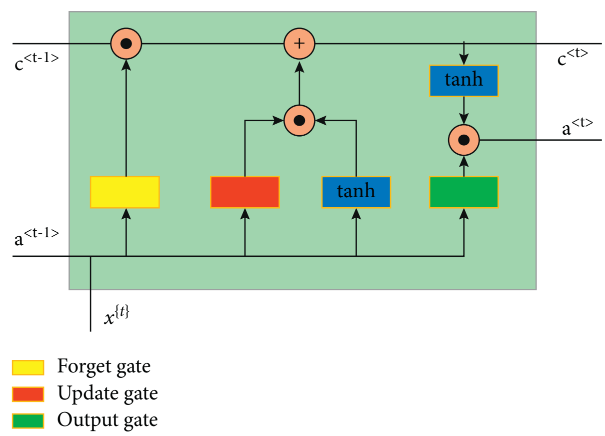


Figure 3: LSTM structure diagram.

3. Construction of Tunnel Deformation Prediction Model of PSO-LSTM

In order to improve the prediction accuracy of surrounding rock deformation of deep buried soft rock roadway in high altitude areas, and make full use of the advantages of a single model, PSO and LSTM neural networks are applied to the prediction of surrounding rock deformation of deep buried soft rock roadway in high altitude area, and a pso-lstm combined model is constructed. The combined model is divided into four stages: data acquisition and processing, PSO optimization, deformation prediction, and prediction result analysis and evaluation.

(1) The first stage is the data preprocessing stage, which refers to the component analysis and dimensionless processing of the surrounding rock deformation data of the deep buried soft rock roadway in the high altitude area by using a series of methods such as principal component analysis, so as to obtain the model input parameters and data;

(2) The second stage is the PSO optimization stage, which refers to using the data processed in the first stage to input the LSTM model, and then using PSO to determine the iteration times, learning rate, and the number of neurons in the hidden layer of the LSTM neural network;

(3) The third stage is the deformation prediction stage, which is also divided into two steps. The first step is the training of the model, that is, the on-site measured data of the roadway are classified by algorithm, one for model training and one for model testing. In this paper, the training data and test data are randomly divided according to 7 : 3;

(4) The fourth stage, the last stage of the model, is the result analysis of the data. By the above steps, the

deformation prediction of the surrounding rock of the deep buried soft rock roadway in the high altitude area is realized, and the final prediction result is obtained. Through the selection of the model evaluation index, the error analysis is carried out.

3.1. Data Acquisition and Processing Stage

3.1.1. Roadway Deformation Data Acquisition. The data set used in this study is from the actual roadway measurement. The data set from each roadway location is analyzed separately, and then those discrete and unreliable data are removed from the data set. Finally, a total of 100 data (measured in a roadway) are selected to develop the pso-lstm deformation prediction model and trial and error training model constructed in this paper. These data sets include the rock mass deformation modulus obtained by the bearing plate method, the rock mass deformation modulus (EM), uniaxial compressive strength (UCS), rock mass quality index (RQD), dry density, the number of joints per unit length (J), and porosity of the complete rock obtained by using a core with a diameter of 40 mm. All parameter data collected in this paper are shown in Table 1.

3.1.2. Data Preprocessing. In many neural network algorithms, it is assumed that the variance of each variable is of the same order. If the variance of a certain factor is particularly large, for example, the basic surrounding rock parameter value of the roadway is several times the surrounding rock density, then this feature will dominate the algorithm. It is difficult for the model to learn rules from features with small orders of magnitude. In order to solve this problem, we map the original data to the specified interval according to certain change rules. The scaling of this data set is data normalization. After standardization, the original data are transformed into dimensionless values on the interval, which facilitates the weighting and comparison of indicators with different units and orders of magnitude. In this paper, min-max standardization method is mainly used for data preprocessing. This method is to linearly change the initial data to 0 to 1, and the conversion function is shown in the following formula:

$$x' = \frac{x - \min A}{\max A - \min A}. \tag{14}$$

3.1.3. Data Feature Selection. For the deformation data of roadway surrounding rock, although through our efforts, we have screened the above six related quantities from dozens of original data. However, not all these variables are independent of each other. Therefore, we can choose some methods to simplify and reduce the dimension of variables with weak attribute significance or factors with high correlation. Pearson correlation coefficient is often used in the processing of engineering data because of its simple operation method and the universality of applicable conditions. It is also the most commonly used linear correlation coefficient at present.

$$\rho_{x,y} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu x)(Y - \mu y)]}{\sigma_X \sigma_Y}, \tag{15}$$

where Cov (X, Y) represents the covariance of the above parameters and σ_X and σ_Y represents the variance of the above parameters.

The correlation between these factors is shown in Figure 4.

In the above thermodynamic diagram, if the correlation between 0 and 0.2 represents no correlation or very weak correlation between the two variables, between 0.2 and 0.4 represents weak correlation between the two variables, between 0.4 and 0.6 represents medium correlation between the two variables, between 0.6 and 0.8 represents strong correlation between the two variables, and between 0.8 and 1 represents very strong correlation between the two factors, it can be seen that the correlation of the above factors for roadway deformation is above 0.6, so when making model prediction input, All variables should be input variables.

3.2. PSO Optimization. The detailed steps of optimizing LSTM using PSO are as follows: (1) initialize particle swarm parameters. Determine the size of the population, the number of iterations, learning factors, search dimensions, and the value range of location and speed. (2) Initialize the position and speed of particles. A particle is randomly generated, $X_i = (n, lr, h_1, h_2)$, n is the number of iterations of LSTM, lr is the learning rate, $h1$ is the number of neurons in the first hidden layer, and $H2$ is the number of neurons in the second hidden layer. The velocity $V_i = (V_{i1}, V_{i2}, V_{i3}, V_{i4})$ of particles produces a set of random sample values with uniform distribution of 0~1. The range of random samples is [0, 1). (3) Determine other parameters of LSTM. Determine the prediction scheme as one-step prediction, that is, use the data of the previous r historical periods $\{x_1, x_2, \cdots, x_n\}$ to predict the data of the next time period x_{n+1}; r value is 10. (4) Determine the fitness function of PSO algorithm. Construct LSTM with initialized particle swarm parameters, and take the mean square error between the measured value and the predicted value of the training set as the fitness function $f(r)$ of particle swarm. $f(r) = 1/N \sum (x_r - \hat{x}_r)^2$, x_r is the measured value of passenger flow, $\hat{x}_r$ is the passenger flow N is the total length of time to be predicted, $n = 210$. (5) Calculate the position of particles in each iteration and calculate the fitness value. By comparing with the fitness value of the initial position, the individual optimal position $P_{i\text{best}}$ is determined, and then the group optimal position $P_{g\text{best}}$ is determined. According to formulas (5) and (6), constantly adjust the position and speed of particles until the fitness function is minimum, and determine the optimal position, that is, determine the optimal parameters of LSTM, so as to build PSO-LSTM model.

3.3. Tunnel Deformation Prediction. PSO-LSTM model is used to predict the deformation data of roadway

Table 1: Statistical description of extracted rock mass attributes.

Data type	E_m (GPa)	UCS (MPa)	GSI	RQD (%)	Density (kg/m^3)	Porosity (%)
Average value	9.8	88.4	36.8	55.2	2.43	5.5
Standard error	0.56	2.4	0.49	0.94	0.02	0.24
Median	8.4	113.4	37.2	44.5	2.47	3.97
Sample variance	51.2	1562.7	57.8	402.3	0.09	10.2
Minimum value	2.3	7.4	16.4	12.4	2.16	0.15
Maximum value	17.5	102.6	46.8	97.5	2.51	13.5

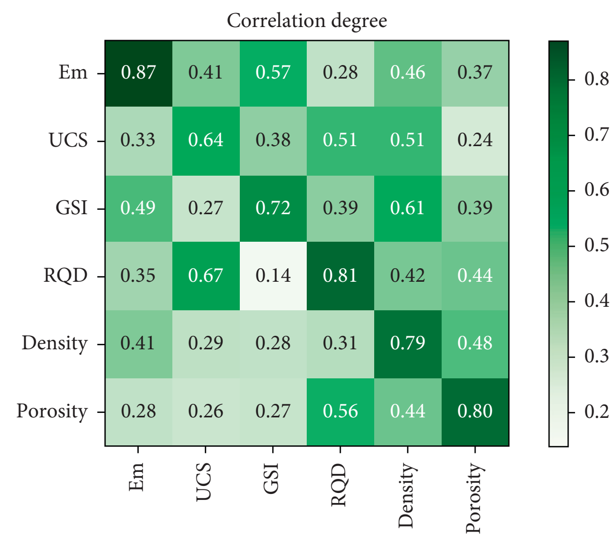


Figure 4: Thermodynamic diagram of correlation among factors.

surrounding rock, and the final prediction result is obtained. The short-term passenger flow prediction process of pso-lstm combined model is shown in Figure 5.

3.4. Prediction Result Analysis and Evaluation Stage. The R^2 value is used as the evaluation index to judge the accuracy of the deformation prediction model of deep buried soft rock roadway in high-altitude areas. The calculation formula of each evaluation index is as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{m} \left(h\left(x_i\right) - y_i\right)^2}{\sum_{i=1}^{m} \left(\mathrm{y_i} - \overline{y}\right)^2}. \quad (16)$$

In the determination coefficient expression, the denominator is the dispersion degree of the original data, that is, the total sum of squares, and the numerator is the error between the true value and the estimated value, that is, the regression square. The interference caused by data dispersion is eliminated by dividing these two parameters, which reflects the goodness of fit of the model. The value range of the determination coefficient is 0 to 1. The larger the value of the determination coefficient (close to 1), the better the simulation effect; mean absolute error is usually used to evaluate the degree of change of data, and it is also used as the loss function of linear regression; the smaller the value, the more accurate the model description; root mean square error is a measure of the deviation between the observed value and the real value. Compared with RMSE, it uses the most reliable value to replace n in the actual data, eliminating the sensitivity of extremely small and extra large values in the mean square error.

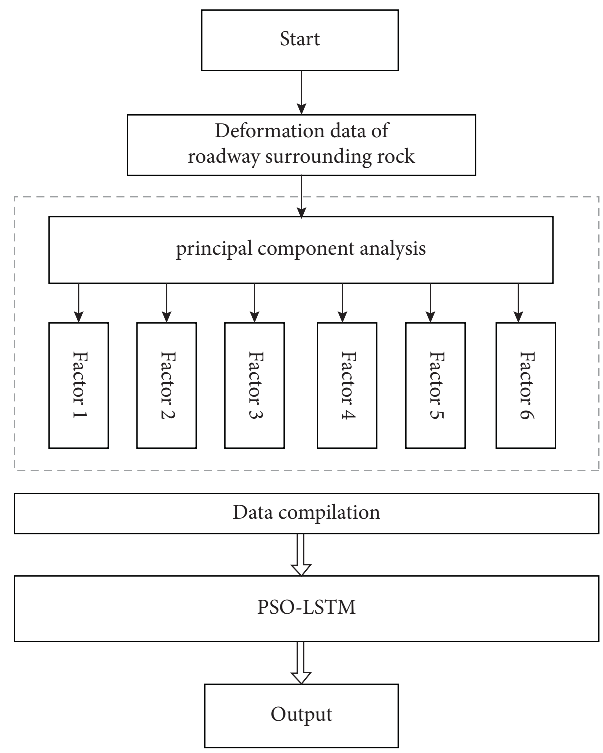


Figure 5: Short term passenger flow prediction process of PSO-LSTM combined model.

4. Engineering Examples and Result Analysis

4.1. Project Overview. In order to verify the reliability of pso-lstm model in engineering practice, this study selects the large deformation data of the excavated section of a coal mine deep roadway project as the learning sample. The structure of the 3# coal seam studied is relatively complex, and the thickness is generally in the range of 3.5~6.2 m, with an average thickness of 5.5 m. The burial depth is basically 463.9 m~633.9 m, with an average of 500 m. The coal seam structure can be mainly summarized into linear banded and layered structures. The joints of the coal seam are relatively developed and complex. The joint density of the coal seam in a specific direction is relatively large. The joint surface is generally not straight, the cracks are tight, and there is no filling. The joints in other directions extend relatively short, the joint surface is not straight enough, and the development

Table 2: Geological conditions of the stratum where the roadway is located.

Cumulative thickness (m)	Layer thickness (m)	Rock name and lithology description
517.6–525.4	7.8	The rock is gray, broken, and semi-hard
525.4–530.5	5.1	Carbonaceous mudstone, thin-layer, rich in plant-based fossils
530.5–541.3	10.8	Black, mainly bright coal, carbonaceous mudstone
541.3–544.9	3.6	The rocks are relatively complete and contain plant fossils
544.9–561.7	16.8	Black mudstone, semi-hard, horizontal texture

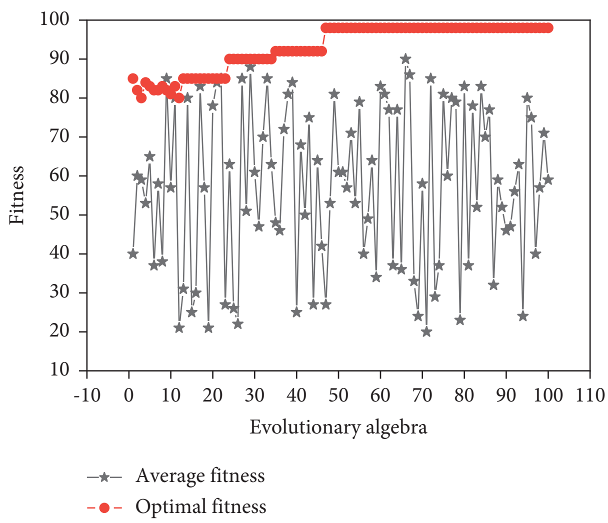


Figure 6: PSO finding the fitness curve of the best parameters.

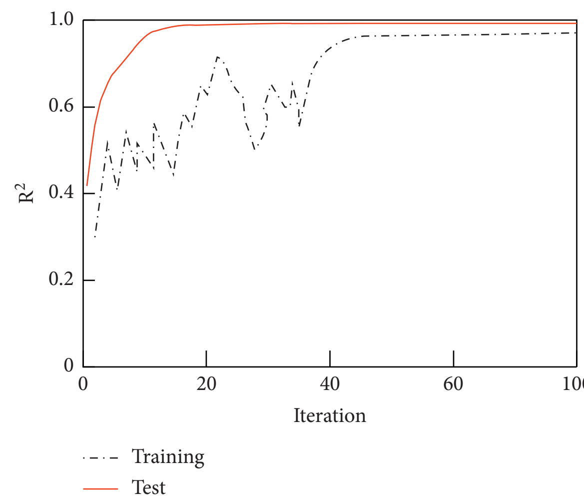


Figure 7: PSO-LSTM model training and testing R^2 value change curve.

density and standardization are not strong. During the excavation process, it is revealed that the cracks of the coal seam and roof rock are developed, and the roof is broken, which is easy to collapse and collapse, coal seams are very easy to chip. The maximum horizontal principal stress in the mining area is in the range of 10~15 MPa, mainly structural stress. The maximum horizontal principal stress is 14.83 MPa, the vertical principal stress is 12.63 MPa, and the minimum horizontal principal stress is 8.12 MPa. The roof strata of the coal roadway are mainly sandy mudstone, then siltstone and local strata are distributed with medium-grained sandstone and fine-grained sandstone. The roadway floor strata are mainly mudstone and sandy mudstone, and the distribution range of medium-sized sandstone, fine-grained sandstone, or siltstone is less. The geological conditions of the mining area where the studied coal seam roadway is located are shown in Table 2.

4.2. Establishment and Training of Improved pso-lstm Large Deformation Prediction Model. The six evaluation indexes of the obtained 100 groups of learning samples are divided into three parts after the pretreatment shown above. The pso-lstm model is trained with 2/3 learning samples. After the training, the reserved 1/3 samples are identified one by one. The parameters of the PSO algorithm are set as: acceleration $c_1 = 1.5$, $c_2 = 1.7$, termination algebra 100, and population number $p = 20$. After multiple iterations, the optimal values of LSTM parameters $c = 12$, $g = 0.31$ are obtained. The change process of parameter optimization fitness is shown in Figure 6. It can be seen from Figure 6 that the model achieves global optimization in 47 iterations.

The variation curve of R^2 value in pso-lstm model training and testing is shown in Figure 7. It can be found from the variation curve of R^2 value of pso-lstm model training and testing in Figure 7 that during the training of this model, due to the different properties of roadway surrounding rock, the fluctuation during the training is large, and the fitting degree of the data is not high. However, when the training times exceed a certain number, the R^2 value of the model tends to be stable, and always remains around 0.95; during the test, because the model is familiar with the deformation law of roadway surrounding rock, it shows good performance and degree in the test. The R^2 value tends to be stable after many times of crossing, close to 1.0, up to 0.9786, indicating that the prediction accuracy of the model for the deformation of roadway surrounding rock is high.

In addition to evaluating the accuracy of the models built in the article, we also need to compare and analyze the models. Therefore, this paper also selects two prediction models commonly used in our prediction research in the engineering field, BP and conventional LSTM, which have prominent advantages and have been well applied in various engineering fields. Similarly, in the comparative analysis, we also use R^2 value as the evaluation result of the comparative analysis model. Through modeling and putting the same data into the training test, it can be found that, different from the prediction model constructed in the above article, when comparing model 1, that is, BP model is used to predict the

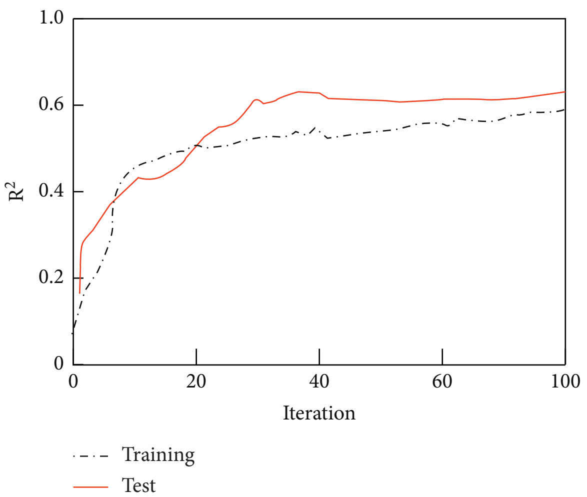


Figure 8: BP model training and testing R^2 value change curve.

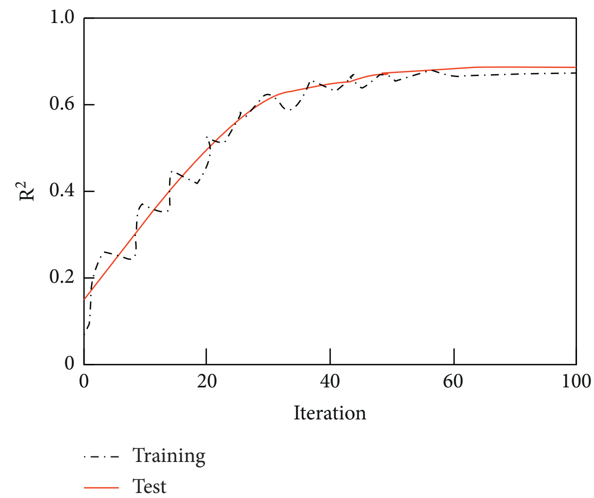


Figure 9: LSTM model training and testing R^2 value change curve.

deformation of the roadway, due to the limitation of the model itself on the processing ability of complex data, the initial value and degree of the model are low, and its initial value and degree of training are less than 0.1, which can be said to be basically not fitted, although with the increase of the number of iterations, The final fitting degree is no more than 0.65, and the maximum fitting degree is only 0.616. Therefore, the prediction accuracy of the roadway deformation comparison model 1 established in this paper is far lower than that of the model established in this paper; Similarly, for comparison model 2, LSTM, this model shows a better fitting curve than comparison model 1. The fitting degree of the test curve is also close to that of the training curve, and the two are always consistent. At the same time, its prediction accuracy is higher than the former model, and the maximum fitting degree is as high as 0.863. However, there are also similar problems with comparison model 1. First, the initial accuracy of the model is not enough, In addition, the convergence speed of the model is also slow, which is difficult to meet the actual work needs for some problems that need rapid prediction.

Based on the training and testing R2 value change curves of the three different models in Figures 7–9, it is found that, first, in the case of small sample size, compared with BP and LSTM models, pso-lstm model has higher prediction accuracy for the deformation of roadway surrounding rock, which proves that the improved pso-lstm combined prediction model constructed in this paper has a good training effect for small samples, which has a good applicability to some projects under complex environmental conditions. Second, from the perspective of the convergence of the model, the model constructed in this paper is obviously superior to the comparison model in terms of convergence speed, and the convergence curve shown by the model is also more in line with the actual process, which can eliminate the overfitting and under fitting phenomena in the process of model training, while the comparison model 2 may have overfitting phenomena. At the same time, this rapid convergence ability is very important for this high altitude, large burial depth The rapid deformation prediction of the complex geology of weak surrounding rock is very helpful. Finally, through comprehensive comparison, it is found that the model constructed in this paper is superior to other comparison models in terms of initial prediction accuracy and convergence. Its time required is shorter and the overall effect is better. It can be judged that the deformation prediction model of LSTM based on PSO optimization constructed in this paper has the advantages of higher accuracy and better generalization of large deformation prediction.

5. Conclusion

With the continuous increase of coal mining depth, the mining depth of many mines has reached thousands of meters, and the rocks show certain rheological properties. Especially for highaltitude areas, under the action of high ground stress, the ground pressure of soft rock roadway shows intense and large deformation, and the support problem of deep buried high-stress soft rock roadway is becoming more and more prominent. Therefore, studying the deformation prediction technology of deep buried soft rock roadway is of great significance to control the stability of roadway surrounding rock. In order to solve the above problems, this study proposes a PSO-LSTM large deformation prediction method for layered soft rock tunnels based on in-situ stress inversion and on-site large deformation monitoring information through machine learning method, which provides a new research idea for hierarchical prediction of large deformation of layered soft rock tunnels. At the same time, the related models are optimized to improve the prediction accuracy, and the feasibility of this method is verified by engineering application, and the following conclusions are obtained.

(1) Through the field monitoring data, it is found that the influencing factors of large deformation of layered soft rock tunnels are complex and changeable. Taking full account of the anisotropy of layered soft

rock, six evaluation indexes for large deformation prediction are selected. The large deformation data of the excavated section of the target tunnel and the field measurement information are used as the learning samples of machine learning, which is more true and accurate than the learning samples obtained from other tunnels.

(2) Using the on-site monitoring technology and combined neural network modeling technology, a combined neural network deformation prediction model of high altitude and deep buried soft rock roadway based on PSO and LSTM is constructed. Through the model training, it is found that the PSO improved by this paper on particle velocity and updating method has more advantages than the traditional PSO.

(3) Through practice, BP, LSTM, and pso-lstmm are used to predict the same sample data. The results show that the accuracy of the test set of pso-lstm model for large deformation prediction is as high as 97.86%, and its prediction results and speed are better than BP and LSTM models.

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] Y. U. Wei-Jian, Y. Yuan, and W. J. Wang, "Deformation mechanism of roadway surrounding rock with large deformation and control technology under difficult conditions," *Coal Science and Technology*, 2015.

[2] J. Wang, M. Lin, D. Tian, and Cl Zhao, "Deformation characteristics of surrounding rock of broken and soft rock roadway," *Mining Science and Technology*, vol. 19, no. 2, pp. 205–209, 2009.

[3] C. Q. Zhu, T. Feng, and S. L. Shi, "Application of neural network in optimal selection of bolt support patterns and deformation prediction of extraction roadway in fully-mechanized caving face," *Journal of China Coal Society*, 2005.

[4] Q. Xu, Y. Li, J. Lu, and L. Zhang, "The use of surrounding rock loosening circle theory combined with elastic-plastic mechanics calculation method and depth learning in roadway support," *PLoS One*, vol. 15, no. 7, Article ID e0234071, 2020.

[5] U. C. Han, C. S. Choe, K. U. Hong, and C. Pak, "Prediction of final displacement of tunnels in time-dependent rock mass based on the nonequidistant grey verhulst model," *Mathematical Problems in Engineering*, vol. 2022, Article ID 3241171, 11 pages, 2022.

[6] H. Zhang, X. Miao, G. Zhang, Y. Wu, and Y. Chen, "Non-destructive testing and pre-warning analysis on the quality of bolt support in deep roadways of mining districts," *International Journal of Mining Science and Technology*, vol. 27, no. 6, pp. 989–998, 2017.

[7] B. Meng, H. Jing, K. Chen, and H. Su, "Failure mechanism and stability control of a large section of very soft roadway surrounding rock shear slip," *International Journal of Mining Science and Technology*, vol. 23, no. 1, pp. 127–134, 2013.

[8] G. F. Zhang, K. H. Zeng, and C. Zhang, "Failure mechanism and support design of interclated roadway surrounding inclined coal seam in Qishan coal mine," *Journal of Mining & Safety Engineering*, vol. 28, no. 1, pp. 22–27, 2011.

[9] X. H. Cui, Y. F. Gao, and Q. Z. Fan, "Optimum calculating method for constant resistant supporting of roadway surrounding rock," *Chinese Journal of Geotechnical Engineering*, vol. 12, no. 1, 2006.

[10] C. Zhao, Y. Li, G. Liu, and X. Meng, "Mechanism analysis and control technology of surrounding rock failure in deep soft rock roadway," *Engineering Failure Analysis*, vol. 115, Article ID 104611, 2020.

[11] L. Wang and G. Zhao, "Analysis of instability mechanism and support technology of surrounding rock in high stress soft rock mining roadway," *Iop Conference*, vol. 199, 2018.

[12] M. A. Abido, "Optimal des'ign of power system stabilizers using particle swarm opt'imization," *IEEE Power Engineering Review*, vol. 22, no. 7, p. 53, 2002.

[13] F. Zeng, C. Zou, and L. T. University, "Numerical simulation of roadway surrounding rock deformation and destruction," *Journal of Liaoning Technical University*, vol. 32, no. 12, 2013.

[14] J. Jiang, G. Qin, and C. Liu, "Dynamic feature study of fully-mechanized caving roadway surrounding rock system," *Chinese Journal of Rock Mechanics and Engineering*, vol. 25, no. 9, pp. 1755–1764, 2006.

[15] Y. F. Xun, "Analysis on affects of in-situ horizontal stress to stability of rock surrounding roadway," *Journal of China Coal Society*, vol. 35, no. 6, pp. 891–895, 2010.

[16] Y. H. Shi and Z. Y. He, "New neural network for pattern recognition," *Electronics Letters*, vol. 27, no. 15, pp. 1353–1355, 1991.

[17] X. Du and J. Yu, "Graph neural network-based early bearing fault detection," 2022.

[18] L. Wei, S. Xv, and B. Li, "Short-term wind power prediction using an improved grey wolf optimization algorithm with back-propagation neural network," *Clean Energy*, vol. 6, no. 2, pp. 288–296, 2022.

[19] M. Anthony and P. L. Bartlett, "Neural network learning: theoretical foundations," *AI Magazine*, vol. 22, no. 2, pp. 99-100, 1999.

[20] H. Chen and R. Wang, "Artificial neural network's application in intelligent displacement back analysis of deep mine roadway surrounding rock," in *Proceedings of the 2010 International Conference on Intelligent Computation Technology and Automation*, Changsha, China, May 2010.

[21] None, B. Conference, An artificial neural network for forecasting the amount of Chinese colliery roadway surrounding rock deformation: yuxiang Zhang & 6 others Applications of computers and operations research in the minerals industries," *International Journal of Rock Mechanics & Mining ences & Geomechanics Abstracts*, vol. 33, no. 5, p. A232, 1996.

[22] Y. Shen and A. Zhang, "The stability classification system of roadway surrounding rock based on VC++ 6.0 and BP neural networks," in *Proceedings of the International Symposium on Electronic Commerc*, Guangzhou, China, August 2010.

[23] H. E. Wei, M. Hong, and X. Meng, "Inversion of rheological parameters of surrounding rocks in a mine roadway based on

BP neural network," *Journal of Mines, Metals and Fuels*, vol. 65, no. 3, pp. 149–155, 2017.

[24] S. Yu, H. Zhang, and Y. Chang, "Application of GRNN in time series prediction for deformation of surrounding rocks in soft rock roadway," in *Proceedings of the 2011 Fourth International Conference on Intelligent Computation Technology and Automation*, IEEE, Shenzhen, China, March 2011.

[25] B. Tan, H. Zhang, and G. Cheng, "Constructing a gas explosion inversion model in a straight roadway using GA-BP neural network," 2021.

[26] G. C. Chen and Y. U. Jin-Shou, "Particle SWARM OPTIMIZATION ALGorithm," *Information and Control*, vol. 186, no. 3, pp. 454–458, 2005.

[27] R. C. Eberhart, "Comparing inertia weights and constriction factors in particle swarm optimization," in *Proceedings of the 2000 IEEE Congress on Evolutionary Computation*, IEEE, La Jolla, CA, 2002.

[28] Y. H. Shi and R. C. Eberhart, "Empirical study of particle swarm optimization," in *Proceedings of the Congress on Evolutionary Computation*, IEEE, Prague, Czech Republic, July 2002.

[29] J. Robinson and Y. Rahmat-Samii, "Particle swarm optimization in electromagnetics," *IEEE Transactions on Antennas and Propagation*, vol. 52, no. 2, pp. 397–407, 2004.

[30] M. Clerc and J. Kennedy, "The particle swarm - explosion, stability, and convergence in a multidimensional complex space," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 1, pp. 58–73, 2002.

[31] S. W. Lin, K. C. Ying, S. C. Chen, and Z. J. Lee, "Particle swarm optimization for parameter determination and feature selection of support vector machines," *Expert Systems with Applications*, vol. 35, no. 4, pp. 1817–1824, 2008.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9813216, 1 page
https://doi.org/10.1155/2023/9813216

Retraction

Retracted: DA-CNN-Driven Innovative Ideological Politics Education Management System

Mobile Information Systems

Received 13 September 2023; Accepted 13 September 2023; Published 14 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] S. Yingkun, C. Liang, and C. Jing, "DA-CNN-Driven Innovative Ideological Politics Education Management System," *Mobile Information Systems*, vol. 2022, Article ID 7959144, 9 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 7959144, 9 pages
https://doi.org/10.1155/2022/7959144

Research Article

DA-CNN-Driven Innovative Ideological Politics Education Management System

Shan Yingkun[1], **Chen Liang**[2] **and Chu Jing**[3]

[1]*Business college, Jiangsu vocational college of electronics and information, JiangSu, Huaian 223003, China*
[2]*School of Education and Science, Yangzhou University, JiangSu, YangZhou 225000, China*
[3]*IFLYTEK CO.LTD, Anhui, Hefei 230093, China*

Correspondence should be addressed to Shan Yingkun; 150814@jsei.edu.cn

Received 22 July 2022; Revised 11 August 2022; Accepted 20 August 2022; Published 6 September 2022

Academic Editor: R. Mo

With the rapid development of the Internet, the network has become an important part of the study, work, and life of all teachers and students in colleges and universities. It also brings great challenges to the traditional way of ideological and political education in colleges and universities. Colleges and universities have launched various beneficial attempts in the ideological and political education into the network. In this context, in order to improve the evaluation accuracy of ideological and political classroom teaching quality in colleges and universities, this paper puts forward an evaluation method of ideological and political classroom teaching quality in colleges and universities based on DA-CNN. On the basis of constructing the evaluation index system of ideological and political classroom teaching quality in colleges and universities by analytic hierarchy process, the scores of 25 evaluation indexes affecting the ideological and political teaching quality in colleges and universities are taken as the input of DA-CNN, and the comprehensive score of ideological and political teaching quality in colleges and universities is taken as the output of DA-CNN. At the same time, an innovative ideological and political education management system based on convolutional neural network is designed and developed, and the functional modules and database of the system are planned.

1. Introduction

At present, with the emergence and wide application of various new technologies such as the Internet, cloud computing, big data, and artificial intelligence [1–9], the environment faced by ideological and political management is becoming more and more complex, which puts forward new requirements for ideological and political management and gives birth to new changes in ideological and political education and management in the new era [10–15]. Therefore, in the era of big data, the reform and innovation of ideological and political management are imminent [16–18].

The significance of building an intelligent ideological and political management system from the perspective of big data mainly lies in the following aspects.

(1) By constructing the intelligent system of ideological and political management, we can realize the interactive exchange of information, work task deployment, emergency reporting, and timely crisis control between the network ideological and political center and ideological and political workers, which can effectively solve the problems of information island and data island between different subjects of ideological and political management, improve the efficiency of ideological and political work, and strengthen the level of ideological and political work.

(2) The intelligent system of ideological and political management provides students with high-quality ideological and political resources, such as excellent courses, excellent cases, tutoring materials, learning, feedback, communication, organization, participation, and feedback of ideological and political activities. Through the construction of student-oriented applications, we can realize the full coverage of ideological and political publicity, education, and teaching and practical activities, and provide new working methods and ideas for constructing a new pattern of network education in the new era.

(3) We can make use of the existing system data resources in colleges and universities, such as learning and work, educational administration, scientific research, all-in-one card, online behavior system, and campus Wi-Fi, so as to realize the early warning and identification function of students' thought, speech, behavior, body, psychology, and consumption, and solve the problems of single tracking means, untimely, and low accuracy in the work of problem students in the traditional working methods.

(4) Building an intelligent system of ideological and political management can deeply carry out research on students' physical and mental health and provide support for improving the scientific research level of colleges and universities.

2. Analysis of Practical Problems

At present, most colleges and universities in China are faced with such problems: the school grass-roots facilities, especially the intelligent information system facilities, are relatively weak, and the problems of information island, process fragmentation, information barriers between departments, and so on are prominent. These problems can be attributed to the following aspects.

2.1. Weak Informatization Support of Ideological and Political Work. The informatization support of ideological and political work in colleges and universities is weak, and the original campus intelligent resources cannot meet the needs of current ideological and political work. There is a lack of information system support for information exchange, key work deployment, key thematic research of Ideological and political work, organization of important activities and tasks, and management and control of major emergencies. Although some colleges' information systems can play a supporting role, there are a large number of information islands and data islands, which cannot realize the accurate management functions of issuing, tracking, monitoring, analysis, and early warning of ideological and political work tasks. Therefore, there is an urgent need to build a working mechanism under the condition of big data technology to meet the current needs of ideological and political work and improve the efficiency of ideological and political work.

2.2. New Challenges of Ideological and Political Propaganda. Facing the information age, the position of ideological and political education management in colleges and universities has gradually moved from offline to online. With the development of mobile terminals and new social media, the multichannel nature of information dissemination and the uncertainty of communication content have brought new challenges to the ideological and political work. There is an urgent need to build a new ideological and political publicity position from the whole school and even a wider perspective, so as to realize the functions of publicity, education, learning, activities, communication, sharing, monitoring, early warning, control, and so on. Using new technological means to strengthen publicity and education has become a powerful means of cultural education and organizational education in colleges and universities.

2.3. Challenges Faced by Student Management. Most of China's college students come from rural areas, many of them are left behind children, and a considerable part of them are children brought out by single parent families and grandparents. In primary school, there is a great lack of emotion and many psychological hidden dangers and improper behavior left by exam-oriented education. After college, the mental health problems represented by anxiety, loneliness, compulsion, and depression show a rising trend.

The initiative, accuracy, and timeliness of traditional monitoring methods need to be improved. The construction of the intelligent system of ideological and political management in higher vocational colleges will effectively improve the scientificity, pertinence, and timeliness of ideological and political management in higher vocational colleges through the integration, processing, and analysis of data and the final utilization of data.

Therefore, with the rapid development of computer information technology, network society has become the second living space of modern people and has a profound impact on people's life with its unique communication mode and powerful function. College students are always a group of young people at the forefront of society. They are active in thinking, like to experience new technology, pursue new things, and become the backbone of the Internet. This makes their ideological culture deeply influenced by the Internet. Therefore, it also brings challenges to the ideological and political education in colleges and universities. The traditional ideological and political education cannot meet the learning requirements of students.

Based on the above analysis, it is extremely important to design and develop an innovative ideological and political education management system. The core of the system is to improve the teaching level and quality of ideological and political theory course in colleges and universities through the research of classroom teaching evaluation index and its system. This paper puts forward an evaluation method of ideological and political classroom teaching quality in colleges and universities based on DA-CNN neural network and applies this method to the design of ideological and political education management system.

3. DA-CNN Neural Network Evaluation Method

3.1. Dragonfly Algorithm. In the dragonfly algorithm (DA) [19–21], dragonfly individuals conduct foraging and optimization through five behaviors, including collision avoidance behavior, pairing behavior, aggregation behavior, foraging behavior, and enemy avoidance behavior. These individual behaviors are described in detail as follows:

The position vector update strategy of collision avoidance behavior is shown in

$$S_i = -\sum_{j=1}^{N} X - X_j. \tag{1}$$

In the formula, X is the position of the current dragonfly individual; X_j is the position of the j-th adjacent dragonfly individual; N is the number of adjoining dragonfly individuals; and the position vector update strategy pairing behavior is shown in.

$$\mathrm{A}_i = \frac{-\sum_{j=1}^{N} V_j}{N}. \tag{2}$$

In the formula, V_j is the individual velocity of the j-th adjacent dragonfly. The update strategy of the position vector of aggregation behavior is as shown in

$$C_i = \frac{-\sum_{j=1}^{N} X_j}{N} - X. \tag{3}$$

The position vector update strategy of foraging behavior is shown in

$$F_i = X^{+} - X. \tag{4}$$

In the formula, X^{+} is the location of the food source (the current optimal solution). The update strategy of the position vector of enemy avoidance behavior is shown in

$$E_i = X^{-} + X. \tag{5}$$

In the formula, the position of the natural enemy is X^{-} (the current worst solution). Synthesizing the five dragonfly group behaviors, the step size vector update strategy of the dragonfly individual is shown in

$$\Delta X_{t+1} = \left(sS_i + a\Lambda_i + cC_i + fF_i + eE_i\right) + w\Delta X_t. \tag{6}$$

In the formula, s、a、c、f、e are the weights of the five dragonfly group behaviors, respectively; w represents the inertia weight; and t is the current number of iterations, as shown in formula (7).

The dragonfly position update strategy is

$$X_{t+1} = X_t + \Delta X_{t+1}. \tag{7}$$

3.2. DA-CNN-Based Teaching Quality Evaluation

3.2.1. Convolutional Neural Network. Convolutional neural network [22–25] is a multilayer feedforward neural network composed of input, hidden, and output layers. Suppose the convolutional network's input and output dimensions are m and 1, respectively, and the number of hidden layers is p. Then, the mapping mathematical expression of the convolutional neural network is shown in

$$x_{i+1} = f\left(X_i\right) = \frac{1}{1 + \exp\left(-\sum_{j=1}^{p} c_j b_j + \varepsilon\right)} \quad j = 1, 2, \cdots, p. \tag{8}$$

In the formula, f is the activation function of the hidden layer; ε is the threshold of the output layer; and c_j and b_j are the connection weights from the hidden layer to the output layer and the output of the hidden layer nodes, respectively. Therefore, the output of the hidden layer node of the convolutional neural network can be expressed as shown in

$$b_j = \frac{1}{1 + \exp\left(-\sum_{i=1}^{m} \omega_{ij} x_i + \theta_j\right)} \quad i = 1, 2, \cdots, m. \tag{9}$$

In the formula, ω_{ij} is the connection weight from the input layer to the hidden layer and θ_j is the threshold of the hidden layer node.

Since the prediction results of the convolutional neural network are easily affected by the initial connection weights c_j, ω_{ij}, and thresholds ε、θ_j, and it is easy to fall into the problem of local extremes, this paper uses DA to optimize the initial connection weights and thresholds of convolutional neural network.

3.2.2. DA-CNN Algorithm Flow. The algorithm flow of teaching quality evaluation of ideological and political classrooms in colleges and universities based on DA-CNN can be summarized as follows.

Step 1. Initialize the convolutional neural network model and determine the network structure. Determine the number of layers, transfer function, and training function type of convolutional neural network and the number of nodes in each layer according to the data samples; read the teaching quality evaluation data of ideological and political classrooms in colleges and universities, and preprocess the data to divide the data into a training set and test set.

Step 2. Coding—the DA algorithm adopts real number coding, and the connection weight c_j、 ω_{ij} and the threshold ε, θ_j are coded as a whole. The search space dimension of the algorithm is m; if the number of nodes in the input layer, hidden layer, and output layer is R、S_1、S_2, then the encoding length S is shown in

$$S = RS + S_1 S_2 + S_1 + S_2. \tag{10}$$

Step 3. DA algorithm parameter initialization: population size N and the maximum number of iterations T.

Step 4. Randomly initialize the step size vector ΔX and randomly generate the initial position X of the dragonfly individual.

Step 5. Set the current number of iterations $t = 1$, input the training set into convolutional neural network, calculate the fitness of all dragonfly individuals according to the fitness function formula (11), and sort and record the current optimal solution.

$$\text{fitness} = \frac{1}{k} \sum_{i=1}^{k} \left(y_i - \widehat{y}_i\right)^2. \tag{11}$$

The mean square error is chosen as the fitness function, shown in formula (11). In the formula, y_i、 $\widehat{y}_i$ is the actual

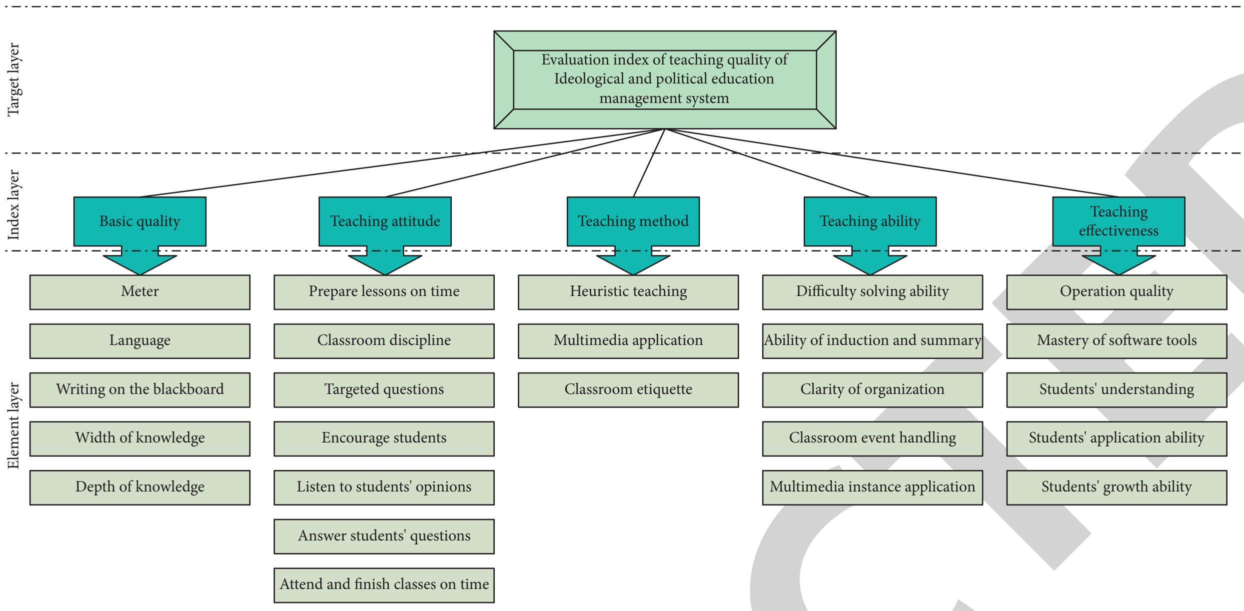


FIGURE 1: AHP model of teaching quality in ideological and political classrooms in universities.

output and expected output of the ith sample, respectively; k is the number of samples.

Step 6. Update the food source position X^+ (the current optimal solution) and the natural enemy position X^- (current worst solution), and update 5 behavioral weights s、a、c、f、e, and inertia weight w.

Step 7. Update S、A、C、E、F according to formulas (3)–(7).

Step 8. Update the step vector and position vector according to formulas (8) and (9).

Step 9. If the number of iterations $t > T$, save the optimal connection weights c_j、ω_{ij}, and thresholds ε、θ_j; otherwise, $t = t + 1$, and return to Step 5.

Step 10. Take the connection weights c_j、ω_{ij}, and thresholds ε、θ_j corresponding to the optimal solution as the initial connection weights and thresholds of the convolutional neural network, train the convolutional neural network, and make predictions.

4. Experimental Analysis

4.1. Quality Evaluation System. Combined with reference literature and teaching experience, the AHP structure model of ideological and political classroom teaching quality evaluation in colleges and universities is constructed using the analytic hierarchy process, as shown in Figure 1. It mainly includes three layers: the target layer, the criterion layer, and the element layer.

Selecting the ideological and political teaching quality data of a Project 211 university from 2008 to 2017 as the research object, this paper uses the maximum value method to standardize the data. The two comparison methods, the constructed physical education quality evaluation index scores, are shown in Table 1, and the evaluation scores are shown in Table 2.

4.2. Evaluation Index. The root means square error (RMSE) and the correlation coefficient R are selected to test the quality evaluation results of ideological and political classroom teaching in colleges and universities. The evaluation index formulas are

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{k=1}^{n}\left(x_k - \hat{x}_k\right)^2}, \tag{12}$$

$$R = \frac{\sum_{k=1}^{n} x_k \hat{x}_k}{\sqrt{\sum_{k=1}^{n} x_k^2}\sqrt{\sum_{k=1}^{n} \hat{x}_k^2}}. \tag{13}$$

Among them, x_k and $\hat{x}_k$ represent the actual value and predicted value of the k-th sample, respectively; n represents the number of sample sets; *RMSE* is mainly used to measure the degree of dispersion of the model; and R is primarily used to describe the degree of correlation between the predicted value and the actual value. The closer the absolute value of R is to 1, the higher the correlation between the predicted value and the actual value.

4.3. Result Analysis. The teaching quality of ideological and political classrooms can be divided into 5 grades, namely, excellent, good, average, poor, and very poor, and the evaluation grades are shown in Table 3. Experts scored the collected data, and a total of 10 sets of data were obtained.

The data are divided into two parts; the first 6 groups of data are used as training sets to establish the DA-CNN

Table 1: Scores of evaluation indicators of ideological and political teaching quality.

Indicator serial number	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
1	0.0274	0.0359	0.0475	0.0604	0.062	0.0639	0.0993	0.1094	0.1372	0.1498
2	0.1118	0.1209	0.1465	0.1683	0.1457	0.13	0.1764	0.1683	0.2029	0.2102
3	0.1677	0.1784	0.1817	0.1861	0.1916	0.1997	0.1996	0.2463	0.2562	0.2895
4	0.1591	0.1842	0.1923	0.1968	0.2013	0.2033	0.2441	0.2527	0.2582	0.2592
5	0.2213	0.2424	0.2176	0.1912	0.1683	0.1583	0.1591	0.1601	0.2797	0.2899
6	0.0456	0.0476	0.0505	0.0719	0.099	0.1059	0.1282	0.1642	0.1816	0.1885
7	0.0382	0.0401	0.0449	0.0582	0.0827	0.1127	0.1409	0.1629	0.1779	0.2211
8	0.0533	0.0547	0.0512	0.0678	0.079	0.0907	0.1003	0.1017	0.1103	0.1151
9	0.0357	0.054	0.0762	0.0731	0.0682	0.0787	0.0946	0.0973	0.0917	0.1054
...	...	...	...	...	...	...	...	...	...	...
24	0.062	0.0769	0.0626	0.0605	0.0631	0.0605	0.0605	0.0854	0.1136	0.1476
25	0.1089	0.1303	0.1204	0.1019	0.0416	0.0458	0.0591	0.07	0.0883	0.1129
26	0.0118	0.014	0.0174	0.0258	0.0379	0.0404	0.0538	0.0674	0.0899	0.1178

Table 2: Final score of ideological and political teaching quality evaluation.

Year	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Score P	1.5444	2.5096	3.5163	3.5209	4.5408	4.617	4.6947	4.8067	4.8852	4.9959

Table 3: Evaluation level divide.

Corresponding score	Evaluation results
[4.2,5)	Very good
[3.4,4.2)	Better
[2.6,3.4)	General
[1.8,2.6)	Poor
[1,1.8)	Very bad

Table 4: Comparison of evaluation results of different algorithms.

Method	Training set		Test set	
	RMSE	R	RMSE	R
DA-CNN	0.0101	0.9968	0.0385	0.9905
GA-CNN [26]	0.0174	0.9798	0.0404	0.9723
PSO–CNN [27]	0.0180	0.9775	0.0507	0.9654
CNN [28, 29]	0.0188	0.9664	0.0484	0.9468

ideological and political classroom teaching evaluation model; the last 4 groups of data are used as test sets to test the DA-CNN ideological and political classroom teaching evaluation model correctness. DA algorithm parameters are dragonfly population size $N = 10$ and the maximum number of iterations $T = 100$.

To verify the accuracy and validity of the DA-CNN model, DA-CNN is compared with GA-CNN, PSO-CNN, and CNN; particle swarm optimization (PSO) algorithm parameters are maximum number of iterations $T = 100$, population size $N = 10$, learning factor $c_1 = c_2 = 2$, and search interval [-1, 1]. Genetic algorithm (GA) parameters are population size $N = 10$, the maximum number of iterations $T = 100$, crossover probability pc = 0.7, and mutation probability pm = 0.1; convolutional neural network parameters are set as follows: input layer. The number of nodes inputnum = 25, the number of hidden layer nodes hiddennum = 50, the number of output layer nodes outputnum = 1, the maximum number of convolutional neural network training is 1 000, the hidden layer and output layer transfer functions are logsig and purelin, the training function is trainlm, the learning rate is 0.01, and the training error target is 0.001.

The comparison results of different algorithms are shown in Table 4.

It can be seen from Table 4 that (1) from the overall ideological and political classroom teaching evaluation results, the evaluation results of DA-CNN are better than GA-CNN, PSO-CNN, and CNN. On the training set and test set, DA-CNN has the smallest RMSE, and the correlation coefficient R reaches the maximum, indicating that the DA-CNN model's ideological and political classroom teaching evaluation value and the actual value of ideological and political classroom teaching evaluation have the highest correlation, and the prediction effect is the best; (2) the evaluation accuracy of DA-CNN, GA-CNN, and PSO-CNN is better than CNN. The main reason is that the swarm intelligence algorithms DA, GA, and PSO optimize the CNN model's parameters and improve the CNN model's evaluation accuracy.

4.4. Discussion and Suggestions. Through the DA-CNN course evaluation, it is of great help to improve the teaching quality of ideological and political classrooms in colleges and universities, so the following countermeasures and suggestions are put forward:

(1) Strengthen the attention of university leaders and increase supervision.

(2) To innovate and improve teaching models and teaching methods.

(3) Fully motivate the main body of students.

(4) Combined with the actual situation, strengthen student guidance.

(5) Clarify the purpose, evaluation index, and teaching evaluation system.

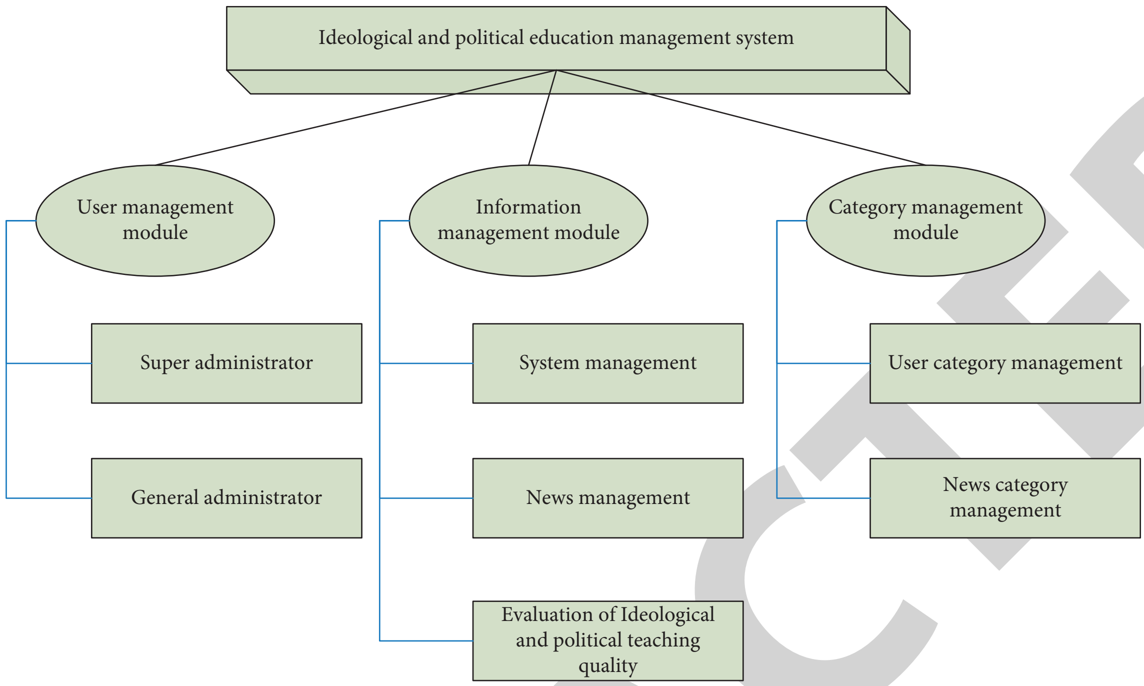


Figure 2: Function module of the designed innovative ideological and political education management system.

(6) Strengthen case teaching and improve students' comprehension and perception.

5. System Design

5.1. Overall Design. The designed innovative ideological and political education management system takes Eclipse as the development platform, MySQL as the background database, and Java as the language to construct the application. Because the system is mainly for university teachers and students, the system adopts B/S architecture. B/S architecture has the following advantages: (1) the architecture has a wider range of applications and can give full play to the advantages of the Internet; (2) the system with this architecture has very good security and high performance: (3) the system with this architecture has fast response speed and transparent data storage and management function.

The designed innovative ideological and political education management system is divided into three modules: user management module, information management module, and category management module. Super-administrator and ordinary administrator have different administrative rights. The function is shown in Figure 2.

Among them, the user management module has the following functions: (1) the super-administrator manages the registration information of their own users and provides functions such as new creation, modification, deletion, and query; (2) ordinary administrators manage their own accounts and provide functions such as new creation, modification, and deletion. The following requirements are considered in the user management module: (1) ordinary administrators add, delete, modify, and query their own accounts; (2) super-administrators have all the functions of ordinary administrators and can also view and manage all management information and add, delete, modify, and query operations.

Table 5: User information table.

Field	Name	Type	Primary key	Is it empty
Username	User name	varchar(20)	No	No
Password	Password	varchar(20)	No	No
Real name	Real name	varchar(10)	No	No
ID	Identifier	int(20)	Yes	No

5.2. Database Design. The system uses MySQL database. It mainly uses five data tables: user information table (tb_users), news information table (tb_articles), role information table (tb_roles), system column table (tb_columns), and news category table (tb_subcolumns).

User information table (tb_users table) is a table used by users (including ordinary administrators and super-administrators) when logging in. It mainly includes user name, user password, real name, number, and other fields. The user ID is the primary key and also the foreign key, which is associated with other tables through this field. Table 5 describes the details of the user information table.

The news information table records all news articles of the website. The fields in the table mainly include number, title, article type, information source, release time, publisher, picture address, etc. This table takes the ID number as the primary key and is also a foreign key. It is associated with other tables through this field. Table 6 describes the details of the news information.

The role information table records the role information of all users in the system. This table has only two fields: number and role name. The number is the primary key and

Table 6: News information.

Field	Name	Type	Primary key	Is it empty
ID	Identifier	int(20)	Yes	No
Title	Title	varchar(20)	No	No
Content	Content	longtext	No	No
info_type	Type of information	varchar(10)	No	Yes
info_from	Information sources	varchar(10)	No	Yes
info_time	Information release time	varchar(20)	No	No
Author	Author of information	varchar(10)	No	Yes
img_url	URL address of image in information	varchar(255)	No	Yes

Table 7: Role information.

Field	Name	Type	Primary key	Is it empty
ID	Identifier	int(20)	Yes	No
Role	Name of role	varchar(50)	No	No

Table 8: System column.

Field	Name	Type	Primary key	Is it empty
ID	Identifier	int(20)	Yes	No
col_name	Column name	varchar(40)	No	No

Table 9: News category.

Field	Name	Type	Primary key	Is it empty
ID	Identifier	int(20)	Yes	No
news_name	Name of news	varchar(40)	No	No
news_type	Type of news	varchar(10)	No	No
column_ID	ID of the column to which it belongs	int(20)	No	No

also the foreign key. This field is used to associate with other tables. Table 7 describes the details of the role information table.

The system column table records all column names in the website. This table contains two fields: column number and column name. The column number is the primary key and also the foreign key. This field is used to associate with other tables. Table 8 describes the details of the system column table.

The news category table records all types of information of news and articles in the website. This table mainly includes four fields: number, column name, column category, and column ID. The number is the primary key and also the foreign key. Through this field, it is associated with other tables. Table 9 describes the details of the news category table.

5.3. System Composition. The designed innovative ideological and political education management system is based on improving the ideological and political management level of colleges and universities, integrating the existing smart campus information resources, relying on the Internet, big data, and other modern information technologies, exploring the construction of an ideological and political portal system, an ideological and political teaching and training system, an ideological and political activity system, and a student behavior monitoring and tracking system, so as to realize the ideological and political work management, the business platforms of superior departments, and the data docking between colleges and universities. Promote the digitization, intelligence, convenience, systematization, and institutionalization of ideological and political management in colleges and universities.

5.4. System Construction. We make use of the big data platform's data collection and exchange capabilities, storage and analysis capabilities, intelligent computing, open sharing capabilities, and data visualization and analysis capabilities to realize the establishment of the ideological and political big data resource base and integrate it into several subject resource bases according to the business type to provide data support for the information collection, management services, and decision support system of ideological and political work. The construction of intelligent system for ideological and political management in colleges and universities mainly covers data collection and exchange system, big data storage and analysis system, big data open sharing system, and data visualization tools.

6. Conclusions

The introduction of intelligent ideological and political management system is the inevitable result of the development of big data, intelligent technology, and mobile

Internet technology. In order to promote the process of informatization and networking of ideological and political education, this paper designs an innovative ideological and political education management system based on convolutional neural network. Aiming at the influence of parameter selection on the evaluation accuracy of convolutional neural network (CNN), this paper applies the dragonfly algorithm (DA) algorithm to the parameter optimization of initial connection weight and threshold of convolutional model, and puts forward an evaluation method of ideological and political classroom teaching quality in colleges and universities based on DA-CNN. Then, this method is applied to the design of ideological and political education management system, and an ideological and political education management system based on convolutional neural network is developed and implemented.

The results show that compared with GA-CNN, PSO-CNN, and CNN, DA-CNN can effectively improve the evaluation accuracy of ideological and political classroom teaching quality in colleges and universities. The innovative ideological and political education management system based on convolutional neural network provides a new method and way for the quality evaluation of ideological and political classroom teaching in colleges and universities.

Data Availability

The dataset can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] S. Liu and Y. Xu, "Exploring innovative development strategies for safety education in the context of big data and the Internet of things," *Wireless Communications and Mobile Computing*, vol. 2022, pp. 1–8, Article ID 8022354, 2022.

[2] L. Li, T. Qu, Y. Liu et al., "Sustainability assessment of intelligent manufacturing supported by digital twin," *IEEE Access*, vol. 8, pp. 174988–175008, 2020.

[3] H. Tian, "Interactive music instructional mode based on cloud computing," *Wireless Communications and Mobile Computing*, vol. 2022, pp. 1–9, Article ID 7493417, 2022.

[4] L. Li and C. Mao, "Big data supported PSS evaluation decision in service-oriented manufacturing," *IEEE Access*, vol. 8, no. 99, pp. 154663–154670, 2020.

[5] N. O. Ogwara, K. Petrova, M. L. Yang, and L. Tan, "Towards the development of a cloud computing intrusion detection framework using an ensemble hybrid feature selection approach," *Journal of Computer Networks and Communications*, vol. 2022, pp. 1–16, Article ID 5988567, 2022.

[6] L. Li, C. Mao, H. Sun, Y. Yuan, and B. Lei, "Digital twin driven green performance evaluation methodology of intelligent manufacturing: hybrid model based on fuzzy rough-sets AHP, multistage weight synthesis, and PROMETHEE II," *Complexity*, vol. 2020, no. 6, pp. 1–24, Article ID 3853925, 2020.

[7] R. Ma, S. W. Kareem, A. Kalra, R. I. Doewes, P. Kumar, and S. Miah, "Optimization of electric automation control model based on artificial intelligence algorithm," *Wireless Communications and Mobile Computing*, vol. 2022, pp. 1–9, Article ID 7762493, 2022.

[8] L. Li, B. Lei, and C. Mao, "Digital twin in smart manufacturing," *Journal of Industrial Information Integration*, vol. 26, no. 9, Article ID 100289, 2022.

[9] Y. Lai, "Application of the artificial intelligence algorithm in the automatic segmentation of Mandarin dialect accent," *Mobile Information Systems*, vol. 2022, Article ID 5116280, 23 pages, 2022.

[10] Y. Li and He Mao, "Study on machine learning applications in ideological and political education under the background of big data," *Scientific Programming*, vol. 2022, Article ID 3317876, 1 page, 2022.

[11] T. Tao and X. Lv, "Construction of ideological and political education in colleges and universities based on the carrier of smartphone," *Security and Communication Networks*, vol. 2022, Article ID 7214844, 2 pages, 2022.

[12] G. Zeng, "Analysis of learning ability of ideological and political course based on BP neural network and improved k -means cluster Algorithm," *Journal of Sensors*, vol. 2022, Article ID 8114971, 2022.

[13] R. Hu and J. Hu, "Construction and analysis of college students' entrepreneurship guidance model from the perspective of ideological and political education under big data," *Mobile Information Systems*, vol. 2022, Article ID 8954317, 23 pages, 2022.

[14] L. Zhang and J. He, "Optimization of ideological and political education under the epidemic via mobile learning auxiliary platform in the era of digitization," *Wireless Communications and Mobile Computing*, vol. 2022, pp. 1–9, Article ID 6149995, 2022.

[15] Bo Gao, "Application of convolutional neural network in emotion recognition of ideological and political teachers in colleges and universities," *Scientific Programming*, vol. 2022, Article ID 4667677, 2022.

[16] X. Zhang, X. Gao, H. Yi, and Z. Li, "Design of an intelligent virtual classroom platform for ideological and political education based on the mobile terminal APP mode of the Internet of things," *Mathematical Problems in Engineering*, vol. 2021, Article ID 9914790, 12 pages, 2021.

[17] J. Zhou, Z. Wei, F. Jia, and W. Li, "Course ideological and political teaching platform based on the fusion of multiple data and information in an intelligent environment," *Journal of Sensors*, vol. 2021, Article ID 1558360, 10 pages, 2021.

[18] Y. Jiao and Y. Liu, "The The Teaching Optimization Algorithm Mode of Integrating Mobile Cloud Teaching into Ideological and Political Courses under the Internet Thinking Modeeaching optimization algorithm mode of integrating mobile cloud teaching into ideological and political courses under the Internet thinking mode," *Scientific Programming*, vol. 2021, Article ID 6492009, 8 pages, 2021.

[19] F. Liu, "Path Path Recognition of the Regional Education Expansion Based on Improved Dragonfly Algorithmecognition of the regional education expansion based on improved dragonfly algorithm," *Mathematical Problems in Engineering*, vol. 2021, Article ID 9928020, 5 pages, 2021.

[20] S. Hameed, Q.-A. Minhas, S. Ahmad et al., "Connectivity of drones in FANETs using biologically inspired dragonfly algorithm (DA) through machine learning," *Wireless Communications and Mobile Computing*, vol. 2022, pp. 1–11, Article ID 5432023, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9782353, 1 page
https://doi.org/10.1155/2023/9782353

Retraction

Retracted: The Implementation Path of Labor Education in Applied Universities Driven by Artificial Intelligence Technology

Mobile Information Systems

Received 13 September 2023; Accepted 13 September 2023; Published 14 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] Y. Luo, "The Implementation Path of Labor Education in Applied Universities Driven by Artificial Intelligence Technology," *Mobile Information Systems*, vol. 2022, Article ID 5375449, 12 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 5375449, 12 pages
https://doi.org/10.1155/2022/5375449

Research Article

The Implementation Path of Labor Education in Applied Universities Driven by Artificial Intelligence Technology

Yingying Luo

Students Affairs Department of Sichuan University of Arts and Science, Dazhou (635000), Sichuan Province, China

Correspondence should be addressed to Yingying Luo; 15040140142@xs.hnit.edu.cn

Received 11 July 2022; Revised 27 July 2022; Accepted 1 August 2022; Published 5 September 2022

Academic Editor: R. Mo

Labor education is an endogenous demand for the construction and development of applied colleges and universities, and it is of great significance for the times to promote the construction of applied colleges and universities. The labor spirit shaping of college students in applied colleges and universities is conducive to helping college students grow up to be the new man of the times, promoting the core task of transformation and development of applied colleges and universities and meeting the construction needs of a modern educational power with the socialism of Chinese characteristics. In view of the outstanding problems of cognitive bias, inaccurate positioning, one-sided means, and disregard of results in the process of shaping labor spirit in applied colleges and universities, the university should grasp the three key words of local, applied, and open system construction in the stage of spirit generation, use specific tactics to promote the combination of educators, educated and educational influence in the stage of spirit strengthening, and form the overall style and action of interpreting labor spirit by the struggle in the stage of spirit sublimation. In the stage of spiritual sublimation, the overall style and action consciousness of interpreting the spirit of labor by struggle are formed. Cultivating students' technical application ability, nurturing quality spirit, and cultivating students' sense of professional identity and responsibility are the value demands of labor education in applied colleges and universities. In view of the problems of the implementation path of labor education in applied colleges and universities, a labor education status assessment model based on artificial intelligence technology is proposed for establishing the value goal of talent cultivation in applied colleges and universities as the guide and promoting the effective labor education in applied colleges and universities in the new era by continuously solving various problems in labor education and cultivating qualified socialist builders and successors.

1. Introduction

Labor education is crucial to the growth of human beings. Cultivating students to establish a correct concept of labor, mastering labor skills compatible with the development of the country and society, and developing good labor habits have always been an important part of China's educational work, as well as an important part of promoting students' overall development in moral, intellectual, physical, social, and aesthetic aspects. As the anchor point of students' literacy education, labor has a key role in feeding moral, intellectual, physical, and aesthetic development. The core of labor education lies in value recognition, i.e., value education, and the spirit of labor itself is a value existence, so shaping the spirit of labor is naturally at the center of labor value education [1–3]. Based on this, from the macro level, strengthening the cultivation of labor spirit among college students in applied colleges and universities is a key move to strengthen the education of labor spirit, which is of pioneering significance to promote college students to enhance the consciousness of revering and respecting labor. It is also conducive to the realization of the goals of labor for moral development, labor for education, labor for physical strength, and labor for beauty. At the micro level, strengthening the labor spirit is conducive to improving the education level of applied higher education and the personal growth of students, which is related to the progress of regional society and the long-term development of the country. The implementation path of labor education in applied colleges and universities is shown in Figure 1.

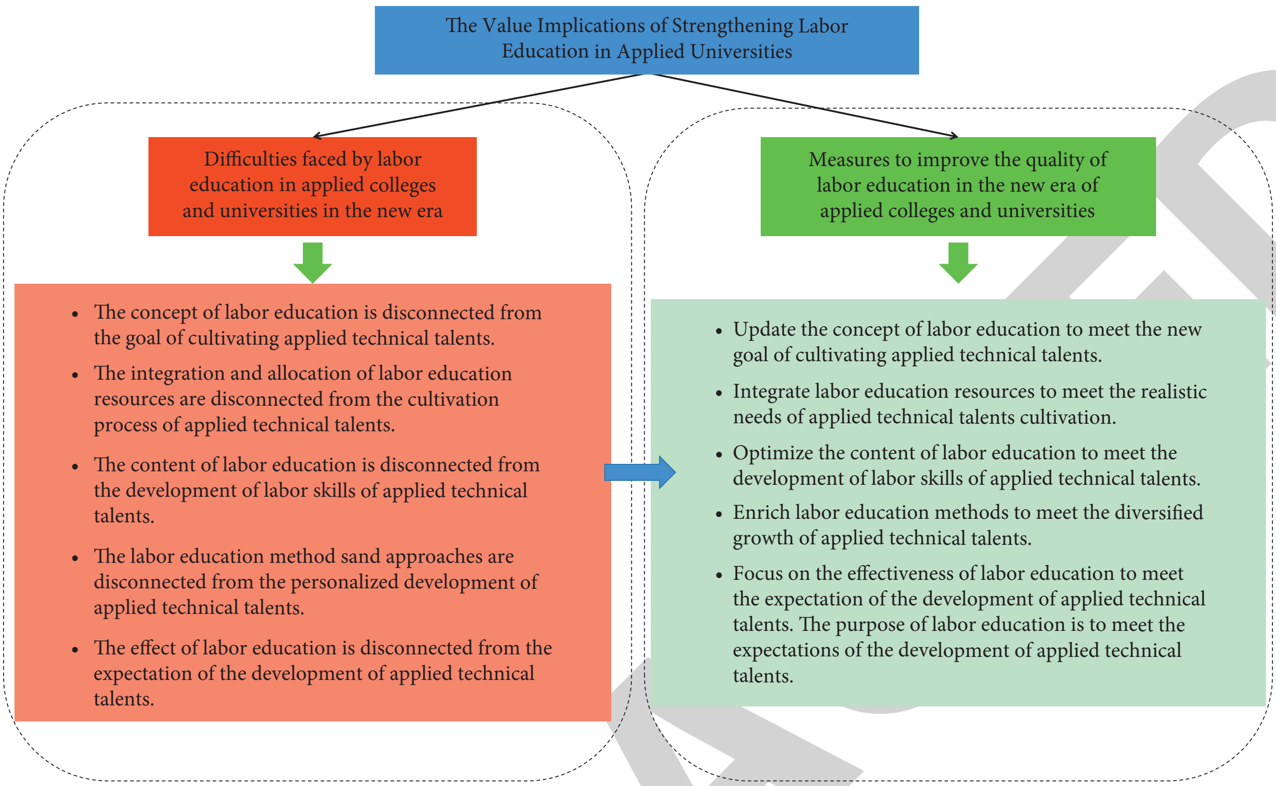


Figure 1: The implementation path of labor education in applied colleges and universities.

Strengthening the labor spirit of college students is conducive to improving the precise teaching level of higher education, which is in line with the strategic requirements of a strong education country [4–7]. It is conducive to helping college students of the new era grow up to be the new man of the times who is brave enough to accept the great responsibility, with the focus on cultivating the new man of the times who takes up the great responsibility of national rejuvenation. To grow into a new man of the times, the core of the new college students of the new era lies in taking up the great responsibility of national rejuvenation, and the fundamental lies in considering their actual situation and aligning with the standards. To accept the great responsibility of national rejuvenation, the new-age college students are required to have firm ideals and beliefs, cultivate strong skills and abilities and practical qualities, and embody the spirit of bearing for the country and the people. Labor is the source of all happiness. The spirit of labor, which is grounded in one's job and down-to-earth, is the concentrated embodiment of firm ideals and beliefs and the spirit of commitment. Therefore, from the core of cultivating the new man of the times, it is necessary to strengthen the shaping of the labor spirit of college students [8].

The new era needs to cultivate talents who can keep pace with the times and meet the development and social needs of the times, and the labor spirit of the new era, with its undertones of struggle, lean and innovation, reflects the characteristic requirements of labor itself in the new historical stage, and is conducive to solving the problem of what kind of people to cultivate. People create history and labor creates the future, and the process of college students' participation in labor in the new era is also the process of their growth into the new man of the times. In conclusion, the labor spirit has positive significance at both general and sub-levels to help college students grow into new men of the times who are brave enough to accept great responsibilities. It is conducive to promoting the core task of transformation development of applied colleges and universities. In order to improve the quality of graduates, we should reduce the employment difficulties, so as to provide application and compound and innovative talents for industrial development. College graduates are increasingly becoming professional and application-oriented direction under the guidance of national macro policy documents. The main tasks of transformation and development of local general undergraduate colleges and universities are carefully analyzed, and each specific task revolves around technical skills, industrial innovation, and other keywords, indicating that while basic skill-based labor is valued, high-level innovative labor should be paid more attention [9].

In this context, strengthening the shaping of labor spirit of college students in applied colleges and universities and vigorously promoting labor education is conducive to two-way efforts to improve practical working ability and industrial docking level and connote pragmatic skills. On the one hand, the labor spirit directly points to concrete labor, which makes the education work dovetail with the actual job, promotes students' work skills with stronger practicality and economic value, and maintains personal interests in employment, career selection, and entrepreneurship. On the other hand, the struggle connotation contained in the labor spirit will make college students realize more value creation possibilities in labor, give more added value to general labor [10–13]. On the other hand, the spirit of labor contains the connotation of struggle, which will make college students realize more value creation in labor, give more added value to general labor, promote the further transformation of concrete labor to abstract labor and realize the transformation of labor to labor. In the interactive collision and consistency of labor force and labor mind, the application characteristics and industrial innovation characteristics of applied colleges and universities will be more prominent, and the task of transformation and development will be better realized.

Artificial intelligence technology-driven labor education implementation path is a technology that uses computer vision technology to extract features from labor education images across cameras for labor education matching, which is widely used in smart security, smart retail, and other fields, and is an important fulcrum for the combination of artificial intelligence technology and industry. In the early research on AI technology-driven labor education implementation paths, a large amount of work was based on traditional computer vision methods, using manually extracted image features into labor education work intelligence technology-driven labor education implementation paths. In recent years, with the successful application of deep learning in many fields, deep learning has been applied to artificial intelligence technology-driven labor education implementation paths and has achieved certain results. The main research idea is to use deep learning methods for feature vector extraction, then use metric learning to discriminate feature vectors, quantify the differences between images, and use a large amount of labeled data for training iterations of the model [14].

The implementation path of labor education driven by artificial intelligence technology is converted into a classification problem for research, and the ID of labor education is used as a label to label the data, and an image classification method is used to classify labor education. In subsequent studies, some scholars inspired by image retrieval methods proposed the CIR method, which uses image pairs as input and finally outputs the similarity between them. The method is based on twin convolutional neural networks, in which two images are input to two independent sub-networks to obtain the feature maps of the images, and then determine whether the image pairs belong to the same labor education by the metric learning method. Applying the twin convolutional neural network to the field of artificial intelligence technology-driven labor education implementation path, the original network of sharing weights is discarded, and the two sub-networks are made independent of each other, considering the difference in background, lighting, and other features of artificial intelligence technology-driven labor education implementation path image data. Based on twin convolutional neural networks, a deep network is proposed for AI technology-driven labor education implementation path, which takes image pairs as network inputs, calculates image feature map differences, and finally determines whether the image pairs belong to the same category. Person-Net deepens the network using multiple 3×3 convolutional layers on the original structure, which makes the effect significantly improved. the CIR method emphasizes the input interrelationship of two images.

The main contributions of this study are as follows. Firstly, applied colleges and universities as an important part of Chinese higher education, are the main position to implement, the combination of education and labor, the basic education policy, and shoulder the noble mission of cultivating high-quality workers and technical skill talents. Therefore, applied colleges and universities should carry out labor education from the elements of cultivating high-quality workers' literacy, knowledge, and ability, so as to cultivate every young student to become advanced applied composite talents with strong adaptability and competitiveness to meet the needs of social and economic development, and finally promote the remarkable improvement of talent cultivation quality of applied colleges and universities, so as to respond to the expectations of the country, society, and high-quality people applied for colleges and universities. In this study, an analysis model of the implementation path of labor education in applied colleges and universities based on an artificial intelligence algorithm is proposed, and the accurate identification of mental state in labor education is achieved through the method of characterization learning is extracting features, combined with extracting local features, and introducing attention mechanism, and the relevant experiments prove the effectiveness of the method.

2. Related Work

2.1. Problems of Labor Education in Applied Colleges and Universities. Application-oriented colleges and universities, as the name implies, are undergraduate colleges and universities with application-oriented rather than research-oriented orientation, focusing on application, which is an important initiative of China to promote the transformation and development of colleges and universities [15–17]. At present, although many local colleges and universities have been transformed into application-oriented colleges and universities according to the requirements, they have not been fully implemented in many aspects such as school philosophy, talent cultivation, faculty construction, scientific and technological innovation, and docking with local enterprises, etc. Some colleges and universities still stay in the old mode, and the progress of transformation is slow. The practical ability of graduates has not been significantly improved, and the current situation of difficult employment and low employment quality has not been effectively improved. There is a deviation in the understanding of the value of labor education, which has been neglected and underestimated to some extent due to the long-term dilution and weakening of labor education.

Although most college students can recognize the importance of labor, they lack the consciousness to practice it. A small number of college students are influenced by the bad social culture, they are greedy for pleasure, good for leisure and bad for work, advocate consumerism, despise laborers, hope to get rich overnight, make quick money and get something without working. Some students mistakenly believe that the purpose of studying in university is to learn professional knowledge and develop the professional ability, not to work, and they are resistant to labor education and do not want to refine their body and mind and sharpen their will through labor, so they just get by. Some teachers also have some deviations in their understanding of the value of labor education, believing that the main function of labor education in college is to cultivate and improve students' professional and technical abilities, and that physical labor and social service are meaningless and a waste of time for college students, and that they lack the value of labor education to lead students [18, 19].

Although the teaching contents of labor education in some universities are in various forms, such as internship training, engineering practice, professional service and job practice, there is no overall planning in the top-level design of labor education, the contents are scattered and not concentrated, and no synergy is formed, and the education of strengthening the Marxist concept of labor is not organically combined with the knowledge of labor science, daily labor habits, and production labor practice. As a result, it is difficult to achieve the overall goal of labor education. At the same time, schools lack a perfect evaluation system and incentive mechanism for labor education, which makes some students lack proper attention and enthusiasm for labor education and teachers lack the enthusiasm to engage in labor education.

Lack of a professional teacher team overall, labor education teachers in colleges and universities are weak, with unreasonable structure and uneven quality, and the construction of a talent team needs to be strengthened [19–21]. At present, most colleges and universities are not equipped with full-time teachers for labor education, and often rely on personnel engaged in ideological and political education and student management, such as counselors or other professional teachers, to teach part-time. In addition, due to the lack of necessary training for teachers, the effectiveness of labor education is reduced.

As for how to carry out labor education in general colleges and universities, the Guidance Outline clearly points out that it should focus on innovation and entrepreneurship, carry out production and service labor in conjunction with academic disciplines, and improve the ability to find problems and solve them creatively in production practice. At present, some colleges and universities only position labor education in the traditional nature of public welfare practice, divide the teaching content of labor education into simply theoretical teaching and labor practice, without deeply exploring the connotation of labor education in applied colleges and universities. They have not yet organically combined the content of labor education with serving local economic and social development, and still have deficiencies in highlighting the integration of industry and education, school-enterprise cooperation, and cultivating applied technical skill talents.

2.2. Strategies for Implementing Labor Education in Colleges and Universities. There are many problems in labor education in applied colleges and universities, both common problems with research universities and individual problems in applied colleges and universities, which should be solved and improved in different and targeted ways. To clarify the goal, run with characteristics and focus on innovation, applied colleges and universities should actively adapt to the new normal of China's economic development and comprehensively improve their ability to serve regional economic and social development and innovation-driven development [11, 12, 14, 15]. Therefore, Liaoning colleges and universities should closely meet the needs of economic and social development of the old industrial base in Northeast China in terms of talent training objectives and labor education curriculum setting, innovating labor education mode according to local natural, economic and cultural conditions, combine with the school's own characteristics, fully explore the available resources such as industry enterprises, run special features, vigorously cultivate innovative talents, meet the requirements of regional economic restructuring and industrial upgrading and serve the old Northeast China.

The university will also be able to meet the requirements of regional economic restructuring and industrial upgrading and serve the revitalization of the old industrial base in the Northeast. In the implementation of labor education, we should always take the inheritance of the red gene of soldier spirit as the core, focus on the institutional mechanism of school-enterprise cooperation to cultivate soldier heirs, introduce enterprises into education, integrate industry and education, integrate soldier spirit into the whole process of labor education, cultivate senior applied talents, and contribute to the enhancement of comprehensive economic strength of northeast China.

Restore the value of labor, reshape the labor culture, and enhance the recognition of the value of labor. The Marxist view of labor holds that labor creates man himself, labor is the first historical activity of human beings, labor creates the world, labor promotes the continuous sublimation of human understanding, and labor promotes the liberation and comprehensive development of human beings. The construction of campus labor culture should be based on the core concept of labor is the most glorious, labor is the noblest, labor is the greatest, and labor is the most beautiful [16–18]. We also make use of new media such as campus micro-letter, school radio station, propaganda board, etc., and various forms such as reports and lectures to widely publicize the advanced deeds of labor role models and labor models, correct the labor dynamics of college students, improve their labor literacy, discard the wrong thoughts of despising and hating labor, establish correct labor values, cultivate the excellent quality of college students who respect labor, respect labor and love labor, stimulate their inner demand and motivation of labor, and realize To achieve the unity of knowledge and action in labor education.

To improve the curriculum system of labor education, the guiding outline points out that schools are the main body of labor education implementation, and they should make overall design and systematic planning of labor education and form the overall implementation plan of labor education, according to the relevant national regulations and combining with the actual situation of local and our school. Therefore, schools should comprehensively sort out the curriculum system according to the talent training objectives and training standards, continuously integrate and optimize the existing teaching resources, integrate labor education into the entire process of talent training, set overall goals of labor education, implement them at all levels, and achieve effective results. To this end, labor education should be conducted from two aspects: theoretical teaching and practical teaching. In terms of theoretical teaching, first, we should set up compulsory courses on labor education in accordance with the requirements of the guiding outline, educate college students systematically on the Marxist concept of labor, and guide students to arm their minds with the Marxist concept of labor and correctly understand the value and significance of labor from a theoretical point of view. Second, we should give full play to the role of the main channel of Marxist basic principles, ideological and moral cultivation, and legal foundation, etc., to shape positive labor dynamics and emotions, encourage students to devote themselves to labor and make their dreams come true with labor.

In terms of practical teaching, a module of labor education should be set up in the curriculum, and the main practical teaching links of labor education, such as internship training, social practice, and work-study, should be clearly defined, while the relevant contents of labor education should be organically integrated into teaching practice and course design and graduation design, so as to realize the organic combination of theory and practice of labor education. The combination of the case study teaching method and group discussion has been proved to be an effective teaching method. The specific practice is to divide students into learning groups and give them 10 minutes to discuss with the teaching content, then each group sends a representative to share the results of group discussion, and finally, the teacher summarizes. The teacher should reflect on the students' learning, identify problems, and make improvements.

The key is to design the discussion questions in a way that is relevant to the students' thinking and not abstract and empty. For example, when teaching the cultivation of positive labor spirit, you can raise questions by watching the video of the advanced deeds of typical figures of great craftsmen, so that students can feel the labor sentiment of model workers and have a more comprehensive and profound understanding of labor spirit. In addition, for students' bad behaviors such as wasting food and not cherishing the fruits of others' labor, they can use the teaching activities of having students create their own school plays to make them understand right and wrong and correct their mistakes. In conclusion, teachers should choose a variety of mixed teaching methods such as participatory, interactive, and debating according to the different course contents of labor education, and continuously optimize the teaching methods to improve the teaching effect.

2.3. Artificial Intelligence Technology. Artificial intelligence technology-driven labor education implementation path is the core technology of long-time, cross-domain multi-objective tracking, whose main goal is to re-identify the same labor education across cameras. In the study, usually given a target labor education image or video clip, the target labor education is identified in the image sequence or video clip to be matched, and the similarity between the target image and the matched image is given [19]. Two surveillance cameras with non-overlapping fields of view, which capture some labor education images at different time periods, select one image from the images captured by camera one as the target image. The goal of the artificial intelligence technology-driven labor education implementation path is to achieve cross-camera tracking of labor education by finding the image with the highest match to the target image and identifying that person again among the images captured by second camera.

For artificial intelligence technology to drive the labor education implementation path, it is necessary to first obtain the feature vector of the labor education image using a feature extractor, and then judge the similarity between the target image and the image to be matched using a metric learning method. Metric learning is a common method of machine learning to determine the degree of similarity of feature vectors through quantitative methods. In the training process of the network, the metric learning method is mainly embodied in the loss function, which is used to update the parameters of the deep network, so as to extract more discriminative image features. The main loss functions used in artificial intelligence technology-driven labor education implementation paths are as follows. Feature extraction at multiple scales is accomplished using null convolution, which makes full use of the contextual information of the image. STN is introduced for image segmentation, local features are extracted using MSCAN networks, and the extracted global features are fused with local feature vectors. SP-ReID architecture uses semantic segmentation methods to replace the commonly used detection methods to determine the target region to avoid the interference of background information for local feature extraction. The recognition performance of the network is improved by fusing global information with multi-scale local information using a multi-task pyramid overlap matching approach and combining the edge information of each chunk. The human body semantic parsing approach is used to localize and classify the labor education body parts at the pixel level by human identity labels only [20].

Let each input image go through two independent convolutional layers to generate two feature maps and fuse them to obtain the respective feature maps of the two images. Combining the single-image feature approach for the input image pairs, extracting their respective features using separate convolutional neural networks, and extracting the feature maps from the feature extraction process for fusion, this approach combines the high efficiency of extracting single-image features with the advantages of the CIR approach for inter-image information extraction. The deep feature fusion AI technique drives the path model of labor education implementation, using convolutional and pooling layers to extract deep features of the network multiple times, and using the fused deep features as feature attributes of labor education, which improves the ability of the network to extract global features. With the same idea, the feature maps extracted from the last three convolutional layers of the feature extraction network are stitched together, and then the deep features are extracted using the convolutional neural network, which improves the deep feature extraction ability of the network [21]. The cross-modal fusion method is used to fuse the visual features as well as the spatial-temporal features of the images, and the method achieves better performance on several datasets.

3. Methods

3.1. Model Architecture. Based on the attention mechanism, this study proposes the diverse local attention network model shown in Figure 2, which contains four modules: backbone network, multi-branch LAN, classification recognition network, and CAP network. In Figure 2, F_g denotes the input of global branch, F_k denotes the input of the k-th LAN branch, F denotes the input of global or local classification recognition network, f denotes the normalized feature, $U*$ denotes the activation graph generated by LAN, $L*$ denotes various loss functions; GAP stands for global average pooling and *BN* stands for batch normalization. The detailed elaboration of each symbol in Figure 2 is described later. In addition, for the purpose of algorithm description, the training set is assumed to be S, where xi denotes the i-th labor education image, y_i is the label of xi, and C and N denote the number of pedestrians and the training set size, respectively. Each module is described in detail below separately.

For deep learning networks, the learning ability of the network theoretically increases as the depth increases. The residual network architecture is easy to optimize and solves the degradation problem well. Since deep features contain rich semantic information, existing methods prefer Resnet as the backbone network, and so does this study, as shown in Figure 2. Resnet-50 contains one convolutional layer (Conv1) and four residual modules, and each residual module contains multiple convolutional layers, BN layers, and ReLU activation functions. To obtain a more comprehensive representation of labor education features, the network is divided into multiple branches after Conv4 as global and local branch inputs, and no further down-sampling operations are performed in F. Specifically, the feature map obtained after inputting the labor education images into the backbone network, D, H and W denote the number of channels, height, and width of the features, respectively.

3.2. Multi-Branch LAN. Unlike the rigid spatial division of the predefined input image, the multi-branch LAN can adaptively locate multiple non-intersecting salient regions in the image. The multi-branch LAN consists of multiple LAN models, and the detailed structure of the LAN is shown in Figure 3. First, each LAN passes through a 3×3 convolutional layer (denoted by Conv1) and a *ReLU* layer to nonlinearly map the features F into a more compact feature space. Next, another 1×1 convolutional layer (denoted by Conv2) is passed to obtain the spatial attention score e_k. Assuming that $e_{k,l}$ is the basic element of e_k, where e_k denotes the spatial location index, then $e_{k,l}$ is the attention score of the lth spatial location of the k-th branch. Then, e_k is subjected to the SoftMax activation function to obtain the L-dimensional spatial attention weight vector w_k. In addition, the reorganization operation is performed on w_k, and each LAN model simultaneously outputs spatial attention weights w_k. Finally, the input features F are weighted and summed according to w_k to obtain the output features for the k-th LAN model.

3.3. Classification Recognition Network. Based on the idea of combining global and local, the classification recognition network is designed as two major branches, global and local. For the global branch, firstly, the feature FERD is obtained by GAP operation on the feature input F_g, and then the final global total loss function is obtained by the top and bottom structure. F_g and F_{gn} belong to different labor education, and the triadic loss L function is

$$L_{tri}^{g} = \frac{1}{N_{tri}} \sum_{\left(F_{g}'', F_{g,p}'', F''\right)} \max\left[0, d\left(\mathbf{F}_{gri}'', \mathbf{F}_{g,p}''\right) - d\left(\mathbf{F}_{g}'', \mathbf{F}_{g,n}''\right) + \delta\right], \tag{1}$$

where T_{tri} denotes the set of triads, N_{tri} is the potential of T_{tri} i.e. the number of triads, $d(-)$ is the Euclidean distance, and δ is the distance interval parameter. The lower branch of the global branching to F first performs the BN operation to obtain the normalized feature, and then uses the fully connected layer and SoftMax function to obtain the probability of the class to which the sample belongs, and further calculates the cross-entropy loss L as the final classification loss. To prevent the model from overfitting, the label smoothing (Label Smoothing) technique can be used during training, so the specific process of the lower branch can be expressed by the formula as

$$\begin{aligned} q_i^g &= \operatorname{soft\,max}\left(\mathbf{W}_g \mathbf{f}_{g,i}\right), \\ p_i^g &= y_i(1-\varepsilon) + \frac{\varepsilon}{C}, \\ L_{\mathrm{id}}^g &= -\sum_{i=1}^{N} p_i^g \lg q_i^g. \end{aligned} \tag{2}$$

where $\mathbf{W}_g$, $\mathbf{f}_{g,i}$, and p_i^g denote the weight matrix of the fully connected layer, the normalized feature of the i-th sample and the smoothed label of the i-th sample, respectively, and ε is the smoothing factor. So far, the total loss function is L_{global}, local classification recognition network is like the global branch, first, the output F of the k-th LAN is subjected to GAP operation to obtain the feature, and the subsequent processing of each local branch is the same as the global branch, so it will not be repeated (just replace the symbol g

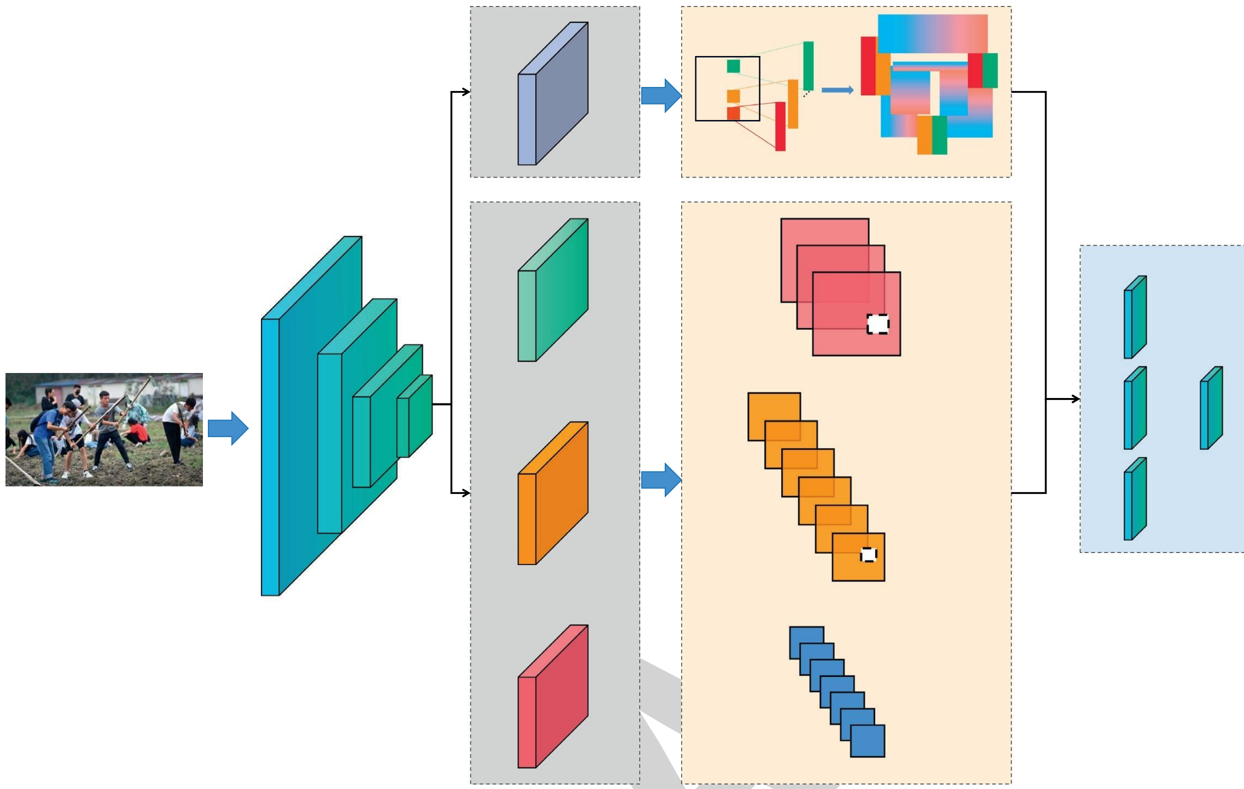


FIGURE 2: Model structure.

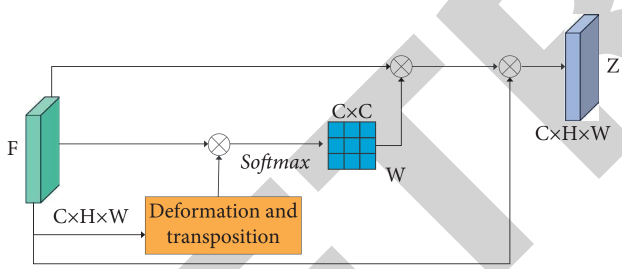


FIGURE 3: LAN structure diagram.

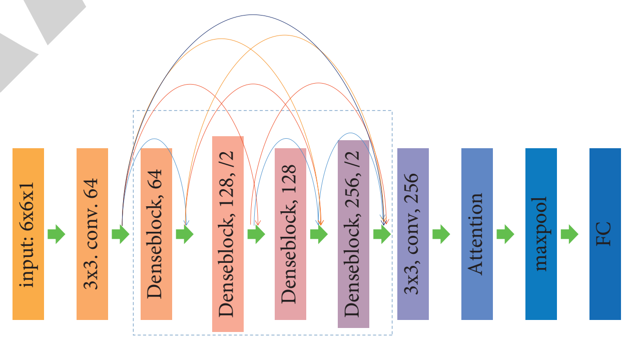


FIGURE 4: CAP structure diagram.

in the above formulas with k). Therefore, the total loss of local branches is $L_{\text{local}} = \sum_{k=1}^{K} (L_{\text{id}}^{k} + L_{tri}^{k})$.

3.4. CAP Network. For a given labor education image, if the multi-branch LAN module is left unconstrained, it will cause the branches to converge, i.e., multiple LAN models can easily focus on the same significant regions and thus ignore other sub-significant regions that also have discriminative power. The CAP structure diagram is shown in Figure 4. Therefore, it is necessary to ensure that each of the K branches focuses on a different region of the image during model training, i.e., each local branch has a different high activation region for the feature response. To this end, a CAP network is proposed to achieve local feature diversity, which is the core of the DLAN model. In short, the CAP network uses the spatial attention weights w_k from the LAN output to guide each local branch to focus on different salient regions of the human body. Specifically, this study uses the Hellinger distance $H(\text{-})$ to measure the consistency of w_i and w_j of any two LAN branch outputs, i.e.,

$$H\left(w_i, w_j\right) = \frac{1}{\sqrt{2}} \sqrt{w_i} - \sqrt{w_j}_2,$$
$$H^2\left(w_i, w_j\right) = 1 - l = 1 \sum_{l=1}^{L} \sqrt{w_{i,l} w_{j,l}}. \tag{3}$$

In order to ensure that the highly activated regions of different attention models do not overlap each other, it is necessary to maximize the distance between w_i and w_j, which is equivalent to minimizing $1 - H^2(w_i, w_j) = \sum_{l=1}^{L} \sqrt{w_{i,l} w_{j,l}}$, and construct the following CAP loss:

$$L_{\text{CAP}} = \mathbf{\Omega}\mathbf{\Omega}^T - I_F^2, \tag{4}$$

where $\|\cdot\|_F^2$ denotes the Frobenius parametrization and I is the K-dimensional unitary array. Obviously, minimizing L_{CAP} is equivalent to minimizing $\sum_{l=1}^{L}\sqrt{w_{i,l}w_{j,l}}$. If the region is a significant feature, the $w_{i,l}$ and $w_{j,I}$ of different branches in the optimization process one tends to be extremely small and the other tends to be extremely large; if the region l is a background feature, the $w_{i,l}$ and $w_{j,I}$ of different branches will both tend to be extremely small values. In other words, smaller loss values will be obtained when the high activation regions of the attention maps obtained from different LAN branches are not consistent. Therefore, the optimized LCAP can control the update direction of w_k, i.e., iterative update along the feature diversification direction.

4. Experiments and Results

4.1. Experiment Setup. This section will verify the effectiveness of the proposed algorithm (DLAN) through various experiments. All experiments are conducted using the PyTorch deep learning framework with a GPU workstation configuration of an Intel Core i7-type CPU, 32 GB of RAM, and a 1080Ti graphics card with 12 GB of video memory.

In this study, experiments were conducted using four commonly used AI technology-driven labor education implementation pathway datasets. Each dataset is divided in advance into a training set and a test set, and the test set is divided into two parts: the image library and the query set. Among them, Market1501 contains a total of 12,936 training images and 23,100 test images from 1501 labor education in 6 cameras, Duke-MTMC-reID contains a total of 16522 training images and 19,889 test images from 8 cameras, CUHK03 contains a total of 7365 training images and 6732 test images from 10 cameras and Partial-REID contains a total of 300 training images and 300 test images. The standard performance evaluation metrics of the artificial intelligence technology-driven labor education implementation path method include the mean average precision (mAP) and the 1st matching rate Rank-1 of the cumulative matching characteristic curve (CMC), so these two metrics are also used in this paper to measure the ability of the DLAN model to retrieve the labor education images to be queried from the image library. For the DLAN model, the size of the input image is 256×128; random horizontal flipping and random erasure are used to achieve data enhancement during the model training. The optimal or better values of the number of branches and hyperparameters were determined. The training settings are shown in Table 1.

The training process loss convergence and performance enhancement are shown in Figures 5 and 6.

4.2. Experimental Results. The proposed method in this study aims to improve the accuracy of AI technology-driven labor education implementation paths on the processed dataset. To verify the effectiveness of the network, the obtained experimental results will be compared with the training results under two basic networks. Network 1 (Baseline1, B1) is fine-tuned on the trained GoogLeNet by

Table 1: Training parameters.

Training parameters	Value
Momentum	0.9
Initial learn rate	0.005
Learn rate drop factor	0.5
Learn rate drop period	10
L2 regularization	0.004
Max epochs	50
Mini batch-size	64
Validation frequency	30

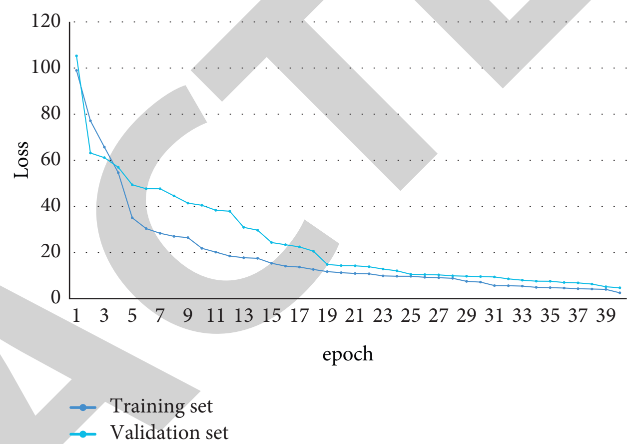


Figure 5: The training process loss convergence schematic.

setting the number of neurons in the last fully connected layer to the number of labor education being used for training. The test process extracts 1024-dimensional feature vectors obtained from pool5 layers for each image in the reference set and test set, and then calculates the Euclidean distance between the vectors. Network 2 (Baseline2, B2), on the other hand, directly uses the comparison method in which only the segment quality of the images is evaluated. Tables 2 and 3 show the results of the experimental evaluation on the datasets iLIDS-VID and PRID2011. The partial cropping of the images in the dataset at the bottom increases the difficulty of the AI technology-driven labor education implementation path, so all three methods have reduced results compared to the direct application on the standard dataset. However, the proposed method in this study is less affected and has a higher accuracy than the other two methods. On the dataset iLIDS-VID, B1 has an accuracy of 50.7% on rank1, and B2 has a 4% increase over the results of B1 due to the addition of the quality assessment module. The proposed method in the study incorporates attribute local features and can achieve an accuracy of 63.3% on rank 1, which is an increase of 12.6% and 8.6% over the results of B1 and B2, respectively. In addition, on the dataset PRID2011, the results of the comparison experiments show that the accuracy of the proposed method increases by 10.0% and 4.4% over B1 and B2, respectively.

Since the dataset MARS was not used for experiments in the comparison method, only the results obtained by the

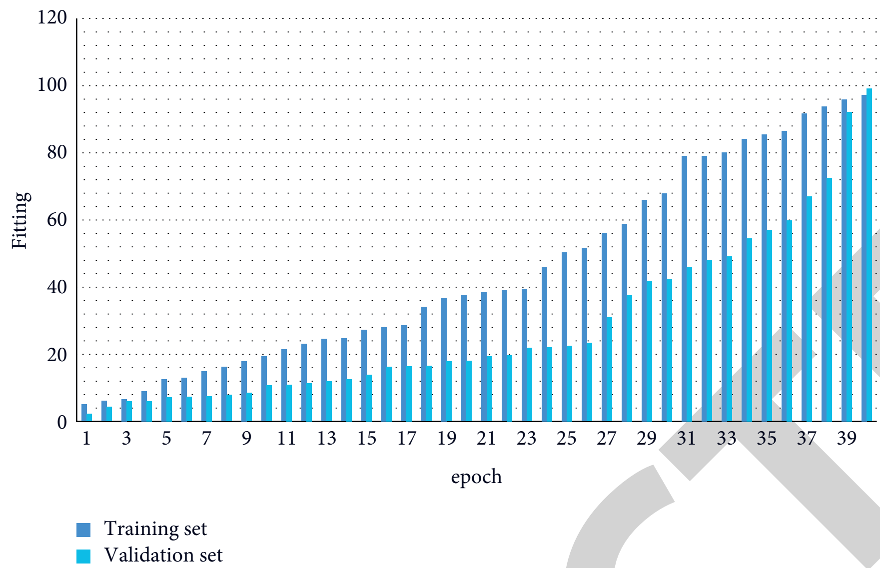


Figure 6: Schematic diagram of training process performance improvement.

Table 2: Recognition rate of related methods on dataset LIDS-VID %.

Methods	Rank			
	1	5	10	20
Baseline 1	50.7	74.7	85.3	91.3
Baseline2 (RQEN)	54.7	80.0	91.3	97.3
Ours	63.3	90.0	94.0	99.3

Table 3: Recognition rate of related methods on dataset PRID2011%.

Methods	Rank			
	1	5	10	20
Baseline 1	80.0	97.8	100.0	100.0
Baseline2 (RQEN)	85.6	100.0	100.0	100.0
Ours	90.0	100.0	100.0	100.0

proposed method on this dataset are compared with those obtained by other existing methods in this paper. From the accuracy of the AI technology-driven labor education implementation paths, the proposed method in this study still has a higher recognition rate than other methods, despite the pre-processing work on the dataset in this study, which crops out some of the labor education bottom and increases the difficulty of AI technology-driven labor education implementation paths. Figure 7 shows that on the dataset MARS, the proposed method in this study improves the accuracy of rank1 by 4.8% over the existing methods. Figures 8 and 9 show that the accuracy of rank1 improves by 5.3% on dataset iLIDS-VID, while the accuracy improves by 9.6% on dataset PRID2011. This result demonstrates that the inclusion of attribute features can effectively improve the effectiveness of AI technology-driven labor education implementation pathways.

A network structure that combines image-based local region quality assessment and attributes recognition can learn both global and local features of images. The method remains dependable in the problem of misalignment due to missing images in some datasets and can address the limitations when using only segmented assessment networks into labor education AI techniques to drive labor education implementation paths. To confirm the effectiveness of the proposed method in this study, labor education in the dataset is labeled separately. The experimental results show that the network with the

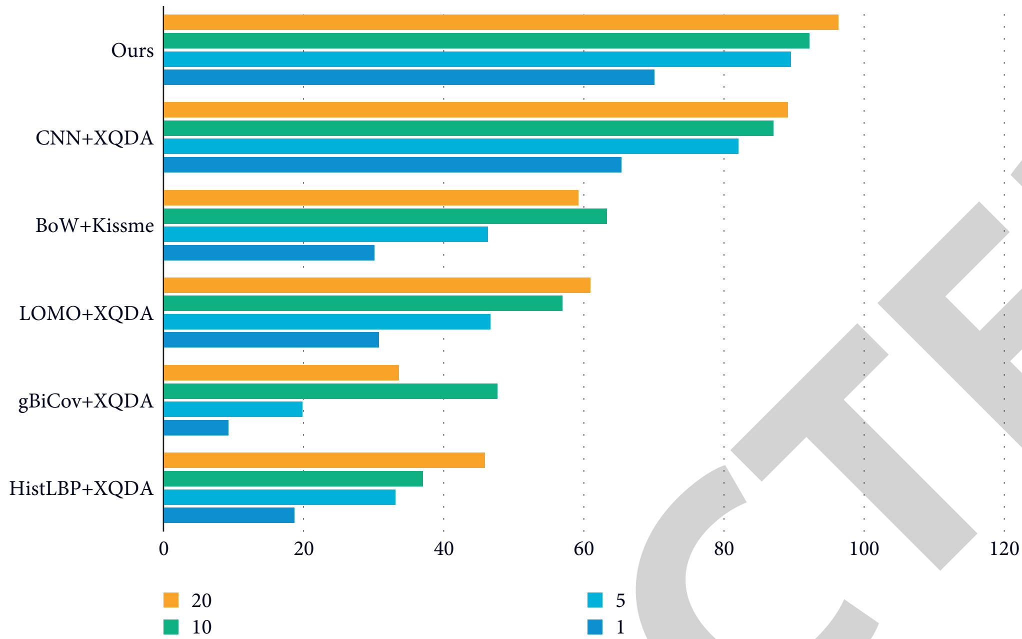


Figure 7: Recognition rate of existing methods on the dataset MARS.

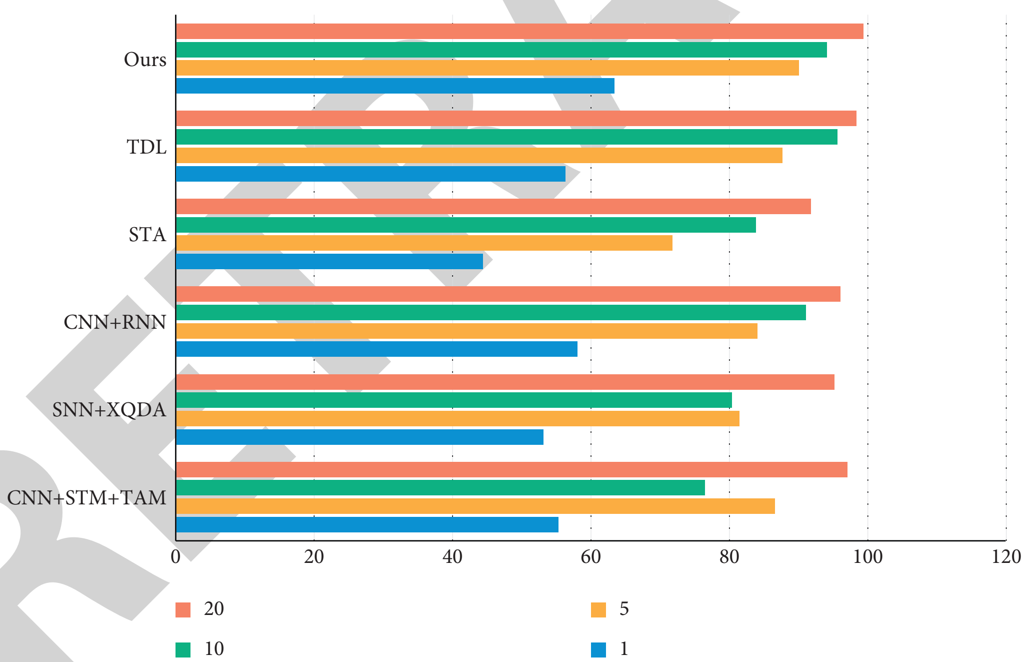


Figure 8: Recognition rate of existing methods on the dataset LIDS-VID.

introduction of attribute features can improve the accuracy of the labor education implementation path driven by artificial intelligence techniques. There are still many areas for further improvement in the method of re-identification using attribute features, such as selecting more representative attributes and assigning larger weights to them during network training, which will be investigated subsequently.

5. Conclusion

Labor education is an important content of the socialist education system with Chinese characteristics, which directly determines the labor spirit outlook, labor value orientation and labor skill level of socialist builders and successors. To carry out labor education in new era of applied colleges and universities, it is necessary to constantly

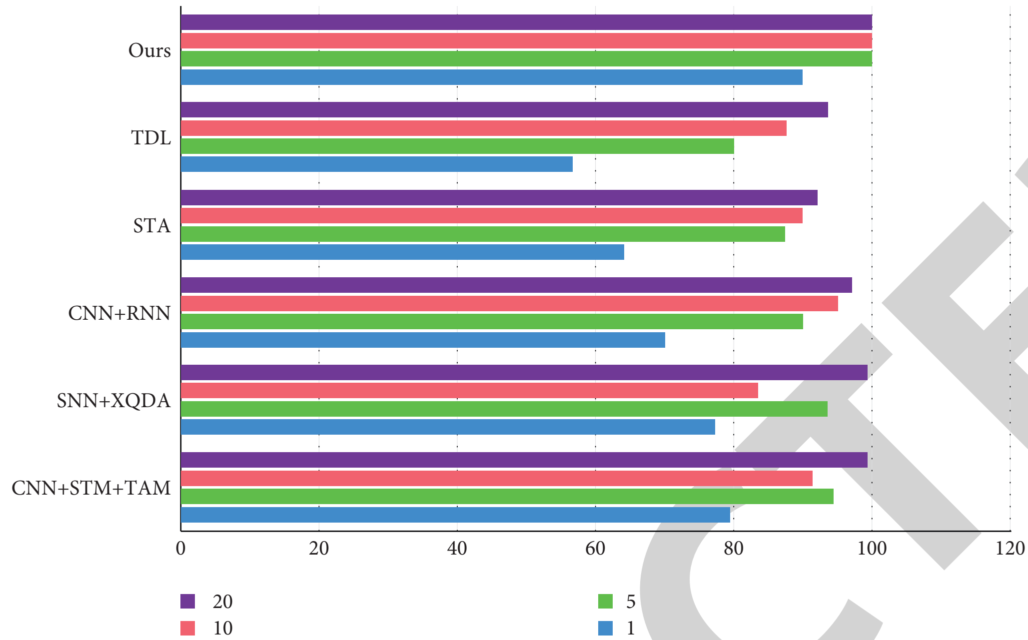


FIGURE 9: Recognition rate of existing methods on dataset PRID2011.

update the concept of labor education, focus on the effectiveness of labor education, grasp the law of students' growth and the law of labor education, fully explore the value of ideological and political education, the value of labor skill cultivation and the value of comprehensive quality improvement in labor education, and build a labor education system supported by daily life labor, service labor and production labor with the advanced concept, rich carrier and distinct levels.

The system of labor education is supported by the content of daily life labor, service labor and production labor, with advanced concept, rich carrier and clear hierarchy, and cultivates the Marxist concept of labor among college students. In the actual application scenario, the AI technology-driven labor education implementation path faces the problems of gestalt change and local obscuration. To this end, this study proposes a diverse local attention network model based on joint global and local learning, which relies on spatial attention networks to locate and enhance activation responses in salient regions and enables each local branch to focus on non-overlapping body parts through diverse canonical constraints, thus improving the re-identification accuracy. On four public datasets, ablation experiments, visualization experiments, occlusion experiments, and comprehensive comparison experiments with existing advanced methods have been carried out successively, fully verifying the robustness and excellent recognition performance of the proposed method. In future work, the consistency constraint will be further utilized to obtain multi-granularity features of global branching and consider learning the spatial relationship among the features to obtain higher accuracy.

Data Availability

The datasets used during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This work was sponsored in part by Fund Project: Key Research Bases of Humanities and Social Sciences in Higher Education in Sichuan Province"Sichuan Labor Education Research Center" project in 2021: Exploration of the implementation path of labor education in applied colleges and universities (Project number: SCLD002).

References

[1] M. Zhao, C. Chen, L. Liu, D. Lan, and S Wan, "Orbital collaborative learning in 6G space-air-ground integrated networks," *Neurocomputing*, vol. 497, pp. 94–109, 2022.

[2] C. Chen, Y. Zeng, H. Li, Y. Liu, and S. Wan, "A multi-hop task offloading decision model in MEC-enabled internet of vehicles," *IEEE Internet of Things Journal*, vol. 8, pp. 53062–53071, 2022.

[3] C. Chen, H. Li, H. Li, R. Fu, Y. Liu, and S. Wan, "Efficiency and fairness oriented dynamic task offloading in internet of vehicles," *IEEE Transactions on Green Communications and Networking*, vol. 10, pp. 156–162, 2022.

[4] C. Chen, J. Jiang, Y. Zhou, N. Lv, X. Liang, and S. Wan, "An edge intelligence empowered flooding process prediction using Internet of things in smart city," *Journal of Parallel and Distributed Computing*, vol. 165, pp. 66–78, 2022.

[5] C. Chen, J. Jiang, R. Fu, L. Chen, C. Li, and S. Wan, "An intelligent caching strategy considering time-space characteristics in vehicular named data networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 99, pp. 1–13, 2021.

[6] L. Wu, Y. Wang, X. Li, and J. Gao, "Deep attention-based spatially recursive networks for fine-grained visual

Hindawi
Mobile Information Systems
Volume 2022, Article ID 4715102, 11 pages
https://doi.org/10.1155/2022/4715102

Research Article

Data Anomaly Diagnosis Method of Temperature Sensor Based on Deep Neural Network

Haifeng Chen

School of Information Engineering, Lianyungang Technical College, Lianyungang 222000, Jiangsu, China

Correspondence should be addressed to Haifeng Chen; 1992050004@lygtc.edu.cn

Received 10 July 2022; Revised 8 August 2022; Accepted 16 August 2022; Published 2 September 2022

Academic Editor: R. Mo

As one of the indicators of whether all kinds of machinery and electrical appliances work normally during use, temperature has an important basis for judging the normal work of related machinery. In order to reduce the probability of safety and quality problems caused by inaccurate temperature measurement in the use of these machines and electrical appliances, this paper uses RBF neural network and EEMD modal analysis two deep neural network models to build a deep neural network-based temperature sensor data anomaly diagnosis method. This method first excavates a large number of historical temperature data samples of temperature control sensors in machinery and electrical appliances and analyzes the change law of relevant sample data, so as to build a data anomaly diagnosis database, and then establish a temperature prediction model based on RBF to predict the temperature of electrical appliances and mechanical components; Second, real-time temperature monitoring and sampling are carried out for the normal temperature sensor. Based on the constructed sample database, automatic identification of various abnormal conditions is realized, and the real measured value of the sensor is reconstructed or estimated under abnormal conditions. EEMD feature extraction is carried out for the difference between the predicted temperature and the actual temperature; Finally, the RBF temperature anomaly diagnosis and classification model is constructed, and the feature vector sets are constructed by variance, variance percentage, energy and energy percentage methods, respectively, or jointly, and these vector sets are used as the input of the fault model for temperature anomaly diagnosis and monitoring. Through the diagnosis of the measured temperature sensor data, the established model has a good ability of fault diagnosis and classification.

1. Introduction

Sensor is an indispensable and important part of measurement and control system, and it is the main tool to collect signals. With the increase and complexity of industrial machinery and civil electrical appliances, the complexity of measurement and control system is gradually increasing, the number and types of sensors are greatly increased, and the working environment is getting worse and worse. In addition, the performance of the sensor is not only directly related to the running state of the equipment but also related to the vital security issues. In particular, the working condition of sensors that provide control signals directly affects the state of the system, so it is more important [1]. The sensors composed of precision components are often in a bad working environment, and failure is inevitable. At this time, the performance of the automation system will be degraded at least, and disastrous consequences will be caused at most. Accidents that cause equipment damage and huge economic losses due to sensor failure also occur from time to time. According to statistics, the sensor failures in general automation systems account for more than 45% of all failures [2, 3]. However, once the sensor fails, the performance of the control system will be degraded at first, and then catastrophic consequences will be caused. Therefore, the demand for on-site fault diagnosis of sensors is becoming increasingly strong, and it is very necessary to carry out fault diagnosis of sensors.

In the field of automatic control, the sensor converts the measured physical signal into electrical signal, which is very important for the operation of the monitoring system. When it fails, it will have a serious impact on the subsequent monitoring, control, fault diagnosis, and other systems [4]. For example, in the locomotive transmission control system, temperature is an important indicator of the working environment and state of the equipment, such as the oil temperature of the traction transformer, the temperature sensor temperature of the water inlet and outlet of the traction converter water cooling system, the air temperature of the cabinet, and the temperature of the traction motor. If the temperature is too high, it is necessary to reduce the power of the locomotive, disconnect the main circuit breaker, and other protective actions to ensure the safe operation of the equipment. The operation environment of traction drive system device is complex, and corrosion, temperature, humidity, electric surge, static electricity, and other factors will affect its operation state [5, 6].

In foreign countries, the research on abnormal diagnosis of temperature sensor data has a long history. As early as the 1950s, scientists used grain monitoring system to ensure food security. Developed countries in Europe and the United States have an early scientific and technological revolution and have always been in the lead in technical research in related fields. They have an efficient temperature monitoring system and standards. On the basis of previous studies, Zhao et al. introduced the concept of dynamic radius support vector data description (DR SVDD) and the idea of kernel space angle to detect turboshaft engine faults based on Dr SVDD [7]. Msi et al. proposed a new fault detection system for large-scale grid connected photovoltaic power stations. The fault detection system makes a cascade comparison between the DC power supply of the actual photovoltaic power station and the simulated photovoltaic power station, distinguishes the fault situation, and identifies the nature of the fault [8]. Hd et al. proposed a fault detection method based on asymmetric pole inductance (APIs), which has certain accuracy in fault detection [9]. Li et al. systematically studied the application of ANN and hybrid ANN models in photovoltaic fault diagnosis. For each application, the target photovoltaic fault, detectable fault, data type and amount used, model configuration, and FDD performance were extracted and analyzed [10]. Duan et al. Used the correlation between atmospheric data and granary historical data to propose SVR method to predict the average temperature of grain pile [11]. Kurpaks et al. established a dynamic model of water content and greenhouse heat flow, and expounded the movement of indoor heat and mass [12]. Chang et al. proposed the finite difference method to establish a model to predict the moisture content and distribution of grain [13]. Szke et al. used the temperature of the last two days in combination with the maximum and minimum temperature given by the weather forecast to predict the temperature in the next 24 hours [14]. Fan et al. described the actual greenhouse system based on the greenhouse temperature prediction model of SVR [15]. These scholars have done a lot of research on fault diagnosis, but the abnormal monitoring of temperature sensor data based on deep neural network is not comprehensive enough, and there is a lack of practical prediction system.

However, in China, the research of sensors is relatively lagging behind. The traditional way to improve the reliability of sensor output signal is to set up multiple redundant sensors. Take the average value or multiple values, of course, this will pay a high cost. With the development of sensor technology, some intelligent sensors have been developed, which makes self-detection and self-diagnosis possible. At present, most sensor neural network fault diagnosis methods use BP neural network [16–20]. Generally, offline training and online work processing mode are adopted in data processing. However, BP neural network has the problem of slow convergence and easy to fall into local minima. In addition, the offline processing method requires that when training the network, a large number of training samples containing all the characteristics of the sensor must be provided to the network. Otherwise, when the field data exceed the coverage of the neural network training samples, the network will often output error information due to the lifting capacity of the network.

In recent years, ANN has been widely used in sensor fault diagnosis, especially for BP neural network. At present, it covers most fields. The principle of BP artificial neural network is relatively simple. It was originally evolved from multilayer perceptron and used for network training with unique back propagation [20–23]. However, when adjusting the weight of BP neural network, it is different from other neural networks, which often adopts the negative gradient descent method. It is precisely for this reason that the network is prone to fall into local extremum and poor convergence speed for some complex training samples, which limits its scope of application. Therefore, people turn their attention to another radial basis function (RBF) neural network that can solve the above shortcomings. It is superior to BP neural network in approximation ability, classification ability, and so on [24, 25]. At the same time, Wang et al. [26] used the limit learning machine method to classify the faults of the fuel system. Guo et al. [27] combined the circular model with the limit learning machine (ELM) to form a fault diagnosis method for linear analog circuits. Xia et al. [28] reported an effective diagnosis method for early faults of water chillers by combining nuclear entropy component analysis (KECA) and voting based ELM (VELM). Li et al. [29] proposed a method based on extreme learning machine and ADABoost Samme's nuclear power plant fault diagnosis framework. Liu et al. [30] proposed a new gear personalized fault monitoring model in combination with two different neural networks (finite element method and elm method) to solve the problem of gear fault when mechanical devices work. The above research combines various methods with limit learning machine for fault diagnosis and has achieved good results, but most of them have poor universality. Moreover, due to the randomness of the weights and thresholds of ELM, the results of each time are unstable, which is easy to affect the prediction results. Therefore, many scholars have studied a fault diagnosis classification model based on empirical mode decomposition (EEMD). After analysis, it is found that the model has strong signal analysis technology ability and is very suitable for dealing with this kind of problems.

In order to solve the above problems, this paper proposes an intelligent temperature data anomaly diagnosis method

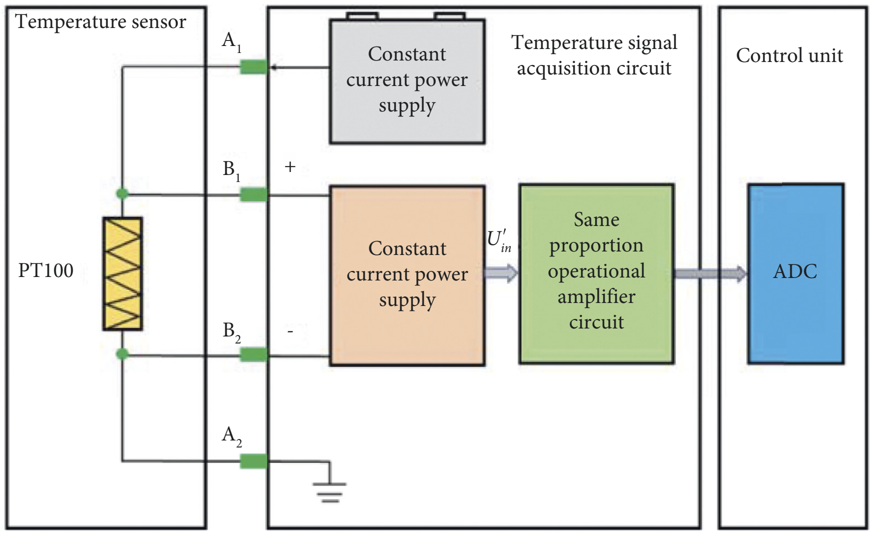


Figure 1: Schematic diagram of temperature signal sampling.

based on RBF neural network and empirical mode decomposition (EEMD), which is based on historical fault data samples and normal historical data of temperature sensors. The rule base of sensor fault diagnosis is established by analyzing the characteristics of fault samples, and the fault-tolerant estimation method of fault conditions is obtained. For the fault conditions that cannot be reconstructed from the fault samples, the normal historical data of the sensor are used to establish the grey prediction model of the measured value of the sensor, so as to realize the fault-tolerant estimation of the nonisolated fault samples, and carry out the research on the abnormal diagnosis method of temperature sensor data based on deep learning.

2. Fundamental Theory

2.1. Brief Introduction of Signal Sampling Circuit of Temperature Sensor. Thermal resistance (such as PT100) is a temperature sensor that converts temperature into resistance based on the principle that its resistance value changes with temperature. The temperature measurement circuit obtains the resistance value (voltage/current) by first applying a known excitation current to the thermal resistance and then measuring the voltage at both ends and converts the resistance value into the temperature value, so as to realize the temperature measurement [31].

There are three connection modes between thermal resistance and temperature measuring circuit: two-wire system, three wire system, and four wire system. Because the four wire measurement method is not affected by the resistance of connecting wires, it has been widely used in the field. This wiring method is also used in most locomotive transmission systems. Therefore, this paper takes the four wire temperature sensor as an example to study its signal sampling circuit principle, fault diagnosis, and fault-tolerant estimation methods.

The typical temperature signal sampling principle is shown in Figure 1.

In Figure 1, both ends of PT 100 are connected to 4 through A1 and A2 ends 9 mA constant current source, and then the temperature signal sampling circuit obtains its differential input voltage U_{in} through B1 and B2. U_{in} generates a single ended voltage signal U'_{in} 'through the voltage follower and then enters the in-phase proportional operation amplification circuit to generate a voltage signal that meets the sampling requirements of the analog to digital converter (ADC) of the control unit.

If the amplification factor of the in-phase proportional operation amplification circuit is k, the relationship between the sampling temperature value t of the temperature sensor (PT100) and the output voltage value u of the temperature signal sampling circuit is shown in formula (1).

$$U = (0.385T + 100) \times 0.0049K. \quad (1)$$

2.2. Basic Principle of RBF Neural Network. MP model describes neurons from the perspective of logic functional devices, which is a mathematical simplification of biological neuron information processing mode and establishes the theoretical research foundation of neural networks.

In 1988, broomhead and low e introduced radial basis function into neural network to form RBF neural network [32]. RBF neural network is a three-layer feedforward network. Its basic idea is to use RBF as the "base" of hidden units to form the hidden layer space and transform the low-dimensional input vector into the high-dimensional space through projection. So that the original linear inseparable problem becomes linear separable. Figure 2 shows the basic structure of RBF neural network.

For the structure of RBF network, its principle is relatively special. The hidden layer space formed by it can directly map the input vector to the hidden layer through neurons, so as to reduce the weight association of the middle layer. Therefore, the association weight from the input layer to the hidden layer of this neural network is also relatively

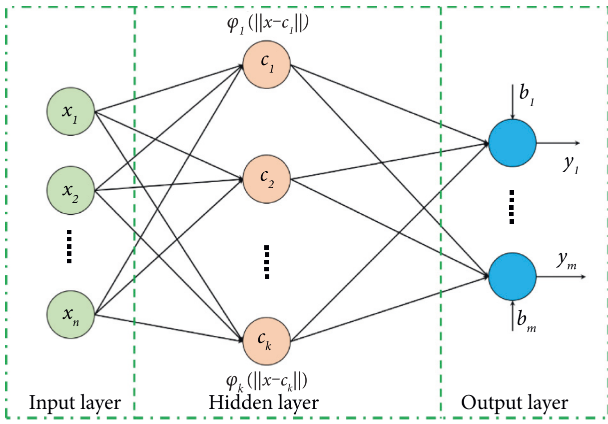


Figure 2: RBF neural network.

special, which is 1. The hidden layer is only responsible for the nonlinear projection of the input vector of the input layer, and the output layer is only responsible for the weighted sum of the values mapped by the hidden layer. The parameters to be learned and optimized in RBF neural network include the center and variance of radial basis function and the connection weight from hidden layer to output layer. The output layer is responsible for optimizing the weights through linear optimization strategy, and the learning speed is usually fast. The hidden layer needs to use nonlinear optimization method to adjust the parameters of the activation function, so its learning speed is relatively slow. In the learning process of RBF neural network model, the learning method of its parameters is not unique, but the selection of radial basis function center is mostly used, mainly including orthogonal least square method and supervised center method, as well as random center method and self-organizing selection method. In addition, the learning process of RBF neural network includes two stages: the first stage is to solve the center and variance of the hidden layer basis function, which is also called unsupervised learning process; the second stage is the supervised learning process, which determines the connection weight between the hidden layer and the output layer. RBF neural network belongs to local approximation network, which omits the learning behavior of hidden layer weights and avoids the time-consuming layer by layer transmission process of errors in the network. Therefore, the learning convergence speed of the network is very fast. Compared with other neural networks, RBF neural network can approach any nonlinear function with any accuracy and has the best approximation performance, classification ability, and global optimization characteristics. Moreover, it has simple topology, small amount of calculation, good applicability of the network, the basic parameters such as the structure of the network, the number of hidden layer units, and constants can also be dynamically adjusted and determined, with fast convergence speed.

Radial basis function is a real valued function whose value only depends on the distance from the fixed point c, and any function satisfying the $\varphi(x,c) = \varphi(\|x-c\|)$ characteristic φ. Both are radial basis functions, which can also be the distance to the origin in a simplified case, that is, $\varphi(x) = \varphi(\|x\|)$. Using Gaussian kernel function as the basis function of radial basis function neural network, the output of hidden unit of radial basis function neural network is

$$\varphi_i(x, c_i) = G(\|x - c_i\|) = \exp\left(-\frac{1}{2\sigma_i^2}\|x_p - c_i\|^2\right), \tag{2}$$

where φ is the radial basis function, x is the sample, c_i is the ith center point of the kernel function, and σ_i is the width of the ith center point of the function. The selection of the central point of kernel function is very critical. The improper central position cannot make the network correctly reflect the actual distribution of the input sample space, and the input space cannot be well fitted. The width of the central point of the kernel function controls the radial range of the function, which is an important factor affecting the performance of RBF neural network. When the width is too small, the dividing line between classes will become blurred, which will reduce the classification accuracy; When the width is too large, the coverage area of the basis function will become relatively small, thus reducing the generalization ability of the network.

Then the output of RBF neural network is

$$y_j = \sum_{i=1}^{h} w_{ij} \exp\left(-\frac{1}{2\sigma_i^2}\|x_p - c_i\|^2\right), j = 1, 2, \cdots, n, \tag{3}$$

where y_j represents the output of the RBF neural network, x_p represents the pth input sample, c_i represents the ith center point, σ_i represents the width of the ith center point of the function, w_{ij} represents the connection weight coefficient between the hidden layer neuron i and the output layer neuron j, h represents the number of nodes in the hidden layer, and n is the number of output samples or classifications.

Based on the above theory, it can be found that the performance index is less than the given error by increasing the number of hidden layer elements, so as to continuously improve the fitting accuracy. However, in practical applications, if the number of hidden layer elements is too large, it may cause the redundancy of the model and numerical ill conditioned. Therefore, effective methods must be adopted to select the network center and determine the network weight.

(1) Cluster analysis method is used to select RBF Network Center. After selecting the center of RBF network, the determination of weight becomes a problem about parameter linearization. The least square method can be used to determine the network weight. The advantage of this method is that the center is easy to determine, the calculation time is short, and it is not easy to appear numerical ill conditioned. The disadvantage is that the selection of RBF center is separated from the determination of weight, and the center obtained cannot be guaranteed to be the best.

(2) The least square problem is orthogonalized by orthogonalization algorithm. When all centers of RBF

neural network are selected, the weights are also determined. The advantage of orthogonalization method is that it can ensure to select the best sample point as the center. Its disadvantage is that when the number of sample points is large, the workload of orthogonalization is large, the calculation time is long, and numerical ill condition often occurs.

(3) The ideal method is to combine the clustering method with the orthogonalization method, that is, the clustering method is used for primary selection, and then the orthogonalization method is used for selection, and the weight of the network is determined at the same time.

2.3. Set Empirical Mode Decomposition (EEMD). Empirical mode decomposition (EMD) is a decomposition method that can reflect the instantaneous frequency of data [33], but because the added noise is random, it is easy to cause modal aliasing, endpoint effect, screening iteration stop standard, and other problems. To solve the above problems, Hafida et al. [34, 35] proposed integrated empirical mode decomposition (EEMD). This method solves the phenomenon of modal aliasing by adding white noise to fill the discontinuous signal segment. In the process of noise signal decomposition, the filtering characteristics of white noise signal are used to solve the average value of intrinsic mode component (IMF) for many times to eliminate the interference of white noise on the original signal at discontinuous points [36].

EEMD is an improved method of empirical mode decomposition EMD [35]. The method of decomposing the original signal into several intrinsic mode functions (IMF), each IMF must meet two conditions:

(1) In all decomposed data sets, the number of extreme points and zero crossings is equal or at most 1 difference;

(2) The average value of the upper and lower envelope determined by the local maximum point and the local minimum point is 0. After EMD decomposition, the original signal $x(t)$ can be expressed as

$$x(t) = \sum_{i=1}^{n} c_i(t) + r_n(t). \tag{4}$$

EMD has the problem of mode aliasing, that is, a certain IMF component has similar characteristics (such as frequency, amplitude, etc.) in other components, which cannot fully reveal the characteristic information of the signal. EEMD adds white noise with a certain frequency and amplitude uniformly distributed to the original signal, which makes the signal centralized and continuous, and weakens the influence of instantaneous pulse on signal decomposition [33, 35]. The specific decomposition steps are as follows:

(1) Add random white noise signal r to $x(t)$:

$$x_j(t) = x(t) + n_j(t), \tag{5}$$

Table 1: Sensor temperature sampling data.

Item	Temperature of different sensors at time t/°C			
Time/min	Sensor 1	Sensor 2	Sensor 3	Sensor 4
1	23.4	23.2	22.9	24.5
2	26.8	27.3	24.1	26.6
3	31.2	29.8	29.6	30.2
4	37.4	36.7	41.5	38.4
5	42.8	41.8	36.4	42.3
⋮	⋮	⋮	⋮	⋮
56	44.7	48.5	45.3	54.9
57	49.8	50.6	44.5	51.4
58	46.8	48.8	56.2	56.9
59	49.6	45.2	51.9	47.3
60	50.2	47.5	53.7	48.4

where: $x_j(t)$ is the signal after adding random white noise, $j = 1, 2, 3 \cdots M$, and M is the number of tests.

(2) $x(t)$ is decomposed into a series of IMFs $(c_{i,j})$, using EMD:

$$x_j(t) = \sum_{i=1}^{n} c_{i,j}(t) + r_{n_j}(t), \tag{6}$$

where $c_{i,j}$ is the ith IMF of the jth test; r_{n_j} is the residual of the jth test, and n_j is the number of IMFs of the jth test.

(3) If $j < M$, repeat steps 1 and 2, adding different random white noise signals each time.

(4) Obtain $I = \min$ (N1, N2, . . ., NM) and calculate the overall average of the corresponding components as the final result:

$$c_i \frac{\left(\sum_{j=1}^{M} c_{i,j}\right)}{M}. \tag{7}$$

3. Design of Temperature Sensor Data Anomaly Diagnosis Model Based on RBF Neural Network and EEMD Modal Analysis

3.1. Data Acquisition and Pretreatment

3.1.1. Data Acquisition. To fully reflect the authenticity of the experiment. The author selected five normal sensors with good performance to collect the experimental data. When the system was put into operation, the sensors were also in normal operation and no fault occurred. At this time, the sampling data sequence within 1 hour of the sensor obtained through time-sharing data acquisition is shown in Table 1.

3.1.2. Data Preprocessing. At present, the longest method in data standardization processing is Z-score standardization, which is also one of the default methods of SPASS. This method is to standardize the data through the mean value (i.e. mean value) and standard deviation (i.e. standard deviation) of the original data set and preprocess the measured data according to the characteristics of the signal transmitted and output by the temperature sensor. The preprocessing

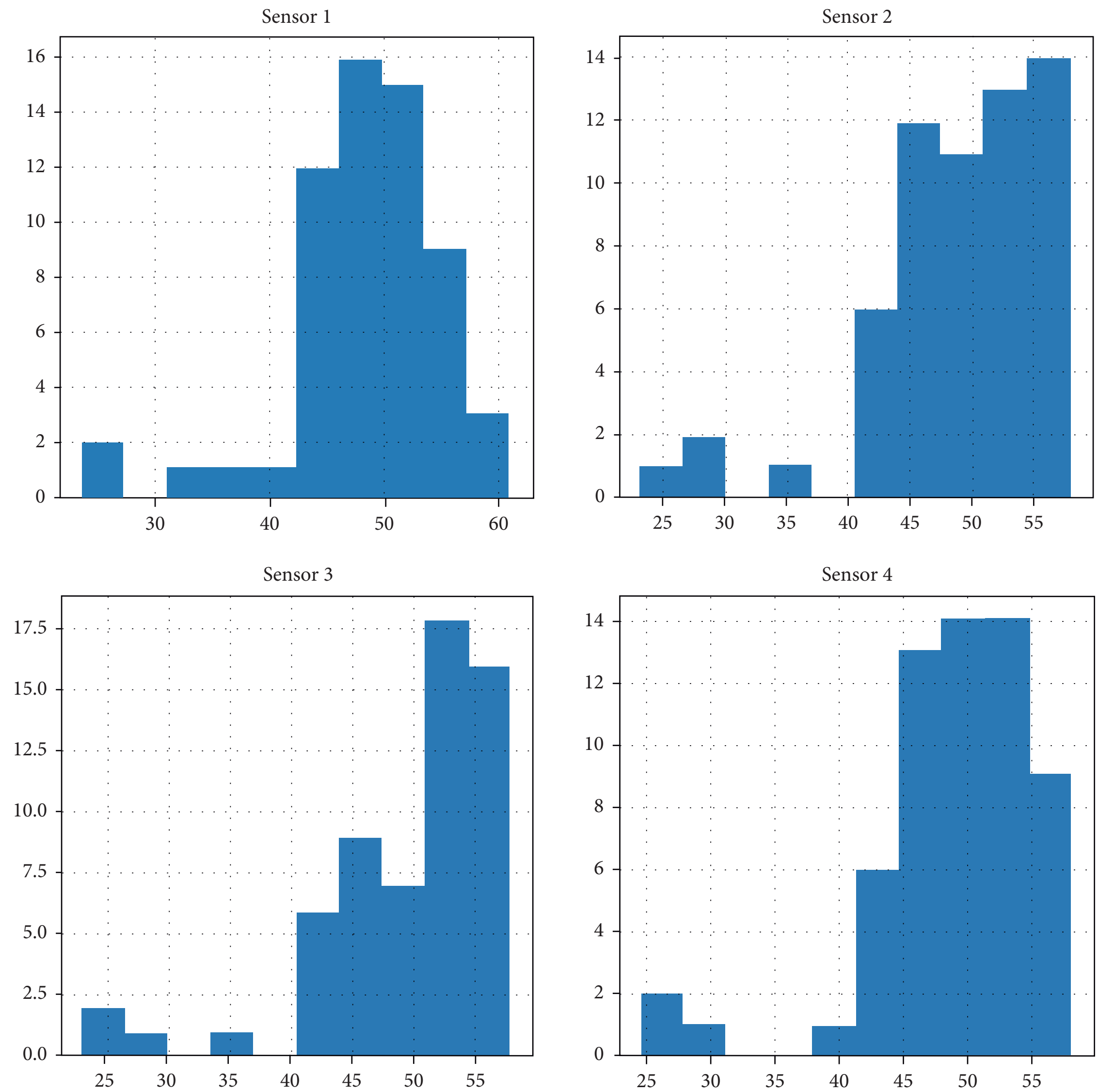


FIGURE 3: Original temperature measurement data.

code is shown in formula (9), and the data processed by this method are shown in Figures 3 and 4. From the law of the data in Figures 3 and 4, it can be seen that the data preprocessed by Z-score standardization only change in the direction of the value of the data, but does not change the size relationship of the data itself, and the data processed by this method can eliminate the influence of the value of the data itself on network training and feature selection, so as to speed up network learning and accurately assign training weights.

$$x_{\text{normalization}} = \frac{x - \mu}{\delta}, \quad (8)$$

3.2. EEMD Abnormal Data Feature Extraction. Taking a sensor with a lot of abnormal data as an example, the residual signal between the predicted temperature and the measured temperature is converted into a continuous time series, as shown in Figure 5. The abnormal data occur in each sequence of sampling points and has the property of random distribution.

The EEMD method is used to decompose the abnormal data features of the results in Figure 5 layer by layer, and the results are shown in Figure 6. Imf1 refers to the high-frequency random noise of the sensor, which fluctuates violently, and has obvious random variability, strong nonlinearity, and insignificant periodicity; Imf2 is the high-frequency periodic component of the output signal of the sensor under external influence; Imf3 refers to the sharp fluctuation of the output signal caused by the abnormal data fluctuation of the sensor; and Imf4 represents the low-frequency periodic component of the sensor output signal caused by the long-term influence of influencing factors.

Follow the same steps to extract the features of the abnormal data of the four temperature sensors used in this paper.

3.3. Model Design. The fault diagnosis modeling of temperature sensor drift fault, accuracy decline fault, impact

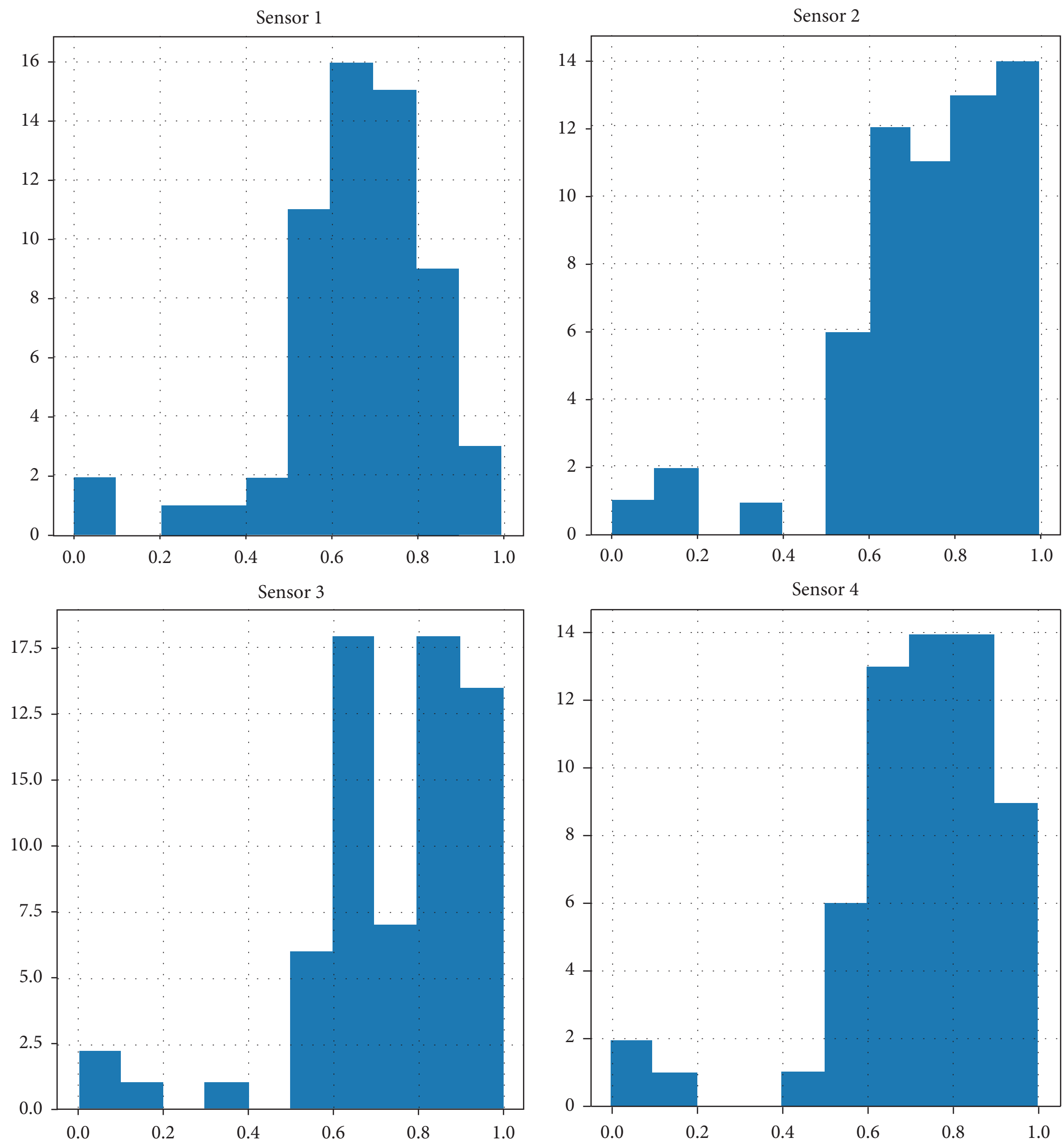


Figure 4: Normalized temperature measurement data.

fault, and fixed deviation fault is mainly composed of three parts: temperature prediction model is based on RBF neural network; EEMD method is used to extract fault features of temperature difference between prediction and measurement; and the anomaly diagnosis scheme based on RBF neural network is shown in Figure 7.

The residual signal between the predicted temperature and the measured temperature still has large fluctuations. When the fault signal is not obvious, it is easy to be hidden by high-frequency fluctuations. The simple threshold discrimination method cannot identify the fault and judge the specific fault type. At the same time, when the temperature prediction accuracy is low, the converted continuous time series fluctuates greatly, which is easy to cause misjudgment. Therefore, the temperature residual signal is decomposed by EEMD, the fault features are extracted from each component, and the fault signal is highlighted to improve the recognition rate.

4. Analysis of Experimental Results

In the test experiment, the normal sensor and the faulty sensor are sampled once every 1 min, respectively, and 240 samples are collected by four temperature sensors within 60 min. Figure 8 reflects the fault transmission signal of an abnormal temperature sensor within 60 min of sampling, and Figure 9 shows the normal output signal of the normal sensor within 60 min of sampling. From the two temperature data transmission signal curves, it can be seen that the failure of the temperature sensor generally occurs 20 minutes after

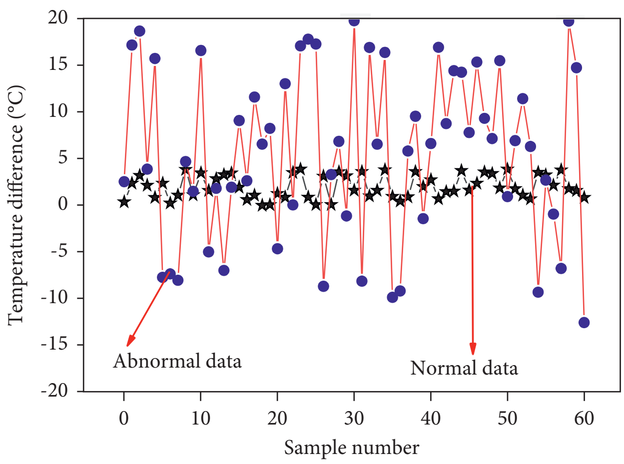


Figure 5: Fixed deviation fault residual signal.

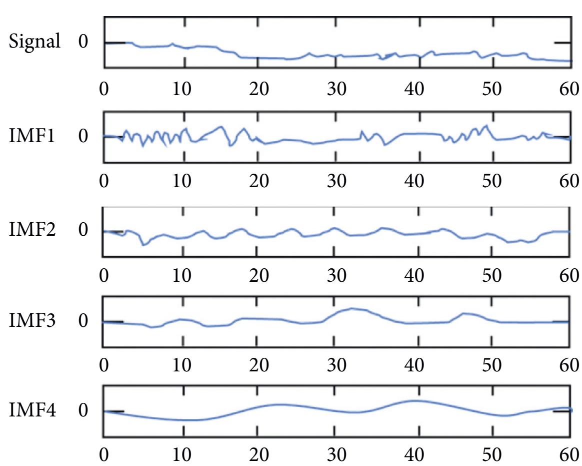


Figure 6: EEMD decomposition effect.

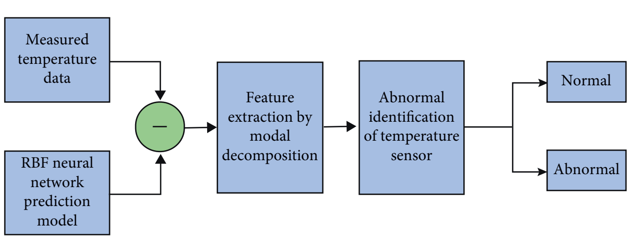


Figure 7: Abnormal diagnosis scheme of temperature sensor based on deep neural network.

the sensor works normally, that is, the sensor temperature rises to a certain temperature, which leads to large fluctuations in the signal value.

At the same time, 240 sets of sample data measured by the above four sensors are used to evaluate the classification effect of combined eigenvectors and RBF neural network. First, input the collected acceleration data and filter, then read the model training parameters, and then quickly classify the data through correlation operation to identify and judge whether the sensor temperature is in an abnormal state. Finally, output the alarm signal according to the actual

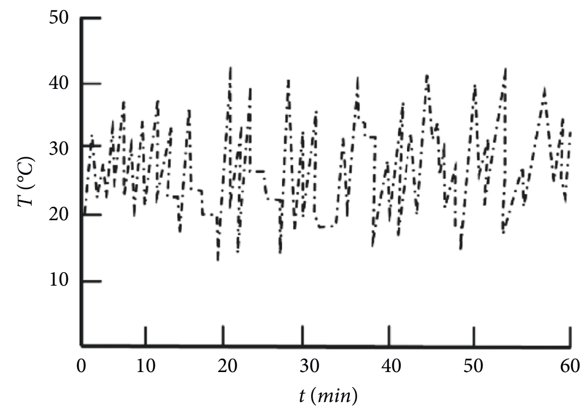


Figure 8: Normal temperature sensor fault transmission signal diagram.

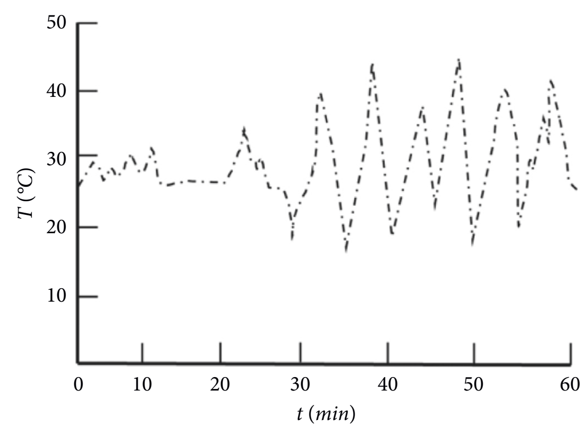


Figure 9: Abnormal temperature sensor fault transmission signal.

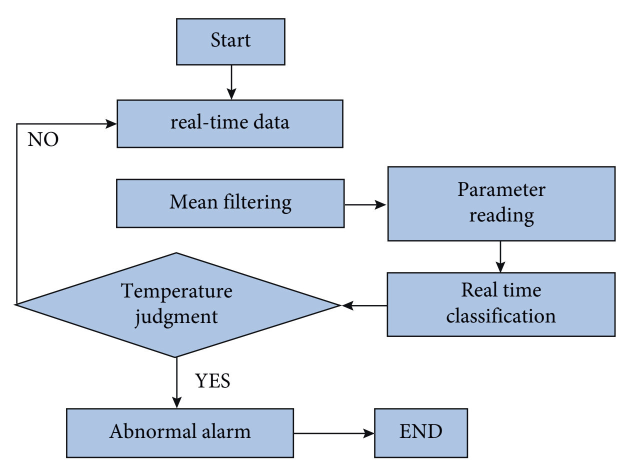


Figure 10: Temperature anomaly detection algorithm flow.

situation of temperature detection. The process of temperature anomaly detection algorithm is shown in Figure 10.

The standard system used to evaluate the effectiveness of temperature sensor anomaly diagnosis algorithm in this paper mainly includes three indicators: accuracy, sensitivity, and specificity. Accuracy refers to the proportion of accurately detecting all abnormal data and nonabnormal data.

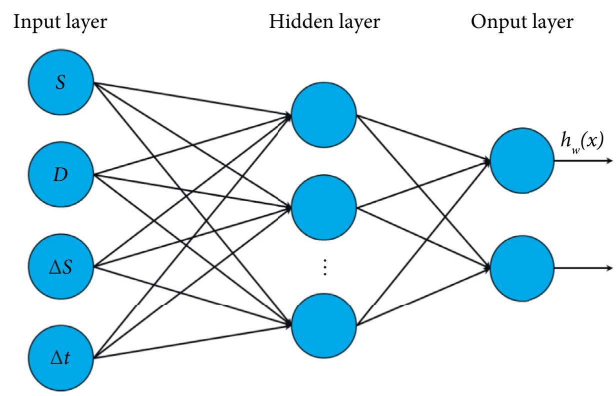


Figure 11: A temperature sensor data anomaly diagnosis model based on the data in this paper.

The higher the accuracy is, the better the effect of this method is; sensitivity refers to the proportion that all abnormal data are accurately detected. The higher the sensitivity, the lower the misjudgment rate; specificity refers to the proportion that all nonabnormal data are correctly detected. The higher the specificity, the lower the misjudgment rate.

$$\begin{aligned} \text{accuracy} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \\ \text{sensitivity} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\ \text{specificity} &= \frac{\text{TN}}{\text{TN} + \text{FP}}. \end{aligned} \tag{9}$$

Among them, TP (real case) is the number of abnormal temperature data events and falls detected, which belongs to correct judgment; FN (false counterexample) refers to the number of abnormal temperature data events but not detected, which belongs to missed judgment; FP (false positive example) is the number of times that no abnormal temperature data event occurs but abnormal temperature data are detected, which is a misjudgment; TN (true counterexample) is the number of times when no abnormal temperature data event occurs and no abnormal temperature data are detected, which belongs to correct judgment. The number of hidden layer nodes of neural network affects the generalization ability and complexity of neural network. If the number of hidden layer nodes is too small, the network will get less useful information, the model description ability is insufficient, and the fault tolerance is poor; too many hidden layer nodes will increase the training time, and the network may store irregular information in the samples, which may lead to the "over fitting" problem and the decline of generalization ability. At present, there is no perfect theory for the selection of the number of hidden layer nodes of artificial neural network, which is mainly verified by numerical value based on previous experience. The input layer of the RBF neural network constructed in this paper contains 4 neurons, the output layer contains 2 neurons, and the number of neurons in the hidden layer is 200 according to the initial value, 20 steps, and 20 iterations. The number of numerical experiments is finally determined to be 280. The abnormal temperature data diagnosis model formed by this is shown in Figure 11. The test results are shown in Table 2.

Table 2: Test experimental results.

Sensor	Number	TP	FN	FP	TN
1	60	0	2	1	57
2	60	0	0	0	60
3	60	54	1	1	4
4	60	58	0	1	1

As can be seen from Table 2, a total of 240 samples were tested. According to the defined formula, the accuracy, sensitivity, and specificity of the algorithm are 97.5%, 96.8%, and 97.9%, respectively. The test results confirm the effectiveness, accuracy, and feasibility of the algorithm.

5. Conclusion

With the promotion of household appliances and industrial machinery, temperature, as one of the indicators of normal operation of various machinery and electrical appliances in the process of use, also occurs from time to time due to equipment damage and huge economic losses caused by sensor failure. This paper analyzes the above problems and draws the following conclusions:

(1) Based on RBF neural network and empirical mode decomposition (EEMD) deep neural network and based on the historical fault data samples and normal historical data of temperature sensors, an intelligent temperature data anomaly diagnosis method is proposed to solve the temperature anomaly problem that is easy to occur in temperature sensors at present.

(2) The sensor in actual work is taken as the object, and the measured data are used as the input of the fault model to diagnose and monitor the temperature abnormality. Through the diagnosis and verification of the measured temperature sensor data, the established model has a good fault diagnosis and classification ability, with an accuracy rate of 97.5%, a sensitivity of 96.8%, and a specificity of 97.9%. It shows that the method constructed in this paper can effectively provide data abnormal diagnosis services for the temperature sensor.

Data Availability

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

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] P. Chen, C. C. Chen, C. C. Tsai, and W. F. Lu, "A time-to-digital-converter-based CMOS smart temperature sensor,"

IEEE Journal of Solid-State Circuits, vol. 40, no. 8, pp. 1642–1648, 2005.

[2] B. Feizifar, Z. Muller, and O. Usta, "A new algorithm for detecting failure to clear mode of circuit breakers using fundamental frequency component of voltage signals," *IEEE Transactions on Power Delivery*, no. 99, p. 1, 2019.

[3] G. Pham and Ritter, "On-chip sensors to detect impending metallization failure of LDMOS transistors under repetitive thermo-mechanical stress," *IEEE Transactions on Semiconductor Manufacturing: A Publication of the IEEE Components, Hybrids, and Manufacturing Technology Society, the IEEE Electron Devices Society, the IEEE Reliability Society, the IEEE Solid-State Circuits Council*, vol. 29, no. 3, pp. 193–200, 2016.

[4] C. D. Breach, F. Wulff, and C. W. Tok, "An unusual mechanical failure mode in gold ballbonds at 50μm pitch due to degradation at the Au–Au4Al interface during ageing in air at 175°C," *Microelectronics Reliability*, vol. 46, no. 2-4, pp. 543–557, 2006.

[5] H. L. Massie, "Channel Utilization by Remote Locomotive Control Systems Using Digital transmission," in *Proceedings of the Vehicular Technology Conference*, IEEE, Detroit, MI, USA, September 1976.

[6] R. Niwas and B. Singh, "Solidstate controller for threephase permanent magnet synchronous generatorbased diesel generator set feeding singlephase loads," *IET Power Electronics*, vol. 8, no. 10, pp. 1844–1852, 2015.

[7] Y. P. Zhao, Y. L. Xie, and Z. F. Ye, "A New Dynamic Radius SVDD for Fault Detection of Aircraft engine," *Engineering Applications of Artificial Intelligence*, vol. 100, 2021.

[8] A. Msi, B. Yakn, A. Uak, and B. W. Lee, "Real-time fault detection system for large scale grid integrated solar photovoltaic power plants," *International Journal of Electrical Power & Energy Systems*, vol. 130, 2021.

[9] A. Hd, A. Mh, and B. Wpa, "DC Grids DC Fault Detection Using Asymmetric Pole inductors," *Electric Power Systems Research*, vol. 193, 2021.

[10] B. Li, C. Delpha, D. Diallo, and A Migan-Dubois, "Application of Artificial Neural Networks to photovoltaic fault detection and diagnosis: a review," *Renewable and Sustainable Energy Reviews*, vol. 138, Article ID 110512, 2021.

[11] S. Duan, W. Yang, X. Wang, S. Mao, and Y. Zhang, "Forecasting of Grain Pile Temperature from Meteorological Factors Using Machine Learning," *IEEE Access*, vol. 7, 2019.

[12] S. Kurpaska, Z. Ślipek, B. Bożek, and J. Fraczek, "Simulation of heat and moisture transfer in the greenhouse substrate due to a heating system by buried pipes," *Biosystems Engineering*, vol. 90, no. 1, pp. 63–74, 2005.

[13] C. S. Chang, H. H. Converse, and J. L. Steele, "Modeling of temperature of grain during storage with aeration[J]," *Transactions of the Asae*, vol. 36, no. 2, pp. 509–519, 1993.

[14] S. Szke, V. Wertz, and E. Persoons, "Outdoor temperature modelling for grain cooling control," *IFAC Proceedings Volumes*, vol. 28, no. 6, pp. 47–51, 1995.

[15] L. Fan, Y. Ji, and G. Wu, "Research on temperature prediction model in greenhouse based on improved SVR," *Journal of Physics: Conference Series*, vol. 1802, no. 4, Article ID 042001, 2021.

[16] Y. Wu, Q. Lan, and Y. Sun, "Application of BP Neural Network Fault Diagnosis in Solar Photovoltaic system," in *Proceedings of the 2009 International Conference on Mechatronics and Automation*, IEEE, Changchun, China, August 2009.

[17] Y. DONG, X. Li, and Cheng, "Application of BP neural network in network fault diagnosis," in *Proceedings of the IOP Conference Series Materials Science and Engineering*, Changchun, China, August 2009.

[18] Y. Jiang and F. Zhang, "Study on BP neural network optimization by improved decay parameter genetic algorithm," *Journal of Physics: Conference Series*, vol. 1621, no. 1, Article ID 012054, 2020.

[19] J. Rong and H. Ge, "Hydroelectric generating unit vibration fault diagnosis via BP neural network based on particle swarm optimization," in *Proceedings of the 2009 International Conference on Sustainable Power Generation and Supply*, IEEE, Nanjing, China, April 2009.

[20] L. Quan and X. Jiang, "Sensor fault diagnosis based on discrete wavelet transform and BP neural network," *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 5998, 2005.

[21] Z. R. Hou, "Rolling bearing fault diagnosis based on wavelet packet and improved BP neural network for wind turbines," *Applied Mechanics and Materials*, vol. 347-350, pp. 117–120, 2013.

[22] Z. Feng, X. Zhang, and Y. He, "Research of pneumatic actuator fault diagnosis method based on GA optimized BP neural network and fuzzy logic," in *Proceedings of the International Conference on Advances in Neural Networks*, Springer, Berlin, Germany, July 2013.

[23] G. Liu, X. Wang, and F. Li, "Fault diagnosis of diesel engine fuel system based on genetic algorithm optimizing BP neural network model and D-S evidence theory," *An International Journal of Research & Surveys.Part B Applications*, vol. 3, pp. 707–716, 2012.

[24] W. M. Lin, C. D. Yang, J. H. Lin, and M. Tsay, "a fault classification method by RBF neural network with OLS learning procedure," *IEEE Power Engineering Review*, vol. 21, no. 8, p. 60, 2001.

[25] X. Li, Z. Tao, and D. Zhe, "A Recognition Method of Plate Shape Defect Based on RBF-BP Neural Network Optimized by Genetic algorithm," in *Proceedings of theThe 26th Chinese Control and Decision Conference*, Changsha, China, May 2014.

[26] H. Wang, W. Jing, Y. Li, and H. Yang, "fault diagnosis of fuel system based on improved extreme learning machine," *Neural Processing Letters*, vol. 53, no. 4, pp. 2553–2565, 2021.

[27] S. Guo, B. Wu, J. Zhou et al., "An analog circuit fault diagnosis method based on circle model and extreme learning machine," *Applied Sciences*, vol. 10, no. 7, p. 2386, 2020.

[28] Y. Xia, Q. Ding, A. Jiang, N. Jing, W. Zhou, and J. Wang, "Incipient fault diagnosis for centrifugal chillers using kernel entropy component analysis and voting based extreme learning machine," *Korean Journal of Chemical Engineering*, vol. 39, pp. 1–11, 2022.

[29] C. Li, R. Yu, and T. Wang, "The Research of Fault Diagnosis of Nuclear Power Plant Based on ELM-AdaBoost. SAMME," *Science and Technology of Nuclear Installations*, vol. 2020, Article ID 6689829, 9 pages, 2020.

[30] X. Liu, H. Huang, and J. Xiang, "A personalized diagnosis method to detect faults in gears using numerical simulation and extreme learning machine," *Knowledge-Based Systems*, vol. 195, Article ID 105653, 2020.

[31] K. L. Man and A. Bermak, "A Time Domain Differential CMOS Temperature Sensor with Reduced Supply Sensitivity," in *Proceedings of the IEEE International Symposium on Circuits & Systems*, IEEE, Seattle, WA, USA, May 2008.

[32] L. Jiao, S. Yang, and F. Liu, "Seventy years of neural network: review and Prospect," *Journal of Computer Science*, vol. 39, no. 8, pp. 1697–1717, 2016.

[33] Z. Liu, H. Wang, and S. Peng, "Texture segmentation using directional empirical mode decomposition," in *Proceedings of the Image ProcessingICIP '04. 2004 International Conference on*, IEEE, Singapore, October 2004.

[34] M. HAFIDA and R. A. I. S. Mahgoun, "Gearbox Fault Diagnosis Using Ensemble Empirical Mode Decomposition (EEMD) and Residual signal," *Mechanics & Industry*, vol. 13, 2012.

[35] A. A. Tabrizi, L. Garibaldi, A. Fasana, and S. Marchesielo, "Ensemble Empirical Mode Decomposition (EEMD) and Teager-Kaiser Energy Operator (TKEO) Based Damage Identification of Roller Bearings Using One-Class Support Vector Machine," in *Proceedings of the Ewshm 2014*, Nantes, France, July 2014.

[36] J. C. Nunes, Y. Bouaoune, E. Delechelle, O. Niang, and P. Bunel, "Image analysis by bidimensional empirical mode decomposition," *Image and Vision Computing*, vol. 21, no. 12, pp. 1019–1026, 2003.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9876931, 1 page
https://doi.org/10.1155/2023/9876931

Retraction

Retracted: Analysis of the Dilemmas and Countermeasures Brought by Data Analysis Based on Short Online Videos on Civic and Political Work

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. Xie, "Analysis of the Dilemmas and Countermeasures Brought by Data Analysis Based on Short Online Videos on Civic and Political Work," *Mobile Information Systems*, vol. 2022, Article ID 3367257, 10 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 3367257, 10 pages
https://doi.org/10.1155/2022/3367257

Research Article

Analysis of the Dilemmas and Countermeasures Brought by Data Analysis Based on Short Online Videos on Civic and Political Work

Junhong Xie

School of Marxism, Beijing Technology and Business University, Beijing 100048, China

Correspondence should be addressed to Junhong Xie; 2031301007@st.btbu.edu.cn

Received 21 June 2022; Accepted 11 August 2022; Published 31 August 2022

Academic Editor: R. Mo

In the new era of Internet development, short video media has emerged and become the second largest network application in China. While bringing opportunities for the propaganda work of civic and political science, short video also brings difficulties and challenges such as complex network information interfering with the attention of netizens, fragmented communication method not conducive to systematic communication, and algorithmic recommendation technology causing barriers to the reception of information by netizens. In the case that the media short video and the political work need to be constantly updated, we should deal with the short video challenge brought by the media, effectively avoid its negative impact, actively deal with the period of time since the media, and make use of the short video to carry out various forms of educational reform. Facing these challenges, this paper focuses on helping the development of political work in the new era from the aspects of talent training, content optimization, brand building, algorithm standardization, etc.

1. Introduction

Short video is an emerging audio-visual media communication method that "uses mobile smart terminals as the communication carrier, relies on mobile social platforms and social chains, and has a playback length between a few seconds and a few minutes" [1–3]. Short video applications are developing rapidly, and various application platforms are vigorously developing network video technology. In August 2013, Sina Weibo's mobile client is built in the "second shot" application, which rapidly increased users [4]. However, the short video applications at this time were still in the exploration stage of software development, and most of the products were in a lukewarm state. In 2016, short video apps such as Jitterbug appeared one after another, and the short video industry was becoming mature at this time, with the advantages of prominent information points, low threshold of user intervention, and large amount of UGC (user-generated content), which quickly set off a short video boom [5, 6]. The booming development of online short video industry has provided great convenience for people to receive information. In this context, whether to make good use of the short video network to do ideological and political work has become a new issue of the times.

Since the media has created a good information dissemination environment, the short video of the media has had a very wide impact on students with its powerful audio-visual effect. Strengthen ideological construction through various forms of innovation, so that every student can firmly establish the ideal of communism. Focusing on cultivating and guiding the core socialist values, we should use the mass media, give full play to the technical and communication advantages of short films by carrying out ideological education for students, make use of the mass media, and create an effective network environment for better ideological exchange among students by carrying out broader ideological education and mode construction [7]. The use of media short films to develop various forms, so that new forms of education can serve higher political education institutions, and the use of media short films to timely understand students' ideological trends learn students' words and deeds, and make full preparations for effective intellectual education.

In today's booming new media technology, it is very important to actively use online short videos to strengthen ideological and political work, promote socialist core values, and build a socialist culture with Chinese characteristics [8]. First of all, short videos on the Internet are conducive to expanding the scope of Marxism's dissemination. Today, short video applications have become the second largest Internet application in China, and short videos have become the main channel for users to access information. Under such circumstances, actively using short videos to disseminate mainstream ideological content is conducive to delivering propaganda content to more users and expanding the scope of mainstream ideology.

Secondly, short videos are good for improving communication effect. Compared with traditional media, short videos combine dynamic video, music, and propaganda text when disseminating information, which has the advantages of prominent information points, low comprehension cost, and strong user participation. The use of short videos to disseminate mainstream ideological content can greatly enhance the fun, interactivity, and timeliness of the propaganda content, thus improving the communication effect.

Finally, short videos are conducive to the establishment of a good network opinion ecology. Nowadays, with the continuous economic growth and the development of social pluralistic values, many kinds of ideas have started to spread new speech through the Internet, and there is no lack of wrong ideas to take advantage of the opportunity to spread, which brings interference for us to do a good job in the field of ideology. In the face of the challenges brought by undesirable ideas, the active use of Internet media, represented by short videos, for ideological propaganda and public opinion guidance is conducive to our seizing the high ground of ideological propaganda and establishing a good Internet public opinion environment [9].

2. Related Work

Short video self-media is a kind of expression of self-media, the application of self-media is becoming more and more extensive, the innovation of self-media based on various new technologies is also developing, and the self-media with short video as the main bearing form has very strong technical advantages, bringing audiences a more intuitive audio-visual effect and a very strong communication influence. In the new context of the rapid transformation of the text era to the graphic era and the video era, short video self-media has become popular, and the coverage of short video self-media represented by Jitterbug, Racer, and Nengxun video number is getting bigger and bigger, gradually replacing the traditional text expression. Today, Racer and Jitterbug have become the main forms for people to better display their lives and socialize, and they have also become one of the most important carriers of information dissemination. The application of smartphones is becoming more and more popular, 4G phones are very mature, and with the rapid construction of 5G communication technology, short video self-media has better technical support and equipment guarantee, video self-media is more widely used, and its communication advantages are becoming more and more prominent, which has replaced microblogs and other as the most popular self-media for college students at present [10].

Firstly, the study of online moral education is the theoretical basis for the study of short online videos. Since short web videos are an emerging medium formed, thus writings on the influence of short web videos on moral education are very rare, but short web videos, as a new-age derivative of network technology, have commonalities between the influence of short web videos on students' moral education and the development of online moral education [11]. The network can link the subject of moral education with the subject. This shows that the network has become a new carrier of moral education [12]. The virtual society of the Internet posing a comprehensive challenge to the content, methods, and approaches of traditional moral education has also brought great development and opportunities for realistic moral education innovation and injected new vitality and vigor into moral education work [13]. Each person is certainly exposed to different online environments due to the influence and limitations of their own or in exchange for several conditions, which inevitably affects the nature, manner, type, and degree of their acceptance of online value influence.

Secondly, the research on the connotation of short online videos is also the starting point of short video moral education impact research. Domestic scholars' research on short Internet video mostly focuses on the following points: 1. defining short video in terms of the length of the video; [14] believes that short video refers to the length of no more than 20 minutes and is collectively known as shooting, editing, uploading, playing, sharing, and interacting through short video platforms, covering short films such as documentary [15].

Thirdly, the research on the impact of short online videos on students is the starting point for studying the issue of short online videos. The problem of students' use of short online videos is the practical reason that has aroused scholars' concern about short online videos. Taking Jitterbug as an example, [16] argued that short video platforms have loopholes in the regulatory mechanism and have created the problem of users making and selling fakes and exporting distorted values. [17] argued that online short videos have a negative impact on the values of youth, creating an atmosphere of "entertainment first," shaking the status of teachers, and delaying the development of students' behavioral norms.

Fourth is the research on the development trend of short online video. In [18], it is pointed out in "The Current Situation, Problems and Trends of Short Video Development in China" that the development of technology will make short videos combine with VR to enhance the immersion of audiences. In [19], it is pointed out in "New Media Communication Innovation and Development Paths" that short video will become a new pole of the "pan-documentary content ecology" as it will move toward boutique and verticalization, and the business model will be further improved and technology-driven factors will be more obvious. In [20], "The Development Status and Research Trends of Netflix Short Video," it is argued that the competition pattern of short video as a whole is gradually forming, content will become the core competitiveness, content will shift toward production specialization, pan-entertainment to vertical fields, and the business model of short video will be gradually clarified in the future.

Fifthly, the research on countermeasures to the problem of short online videos for adolescents, which is the ultimate landing point and destination for conducting research on the impact of short online videos on moral education [21] in "Countermeasures for College Students' Internet Moral Failure," pointed out that college students should be guided to treat and use short online videos correctly and provided countermeasures by further consolidating the basic position of family education, further strengthening self-education, strengthening school management and education, further strengthening and improving government supervision, and strengthening and improving self-management of Internet service platforms. In [22], the problems, countermeasures, and directions of short video development emphasize that all audiences should strengthen the supervision of short video content, highlight the role of leaders, and face up to the role of the third person.

In summary, it can be seen that although the current academic community has a certain basis, a complete system has not yet been formed. In addition, Chinese scholars still have the following problems when conducting research: first, lack of knowledge connection with the original theory. Most scholars in the process of conducting research on the problem of short online videos only single from the current situation to analyze the current situation and causes of the problem of moral education of short online videos at the present stage are not able to analyze and link the research on moral education of short online videos with the original scholars' research on moral education of the network in a longitudinal way through historical logical analysis, which makes the research on moral education of short online videos too isolated and lacking the proper theoretical basis [23]. Second, the research on short online videos and students' moral education are separated from each other, and they are not able to accurately discuss the relationship between them. Among the existing research results, the research on short online videos is mostly focused on the perspective of media science, and there are fewer research results in the direction of students' moral education [24]. Third, the students' short online videos are mainly theoretical research, lacking relevant data and practical research, lacking quantitative analysis, and with little credibility and persuasiveness [25]. Meanwhile, the previous studies in the results often simply attribute online problems to the environment or students' own subjective factors, without being able to argue the deep-seated causes of these problems in the context of the general social background and economic development changes.

3. Data Analysis to Study the Changes of Civic Science and Politics Work in the Short Video Environment

Figure 1 shows the ideas of political construction work under the short video environment.

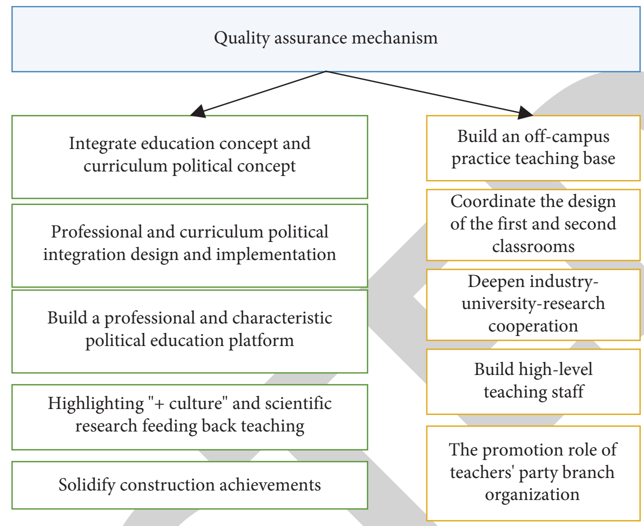


Figure 1: Ideas for political construction work.

In the face of the impact of short online videos is that each learner's learning behavior has its own unique characteristics as well as learning strategies, which are generally influenced by both intrinsic factors (e.g., frequency of retrieving and reading course materials, degree of participation in discussions, etc.) and extrinsic factors (e.g., sociodemographic profile). In this study, student sociodemographic data were provided by the Student Information System (SIS), and learning behaviors were obtained from the LMS. In this paper, we use exploratory data analysis (EDA) and machine learning methods to obtain the same characteristics in the student population patterns in order to discover the patterns of student behavior characteristics. Exploratory data analysis (EDA) is a common data processing method used in big data technology to analyze data sets to generalize their main features. EDA is a data analysis method that explores existing data with as few a priori assumptions as possible, exploring the structure and patterns of the data by graphing, tabulating, and calculating feature quantities. In this paper, we use clustering, association, and network graphs to obtain knowledge about the relevant features of a data set to obtain patterns of student behavior.

Based on the study of a large number of relevant literature, this paper fully understands the development trend of online video, online ethics education, students' moral education, and the current situation of students' use of online video, and analyzes its impact on online video. It is divided into four dimensions: "students currently use video clips, the impact of video clips on students' development, the negative impact of video clips on children's development." The detailed sample of this analysis is shown in Table 1. According to the statistical results, in this survey, 47.0% of the gender samples are male and 53.0% are female, with a more balanced overall distribution: in terms of grade distribution, 37.1% of the senior class, 41.2% of the sophomore class, and 21.7% of the senior class, with the largest proportion of the sophomore class and the least of the junior class. Among the students, 37.9% are undifferentiated, 32.7%

Table 1: Sample situation of the study.

Gender	Male	Female		Total
Quantity	234	264		498
The proportion	47.0%	53.0%		100%
Grade	Senior year	Sophomore	Senior	Total
Quantity	185	205	108	498
The proportion	37.1%	41.2%	21.7%	100%
Family	Liberal arts	Science	Unspecialized	Total
Quantity	163	146	189	798
The proportion	32.7%	24.4%	37.9%	100%

are arts students, and 24.4% are science students. The above situation shows the distribution of gender, grade distribution, and student categories in this survey, which can reflect the representativeness of the sample, and the survey results can reflect the current situation of students' use of online short video and its influence on students at this stage.

Through the analysis of the data on the whole, it can reflect the current situation of contemporary students' use of short videos in many aspects and at many levels, and reflect the disadvantages and effects produced by students' use of short online videos, so that through the analysis of these data, effective countermeasures can be sought. The diversity of information contained in network short films also brings severe challenges to high school moral education. Through empirical analysis and investigation, this paper uses the method of network short film to analyze and describe the situation of students in detail. By compiling the data on students' viewing of short online videos, it can be found that students' moral education in general remains in the mainstream direction in the use of short videos as in the following aspects.

3.1. Students' Character Has Some Self-Control Ability. When examining the time students watch short videos, there is this one to question: do you watch short videos in class. See Table 2.

According to the data shown in Table 2, 83.1% of the surveyed students chose "no," which indicates that contemporary students have a strong self-control and sense of responsibility for their own learning, and only 3.8% of the students said they would watch short videos in class, which indicates that contemporary students have a rational perception of current learning and entertainment, can fully understand the relationship between current learning and their future development goals, and have a high self-control ability.

3.2. Positive Development of Values. When examining students' value judgments, the question "Which type of short videos do you like more (multiple choice)" was set, taking into account students' characteristics and the classification of short videos. The survey results are shown in Table 3. 42.7% of the students chose film and television, 48.2% chose spoof, 63.5% chose positive energy, 71.9% chose situational skits, 53.8% chose skill sharing, and 39.2% chose life sharing. This result fully indicates that contemporary students have more correct values and ability to distinguish, and can choose positive short videos that are beneficial.

In addition, the question "What kind of content do you want the short video platform to present more (multiple choice)" was set for students' suggestions on short videos. In the news reports, the proportion of news comments is 53.3%, that of traditional culture is 30.2%, and that of international diplomacy is 16.5%. This result shows that students have a strong sense of home and country, and can relate themselves to national development and international diplomacy. At the same time, there is a strong learning initiative.

The above data analysis fully illustrates that contemporary students' values have a benign development state. They have certain ability to discriminate short videos with diversified contents and forms, have strong learning power, have certain family sentiment, and can link personal development with national development.

When confronted with the question, "What type of videos do you want the short video platform to provide (multiple choice)," 75.8% of students chose "news and commentary," 65.3% chose "entertainment and funny," 78.1% chose "knowledge and explanation," and 69.6% chose "movie and entertainment." This result shows that as high school students, they have a strong motivation to acquire knowledge and grow their ability through short video platforms. This motivation is also a prerequisite for guiding students' correct values through the short video platform.

A question was designed in terms of the influence on students in short videos, "Do you think that socialist core values can be promoted and achieve good influence through the short video format?," the results of the survey as in Table 4.

According to the results in Table 4, 57.0% of students think that short videos can be used to promote socialist core values and think that they will have a good impact on students' moral education. It can be seen that at this stage, most students in the student body are willing to accept the guidance of core values and are able to use short video resources to watch video contents that are beneficial to their physical and mental growth.

Ideals and beliefs are the motivation that guides students to study hard and strive for advancement at this stage, and they are also the important support that guides students to make unremitting efforts to achieve the great rejuvenation of the Chinese nation. In order to examine the students'

Table 2: Statistical results.

Question Options	Do you watch short videos in class? Content	Number of people (person)	Percentage
1	Won't	414	83.1%
2	I would like to watch it, but I will insist on watching it after class or during holidays	65	13.1%
3	Yes, and can't help but watch it in class	19	3.8%

Table 3: Statistical results.

Question Options	Which type of short videos do you like more (multiple choice) Content	Number of people (person)	Percentage
1	Film and television	215	43.2%
2	Spoof	240	48.2%
3	Positive energy	316	63.5%
4	Situational skits	358	71.9%
5	Skill sharing class	268	53.8%
6	Life sharing	195	39.2%

Table 4: Statistical results.

Question Options	Do you think the socialist core values can be publicized in the form of short videos and have a good impact? Content	Number of people (person)	Percentage
1	Yes, it will have a big impact	284	57.0%
2	No, it will not affect	65	13.1%
3	Not sure	149	29.9%

firmness of ideal beliefs, the question was set as "Do you think the short video platform can provide the right guidance for people to establish ideal beliefs" as in Table 5.

According to the results in Table 5, 59.4% of students chose "yes," 24.1% chose "no," and 16.5% chose "not sure." The short video platform is important in promoting positive energy and guiding students to firm ideals and beliefs.

In the survey of the question "Do you understand the promotion of national policies through short videos," 71.3% of the students chose "Yes, very much." This shows that the younger generation of students has a strong desire to understand the national policy and wants to show the national policy in a new way.

The above survey questions show that contemporary students are able to maintain a firm political direction, firm ideals and beliefs, and a sense of responsibility for the country, and are able to use the short video platform for effective political learning.

Legal consciousness is people's heartfelt recognition, reverence, compliance, and obedience to the law. There are 2 questions designed for this issue.

According to the survey on students' awareness of rights, we set the question "What would you do if there is an illegal phenomenon in the Internet short video." As shown in Table 6 24.9% of students chose to "tell their parents." These data show that contemporary students have a stronger sense of the legal system and are able to use their legal rights to defend their rights in matters that are closely related to them.

3.3. Confidence in the Future Construction of the Rule of Law in China. When examining the students' confidence in the future construction of the rule of law in China, we set the question "Do you think the country can control the illegal phenomenon of short video platforms well?"

From Table 7, 79.3% of the students chose yes, 2.4% chose no, and 18.6% chose not sure. This result shows that the majority of students at this stage have strong confidence in our country's national governance short video platform.

3.4. Agree with the Measures of National Governance of Short Videos at This Stage. Due to a series of problems with short videos, the state has taken some control measures, which are important for regulating the behavior of short video platforms, creating a healthy network environment for people, and forming healthy values for people. To address this issue, the following question was set: "Do you agree with the current measures taken by the state to regulate short video platforms" as in Table 8.

From Table 8, 63.7% of the students chose to agree, 13.3% chose to disagree, and 23.1% chose to be unsure. This result shows that students at this stage generally agree with the current measures of state governance of short video platforms and are confident about the future development direction of short video under state governance.

With the strengthening of China's legal system, students have a certain degree of popularity for the law, and there is also some popular legal knowledge in short videos and more

Table 5: Statistical results.

Question Options	Do you think the short video platform can provide correct guidance for people to establish their ideals and beliefs?		
	Content	Number of people (person)	Percentage
1	Yes	296	59.4%
2	No	120	24.1%
3	Not sure	82	16.5%

Table 6: Statistical results.

Question Options	What would you do if you broke the law in a short video on the Internet?		
	Content	Number of people (person)	Percentage
1	Report according to law	298	59.9%
2	Tell parents	124	24.9%
3	Ignore it	76	15.3%

Table 7: Statistical results.

Question Options	Do you think the state can control the illegal phenomenon of short video platforms?		
	Content	Number of people (person)	Percentage
1	Yes	395	79.3%
2	No	12	2.4%
3	Not sure	91	18.6%

Table 8: Statistical results.

Question Options	Do you agree with the measures taken by the current national standard short video platform?		
	Content	Number of people (person)	Percentage
1	Agree	317	63.7%
2	Disagree	66	13.3%
3	I don't know	115	23.1%

on the popularity of traffic knowledge. According to the survey, there are still some problems, combined with the actual situation analyzed as follows.

3.4.1. Insufficient Self-Control and Lack of Self-Discipline. This is the most serious problem of students in the process of using short videos at this stage. According to the questionnaire data, students use short videos because of the settings of short video software, vertical screen, and direct display of new contents through the slide of fingers, and the content that appears is completely different from the former. This makes students refresh constantly driven by curiosity on the one hand and makes them indulge in the entertainment of short videos in pursuit of mental relaxation on the other.

3.4.2. Diverse Short Video Contents Lead to Confusion of Students' Values. Due to more short video applications, publishers, censorship mechanism has not been fully established, etc., students are prone to see some contents that are contrary to the mainstream values in the process of watching short videos; as the table shows, there are mainly early marriage and pregnancy, thick and dark school of interpersonal relationship handling, violence, and pornography. These short videos with poor content seriously erode the students' ideological health. And some videos that promote foreign values are easy to make students worship foreign values, leading to the weakening of socialist core values among students.

To further analyze the negative impact of online short films on students and ensure that they conform to the reality of students' modern life, a questionnaire survey on improving students' moral education methods is formulated. It is hoped that through the analysis of this form, combined with the characteristics of students and the current situation of the use of short videos, the best strategy to solve the negative impact of network short videos on students will be discussed. See Table 9.

4. Case Analysis

4.1. Data Sources. In order to further investigate whether the countermeasures in this paper are real and effective, we selected the situation of students' civic learning after the influence of short online videos to conduct the analysis. In recent years, studies have been conducted to show that data quality is important. This demonstrates the importance of

Table 9: Survey results.

Question Options	What measures do you think should be taken to deal with the adverse effects of short videos? (multiple choice)		
	Content	Number of people (person)	Percentage
1	Strengthening the construction of macro-environment for the moral development of high school students	319	64.1%
2	Attach importance to the main position of moral education for high school students and strengthen the guidance of school values	368	73.9%
3	Pay attention to guide students to educate themselves	417	83.7%

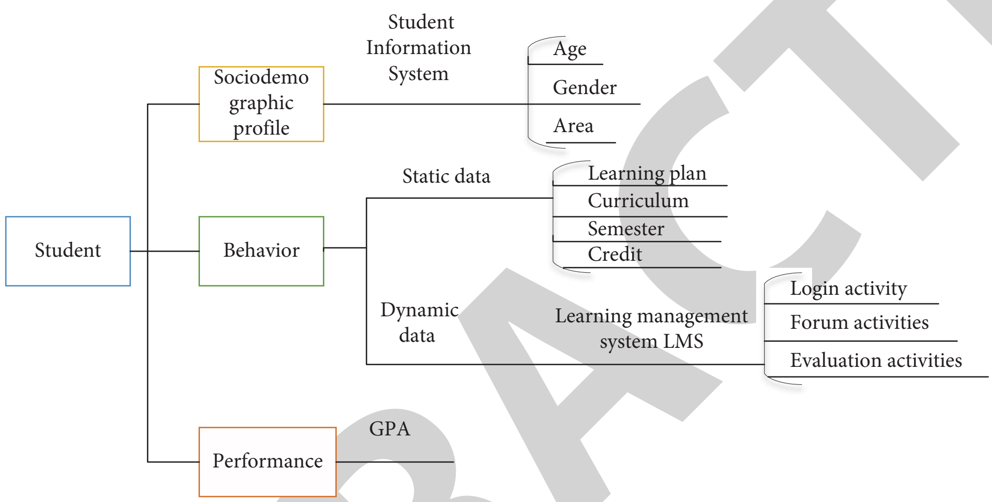


Figure 2: Structure of the data set.

data sets that provide accurate process presentations and comparisons of results. The data set in this paper is derived from data of more than 300 students from four classes of a course, from different regions, ages, and grades. The first step of EDA is to analyze the LMS logs, which are composed of various data and contain data. In this study, we select only the operation logs of student users and do not study the operation logs of other users.

The data collected and analyzed in this study are from a course under the influence of short online videos, selected from more than 300 students in four classes. The large and diverse nature of this student online learning data requires the use of special technical methods to analyze the data. Figure 2 represents the structure of the data set used in this study.

In this paper, the online learning activities of these 300 students in the first sixteen weeks were selected as the sample. The students' final grades depend on the number of times they participated in the online course, the discussion activities, and the completion of the assignments.

4.2. Data Analysis. The data are obtained through the LMS and downloaded from the Student Information Management System. Three different types of data are usually distinguished. Data are divided into static data and dynamic data. Table 10 lists the data sets used as inputs in this study.

In the first phase of this research experiment, we found associations between sociodemographic characteristics and students' online learning activities in the quantitative data, and based on the collected data, we used the Spearman correlation method. Based on the correlation analysis, we found some patterns in students' weekly logins: students were more active in the first week of online learning. The students' activity tends to be negatively skewed from the ninth week to the twelfth week, and the worst activity is in the twelfth week, as shown in Figure 3. Then, we further analyzed the data using correlations to find the link between student activity and student achievement.

Next, we use clustering algorithm to obtain potential standard features from the collected data. Multilevel algorithm is the simplest uncontrollable algorithm, which is often used to deal with big data. K-means algorithm is used in this work, because it is efficient and easy to use. Before using the k-means algorithm, the data need to be converted into digital values and normalized. In terms of data standardization, the most important step is to determine the k. If the initial manifold is bad, then the manifold will be wrong. In this paper, the elbow method is used to determine the optimal number of classes in the unit comparison. As shown in Figure 4, the result of this process shows that the optimal value of the cluster is 3, so the initial value of the cluster in this paper is 3 ($k = 3$). In addition, Euclidean distance can be used to determine the

Table 10: Input data collection.

Type	Feature	Data	Describe
Input data	Sociodemographic profile	Age Gender Area	Student age Student gender Student area
Sperm data	Affiliated college	Student's department Course Semester	 Student course Semester
	Dynamic data	Update activity View courses Browse forums View course reports Answer question View issues Login score Forum score Job task score	Number of updates View the number of courses Browsing forums Viewing course reports Number of questions answered View the number of questions Login system score Participate in forum discussions Task score
Output data	Performance	GPA	Student's final performance

Figure 3: Number of student logins from week 1–16.

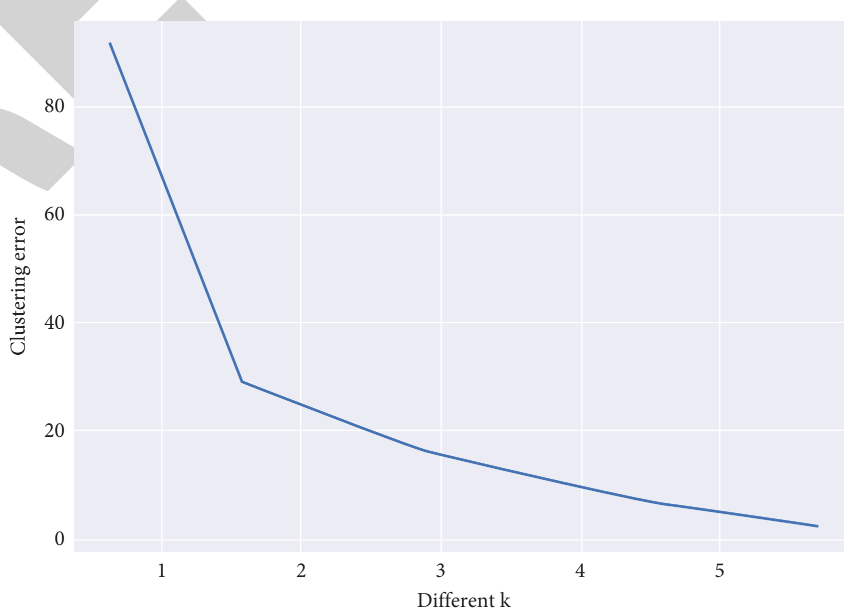


Figure 4: Elbow method to determine the optimal cluster value.

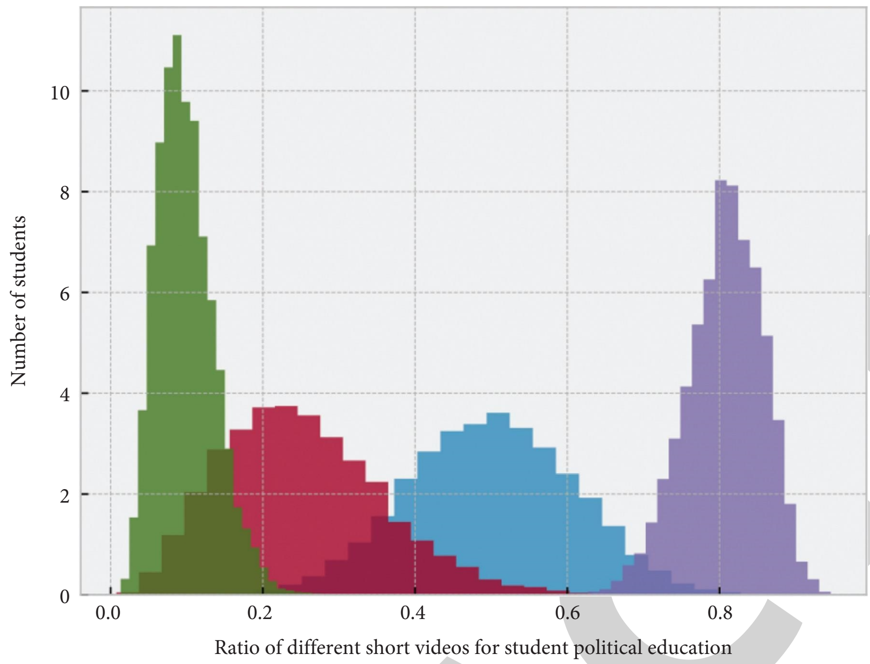


Figure 5: Distribution of students' political learning using different short videos.

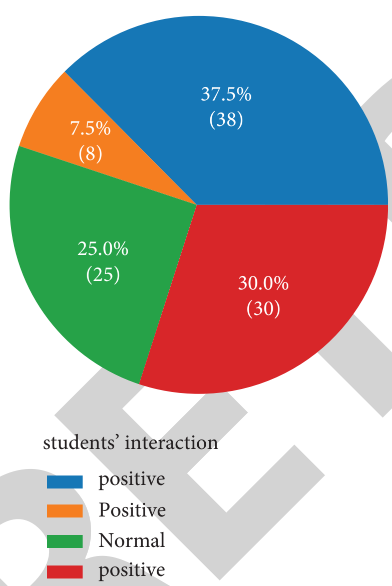


Figure 6: Students' interaction.

clustering used in the data. The value of k is 3 ($k = 3$). After defining the value, use the k-means method to execute the class.

Then, the k-means method is used to group the values defined by the above two methods. One problem we have to face in the experiment is the high dimension of data, so we use the core components to analyze PCA to reduce the amount of data measurement. Then, we analyze the distribution of students' political learning using different short videos in Figure 5. Next, we want to know how students interact in the learning process. In the e-learning process, all interactive data are stored in the system database log. The quality of students' interaction with students is an indicator to measure students' efforts to deepen their understanding to achieve learning results, as in Figure 6.

5. Conclusion

Generally speaking, in the short video starting from the media, the political education in colleges and universities should fully understand the existing problems, see the short video from the various advantages and convenience of the media, take its advantages, and effectively avoid its shortcomings. Promote the innovation and reform of political education in colleges and universities. Carry out transformation in the field of education and talent training, strengthen the orientation of students' adaptation to students' psychology and needs, strengthen basic ideological construction, stimulate university political education in the new era of innovative development, and cultivate outstanding talents that meet the needs of China's modern construction.

Data Availability

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

Conflicts of Interest

The author declares that there are no conflicts of interest to report regarding the present study.

Acknowledgments

The study was supported by Research Capacity Improvement Project of Beijing Technology and Business University

Hindawi
Mobile Information Systems
Volume 2022, Article ID 3052469, 12 pages
https://doi.org/10.1155/2022/3052469

Research Article

Big Data Analysis Technology for Artificial Intelligence Decision-Making Platform Construction and Application

Liangfang Liu [1] and Zhiyi Hu [2]

[1] *School of Information Engineering, Zhongshan Polytechnic, Zhongshan, Guangdong 528400, China*
[2] *University of Electronic Science and Technology of China Zhongshan Institute, Zhongshan, Guangdong 528400, China*

Correspondence should be addressed to Liangfang Liu; 2009013@zsc.edu.cn

Received 11 July 2022; Revised 27 July 2022; Accepted 2 August 2022; Published 31 August 2022

Academic Editor: R. Mo

With the development of the information age, it is the opportunity and challenge for enterprises to apply big data analysis technology to make decisions and better solve the major problems of global and sustainable development. Decision-making is crucial for enterprises, and a correct decision can improve the development potential and competitiveness of enterprises. However, traditional decision models have certain limitations, and it is difficult to handle the massive, polymorphic, and changing decision data. To address these problems, we propose a combination of temporal data anomaly detection and width learning big data analysis technology for building an intelligent decision platform to assist enterprises to better solve major decision problems. First, the goal of time series data anomaly detection is to correctly determine whether the data points in each moment of the time series are abnormal. The variation of time series data is affected by various factors, and the fluctuation of data caused by some nonanomalous factors can increase the difficulty of anomaly detection. To address the above problems, we propose an anomaly detection algorithm for time series data based on time series decomposition method. In this algorithm, the time series are decomposed by STL method and HP filtering method according to whether the time series is periodic or not, and then the components of the time series that are relevant to anomaly detection are retained and anomaly detection is performed on the processed time series using a cyclic model. Then, based on the call text and signaling data in enterprise decision-making, an improved width learning model called coding width learning is proposed. The coding width learning model is used to identify decision problems and make comprehensive decisions to improve the model training time and the accuracy of identification. At the same time, an integrated learning method with parallelized training is proposed for width learning in order to further improve the efficiency of coding width learning and prevent the potential memory explosion problem. Finally, the experimental results show that the proposed anomaly detection method effectively improves the anomaly detection performance of the model, and its performance is better than the existing time series anomaly detection algorithm based on variation self-encoder; combined with the improved width learning model, it can make fast decisions and analysis without using scenarios and targets, and help and guide the work of related personnel.

1. Introduction

The current rapid development of big data technology has added new impetus to the development of various industries, but also makes the enterprise management decision-making environment change dramatically, which has a greater impact on decision-making data and decision-making participants [1]. At the same time, the difficulty of identifying decision-making information has increased, and the traditional decision-making methods are in urgent need of innovation, which can easily affect the normal operation of enterprises once mishandled. In this regard, the leadership should establish data awareness, apply big data to management decision-making, ensure that the human resources management approach is more comprehensive and diversified, can fully respond to environmental changes, and innovate data mining and analysis technology to collect more information related to management decision-making, relying on many new ways to avoid and reduce decision-making risks.

The meaning of big data is divided into broad and narrow sense. Broad sense represents the collection of the data, which not only can analyze and process the data quickly and accurately, but also has important implications for decision-making; narrow sense represents the combination of massive information with each other. Usually, big data has the following characteristics. First, it has large capacity, there are several storage units available in the data integration, the minimum is TB, and then in this unit there are PB and EB; second, processing efficiency is fast. Compared with the previous processing technology, big data technology has significant advantages in processing efficiency, which can quickly find favorable resources in the huge amount of data and show enterprise capabilities; third is the type of diversity, such as text, images, and video [2]. Under the influence of this feature, it can provide strong support for data capacity expansion, and when the capacity becomes large, it can provide many facilities for more people to collect relevant information, so that the data needs of more groups can be fully satisfied, and it can also promote the efficiency of data processing and analysis, and provide positive auxiliary effects for scientific management decisions.

Decision-making as the central line of enterprise management, including strategic decisions and related decisions, belongs to a highly dynamic and complex management behavior, which has a decisive role in the state of enterprise operation and development, as shown in Figure 1:

(1) *The Influence of the Decision-Making Environment.* The information content in the era of big data is in dynamic change, in which the volume of data has changed dramatically, from TB, PB to EB in the past, the storage volume has also increased rapidly, the decision-making environment has been greatly affected, and the decision-making mode has changed under the data drive [3]. According to the current situation of the application of big data in decision-making, it can be seen that the information processing efficiency of most enterprises is low, which affects the effectiveness of big data.

(2) *The influence of Decision-Making Data.* Through long-term development, data have been greatly improved in terms of type, quantity, and structure [4]. In business operation, data collected through the information platform should be uncluttered and organized, during which it should be purposefully selected and screened, then the data and information should be continuously optimized, and finally the existing information processing system should be comprehensively upgraded. In the fickle information dissemination environment, strong technical assistance can be provided for real-time data processing and focus on the connection between big data and information, thereby mining information closely related to the enterprise to achieve the goal of sound development.

(3) *The impact of Decision-Making Participants.* On the one hand, the role of participants has changed. After the birth and application of big data, traditional decision-making schemes cannot adapt to the development of the times and the development of enterprises, so decisions need to be made on the basis of more refined and reasonable analysis. For the senior leaders, they should change the previous wrong decision-making methods, not simply according to past experience to give orders, but to comprehensively collect data and information, and combined with the actual situation of the enterprise, the deployment of tasks as the key elements of decision-making, carefully arranged and arranged to ensure the maximum use of human resources.

(4) *Decision-Making System Impact.* The system mainly includes two elements: one is the basis for decision-making, and the other is the decision-making process. In the past, enterprise decision-making was mainly based on the data of internal information system and report data, with strong one-sidedness and subjectivity, which could only reflect the operation and financial management of the enterprise [5]. With the rapid development of the network, enterprises can quickly and easily collect information from other enterprises so as to understand the market price fluctuations, market demand, consumer evaluation, and other information. And the above information is applied to the decision, to make the decision more objective and comprehensive, which can help help enterprises clear the development direction, timely avoid market risks, and enhance their core competitiveness.

At the same time, enterprises can also create decision management systems supported by big data and create integrated systems corresponding to different departments to fully demonstrate the practicality, expandability, and comprehensive functions [6]. Relying on the integrated system, enterprises can open up channels to collect relevant data sources, grasp user behavior and feedback, and track and collect user behavior as a basis for optimizing product design, which helps products better meet consumer expectations, increase sales, and gain more economic benefits. In addition, the content and form of decision-making are increasingly complex due to big data.

In enterprise decision-making, the content of data analysis should be closely focused on promoting the full play of big data technology. Data analysis results are affected by the ability of employees, such as staff analysis ability is weak, it is easy to collect a lot of worthless information, and cannot use the data correctly, resulting in a waste of data resources, while employees with strong analytical ability can accurately grasp the valuable information to promote the effectiveness of data resources [7]. In this regard, a platform should be created for data analysis during enterprise development to achieve higher efficiency with lower cost as possible [8].

The main contributions of this paper are as follows. In this paper, we propose an artificial intelligence decision-

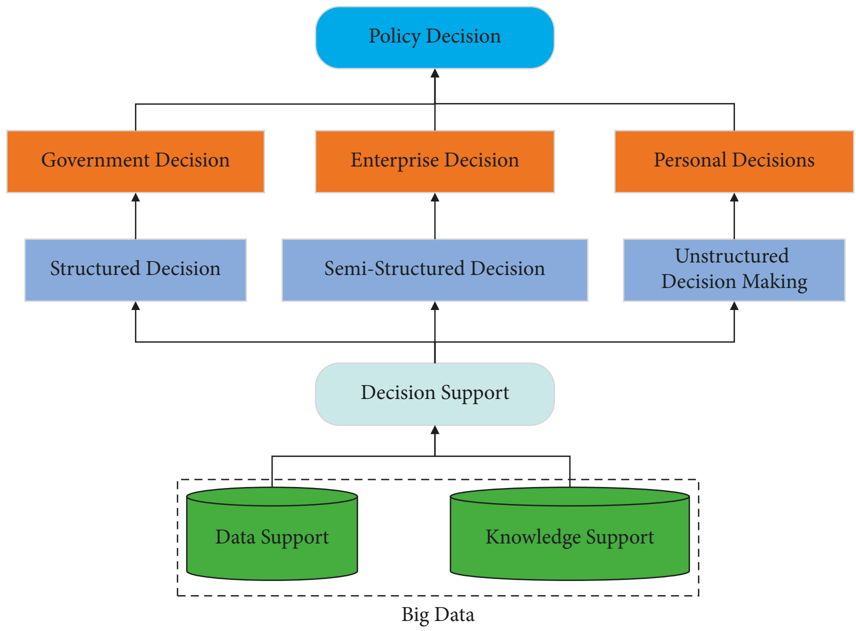


Figure 1: Enterprise decision based on big data.

making platform in which anomalous data detection methods and width learning methods are combined to better utilize big data technologies to solve major decision-making problems for enterprises. The model can obtain information related to time series from time series and learn the patterns of normal data. Then, based on previous research, the code width learning network algorithm is proposed and applied to multi-scene and multi-objective decision-making tasks. Combined with the corresponding parallelized training algorithm, the feasibility and efficiency of the algorithm are discussed at the level of algorithmic principles. Finally, it is experimentally verified that the proposed method can detect abnormal data, improve the accuracy and reliability of decision-making, and assist the relevant personnel in decision analysis.

2. Related Works

2.1. Current Status of Big Data Technology Research. To respond to users' needs in a short time, accurately complete data analysis tasks, and visualize the results to users is not encountered in traditional data analysis and processing. At present, a lot of researches are aimed at solving the problems faced by big data in various stages of generation, collection, storage, analysis and mining, and visualization.

Among them, big data collection is everywhere, and its sources cover finance, medical, Internet, transportation, communication, education, scientific research, and other fields [9]. Visual presentation: In order for users to better understand the results of data analysis and mining, the knowledge or patterns mined need to be visually displayed to users at the terminal in a friendly and easy-to-understand way to provide advice or support for user decision-making.

Big data analysis and mining is an important technology to transform massive, complex, high-speed, and low-density big data into knowledge or patterns for human production and life services. Second, the clustering algorithm module uses the density k-means algorithm as the basis for selecting the initial clustering centers; finally, the CUDA architecture and MPI message passing interface are used to achieve parallelism and reduce the time overhead of the algorithm [10]. On this basis, some researchers have focused their research on textual big data on the semantics of big data and given a constraint model based on clinical document standards and user use case consistency, which solves the semantic loss problem in the process of traditional medical big data document division.

On this basis, some researchers have worked on overlapping community structures and discovered community structures based on complete subgraph percolation, which has been successfully applied in biological, information, and social networks; further, some researchers have proposed a new community discovery method using aggregated hierarchical clustering techniques, which can reveal both network hierarchies and overlapping community structures [11]. Some researchers choose a series of important video clips to represent the original video and then use the features of the original video to smooth the video clips to get a smoother video summary. Based on this, some researchers construct a video hypergraph model and use hypergraph sorting to classify videos according to different contents, and finally generate video summaries by function optimization.

In order to analyze the semantics of mobile data to detect anomalies in mobile object activities, some researchers have studied both temporal and spatial aspects of mobile object trajectory data. For the mobility prediction problem in mobile data, a new evolutionary algorithm has been proposed, which predicts the next movement of a mobile user in a personal communication system through three stages: movement pattern mining, movement rule extraction, and mobility prediction [12]. To improve the security of mobile data, a researcher has proposed a framework to collect real-time information and alert in real time.

2.2. Current Status of Anomaly Data Detection Research. In order to better and fully utilize the value of data, it is usually necessary to govern the huge amount of data. One of the more important tasks is anomaly data detection. In many data mining and statistical literature, anomalies are also referred to as inconsistencies.

Almost all anomalous data detection methods nowadays create models of the normal patterns of the data and then calculate or measure data anomalies based on their own deviations from these models. The principle of anomaly data detection can be broken down into two specific sub-problems [13]. The first problem is the problem of how to define anomalous data for a given dataset. The second problem is how to select an effective anomaly data detection method for anomaly data based on the characteristics of the dataset.

The basic problems of anomalous data detection have been effectively solved from different perspectives. The global anomaly detection model gives a binary label to the data to be tested whether it is anomalous or not. The local anomaly detection model calculates the anomaly score of each data object, which indicates the probability that each object is anomalous data [14].

Since this method appeared before the emergence and popularity of computer technology, it does not take into account the various problems encountered in the actual use of data representation, computational efficiency, and computational complexity in the process of anomaly detection [15]. Despite the problems of this method, the idea of mathematical modeling is very useful in many computational scenarios. One of the more commonly used statistical model-based anomaly detection methods is to detect anomalous data by detecting extreme univariate values.

The distance-based anomaly data detection method is an anomaly data detection method that can span various data domains, which uses the nearest neighbor distance to define anomaly measurement criteria.

Later, techniques based on unsupervised representation learning began to emerge, such as subspace feature selection methods, neural networks, and stream learning methods. Subspace-based feature selection methods reduce the impact of irrelevant features by finding subsets of features that are relevant to the anomalous data, and then perform regular anomaly detection on these feature subsets. This approach usually separates subset selection and anomaly detection, which results in features that are irrelevant to the anomaly data being used for anomaly detection [16]. Therefore, this approach results in reduced accuracy and large bias in anomaly detection. Neural network and stream learning-based approaches focus on retaining regular information about the data, which is then used for learning tasks such as clustering and data compression.

In summary, in the field of anomaly data detection, the problem of anomaly detection for basic data types or low-dimensional data is relatively mature. However, there are still many problems in detecting anomalous data when facing high-dimensional data. In addition, with the continuous development of various industries, data have been fully accumulated, and the data in many fields have become very large.

2.3. Current Status of Research on Artificial Intelligence Decision-Making Platforms. Rapidly developing artificial intelligence (AI) technologies have enabled intelligent decision-making applications to rapidly penetrate into various fields, which have a significant impact on socioeconomic and people's lives. The use of AI technology to empower existing complex decision-making systems to improve their intelligent decision-making capabilities is known as AI-enabled systems [17]. At present, AI has become an important development strategy for major countries and regions in the world, such as China, the USA, the European Union, and Japan, and the application of AI and competition in AI-based decision-making will directly affect the future evolution of the international landscape.

However, with its deeper integration in related industries, accidents of AI-enabled system decision-making frequently occur. It has been found that AI-enabled systems have endogenous risks such as black box, bias, security, and unaccountability, along with superhuman performance, and trust risks in the process of people's interaction with AI-enabled systems, which together lead to a crisis of trust in AI decision-making [18]. Especially in high-risk scenarios, the wrong predictions and bad decisions of AI-enabled systems will lead to unbearable consequences.

In the early days of intelligent decision-making applications, relatively simple models such as linear regression algorithms and decision tree models were mainly used, and humans could easily understand the logic and make decisions directly. From the perspective of user needs, explainable AI methods can be classified into four categories: visual explanation for intuitive detection of the interior, exploration explanation from external perturbation, knowledge explanation based on user common sense, and causal explanation reflecting decision causality, which is expected to explain AI black box decisions into transparent ones [19]. In terms of application research, the application research of explainable AI in the fields of intelligent medical care, unmanned driving, intelligent finance, intelligent justice, etc., has also been carried out, with the goal of improving the reliability of AI-enabled system predictions or decisions.

Big data-driven AI systems should not be influenced by human subjectivity, but flawed data can lead to biased and unfair decisions, and adversarial data can lead to serious decision errors, which seriously affects the credibility of AI-enabled system applications and decisions. At the methodological level, research on how to address inequity has evolved through three phases. The first stage is perceived fairness, which investigates how to deal with protected attributes directly to obtain fairness, using differential treatment such as directly excluding protected attributes such as race and gender in the decision-making process.

In a multi-risk environment, AI-enabled systems raise uncertainty in both data and model dimensions. Data risk will lead to changes in prediction or decision results, model risk can present incorrect results for the inputs, further affecting the continuity of trust, and the above risks lead to increased uncertainty as shown in Figure 2. AI-enabled systems under multiple uncertainties make trustworthy decisions for information situations where the data sources are cross-domain and the data are no longer perfect, but may be biased or adversarial risk data, leading to unfair or incorrect results of the decision model. In the process of human-machine collaborative decision-making, the results at the case level need to be calibrated for trust, for the purpose of enhancing human or AI single decision-making.

In high-stakes decision-making, many human experts rely on the output of AI to form the final decision, forming human-machine teams. Research suggests that these human-machine teams may perform better than human-only or AI-only teams, and that in order to achieve this, humans must have a moderately calibrated level of trust; otherwise, the AI's trust level will be mist-calibrated and the human-machine team will not perform as well as the human decision alone [20]. For trust calibration, some researchers have suggested the need for trust calibration that understands the capabilities of the system and the reliability of the system output. Thus, enhancing the understandability, responsiveness, and ability to resolve conflicting goals of AI-enabled systems may be far more meaningful than simply improving the accuracy of AI.

Metrics and standards for AI decision-making as research on AI-enabled system decision-making progresses: the problems of different principles and methods in industry applications will gradually be exposed, and metrics and standardization of decision-making are the way to go. On the one hand, we need to design AI decision-making from the perspective of developer training, testing and experimentation, deployment and operation, and supervision, and on the other hand, we need to develop standards and specifications from hardware, algorithms, and systems such as chips. Future research should focus on the metrics and standards of AI-enabled system decision-making, coordinate guidance, and regulation of decision-making, and promote the healthy and sustainable development of AI decision-making.

3. Algorithm Design

3.1. Abnormal Data Detection Method. In anomaly detection, whether the target data are anomalous at a certain moment is related to its timing information, contextual information, and the numerical information of the data itself. Some characteristics of the anomaly detection problem itself also make it different from the conventional classification problems, such as the fact that the positive and negative samples of the data corresponding to anomaly detection are usually severely unbalanced, and there are sometimes unlabeled anomalous data in the data.

Recent research work on time series anomaly detection has made more use of modeling methods related to deep learning. In this thesis, we focus on analyzing recurrent neural network-based and vibrational self-encoder-based anomaly detection algorithms.

In anomaly detection problems, there are usually far more normal data than anomalous data, so the anomaly detection problem cannot be solved using the conventional classification problem solution [21]. Due to this nature of the anomaly detection problem, solving the anomaly detection problem using unsupervised learning algorithms has become an idea for researchers. According to existing time series decomposition methods, such as STL methods, the time series can be decomposed into several combinatorial parts. The decomposition of the time series is represented using the following formula:

$$x_t = \tau_t + c_t + s_t + i_t. \tag{1}$$

In the formula, x_t denotes the numerical vector of the time series at moment t, τ_t is the trend part, c_t is the periodic part, s_t is the seasonal part, and i_t is the irregular part. The retained part of the time series will be used as the object of anomaly detection model processing. The following explains the reasons for using the time series decomposition method to process the sequences. In many cases, the amount of data for training anomaly detection models is not sufficient, so anomaly detection models based on generative models do not learn the pattern information inside the data well. There may be noisy parts in the time series and some parts that are not related to anomaly detection, and these parts can interfere with anomaly detection to some extent. By decomposing the time series using the time series decomposition method, it will be less difficult for the model to perform anomaly detection on the processed time series, which is equivalent to the simplification of the problem and is beneficial for the model to perform better anomaly detection.

The complex distribution of data can be learned by unsupervised learning, the structural features of recurrent neural networks (RNN, LSTM) make them suitable for scenarios involving sequences, recurrent neural networks are able to obtain time series-dependent information from time series, and this thesis uses recurrent neural networks in time series anomaly detection. There have been research works applying LSTM to time series anomaly detection and achieving better results in some scenarios. In order to better obtain time series-dependent information, we consider adding more neural network processing layers for processing time series in VAE, but the LSTM layers have more parameters, and the presence of too many LSTM layers in the model will increase the difficulty of model training, so RNN layers are introduced. The overall structure of the anomaly detection model is consistent with that of VAE, and the algorithm performs model training in an unsupervised learning manner, using normal time series data to train the model and determine anomalies based on reconstruction errors [22]. The general process of solving the problem is to process the time series using the time series decomposition method and subsequently perform anomaly detection on the processed time series using the cyclic VAE model.

Figure 3 shows the general framework of D-R-VAE, which can be divided into 3 parts. The anomaly detection

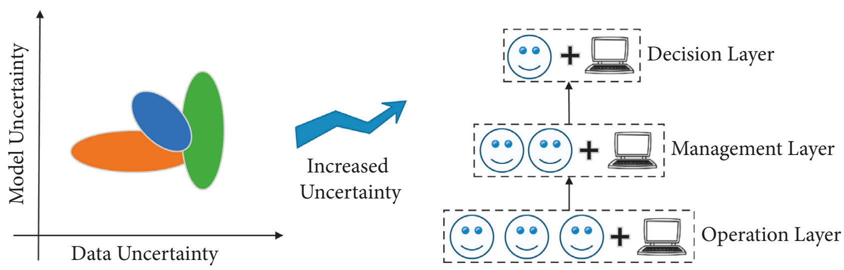


Figure 2: Decision-making process of uncertainty increase and human-machine cooperation.

model uses a combination of recurrent neural network and VAE. Each module is described below.

For nonperiodic time series, we utilize the HP filtering method. This decomposition method can decompose the time series data into two components, the trend part and the remaining part. We retain the trend part of the time series decomposed by the HP filtering method for subsequent anomaly detection. The trend part of the time series becomes insensitive to short-term fluctuations compared to the original time series, which is equivalent to removing the interference of noise and other factors to some extent. The sensitivity of the trend part to short-term fluctuations can be adjusted using the parameters in the decomposition function, and suitable parameters are selected for the HP filter decomposition function according to the nature of the dataset during the experiment.

D-R-VAE decomposes the time series into multiple components using the time series decomposition method and removes the unnecessary components of the series for the anomaly detection task. The cyclic VAE model performs anomaly detection on the retained time series, and a combination of VAE, LSTM, and RNN is used in the model.

Figure 4 shows the framework flowchart of the algorithm. The execution process of the algorithm can be divided into two main parts: firstly, the time series are decomposed using the classical time series decomposition methods in statistics and mathematics. Time series with obvious periodicity are decomposed by the STL method, and time series without periodicity are decomposed by the HP filtering method. After the sequence is decomposed by STL method, the remaining part is retained, and after the sequence is decomposed by HP filter method, the trend part is retained. The retained part of the time series is used as the new time series. At the end of the first part, the new sequence is divided into a set of fixed-length subsequences and further divided into a training set and a test set. In the second part of the algorithm, the D-R-VAE model performs anomaly detection on the processed time series [23]. The D-R-VAE model is first trained using normal time series, and subsequently the anomaly score output from each subsequence in the training set data after processing by the model is retained and used to calculate the anomaly threshold for the anomaly scores. In the testing phase, the time series in the test set are reconstructed using the D-R-VAE model, the abnormal scores are calculated, and the abnormal scores are compared with the abnormal thresholds to determine whether they are abnormal or not.

The first judgment is based on the time series' own attributes. When the origin of the time series is not clear, it is necessary to divide the multivariate time series into a set of univariate time series, visualize each univariate time series using a visualization function, and observe whether it is cyclical or not. If the periodicity cannot be judged with certainty, it is necessary to select an approximate period for the series based on the observation of the series and segment the time series according to this period, and if the distance of adjacent segments is less than the threshold, it is periodic.

The model is expanded to represent according to the length of the fixed-length window, and the structure and internal parameters of each unit after expansion are the same. The overall structure of the model is the same as that of VAE, the first neural network processing layer of both the encoder part and the decoder part of the model is the LSTM layer, and the processing layer for generating the parameters of the hidden variable distribution of the model is the RNN layer. Due to the introduction of LSTM and RNN, the D-R-VAE model can obtain the time-dependent information in the sequence when processing the time series data. Substituting the variables into the objective function of VAE, the objective function of the D-R-VAE model is

$$\sum_{t=1}^{w}\left(KL\left[q_{\phi}\left(z_t \mid \overline{x}_t\right)\|p\left(z_t\right)\right]\right). \tag{2}$$

The objective function is maximized during the model training. The right-hand term calculates the KL scatter of the two Gaussian distributions. In the anomaly detection problem, it is necessary to define an anomaly score for the object to be detected and determine the anomaly according to the anomaly score. For example, the data x_t at time t as the object to be detected are expressed by the following equation:

$$\begin{aligned}\left\{\overline{x}_{t-w+1},\overline{x}_{t-w+2,..,}\overline{x}_t\right\} &\longrightarrow D-R-VAE\\ &\longrightarrow \left\{\overline{x}_{t-w+1,}'\overline{x}_{t-w+2,..,}'\overline{x}_t'\right\}as_t = mse\left(\overline{x}_t,\overline{x}_t'\right).\end{aligned} \tag{3}$$

The as_t in the formula is the anomaly score. After the model training is completed, the anomaly scores corresponding to each moment in the training set are calculated for use in the anomaly threshold selection phase. Anomaly threshold selection: After the model training is completed, the model is used to reconstruct the time series in the training set, and we get a set of anomaly scores.

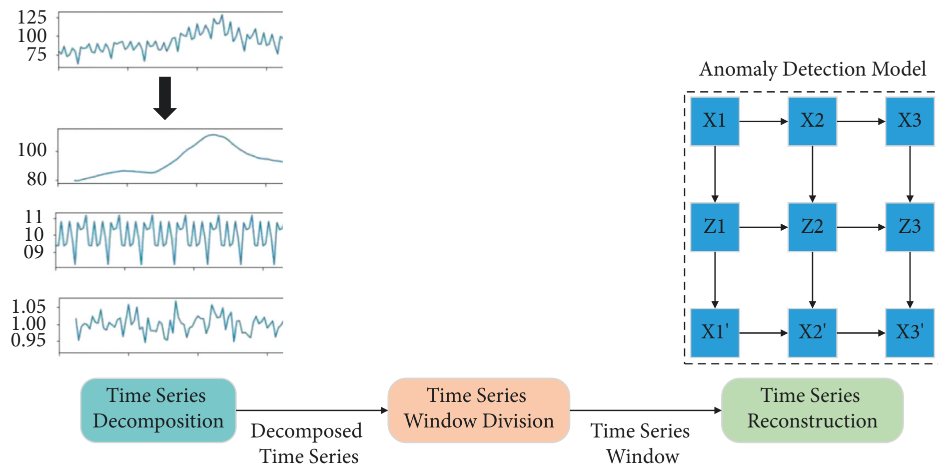


Figure 3: Overall framework of D-R-VAE.

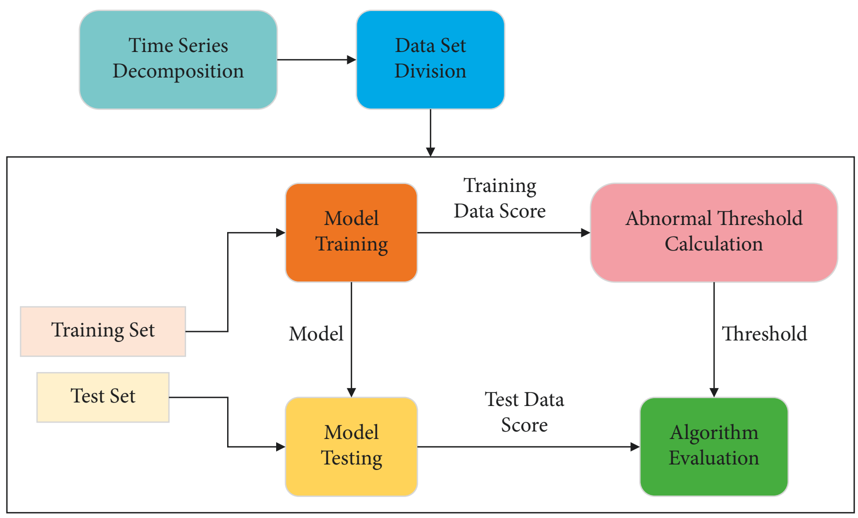


Figure 4: D-R-VAE framework flowchart.

Anomaly thresholds are selected based on density. We place the anomaly scores into a one-dimensional space, and we also have a set of points distributed in the one-dimensional space. Anomalous data are data that are significantly different from normal data. The model is trained using normal data, so the reconstruction error will be larger when the model is reconstructed for the abnormal time series. Since the training data are all normal data, the scores in the set of abnormal scores corresponding to the training set are concentrated in a certain normal range, and the density of abnormal scores in the normal range is the largest [24]. However, there are a small number of abnormal scores in the training set that deviate from the normal range, and the density of these abnormal scores in the space is small. We select the threshold value among these abnormal scores. The smaller density of points at the locations of the anomaly scores leads to the distances between the points and their immediate neighbors, so the sum of the differences between the distances of the anomaly scores and their immediate neighbors in the one-dimensional space is used as a measure of point density.

3.2. Artificial Intelligence Decision-Making Platform. In order to better meet the requirements of accuracy and timeliness of enterprise decision-making, the width learning system is referred to as the training model, which finds the weight matrix by generating feature nodes and enhancement nodes and calculating the pseudo-inverse directly, which greatly reduces the prediction time [25–28].

The signaling features and the first 15 bits of the voice call are extracted as the text, the text is preprocessed accordingly to build a word vector space, and a decision method based on the width learning network is proposed, which together constitute the data preprocessing part. The proposed decision model is shown in Figure 5.

However, since the feature nodes are initialized in a completely random way for width learning, there may be a large number of inefficient node generation and a certain amount of feature information missing [29, 30]. To solve this problem and to learn the vector representation better, the denoising self-encoder adds noise to the original data in order to solve this problem:

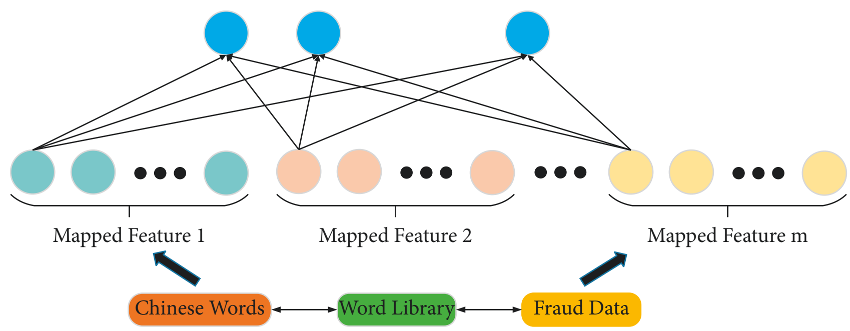


FIGURE 5: Decision model based on width learning network.

$$noisy(x, y) = \frac{\overline{x}_{random} - \mu}{\sigma/\sqrt{m}}. \quad (4)$$

The new variables with Gaussian noise are then fed to the noise reduction auto-encoder, which is set to zero with a certain probability, equivalent to the loss of some features of the data, to improve the noise immunity of the data, and to generate a new representation:

$$x_{new} = P(x + noisy(x, y)). \quad (5)$$

Next, we feed the learned features into the width learning network as coded vectors. In this way, the feature nodes can be represented as

$$Z_o^i = \phi\left(x_o w_o + \beta^i\right). \quad (6)$$

In deep learning, the common way of parallelized training is that at the beginning of each round of training, these processes or devices will first read the current parameter values uniformly and get the data.

But width learning is different, because it does not back propagate the gradient for training, so we need to partition at the data level more, cut into smaller matrices to reduce the capacity, and calculate the pseudo-inverse. To better solve this problem and to try to further optimize the performance of the width learner, the integration learning approach is borrowed. The integrated learning essentially trains multiple weak learner models and combines the learning results of each learner in some way to obtain better results than a single weak learner. For the width parallelization algorithm, we can train multiple weak width learners with small data size and combine the learning results in an integrated way, which can well solve the memory explosion problem and improve the generalization performance of the model.

4. Experiments

4.1. Abnormal Data Detection Experiment

4.1.1. Experimental Parameter Settings. Three algorithms are selected for comparison with the algorithms proposed in this thesis, and the benchmark algorithms are briefly described below: (1) LSTM-based anomaly detection algorithm, which uses LSTM to predict time series and determine anomalies based on the prediction error; (2) VAE-based anomaly detection algorithm, which uses VAE to reconstruct input variables and determine anomalies based on the reconstruction error; And (3) VAE-LSTM anomaly detection algorithm, which uses a combination of VAE and LSTM in the anomaly detection model and adapts the first fully connected processing layer of the encoder part and decoder part of VAE to the LSTM layer. The experimental environment is as follows: Intel i7 3.30 GHz, RAM 256 GB, and Nvidia 2080 GPU; operating system and software platform are Ubuntu 18.04, TensorFlow 1.12, and Python 3.8. The training process loss convergence curve and performance improvement are shown in Figures 6 and 7.

The experimental data in this paper were obtained from internal nonpublic data of the China Decision Evaluation Research Institute. The test datasets include (1) Server Machine Dataset (SMD), which is collected by the researchers from an Internet company and contains time series data from 3 sets of physical machines, for a total of 28 multivariate time series datasets from different machines, which need to be used independently; (2) Yahoo Benchmark Dataset (Yahoo), which contains 4 benchmark datasets, the first of which is selected to evaluate the model algorithm, and the remaining 3 datasets (synthetic datasets, where anomalies are easily identified by the model).

In the anomaly detection problem, there are usually more normal data than anomalous data in the dataset, and the data are unbalanced, so we choose the criterion for evaluating the algorithm performance when it is unbalanced. We use abnormal data as positive data and normal data as negative data, and classify the data judged by the model into four categories according to the nature of the data itself and the model's judgment of the nature of the data. True positive (TP) is the abnormal data judged by the model as abnormal, false negative (FN) is the abnormal data judged by the model as normal, false positive (FP) is the normal data judged by the model as abnormal, and true negative (TN) is the normal data judged by the model as normal.

4.1.2. Experimental Results and Analysis. Firstly, we analyze the effects of parameters and window length in the objective function on the anomaly detection performance of the D-R-VAE model by increasing the parameters from 0.2 to 1.0 (each time increasing by 0.2), and evaluate the model performance on the SMD dataset with the $F1$ score metric. The red curve in Figure 8 shows the experimental results. The model performance is optimal when the parameters are

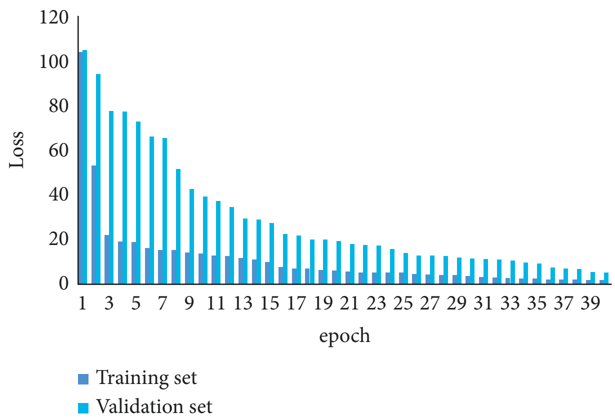


Figure 6: Training process loss convergence curve.

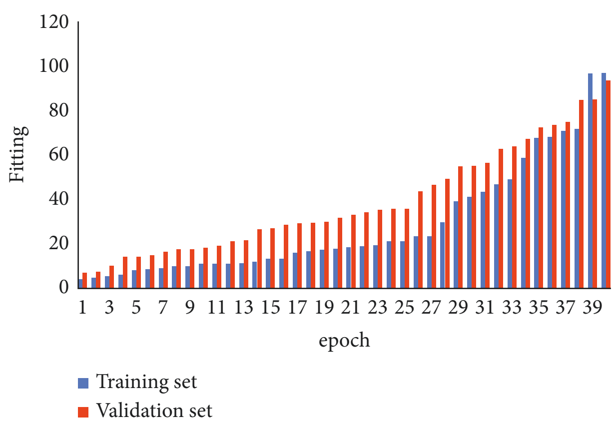


Figure 7: Training process performance improvement diagram.

between [0.2, 0.6], and the model anomaly detection performance tends to decrease as the parameters are increased. For model training of the D-R-VAE model and anomaly detection using the model, a fixed-length sub-series is required as model input. The sub-series length is the time series window length, and the window length has an impact on the anomaly detection performance of the model. The blue curve in Figure 8 shows the experimental results of the D-R-VAE model with different window lengths for anomaly detection on the SMD dataset, and the anomaly detection performance of the model gradually improves as the window length increases from 80 to 140 and then stabilizes.

Table 1 shows the effect of time series decomposition on the anomaly detection performance of the D-R-VAE model prior to anomaly detection by the D-R-VAE algorithm, where the D-R-VAE model performs anomaly detection on the retained portion of the time series decomposed by a specific time series decomposition method, and the R-VAE algorithm uses the same model to perform anomaly detection on the original time series. The experimental results show that the performance of the D-R-VAE algorithm is significantly better than that of the R-VAE algorithm, thus indicating that the anomaly detection performance of the D-R-VAE model can be effectively improved by using the time series decomposition method to process the time series appropriately.

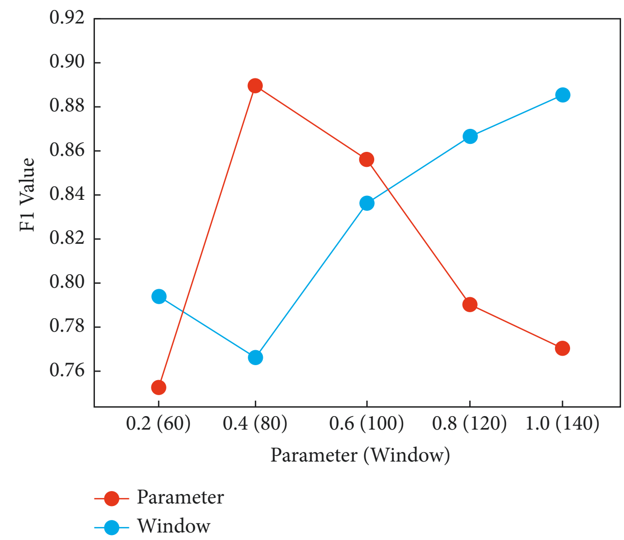


Figure 8: Influence of parameters and window length model.

Table 1: Experimental results of different detection methods.

Dataset	Method	Precision (%)	Recall (%)	*F*1 score (%)
SMD	LSTM	67.22	72.61	63.92
	VAE	49.56	67.48	48.55
	VAE-LSTM	56.31	82.34	57.87
	D-R-VAE	75.83	87.06	76.28
Yahoo	LSTM	81.93	77.66	55.70
	VAE	73.46	71.52	47.64
	VAE-LSTM	86.37	82.93	62.75
	D-R-VAE	91.29	89.64	68.91

Figure 9 shows the visualization results of the D-R-VAE algorithm for anomaly detection on the SMD dataset, where the red dashed line corresponds to the anomaly threshold, the red points correspond to the anomalous points, and the blue points correspond to the normal points. The number and density of red abnormal points and blue normal points are based on the abnormality threshold above. It can be seen from Figure 9 that the D-R-VAE model performs better, the red anomalous points above the anomaly threshold are more intensive, and the number of red anomalous points above the threshold is significantly more than the blue anomalous points, thus indicating that the anomaly scores generated by the D-R-VAE model can better distinguish between anomalous and normal points.

4.2. Decision Platform Experiment

4.2.1. Experimental Parameter Setting. A total of about 500,000 pieces of labeled data are used to verify the algorithm effect. The data are divided into 6 categories according to the decision type. At the same time, the training set and the validation set were divided into 8 : 2 ratios according to the data time series, and the feature generation process was completed as described above.

First, the features used are modeled accordingly to fill in the missing values for continuous and discrete numerical

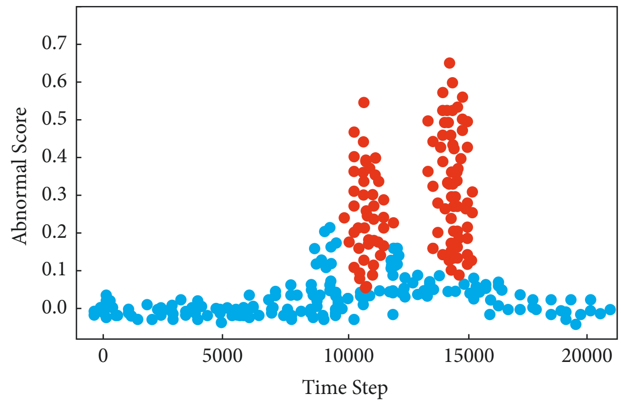


Figure 9: Visualization results of anomaly detection.

features, and a large number of missing fields are removed. Based on the a priori knowledge, statistical features are generated, such as the number of calls and the average call duration of each call in a certain period of time. The discrete features are combined to construct cross features; the continuous features are varied to transform the value range. The features are evaluated by the feature importance evaluation method mentioned above, and the feature selection is completed. The preprocessed feature set is imported into the code width learning model classifier, and the effect of the model is adjusted by adjusting the number of feature nodes and enhancement nodes.

To verify the effectiveness of code width learning network in enterprise decision-making, it is compared with several classical neural network algorithms including convolutional neural network (CNN), recurrent neural network (RNN), and deep confidence network (DBN) without using our proposed parallel algorithm.

4.2.2. Experimental Results and Analysis. First, select four groups of parameters with high accuracy and draw the AUC characteristic curve, as shown in Figure 10. It can help us better judge whether the classifier is effective enough and find the best decision data parameters in the model. It can be seen from Figure 10 that the AUC is greater than 0.9, which shows that the classifier has a good effect in enterprise decision-making. In this figure, the maximum value of AUC is 0.94. When 400 feature nodes and 2000 enhancement nodes are set, better results can be obtained. The experimental results also show that it is necessary to reasonably adjust the parameters of the width learning network. Increasing the number of feature nodes and enhancement nodes can help the module better interpret features and improve the accuracy of the model. However, if the number of feature nodes and enhancement nodes is too large, the accuracy will be reduced and the model will be over fitted. Using the existing data, the maximum accuracy can reach 94.38%.

Table 2 gives the experimental results of comparing the coding width learning model with the classical algorithm. From Table 2, we can see that the accuracy of our proposed coded width learning network and recurrent neural network

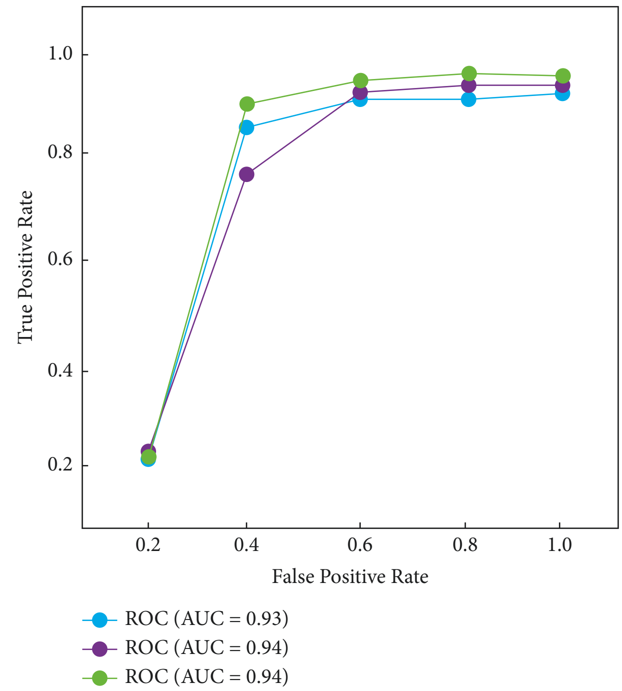


Figure 10: ROC curve of coding width learning.

Table 2: Comparison of coding width learning model and classical algorithm.

Method	Accuracy (%)	Recall (%)	Training time (s)
CNN	91.62	92.17	975.39
RNN	91.86	91.28	4214.65
DBN	92.75	91.05	3618.07
Ours	94.38	93.46	38.29

is high, reaching 94.38% and 92.75%, respectively, while the original width learning is slightly lower compared to the improved one, but its accuracy is also high compared to most models. It reached 91.86%. From the perspective of recall, the coded width learning network has the highest recall rate of 93.46%. In terms of training time, the original width learning network is the fastest. The whole training process takes only 38.29 seconds, which is hundreds of times faster than other algorithms. Our proposed coded width learning network model takes slightly longer time due to the addition of coded feature extraction, but it also has obvious advantages over other models. In general, the code width learning network has higher accuracy and highest recall compared to several classical models, and the training time is very short. Therefore, it has excellent performance in business decisions that require very high timeliness and accuracy.

5. Conclusions

In the era of big data, enterprise management and decision-making behavior has changed greatly, mainly in terms of decision-making environment, data, participants and system, etc., which put forward higher requirements for the leadership and managers, who should actively establish data awareness, rely on big data to comprehensively grasp the current market situation change pattern, and combine with their own situation to scientifically develop decision-making solutions. To this end, combined with big data analysis technology, this paper proposes an artificial intelligence decision-making platform to assist relevant personnel in solving major decision-making problems. First, for the problem of anomaly detection of time series data, the D-R-VAE algorithm is proposed to decompose the time series and retain the components of the time series related to anomaly detection. The algorithm uses the model structure of VAE to learn the normal patterns of time series data, the LSTM and RNN layers included in the model make the model more suitable for processing sequence-like data, and the model is able to utilize the time series-dependent information of time series in the process of processing data. Then, an artificial intelligence decision platform is constructed based on width learning, which includes a decision algorithm based on coded width learning network and a parallelized training algorithm based on width learning model. The former draws on denoising self-coding and width learning to achieve the requirement of high timeliness and high accuracy in decision-making; the latter combines width learning features with integrated learning methods for parallelized training to further improve efficiency and solve the memory explosion problem. Finally, the experiments show that the D-R-VAE model can effectively detect anomalous data and provide data reference for the subsequent decision platform; the width learning-based decision platform requires short practice time and high accuracy, and can meet the real-time decision requirements. In the future, we plan to use graph convolutional neural network for artificial intelligence decision platform building and application of big data analysis technology.

Data Availability

The datasets used during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] H. Lin, R. Li, Z. Liu et al., "Diagnostic efficacy and therapeutic decision-making capacity of an artificial intelligence platform for childhood cataracts in eye clinics: a multicentre randomized controlled trial," *EClinicalMedicine*, vol. 9, pp. 52–59, 2019.

[2] E. Long, H. Lin, Z. Liu et al., "An artificial intelligence platform for the multihospital collaborative management of congenital cataracts," *Nature biomedical engineering*, vol. 1, no. 2, pp. 1–8, 2017.

[3] T. Sourdin, "Judge v Robot?: artificial intelligence and judicial decision-making," *University of New South Wales Law Journal The*, vol. 41, no. 4, pp. 1114–1133, 2018.

[4] R. Vaishya, M. Javaid, I. H. Khan, and A Haleem, "Artificial Intelligence (AI) applications for COVID-19 pandemic," *Diabetes & Metabolic Syndrome: Clinical Research Reviews*, vol. 14, no. 4, pp. 337–339, 2020.

[5] K. Colchester, H. Hagras, D. Alghazzawi, and G Aldabbagh, "A survey of artificial intelligence techniques employed for adaptive educational systems within E-learning platforms," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 7, no. 1, pp. 47–64, 2017.

[6] 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.

[7] Y. R. Shrestha, S. M. Ben-Menahem, and G. Von Krogh, "Organizational decision-making structures in the age of artificial intelligence," *California Management Review*, vol. 61, no. 4, pp. 66–83, 2019.

[8] T. Araujo, N. Helberger, S. Kruikemeier, and C. H. de Vreese, "AI we trust? Perceptions about automated decision-making by artificial intelligence[J]," *AI & Society*, vol. 35, no. 3, pp. 611–623, 2020.

[9] N. Singh, "Big data technology: developments in current research and emerging landscape," *Enterprise Information Systems*, vol. 13, no. 6, pp. 801–831, 2019.

[10] L. W. Santoso and Yulia, "Data warehouse with big data technology for higher education," *Procedia Computer Science*, vol. 124, pp. 93–99, 2017.

[11] L. T. Wright, R. Robin, M. Stone, and D. E Aravopoulou, "Adoption of big data technology for innovation in B2B marketing," *Journal of Business-To-Business Marketing*, vol. 26, no. 3-4, pp. 281–293, 2019.

[12] W. S. Lee, E. J. Han, and S. Y. Sohn, "Predicting the pattern of technology convergence using big-data technology on large-scale triadic patents," *Technological Forecasting and Social Change*, vol. 100, pp. 317–329, 2015.

[13] J. Zhao, K. Liu, W. Wang, and Y Liu, "Adaptive fuzzy clustering based anomaly data detection in energy system of steel industry," *Information Sciences*, vol. 259, pp. 335–345, 2014.

[14] S. Agrawal and J. Agrawal, "Survey on anomaly detection using data mining techniques," *Procedia Computer Science*, vol. 60, pp. 708–713, 2015.

[15] S. Ahmad, A. Lavin, S. Purdy, and Z Agha, "Unsupervised real-time anomaly detection for streaming data," *Neurocomputing*, vol. 262, pp. 134–147, 2017.

[16] A. Blázquez-García, A. Conde, U. Mori, and J. A Lozano, "A review on outlier/anomaly detection in time series data," *ACM Computing Surveys*, vol. 54, no. 3, pp. 1–33, 2022.

[17] T. J. Loftus, P. J. Tighe, A. C. Filiberto et al., "Artificial intelligence and surgical decision-making," *JAMA surgery*, vol. 155, no. 2, pp. 148–158, 2020.

[18] H. Wang, Y. Huang, A. Khajepour, D. Cao, and C Lv, "Ethical decision-making platform in autonomous vehicles with lexicographic optimization based model predictive controller," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 8164–8175, 2020.

[19] J. Qin, X. Zhu, Z. Wang, J. Ma, S. Gao, and C. Hu, "Substation decision-making platform based on artificial intelligence," *Distributed Generation & Alternative Energy Journal*, pp. 151–172, 2020.

[20] S. Tripathy, P. Supraja, S. Mohanty et al., "Artificial intelligence-based portable bioelectronics platform for SARS-CoV-2 diagnosis with multi-nucleotide probe assay for clinical decisions," *Analytical Chemistry*, vol. 93, no. 45, pp. 14955–14965, 2021.

[21] C. Liu and Y. Wang, "Design of abnormal data detection system for protein gene library based on data mining

technology," *Cellular and Molecular Biology (Paris, France, Print)*, vol. 66, no. 7, pp. 103–110, 2020.

[22] X. Sun, "Similarity detection method of abnormal data in network based on data mining," *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 1, pp. 155–162, 2020.

[23] M. Bang, S. S. Engelsgaard, E. K. Alexandersen, M. Riber Skydt, H. R. Shaker, and M Jradi, "Novel real-time model-based fault detection method for automatic identification of abnormal energy performance in building ventilation units," *Energy and Buildings*, vol. 183, pp. 238–251, 2019.

[24] L. Li, S. Das, R. John Hansman, R. Palacios, and A. N Srivastava, "Analysis of flight data using clustering techniques for detecting abnormal operations," *Journal of Aerospace Information Systems*, vol. 12, no. 9, pp. 587–598, 2015.

[25] B. Aikemu, P. Xue, H. Hong et al., "Artificial intelligence in decision-making for colorectal cancer treatment strategy: an observational study of implementing watson for oncology in a 250-case cohort," *Frontiers in Oncology*, p. 10, 2021.

[26] J. C. Pomerol, "Artificial intelligence and human decision making," *European Journal of Operational Research*, vol. 99, no. 1, pp. 3–25, 1997.

[27] C. C. Bennett and K. Hauser, "Artificial intelligence framework for simulating clinical decision-making: a Markov decision process approach," *Artificial Intelligence in Medicine*, vol. 57, no. 1, pp. 9–19, 2013.

[28] U. Cortés, M. Sànchez-Marrè, L. Ceccaroni, I. R-Roda, and M. Poch, "Artificial intelligence and environmental decision support systems," *Applied Intelligence the International Journal of Artificial Intelligence Neural Networks & Complex Problem Solving Technologies*, vol. 13, no. 1, pp. 77–91, 2000.

[29] S. P. Joshi, J. R. Umaretiya, and S. B. Joshi, "Decision-making in preliminary engineering design," *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, vol. 5, no. 1, pp. 21–30, 1991.

[30] V. U Vincent, "Integrating intuition and artificial intelligence in organizational decision-making," *Business Horizons*, vol. 2021, no. 3.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 6323811, 10 pages
https://doi.org/10.1155/2022/6323811

Research Article

The Digitisation of Intangible Cultural Heritage Oriented to Inheritance and Dissemination under the Threshold of Neural Network Vision

Qin Wang

School of Design Arts,Wuxi Institute of Technology, Wuxi 214121, China

Correspondence should be addressed to Qin Wang; wangqin@wxit.edu.cn

Received 7 July 2022; Revised 25 July 2022; Accepted 1 August 2022; Published 31 August 2022

Academic Editor: R. Mo

As economic transformation and social change have taken place, digital media and Internet technology have developed rapidly, and the need for the development of digital research on nonheritage oriented to transmission and dissemination is becoming increasingly strong. On the basis of cultural identity and cultural sovereignty, intangible cultural heritages (ICHs) are disappearing at an alarming rate, ICHs are facing an existential crisis, and their development prospects are worrying. With the emergence of neural network models, the development of digital technology has revived many things that were on the verge of extinction. Traditional cultures and industries that originally seemed unrelated can take on new forms with the help of digital technology, thus enabling ICHs to find new ideas for development. This paper takes Huizhou ICH as an example and tries to design and construct a Huizhou ICH database and a digital map of Huizhou ICH, and it establishes a database for management and operation. The paper applies the information space theory to study the use of digital resources of ICH under the threshold of the neural network and employs digital information technology to recode, reconstruct, and interpret ICH. As a result, traditional ICH items are displayed in a digital way, which improves the public's recognition of digital ICH items, thereby promoting the inheritance and dissemination of ICH.

1. Introduction

With the rapid development of information technology, digitalization has gradually become the development trend of ICHs' protection and dissemination due to its advantages of nondestructiveness and wide dissemination [1]. Due to the large scale, various types, and diverse contents of ICH resources, some valuable knowledge cannot be fully revealed and utilised because of the single way of describing and organising existing resources, and the knowledge that cannot be expressed in a visible way is precisely the important content and difficulty of ICH heritage preservation and dissemination.

In the 20th century, with the continuous development of industrial civilization, both urbanization and commercialisation have ushered in great changes, but the preserved cultural heritage is disappearing at an alarming rate, and even on the verge of extinction. The natural environment for survival is being increasingly destroyed, and the Earth's environment is at risk. The survival and development of ICH has become an important issue in front of the world.

The traditional oral transmission and inheritance method has hindered the modern transmission and development of ICHs. Many ICHs have no corresponding written records or transmission methods, relying entirely on oral transmission. As most of the inheritors are not well educated and economically disadvantaged, and some are already old, young people go out to work and often lose interest in the work of the inheritors. Under the impact of modern digital technology, information technology, and network technology, various memory-based and skill-based ICHs that rely on a unique cultural space are mainly taught by the inheritors themselves are at risk of being lost. Moreover, excessive commercialisation and industrial development have, to some extent, compromised the integrity and authenticity of the cultural ecology, and ICHs are not protected and passed on in a scientific and effective manner. Urbanisation, commercialisation, and industrialisation have cut into the harmonious relationship formed by the long-term interaction between man and nature, destroying the ecological environment of ICHs, and the cultural space on which they depend is gradually disappearing, with many ICHs facing a serious crisis of survival. Then, there is the lack of localised excavation of ICHs. The cultural connotations, cultural characteristics, and cultural values of many ICHs are unclear, and some even appear homogenised.

In response to the abovementioned heritage environment, ICH is facing an existential crisis, and its development prospects are worrying. In particular, the transmission of the message of originality and integrity has been neglected by economic interests. As the continuation of human civilization as a proof, only inheritance and dissemination will not let our history disappear. But with the passage of time, any heritage will mutate or die, any means of heritage protection is extending the life cycle of heritage, and of course, digital means is no exception.

Technological change is not a quantitative change of gain or loss, but an overall ecological change [2]. Digital media technology is conducive to promoting the development of traditional ICH in the direction of digitisation, thereby promoting the digital theory and practical research of ICH projects. The development of digital and information network technologies has promoted the survival and development of culture and given rise to new business models and modes of digital cultural industries. The 39th statistical report released by the China Internet Network Information Centre shows that, as of December 2016, the size of China's Internet users reached 731 million, the Internet penetration rate reached 53.2%, and the number of mobile phone users reached 695 million, accounting for 95.1% of the total. The impact of digital technology is very broad, and in addition to the dramatic impact on the development of the cultural industry itself, the impact on cultural consumers cannot be ignored. In the digital age, the cultural consumption behaviour and content preferences of consumers, especially those of the digital native generation, have changed radically, with personalised consumption, customised content, borderline access, and more.

The social environment inevitably changes, and with it, the cultural memories rooted in these social environments are forgotten, and texts from the past lose their self-evidentness and become subject to interpretation. In this new era, texts cannot continue to be self-evident and are caught in the tension created by the time and space gap between the text and the present. For ICH, in addition to traditional methods of transmission and conservation, new conservation concepts have emerged in response to technological change. Digital technology has revived many things that were on the verge of extinction. Traditional culture and industry, which originally seemed unrelated, can form new links and cross borders with the help of digital technology, expanding many new business models and finding new ideas for heritage development. New digital media technologies can be used to capture, process, reproduce, interpret, preserve, and disseminate cultural heritage, shifting it from uniqueness, subjectivity, and exclusivity to plurality, digitality, and sharing. More importantly, the intervention of digital technology not only changes the way cultural heritage exists but also changes the public's perception of cultural heritage and the way it is passed on.

For a thousand ICHs, it is not only necessary to pass them on but also to use modern technology and discourse systems to reinterpret them, giving them a new cultural connotation and making them relevant to modern life. In other words, digital information technology can be used to recode, reconstruct, and interpret ICH and make it accessible to the public at the cultural level of digital technology. While reconstructing ICH, digitisation also brings about changes in the public's behaviour of perceiving, consuming, and using heritage, i.e., humanistic perception, reflection, and action on the digital dimension of technology, which is a universal value orientation. Digital heritage resources can vividly reflect the cultural forms of various regions. There are many studies on language and rituals, and it is more helpful to promote dialogue with the public with different cultural psychology. It guides people to love and learn about traditional culture at a large level.

In 1954, Japan established a record-keeping system for nonheritage in the revision of the Cultural Property Protection Law [3], in which the University of Oz in Japan carried out digital preservation of the lion dance, a non-heritage of the Oz region [4]. In 1976, the Library of Congress in the United States established the folklore preservation centre, which has been engaged in the digital preservation of folklore materials. The Traditional Music Archive at Indiana University and the World Music Archive at Harvard University use digital technology to preserve music archives. The European Union Cultural Heritage Online (ECHO) makes full use of the potential of new media for archival preservation, academic and educational exploration, cultural heritage sharing, etc. [5]. The Asia-Pacific Cultural Centre for UNESCO has established the Asia-Pacific database on African heritage to comprehensively document the cultural messages of African heritage. The Louvre in France, the Uffizi museum in Italy, the British Museum in the UK, the National Museum of Ethnology in Japan, and a number of other museums and libraries have transformed and preserved heritage information digitally. It is worth mentioning UNESCO's Memory of the World project, the European Union's Content Creation Initiative, the United States' American Memory, and the Microsoft Asia Research Institute's 'eHeritage' programme in the United States.

Foreign theoretical research on the digital aspects of cultural heritage predates that of China. Weiss [6] argues for community-based documentation, dissemination, and production of ICHs and the complexity of digital heritage in preserving heritage in the digital age. Ginsburg et al. [7] argues that the design of applications for visual interfaces and foreground browsing of digital cultural heritage facilitates the understanding of cultural heritage. Li [8] discusses the Indian government's digital preservation measures for cultural heritage resources, etc. Salins et al. [9] emphasise the need to use digital technology for the preservation and use of cultural heritage to ensure that cultural heritage should be made available to the public, i.e., for recreation, education etc. While the above analyses the digitisation of cultural heritage in terms of regional or project or technological realisation, Cameron and Kenderdine and Kalay et al. [10] theorizes the digitisation of cultural heritage from a cultural and media-critical perspective. In addition, the annual conferences related to multimedia and virtual world heritage have created a topic on the digitisation of cultural heritage, and scattered studies have been published from time to time.

Most of the current research on the digitisation of nonheritage is concerned with the digital framework of heritage, the language of digital technology, database issues, digital museums, digital libraries, etc., and rarely with the issue of the discourse of cultural subjects. There are more discussion on the digitisation technology of cultural heritage and few discussion on the cultural issues, ethical issues, intellectual property issues, rights and interests of cultural subjects, and issues of the interpretative power of heritage in the process of digitisation of cultural heritage. Domestic academic research on the digital conservation of ICH and ICH databases is still in its infancy, but both academics and practitioners have realised the importance and practical significance of digitisation of ICH. However, due to the multidisciplinary design and knowledge span of ICH digital research, many of the existing domestic studies face the dilemma of systematic and indepth research, the disconnection between technology and practice, and the lack of in-depth communication between theoretical scholars and technical personnel. In order to change this situation, this paper takes Huizhou ICH as an example for research, tries to design and construct Huizhou ICH database and Huizhou ICH digital map, and discusses the management and operation of the database. Using information space theory in the context of neural networks, this paper studies the use of digital resources of ICHs, uses digital information technology to recode, reconstruct, and interpret ICHs, and makes them accessible to the public at the cultural level of digital technology, thus promoting the transmission and dissemination of ICHs.

2. Concepts Related to Nonheritage

2.1. Cultural Heritage. "Cultural heritage" is a composite term, with "culture" as a qualifier and "heritage" as a central term to distinguish it from other heritage such as natural heritage and agricultural heritage. The term "cultural heritage" is commonly used today and has its origins in the western expressions for human cultural heritage in its material form, which was officially used in 1972 when the 17th UNESCO World Heritage Convention was adopted in Paris. It was only in 1985 that the internationally accepted term "cultural heritage" was universally used in China. Following the recommendations of some UNESCO member states, and along with the deepening of heritage conservation and awareness worldwide, the international community has formed a consensus that cultural heritage. The authors of [11] includes both tangible and intangible cultural heritage dimensions. The tangible cultural heritage refers to the materialised and physical remains of human culture, and the scope and classification of the human heritage system is given in Figure 1.

2.2. Intangible Cultural Heritage. UNESCO established a Folklore Expert Committee in 1982 to conduct a research on ICH through the establishment of an ICH department, and the concept of "intangible cultural heritage" has been gradually adopted by the socially accepted international community. The intangible heritage includes arts, crafts, languages, literature, folklore, myths, customs, beliefs, rituals, and other traditional cultures. Influenced by Japan's use of the concept of "intangible cultural property," UNESCO changed the title of "Nonphysical Heritage" to "Intangible Heritage" in 1992. In 2011, UNESCO adopted the Universal Declaration on Cultural Diversity and other important documents, officially using the term intangible cultural heritage and emphasising the importance of intangible heritage to the world's cultures and to the cultural diversity and creativity of humanity. In 2003, UNESCO adopted the Convention for the Safeguarding of the Intangible Cultural Heritage, the revised version of which states that the term "intangible heritage" refers to the heritage of communities and groups, and sometimes to individuals. The convention emphasises the heritage, ecological, and creative aspects of ICH as well as the sense of cultural identity and continuity. Specific information on the conceptual classification of UNESCO's ICH is given in Table 1.

The concept of ICH and the related elaboration used in this paper follow the content of the ICH Law adopted in 2011, as shown in Table 2. In terms of broad categories, they are divided into 5 major categories, but in terms of subdivision, they are divided into 6 subcategories or 10 categories, as shown in Figure 2.

The representative items of ICH are classified as shown in Figure 2.

2.3. Digitisation. In this paper, Huizhou is not only a geographical concept but also a cultural concept, which is an important area in the sense of traditional Chinese culture. Due to the objective and subjective factors, the objects examined and researched in this paper are the nonheritage in the Huizhou region, which refers to the whole territory of Huangshan City (Huangshan district, Huizhou district, Tunxi district, She county, Man county, Qimen county, Xuancheng city, and Jixi county), and the Huizhou cultural region has established a four-level list system of nonheritage. Most of the unique and representative ICH projects in Anhui Province originate from the Huizhou region. The region's nonheritage projects cover the ten categories of the state council's nonheritage list, and all types of nonheritage projects have different characteristics.

With the development of mathematics and computer science, digitisation [12, 13] has become a highly modern term of art. The original meaning of digitisation was to replace the traditional decimal system with binary. Binary was initially used for data processing, i.e., the digital

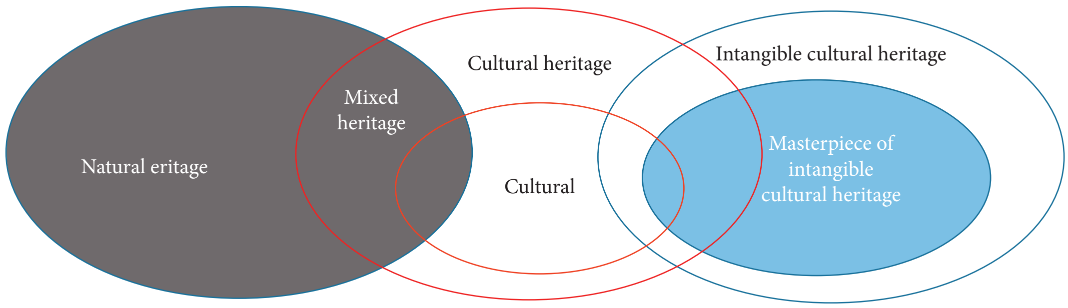


Figure 1: Scope and classification of the human heritage system.

Table 1: Classification of UNESCO's ICH concepts.

Serial number	Category
1	Oral traditions and expressions
2	Performing arts
3	Social practices, rituals, and festivals
4	Knowledge and practices about the natural world and the universe
5	Traditional handicrafts

Table 2: Classification of the ICH Law of the People's Republic of China.

Serial number	Category
1	Traditional oral literature
2	Traditional fine arts, calligraphy, music, dance, drama, and acrobatics
3	Traditional techniques, medicine, and calendars
4	Traditional etiquette, festivals, and other folk customs
5	Traditional sports and recreation
6	Other intangible cultural heritage

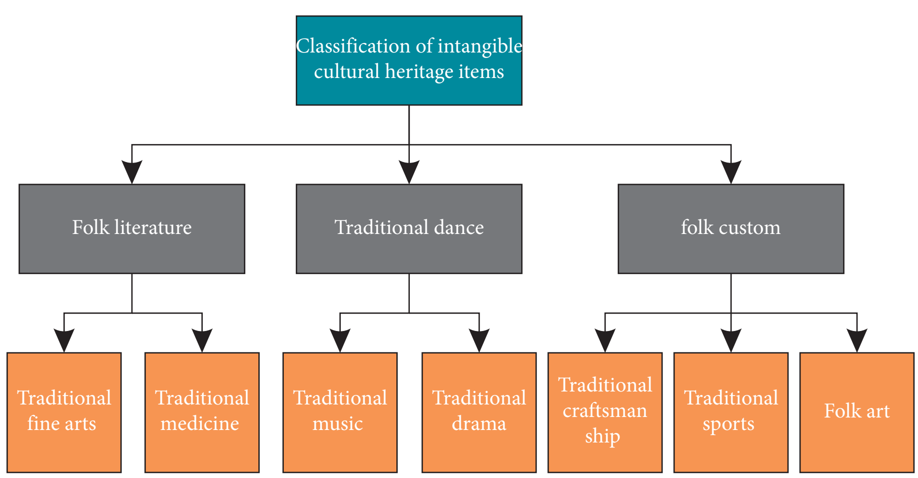


Figure 2: Classification of representative items of ICH.

representation of text, language, images, sound, etc. Through data conversion and processing, in the binary system, each 0 or 1 is a bit, and the bit is the smallest unit of data storage. Today, the concept of digitisation is much more than a combination of bits—one is no longer a static and intuitive symbolic meaning.

Digital technology has permeated all areas and levels of society and has greatly influenced how people learn, work, and live. However, it is not easy to define the term digital not only because digitisation is a new phenomenon but also because it is changing rapidly and is clearly dynamic, making it difficult to define a fixed concept.

This article will focus on digitisation at the cultural level, and the concept of digitisation covered below refers mainly to the various digital technologies using the digital 0 s and 1 s as a medium. As there are many different kinds of ICHs, there are real difficulties in using a unified digital technology for digital preservation, transmission, and industrialisation.

In the real world, certain ICHs have lost their audiences, soil, and space for survival, but the virtual reality built by digitisation opens up a new space for the survival and development of ICHs, and ICHs can find their own way of communication.

3. Neural Network Models

3.1. Principle of Artificial Neural Networks. The principle of a neural network [14, 15] is as follows: a neural network has multiple neurons connected, somewhat similar to the Internet, which is an interconnection of servers. The neurons in a neural network are connected in a certain hierarchy and rules. In general, a neural network consists of three layers, namely, the input layer, the hidden layer, and the output layer, while the hidden layer can choose more than one neuron, as shown in Figure 3. Each neuron processes the input data, and the specific processing operations are based on the transfer function and learning function. Neurons between different convolutional layers are connected to each other, and there are no neurons between the same layers. When the weight of the connection layer changes, the neural network completes the training process.

We think of each connection in a neural network as a processing unit, which transforms and aggregates the input data. Each processing unit has a limited capacity, but when there are multiple processing units, it is more powerful. The individual processing units work in concert to adjust the weights on each connection and eventually simulate the data accurately. This adjustment process is a learning process that relies on a series of learning algorithms. The process is complex and may take place thousands of times. Learning ends when the error is within tolerance or a predetermined number of times has been reached. If the trained network model achieves a predetermined accuracy, the data can be processed.

There are various artificial neural networks currently available, such as BP neural networks [16, 17], adaptive resonance theory [18, 19], avalanche networks [20, 21], bidirectional associative memory [22, 23], and Boltzmann machines, but the more widely used is the BP network model. The BP neural network model is described in detail in Section 3.2.

3.2. The BP Neural Network. The error back-propagation algorithm is often used in the training of multilayer feedforward networks, and this method is needed to adjust parameters during the training process, which is called the BP network. The basic idea of the BP algorithm is that there is a forward learning signal and a reverse error signal in the learning process. In the forward learning signal, it is transmitted from the input to the hidden layer and then to the output layer, whereas the reverse signal is in the opposite direction. The reverse error signal is passed from the output layer, and when there is a large error between the network output data and the desired data, that error is passed to the hidden layer. The error is assigned to different neurons depending on the weights of the connections as it passes and finally passes out of the input layer.

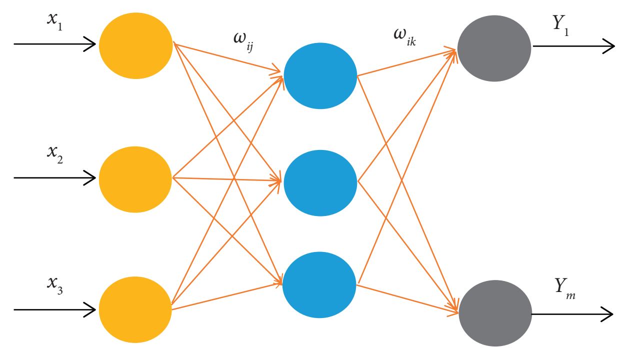


FIGURE 3: Schematic diagram of the neural network structure.

Figure 4 depicts a portion of a multilayer forward network in which two signals are being transmitted:

(i) The working signal (indicated by a solid line), which is passed forward from the input to the output, which is a function of the input and the weight

(ii) The error signal (indicated by a dashed line), which is passed backwards from the output to the input, which is a function of the error, and which is assigned to each neuron during the transfer according to the connection weights

The specific process of the BP neural network [24] is as follows:

(i) Determine the input, output, number of layers, and the number of neurons in the hidden layer of the neural network; there can be more than one hidden layer, but there is no theoretical basis for determining the exact number of layers.

(ii) Setting up the transformation function and the learning function in the neural network, both of which are described in detail.

(iii) The selected sample size, generally obtained according to an empirical formula.

(iv) The initialization of the weights, which can be given randomly but may lead to an increase in learning time, is given in this paper to calculate the initialization of the weights.

(v) Enter samples in sequence.

(vi) Calculate the output of each layer in turn.

(vii) Finding the back-propagation error for each layer.

(viii) Correction of individual weights and thresholds by the weight adjustment formula.

(ix) The output of each layer is calculated by the new weights until the error is less than a predetermined threshold. The following transformation functions in the neuron can be used [25]:

(i) Step function:

$$f(x) = \begin{cases} 1, & x \geq 0, \\ 0, & x < 0. \end{cases} \tag{1}$$

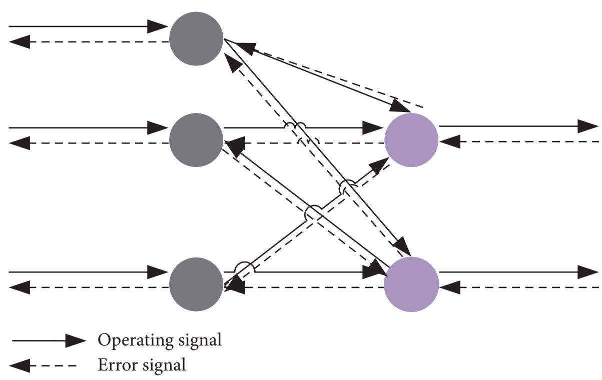


Figure 4: Schematic diagram of the forward working signal and the backward error signal.

(ii) S-type function: usually in [0, 1] or [−1, 1] within the continuous value of the monotonic differentiable function. Commonly used response functions are S-type sigmoid functions.

$$f(x) = \frac{1}{1+e^{-\mu x}}. \tag{2}$$

(iii) Proportional functions:

The proportional function is a linear function, often used between the hidden and output layers, which has a filtering effect while the linear propagation function does not change the nonlinear relationship of the output of the hidden layer. It is expressed as follows:

$$f(x) = kx. \tag{3}$$

(iv) Symbolic functions:

$$f(x) = \begin{cases} 1, & x \geq 0, \\ -1, & x < 0. \end{cases} \tag{4}$$

(v) Saturation function:

$$f(x) = \begin{cases} 1, & x > \frac{1}{k}, \\ kx, & -\frac{1}{k} < x < \frac{1}{k}, \\ -1, & x < \frac{1}{k}. \end{cases} \tag{5}$$

(vi) Hyperbolic functions:

$$f(x) = \frac{1-e^{-\mu x}}{1+e^{-\mu x}}. \tag{6}$$

The transformation function is determined by the complexity of the problem; if the problem is more complex, the S-shaped function or hyperbolic function is used because their nonlinear mapping ability is relatively strong for simple problems can be used linear function or step function.

BP is the most widely used network model of any neural network. And among the BP neural networks, the more commonly used is the single hidden layer network, which is also known as a three-layer feedforward network or a three-layer perceptron.

The formula for the output layer in a BP neural network is as follows:

$$\begin{aligned} o_k &= f(\mathrm{net}_k), \quad k = 1, 2, \ldots, l, \\ \mathrm{net}_k &= \sum_{m}^{j=0} w_{jk} y_j, \quad k = 1, 2, \ldots, l, \end{aligned} \tag{7}$$

where f is the activation function, w is the weight, and y is the corresponding eigenvector.

The implicit layer is calculated as

$$\begin{aligned} y_j &= f(\mathrm{net}_j), \quad j = 1, 2, \ldots, m, \\ \mathrm{net}_j &= \sum_{n}^{r=0} v_{rj} x_r, \quad j = 1, 2, \ldots, m, \end{aligned} \tag{8}$$

where y_j represents the output of the hidden layer, f is the activation function, and v represents the weight corresponding to the vector x.

In the above equation, the transfer functions f(x) are all unipolar sigmoid functions.

$$f(x) = \frac{1}{1+e^{-x}}. \tag{9}$$

Depending on the application, a bipolar sigmoid function can also be used.

$$f(x) = \frac{1-e^{-x}}{1+e^{-x}}, \tag{10}$$

where e is the logarithmic base and x is the vector.

Together, the above equations form a mathematical model of a three-layer feedforward network.

4. Experiments in Digitisation of Nontraditional Heritage

4.1. Digital Indicators Established. This paper examines and researches the ICHs in the Huizhou region, which refers to the whole of Huangshan City, including Huangshan district, Huizhou district, Tunxi district, She county, Man county, Qimen county, Xuncheng city, and Jixi county. Digitize all intangible cultural heritage items to establish a database, recode, reconstruct, and interpret all traditional intangible cultural heritage items for digital display.

Digitised ICH items from four different regions were randomly pushed out to the general public, an online questionnaire was created to test the classification of ICH items from different regions, and the dissemination of ICH was promoted in the campaign. Data from different users were collected, the results of the questionnaires were analysed, and dissemination measures were improved based on the results.

4.2. Digital Research. The results of the online questionnaire collection, a total of 1,000 online questionnaires, were distributed in this experiment, 856 were returned, and the information of all the data collected was tallied. The results of the four regions' public favourite nonheritage items are shown in Table 3.

A comparison of the public's favourite ICH items in these four regions is shown in Figure 5. As can be seen from Figure 5, there is not much difference between the number of popular favourite ICH items and the most impressive items, which shows that the favourite items are more associated with the most impressive items.

The 856 online questionnaires were also categorised into different regions of the ICH, and the data information for each of the five scenarios, including all correct and only one wrong, was tallied as shown in Table 4.

The breakdown of all questionnaires is shown in Figure 6. The data in Table 4 are obtained from an electronic questionnaire withdrawn from the general public, which consists of two parts. The first part is to let users choose their favourite ICH projects, so as to deepen the user's impression of all ICH projects, indirectly for the purpose of dissemination. The other part is to push different ICH projects to users, after the user understands the background and source of all projects, and finally set up a question for the user to classify the ICH projects browsed; there are a total of four categories here, so as to collect the user's understanding information data and facilitate further research.

As can be seen in Figure 6, the digitisation of ICH items is more impressive to the general public, as the number of misclassified items is inversely proportional to the size of misclassified items, indicating that the digital dissemination of ICH plays an important role in the expansion of traditional ICH and contributes to its revival.

The next step was to use BP neural networks to predict the classification results of the questionnaire data and test the prediction results under different parameters. Six hundred traditional ICH data based on recoding, reconstruction, and interpretation were selected as the training set, and the remaining questionnaire data were used as the test set for prediction. To ensure the accuracy of the results, all experiments were repeated five times, and the classification accuracy data under different iterations are shown in Table 5.

A visual comparison of the classification accuracy data for different numbers of iterations is shown in Figure 7. From Figure 7, it can be seen that different numbers of iterations have a greater impact on the prediction results, with the best classification results being achieved when the number of iterations is 5.

After the network was trained, we tested it with 20 test samples, and the results and error analysis are shown in Tables 6 and 7.

A visual comparison of the test results and error data for the test set is shown in Figure 8.

As can be seen in Figure 8, the actual discrepancies between the sample values and the test values are not very different; the trends are roughly the same for both, and the errors always remain within a certain range. The errors of the 20 sets of samples were next analysed, and the results of the comparison are shown in Figure 9.

Table 3: Data on the public's favourite ICH projects in the region.

Category	Favourite items	The most impressive project
National	325	256
Provincial	260	301
Municipal	150	128
County	121	171

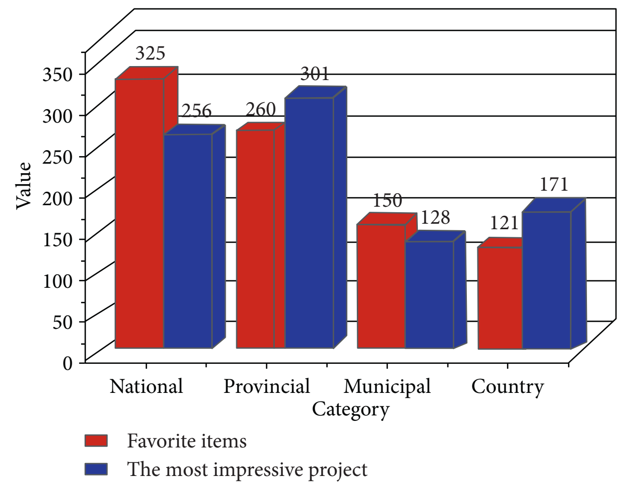


Figure 5: Comparison of popular favourite ICHs in the four regions.

Table 4: Classification of ICHs from 856 questionnaires.

Category	Value
All correct	288
Divide the wrong one	216
Divide the wrong two	198
Divide the wrong three	103
All errors	51

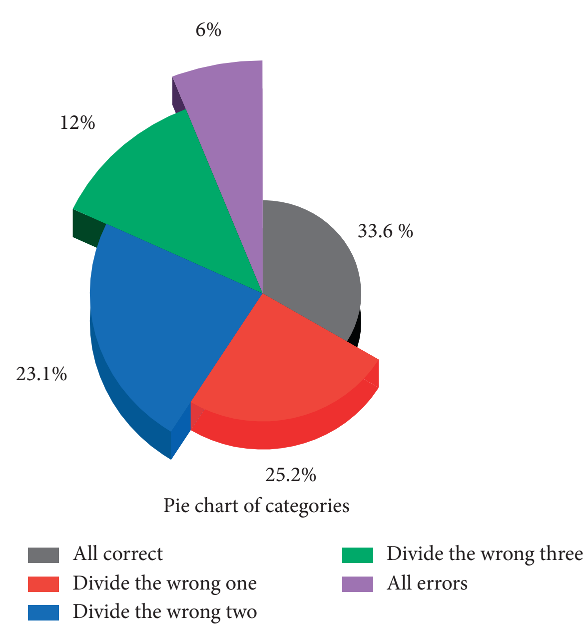


Figure 6: Visual comparison pie chart of the results of the online questionnaire ICH classification.

Table 5: Classification accuracy data with different number of iterations.

Number of iterations	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Experiment 5	Average value
3	0.86	0.84	0.88	0.86	0.86	0.86
4	0.90	0.88	0.92	0.93	0.91	0.91
5	0.92	0.94	0.93	0.93	0.94	0.93
6	0.91	0.89	0.90	0.88	0.91	0.90
7	0.87	0.89	0.90	0.86	0.90	0.89

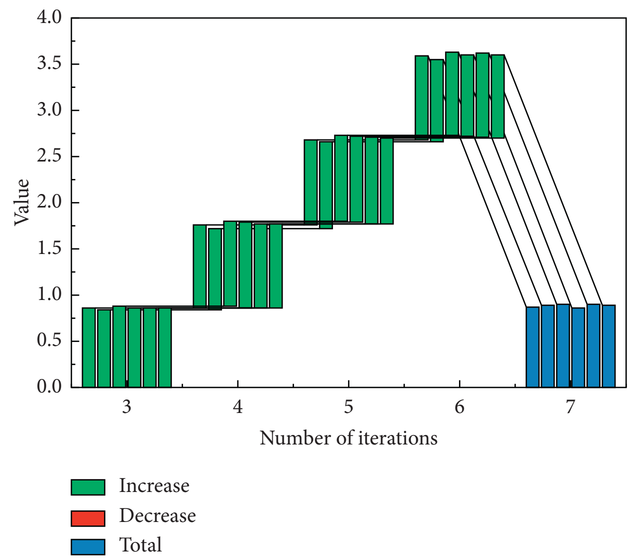


Figure 7: Visual comparison of classification accuracy data for different iterations.

Table 6: 1–10 test results and error analysis data.

Serial number	1	2	3	4	5	6	7	8	9	10
Sample values	0.44	0.35	0.28	0.54	0.34	0.38	0.39	0.41	0.64	0.60
Test value	0.48	0.38	0.29	0.55	0.38	0.36	0.42	0.38	0.69	0.65
Error	0.04	0.03	0.01	0.01	0.04	0.02	0.03	0.03	0.05	0.05

Table 7: 11–20 test results and error analysis data.

Serial number	11	12	13	14	15	16	17	18	19	20
Sample values	0.61	0.63	0.64	0.79	0.73	0.87	0.89	0.93	0.98	0.92
Test value	0.64	0.66	0.70	0.83	0.77	0.90	0.95	0.97	0.95	0.94
Error	0.03	0.03	0.04	0.04	0.04	0.03	0.04	0.04	0.03	0.02

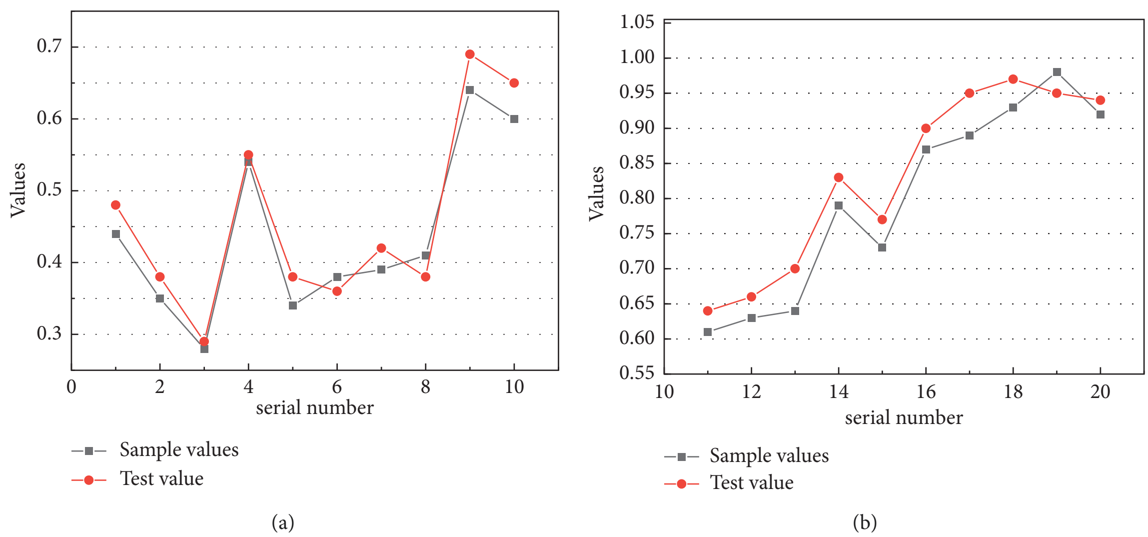


Figure 8: Visual comparison of sample and test value data for the test sample. (a) 1–10 samples and test data. (b) 11–20 samples and test data.

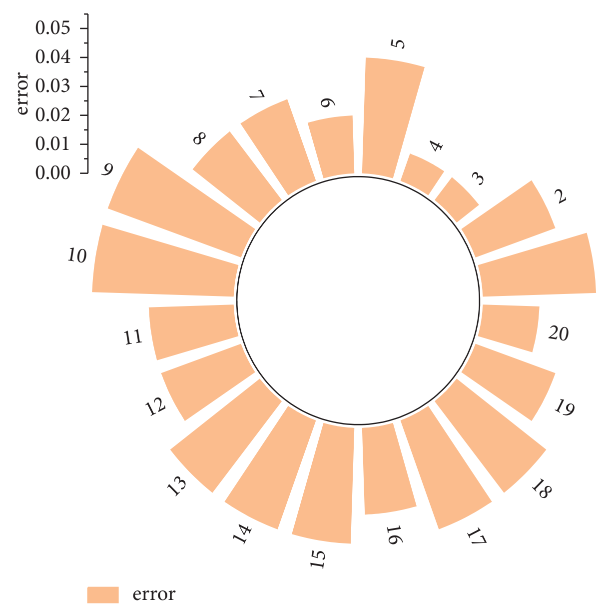


Figure 9: Comparison of test sample errors.

As can be seen in Figure 9, the closer to the centre the smaller the error, and vice versa, the larger the error. As a whole, it can be seen that the error between all samples and the number of tests is not very different, indicating the accuracy of the digital description of all the nonheritage in the Huizhou region, which is conducive to the widespread dissemination of traditional nonheritage.

5. Conclusion

This paper studies the impact of digital technology on the dissemination of intangible cultural heritage based on the current situation of Huizhou's intangible cultural heritage through on-the-spot research. The paper uses digital information technology to recode, reconstruct, and interpret ICH items, present them digitally, and establish an online questionnaire survey to collect data on the public's liking of digitised ICH items and the comparison of different ICH items by category. Predictive analysis of the ICH test sample was carried out under the BP neural network, and the error between the sample data and the test data was controlled within a reasonable range, making the digital ICH items more acceptable to the public and, at the same time, raising people's attention to the ICH items during the data collection process, further promoting the dissemination of ICH, and enabling the innovative development of the traditional ICH dissemination channels. The conclusions of this paper are mainly derived from the collected digital information data, and the dissemination of ICH projects is still facing a huge test for a wider range of personnel needs.

Next, we will gather more public information so that the ICH project can be more widely disseminated.

Data Availability

The datasets used during the current study are available from the author on reasonable request.

Conflicts of Interest

The author declares no conflicts of interest.

Acknowledgments

This work was sponsored in part by the Soft Science Research Topic of Wuxi Association for Science and Technology taking “Huishan Clay Man” in Wuxi as an example (subject no. KX-21-C168).

References

[1] D. Peng, X. Liu, and S. Sun, “Information perspective: digital theory of intangible cultural heritage protection,” *Journal of Computer-Aided Design and Graphics*, vol. 21, no. 1, pp. 117–123, 2008.

[2] A. Asman, *Memory Space: The Form and Change of Cultural Memory*, Peking University Press, Beijing, China, 2016.

[3] L. C. Baudrillard, *Quan Zhigang Translated Consumer society*, Nanjing University Press, Nanjing, China, 2000.

[4] P. Bourdieu, *On Television*, Shenyang Liaoning Education Publishing House, Shenyang, China, 2000.

[5] M. Marshall, *He Daokuan’s Translation of Understanding the Medium: On the Extension of Man*, The Commercial Press, Shanghai, China, 2000.

[6] E. B. Weiss, *Treats Future Humanity Fairly*, Beijing Law Press, Beijing, Chjna, 2000.

[7] F. Ginsburg, L. Abu-Luhd, and B. Larkin, *Ding Huimin in Translation of Media World: A New Frontier of Anthropology*, Beijing Commercial Press, Beijing, China, 2015.

[8] P. Li, *Cultural Anthropology in Translation by Marvin Harris*, Shanghai Oriental Publishing House, Shanghai, China, 1988.

[9] M. Salins and M. Wang, “Zongze Hu Translating Sweet Sorrow,” *Beijing: Life. Study*, Xinzhi Sanlian Bookstore, Beijing, China, 2002.

[10] N. Bozeman, *Translated by Zhang Yan entertaining to death. The Disappearance of Childhood*, Guilin Guangxi Normal University Press, Guilin, China, 2009.

[11] N. Bozeman, *He Daokuan Translation Technology Monopoly*, Peking University Press, Beijing, China, 2017.

[12] Z. Zheng, “The digital development trends of circular economy in China,” *Environment, Resource and Ecology Journal*, vol. 5, no. 4, pp. 25–32, 2021.

[13] B. Barłusiak, “The PERFECT DIGITAL DEVELOPMENT,” *Dental Products Report*, vol. 55, no. 9, p. 16, 2021.

[14] B. Belhassen and S. C. Matthew, “Design principles of convolutional neural networks for multimedia forensics,” *Electronic Imaging*, vol. 29, pp. 77–86, 2017.

[15] C. Zhang, “Research on principle and application of convolutional neural networks,” *Earth and Environmental Science*, vol. 440, no. 4, pp. 561–567, 2020.

[16] X. You, C. Su, and Y. Wang, “A review of BP neural network algorithm improvement,” *Private Science and Technology*, vol. 3, no. 4, pp. 146-147, 2018.

[17] L. Chen, “A review of research on neural network BP algorithm,” *Computer Knowledge and Technology*, vol. 6, no. 36, pp. 10364-10365, 2010.

[18] R. Naga, D. Satya Siva Kiran, C. Aparna, and S. Radhika, “Classification of groundwater by applying simplified fuzzy adaptive resonance theory,” *IJDNE*, vol. 16, no. 2, pp. 167–176, 2021.

[19] B. Anitarani and P. Suvasini, “Role-based profiling using fuzzy adaptive resonance theory for securing database systems,” *International Journal of Applied Metaheuristic Computing*, vol. 12, no. 2, pp. 36–48, 2021.

[20] C. Hooper, “Law enforcement organisations take down Avalanche cybercrime ring - ScienceDirect,” *Computer Fraud & Security*, vol. 2016, no. 12, p. 13, 2016.

[21] S. Ji and M. Deng, “Avalanche dynamics on complex networks,” *American Physical Society*, vol. 12, no. 5, pp. 54–59, 2005.

[22] J. Cao, “Global asymptotic stability of delayed bi-directional associative memory neural networks,” *Applied Mathematics and Computation*, vol. 142, no. 2-3, pp. 333–339, 2003.

[23] J. Cao and M. Dong, “Exponential stability of delayed bi-directional associative memory networks,” *Applied Mathematics and Computation*, vol. 135, no. 1, pp. 105–112, 2003.

[24] Y. Wang and L. I. Tao, “Generalized neural network system and generalized BP algorithm,” *Small Microcomputer Systems*, vol. 5, no. 8, pp. 967–970, 2001.

[25] L. I. Pan-chi and X. U. Shao-hua, “Neural network training of processes based on continuous Walsh transform,” *Computer Engineering and Design*, vol. 6, no. 3, pp. 702-703+713, 2018.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 7666615, 11 pages
https://doi.org/10.1155/2022/7666615

Research Article

Intelligent Physical Education Teaching Tracking System Based on Multimedia Data Analysis and Artificial Intelligence

Feng Cao,[1] Maojuan Xiang,[1] Kaijie Chen,[2] and Ming Lei[3]

[1]*Department of Sports, Xidian University, Xi'an 710126, China*
[2]*Department of Physical Education, Northwest University, Xi'an 710069, China*
[3]*Department of Physical Education, Institute of Humanities and Social Sciences, Xi'an International Studies University, Xi'an 710128, China*

Correspondence should be addressed to Ming Lei; 201904217401@stu.zjsru.edu.cn

Received 14 July 2022; Revised 31 July 2022; Accepted 5 August 2022; Published 31 August 2022

Academic Editor: R. Mo

The education system begins a significant dimension characterized by continuous improvement and impacted by technology, society, and cultural developments. This pattern shows the need to enhance physical and athletic scientific training methods. To make the teacher's role more successful is the usage of computer systems and other computing infrastructure. This article aims to show the use of Information and Communication Technology (ICT) in the national association for physical education. The digitalization consequences in the profession can be synthesized into the following elements: an operating system for learning, application essential to finish actions, findings recording, movement monitoring, video processing, efficiency correlation and synchronization, object tracking and duration measurement systems, and exercise assessment. While physical activity and athletics are realistic activities, current instructional technology is not sufficient. An intelligent physical education tracking system (IPETS) is proposed in this research. This study analyses and investigates the methodology for the formative assessment of athletic knowledge in computer evaluation. During the first segment, the evaluation technique for the athletic education program was presented. The second step of the paper is to understand the contents of the mathematical formula of particular activity based on the complete approaches proposed for the theory of assessment. A different phase of modernizing teaching activities using the computerized quality education for Artificial Intelligence (AI) technology is established in this article. The experimental findings are high in identifying university students' regular exercise.

1. Introduction to the Physical Education System

With proper and effective teaching and learning environment, a physical training organization is designed to establish a rational and coherent learning method in the undergraduate syllabus [1]. When a person engages in physical exercise, the flow of blood to the brain rises. It enhances brain activity, care, memory, thinking, and intellectual abilities, the key to a kid's development [2]. By natural, overall conditioning, the physical activity seeks to develop physiologically, socially, intellectually, and creative persons at play [3, 4].

The curriculum is arranged according to the substance of the physical training. Consequently, the main components of materials of physical education organization can directly change the durability and characteristics, impacting the learning technique of the students [5, 6]. An effective school organization demands improved job allocation, appropriate management supervision, and the correct interpretation of information to alter pupils' functioning and characteristics [7].

It includes several activities to make it easier and more accessible for students to grasp via improved teacher co-ordination. Learners participate in physical workouts due to a low student/physical educator proportion [8]. Every availability is predicted to influence the access to proper management immediately; the specific focus is on physical tasks. The administration of special activities is vital to achieving instructional objectives [9].

The components of the sports education program are divided into so many subdivisions and graded into numerous elements to coordinate the instructional media of primary training and create them into a complete theoretical framework for strength and conditioning materials [10, 11]. Fitness conditioning directly teaches and enhances the skills and talents need to connect and communicate with everyone. It enables team and management skills to be developed and allows students to collaborate expertise in different areas of study [12].

For instance, this is accomplished by encouraging children to collaborate with others or participate in sports groups and community organizations as members in their schools and outside academia [13]. In addition, vertical and lateral organizational techniques for the general activity curriculum design are used to design the arrangement of information. Those two primary forms of organization and their specific contents have several essential and microfilm of organization [14].

The vertical arrangement emphasized that the sequence, coherence, and horizontal correlation gradually assessed the difficulties and depths of the material. The horizontal organization, content breadth, and scope richness underline material capabilities and side combinations [15, 16]. There are issues with the course content structure for strength conditioning and collaboration within the two organizations. The organization of the contents of the curricula should be rational and complete overall [17]. It shows that aligning the connection between the vertical and lateral organizations and guaranteeing that subtle educational content is converted into a comprehensive technical, conceptual framework is typically seen as the meaningful measure of education administration.

The structural complexity with numerous interrelated components is not usually a feature. Sophisticated, transparent processes can contain several elements and could be defined as simple systems capable of disassembling different pieces with predictable and accurate operations management. In other aspects, complex structures contain several elements, but each element is related to another. Any component is not inert but dynamic or adaptable.

The physical training goals in this article are given as follows:

(i) To the experimental nature of fitness exercise.

(ii) Practical repetition of the teaching content for physical exercise.

(iii) Design of the search rankings for university learners' physical activity (PA) depending on the complex analyses provides the grading and classification procedure.

The remaining of the paper is as follows: Section 2 denotes the background of the physical education systems. The proposed intelligent physical education tracking system (IPETS) is designed and developed in Section 3. Section 4 depicts the software analysis and performance measurements. Section 5 illustrates the conclusion and findings.

2. Background to the Physical Education Systems

Internet of Things (IoT) provided educational knowledge to pupils. The learner and teacher were allowed to attain zero verbal interaction using this interface. Teachers can upload activities on their cellular phones with specific timelines to complete the tasks [18]. In addition, employ Wireless sensors, which allow smartphone apps to monitor and respond to information from signals in the actual environment. Strategic and inventive position-oriented developments might raise the issues of a pupil.

Learning strives to resolve the connection gaps among learners and educators in their connections [19]. This article highlighted the need to safeguard such programs, considering the number of protection defects in hardware implementations to keep the process secure. In this study, the protection of such systems was enhanced through two document analyses, complex automated development and factory equipment, which examined possible risks and dangers in this sector [20].

Wyant and Baek suggested a noninvasive brain method for a physical intelligence activity coaching component and outlined an Intelligent Teaching System (ITS) experimentation learner module utilized for reliable evaluations of instruction in PA [21]. Inputs utilized for quantification are the power evaluated for the regular exercise function, as well as the actual energy costs were dependent on the same action. At the same time, an evaluation at a particular sixth grade was dependent on the education grade.

Without interaction from teachers, the students tested their physical, educational capacity, and other areas. A direct objective technique was employed in measuring PA, and its precision could be more than regarded as suitable. Pocock and Miyahara established the identification of human activities through IoT devices [22]. They utilized algorithms to evaluate four groups' behavior (wait, sit, stroll, and jog). Meanwhile, the programming of alerts before and afterward activities were delivered via a telemonitoring function with distant access [23].

This technique had been deployed with a favorable 75.8% success percentage. The method was meant to ensure that each patient followed every day the healing process through routine, practice, and physiotherapy [24]. Even though a safety analysis was required for an IoT methodology, the IoT method didn't depend on a network of wireless sensors; the relevant data were not susceptible to a particular audience.

The sensors-based PA identification and surveillance model utilizing IoT were suggested by Bertills et al. [25]. The advent of the Internet of things had turned PA research into an uncontrolled, open, and interconnected environment by interacting with diverse wearable applications and wearables. It can meet the new difficulties of IoT settings by using traditional PA software or how these techniques are used and upgraded effectively [26]. A structured assessment was done on PA investigations from a conventional level to explain the usage of IoT.

Initially, it covered the cutting-edge techniques employed by traditional PA methodology in healthcare care, such as vision, extracting, and recognition processes suggested by Shaik and Patil [27]. The report identified and addressed specific critical approaches and novel patterns of investigation and difficulties inside the physical activity recognition model (PARM) investigation in IoT contexts. Finally, this article discussed some situations in which PARM's intelligent health care had been effective in this area and potential future industrial uses. Sariyska and Montag started the notion of self for incentive and the connection between students and teachers throughout schooling [28].

The hypothesized connections between the research characteristics were explored using latent analysis of structural equation components [29]. The most significant change research resulted using the bootstrapping procedure has demonstrated excellent compatibility with the data. The approach had demonstrated increased autonomy, expertise, and connection and a more excellent self-determination rating among students experiencing an autonomous supporting environment [30]. Children's self-reported motivation levels had predicted physical education (PE) efforts and teachers' perseverance favorably. In PE contexts, the findings were examined to enhance student enthusiasm.

Giese and Ruin proposed identifying group activities and sensitivity analyses as a service (GASAaaS) [31]. Suppose detector information was compiled using a mechanism for sharing information across the different devices acquired via wearable sensors, Wristwatch sensors, and GASAaaS-specified built-in sensors [32]. This approach was exemplified by the solutions they offered to mudlarks and bushwalkers in a company. The model was shown as feasible and expressive as the recommended model. The method incorporates built-in sensors for mobile devices to identify activities or scenarios as customer services.

The inertial sensors generator had modeled and developed a variety of settings for demonstrating and evaluating the actions of the sportsperson, and the PE framework at the IPETS was presented in this study to solve these questions. Application of fuzzy inference system and output variable, in particular, can help to detect intensity inexactitude. The true importance of the notion of PA intensities language ambiguity was shown. Fuzzy sets were a suitable difficulty metric that translated the zone of potential misinterpretation into PA effectively.

3. Proposed Intelligent Physical Education Tracking System (IPETS)

This article proposes IPETS based on computer Information Technology (IT) for university physical education training and studies [33–37]. The PAs could be better defined by using fuzzy logic to a more robust analytical approach in this research amongst university graduates. The new schooling regime, a teaching notion focused on the learner, becomes vital for the concept of instructors throughout education. The notion based on each pupil must be appreciated. The mode of instruction involves in-time teacher-student contact. Just in this manner can instructors take into account the evaluation of PE training. This article created the computer assessment system for a particular activity to evaluate the course.

Figure 1 depicts the architecture of the proposed IPETS. It shows the PE and evaluation program's internet servers and client-side configuration. The program's primary purpose is to evaluate and analyze the instructor syllabus online, fix the management reports, examine the students, and evaluate them. During the evaluation, the application is executed by the administration.

The modern reform of education necessitates that the student is recognized as the center of learning [38–42]. The educational process must be participatory in real-time for instructors and learners. The regular *C* applications mode is a conventional computer education evaluation process which is a confined mode to evaluate the teaching method. By the objectives of the new educational reform, this article established a computerized assessment method for the take steps to address, which was controlling the servers and the customer. This method enables students to view and evaluate the evaluation findings and instructor effectiveness.

Teachers watch learners' actual insight and instructors and learner commitment in actuality. It is a practical computer training assessment system using the shared foundation type of the Website and the virtual servers. This procedure permits actual teacher-student contact and reduces customer-server shuttered redundant information, and enhances computing-related performance levels. The framework for an examination is more significant and much more reliable.

The wireless model of the proposed IPETS system is shown in Figure 2. It has a speed acquisition module, wireless transmission modules, video playback modules, sports control unit, and training data display module. It demonstrates that several outcomes were created by the learners who submitted their teacher assessments. The IPETS, sophisticated analyses based on a computerized evaluation process and complete evaluation findings, can examine such assessments.

3.1. Analysis on the Application Advantages of Artificial Intelligence Technology in Physical Education. Constitute a smart environmental education sector, the time slice limitations for learning activity and the communicating of advanced educational assets, etc., about the issues in traditional colleges special activity and taking into account the significant benefits of AI technology in effectively integrating atmosphere and virtuality [43–48]. The conventional college primary fitness emanates fresh life via further discussions on the specific use of AI technology in special university activities.

3.1.1. Deep Integration of Internet Plus Technology with Virtual Reality (VR). The Online plus focuses on the effective integration of the Website's easy and diverse benefits with conventional industries, modifying the structural base of old systems to improve production effectiveness. A simulation model that can replicate actual worlds is the core

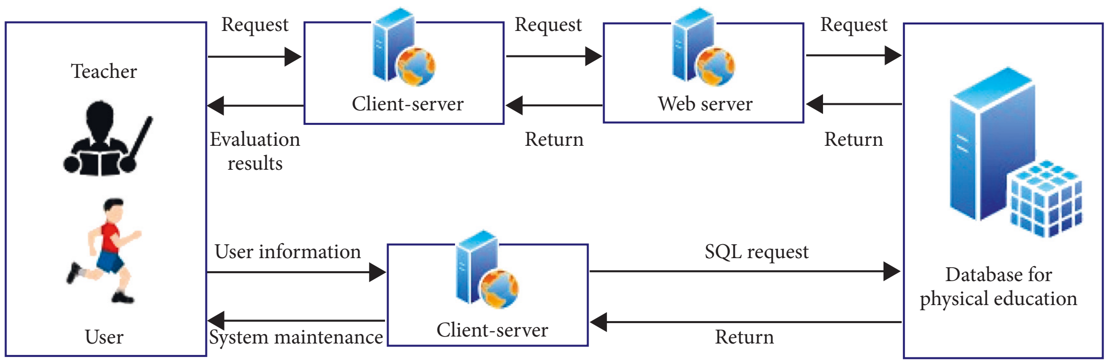


Figure 1: The architecture of the proposed IPETS.

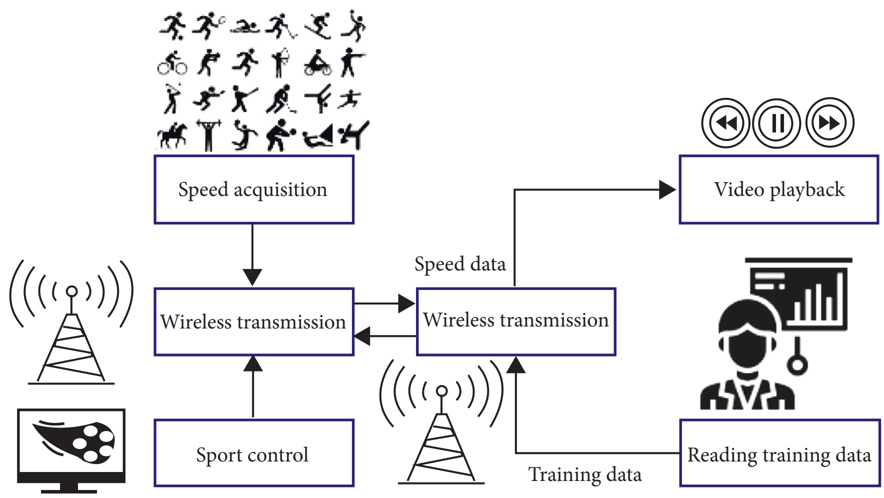


Figure 2: Wireless model of the proposed IPETS system.

of Virtual Reality technology. It can reduce the gap in teaching methods between professors and students. Physical activity is brief and poor teachers are unable to offer additional professional support to kids. Employing more broad and advanced Internet technologies and Virtual worlds can tackle this problem.

The atmosphere for physical training is thus more colorful and three-dimensional in academic institutions and can be achieved using the following example. Create an effective system for athletics education and advertise it as an element of the curriculum content to be used by learners. Second, educators encourage to submit information connected to sporting instruction on the system, such as videos, knowledge gained, and other material on the framework, so that the students search and study. Knowledge gained is available to people. Third, 3D animation is integrated into video instruction given the rigorous criteria for the primary training level of sports motions. In addition, it allows students to study more academic and practical information and to have a greater immersive storyline.

3.1.2. Intelligent Recognition Technology. Smart identification is a technological method by which information is gained by identifying apparatus, data processed by the computer science department is automatically identified, and feedback is provided. Smart identification technologies can contribute to enhancing the issue of unilateral instruction in conventional physical training. First, the foundation of classroom instruction is the kids coming to school. Smart identification technologies can thus be used in classes.

The face identification system could profoundly remove signing options on everyone else, relative to conventional testing techniques such as questions reply and protocol inspection. It makes sure that inspection processes can be finished over a very brief time and therefore consider the consequences of class restraint and effectiveness. Second, smart identification technologies might also be used in the specialized athletic school curriculum. If pupils have problems throughout the education process, people typically get perplexed or do irregular sporting activities repeatedly.

The Smart Identification System analyses and processes the material acquired through voice commands and identification of the face to better understand human behavior and the stage of education and retrieve the data to the teacher. To create more contact among professors and students better understood, the educators must better understand the challenges faced by the learners in physical training and enhance their educational consequences of making the sporting schooling more entertaining.

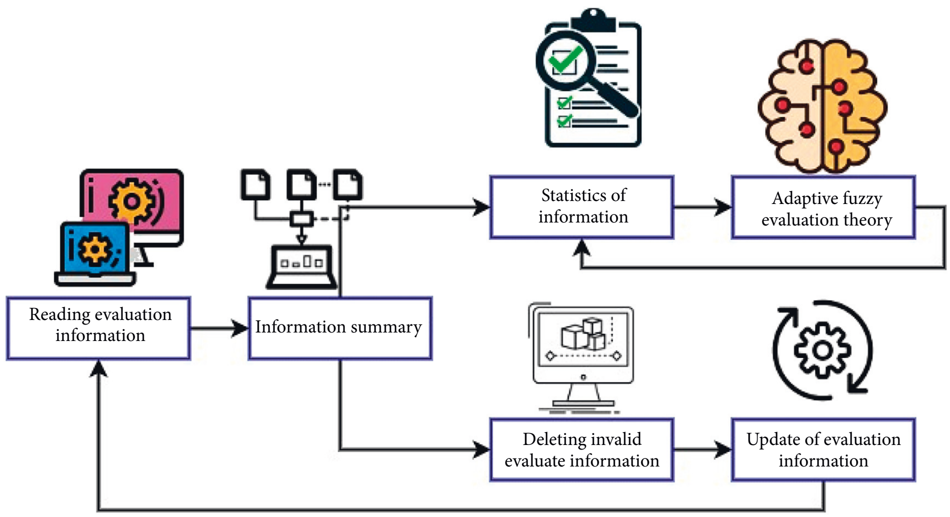


Figure 3: Visualization module of the proposed IPETS.

3.1.3. Big Data Analysis Technology. The technique of large datasets comprises the display of data, information gathering, and predictive analytics essentially. Data visualization technology might also display data to consumers intuitively as a core necessity of analytical instruments. The rising economy can be assessed by hypothesis testing associated with visual analysis findings.

Figure 3 shows the visualization module of the proposed IPETS system. It shows the processes of application of information visualization analytics technology comprising data collecting, access to data, collection of information, statistics, and evaluation of correspondences. Applying big data analytics technologies in university physical training increases the scientific, standardized, and effective instruction and administration. Instructors can put their assignments on the educational platform and encourage their pupils to send pictures or films. The endpoint of the learning management system is utilized to extract and process information from students with VR, computer vision, and other technology.

Then the learner's learning outcomes are returned to their professors in information via visualization technology. Instructors can thus change the direction of the instruction and its development in time by using the data gathered. At the very same moment, the technologies of large datasets might also enable pupils to study more. Analyzing students' progress, participation, and workplace presentation on the learning system can identify their interests and learning patterns to present a specific picture of students. It offers students better adapted physical, educational content, and techniques to reach the objective of pupils through additional predictions and analyses.

3.2. Fuzzy-Level Evaluation. Learners can use the proper vector table activities $p_y(y = 1, 2, \ldots, n)$ to calculate cognitive function $s(p_y)$, and it is expressed in equation (1):

$$s\left(p_y\right) = \frac{V_{p_y}(1)}{V_{p_y}(1) + V_{p_y}(-1)}. \tag{1}$$

As demonstrated in equation (1), which amount of incorrect uses of subcoagnitifist abilities in the actions are represented $V_{p_y}(1)$ and $V_{p_y}(-1)$, respectively, the proper vector of utilization is derived in equation (2).

$$S = \left(s_1, s_2, \ldots, s_n\right) \in (0, 1). \tag{2}$$

The elements of the vector are denoted s_x. From the assessment process, the right rate vector is determined in equation (3).

$$S = \left\{\frac{2/3}{2, 1/2}, 1, 0, 1\right\}. \tag{3}$$

Therefore, the area M varies from $(0, 1)$, which is the correct range for use. The fuzzy M sub-sets are student performance evaluations (great, good, and bad). M is the "good" fuzzy subset; $M_w(m)$ is the W member feature, and it is expressed in equation (4).

$$\mathrm{M_w(m)} = \begin{cases} \dfrac{1}{1 + \mathrm{m}/0.14}, & \mathrm{m} > 0, \\ \\ 0, & \text{else}. \end{cases} \tag{4}$$

In the application of the vectors and the mathematical explanations of the fundamental component activity (m): the participation rank of the sub-cognitive abilities is established using the suitable vector of use, and it is expressed in equation (5).

$$W = \left\{w_1, w_2, \ldots, w_n\right\} \in (0, 1). \tag{5}$$

Y's enrollment is highly qualified as a sub-cognitive and it is expressed w_x. The notion is that the activities are P groups and that there are multiple assessment vectors. The assessment matrix is expressed in equation (6).

$$HM = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ w_{p1} & w_{p2} & \cdots & w_{pn} \end{bmatrix}. \quad (6)$$

The elements of the evaluation matrix are denoted w_{ij}. The rows in the formulation of the matrices HM correspond to an evaluation vector for a series of questions. As each P activity plays different roles (some practice is typical in specific unit tests), different training sets fit different weight levels. A weight matrix is denoted in equation (7).

$$B = \{b_1, b_2, \ldots, b_p\} \text{ and } \sum_{y=0}^{p} b_y = 1. \quad (7)$$

The elements of the weight matrix are denoted b_i. The sum of all the elements of the weight matrix is always one. The weighted mean algorithm provides an exhaustive assessment of "a" vector. The learning capacity of the pupils is denoted in equation (8).

$$k = B \times HM = \{k_1, k_2, \ldots, k_n\}. \quad (8)$$

Test scores, $k_1, k_2, \ldots, k_n$ denoted age, acquisition effectiveness, psychological circumstances, training, mood, etc., are considered the input nodes of the multimodal network, the values of outputs of network C where multivariate and three-dimensional spaces are mapped nonlinearly and expressed in equations (9) and (10).

$$\text{Output} = \{\text{grade}, k_1, k_2, \ldots, k_n, w_1, w_2, \ldots, w_n\}, \quad (9)$$

$$C = \{C_1, C_2, C_3\}. \quad (10)$$

As indicated above in C_1, C_2, C_3, the masters of skills, ideas, and implementation are referred to by the master of pupils. The test scores are denoted $k_1, k_2, \ldots, k_n$, the weight matrix elements are denoted $w_1, w_2, \ldots, w_n$, it would be helpful to integrate the idea into computers' digital technologies to make the program more efficient; and the assessment model's design uses the fuzzy set theory to evaluate the computational formula.

The fuzzy technique was used in the PA of university students to grade the issue. Whenever massive datasets are accessed, the advanced clustering technique updates incomplete data and utilizes prior classification results to enhance efficiency. The rankings hypothesis is that the quicker the group of the team approaches the grade. The consequence is an excessive dependency on the first clustering center and the fall into the grouping state because of the randomized clustering centers, leading to grouping outcomes and the instabilities in a statistical distribution.

3.3. Physical Education and Training System Based on Machine Learning and Internet of Things. The physical education training model based on machine learning is designed in this section. The proposed IPETS system uses IoT to support physical education training and machine learning to enhance system performance.

The workflow of the proposed IPETS system is shown in Figure 4. It uses a human-machine interface module and controller to track the performance of the athlete. It depicts the physical school curriculum data gathering architecture based on machines and the IoT. Whenever the client is training for sports, the sensor transfers to the actual training systems through the transmitter via the velocity information recorded from the speed collecting module. The proper training system receives the velocity data and directs the media player to move the image appropriately. The system delivers in live time the tracking information to the trainer accordingly. The platform gathers data on sports education, modifies the size of the module, and mimics the sporting tea procedure. This study presents a true athletic training systems control method. The algorithms innovate the essential phases of the early data processor and athletics computation of the authentic athletic training program principally. This approach can increase the accuracy of training examples, enhance the teaching environment and eventually enhance the science and efficiency of the complete system. Dynamically mounting incline data, partially polynomial fitted energy, and start streaming video speed are all three phases of the genuine motion control approach.

Multimedia technology provides a good visual actuality feeling and delivers a distinct sports encounter from 3D virtual athletics. Virtual items can be included in actual films. To accomplish smooth connectivity, it first creates virtual celestial objects in line with the actual geometries. The imaginary character motion should also be displayed in the actual video in the immediate environment.

This issue is a spatial problem, since owing to the presence of 3D objects, the geometrical restrictions imposed through the real world vary in time. The motive is to add imaginary elements to actual films. The synthesis of faithful animation, i.e., its responsiveness to virtual item movements, should achieve an acceptable standard compared to genuine human behavior. Once the new foot position has been identified in the camera, the increasingly virtual character must forecast these walkers' future movement locations and modify their movements to prevent a crash with actual objects.

It has devised a technique in this article to incorporate imaginary characters smoothly into certain actual online scenarios. This work gives an algorithm for real-world sports training that replicates the effects of genuine sports activities in real-world images by innovating pre-data treatment and athletics methods. It then offers a Unique Path Planner. The online avatar can change its movements to maintain peace and achieve the predetermined goal, some statically indeterminate restrictions established by physical and virtual worlds.

Second, it presents an enhanced 3D virtual segmentation leveraging the camera prototype's information and the closed environment's topology. It removes several visual artifacts, ensures the final blended image quality, and

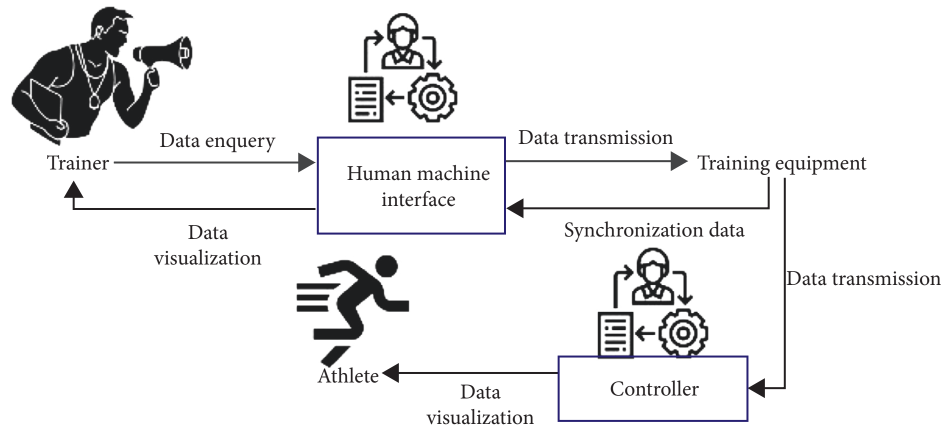


Figure 4: Workflow of the proposed IPETS system.

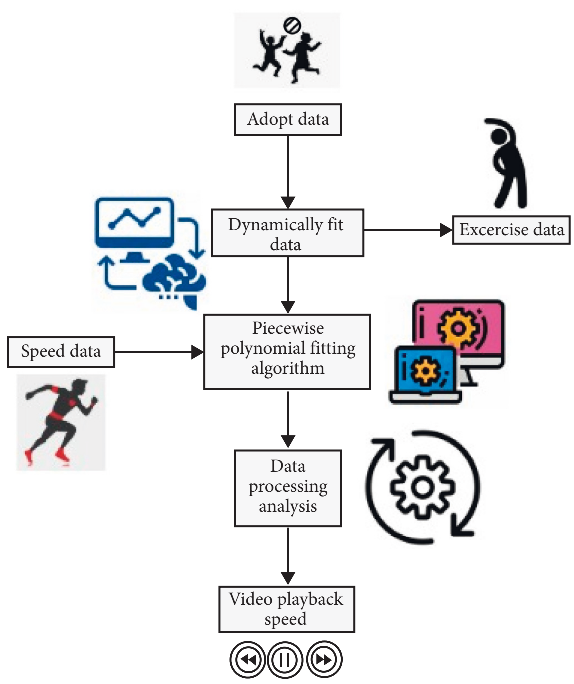


Figure 5: Visualization module of the proposed IPETS system.

incorporates imaginary characters into an actual online video. This article develops an exact hybrid translation based on each imaginary organism's 2D matrix to integrate seamlessly.

The screen module is the main component of the whole system. Meaningful variations in athletic training examples can be reflected in the viewing chart created by the display unit. The complete display unit consists of five pieces: diagram 1, diagram 2, histogram, dispersion diagram, and pie diagram. A range of different indications is provided by the subsystem and diagrams of different measures per the demands.

The style and behavior of the user interface are equilibrated. "Consumer experience" is defined as the subjective emotions of the user whenever a product is used. That demonstrates significant subjective adaptability dependent on the user firm's age, technology, geography, and culture. Furthermore, the consumer user can't interface only to end-users at requirement specification.

Figure 5 indicates the visualization module of the proposed IPETS system. It gets adaptive data from the camera and fits the data with the previously trained exercise data. User speed is found based on the calculation, and video playback is given to the display unit. Each extended data doesn't interact with one another from the point of view of information transfer. There is no network connectivity between both the expansions, and every extension interacts with only the hosting. Every extender sends the information to the host in the programmed training stage between both the hosts and the extender. Only directed extensions are free workouts and unexpected activity data kept, and data retroactivity to ordinary people after a series of exercises displayed.

This component of the information cannot be requested twice and is destroyed locally regularly. The hosts are accountable for extending registrations, managing customer data, and publishing normal training plans again for the extensions plan to learn program. The total network customer module design is displayed. The client needs no big number of system data from the servers, and a single instance can satisfy performance expectations. It got a message in the messaging buffer list in sequence in the asynchronous socket callbacks procedure and then used the list to process and understand them. The processor thread has first to enroll the subscriber and record the data into the listening table. An update calls the relevant function to process the respective messages according to the audience table and protocols name.

Preparing in sports is an exceedingly difficult business. Even though the nation's urbanized places are rich with a machine learning algorithm, it does not suffice to digitize and automate, just limited current research fully. It enhanced the initial inertial motion control system due to the connection judgment and posture corrections based on contact location restrictions. The restoration of postures on the part of the trainee is further accurate. It also puts reflection points at the outer border of the training ground to get the athletic trainer data directly.

Table 1: Average participation ratio analysis of the proposed IPETS.

Number of dataset	Male (%)	Female (%)
50	79	69
100	81	72
150	83	75
200	79	71
250	75	68
300	72	65
350	71	62
400	68	59
450	64	57
500	62	55
550	59	52
600	57	49

Table 2: Evaluation score analysis of the proposed IPETS.

Number of dataset	Male (%)	Female (%)
50	42	76
100	46	78
150	49	75
200	51	72
250	48	68
300	46	71
350	45	75
400	48	78
450	50	81
500	53	78
550	51	75
600	49	72

The connection among the vertical structure shows the relationship among certain components in various fitness training stages following the initial curriculum. The significance of lateral connection in the syllabus of physical training increased the consistent link between the subject matter of the physical learning environment at different stages or test scores. It brings the syllabus content in line with the subject's cohesive manner, the cognitive and physical phases, and the formation of the young person.

Teacher exercise increases children's capability and willingness to participate in a range of physical exercises both within and without schools. A strong athletic education curriculum strongly encourages all kids to do a lot of physical exercises. They create a broad variety of skills and apply techniques, approaches, and composing ideas effectively. The students think through what they do, assess the circumstances, and make decisions when they do. Indeed, they are looking for strategies to enhance their outcomes as well as other outcomes.

4. Software Analysis and Findings

The relationship between strength and PA is complicated and bilateral. Several research studies have shown a substantial correlation between organizational endurance and PA, allowing reasonably active people to surpass PA, or the two. Empirical work has shown that exercise enhances health. All sporting activities aim to improve the health and wellbeing of the pupils. The simulation outcomes of the proposed model are verified in this section, with the dataset containing 100 students and five physical education students.

Table 1 depicts the average participation ratio analysis of the proposed IPETS. The simulation is done by varying the number of datasets from a minimum of 50 to a maximum of 600 with a step size of 50. The respective simulation outcomes, such as the average participation ratio, are analyzed separately for the male and female participants. As the size of the dataset increases, the respective performance of the user decreases. The proposed IPETS with fuzzy-based event tracking and performance evaluation model exhibits higher performance in all situations.

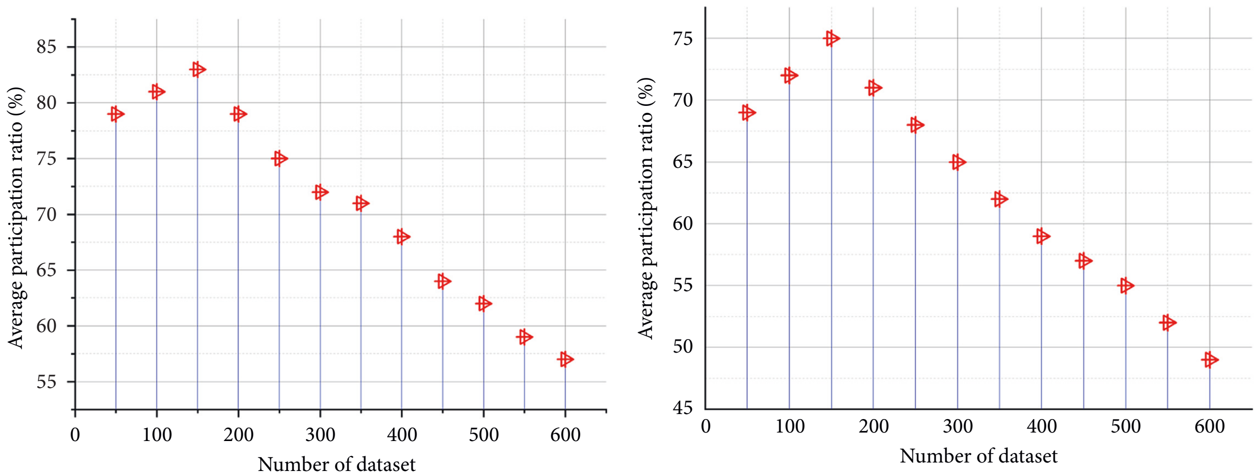
Table 2 indicates the evaluation score analysis of the proposed IPETS. The simulation outcomes for the male and female candidates are analyzed separately for the evaluation process, and their scores are compared. As the size of the dataset increases, the respective evaluation scores also vary. The proposed IPETS with fuzzy evaluation model and event tracking system accurately evaluates the efficiency of the participants and hence produces better results than others models. The male candidates perform well than the female candidates.

Figures 6(a) and 6(b) indicate the average participation analysis of the male and female candidates of the proposed IPETS, respectively. The simulation is done by varying the size of the dataset from the minimum level to the maximum level. As the dataset varies, the respective outcome of the participants also decreases. The proposed IPETS with fuzzy evaluation and event tracking models exhibit higher performance for all the criteria. The male candidate's performances well in all physical education activities than the female candidates because of their high stamina and fitness.

Figures 7(a) and 7(b) show the evaluation score analysis of the male and female participants using the proposed IPETS. The simulation analysis of the male and female candidates is done using the proposed IPETS and the given dataset. The size of the dataset is varied from minimum to the maximum level for the simulation analysis. The male candidates exhibit higher performance than the female candidates because the male candidates have a higher fitness level and higher stamina for long-time workouts. The proposed IPETS with a fuzzy evaluation system produces results with higher accuracy, which helps to better coaching and training of the students.

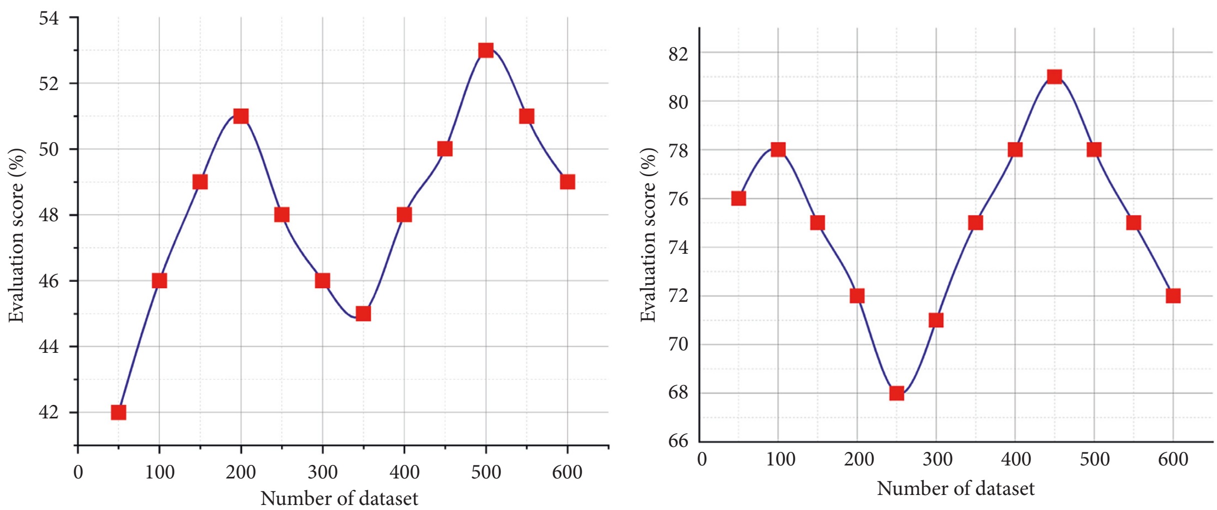
The data processing evaluation analysis and the data processing speed analysis of the proposed IPETS are shown in Figures 8(a) and 8(b). The proposed IPETS is designed, and the performance of the system is evaluated using the given dataset. There are 25 candidates (male and female) taken for the analysis from the given dataset. Their respective simulation outcomes, such as data processing score and speed, are analyzed and plotted. The proposed IPETS with a fuzzy evaluation system exhibits higher accuracy in evaluation with lower complexity.

The proposed IPETS is designed, analyzed, and tested in this section. The simulation outcomes such as processing



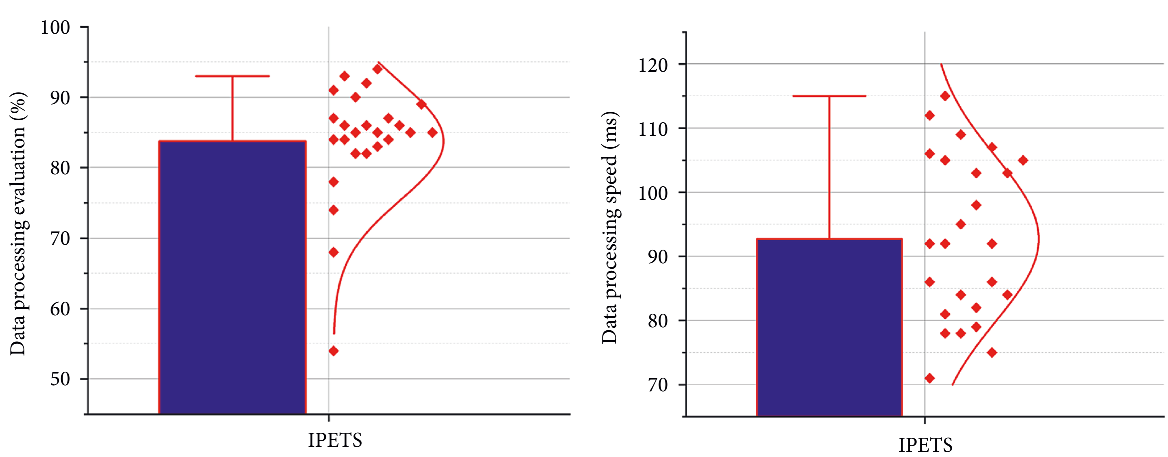
(a) (b)

Figure 6: (a) Average participation ratio analysis of male candidates. (b) Average participation ratio analysis of the female candidates.



(a) (b)

Figure 7: (a) Evaluation score analysis of the male participants. (b) Evaluation score analysis of the female participants.



(a) (b)

Figure 8: (a) Data processing evaluation analysis of the proposed IPETS. (b) Data processing speed analysis of the proposed IPETS.

speed, evaluation score, and the proposed IPETS with the given dataset are analyzed. The results show that the proposed IPETS with event tracing and fuzzy-based evaluation models produces higher results with lower system complexity.

5. Conclusion and Findings

This article provides an intelligent physical education tracking system (IPETS) for undergraduate training and education sciences with computer-based digital technologies. The assessment of athletics courses is performed using the computing growth of IT and the improvement of the PE systems. Instructors and learners must communicate information promptly in physical instruction, improving their instructional performance and efficiency. The information systems assessing mode of the PE curricular and the development of a PE mathematics evaluation method relies on IPETS. Finally, the study introduced the programmed computer code of the PE curricular evaluation framework using the *C* program and the expert's weighting factor for the PE curricula appraisal. It develops a computerized assessment instrument for physical training, which offers a theoretical basis for the technique of physical training. In the future, the system outcomes can increase by using the deep learning model.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by Research on "The 14th National Games" Boosting the Construction of "Internet Plus Sports" Culture in Shaanxi (2021ND0314) and IT-Based Teaching Reform and Practice of All Members in Sports Club participating in the competition (B21040).

References

[1] H. Song, C. E. Montenegro-Marin, C. E. Montenegro-Marin, and S. krishnamoorthy, "Secure prediction and assessment of sports injuries using deep learning based convolutional neural network," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 3, pp. 3399–3410, 2021.

[2] Y. Wang, B. Muthu, and C. B. Sivaparthipan, "Internet of things driven physical activity recognition system for physical education," *Microprocessors and Microsystems*, vol. 81, p. 103723, 2021.

[3] W. Huifeng, A. Shankar, and G. N. Vivekananda, "Modelling and simulation of sprinters' health promotion strategy based on sports biomechanics," *Connection Science*, vol. 33, no. 4, pp. 1028–1046, 2020.

[4] P. F. Sheron, K. P. Sridhar, S. Baskar, and P. M. Shakeel, "Projection-dependent input processing for 3D object recognition in human robot interaction systems," *Image and Vision Computing*, vol. 106, p. 104089, 2021.

[5] K. Cengiz, R. Sharma, K. Kottursamy, K. K. Singh, T. Topac, and B. Ozyurt, "Recent emerging technologies for intelligent learning and analytics in big data," in *Multimedia Technologies in the Internet of Things Environment*, pp. 69–81, Springer, Singapore, 2021.

[6] A. Gunasekaran and P. Narayanasamy, "Analyzing the network performance of various replica detection algorithms in wireless sensor network," *Journal of Computational and Theoretical Nanoscience*, vol. 15, no. 3, pp. 989–994, 2018.

[7] G. Manogaran, P. M. Shakeel, H. Fouad et al., "Wearable IoT smart-log patch: an edge computing-based Bayesian deep learning network system for multi access physical monitoring system," *Sensors*, vol. 19, no. 13, p. 3030, 2019.

[8] S. Bevinakoppa, A. Alazab, and T. Jan, "Design of computer networking courses with major in cyber security," *International Journal of Education and Literacy Studies*, vol. 3, 2018.

[9] G. Amudha, "Dilated transaction access and retrieval: improving the information retrieval of blockchain-assimilated internet of things transactions," *Wireless Personal Communications*, pp. 1–21, 2021.

[10] J. Gao, H. Wang, and H. Shen, "Machine learning based workload prediction in cloud computing," in *2020 29th International Conference on Computer Communications and Networks (ICCCN)*, pp. 1–9, IEEE, 2020, August.

[11] M. K. Chowdary, T. N. Nguyen, and D. J. Hemanth, "Deep learning-based facial emotion recognition for human–computer interaction applications," *Neural Computing & Applications*, pp. 1–18, 2021.

[12] G. Manogaran, S. Baskar, P. M. Shakeel, N. Chilamkurti, and R. Kumar, "Analytics in real time surveillance video using two-bit transform accelerative regressive frame check," *Multimedia Tools and Applications*, vol. 79, no. 23-24, pp. 16155–16172, 2020.

[13] P. M. Kumar, C. S. Hong, F. Afghah et al., "Clouds proportionate medical data stream analytics for internet of things-based healthcare systems," *IEEE Journal of Biomedical and Health Informatics*, p. 1, 2021.

[14] S. K. Sahu, D. P. Mohapatra, J. K. Rout, K. S. Sahoo, and A. K. Luhach, "An ensemble-based scalable approach for intrusion detection using big data framework," *Big Data*, vol. 23, p. 3423, 2021.

[15] D. J. Samuel R, G. Manogaran, G. Manogaran et al., "Real time violence detection framework for football stadium comprising of big data analysis and deep learning through bidirectional LSTM," *Computer Networks*, vol. 151, pp. 191–200, 2019.

[16] V. T. Pham, T. N. Nguyen, B. H. Liu, and T. Lin, "Minimizing latency for multiple-type data aggregation in wireless sensor networks," in *2021 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, IEEE, 2021, March.

[17] G. Manogaran, N. Chilamkurti, and C. H. Hsu, *Editorial Note: Machine Learning for Visual Analysis of Multimedia Data*, Springer, China, 2020.

[18] Y. Zhai, D. Li, C. Wu, and H. Wu, "Urban park facility use and intensity of seniors' physical activity–An examination combining accelerometer and GPS tracking," *Landscape and Urban Planning*, vol. 205, p. 103950, 2021.

[19] W. M. Roberts, D. J. Newcombe, and K. Davids, "Application of a constraints-led approach to pedagogy in schools: embarking on a journey to nurture physical literacy in primary physical education," *Physical Education and Sport Pedagogy*, vol. 24, no. 2, pp. 162–175, 2019.

[20] Y. Hutzler, S. Meier, S. Reuker, and M. Zitomer, "Attitudes and self-efficacy of physical education teachers toward inclusion of children with disabilities: a narrative review of international literature," *Physical Education and Sport Pedagogy*, vol. 24, no. 3, pp. 249–266, 2019.

[21] J. Wyant and J. H. Baek, "Re-thinking technology adoption in physical education," *Curriculum Studies in Health and Physical Education*, vol. 10, no. 1, pp. 3–17, 2019.

[22] T. Pocock and M. Miyahara, "Inclusion of students with disability in physical education: a qualitative meta-analysis," *International Journal of Inclusive Education*, vol. 22, no. 7, pp. 751–766, 2018.

[23] S. Yeon-hee, F. Miao, and R. Li, "Retracted article: analysis of PM2.5 air pollution sources based on image recognition and tracking of sports training process," *Arabian Journal of Geosciences*, vol. 14, no. 16, pp. 1630–1712, 2021.

[24] S. D. Wilkinson, D. Penney, L. Allin, and P. Potrac, "The enactment of setting policy in secondary school physical education," *Sport, Education and Society*, vol. 26, no. 6, pp. 619–633, 2021.

[25] K. Bertills, M. Granlund, Ö Dahlström, and L. Augustine, "Relationships between physical education (PE) teaching and student self-efficacy, aptitude to participate in PE and functional skills: with a special focus on students with disabilities," *Physical Education and Sport Pedagogy*, vol. 23, no. 4, pp. 387–401, 2018.

[26] K. Schenker, "Health (y) education in health and physical education," *Sport, Education and Society*, vol. 23, no. 3, pp. 229–243, 2018.

[27] R. Shaik and L. V. Patil, "Human tracking mechanism for institutions using RFID and facial recognition," in *Techno-Societal 2020*, pp. 355–364, Springer, Cham, 2021.

[28] R. Sariyska and C. Montag, "An overview on doing psychodiagnostics in personality psychology and tracking physical activity via smartphones," *Digital Phenotyping and Mobile Sensing*, pp. 45–63, 2019.

[29] A. Steinert, M. Haesner, and E. Steinhagen-Thiessen, "Activity-tracking devices for older adults: comparison and preferences," *Universal Access in the Information Society*, vol. 17, no. 2, pp. 411–419, 2018.

[30] W. Guo and J. H. Kim, "Designing augmented reality learning systems with real-time tracking sensors," in *International Conference on Applied Human Factors and Ergonomics*, pp. 269–276, Springer, Cham, 2021, July.

[31] M. Giese and S. Ruin, "Forgotten bodies–an examination of physical education from the perspective of ableism," *Sport in Society*, vol. 21, no. 1, pp. 152–165, 2018.

[32] Q. Hui, "Motion video tracking technology in sports training based on Mean-Shift algorithm," *The Journal of Supercomputing*, vol. 75, no. 9, pp. 6021–6037, 2019.

[33] L. Zhou and Q. Tang, "Construction of a six-pronge intelligent physical education classroom model in colleges and universities," *Scientific Programming*, vol. 2022, p. 202, 2022.

[34] M. Liu, "Delphi method combined with computer-assisted teaching of information fusion to explore intelligent physical education in colleges and universities," *Mobile Information Systems*, vol. 2021, p. 221, 2021.

[35] L. h. Li, J. c. Hang, Y. Gao, and C. y. Mu, "Using an integrated group decision method based on SVM, TFN-RS-AHP, and TOPSIS-CD for cloud service supplier selection," *Mathematical Problems in Engineering*, pp. 1–14, 2017.

[36] L. H. Li, J. C. Hang, H. X. Sun, and L. Wang, "A conjunctive multiple-criteria decision-making approach for cloud service supplier selection of manufacturing enterprise," *Advances in Mechanical Engineering*, vol. 9, no. 3, p. 168781401668626, 2017.

[37] Y. Dong, "Analysis of intelligent physical education teaching scheme based on 5G communication + VR technology," *Mobile Information Systems*, vol. 2022, p. 22, 2022.

[38] Y. Wang, Y. Han, and Q. Wang, "The application of motion trajectory acquisition and intelligent analysis technology in physical education teaching in colleges and universities," *Security and Communication Networks*, vol. 2022, p. 22, 2021.

[39] L. Li and C. Mao, "Big data supported PSS evaluation decision in service-oriented manufacturing," *IEEE Access*, vol. 8, pp. 154663–154670, 2020.

[40] L. Li, C. Mao, H. Sun, Y. Yuan, and B. Lei, "Digital twin driven green performance evaluation methodology of intelligent manufacturing: hybrid model based on fuzzy rough-sets AHP, multistage weight synthesis, and PROMETHEE II," *Complexity*, vol. 2020, no. 6, pp. 1–24, 2020.

[41] M. Liu and J. Bu, "Deep integration of physical health education based on intelligent communication technology," *Journal of Healthcare Engineering*, pp. 2021–6, 2021.

[42] K. Hao, K. Zhao, and H. Cao, "Intelligent analysis on the rationalization of children's physical education curriculum based on recurrent neural networks," *Scientific Programming*, vol. 2022, 2022.

[43] L. Li, B. Lei, and C. Mao, "Digital twin in smart manufacturing," *Journal of Industrial Information Integration*, vol. 26, no. 9, p. 100289, 2022.

[44] L. Li, T. Qu, Y. Liu et al., "Sustainability assessment of intelligent manufacturing supported by digital twin," *IEEE Access*, vol. 8, pp. 174988–175008, 2020.

[45] Ge Wang, "Feasibility analysis and discrete dynamic modeling of physical education teaching strategy based on intelligent computing," *Computational Intelligence and Neuroscience*, p. 987, 2017.

[46] Q. Xu and Q. Jia, "The cultivation and training effect of the subconscious mind in physical education and training by intelligent internet of things network computing," *Security and Communication Networks*, p. 2022, 2022.

[47] G. Wu and X. Zhang, "Realization of wireless sensors and intelligent computer aided teaching in physical education and training," *Wireless Communications and Mobile Computing*, vol. 24, p. 32864, 2022.

[48] H. Hao and S. Hu, "Recommendation optimization of physical education for developing the intelligence of autistic children following intelligent collaborative filtering algorithm," *Mobile Information Systems*, vol. 202, p. 2022, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 1551264, 10 pages
https://doi.org/10.1155/2022/1551264

Research Article

Research on the Innovative Marketing Model of E-Commerce Platform Driven by Big Data in the Era of Network Economy

Xiaofeng Zhou and Ying Zhou

Jiangxi University of Technology, Nanchang 330098, China

Correspondence should be addressed to Ying Zhou; zhouying@jxut.edu.cn

Received 10 July 2022; Revised 2 August 2022; Accepted 13 August 2022; Published 29 August 2022

Academic Editor: R. Mo

With the booming development of Internet information technology, e-commerce platforms in the era of network economy have undergone great changes, triggering a new marketing model change. Innovative research on marketing models can help the transformation and development of small and medium-sized e-commerce companies, which has important practical significance and theoretical value. The prediction of e-commerce sales is one of the key aspects of the evaluation of innovative marketing models, and only an accurate prediction of future sales can lead to a reasonable marketing plan. Therefore, a big data-driven e-commerce sales forecasting method is proposed. First of all, for 1703 real e-commerce companies, a large number of relevant data that affect sales are selected, including sales records, product information, product evaluation, and other information. A knowledge graph was then used to preprocess the data samples to produce a sample set containing concepts, entities, and relationships. Next, the knowledge graph K-modes clustering model is established. By fixing the affiliation matrix and the clustering cluster matrix in turn, the minimum of the objective function is continuously solved to obtain the cluster centres. Finally, sales prediction is achieved based on the clustering results. The experimental results show that the proposed clustering model is able to obtain better performance in terms of cluster purity, NMI, and *F*-value. The proposed clustering model has high sales prediction accuracy and has certain reference value for e-commerce enterprises of different scales to formulate innovative marketing models.

1. Introduction

Of all the innovative modes of business management, the innovation of the marketing model is very important. Finding the right marketing model is fundamental to a company's survival. Most scholars see the marketing model as a synergistic system consisting of several elements. The ultimate goal of marketing is to gain competitive advantage and achieve profitability. The marketing system therefore has to specify a holistic solution to achieve this goal [1–4]. The information technology revolution has provided space and impetus for marketing model innovation. In the current era of network economy, network and mobile communication technologies are changing the way of production and life of human beings in an all-round way. Therefore, marketing model innovation must also be coupled with information technology to play an important role. The Internet economy is developing rapidly and big data-driven innovation models have become a popular research direction at present [5–8]. Many companies are now using big data technology to achieve knowledge mining. At the same time, the full sharing of off-site resources is achieved through the network, shortening the distance between time and space. Real-time knowledge mining driven by big data can maintain the continuity and uninterrupted nature of innovation.

The China Internet Information Network Centre (CIINC) has published the 41st Statistical Report on the Development of the Internet in China [9–11]. This report showed that by December 2020, the number of online consumers in China would reach 772 million, with 40.74 million new people. Compared to 2019, the number of online consumers has increased by 2.6%. The Internet penetration rate has reached 55.8%, as shown in Figure 1. According to the China Electronic Commerce Research Centre, e-tailing transactions reached 3.1 trillion yuan in 2020, up by 34.8% year on year compared to 2.3 trillion yuan in 2019, as shown in Figure 2.

In recent years, the number of small e-commerce companies engaged in marketing via the Internet is increasing year on year. E-commerce platforms such as Alibaba, Amazon, and eBay are home to a huge number of small e-commerce companies. These small e-commerce companies, like other traditional businesses, are faced with the problem of marketing model innovation. Among all the innovations, the innovation of the marketing model is fundamental to the business. There are three broad categories of marketing models [12, 13]: value creation models, ecosystem models, and profitability models. The three types of theories discuss the connotation of the marketing model from different perspectives. A marketing model is an ecosystem consisting of many elements. This study considers the core element of the marketing model to be the profitability model. Profitability is a necessary condition for the existence of a market player. A business cannot sustain its basic survival without profit. For the many small e-commerce companies in e-commerce platforms, profit model innovation is their main concern.

The inadequate information management systems of small and medium-sized e-commerce businesses make it more difficult for them to innovate their profit models. In addition, small and medium-sized e-commerce enterprises face huge risks in the process of marketing reform because they cannot afford the larger capital costs. As a result, small and medium-sized e-commerce enterprises can only rely on forecast data of future sales to monitor their own business risks in real time [14, 15]. Sales forecasting can greatly improve the flow of capital for small and medium-sized e-commerce enterprises, allowing them to use these funds to develop broader sales channels or to cope with turbulent changes in the market environment. The accuracy of sales forecasting is therefore relevant to all aspects of an e-commerce business and will have a direct impact on its profit and loss, its next steps in financing, and its survival.

There are many models on sales forecasting, such as linear regression models, time series exponential smoothing forecasting models, logistic regression models, convolutional neural networks, and clustering mining [16–18]. In the preparation stage of each of these forecasting models, a sufficient amount of historical data has to be collected in order to predict the results with high accuracy. Therefore, these prediction models are all big data-driven models. Clustering analysis is an important application technique in data mining and is widely used in practical problems. Due to the development of information technology, consumer business data have become very large. With the continuous accumulation of data and the emergence of new business behaviours, it becomes difficult to classify data based on a priori experience. Therefore, Erenko et al. [19] proposed the application of cluster analysis to business marketing problems. Kingsland et al. [20] proposed the use of cluster analysis for data mining of consumer behaviour to give valuable guidance for future business decisions. Elizabeth et al. [21] proposed a cluster analysis-based model for e-commerce platform sales risk prediction model. From the perspective of microfinance companies, the types of e-commerce financial risks and the causes of their formation are analysed.

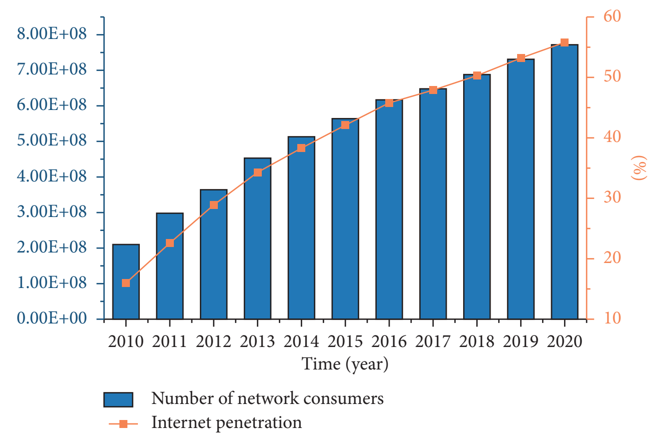


Figure 1: Size of Internet users in China and Internet penetration rate.

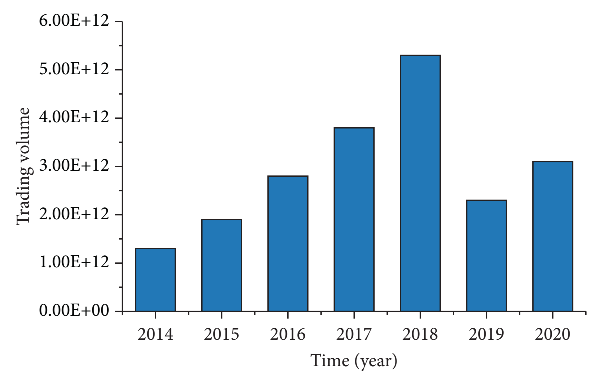


Figure 2: China's online retail market trading size.

Overall, clustering analysis can uncover new patterns in large amounts of data, which can be quite a powerful tool for data processing problems in the marketing field. By applying clustering algorithms, new patterns can be obtained that are not influenced by previous experience, thus allowing a more comprehensive exploitation of the information contained in the data. One of the most widespread applications of clustering analysis is the classical K-means algorithm [22]. However, the K-means algorithm is only applicable to datasets with ordered categorical attributes. To solve the clustering problem for unordered categorical attributes, Huang et al. [23] proposed the K-modes algorithm based on the classical K-means algorithm. Under the framework of the original K-modes algorithm, Saha et al. [24] proposed the genetic fuzzy K-modes algorithm, which improved the accuracy of data mining to a certain extent.

The aim of this study is to use the K-modes algorithm to mine the historical data of small and medium-sized e-commerce businesses to achieve an accurate forecast of future sales, thus providing data support for e-commerce businesses of different sizes when developing innovative marketing models. In the sales forecast of goods, the demand for products in the e-commerce industry is unstable as people's hobbies, consumption habits, and other factors are changing all the time. This phenomenon leads to no clear pattern in sales volume trends. In addition, sales are influenced by various external social factors such as climate and fashion trends. These changes can often affect the predicted outcome of a product, so it is essential to extract the key attributes of the data before clustering analysis.

The main innovations and contributions of this paper include the following:

(1) The K-modes algorithm in cluster analysis is introduced into the field of e-commerce marketing forecasting research, thus providing data support for e-commerce enterprises of different sizes in developing innovative marketing models.

(2) To further improve the clustering accuracy, the knowledge graph technique [25, 26] was used to extract the key attributes of the historical sales data before the clustering analysis of the historical sales data using the K-modes algorithm. After the knowledge graph analysis, K-modes algorithm was able to obtain a higher clustering accuracy.

The rest of the paper is organized as follows. In Section 2, the knowledge graph is studied in detail, while Section 3 provides the proposed knowledge graph K-modes clustering algorithm. In Section 4, the KGK-modes based e-commerce sales volume prediction model is studied in detail, while Section 5 provides experimental results and analysis. Finally, the paper is concluded in Section 6.

2. Knowledge Graph

For the marketing model of the e-commerce industry, accurate clustering of relevant sales data samples is more difficult due to the variability of the merchandise. This is because people's preferences, consumption habits, and other factors change all the time when forecasting sales of products. This phenomenon leads to no clear pattern in sales volume trends. Sales are also influenced by external social factors such as weather, trends, and so on. These changes ultimately lead to highly unstable forecasting results for commodities. In addition, if the key attributes of the sample are not extracted correctly, this may lead to incorrect clustering results. The above analysis shows that if the core attributes contained in the data sample can be identified, the scope of the clustering can be further reduced, thus effectively improving the accuracy of the clustering. Therefore, before using the K-modes algorithm to cluster the historical sales data, this paper uses the knowledge graph technique to extract the key attributes of the historical sales data.

The knowledge graph uses quadruples to represent knowledge [27], mainly containing concept, entity, relation, and attribute. The structure of the knowledge graph is shown in Figure 3.

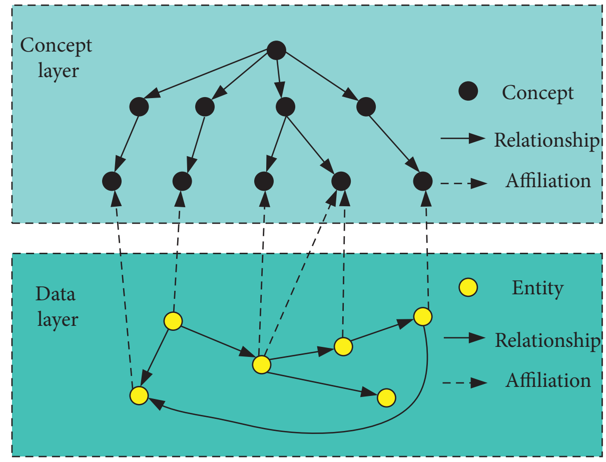


Figure 3: Structure of knowledge graph.

Suppose the set of all knowledge elements in knowledge domain d is KE_{d}:

$$KE_{\mathrm{d}} = \{ke_1, ke_2, \cdots, ke_i, \cdots, ke_n\}, \tag{1}$$

where ke_i represents the i-th knowledge element. Each knowledge element contains the concept knowledge c_i, entity knowledge e_i, relation knowledge r_i, and attribute knowledge a_i.

The set of concepts, entities, and relationships within knowledge domain d is denoted as C_d, E_d, and R_d, respectively.

$$\begin{aligned} C_d &= \{c_1, c_2, \cdots, c_k, \cdots, c_{nc}\}, \\ E_d &= \{e_1, e_2, \cdots, e_k, \cdots, e_{ne}\}, \\ R_d &= \{r_1, r_2, \cdots, r_k, \cdots, r_{nr}\}, \end{aligned} \tag{2}$$

where *nc*, *ne*, and *nr* represent the total number of concepts, entities, and relations, respectively.

The set of attributes corresponding to each concept A_{c_i} is A_{c_i}:

$$A_{c_i} = \left\{a_1, a_2, \cdots, a_j, \cdots a_{na}\right\}, \tag{3}$$

where *na* indicates the total number of attributes.

First, the complex commodity data are classified into knowledge sets. Next, knowledge unit parsing is performed. Finally, the knowledge elements and graphs contained in the knowledge units are extracted [28]. The knowledge graph is obtained by layer-by-layer analysis, where the scale structure of knowledge is shown in Figure 4.

3. The Proposed Knowledge Graph K-Modes Clustering Algorithm

3.1. Principle of the K-Modes Algorithm. The K-modes algorithm is a divisional clustering algorithm used to solve the problem of clustering categorical attributes [29–31]. The basic idea of the K-modes algorithm is the same as the classical K-means algorithm, but it introduces a different distance metric and a centroid selection method. Suppose the set of sample points for a classification attribute is

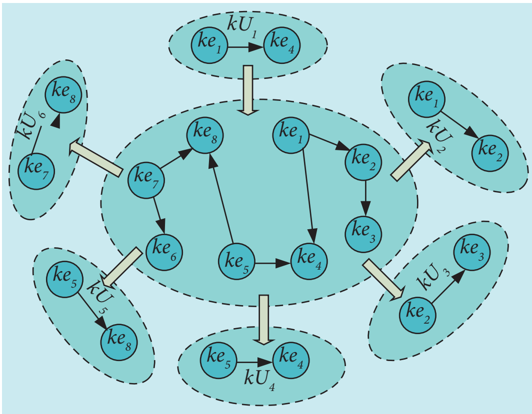


Figure 4: Knowledge structure composition.

$X = \{x_1, x_2, x_3, \ldots, x_n\}$. Each sample point contains m attributes. The samples are divided into k clusters z; then, the minimisation objective function of K-modes clustering is $P(\mathbf{W}, \mathbf{Z})$:

$$\begin{aligned} P(\mathbf{W}, \mathbf{Z}) &= \sum_{l=1}^{k} \sum_{i=1}^{n} \omega_{i,l} D(x_i, z_l), \\ \sum_{l=1}^{k} \omega_{i,l} &= 1, 1 \le i \le n, \end{aligned} \tag{4}$$

where $\mathbf{W}$ is the binary subordination matrix of $n \times k$ and $\mathbf{Z} = \{z_1, z_2, \ldots, z_k\}$ is a matrix containing k centroid [32].

$D(x_i, z_l)$ is the distance from a sample point to a centroid. The distance is calculated in the K-modes algorithm.

$$\begin{aligned} D(x, y) &= \sum_{j=1}^{m} d(x_j, y_j), \\ d(x_j, y_j) &= \begin{cases} 1, \; if \; x_j = y_j, \\ 0, \; f \; x_j \neq y_j. \end{cases} \end{aligned} \tag{5}$$

The objective function problem P is transformed into two subproblems. Let $\widehat{\mathbf{Z}}$ and $\widehat{\mathbf{W}}$ be the current optimal solutions. After each solution, we need to update ($\mathbf{W}$ and $\mathbf{Z}$) and save the result of this update to the database in order to continue solving for the minimum of $P(\mathbf{W}, \mathbf{Z})$.

$$\omega_{i,l} = \begin{cases} 1, if \; D(x_i, z_l) \le D(x_i, z_t) \; for \; 1 \le t \le k, \\ 0, \qquad\qquad\qquad\qquad for \; t \neq l. \end{cases} \tag{6}$$

When $\mathbf{W} = \widehat{\mathbf{W}}$, set $z_{l,j}$ to the attribute value of the j-th component of the cluster z_l.

$$\begin{aligned} \sum_{i=1}^{n} D(x_i, z_l) &= \sum_{i=1}^{n} \sum_{j=1}^{m} d(x_{i,j}, z_{l,j}) \\ &= \sum_{j=1}^{m} \left(\sum_{i=1}^{n} d(x_{i,j}, z_{l,j}) \right) = \sum_{j=1}^{m} n \left(1 - \frac{n_{z_l,j}}{n} \right), \end{aligned} \tag{7}$$

where $\mathbf{Z}_1$ is the attribute value of the cluster $\mathbf{Z}_1$. Solve for the distance values in turn until the minimum value of $\mathbf{Z}_1$ is found, thus obtaining the centroid and the class of each centroid.

The original K-modes algorithm takes a simple matching similarity measure based on the Hamming distance [33]. When the values of an attribute of two sample points are equal, their similarity on that attribute is 1; otherwise, the similarity is 0. This distance metric method may randomly assign less similar objects when assigning sample points, resulting in weaker intra-cluster similarity [34]. Thus, the original K-modes algorithm has limitations when mining sales-related data of small and medium-sized e-commerce companies. This is because there is no obvious pattern in the trend of sales volume changes. In addition, the sales volume of a product is influenced by various external social factors, such as climate, fashion trends, and so on. This problem causes the original K-modes algorithm to rely heavily on a priori experience. Therefore, in order to further improve the clustering accuracy, this paper uses the knowledge graph technique to extract the key attributes of historical sales data before using the K-modes algorithm for clustering analysis.

3.2. The Proposed Knowledge Graph K-Modes (KGK-Modes) Algorithm. Firstly, the proposed KGK-modes algorithm will analyse the data related to e-commerce sales and generate a new sample set containing the knowledge graph quaternions. Then, K-modes clustering is performed on the new sample set. Firstly, k cluster centres are determined, and k clusters are randomly selected from the sample as cluster centres, thus forming the initial cluster set $\mathbf{Z}_1$. A suitable $\mathbf{W}_1$ is found so that $P(\mathbf{W}_1, \mathbf{Z}_1)$ is minimised. During the updating process, stop updating if $P(\mathbf{W}_t, \mathbf{Z}_{t+1}) = P(\mathbf{W}_{t+1}, \mathbf{Z}_{t+1})$ is satisfied; otherwise, continue updating $\mathbf{W}_{t+1}$. The steps of the KGK-modes algorithm implementation are shown in Figure 5.

4. Big Data-Driven KGK-Modes-Based E-Commerce Sales Volume Prediction Model

4.1. Source of Data and Preprocessing. The data source for this study is the daily order volume, sales, number of customers, and number of reviews for 2,700 e-commerce companies on the Taobao website (https://www.taobao.com/). The time frame is from August 3, 2019, to April 30, 2020. The dataset was filtered using desensitisation rules as the data needed to be collected in a way that protected the privacy of the customer. The data deviate from the real e-commerce business data. However, this deviation did not affect the exploration and research of this solution. First, the three tables required (order, review, and product information) were merged into one table based on the e-commerce number in the SQL database. Then, the data were cleaned. The missing values in the fields were replaced by the median and mean values of the column. E-commerce numbers with only one or two items are removed directly. After a hierarchical process, the final selection of 1703 e-commerce companies with at least seven or more types of products was made.

Correction of missing data, duplicate data, and incorrect data contained in datasets related to e-commerce sales volume was done according to statistical methods. Records

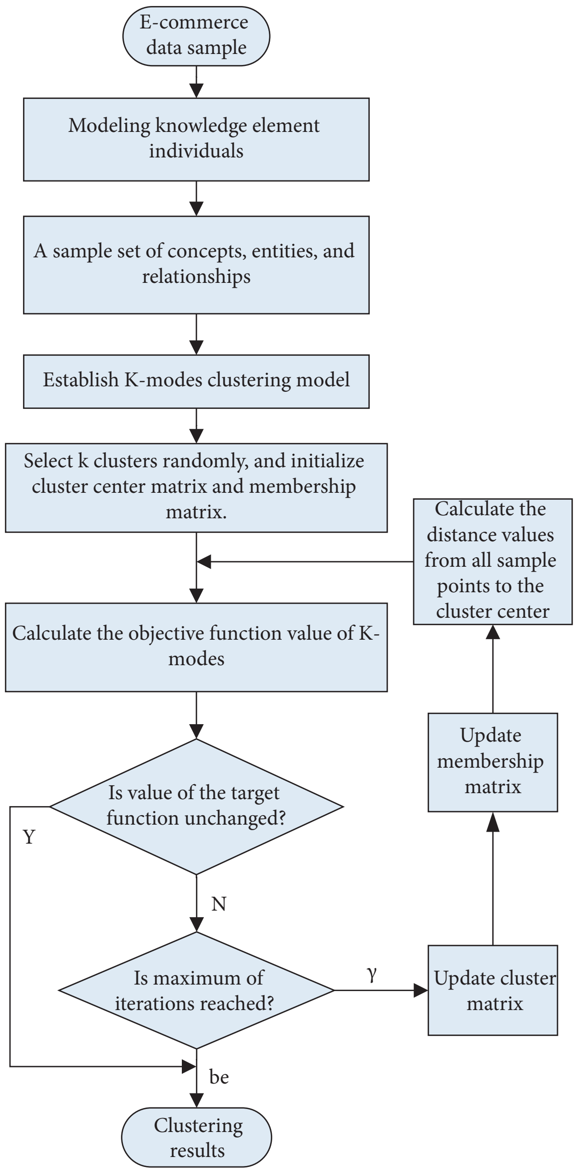


Figure 5: Clustering flow of KGK-modes.

with the same attribute value in the database are called duplicate records. Two records or two variables that are identical are combined into one record. Data merged from multiple data tables may have semantic conflicts. Inconsistent data can be transformed into consistent data by analysing the links between the data.

4.2. Sales Forecasting Using KGK-Modes Clustering. In this paper, the proposed KGK-modes algorithm is used to cluster the data of 1703 e-commerce companies, so as to complete the sales forecast. Eight variables are first set in the data of these e-commerce companies: sale_amt (sales), offer_amt (offer amount), offer_cnt (number of offers), rtn_cnt (number of returned orders), rtn_amt (amount of returned orders), ord_cnt (number of orders), bad_num (number of negative comment), and good_num (number of positive comment). Based on these 8 variables, the total data of 1703 e-commerce companies were calculated. The purpose of clustering with all variables is that the same type of e-commerce can be combined together, so as to forecast the sales of each type of e-commerce. The total daily data of the e-commerce companies are shown in Table 1.

Before clustering, we first performed a data check, i.e., descriptive statistics of the data [35], such as mean and standard deviation. It is found that the mean and variance are very small, so there is no need to standardize the data.

After analysing 1703 e-commerce companies by KGK-modes clustering algorithm, these e-commerce companies can be roughly divided into three categories. The first category is the e-commerce company with the largest variety of goods and a favorable rate of over 99%. In the first category, each company has more than 30 kinds of commodities. The number of the first type of e-commerce companies is 39, accounting for 2.3% of the total number. The second category includes e-commerce companies with 10~30 commodity types and 93%~98% favorable rate. The number of the second type e-commerce companies is 1,129, accounting for 66.3% of the total number. The third category includes e-commerce companies with less than 10 kinds of goods and a favorable rate of less than 93%. The number of the third type e-commerce companies is 535, accounting for 31.4% of the total number.

5. Experimental Results and Analysis

In order to verify the performance of KGK-modes in e-commerce sales forecasting, standard dataset tests and real case tests were conducted. The experimental hardware environment was a desktop computer with 64-bit Windows 10 operating system, Intel 7 CPU, 8G RAM, and GTX3060 graphics card, and the software used for the experiments was MATLAB R2018b. Firstly, the effect of the knowledge graph on the clustering of KGK-modes was verified on the commonly used machine learning dataset. Secondly, the clustering performance was compared between the commonly used clustering algorithms and the proposed KGK-modes algorithm, respectively. Finally, the effectiveness of the proposed KGK-modes algorithm was analysed on a dataset of historical sales of 1703 e-commerce companies. The main clustering evaluation metrics [36] were purity (P), standard mutual information (NMI), and F-value (F). Commonly used machine learning data come from the published UCI dataset and Sogo laboratory news dataset, which are shown in Tables 2 and 3, respectively.

5.1. The Influence of the Knowledge Graph on K-Modes

5.1.1. Clustering Performance on the UCI Dataset. To verify the effect of the knowledge graph on K-modes, the UCI dataset was tested using the K-modes and KGK-modes algorithms, respectively, and the results are shown in Table 4.

It can be seen that the KGK-modes algorithm shows better performance for all four different datasets. The cross-sectional comparison shows that the K-modes algorithm has the highest clustering purity of 0.8061 on the Seeds dataset

Table 1: Total daily data for e-commerce.

ale_amt	offer_amt	offer_cnt	rtn_cnt	rtn_amt	ord_cnt	bad_num	good_num
5324.47	27.55	23	2	187.25	67	0	0
6000.27	132.37	34	3	112.37	84	0	1
2957.59	48.13	17	1	9.2	52	0	17
4247.57	64.74	26	1	23.91	72	6	38
2888.33	77.01	18	0	0	60	6	122
6156.68	131.39	27	0	0	84	9	214
3286.77	84.31	31	3	132.07	65	17	210
3356.83	41.99	20	5	238.76	60	28	277
4227.12	129.61	34	2	30.53	80	23	270
6046.86	1201.02	32	2	119.19	73	3	315
2911.44	128.57	19	1	10.3	55	9	272
4245.03	57.82	19	1	0	62	31	232
3944.52	141.86	22	3	127.98	64	30	257
3706.83	127.55	29	3	86.82	71	17	516
4145.74	135.45	33	3	122.13	68	12	305
4138.08	606.01	31		39.73	65	11	472
4747.22	99.07	24	2	672.12	8	5	547
3587.7	83.18	17	1	146.78	61	10	157
6683.14	90.29	20	5	127.32	61	20	755
4328.59	1357.18	44	1	121.03	89	12	462
4431.47	255.46	56	3	268.55	85	10	556
3850.74	894.16	32	4	192.04	67	9	464
3887.83	203.74	25	1	61.07	57	7	638
3446.55	184.1	33	2	142.37	61	11	668
3080.63	107.5	20	2	41.94	51	6	862
3990.54	82.08	19	1	73.21	60	10	317
2918.88	94.95	21	0	0	55	17	731
2679.41	73.51	16	0	0	47	13	583
2957.46	48.71	16	2	60.33	61	13	548
4354.07	128.83	35	2	63.28	76	10	666

Table 2: UCI simulation set.

Datasets	Sample size	Number of attributes	Number of clustering categories
Flowers	385	34	10
Wine	181	16	5
Iris	1024	21	8
Seeds	676	30	6

Table 3: News text set.

Datasets	Sample size	Number of document categories	Category
1	100	5	Finance, military
2	500	7	Education, sports
3	800	7	Tourism, culture
4	1000	7	IT

Table 4: Clustering performance of two algorithms (UCI set).

Datasets	Algorithms	P	NMI	F
Flowers	K-modes	0.7662	0.6311	0.7283
	KGK-modes	0.8824	0.7017	0.8592
Wine	K-modes	0.7713	0.6621	0.7356
	KGK-modes	0.8632	0.7112	0.8572
Iris	K-modes	0.7894	0.7177	0.7812
	KGK-modes	0.8927	0.7809	0.8798
Seeds	K-modes	0.8061	0.7249	0.7651
	KGK-modes	0.9046	0.7964	0.8955

and the lowest clustering purity of 0.7662 on the Flowers dataset. This indicates that both algorithms obtained optimal performance on the Seeds dataset and the worst clustering on the Flowers dataset. Comparing the NMI and F performance, the K-modes algorithm showed better clustering performance after the knowledge graph analysis. This is because after the knowledge graph analysis, the data samples are accurately delineated in terms of concepts, entities, and relationships. The delineation of the data samples helps to determine the sample categories to a certain extent, thus reducing the difficulty of subsequent K-modes clustering.

KGK-modes showed good performance in terms of P, NMI, and F-values in clustering the four-class dataset of UCI. To analyse the stability of clustering purity, the RMSE performance of clustering purity [37] was tested. A random sample of 1000 from the UCI dataset was tested for clustering, and the results are shown in Figure 6.

The RMSE of both the K-modes and KGK-modes algorithms gradually decreased as the clustering time increased. In comparison, it was found that the RMSE of the clustering purity obtained by KGK-modes decreased more rapidly. Eventually, the RMSE of KGK-modes converges to 0.5, while that of K-modes converges to 0.75.

In addition, the clustering times of the K-modes algorithm and the KGK-modes algorithm were further compared on the UCI dataset, and the statistical results are shown in Table 5.

It can be seen that the clustering time of the K-modes algorithm and KGK-modes algorithm on the UCI dataset is directly related to the sample size. The Iris dataset with the largest sample size required the longest clustering time, while the Wine dataset with the smallest sample size required the shortest clustering time. The comparison revealed that the knowledge graph analysis consumed some time, and therefore the clustering time for KGK-modes was longer than that for K-modes, but the difference between the two was smaller.

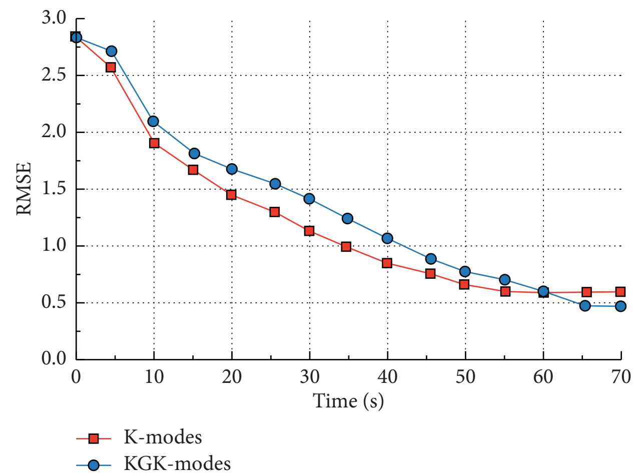


Figure 6: RMSE of clustering purity.

Table 5: Clustering time of two algorithms (UCI set).

Sample set	Algorithms	Running time (s)
Flowers	K-modes	25.13
	KGK-modes	29.25
Wine	K-modes	11.29
	KGK-modes	14.76
Iris	K-modes	61.52
	KGK-modes	67.19
Seeds	K-modes	42.67
	KGK-modes	48.73

5.1.2. Clustering Performance on News Datasets. To further validate the effect of the knowledge graph on the K-modes algorithm, the performance of the news dataset was tested using the K-modes and KGK-modes algorithms, respectively, and the results are shown in Table 6.

Table 6: Clustering performance of two algorithms (news sets).

Datasets	Algorithms	P	NMI	F
1	K-modes	0.8646	0.6473	0.8064
	KGK-modes	0.9275	0.7014	0.9012
2	K-modes	0.8796	0.6725	0.8355
	KGK-modes	0.9418	0.7368	0.9137
3	K-modes	0.8821	0.7073	0.8382
	KGK-modes	0.9437	0.7716	0.9176
4	K-modes	0.8946	0.7312	0.8468
	KGK-modes	0.9526	0.8192	0.9235

It can be seen that as with the UCI dataset, after the knowledge graph analysis, the K-modes algorithm showed better NMI and F performance. The analysis of the stability of the clustering purity on the news dataset is given in the following, and the results are shown in Figure 7.

It can be seen that the RMSE of the K-modes and KGK-modes algorithms decreases significantly as the number of clustering iterations increases. The RMSE of the clustering purity obtained by KGK-modes decreases faster than that of the K-modes algorithm and eventually converges to about 0.4, while that of K-modes converges to about 0.5.

Next, a comparison of clustering time performance was performed. The clustering times of the K-modes and KGK-modes algorithms on the four news datasets are shown in Table 7.

For the news set with the same sample size, the KGK-modes clustering time was slightly longer than the K-modes clustering time. A comprehensive analysis of the above results shows that the KGK-modes algorithm shows higher performance on the news dataset than the UCI dataset, mainly because the news dataset has significantly fewer feature dimensions than the UCI dataset, and therefore the data clustering effect is more significant.

5.2. Sales Forecast. To verify the effectiveness of the KGK-modes algorithm in e-commerce sales prediction, the sales-related datasets (containing 8 variables) of 1703 e-commerce businesses were tested using K-means, K-medoids, CNN [38], and KGK-modes algorithms, respectively. 90% of the dataset was used as the training set, and the remaining 10% was used as the test set. Eight variables are used as inputs, and the above four prediction models are, respectively, used for fitting. The loss function was mean squared deviation, the learning rate was 0.01, the maximum number of iterations was 1000, and the resampling rate was 50%.

Taking the first type of e-commerce as an example, the KGK-modes prediction model reached optimality when the optimal number of iterations was 367, as shown in Figure 8, by continuously performing iterative calculations. This indicates that the error has been minimised when the number of iterations reaches 367. The predicted values of the model obtained by the KGK-modes algorithm are more accurate at this point. The results in Figure 9 show that in the first

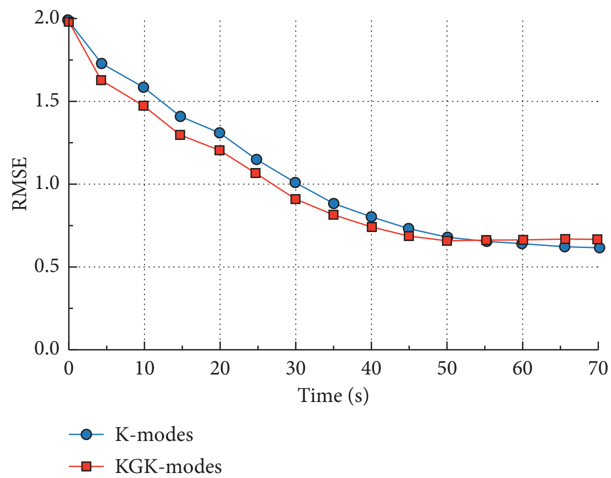


Figure 7: RMSE of clustering purity (news sets).

Table 7: Clustering time of two algorithms (news set).

Sample set	Algorithms	Running time (s)
1	K-modes	9.53
	KGK-modes	11.05
2	K-modes	31.75
	KGK-modes	33.94
3	K-modes	48.67
	KGK-modes	51.28
4	K-modes	60.29
	KGK-modes	65.13

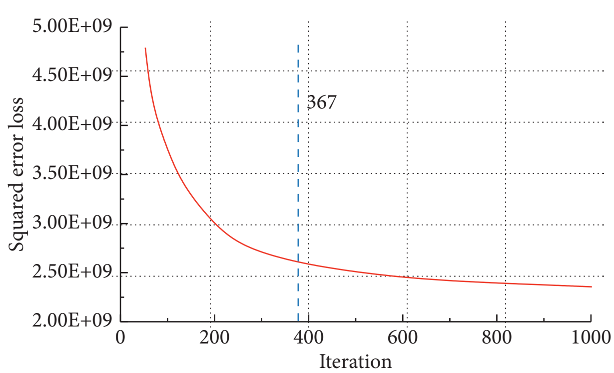


Figure 8: Iteration error curve.

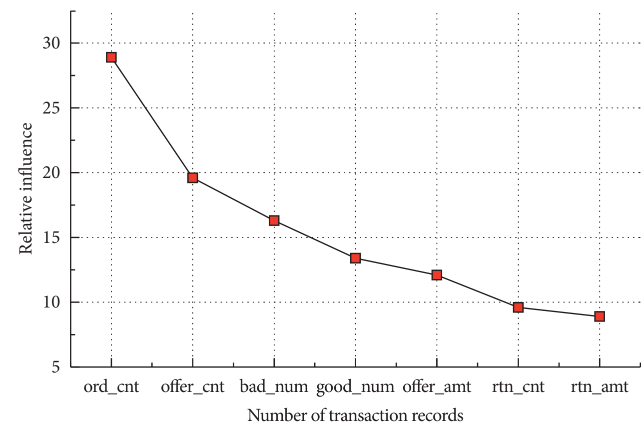


Figure 9: Extent of influence of 8 variables on sales.

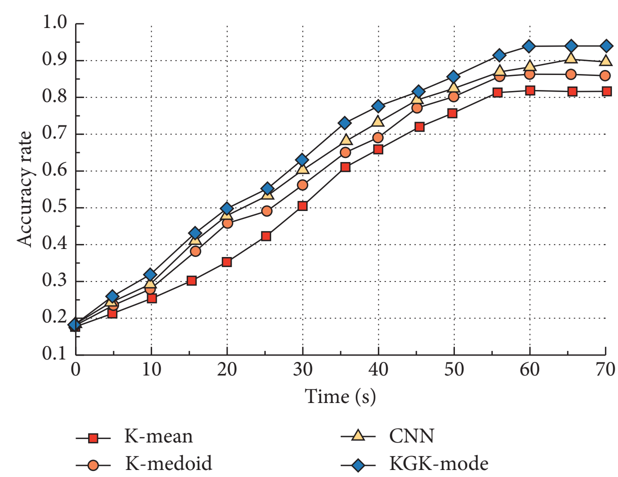


Figure 10: Comparison of the accuracy of the four prediction models.

category of e-commerce, the magnitude of the impact on sales is ord_cnt, offer_cnt, bad_num, offer_amt, rtn_cnt, and rtn_amt in that order.

The fitted model was applied to the test set, and the error was compared between the true and predicted values of the test set, and the results are shown in Figure 10.

It can be seen that the KGK-modes algorithm has the highest sales prediction accuracy of about 0.94. The CNN algorithm has the second highest sales prediction accuracy after the KGK-modes algorithm at about 0.91. The K-means algorithm has the worst sales prediction accuracy at about 0.81. In terms of running time, the K-means and K-medoids algorithms are the most efficient, reaching stability at around 60 s. On the other hand, the KGK-modes and CNN algorithms took 65 s and 70 s, respectively, to reach stability, and on balance, the KGK-modes algorithm had the highest sales prediction accuracy and the K-means algorithm had the best runtime.

5.3. Adjustment Strategies for Innovative Marketing Models. For the first e-commerce, the main influencing factors for sales are ord_cnt, offer_cnt and bad_num. The higher the order volume, the higher the sales amount. For the second category of e-commerce, the main influencing factors of sales are ord_cnt, rtn_amt, bad_num and offer_amt. For the third category of e-commerce, the main influencing factors of sales are ord_cnt, bad_num and good_num. Therefore, for the third category of e-commerce, which has the smallest range of products, they should focus more on the marketing model innovation in positive comment. The number of bad reviews has too great an impact on third category e-commerce, and some reasonable marketing strategies should be prepared in advance to reduce bad reviews. This is because the first thing that consumers browse for when they enter a third category e-commerce is the reviews of the goods. It is recommended to offer some economical and affordable goods from the customer's point of view and to improve the logistics management.

After a comprehensive analysis of the characteristics of these three types of e-commerce, it is found that ord_cnt is the most important influencing factor on e-commerce sales. Therefore, we can directly use ord_cnt to simply forecast the change of sales. The role of marketing activities such as appropriate price reductions is more obvious for the second category of e-commerce. For the third category of e-commerce, the impact of bad reviews is much greater than that of good reviews. This is because in the process of online shopping, in order to buy what they want, people will refer to other people's reviews of this product. If this shop has more bad reviews, it will affect the consumer's purchase intention. For e-commerce companies of different scales, the most important sales influencing factors are different. In the future, according to the priority of the factors affecting sales, different sizes of e-commerce companies need to adjust the innovative marketing model to ensure maximum profit.

6. Conclusion

In this paper, a big data-driven KGK-modes clustering algorithm is proposed and applied to e-commerce sales volume prediction, so as to reasonably adjust the required innovative marketing mode. knowledge graph was used to preprocess the data samples to generate a sample set containing concepts, entities, and relationships so that the key attributes of the historical sales data could be extracted. After the knowledge graph analysis, K-modes algorithm was able to achieve better clustering performance. The KGK-modes clustering model was used to achieve e-commerce sales prediction. The experimental results show that the KGK-modes clustering model has high sales prediction accuracy. The proposed method has certain reference value for e-commerce enterprises of different scales to formulate innovative marketing models. Subsequent research will be carried out in terms of the efficiency of clustering, and invoking the Spark parallel platform should be considered to reduce the running time of KGK-modes clustering. On the other hand, an attempt is made to optimize the objective function of KGK-modes clustering in order to improve the efficiency of the clustering algorithm.

Data Availability

The experimental data 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 study was supported by the Humanities and Social Sciences Projects of Universities in Jiangxi Province (Research on the Compounding Mechanism of Changes in College Students' Political Trust (issue no. SZZX21019); Practical Reflection and Strategy Optimization Research on Rural Cultural Construction from the Perspective of the Peasant (issue no. SH19104)).

References

[1] L. Y. Chen, "A knowledge marketing model: determinants of organisational resource-based capabilities on e-retail performance," *International Journal of Business Information Systems*, vol. 9, no. 1, pp. 89–107, 2012.

[2] L. V. Lorente-Bayona, E. Gras-Gil, and M. Moreno-Enguix, "Internet penetration and international travel and tourism expenditure: the role of foreign exchange control," *Tourism Economics*, vol. 2021, no. 2, p. 135481662110278, 2021.

[3] D. Shah and B. P. S. Murthi, "Marketing in a data-driven digital world: i," *Journal of Business Research*, vol. 125, pp. 772–779, 2021.

[4] H. Q. Vu, G. Li, and R. Law, "Discovering implicit activity preferences in travel itineraries by topic modeling," *Tourism Management*, vol. 75, no. 12, pp. 435–446, 2019.

[5] L. Liu, F. Zhang, H. Li, and C. Wen, "Research on the community group buying marketing model of fresh agricultural products in Jilin province from the perspective of internet marketing and retail," *IOP Conference Series: Earth and Environmental Science*, vol. 769, no. 2, p. 022054, 2021.

[6] D. Klimanov and O. Tretyak, "Linking business model research and marketing: new network-based approach to business model analysis," *The Journal of Business and Industrial Marketing*, vol. 34, no. 1, pp. 117–136, 2019.

[7] D. Xu, T. Chen, J. Pearce, Z. Mohammadi, and P. L. Pearce, "Reaching audiences through travel vlogs: the perspective of involvement," *Tourism Management*, vol. 86, no. 2, p. 104326, 2021.

[8] H. Gebauer, A. Arzt, M. Kohtamäki et al., "How to convert digital offerings into revenue enhancement – c," *Industrial Marketing Management*, vol. 91, pp. 429–441, 2020.

[9] R. Dew and A. Ansari, "Bayesian nonparametric customer base analysis with model-based visualizations," *Marketing Science*, vol. 37, no. 2, pp. 216–235, 2018.

[10] E. Kappe, A. Stadler Blank, and W. S. Desarbo, "A random coefficients mixture hidden Markov model for marketing research," *International Journal of Research in Marketing*, vol. 35, no. 3, pp. 415–431, 2018.

[11] J. Martins, R. Gonalves, and F. Branco, "A multisensory virtual experience model for thematic tourism: a Port wine tourism application proposal," *Journal of Destination Marketing & Management*, vol. 6, no. 2, pp. 91–106, 2017.

[12] J. J. Yang, H. W. Lo, C. S. Chao, C. C. Shen, and C. C. Yang, "Establishing a sustainable sports tourism evaluation framework with a hybrid multi-criteria decision-making model to explore potential sports tourism attractions in taiwan," *Sustainability*, vol. 12, no. 4, p. 1673, 2020.

[13] S. Osama, N. Mohsen, H. Walayat, and M. Luis, "Cloud computing model selection for e-commerce enterprises using a new 2-tuple fuzzy linguistic decision-making method," *Computers & Industrial Engineering*, vol. 132, pp. 47–58, 2019.

[14] H. Y. Lee and N. J. Wang, "Cloud-based enterprise resource planning with elastic model–view–controller architecture for Internet realization," *Computer Standards & Interfaces*, vol. 64, no. 5, pp. 11–23, 2019.

[15] H. Xie, R. T. B. Ma, and J. C. S. Lui, "Enhancing reputation via price discounts in E-commerce systems: a data-driven

approach," *ACM Transactions on Knowledge Discovery from Data*, vol. 12, no. 3, pp. 1–29, 2018.

[16] S. Mu, Y. Wang, F. Wang, and L. Ogiela, "Transformative computing for products sales forecast based on SCIM," *Applied Soft Computing*, vol. 109, no. 1, p. 107520, 2021.

[17] J. Zhao and C. Zhang, "Research on sales forecast based on prophet-SARIMA combination model," *Journal of Physics: Conference Series*, vol. 1616, no. 1, p. 012069, 2020.

[18] A. Kumar, A. Adlakha, and K. Mukherjee, "Modeling of product sales promotion and price discounting strategy using fuzzy logic in a retail organization," *Industrial Management & Data Systems*, vol. 116, no. 8, pp. 1418–1444, 2016.

[19] O. Erenko and O. Koppel, "Formal and informal macro-regional transport clusters as A primary step in the design and implementation of cluster-based strategies," *Transport and Telecommunication Journal*, vol. 16, no. 3, pp. 207–216, 2015.

[20] M. Kingsland, L. Wolfenden, J. Tindall et al., "Tackling risky alcohol consumption in sport: a cluster randomised controlled trial of an alcohol management intervention with community football clubs," *Journal of Epidemiology & Community Health*, vol. 69, no. 10, pp. 993–999, 2015.

[21] G. T. Elizabeth, W. Ian, and V. Bas, "Householders' mental models of domestic energy consumption: using a sort-and-cluster method to identify shared concepts of appliance similarity," *PLoS One*, vol. 11, no. 7, p. e0158949, 2016.

[22] Y. Fan, Y. Liu, H. Qi, F. Liu, and X. Ji, "Anti-interference technology of surface acoustic wave sensor based on K-means clustering algorithm," *IEEE Sensors Journal*, vol. 21, no. 7, pp. 8998–9007, 2021.

[23] Z. Huang and M. K. Ng, "A fuzzy k-modes algorithm for clustering categorical data," *IEEE Transactions on Fuzzy Systems*, vol. 7, no. 4, pp. 446–452, 1999.

[24] A. Saha and S. Das, "Categorical fuzzy k-modes clustering with automated feature weight learning," *Neurocomputing*, vol. 166, no. 10, pp. 422–435, 2015.

[25] Y. Jia, Y. Qi, H. Shang, R. Jiang, and A. Li, "A practical approach to constructing a knowledge graph for cybersecurity," *Engineering*, vol. 4, no. 1, pp. 53–60, 2018.

[26] L. Shen, R. He, and S. Huang, "Entity alignment with adaptive margin learning knowledge graph embedding," *Data & Knowledge Engineering*, vol. 139, pp. 101987–101996, 2022.

[27] W. Zheng, L. Yan, C. Gou, Z. C. Zhang, and F. Y. Wang, "Pay attention to doctor-patient dialogues: multi-modal Knowledge Graph attention image-text embedding for COVID-19 diagnosis," *Information Fusion*, vol. 75, no. 8, pp. 108–116, 2021.

[28] H. Wang, Y. Wang, J. Li, and T. Luo, "Degree aware based adversarial graph convolutional networks for entity alignment in heterogeneous knowledge graph," *Neurocomputing*, vol. 487, pp. 99–109, 2022.

[29] C. Liu, X. Wang, Y. Huang et al., "A moving shape-based robust fuzzy K-modes clustering algorithm for electricity profiles," *Electric Power Systems Research*, vol. 187, p. 106425, 2020.

[30] H. H. Nguyen, "Privacy-preserving mechanisms for k-modes clustering," *Computers & Security*, vol. 78, pp. 60–75, 2018.

[31] C. Liu, X. Wang, Y. Huang et al., "A moving shape-based robust fuzzy K-modes clustering algorithm for electricity profiles," *Electric Power Systems Research*, vol. 187, p. 106425, 2020.

[32] M. Faezirad, A. Pooya, Z. Naji-Azimi, and M. Amir Haeri, "Preventing food waste in subsidy-based university dining systems: an artificial neural network-aided model under uncertainty," *Waste Management & Research: The Journal for a Sustainable Circular Economy*, vol. 39, no. 8, pp. 1027–1038, 2021.

[33] R. J. Kuo, Y. R. Zheng, and T. P. Q. Nguyen, "Metaheuristic-based possibilistic fuzzy k-modes algorithms for categorical data clustering," *Information Sciences*, vol. 557, pp. 1–15, 2021.

[34] R. J. Kuo and T. P. Q. Nguyen, "Genetic intuitionistic weighted fuzzy k-modes algorithm for categorical data," *Neurocomputing*, vol. 330, pp. 116–126, 2019.

[35] A. Ichenihi, W. Li, Y. Gao, and Y. Rao, "Feature selection and clustering of damage for pseudo-ductile unidirectional carbon/glass hybrid composite using acoustic emission," *Applied Acoustics*, vol. 182, no. 2, p. 108184, 2021.

[36] J. C. Forte, G. Yeshmagambetova, M. Grinten, and B. Hiemstra, "Identifying and characterizing high-risk clusters in a heterogeneous ICU population with deep embedded clustering," *Scientific Reports*, vol. 11, no. 1, p. 12109, 2021.

[37] X. Ran, X. Zhou, M. Lei, W. Tepsan, and W. Deng, "A novel K-means clustering algorithm with a noise algorithm for capturing urban hotspots," *Applied Sciences*, vol. 11, no. 23, p. 11202, 2021.

[38] F. Sun, G. Xie, and S. Li, "An artificial-neural-network based prediction of heat transfer behaviors for in-tube supercritical CO2 flow," *Applied Soft Computing*, vol. 102, p. 107110, 2021.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9832807, 1 page
https://doi.org/10.1155/2023/9832807

Retraction

Retracted: The Practical Dilemma and Path Optimization of Big Data Integration into Ideological and Political Teaching in Private Colleges and Universities

Mobile Information Systems

Received 13 September 2023; Accepted 13 September 2023; Published 14 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] X. Li, "The Practical Dilemma and Path Optimization of Big Data Integration into Ideological and Political Teaching in Private Colleges and Universities," *Mobile Information Systems*, vol. 2022, Article ID 7417406, 10 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 7417406, 10 pages
https://doi.org/10.1155/2022/7417406

Research Article

The Practical Dilemma and Path Optimization of Big Data Integration into Ideological and Political Teaching in Private Colleges and Universities

Xihong Li

United Front Work Department of Party Committee, Yellow River University of Science and Technology, Zhengzhou 450063, China

Correspondence should be addressed to Xihong Li; 201103022@hhstu.edu.cn

Received 13 July 2022; Revised 28 July 2022; Accepted 8 August 2022; Published 28 August 2022

Academic Editor: R. Mo

The promotion and application of big data have promoted the reform of quality classes in Colleges and universities, enriched the teaching content, and changed the traditional education mode. However, at present, due to the difficulties in funding, understanding teachers, and system construction, the integration of data into teaching often presents an embarrassing situation of separation between form and content, the separation between theory and practice, and the distance between teaching results and teaching objectives. For this reason, many private colleges and universities should make full use of data technology to change the quality of education mode according to the new characteristics of college students in the data environment, use the network information carrier to strengthen the interaction and exchange with college students, do a good job in guiding the ideological dynamics, actively promote positive capacity, and improve the networking level of quality education. At the same time, in order to change the current situation of incomplete integration and break through the dilemma, private colleges and universities should change their ideas and strengthen their attention to the practical ideological and political teaching. In view of the above problems, this study analyzes the internal logic of the integration of private colleges and big data, combines the actual difficulties, and uses LSTM neural network to put forward reasonable optimization strategies and suggestions, aiming to expand and cultivate high-quality practical teaching teams of Ideological and political courses, so as to improve the system and establish and improve the relevant teaching system and mechanism.

1. Introduction

Universities are the main positions for talent training and output. At present, there are more than 3000 universities in China. The distribution of universities in various provinces is shown in Figure 1. However, with the increasing difficulty of higher education management, education and teaching problems are becoming increasingly prominent. To better implement educational work, universities must introduce and apply data technology, combine big data with education management, constantly improve the original management mode, create a new path of university education work, improve the quality and efficiency of university education work, and provide more modern system talents for the unit [1, 2].

In February 2017, China issued the reply document on Strengthening Quality classroom education for college students in the socialist environment of the new era. The document pointed out that it is essential to strengthen the framework of students' Internet Quality work carriers, strengthen the construction of students' interactive communities, theme education websites, professional academic websites, and "two micro ends", carry out relevant education work by using the expression methods that students like, and innovate education carriers based on students' interests and hobbies, Enrich education channels, make full use of Weibo, WeChat, etc., push information, and fundamentally change the previous education methods [3–5]. By integrating modern data means, we can innovate classroom education resources, continuously improve the effectiveness and

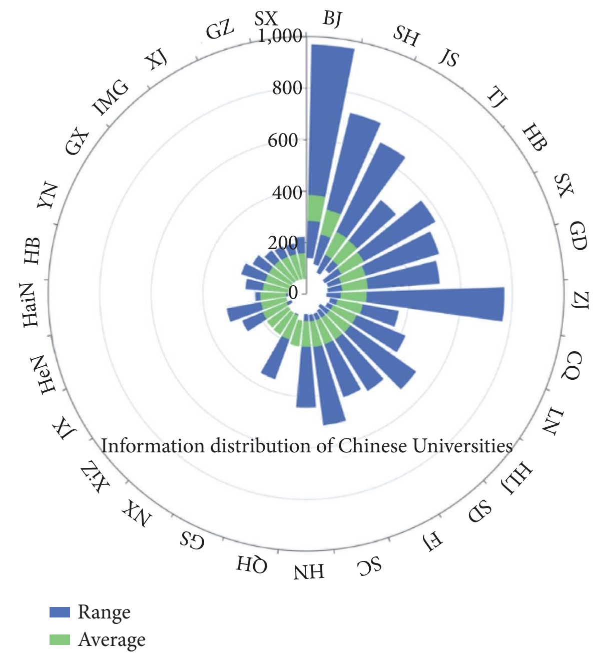


Figure 1: Information distribution of Chinese Universities.

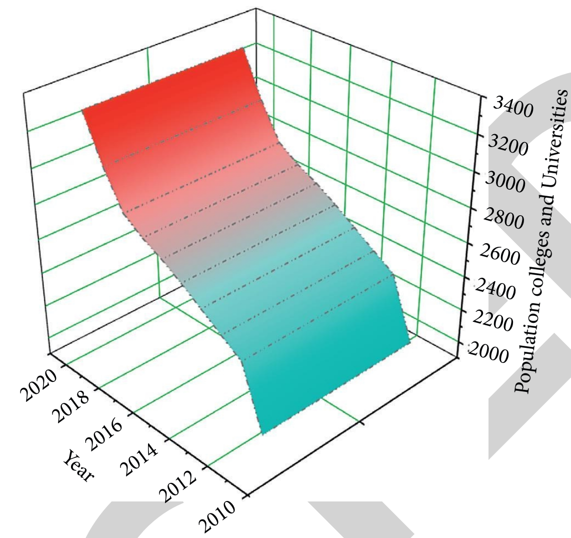


Figure 2: Statistical law of the number of College Students.

effectiveness of education, strengthen the teaching hardware of the school, change the teaching methods of school teachers, provide flexible and diverse guidance to students, and play the role of big data in modern school education [6–8]. However, in the background of data development, university education also faces many problems and deficiencies, such as weak infrastructure, ineffective sharing of education data information, and security management risks, which affect the effect of big data application in university basic course education [9, 10]. Therefore, university education managers should integrate big data into the internal logic of classroom education, analyze the existing difficulties and take targeted solutions [11, 12].

2. Big Data and the Internal Logic of Ideological and Political Teaching in Private Colleges

With the vigorous development of education in our country, the number of college students in China has also shown exponential growth. Figure 2 reflects the changes in college students in China since 2010. From the change of data, the number of college students has changed greatly since 2010, especially after 2014, it has almost multiplied every year. Currently, the integration of big data into university classrooms also has its internal logic, enriching educational resources, providing data level basis for education, and improving the technical level of Quality education. But yet, according to incomplete statistics, the quality of education of college students is not optimistic. The quality classes in most schools are still in the early stage of education, which seriously hinders the healthy development of education [13–16].

From the above Figure 2, we can find that since 2014, the number of students in Colleges and universities in China has increased year by year, showing a trend of class index expansion. It can be seen that the development of colleges and universities in China has changed greatly, the number of college students has expanded, and the development of the Internet in recent years has been more rapid. It can be predicted that in the next few years, The contradiction between the ideological and political education of college students in China and the expansion of big data will become more and more serious.

2.1. Big Data Provides Important Data Information for College Students' Education. Before 2010, affected by external environmental factors such as environment and human activities, it was difficult to obtain accurate data and statistical information in the quality classes of college students, resulting in varying degrees of errors in the educational conclusions [17]. However, today, the arrival of the digital age can solve these problems. First, with the help of digital technology, mobile terminals, apps, and other innovative teaching means and methods to quantify relevant educational information and data, so that teachers can re-examine and analyze the characteristics of quality education from many angles. Second, the innovation and development of digital technology provide convenience for teachers to conduct teaching research. Teachers can do a good job in data screening, processing, storage, and utilization in combination with actual educational needs, ensure the integrity of teaching information, improve the education level of relevant classes, provide important data basis for subsequent teaching, make daily education and teaching more accurate and improve the predictability of education. Teachers can analyze and study students' behavior according to digital technology, so as to formulate more scientific and accurate teaching programs.

To effectively mine all kinds of teaching data, we can reduce the workload of data collection by establishing relevant data mining models. Taking the long and short-term memory neural network data mining model as an example, this paper designs a teaching data mining model based on LSTM.

2.1.1. LSTM Theory. The long-term and short-term memory network model is widely used in time series problems. It is a variant of a recurrent neural network (RNN), which processes data through a gating mechanism. The LSTM neural unit consists of CEC, input gate, output gate, forgetting gate, and peephole. The relationship of each component is shown in Figure 3. The data transmission formula followed is as follows:

Initial input gate i_t:

$$i_t = \sigma\left(W_i x_t + U_i h_{t-1} + b_i\right). \tag{1}$$

Initial output gate o_t:

$$o_t = \sigma\left(W_o x_t + U_o h_{t-1} + b_o\right). \tag{2}$$

Initial forgetting gate f_t:

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right). \tag{3}$$

Input gate after the introduction of peephole i_t:

$$i_t = \sigma\left(W_i x_t + U_i h_{t-1} + V_i c_{t-1} + b_i\right). \tag{4}$$

Output gate after the introduction of peephole o_t:

$$o_t = \sigma\left(W_o x_t + U_o h_{t-1} + V_o c_t + b_o\right). \tag{5}$$

Forgetting door after introducing peephole f_t:

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + V_f c_{t-1} + b_f\right). \tag{6}$$

Of which:

$$\begin{aligned} c_t &= i_t \odot a_t + f_t \odot c_{t-1}, \\ a_t &= f\left(W_a x_t + U_a h_{t-1} + b_a\right), \\ h_t &= o_t \odot g\left(c_t\right). \end{aligned} \tag{7}$$

Under the joint action of CEC, peephole and three gating mechanism units, it makes:

$$eh_{t-1} = eh_t \odot \left(W f'\left(p(t-1)\right) = 1.\right. \tag{8}$$

2.1.2. LSTM Data Mining Model Architecture. Data mining pattern recognition can be divided into three functional modules: data preparation, LSTM-based learning model training, and data pattern recognition. The main functions of the modules are as follows:

(1) Data preparation: use the window sliding method to divide the original sequence to be tested into subsequences.
(2) Data mining based on LSTM learning model: This module generates LSTM network data mining model. The pattern recognition module can input data features into the model and output them as pattern labels of data.
(3) Data pattern recognition: input the sequence to be tested into the well-trained network model, and automatically recognize the data pattern. The workflow of the data pattern recognition model is shown in Figure 4.

2.2. It Has Promoted the Reform of Educational Ideas in Colleges. With the support of digital technology, the university curriculum education reform process can retain the original text, video, sound, pictures, etc., can form different information from different channels, and can quickly update, which can create a more suitable learning environment for the change of curriculum education ideas of relevant professional teachers [18, 19]. The application of big data has broken through the geographical and spatial constraints, and the speed of data transmission and update is faster. It has enriched the channels for teachers to obtain relevant educational resources and formed a rich educational information database. The use of relevant data technology can not only bring teaching benefits to educators, but also to the curriculum process. To meet the needs of the rapid development of big data, teachers continue to innovate educational ideas, effectively improve the single ideological path of previous education objects, and better stimulate students' awareness of autonomous learning. Teachers can make full use of education carriers rich in big data, such as video teaching or online education, to view, review and reflect on the education process, so as to enhance the intuitiveness of education.

2.3. Broaden the Horizon of Education. Big data not only provides massive data for education managers, but also provides personalized data analysis for each student. The application and expansion of big data have played an vital role in improving the level of educational technology in university courses. First, big data, as a symbolic product of the information age, it provides a new perspective for related college curriculum education, and also provides teachers with a variety of network data. Teachers can further broaden the educational space by mining the internal related data. Second, teachers use digital technology to analyze the new characteristics of current university curriculum education as a whole, flexible application of big data and methods to daily teaching in combination with the actual environment of the school, analyze the learning habits and behavior patterns of education subjects and objects, and promote the improvement and innovation of education methods; Third, using data technology, teachers can model and simulate future teaching practice activities, scientifically predict education, analyze potential problems and adverse factors, more accurately grasp internal laws, and make education more targeted and accurate.

3. Existing Dilemma

3.1. Problems in Teachers' Ideological Understanding of Related Education. In the current era, big data has not only developed rapidly in the network society, but has also begun

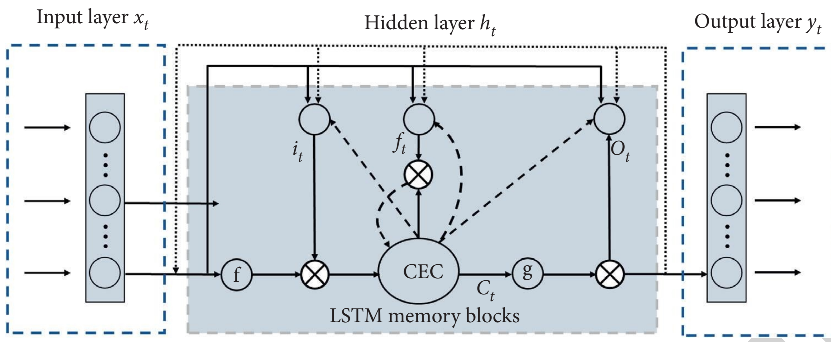


Figure 3: Structure diagram of LSTM basic neural network.

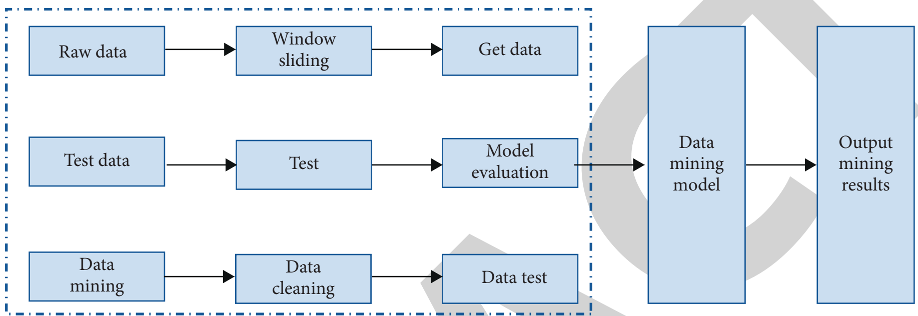


Figure 4: Workflow of data pattern recognition model.

to penetrate into all fields of the school, promoting social innovation and progress. But yet, based on the development status of big data integration into university curriculum education, there are still problems in all aspects. First, as an important force of young netizens, college students should not only be limited to the teaching of books and textbooks but also play the role of big data in teaching to improve the timeliness of education. Some course teachers have not given enough attention to innovation in the data field, and their positioning is unscientific, so they can not give full play to the important role of big data. Second, there is a lack of big data application concept in the education evaluation and assessment of relevant courses, and there are insufficient problems in the collection, sorting, screening, and utilization of big data information, which reduces the effect of education. Third, in view of the impact of school data on college classroom education in the data age, no targeted solutions have been taken. It is difficult to control reform and innovation, and it is difficult to achieve effective use of big data due to the lack of specific educational practice.

Moreover, in some private colleges, some top managers have a one-sided understanding of Quality teaching. Especially in the 21st century, education practitioners are not aware of the internal relationship between students' professional practice and quality teaching practice. Some elderly educators are divorced from the data age, and their awareness deviates greatly from students. Because of their habitual thinking, they often think that the purpose of school classes is to deliver applied talents that meet the market demand to the society. Therefore, the employment rate of the school is taken as the final measure of education. Under the background of fierce talent competition, students should strengthen the learning and practice of relevant professional knowledge, so as to master more knowledge and skills. The quality courses and their practical teaching, which do not involve professional knowledge and skills, are useless for students' employment. Therefore, as long as the credits and class hours can meet the requirements of national documents. This concept makes it difficult for quality-oriented teaching in China to get rid of the current situation of formalism, and it is even more difficult to connect with the data age.

3.2. Information Data Filtering and Management Issues. From the basic situation of big data being integrated into quality education in universities, there is a huge amount of information, showing fragmentation and diversity, and positive and negative information coexist, which will have an impact on students. First, as the young generation of Internet users, college students are energetic and flexible, and are in an important stage for young college students, and their concept of life has just been gradually formed. However, the application of the network is also more than that of ordinary people. Due to the diversity of data information they daily contact, some information is complex, and it is difficult to distinguish the good from the bad. If the data use of college students in this period is not well guided, it is bound to have an adverse impact on their behavior values. Second, the application of big data has strengthened people's communication and interaction. Driven by Internet information, the network is full of all kinds of bad information and ideas, which has a certain impact on the main melody and positive energy of relevant education and publicity, and increased the difficulty of follow-up education [20–23]. Third, From the perspective of big data's elimination of useless information and retention of useful information, the whole process of education should be controlled. Due to the relevance of data and the innovation and application of relevant technologies, if not strictly controlled, it is very easy bring about the risk of private information being leaked, posing a threat to students' personal physical and mental health.

3.3. Lack of Specialized Big Data Management Talents. At present, Big data has been highly recognized by colleges and developed in university classes, but the effect is not satisfactory. From the actual education situation, there is a shortage of management talents in related fields, which can be reflected from three different angles: first, although the application of big data has been recognized by both college students and university teachers, it also plays an irreplaceable role, in fact, in the field of big data combined with classroom education, there are not enough relevant technical talents, The daily management and maintenance of digital technology and data technology are not in place, which is difficult to fulfil the basic needs of people and universities for education in real society. Second, in view of the rapid development of digital technology and data technology, in the course of carrying out theoretical education, various universities should formulate scientific talent introduction plans according to specific talent needs to meet the actual educational requirements. But yet, Based on the basic situation of these schools in major provinces in China, under the influence of the data environment, the talent gap is relatively large, especially the number of employed teachers with big data processing ability is relatively small. Third, from the perspective of the university quality education team, teachers' professional quality and ability need to be improved urgently. In the application of new technologies, there is a problem of inadequate utilization. Especially in the entry point of integrating big data and quality education, teachers' actual professional ability and the actual selection and useability are still lacking.

4. The Key and Focus of Integrating Big Data into Quality Teaching in Private Colleges

After China's education reform, the number of universities and big data has shown a blowout, and the integration of big data into school teaching has become an irreversible trend. Especially in the post-epidemic era after 2019, both primary and secondary education and university education have been deeply bound with big data, and there are more and more online teaching platforms and online learning resources [24, 25]. The contradiction between the continuously improving digital education bundle, the gradually increasing number of young educators and the limited administrative ability of the university has become increasingly significant, driving the innovation and reform of the internal education model of the University. At present, many universities, in their own actual situation, should vigorously develop the work of combining data education with classroom, so as to realize the education revolution in the data age [26–28]. However, in view of the actual situation of the current school, we should also consider several issues.

4.1. Modelling of Individual Goals for Older People. Today, with the development of big data, the integration of technology into education is no longer a freshman's idea, but goes deep into the heart of school development. Everyone knows it and can use it. Therefore, in terms of the concept of "integration", today's universities should be alert to relevant technologies' relationship between basic digital teaching and comprehensive digital teaching, so as to avoid the fragmentation, lack of integrity and pertinence of "integration" of data technology [29]. Based on this reason, this study suggests a new model of integrating big data into quality education to change the traditional concept of integrating education data.

The fundamental difference between ideological quality education in the era of big data and that in the era of traditional data is that the former pays attention to the correlation between full data and the usefulness of data information; The latter focuses on the causal relationship between data samples and the accuracy of results [30]. In general, the integration of data technology into the curriculum education of private universities needs to shift from the concept of sampling analysis to the concept of digital statistics. Specifically, the change in the concept of "integration" is mainly reflected in two places.

The first is the transformation from a small amount of sample data to a full amount of data. Before the data technology was applied to the field of quality education in private colleges, traditional digital education mainly used sampling to infer the overall situation due to the difficulties in obtaining data and other factors. After the popularization of big data, full sample data analysis has gradually become the mainstream, and the educational philosophy of relevant education practitioners has also changed.

Second, the concept of causality among digital educators has shifted to the concept of correlation [24, 31]. The big data is applied to discuss and compare the data correlation between teachers and students of private university education, screen invalid education projects, break the limitations of the artificial dominated education method that uses causality to screen the doubtful points of the education concept, reduce the subjective judgment of relevant personnel, and expand the channels for university education to find problems.

4.2. Innovative "Inclusion" Approach. In terms of the "integration" method, attention should be paid to the comprehensive shift from rote copying to organic integration to avoid fragmentation, isolation, lack of penetration, and integration after the "integration" of data technology. Hard copy of big data is the most frequent and common teaching method used in the internal classroom of many private schools. From the perspective of teaching efficiency, although this teaching method has a small workload and simple operation in the early stage, especially for private colleges that have not established a database. However, from the perspective of teaching quality, since teachers mainly adopt data application to integrate relevant teaching classes, it is hard to ensure their teaching quality [29]. If university education is carried out mainly by means of data fusion education based on rote learning, there will inevitably be some problems, such as the lack of teaching supervision links, the insufficient depth of teaching ideas, and the inability to expose systemic risks. Therefore, to souped-up the efficiency and effect of quality classes, it is essential to change the teaching method based on traditional data fusion, open up new "integration" methods with information sharing and communication as the starting point, especially for the key areas and key links of university quality education, build teaching platforms and databases relying on data technology, and excavate the new information contained in these data in a parallel way based on online data and supplemented by traditional teaching methods, We should have a more comprehensive and profound understanding of the relevant classroom education objects, and deal with the problems found in the university classroom promptly, so as to increase the effectiveness of integrating data technology into the internal quality education of colleges and universities. Based on the above problems, the quality education teaching platform integrated with big data built in this test based on the data new era database is shown in Figure 5.

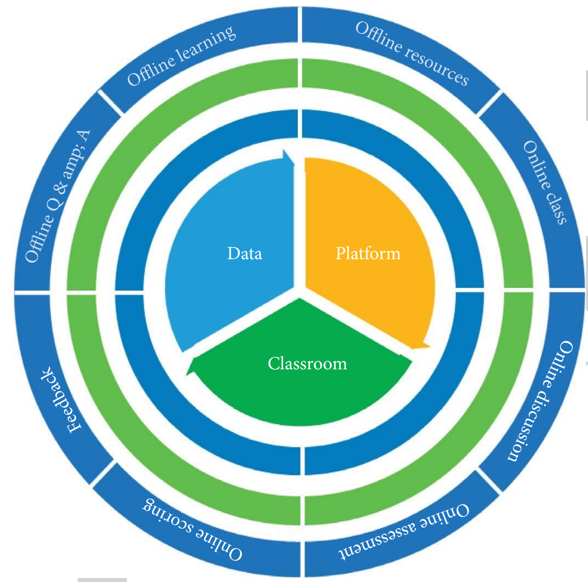


Figure 5: Teaching platform integrated with big data.

4.3. Adjust the "Integration" Approach. In the way of integration, we should change the focus of attention, and comprehensively shift the way of post supervision to the way of pre-event, in-event, post-event and tracking, so as to avoid the situation that the "integration" of digital technology is too narrow and lacks breadth and depth. Most of the quality education models within universities are formulated or evolved by the Ministry of education. Especially for private universities, they are evolved from the traditional universities. This hard copy or evolved big data "integration" method not only causes the course education in these universities to be a little dull, students are not active in class, teachers are not serious in teaching, and other problems. At present, with the promotion of the application of big data, although the curriculum education task of integrating big data has been promoted to a new level by major private universities under the promotion of the situation, this forced development still lags behind, which it is difficult to timely disclose relevant information that is not conducive to offline classes, and it is even more difficult to improve school teaching quality, and seriously affects the teaching classroom effect of students. Therefore, we should try our best to make flexible use of big data to open up more new forms that can improve teaching efficiency, and integrate data technology into the internal curriculum construction of private universities more fully, deeply, and comprehensively. In the stage of preparing lessons in advance, education practitioners collect and share quality data and information through data technology to avoid the inefficient work of manually sorting out data. When the teacher is in class, the data technology is used to quantify the classroom risk value in real-time, and a good classroom feedback mechanism as shown in Figure 6 is established to reduce the deviation and misjudgment of the instructors to a certain extent. After the event, timely track the teaching results, use the technical advantages to establish a student evaluation system on the information platform and the implementation of teachers' feedback on the classroom, and then vigorously improve the teaching quality of private schools. From the perspective of continuous information sharing, real-time monitoring, and effective evaluation, it is imperative to cultivate the "integration" approach to data technology.

5. Path Optimization of Integrating Big Data into Quality Teaching in Private Colleges

5.1. Actively Seek Internal Connection and Create a New Education Platform. As shown in Figure 7, the essence of university quality education is to study the growth process of people and realize the role of personality education. In the teaching process of university quality course, teachers should respect the background of all media, and actively seek the internal relationship between university course education and modern media technology from the professional nature and teaching environment. First of all, global communication in the all-media mode is beneficial to promote the concept updating and content innovation of university curriculum education. When imparting knowledge to students, teachers had better create a new education platform based on the students' growth needs at this stage and their favorite audio-visual forms to spread and partially cover university course education on the Internet. Second, in order to realize the long-term development of curriculum quality education in the form of all media, colleges need to innovate the content of university education and enrich its deep connotation. In the Internet age, college students' interest in

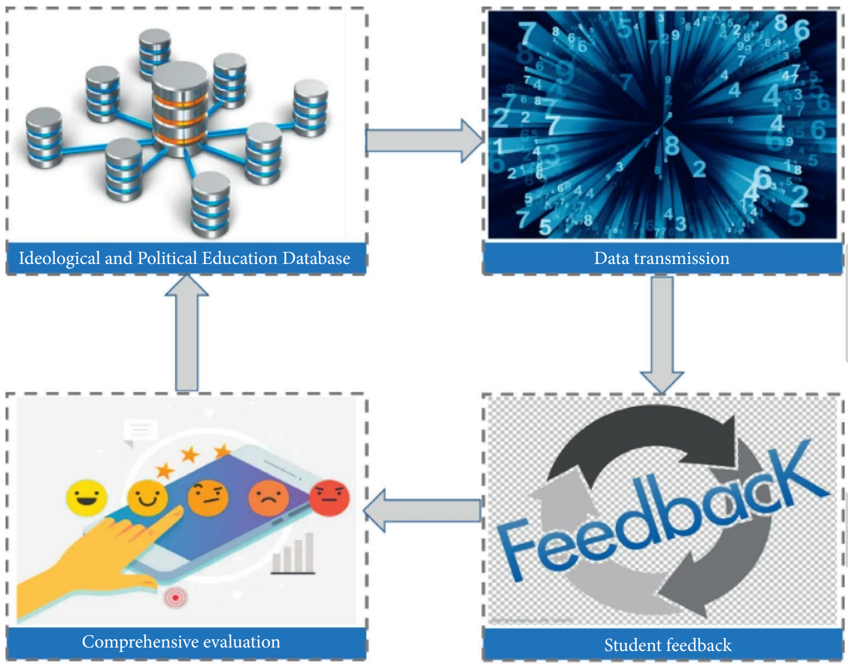


Figure 6: Classroom feedback mechanism.

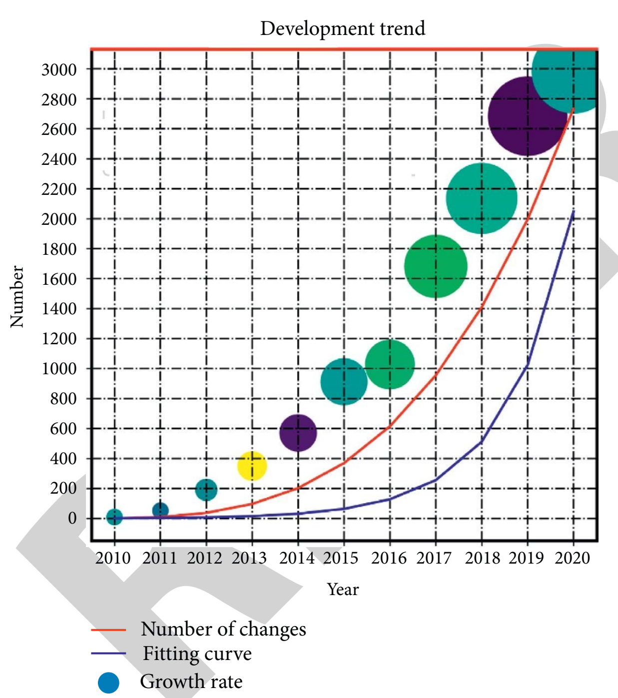


Figure 7: Changes of new education platforms for quality creation at home and abroad.

emerging culture will weaken the influence of traditional culture to a certain extent. In this case, university quality education should start with the spread of network culture, carry out connotation innovation of scientific and healthy values, and ensure the correct cognition of college students on the education of related courses. Thirdly, the creation of a new education platform is a key step for the modernization of university curriculum education. The setting of this education platform has greatly increased the proportion of quality education of college students in daily learning. Students can learn and explore these data on mobile phones, computers, and other electronic devices, greatly improving the enthusiasm of the whole people for political participation. At the same time, the setting of curriculum quality education platform also needs to strengthen the integration with modern network ideas, such as event comments in news media. Educators should grasp their ideological values from the comments of college students in time, and carry out targeted ideological education and personality improvement.

5.2. Positively Carry Forward the Value Concept and Promote the Optimized Development of Quality Work. While optimizing curriculum education, school teachers should keep up with the development of the times when teaching modern courses, and realize the positive promotion of socialist values through all media communication. First of all, innovate the teaching concept of curriculum quality. Under the dual requirements of diversified background and Internet thinking, curriculum quality education should boldly integrate the modern values recognized by students, take the essence, and effectively combine the traditional values with the new cultural trend of thought. For example, the shaping of patriotism values. During the National Day military parade, all media technology ran through the whole process of curriculum quality education. Students witnessed the strength of the country from multiple angles through online viewing. This intuitive curriculum quality education has greatly improved the national identity of college students and eased the current belief crisis of new youth. Second, create quality education products. In the process of university education, Internet technology has become an important supplementary channel for classroom resources. However, due to its large amount of information and complex sources, strict product control and knowledge screening should be carried out in the introduction stage of university classroom education, so as to vigorously promote the positive information beneficial to the shaping of College Students' values. Third, timely disclosing social problems to improve students' independent reflection ability and vigilance awareness. For example, the Central of China is tracking and reporting social emergencies in real-time. This open classroom education caters to the development needs of young scholars era, enables students to evaluate values from the height of national politics, and gradually defines the practical significance of mainstream values through the fusion and identification of multi-party information in the process of practice.

5.3. Enhance Teachers' Modern Consciousness and Strengthen the Integration of Information Technology. Especially in the process of quality education in the university classroom, what kind of talents to cultivate is the primary problem that university education faces. Therefore, strengthening the educational background and quality construction of school teachers have played a fundamental role in the education itself. First of all, we should constantly improve the teaching and self-cultivation level of teachers, strengthen the training of curriculum education for teachers, strengthen the ideological awareness of teachers from the perspective of firm ideals and beliefs, and cultivate patriotism [26], so that teachers can adhere to the direction of socialist school management in teaching, cultivate talents from the Perspective of dialectical materialism, and improve students' personality. Secondly, we should constantly enhance the modernization awareness of teachers. For the impact of global ideology, teachers should guide and care for individual students with modern thinking mode. In the meantime, educators should change the traditional teacher-student education model in the process of education, and guide students out of psychological difficulties with a modern and equal attitude, so as to grow up healthily. Thirdly, enhance the ability of teachers' technical integration. The daily activities of contemporary college students are basically concentrated on the whole Internet. The teaching dilemma brought by this new technology makes education informatization an inevitable trend. Therefore, Improving the teaching ability and data application ability of educators in school is an inevitable choice to effectively meet the needs of students and implement curriculum quality education at the same time.

5.4. Improving the Educational Environment in Colleges and Universities and Paying Attention to the Students' Sense of Autonomy. Campus environment is a comprehensive manifestation of university teaching resources and cultural connotation. It has a significant influence on the improvement of students' humanistic quality and the transformation of their spiritual outlook. Primarily, colleges should build a healthy and positive education environment, integrate the goal concepts of "The aim of education is to educate people" and "Teaching should uphold the advantages of people" into all links of university education, and pay attention to the development of students' self-awareness in an equal way. At the same time, to make more college students get more good jobs when they apply for jobs and have a better life outlook in life, all kinds of universities need to constantly create an atmosphere of socialist values in the campus environment, so that students can constantly motivate themselves and improve their personality. Secondly, all kinds of universities should build an open and free learning environment, actively carry out seminars on Marxist thought, organize quality and moral education courses inside and outside the campus, and cultivate a group of young Party members with firm political positions and both virtue and talent among the student groups through personal experience and practical experience. Through the form of students' mutual assistance, carry out the dissemination and innovation of quality education in the form of all media. Thirdly, colleges should build a "Trinity" educational environment. Fully combine the advantages of the three aspects, integrate and spread in the socialist environment, to urge the school to run the school in the direction of high quality. Therefore, in order to build a modern environment for quality education, colleges and universities should edify students' values from multiple perspectives, So that our students can absorb the fresh learning air and give full play to all their talents in this clean learning environment.

6. Conclusion

The application of big data in higher education management has not only changed the traditional management concept, improved the scientificity and effectiveness of education decision-making, but also optimized the university education management mode including private universities, and accelerated the process of information management. However, under the current situation, college students' education management and teaching work have been inseparable from the support of big data. The rapid development of data technology has had a profound impact on the current university curriculum quality education. Based on the above problems, this paper analyzes the practical dilemma of big data integrating into the quality education of private universities based on the current situation of quality education. And draw the following conclusions:

(1) At present, all kinds of universities have applied data technology to carry out teaching work to meet the needs of teaching development. In practice, university administrators should integrate data technology into the concept, methods and approaches of quality teaching process on university campuses for all-round replacement and improvement.

(2) In the process of modernization transformation of university quality class, in order to get rid of the dilemma of big data integration, teachers should constantly improve their big data thinking, create a

good data exchange platform, establish a high-quality education team, reconstruct the quality education mode, and promote the innovation and reform of University Quality Education.

(3) In order to achieve the goal of optimizing the management of curriculum quality education path, private universities must pay attention to the application of data technology, constantly improve the ways of using big data, improve the university education environment, and attach importance to student's sense of autonomy; Enhance teachers' modern consciousness and strengthen the integration of information technology; Positively carry forward the value concept and promote the optimized development of quality class; Actively seek the internal connection and create a new education platform; Strengthen the effect of personalized education and inject impetus into the cultivation of comprehensive talents in universities.

Data Availability

The experimental data 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 to report regarding the present study.

Acknowledgments

This work was supported by 2021 Party Construction Innovation Project of Education System of Henan Province, "Research on Strengthening the United Front Work of Private Colleges and Universities," project number: 2021--DJXM-135.

References

[1] Y. Wang, "Big data era influence on college students' ideological and political and innovation strategy," in *Proceedings of the Eighth International Conference on Measuring Technology & Mechatronics Automation*, IEEE, Macau, China, March 2016.

[2] Z. Du, "Cultural development of ideological and political under the background of big data," in *Proceedings of the ICIMTECH 21: The Sixth International Conference on Information Management and Technology*, Huai An City, China, October 2021.

[3] F. Wang, "Analysis on the application of ideological and political education in student management based on big data," *Journal of Physics: Conference Series*, vol. 1744, no. 3, Article ID 032069, 2021.

[4] Q. Gao and L. Ai, "Problems of ideological and political education of college students in the era of big data," *Journal of Physics: Conference Series*, vol. 1852, no. 3, Article ID 032037, 2021.

[5] T. Chen, "Research on the dilemma and breakthrough path of ideological and political education in colleges and universities in the era of big data," *Journal of Higher Education Research*, vol. 3, no. 2, pp. 203–206, 2022.

[6] H. Wang, "Research on improving computer teaching quality in colleges and universities in the context of big data[J]," *Journal of Physics: Conference Series*, vol. 1738, Article ID 012049, 2021.

[7] C. Zhai, L. Sun, H. Xu, and L. Zhou, "Thoughts on "curriculum ideological and political " in universities based on big data," in *Proceedings of the 2021 10th International Conference on Educational and Information Technology (ICEIT)*, Chengdu, China, March 2021.

[8] L. L. Hou and J. Lin, "Application of big data in ideological and political," *Journal of Physics: Conference Series*, vol. 1744, 2017.

[9] Z. Zhu and J. Zhang, "Research on the innovation of ideological and political mode in colleges and universities based on big data era," in *Proceedings of the 2017 9th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, Changsha, China, January 2019.

[10] X. Ping, "Research on the effects of ideological and political in colleges in the era of big data," *Journal of Physics: Conference Serie*, vol. 1, 2018.

[11] W. Di, *Research on the innovation and development model of ideologicaland political education in we-media environment based on big data*, NY China, 2017.

[12] Z. Yandong, "On the influence of big data on the Ideological and Political in colleges and universities," in *Proceedings of the 2015 International Conference on Social Science and Technology Education*, Amstelkade, July 2015.

[13] W. Yang, C. Ling, and L. Fan, "Exploration of new methods of ideological and political for college students: integration of new media and construction of micro-environment[J]," *Geophysical and Geochemical Exploration*, vol. 34, no. 1, pp. 108–110, 2016.

[14] J. Y. Zheng and Z. B. Wang, "Research on integrating the thought of ruling by law into ideological and political education in higher vocational colleges under the new normal," *Journal of Heilongjiang College of Education*, vol. 13, 2017.

[15] X. Lv, "Big data in ideological and political in colleges and universities application and reflection," in *Proceedings of the ICBDE '20 2020 The 3rd International Conference on Big Data and Education*, Ottawa, ON, Canada, October 2020.

[16] Y. Zhou and Y. Cao, "Study about impact of big data technology on the diversity of ideological and political teaching methods," *Journal of Physics: Conference Series*, vol. 1852, no. 3, Article ID 032019, 2021.

[17] T. Zhang, "Construction and application of big data analysis platform for ideological and political in colleges," in *Proceedings of the 2019 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, Hohhot China, July 2019.

[18] Q. Qianchao, "Exploration on the reform of ideological and political in colleges and universities in the age of big data," in *Proceedings of the 2019 5th International Conference on Education Technology, Management and Humanities Science (ETMHS 2019)*, Kunming, Yunnan, June 2019.

[19] R. Zheng, X. Liu, S. Tang, and T. Chen, "Research on quality education in universities in the era of big data," in *Proceedings of the 2020 12th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, NY China, July 2020.

[20] S. Liu, "Research on university ideological and political based on big data information teaching mode," in *Proceedings of the 2020 International Conference on Computers, Information*

Hindawi
Mobile Information Systems
Volume 2022, Article ID 6943120, 11 pages
https://doi.org/10.1155/2022/6943120

Research Article

Intelligent Tourism Marketing and Publicity Methods for Revenue Enhancement

Ying Lu [1] and Bo Cui [2]

[1] *School of Applied Technology, University of Science and Technology Liaoning, Anshan 114051, China*
[2] *Engineering & Research Institute of Ansteel Corporation, Anshan 114000, China*

Correspondence should be addressed to Ying Lu; z51414013@stu.ahu.edu.cn

Received 14 July 2022; Revised 1 August 2022; Accepted 8 August 2022; Published 28 August 2022

Academic Editor: R. Mo

For online e-commerce platforms, big data intelligent marketing is an essential tool for promoting companies, creating pictures, participating in competitions, and engaging customers. This also gives traditional marketing a new sense of precision and a new approach to increasing revenue. Marketing and public relations based on big data analytics and intelligence should be considered important and profitable for the travel and tourism business. Tourism enterprises need to learn to use intelligent technology to guide marketing activities and formulate comprehensive and accurate marketing strategies. Infiltrate the advantages of intelligence into all aspects of tourism marketing, so as to form an intelligent, precise, and modern tourism marketing and publicity model, and help enterprises to improve their income. This work focuses on the research on intelligent tourism marketing and publicity. The main research contents include the following aspects. First, this work proposes a revenue-enhancing tourism marketing and publicity method, which is mainly divided into the use of intelligence to enrich the form of publicity and marketing and the use of intelligence to refine the content of publicity and marketing. Second, this work proposes an IWOA-BP network for evaluating a revenue-enhancing intelligent tourism marketing promotion method. It improves the WOA algorithm by introducing a nonlinear convergence factor and adaptive crossover mutation to construct the IWOA algorithm. Then IWOA is applied to initial weights and thresholds for optimization, which can solve the drawbacks of the traditional BP network and improve the performance and reduce the training time. Third, this work conducts a comprehensive evaluation of the proposed revenue enhancement-oriented tourism marketing and publicity methods and IWOA-BP, and the experimental results verify the feasibility of these methods.

1. Introduction

With the improvement of residents' living standards, the demand for tourism is becoming more and more strong, and the tourism industry developed from this has already played a very important role in the national economy. The tourism industry has shown a vigorous development momentum, and the proportion of tourism in the economic development will be further enhanced. The rapid development of information technology is significant in the vigorous development of tourism. In particular, the development of intelligent tourism not only provides effective information consulting services for the majority of tourists but also provides technical support for the innovative marketing of tourism enterprises. Due to the enhancement of economic strength and the acceleration of the globalization process, not only the international competition has been strengthened, but also the competition among domestic tourist attractions has become more and more fierce. If tourist attractions want to improve their competitiveness, they need to consider making a fuss about communication and marketing methods. The concept of intelligent tourism marketing and publicity is valued by more and more tourist attractions in today's era. The connotation of this concept is to regard scenic spots as enterprises, tourists as customers, and through intelligent marketing and publicity means, the scenic environment, including hard and soft environments, are manufactured as products. The purpose of tourism marketing and publicity is to make tourist attractions better, through cultural publicity, tourism development, etc., to make consumers pay the bill. To achieve this purpose, it is necessary to utilize intelligent tourism marketing and publicity methods [1–7].

On the one hand, the social needs of groups are gradually changing from material needs to spiritual needs, and consumption levels and needs are becoming more and more diverse. The quality requirements for goods and services are also constantly escalating, and factors such as consumer personalization, customization, and group belonging have gradually become the leading consumer decision-making for many emerging consumer groups. On the other hand, in the context of the Internet, intelligent transformation is an inevitable trend of economic and social development. Under the background of digitalization and intelligence, the tourism market structure has changed, and market competition has become more fierce in the digital age. In order to establish a brand image in the competition and seize market share, major tourism platforms have used big data to carry out various precise marketing activities. In such a market environment, learning to make good use of big data to carry out marketing, mining user needs, and carrying out targeted promotion is very important to enhance the core competitiveness of enterprises [8–14].

Intelligent tourism marketing is conducive to changing the traditional concept of scenic spot management and comprehensively improving the image of tourist scenic spots. Intelligent tourism marketing is different from the past tourism management in terms of word creation, and also has a big difference in meaning. The difference between the two is that the management of the latter places too much emphasis on administrative tendencies, while the former is more inclined to guide the construction of scenic spots with theories of marketing. In terms of purpose, the main purpose of scenic spot management is to strengthen interaction and ensure the harmony of scenic spots. Tourism marketing and publicity can bring changes to the scenic environment, enhance the attractiveness of tourist attractions, and enhance the overall image of the scenic spots. Intelligent tourism marketing is conducive to building a tourism brand and establishing a competitive advantage in tourist attractions. Building a tourism brand can bring various benefits to tourist attractions, such as enhancing the attractiveness of business investment, enhancing the pride of residents in the scenic area, and ensuring the interests of residents. And it is conducive to the scenic spot to participate in the competition and enhance its influence in the country. The building of tourism brands is also of great significance. Through brand building, the scenic spot can have a clear positioning. And it has a certain degree of recognition nationwide, attracting foreign business investment and social resources, etc. Many scenic spots have noticed this, and actively promote intelligent tourism marketing and publicity through various means to create a unique brand advantage for tourism. Intelligent tourism marketing is conducive to promoting the innovation of tourism taste and realizing the sustainable development of scenic spots. The formulation of tourism marketing promotion strategy can refer to the marketing strategy. Taking tourists as consumers, what efforts should be made to meet the needs of consumers is what tourism marketing needs to pay attention to. By analyzing the development status of the scenic spot and its possible future development trend, it is necessary to use a growth-oriented development strategy by analyzing scenic spots [15–20].

To promote the high-quality development of the tourism industry, this work studies tourism marketing and publicity methods for revenue enhancement. First, this work proposes a revenue-enhancing tourism marketing and publicity method, which is mainly divided into the use of intelligence to enrich the form of publicity and marketing and the use of intelligence to refine the content of publicity and marketing. Second, this work proposes an IWOA-BP network for evaluating a revenue-enhancing intelligent tourism marketing promotion method. It improves the WOA algorithm by introducing a nonlinear convergence factor and adaptive crossover mutation to construct the IWOA algorithm. Then IWOA is applied to initial weights and thresholds for optimization, which can solve the drawbacks of the traditional BP network and improve the performance and reduce the training time. Third, this work conducts a comprehensive evaluation of the proposed revenue enhancement-oriented tourism marketing and publicity methods and IWOA-BP, and the experimental results verify the feasibility of these methods.

2. Related Work

Literature [21] believed that if a region wanted to successfully use destination tourism marketing, it should plan a detailed marketing strategy in the scenic spot development stage. Moreover, if a region wanted to develop a tourist destination, it not only depended on the planning and marketing of the scenic spot itself, but also on the support of the local government. This also involved issues such as whether the development of tourism destinations harmed the interests of local residents, and whether it could bring sustainable economic development to residents. These all affected how companies used tourism marketing methods scientifically. Literature [22] proposed that tourism enterprises could increase the exposure of tourism enterprises according to the influence of the Internet, so as to convert the traffic obtained by enterprises into the competitiveness and productivity required by enterprises. Literature [23] believed that with the continuous progress of the Internet era and marketing theory, the importance of intelligent marketing to scenic spots is self-evident. It proposed that the scenic spot used intelligent media to strengthen the interaction with tourists, which could invisibly strengthen the publicity of the scenic spot and increase loyalty to the scenic spot in the interaction with the tourists. Literature [24] studied the role of mobile marketing in tourism marketing in the era of intelligent marketing. The research results showed that the status of consumers in the process of searching for travel information was the most important reason for triggering behavior, and videos with rich content could fully exert the emotional triggering effect. The more participants like video, the more it would affect their engagement intention and resonance, and the effect of media content richness was positively related to marketing effectiveness. Literature [25] studied the marketing strategy of scenic spots and proposed a sustainable development marketing model. On the basis of sorting out its tourism development process, combined with SWOT analysis and STP analysis framework, a tourism marketing strategy was constructed using the 4P marketing theory system. Literature [26] pointed out that under the political background of supply-side reform, the tourism industry was the gold industry of the tertiary industry, and various regions were vigorously developing tourism. With the continuous development of the tourism market, the number of potential tourists brought by festivals was also very considerable. Reference [27] took scenic spots as an example, used questionnaires and interviews to analyze the current development status of scenic spots and the existing problems in marketing, and improved the tourism marketing strategy of scenic spots according to their existing problems. Literature [28] believed that the growth point of the future economic income would largely depend on the experience of tourists. In this paper, factors that may affect the tourists' sense of experience are expounded, and suggestions on how to enhance the tourists' sense of experience and comfort in the scenic spot were put forward.

Literature [29] pointed out that when tourists choose tourist destinations, they pursued the cultural experience brought by scenic spots. According to the current preference of tourists in pursuit of a sense of cultural experience, the scenic spot could try to integrate the local characteristic culture with the tourism of the scenic spot, launched the marketing strategy of the region, and put forward the concept of cultural tourism. Literature [30] proposed that with the continuous development of tourism, consumers enjoyed more tourism with cultural elements and enjoyed the sense of experience brought by tourism. The trend of taking food culture as the center of tourism marketing had begun to appear in the tourism market. Literature [31] proposed the current marketing problems, and proposed a tourism marketing strategy centered on food culture according to the current problems. Literature [32] expounded on the development trend of the current tourism market, and the current competitive trend of the tourism market had been separated from the way of relying on price wars. With the improvement of people's material living standards, the competition to build scenic spots had begun. Most scenic spots did not have brand awareness. In the development and exploration of scenic spots, according to the actual situation of their own scenic spots, they could create a brand strategy suitable for the actual operating conditions of the scenic spots and improved popularity. Literature [33] pointed out that with the improvement of living standards, consumers were more inclined to choose comfortable and enjoyable travel modes. Gone was the situation where group tours did not pay attention to the diet of the tourist destination in the past. Now more young people usually made a strategy for the diet of the tourist destination before choosing a tourist destination. Tourism food marketing had begun to have an impact on the profitability of scenic spots, and food and beverage marketing had begun to affect scenic marketing. Literature [34] believed that the emergence of new media had made the connection between people more convenient, and people's opinions and evaluations on a certain thing through new media have diversified voices. Many industries had seized the characteristics of new media and used the characteristics of new media as the main focus to develop their own industries. Literature [35] believed that although many scenic spots were using marketing methods with good performance at this stage, due to a large number of tourism resources, some tourism resources were inherently insufficient or the scenic spots did not pay enough attention to intelligent marketing. Literature [36] studied the intelligent strategy of tourist attractions, analyzed the current situation of intelligent marketing of scenic spots, and gave suggestions for improving tourism intelligent marketing. Reference [37] introduced in detail the contribution of intelligent marketing to the marketing of scenic spots in the Internet era. It allowed potential users to truly see the characteristic scenic spots of the scenic spot, thus generating the idea of who want to experience it themselves and attracting tourists to the scenic spot. Literature [38] believed that the intelligent marketing of scenic spots should be combined with traditional marketing methods, which was more beneficial to the economic benefits of scenic spots. Through the above analysis, the popularity of the Internet had brought us intelligent new media, and intelligent marketing in tourism marketing was also an important marketing method.

3. Method

First, this work proposes a revenue-enhancing tourism marketing and publicity method, which is mainly divided into the use of intelligence to enrich the form of publicity and marketing and the use of intelligence to refine the content of publicity and marketing. Second, this work proposes an IWOA-BP network for evaluating a revenue-enhancing intelligent tourism marketing promotion method. It improves the WOA algorithm by introducing nonlinear convergence factor and adaptive crossover mutation to construct the IWOA algorithm. Then IWOA is applied to initial weights and thresholds for optimization, which can solve the drawbacks of the traditional BP network and improve the performance and reduce the training time.

3.1. Intelligent Tourism Marketing Promotion Strategy. The first is to use intelligent technology to enrich the forms of marketing and publicity, and intelligence provides innovative forms of tourism marketing and publicity. This article will discuss how to combine intelligence with different forms of marketing promotion from the level of membership marketing and opinion leaders, break the limitations of traditional marketing models, and make promotional activities more in line with tourism characteristics and consumer psychological needs. Like other e-commerce platforms, travel products pay more attention to the marketing management of members. The more common affiliate marketing methods are push messages and e-mail advertisements, which are based entirely on user data within the platform. This can directly connect with the target audience, but there are also disadvantages such as disturbing the user, causing the user to be bored and numb, and even causing the user to uninstall the software. Therefore, in the process of carrying out tourism marketing and publicity, we must make full use of intelligent analysis and grasp the scale of publicity. When carrying out marketing promotion, it is first necessary to establish a user database to subdivide audience characteristics, regions, and preferences. This part can be completed by the customer's own choice, or a user model can be established through big data analysis of user behavior. At the same time, in the process of marketing and publicity, we must grasp the publicity time, which will directly affect the acceptance of the audience's publicity information, and set corresponding time nodes according to different groups of people. For travel, the travel tips and experiences of travel experts can undoubtedly play the role of opinion leaders. The sharing of these experts from social media on the Internet does not directly involve product promotion and brand marketing, but explores each different travel location from the perspective of travelers. This kind of communication form of putting oneself in the shoes, through the reprinting and promotion of social media and special websites, will form an amplification effect, and over time, the travel notes of these talents will become more authoritative. Prediction is an important function of intelligence, and it is also one of the attributes of opinion leaders. Traditional opinion leaders may gain experience based on personal behavior, while data is epoch-making, and these experiences and future expectations can be revealed through data. The predictive information detected by these data is then combined with the opinions of opinion leaders, thereby expanding the professionalism and authority of its marketing promotion content.

The second is to use intelligent technology to refine the promotion content. In the presentation of marketing promotion content, advertising copy and visual presentation are two important contents. This article talks about how to use intelligent technology to make the copy content more targeted, the product display more representative, and the visual presentation more experiential. The first is to create targeted promotion copywriting. In the process of using the media, the audience gives up the opportunity to actively interpret and replaces the active establishment of self-awareness with passive sensory enjoyment. That is to say, the communication content presented by the carrier can exceed the rational influence on the audience to a certain extent. There are large differences between the groups of users to which different types of headlines are intended. On the whole, title copy can be divided into two categories: rational copy and emotional copy. Most of the keywords involved in the title of the rational copy are numbers and rational words, which can directly show the interests in tourism products, and this type of copy is often more attractive to rational consumers. Perceptual vocabulary is divided into two directions: curiosity and emotion. The starting point of perceptual vocabulary is to stimulate the emotional nerves of the audience, thereby generating emotional resonance. A copywriting style that combines sensibility and rationality is easier to catch the user's attention. For the same target group and the same series of travel products, the copywriting style that combines the virtual and the real will eventually get a higher click-through rate and order conversion. The second is to create impactful visual displays. Society has already entered the era of the eyeball economy, and audiences are reluctant to spend a lot of time reading advertising copy. On the contrary, the display of advertisement images or video content has a greater visual impact and influence on users. In the process of big data marketing, it is necessary to be able to recognize the importance of visual presentation, and to improve the visual experience as the ultimate goal. For the content of pictures, through data statistics, the pictures with the highest ranking of users' clicks are the most used for marketing and promotion, and the effect will be better in attracting the attention of similar consumers. The third is to set up a dedicated product page. Using data can make the presentation of tourism product information more representative. Through this customized advertising page, on the basis of attracting potential consumer groups, the conversion rate of tourism product sales can be greatly improved. The purpose of setting up a dedicated travel product page is to make travel products more targeted. Through personalized copywriting and creative visual impact, segmented audiences can be diverted to the tourism product display page, and whether these more accurate target users can make purchases depends on the product in the final analysis. These exclusive travel products are based on the answers obtained by users' search behaviors and transaction traces.

3.2. BP Algorithm. The signal of the BP network is propagated forward, and the error is propagated backward. After the forward propagation is processed by each layer, when there is an error between output and expected result, it enters the backpropagation. The weight of each layer is adjusted one by one, which can not only solve the nonlinear problem between the input and output but also complete the function of self-learning. The BP network can also learn and store the mapping relationship between input and output without knowing that there is a certain mapping relationship between functions. Through the comparison of input samples, the network weights and thresholds are continuously adjusted.

$$y = f\left(\sum w_i x_i + b\right), \tag{1}$$

where w is weight, x is the feature, and b is the bias.

The input layer, the hidden layer, and the output layer are all components of the BP network. In general, the number of nodes in the input layer and the output layer corresponds to the dimensions of the input vector and the output vector. You have complete freedom in determining the number of nodes and layers in the hidden layer, and the ideal number can be chosen based on the actual model requirements. Generally speaking, the higher the accuracy required for model training, the more hidden layers and nodes are required, but the computational complexity will also increase. BP structure is demonstrated in Figure 1.

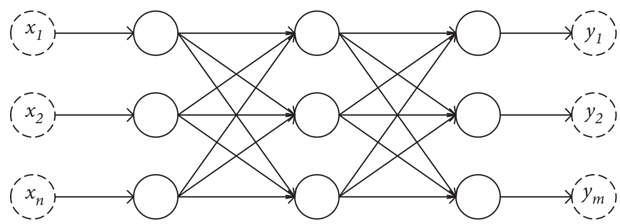


FIGURE 1: BP structure. The commonly used activation function in BP network is Sigmoid.

$$\text{Sigmoid}(x) = \frac{1}{(1 + \exp(-x))}, \tag{2}$$

where x is the input.

The application of the BP network is very extensive, and it has many advantages. First, it can store information in a distributed manner. When it comes to neural networks, information is not merely stored in a single location, but rather spread over all of the processors and the connections between them. Instead, it is dispersed throughout the network based on the type of data being transmitted. When the input signal of the neural network changes suddenly, it does not affect the correct output of the network, so the neural network has a strong ability to defend against faults and fault tolerance. Second, it can process information in parallel. For the input and output information, each neuron processes and outputs in a massively parallel manner, and each neuron in the same layer can process the information at the same time. This method of parallel processing and information storage greatly improves the emergency response capability of the system, improves the computing speed of the system, and improves the real-time performance of the neural network. Third, it combines information storage and processing into one. The storage and processing of BP network information are integrated, and each neuron not only has the function of storage but also has the function of information processing. The change of connection weights between neurons not only reflects the memory of information, but also reflects the process of information processing. Fourth, its self-organization and self-adaptive ability are strong. The neural network can learn and quickly adapt to the characteristics of the system with uncertainty, and can automatically find the law according to a certain relationship between the input and output of the system, and memorize it into each weight of the neural network.

BP network exposed more and more shortcomings, mainly in the following aspects. The first is the contradiction between learning rate and stability. The gradient descent algorithm has a small learning rate during the learning process, resulting in a slow convergence rate. The additional momentum method is usually faster than the simple gradient method because it takes into account not only the effect of the changing trend on the error surface but also the effect of the error on the gradient. However, to ensure the stability of the learning process, a higher learning rate is usually selected. The second is that it is difficult for nonlinear networks to choose the learning rate. For the learning rate in the linear network, if the setting is too large, the learning process will be unstable. On the contrary, setting too small will prolong the learning time and affect the efficiency of learning. The problem that the learning rate of the nonlinear network cannot be valued is still to be solved. The third is the local minima problem. One of the main shortcomings of traditional BP network is that the training process is easy to fall local minima. The generation for local minima is closely related to the selection of network weights. When two initial points are very close together may eventually one will iterate to a local minimum point. When there is a function curve with multiple minima, the algorithm is easy to fall into it and cannot escape, which leads to the failure of network training. The fourth is that the initial weight dependence is too strong. The BP network has a strong dependence on the initial weight. The initial weight has a certain influence on the training time of the neural network and the reliability of the final training result. Selecting different initial weights for training will often converge to different local minima, which is why the results obtained from each training will be different. However, in most cases, random numbers are used as initial weights, which not only affects the learning efficiency of the BP neural network, but also has low practical value.

3.3. WOA Algorithm. The WOA algorithm is a population-based optimization algorithm proposed by simulating the hunting behavior of humpback whales. It realizes the local development function of the algorithm through the shrinking and wrapping or spiral upward mechanism and realizes the global exploration function of the algorithm through the random search mechanism. The algorithm mainly has three behaviors: surrounding the prey, attacking the prey, and randomly searching for the prey.

An approach to a solution can be approximated as a humpback whale circling its prey, and this process can be compared to the computer process of searching for a solution. Updated positioning surrounds the target if it's judged that it's the best prey available at the moment.

$$\begin{aligned}
D &= \left|X(t) - \mathrm{CX}_{\text{best}}(t)\right|, \\
X(t+1) &= X_{\text{best}}(t) - \mathrm{AD}, \\
A &= a(2r_1 - 1), \\
C &= 2r_2,
\end{aligned} \tag{3}$$

where t is the iterations, A and C are the coefficient vectors, $X_{\text{best}}(t)$ is the best whale position, and $X(t)$ is the current position.

When humpback whales attack their prey, there are mainly two ways of shrinking and surrounding or spiral surrounding. The shrinking and surrounding mechanism is realized by reducing the parameter value. The spiral encirclement mechanism is to update the position through the spiral position formula, so as to hunt the selected target prey. During the actual hunt, humpback whales swim around their prey in a shrinking circle while simultaneously swimming along a spiral path. In order to simulate these two simultaneous hunting behaviors, the algorithm assumes a probability of 50% to choose between the shrinking encirclement mechanism and the spiral encirclement mechanism, so as to update the position vector.

$$X(t+1) = \begin{cases} X_{\text{best}}(t) - \text{AD}, & p < 0.5, \\ D * \exp(bl) * \cos(2\pi l) + X_{\text{best}}(t), & p \geq 0.5, \end{cases} \tag{4}$$

where b is the logarithmic spiral shape constant, l is a random number.

Whales' hunt for prey can be compared to the process of finding the best solution to a problem. The whale algorithm can also be used to search for prey using the same method based on the change of vector A. By changing the value of A to be greater than 1 or less than −1 to force away from the current best position vector, allowing for a global random search.

A basic algorithm, a few adjustments, and a straightforward implementation are all WOA's advantages. As a result, there is no way to prevent the whale algorithm from slipping into a local best. 'It still has flaws like slow convergence, limited convergence accuracy, and an easy fall into local optimum, like other clever algorithms.

3.4. Improved WOA and IWOA-BP Algorithm. The search technique is largely determined by parameter value adjustment in the WOA algorithm. The global search method is used when the absolute value is larger than 1, while the local search technique is used when the absolute value is less than 1. The convergence factor dictates the parameter value's magnitude. Convergence factor is a measure of how well the algorithm can search globally. When the convergence factor is low, its capacity to do local searches is more powerful. When the algorithm reaches its middle and late stages, the convergence factor drops linearly from 2 to 0, making it easy to fall into a local optimal solution. A nonlinear convergence factor (NCF) is utilized to solve these issues. Its value is kept high in the beginning and low in the latter stages of the process. The algorithm's convergence speed is boosted as a result of balancing global and local search abilities in the early stages.

$$a = 2 - \frac{1.7}{1 + \exp\left(-\alpha\left(2t - t_{\max}\right)/2t_{\max}\right)}, \tag{5}$$

where α is the control factor.

Combining two individuals from separate parents to produce a new person is known as the crossover process in genetic algorithms and can significantly improve the algorithm's global optimization capabilities. The process of mutation refers to the creation of a new individual through the alteration of a specific gene on the individual, allowing the algorithm to perform local optimization. To avoid early convergence in an iterative process, it can keep population variety intact. This study provides the crossover and mutation technique to the genetic algorithm in order to improve its global optimization ability to further improve the performance of the whale algorithm.

For the crossover operation, individuals in the population are first selected and paired with each other at random. Then let the paired individuals perform crossover operations according to a predetermined probability. For crossed individuals, the calculation of the crossover operation is as follows.

$$\begin{aligned} y_i^{k+1} &= \beta y_i^k + (1-\beta) y_j^k, \\ y_j^{k+1} &= \beta y_j^k + (1-\beta) y_i^k, \end{aligned} \tag{6}$$

where β is a random number.

For mutation operation, let individuals perform mutation operation according to a predetermined probability.

$$y_i^d = y_i^d (1 + 0.5\nu), \tag{7}$$

where ν is a random number.

The algorithm's ability to optimize itself was profoundly impacted by his work on establishing crossover and mutation probabilities. When the likelihood of crossover is low, the rate at which new people are generated is slower, and the capacity to perform a global search is diminished. The frequency of acquiring new individuals through mutation operation decreases when the mutation probability is low, and the variety of the population cannot be maintained. Individuals close to the optimal answer may be wiped out if the mutation chance is high enough, resulting in a nearly random search. This research uses an adaptive crossover and mutation (ACM) system, which allows the likelihood of crossover and mutation to be changed with the population's evolution.

$$\begin{aligned} p_c &= \begin{cases} \dfrac{k_1\left(f_{\max} - f\right)}{f_{\max} - f_{av}}, & f \geq f_{av}, \\ k_2, & f < f_{av}, \end{cases} \\ p_v &= \begin{cases} \dfrac{k_3\left(f_{\max} - f'\right)}{f_{\max} - f_{av}}, & f' \geq f_{av}, \\ k_4, & f' < f_{av}, \end{cases} \end{aligned} \tag{8}$$

where $f_{\max}$ is the maximum fitness, f_{av} is the average fitness, f is the larger fitness of the two individual performing crossover, and f' represents the fitness of the mutant individual.

Crossover and mutation operations are less likely to occur when an individual's fitness is better than the average, allowing it to be passed down to future generations. When an individual's fitness is lower than the average, crossover and mutation operations are more likely, making it easier for the individual to be eliminated.

An upgraded version of the IWOA algorithm is used to create IWOA-BP. Optimizing the BP network's initial weights and thresholds with the revised IWOA method relies on using the training error of the BP network as a measure of population fitness. A final step is to apply these optimum values as initial values for the BP network. The pipeline of IWOA-BP is demonstrated in Figure 2.

The specific operation steps of the IWOA-BP network are divided into some steps. The first step is to initialize the BP network, determine the main structure of the network, and generate initialization weights and thresholds. The second

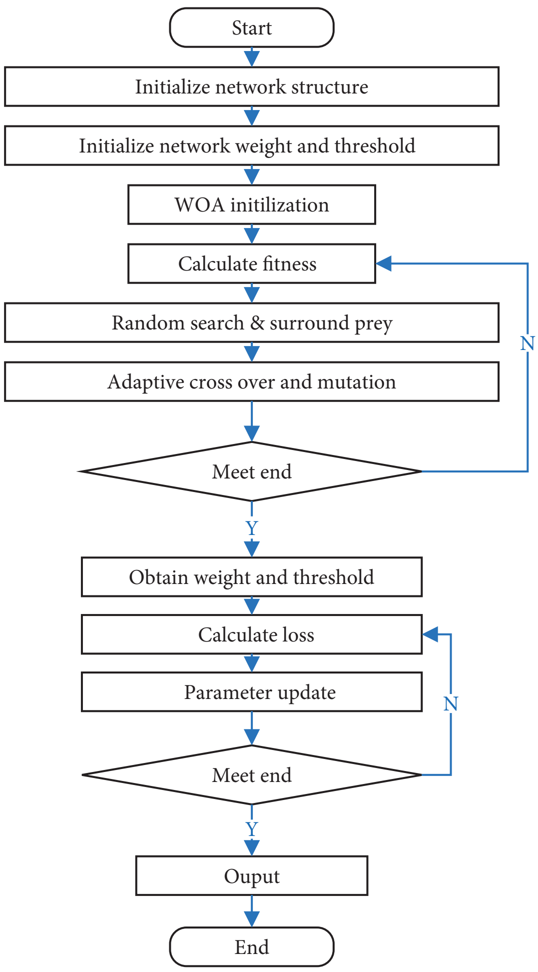


Figure 2: IWOA-BP pipeline.

step, the initialization of IWOA, converts the initialization weights and thresholds in the above steps into position vector of the improved IWOA, and initializes other parameters of the IWOA algorithm. At the same time, MSE is utilized as the fitness of the improved IWOA. The third step is to calculate fitness, find the position of optimal fitness, record the position vector and use it as the current optimal individual position, and calculate average fitness. The fourth step is to choose a strategy to surround or randomly search for prey. The fifth step is the individual crossover, individual crossover operation is performed according to the established adaptive crossover probability. The sixth step is an individual mutation, and the individual mutation operation is conducted via established adaptive mutation probability. The seventh step is to output the results. When the maximum number of iterations of the algorithm is met or the required error accuracy is reached, the algorithm ends, and the optimized weights and thresholds are output to the BP network.

4. Experiment and Analysis

4.1. Dataset. This work collects the tourism revenue data under the corresponding intelligent tourism marketing publicity method to construct the data required for the training and testing of the IWOA-BP network. The dataset contains a total of 29,187 samples, of which 19,038 are training samples, and the rest are test samples. Each data sample contains 8 feature indicators, as demonstrated in Table 1, and the corresponding label is the income level. The evaluation indexes in this work are accuracy and precision.

$$Acc = \frac{(\mathrm{TP} + \mathrm{TN})}{(\mathrm{TP} + \mathrm{FN} + \mathrm{FP} + \mathrm{TN})},$$
$$Rec = \frac{\mathrm{TP}}{(\mathrm{TP} + \mathrm{FN})}. \tag{9}$$

4.2. IWOA-BP Training Experiment. This work analyzes the training of the IWOA-BP network. The main analysis objects are the training loss, training accuracy and training recall during the network training process. The data are demonstrated in Figure 3 and Figure 4.

With the increase in training times, the loss of the IWOA-BP network shows a downward trend, and the correct rate and recall rate show a downward trend. When the number of training reaches 50 epochs, all three tend to stabilize, at which point the network training has converged.

4.3. Method Comparison. To verify the feasibility of IWOA-BP for evaluating the income effect of the intelligent tourism marketing and publicity method, this work compares it with other methods. The compared methods include SVM, KNN, and DBN. The comparison data is demonstrated in Table 2.

IWOA-BP method can obtain the highest accuracy rate and recall rate. Compared with other machine learning methods, this method can achieve different degrees of performance improvement in both accuracy and recall.

4.4. Nonlinear Convergence Factor Analysis. This work uses nonlinear convergence factor (NCF) to improve the traditional WOA algorithm. To verify the superiority of this improvement, the performance of traditional convergence factor (TCF) and NCF are compared respectively. The comparison data is demonstrated in Figure 5.

Compared with the traditional convergence factor, after using the improved nonlinear convergence factor, the IWOA-BP network can achieve a 1.9% and 1.6% improvement in accuracy and recall.

4.5. Adaptive Crossover and Mutation Analysis. This work uses adaptive crossover and mutation (ACM) to improve the traditional WOA algorithm. To verify the superiority of this improvement, the performance of traditional crossover and mutation (TCM) and ACM are compared respectively. The comparison data is demonstrated in Figure 6.

Compared with the traditional crossover and mutation, after using the improved adaptive crossover and mutation, the IWOA-BP network can achieve a 1.6% and 1.1% improvement in accuracy and recall.

Table 1: Feature indicators of data sample.

Index	Feature
x_1	Operating profit margin
x_2	Return on assets
x_3	Return on capital
x_4	Capital appreciation rate
x_5	Account recovery rate
x_6	Inventory turnover
x_7	Profit per capita
x_8	Reception per capita

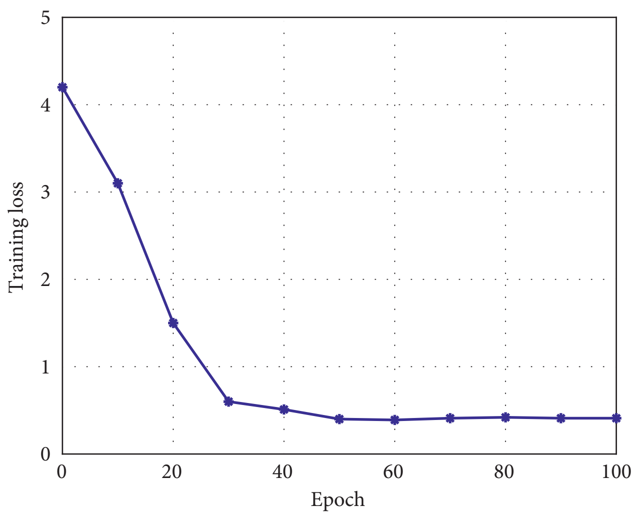


Figure 3: Training loss analysis.

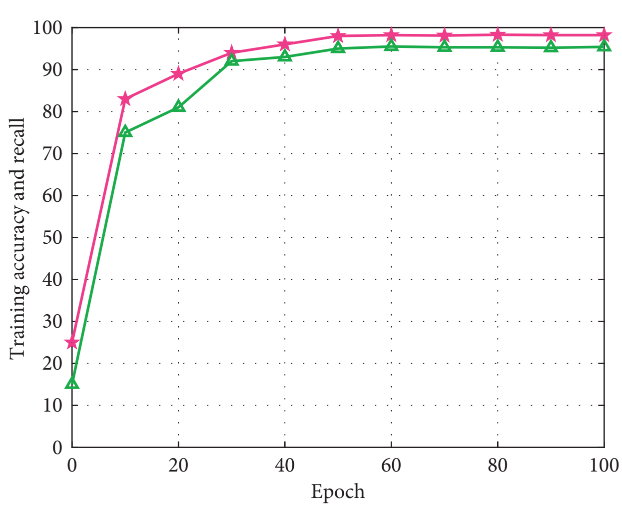


Figure 4: Training accuracy and recall analysis.

4.6. IWOA Analysis. This work uses the IWOA algorithm to optimize the BP network. In order to verify the superiority of this optimization measure, the accuracy and recall rate of the traditional BP, WOA-BP and IWOA-BP methods are compared respectively. The comparison data is demonstrated in Table 3.

Table 2: Comparison with different method.

Method	Accuracy	Recall
SVM	90.3	87.8
KNN	92.1	89.6
DBN	93.3	91.1
IWOA-BP	95.1	92.7

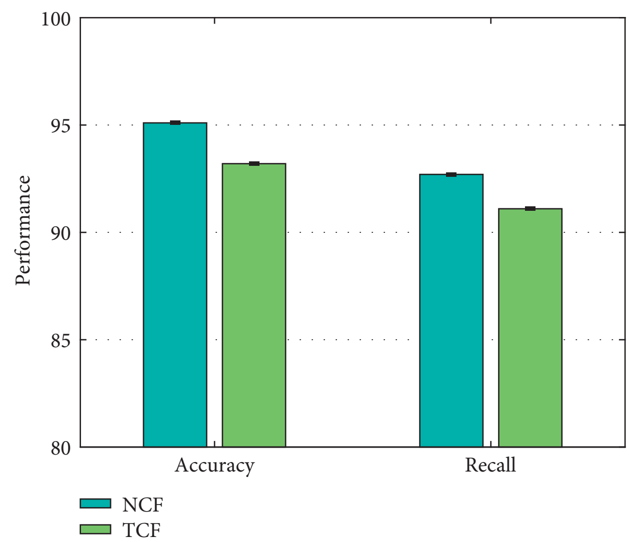


Figure 5: Nonlinear convergence factor analysis.

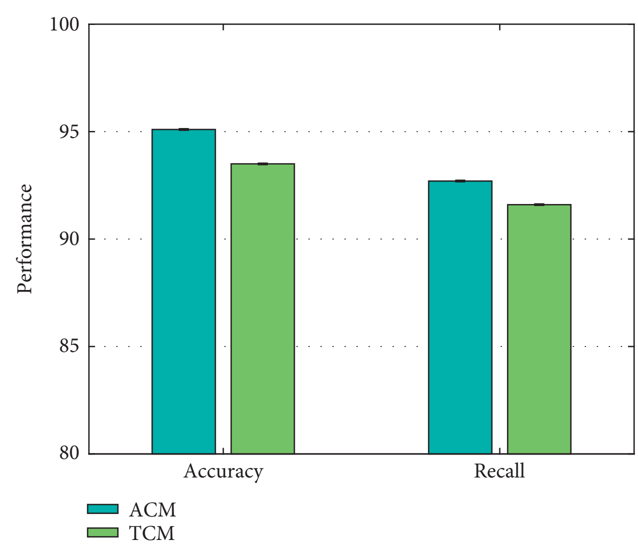


Figure 6: Adaptive crossover and mutation analysis.

Table 3: Comparison of BP, WOA-BP and IWOA-BP.

Method	Accuracy	Recall	Training time
BP	91.9	90.3	297 min
WOA-BP	93.7	91.4	238 min
IWOA-BP	95.1	92.7	189 min

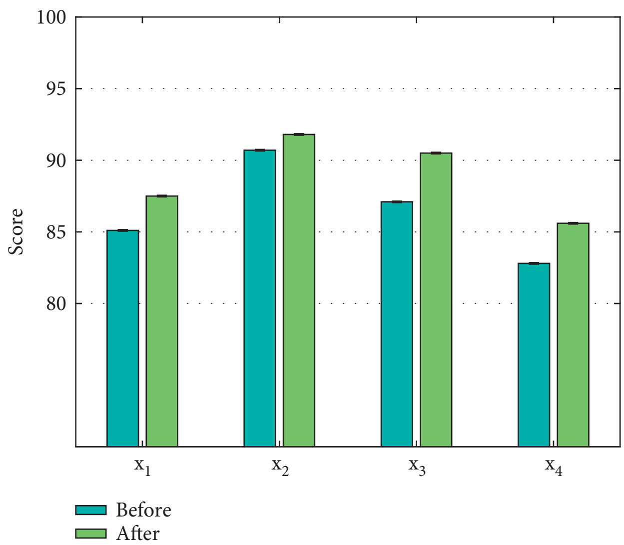


Figure 7: Comparison of scores for the first four indicators.

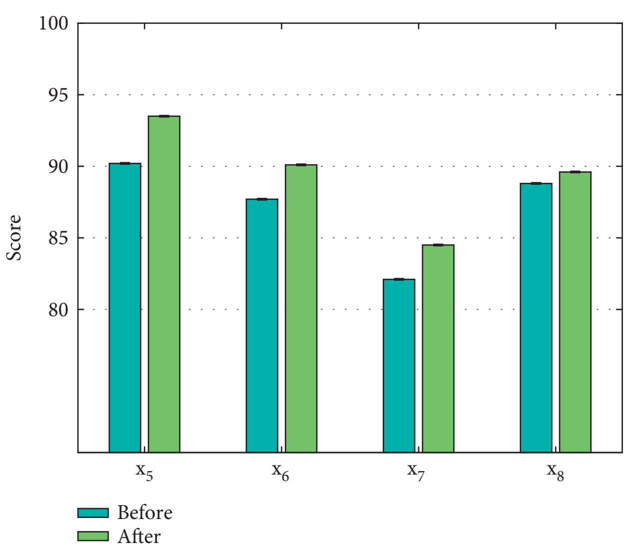


Figure 8: Comparison of scores for the last four indicators.

As illustrated in the comparison data in the table, the performance of the traditional BP network is the lowest, and the training time is the longest. After the introduction of WOA for optimization, the accuracy and recall rate of the network have improved, and the training time has decreased. These performances can be further improved after BP is optimized with the improved IWOA.

4.7. Intelligent Tourism Marketing Promotion Strategy Analysis. This work proposes a series of intelligent tourism marketing promotion strategy, to verify that these strategies can promote tourism revenue, this work compares the tourism efficiency indicators before and after using these strategies, which are the same as those in Table 1. The comparison results are demonstrated in Figure 7 and Figure 8.

According to the data shown in the figure, after using the intelligent tourism marketing and publicity strategy proposed in this work, each indicator of tourism revenue can get a corresponding score improvement. This verifies the feasibility and superiority of the strategy designed in this work.

5. Conclusion

With the improvement of residents' living standards, the demand for tourism is becoming more and more strong, and the tourism industry developed from this has already played a very important role in the national economy. The tourism industry has shown a vigorous development momentum, and the proportion of tourism in the economic development will be further enhanced. The tourism industry is in a period of rapid development, and the level of tourist demand is gradually rising. Information technology is constantly changing and upgrading, and the direction and method of tourism marketing and publicity are in urgent need of research and innovation. The traditional tourism marketing propaganda is mainly based on offline travel agencies, the marketing strategy is to attract tourists with low-cost marketing, the marketing channels and marketing methods are single, and the design of tourism products cannot meet the needs of tourists. With the continuous updating of scientific and technological information in modern society and the rise of emerging technologies such as big data intelligence, the era of intelligence has emerged, which has special social research and application value for intelligent tourism marketing. To promote the high-quality development of the tourism industry, this work studies tourism marketing and publicity methods for revenue enhancement. First, this work proposes a revenue-enhancing tourism marketing and publicity method, which is mainly divided into the use of intelligence to enrich the form of publicity and marketing and the use of intelligence to refine the content of publicity and marketing. Second, this work proposes an IWOA-BP network for evaluating a revenue-enhancing intelligent tourism marketing promotion method. It improves the WOA algorithm by introducing nonlinear convergence factor and adaptive crossover mutations to construct the IWOA algorithm. Then the IWOA algorithm is applied to the initial weights and thresholds of the BP network for optimization, which can solve the drawbacks of the traditional BP network and improve the performance and reduce the training time. Third, this work conducts a comprehensive evaluation of the proposed revenue enhancement-oriented tourism marketing and publicity methods and IWOA-BP, and the experimental results verify the feasibility of these methods.

Data Availability

The datasets used during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] D. Li, "Research on the application of big bata technology in tourism industry," *International Journal of Frontiers in Engineering Technology*, vol. 1, no. 1, 2019.

[2] J. Du, "Research on Intelligent Tourism Information System Based on Data Mining Algorithm," *Mobile Information Systems*, vol. 2021, Article ID 5727788, 10 pages, 2021.

[3] E. Sigalat-Signes, R. Calvo-Palomares, B. Roig-Merino, and I Garcia-Adan, "Transition towards a tourist innovation model: the smart tourism destination," *Journal of Innovation & Knowledge*, vol. 5, no. 2, pp. 96–104, 2020.

[4] X. Wang, F. Zhen, J. Tang, L. Shen, and D. Liu, "Applications, experiences, and challenges of smart tourism development in China," *Journal of Urban Technology*, pp. 1–26, 2021.

[5] D. Ma, J. Hu, and F. Yao, "Big data empowering low-carbon smart tourism study on low-carbon tourism O2O supply chain considering consumer behaviors and corporate altruistic preferences," *Computers & Industrial Engineering*, vol. 153, Article ID 107061, 2021.

[6] O. H. Chi, D. Gursoy, and C. G. Chi, "Tourists' attitudes toward the use of artificially intelligent (AI) devices in tourism service delivery: moderating role of service value seeking," *Journal of Travel Research*, vol. 61, no. 1, pp. 170–185, 2022.

[7] H. Zhang and M. Li, "Integrated design and development of intelligent scenic area rural tourism information service based on hybrid cloud," *Computational and Mathematical Methods in Medicine*, vol. 2022, Article ID 5316304, 9 pages, 2022.

[8] Y. Liu, "Research on the marketing strategy of rural characteristic tourism based on the analysis of big data," *Journal of Physics: Conference Series*, vol. 1744, no. 4, Article ID 042081, 2021.

[9] G. Gupta, *Inclusive Use of Digital Marketing in Tourism industry Information Systems Design and Intelligent Applications*, pp. 411–419, Springer, Singapore, 2019.

[10] C. Halkiopoulos, E. Dimou, A. Kompothrekas, and T. Teywvno, *The E-tour facilitator platform supporting an innovative health tourism marketing strategy*, pp. 609–623, Culture and Tourism in a Smart, Globalized, and Sustainable World. Springer, Cham, 2021.

[11] P. L. Pearce, "Smart tourists and intelligent behaviour," *Handbook of e-Tourism*, pp. 1–17, 2020.

[12] X. Shi, "Tourism culture and demand forecasting based on BP neural network mining algorithms," *Personal and Ubiquitous Computing*, vol. 24, no. 2, pp. 299–308, 2020.

[13] W. Zhou, Y. Liu, J. Wang et al., *Journal of Separation Science*, vol. 43, no. 1, pp. 87–104, 2020.

[14] M. P. Peñarrubia-Zaragoza, M. Simancas-Cruz, and G. Forgione-Martín, "An application of geomarketing to coastal tourism areas," *Tourism & Management Studies*, vol. 15, no. 4, pp. 7–16, 2019.

[15] W. Wang, N. Kumar, J. Chen et al., "Realizing the potential of the internet of things for smart tourism with 5G and AI," *IEEE network*, vol. 34, no. 6, pp. 295–301, 2020.

[16] S. Naraparaju, P. Jalapati, and K. Nara, *Smart poster for tourism promotion through NFC technology*, pp. 507–519, Integrated Intelligent Computing, Communication and Security. Springer, Singapore, 2019.

[17] P. Tavitiyaman, H. Qu, W. s L. Tsang, and CwR. Lam, "The influence of smart tourism applications on perceived destination image and behavioral intention: the moderating role of information search behavior," *Journal of Hospitality and Tourism Management*, vol. 46, pp. 476–487, 2021.

[18] N. Wise and H. Heidari, *Developing smart tourism destinations with the Internet of Things*, pp. 21–29, Big Data and Innovation in Tourism, Travel, and Hospitality. Springer, Singapore, 2019.

[19] C. W. Lu, J. C. Huang, C. Chen, M. H. Shu, C. W. Hsu, and B. Tapas Bapu, "An energy-efficient smart city for sustainable green tourism industry," *Sustainable Energy Technologies and Assessments*, vol. 47, Article ID 101494, 2021.

[20] U. Akdu, "Smart tourism: issues, challenges and opportunities," *The Emerald handbook of ICT in tourism and hospitality*, vol. 1, 2020.

[21] C. M. Hall, *Wine, Food, and Tourism marketing*, Routledge, England, UK, 2013.

[22] G. Büyüközkan and B. Ergün, "Intelligent system applications in electronic tourism," *Expert Systems with Applications*, vol. 38, no. 6, pp. 6586–6598, 2011.

[23] L. Sheehan, A. Vargas Sánchez, A. Presenza, and T. Abbate, "The use of intelligence in tourism destination management: an emerging role for DMOs," *International Journal of Tourism Research*, vol. 18, no. 6, pp. 549–557, 2016.

[24] P. Del Vecchio and G. Passiante, "Is tourism a driver for smart specialization? Evidence from Apulia, an Italian region with a tourism vocation[J]," *Journal of Destination Marketing & Management*, vol. 6, no. 3, pp. 163–165, 2017.

[25] D. Buhalis, "Tourism and information technologies: past, present and future," *Tourism Recreation Research*, vol. 25, no. 1, pp. 41–58, 2000.

[26] J. K. Gerrikagoitia, I. Castander, F. Rebón, and A. Alzua-Sorzabal, "New trends of intelligent E-marketing based on web mining for E-shops," *Procedia-Social and Behavioral Sciences*, vol. 175, pp. 75–83, 2015.

[27] K. Lau, K. Lee, P. Lam, and H. Ying, "Web-site marketing for the travel-and-tourism industry," *Cornell Hotel and Restaurant Administration Quarterly*, vol. 42, no. 6, pp. 55–62, 2001.

[28] K. A. Paskaleva-Shapira, "New paradigms in city tourism management: redefining destination promotion," *Journal of Travel Research*, vol. 46, no. 1, pp. 108–114, 2007.

[29] Y. Guo, H. Liu, and Y. Chai, "The embedding convergence of smart cities and tourism internet of things in China: an advance perspective[," *Advances in Hospitality and Tourism Research*, vol. 2, no. 1, pp. 54–69, 2014.

[30] B. Doolin, L. Burgess, and J. Cooper, "Evaluating the use of the Web for tourism marketing: a case study from New Zealand," *Tourism Management*, vol. 23, no. 5, pp. 557–561, 2002.

[31] B. Seetanah and R. V. Sannassee, "Marketing promotion financing and tourism development: the case of Mauritius," *Journal of Hospitality Marketing & Management*, vol. 24, no. 2, pp. 202–215, 2015.

[32] E. Nel and T. Binns, "Place marketing, tourism promotion, and community based local economic development in post-apartheid South Africa: the case of Still Bay the Bay of Sleeping Beauty," *Urban Affairs Review*, vol. 38, no. 2, pp. 184–208, 2002.

[33] B. Bramwell and L. Rawding, "Tourism marketing images of industrial cities," *Annals of Tourism Research*, vol. 23, no. 1, pp. 201–221, 1996.

[34] K. S. Alkharabsheh, M. N. ALsarayreh, M. A. Rumman, and A. H. Al Farajat, "The impact of viral marketing via internet on promotion of tourism products in Jordan," *International Research Journal of Finance and Economics*, vol. 80, pp. 138–147, 2011.

[35] S. A. K. Pramanik and M. R. H. K. Rakib, *Conceptual analysis on tourism product and service promotion with special*

reference to Bangladesh, pp. 109–126, Tourism Marketing in Bangladesh. Routledge, England, UK, 2020.

[36] Y. S. Lin and J. Y. Huang, "Internet blogs as a tourism marketing medium: a case study," *Journal of Business Research*, vol. 59, no. 10-11, pp. 1201–1205, 2006.

[37] Z. M. Sadq, B. Othman, and R. K. Khorsheed, "The impact of tourism marketing in enhancing competitive capabilities," *African Journal of Hospitality, Tourism and Leisure*, vol. 8, no. 5, pp. 1–11, 2019.

[38] E. Ciriković, "Marketing mix in tourism," *Academic Journal of Interdisciplinary Studies*, vol. 3, no. 2, p. 111, 2014.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 8739097, 10 pages
https://doi.org/10.1155/2022/8739097

Research Article

Research on Construction Cost Control Technology of Construction Project Based on Big Data Analysis Theory

Xiaobing Yu[1] and Hengzhong Zuo[2]

[1]*Hunan Vocational College of Engineering, Hunan 410151, China*
[2]*Changsha University of Science &Technology, Hunan 410076, China*

Correspondence should be addressed to Hengzhong Zuo; 002923@csust.edu.cn

Received 27 June 2022; Revised 27 July 2022; Accepted 6 August 2022; Published 27 August 2022

Academic Editor: R. Mo

The current methods proposed in construction cost control of construction projects have problems such as applicability and low accuracy. Therefore, this paper proposes a method for evaluating engineering cost control based on the big data Analytic Hierarchy Process (AHP) method. Based on the big data technology, the construction cost control index of the construction project is selected, and the construction cost control index model of the construction project is constructed according to the selection results, and the evaluation set is constructed. The corresponding weights of the construction cost control indicators of each construction project are obtained through the analytic hierarchy process, and a weight set is constructed. The membership relationship between the index set and the evaluation set is determined, a fuzzy relationship matrix is constructed, and the construction cost control result of the construction project is obtained according to the calculated weight and the constructed fuzzy relationship matrix. The test results of examples show that the risk assessment accuracy of the proposed method is higher than that of ordinary methods, and the quantitative processing of audit risk assessment is realized.

1. Introduction

With the rapid development of the Internet and cloud computing, all kinds of data have exploded, and mankind has since entered the era of big data [1, 2]. At present, many companies have used big data for business analysis and processing, which is conducive to companies to more accurately specify future development strategies and improve their competitiveness. With the advent of big data, the development of information technology has also had a huge impact on the cost control of construction projects. With the establishment and continuous improvement of the socialist market economic system, construction enterprises have fully implemented the cost control of construction projects. However, how to use big data theory to effectively control construction costs in the construction process of construction projects is a problem that all enterprises and project managers must face.

Scholars at home and abroad have carried out a lot of research on the problem of construction cost control. As early as the beginning of the 14th century, foreign scholars have made preliminary explorations on construction cost control technology. The classical cost control theory has gone through three stages: cost measurement and control, standard cost control, and activity cost control. Early cost control was mainly realized by controlling a single object, ignoring the fact that cost control is actually a complete system [3]. At the end of the 19th century, the scientific management represented by Taylor came into being [4]. After abandoning the traditional experience management, he proposed a relatively complete scientific management method to improve the management quality [5]. Until the beginning of the last century, some scholars put forward the basic concept of standard cost for the first time and implemented it after further improvement [6, 7]. Driven by the trend of the times, in the 1980s, western scholars gradually integrated scientific concepts into cost control theory, thus forming activity-based cost control [8–10]. The above theories have promoted the development of cost control technology in various industries to a certain extent. In recent years, some scholars have also carried out relevant research on construction cost control in the field of construction engineering. Zhou pointed out that the cost control in the construction phase of the construction project is an effective means to reduce the investment consumption of the construction project and control the investment waste in the construction phase [11]. The cost management of construction enterprises is an important part of project cost control in the whole process of project construction. Feng et al. expounded the details and steps of the cost management of construction enterprises in the construction stage. This paper divided the project cost management into three stages: project preparation stage, implementation stage, and completion stage. Construction companies adopt the methods of precontrol, process control, and post-event control, respectively, for the cost management of each stage [12]. Li established a project cost control case database and used the weighted grey correlation method to realize case retrieval, so as to find the best solution for the new project cost management control problem [13]. Zhang et al. (2014) described the connotation and significance of cost control of engineering construction projects, systematically discussed the main procedures and work content of project cost management, and, combined with engineering practice, emphatically analyzed the cost management and cost control measures in the construction stage [14]. Oyegoke et al. identify and analyze the factors influencing the choice of effective cost control techniques for the UK construction industry and assess their importance [15]. Abbas et al. analyzed the unfavorable factors of cost control in Iraqi construction projects by carrying out a questionnaire. The results show that modern construction concepts, methods, and technologies are the main factors restricting the cost control of the project [16]. After analyzing the characteristics and costs of construction projects, Dai et al. constructed a cost early warning system suitable for construction projects based on the earned value management method [17]. Xie et al. put forward a method to control the cost of construction projects by combining artificial intelligence technology and wireless communication after analyzing market fluctuations and national policies in the construction industry [18]. Zhang et al. proposed a dynamic control method of construction cost for coastal engineering projects based on rough set theory, and the results show that the amount of data is the main factor to objectively judge cost control [19]. Korke et al., from the perspective of potential owners of construction assets, finally proposed a technology for construction cost control of construction projects by analyzing the financial feasibility of the investment and analyzing all aspects of cost control [20]. Ai put forward corresponding measures to control the cost of construction projects from the aspects of the preliminary planning of project cost, the control of labor costs in the process of project cost, the guarantee of bidding and procurement of bulk materials, the implementation of the quota system, and the management of turnover materials [21]. The above research has focused on the area of analysis of cost control techniques in engineering construction. However, the current data sources focus on individual construction projects, which results in the proposed cost control techniques not being prevalent in the construction projects concerned. It is also impossible to reasonably guide the project construction and the corresponding project cost control. In addition, most of the above studies are considered from the perspective of cost control in the whole process of the project. Although they can provide guidance for cost control in the whole process of the project, they cannot specifically guide the cost control in the construction stage due to the long project cycle.

In this paper, after collecting and sorting out the current data in the field of construction engineering through big data technology, based on the principles of comparability, independence, operability, and comprehensiveness, the impact indicators of cost control in the construction process of construction projects are selected. After the judgment matrix of the influencing factors of construction cost control, the relative weight of each influencing factor evaluation index is obtained by expert evaluation, and the method of construction cost control of construction engineering is proposed. On this basis, taking a construction project as an example, the feasibility and accuracy of the construction cost control method established based on the big data analysis theory were tested.

2. Principles and Basis of Construction Cost Control of Construction Projects

The construction phase of a construction project is long, and there are many uncertain factors affecting the cost control. For example, due to the strengthening of environmental protection policies, the prices of raw materials such as sand, gravel, and cement have risen sharply in some areas. In addition, when a construction project is undergoing underground construction, the underground situation is relatively complex, and the preliminary survey may not be comprehensive and accurate, resulting in a large number of design changes and engineering visas. If the construction unit does not pay enough attention to project changes and visa work, or the work procedures are not rigorous, it will also bring great uncertainty to the cost control of construction projects. For example, when a visa occurs, the construction unit may exaggerate the amount and scope of the project for the project visa. The supervisors have poor business ability, are not familiar with contracts and related regulations, and blindly sign their certification. Since the construction unit is a follow-up audit, it may not have a good understanding of the actual situation on-site. The above problems all bring hidden dangers to the cost control of construction projects and cause financial investment losses.

2.1. Principles of Construction Cost Control of Construction Projects. The cost control in the construction stage is a dynamic control process that is constantly cyclical. The so-called dynamic cost control of engineering projects is the collection of a large amount of data and information generated by the project manager at any time during the construction stage of the construction project, and the comparison between the actual cost and the target cost is carried out in real time. If cost deviations are found to be outside the permissible range, corrective measures are taken in a timely manner, while, at the same time, projections are made for future phases based on the actual costs of the project. In addition, comparisons are made with target costs, and deviations are constantly identified and corrected, so that actual costs are kept close to cost control targets. Figure 1 shows the schematic diagram of construction cost control of construction projects.

2.2. Basis for Construction Cost Control of Construction Projects. Basis for cost control in construction stage Basis for cost control in construction stage includes the following: (1) engineering contract: the cost control in the construction stage should be based on the project contract, around the goal of reducing the project cost, from both the budget revenue and the actual cost, and strive to tap the potential of increasing revenue and reducing expenditure, in order to obtain the greatest economic benefits. (2) Construction cost plan: the construction cost plan is a cost control plan in the construction stage formulated according to the specific conditions of the construction project. It includes not only the predetermined specific cost control goals, but also the measures and plans to achieve the control goals. It is the guiding document for the cost control in the construction stage. (3) Progress report: the progress report provides important information such as the actual completion of the project at each moment and the actual payment of the cost during the construction stage. The construction cost control work is to compare the actual situation with the cost plan of the construction stage, find out the difference between the two, analyze the reasons for the deviation, and take measures to improve the future work. In addition, the progress report also helps managers discover hidden dangers in project implementation in a timely manner and take effective measures to avoid losses as much as possible before the situation causes major losses. In addition to the above-mentioned main bases for cost control in the construction stage, project changes, related construction organization design, subcontract documents, etc. are also the basis for cost control in the construction stage.

In this paper, under the condition of strictly following the principles and basis of construction cost control proposed above, the hierarchical analysis method based on the big data environment is used to propose the construction cost control method for construction projects. The proposed method can provide reference and guidance for the construction of building projects in a more accurate and reasonable manner.

3. Establishment of Construction Cost Control Objectives and Index System under Big Data Environment

3.1. Objectives of Cost Control. The cost control objective must be clear and scientific, and only the effective control behavior can be guaranteed. If there is no goal, the control will lose its direction; if the goal setting is unscientific, the control behavior will deviate from the normal running track. The ultimate goal of cost control is to reduce costs. The resources (human, financial, and material) owned by any organization are limited, and any economic activity is inseparable from the role of economic laws. However, the role of the law of value prompts people to continuously pursue the least resource consumption, obtain better quality and more quantitative results, and try their best to achieve economic or social benefits that are higher than the average social consumption level. Therefore, it is necessary to strictly control the consumption of resources under the premise of satisfying a certain quality and function. The role of this value law and benefit mechanism is the economic basis for cost control. Use big data analysis theory to analyze various costs that occur in the production and operation process, find unnecessary economic expenditures and improve them, reduce any costs that do not increase the value of the enterprise, and thus reduce costs.

3.2. Data Sources. Based on the data collection method under the big data technology, the construction cost information of related enterprises in the construction industry is collected. The data set includes the construction cost information of the construction engineering field in different provinces in the past 10 years. By analyzing the collected information with the help of big data technology and following the index selection principles described in the following section, the evaluation indicators are organized.

3.3. Principles of Indicator Selection. The establishment of the evaluation index system should be determined according to the actual situation, and the establishment of the evaluation system is also different for different research purposes. Since the functions and costs in value engineering always change at the same time, an appropriate evaluation system should be selected to achieve the optimal matching between the two while quantitatively evaluating the indicators. Therefore, the establishment of the construction cost control index evaluation system for construction projects should follow the following principles: (1) typicality: the index should have a certain typical representation, that is, reflecting the specific industry as accurately as possible. (2) The comprehensiveness of the comprehensive index is aimed at the whole evaluation system. For example, the established index evaluation index should cover all aspects of the construction process related to the construction cost, and on the premise of ensuring that the evaluation index can truly affect the construction cost, the evaluation index should be refined and simplified as far as possible. (3) Systematism means that there must be a certain logical relationship between the indicators. Each subsystem consists of a set of indicators. The indicators are independent and related to each other. (4) Objectivity: the establishment of indicators should respect objective facts, truly and objectively reflect the impact on construction costs of construction projects, minimize the impact of subjective factors on the establishment of indicators, and establish indicators based on the principle of

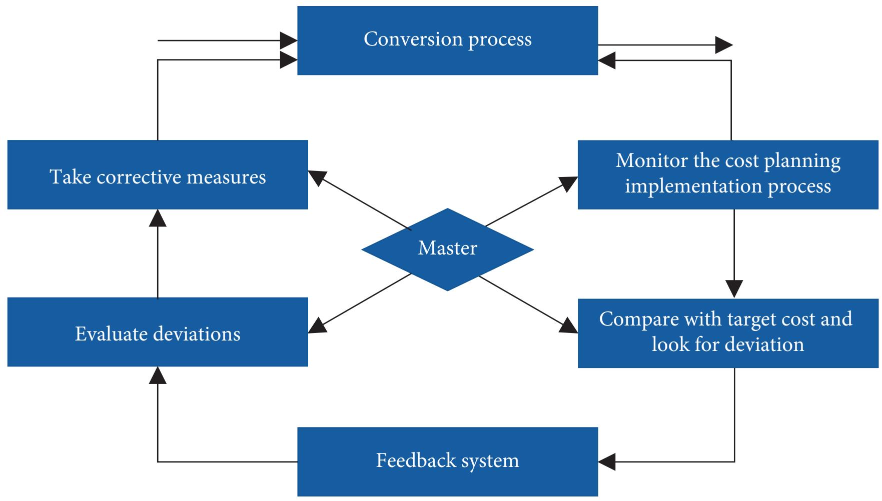


Figure 1: Schematic diagram of construction cost control of construction projects.

objectivity. (5) The establishment of operable and quantifiable indicators should have certain operability and be as simple as possible and easy to use in practice. At the same time, the principle of combining quantitative and qualitative is adopted to quantify each functional index, which is convenient for mathematical calculation and analysis.

3.4. Determination of Evaluation Indicators. The evaluation indicators of construction cost control of construction projects mainly include direct cost control and indirect cost control. The potential costs and importance involved in each factor are as follows: direct costs refer to the expenses that constitute or contribute to the formation of the engineering entity during the construction process and are expenses directly related to the production process. The results of the big data analysis show that direct cost control mainly includes production workers' wages, raw materials, fuel, electricity costs, and depreciation of machinery and equipment, while indirect cost control mainly includes managers' wages, office expenses, travel and transportation costs, labor insurance costs, inspection, and testing costs.

The construction cost control of construction projects is affected by many factors. The importance and influence of each factor are different. If all factors are used as evaluation indicators, unnecessary calculations will be added; if too few factors are used as evaluation indicators, it will be difficult to accurately reflect Reliability and accuracy of construction cost control. Therefore, based on the construction cost control index obtained from big data statistics, the key factors were identified, the main index parameters were extracted, and the evaluation index system as shown in Figure 2 was established. There are 2 first-level indicators, namely, direct cost U_1 and indirect cost U_2. There are 5 second-level indicators: U_1 includes U_{11}~ U_{13}, including labor cost, material cost, and machinery cost. There are 13 three-level indicators in total. U_{11} includes U_{111}~ U_{114}, including time fee, bonus, allowance, and expenses paid under special circumstances, respectively. U_{12} includes U_{121}~ U_{123}, including raw material fee, material transportation fee, and material transportation loss fee, respectively. U_{13} includes U_{131}~ U_{132} for mechanical usage fee and instrument usage fee; U_{21} includes U_{211}~ U_{212} for fuel power fee and other loss fee; U_{22} includes U_{221}~ U_{222} for data acquisition fee and testing and laboratory processing fee, respectively.

4. Establishment of Construction Cost Control System under Big Data Environment

4.1. Building the Hierarchy. Analytic Hierarchy Process (AHP) expresses the problem to be solved by establishing a hierarchical structure, which generally includes the target layer, the criterion layer, and the substandard layer. According to the evaluation index system established in Figure 2, the goal is the construction cost control of construction projects, which includes two first-level indicators U_1~U_2, and three and two second-level indicators under the first-level index, respectively. Similar to the second-level index, the next time contains 4, 3, 2, 2, and 2 tertiary indicators.

4.2. Constructing the Judgment Matrix. Judgment matrix is used to judge the relative importance of elements in the same layer, which is an important step of AHP. A judgment matrix is constructed for the elements of each layer in Figure 2, and the relative importance of the elements of the same layer is judged by the pairwise comparison between the elements. Tables 1–3, respectively, give the judgment matrix of construction cost control and direct and indirect cost control. The evaluation index model corresponding to the construction cost control index of the construction project is shown in Table 4.

4.3. Hierarchical Consistency Test. Calculation of relative weight: based on the feature vector method, the feature vector of the judgment matrix can be obtained:

$$AW = \lambda_{\max} W, \tag{1}$$

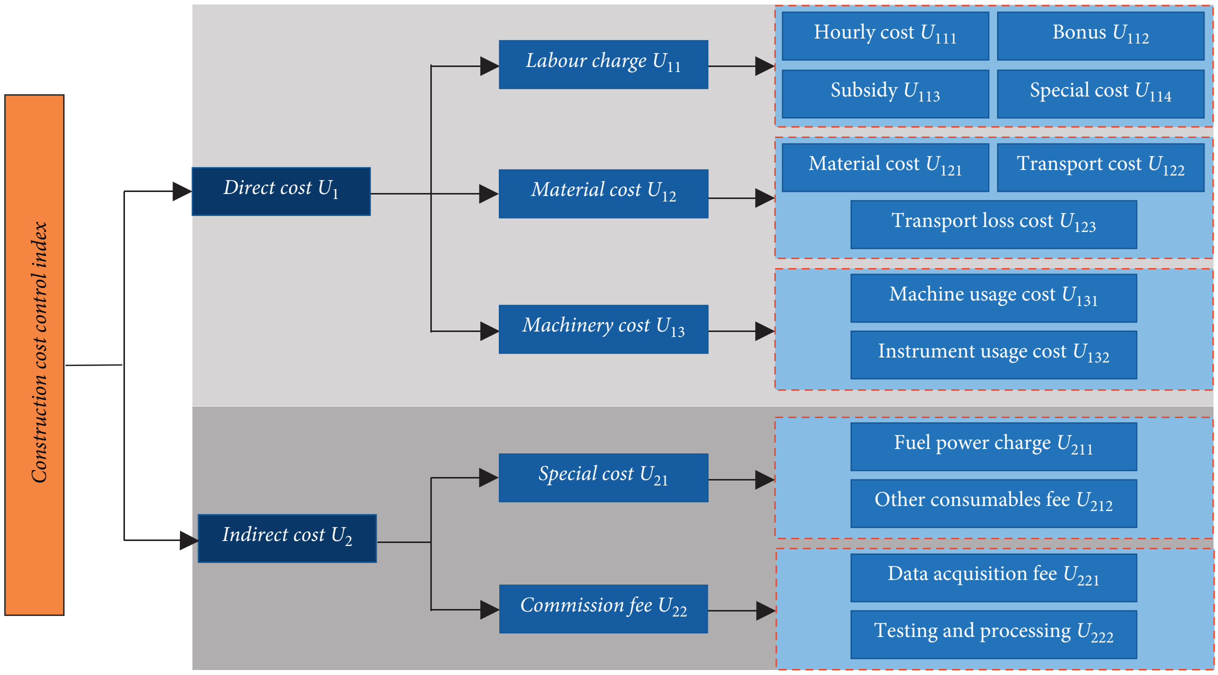


Figure 2: Index system of construction cost control based on big data theory.

Table 1: Construction cost control judgment matrix A.

A	U_1	U_2
U_1	1	a_1
U_2	a_2	1

Table 2: Direct and indirect cost control judgment matrix A_1.

Factor index	A_1	U_{11}	U_{12}	U_{13}
Direct cost U_1	U_{11}	1	a_{12}	a_{13}
	U_{12}	a_{21}	1	a_{23}
	U_{13}	a_{31}	a_{32}	1

Table 3: Indirect cost control judgment matrix A_2.

Factor index	A_2	U_{21}	U_{22}
Indirect cost U_2	U_{21}	1	a_{12}
	U_{22}	a_{21}	1

where $\lambda_{\max}$ denotes the maximum eigenvalue of A. The eigenvector W corresponding to $\lambda_{\max}$ is normalized to obtain the weight vector $\overline{W}$ between elements of the same layer.

When there are more elements to be judged (more than two), it is easy to construct a contradictory judgment matrix. Therefore, Saaty proposed to use the consistency ratio "*CR*" to test the consistency of the judgment matrix [22]. Consistency index "*CI*" is calculated in Formula (2), where n represents the order of judgment matrix:

$$\mathrm{CI} = \frac{\lambda_{\max} - n}{n - 1}. \tag{2}$$

For problems with three or more layers in the hierarchical structure, the relative weights between the elements of each layer from low to high should be calculated first through the judgment matrix, and then the composite weight of the lowest-level elements to the total target should be calculated, and the overall consistency test should be carried out. The consistency test of the total ranking needs to consider the consistency index "*CI*" and the average consistency index *RI* based on the element weights of each layer, and the consistency ratio is

$$\mathrm{CR} = \frac{\sum_{i=1}^{n} w_i \mathrm{CI}_i}{\sum_{i=1}^{n} w_i RI_i}. \tag{3}$$

For the hierarchy established in Figure 3, w_i is the relative weight value of each criterion, and CI_i and RI_i are the consistency index and average consistency index of corresponding judgment matrix $A_1 \sim A_2$.

4.4. Evaluation of Construction Cost Control. The constructed first-level indicator set is the expected audit risk indicator set:

$$U = \{U_1, U_2\}, \tag{4}$$

where U represents the set of construction cost control indicators; U_1 represents the direct cost control indicator of construction; U_2 represents the indirect cost control indicator of construction. The constructed secondary index sets are the direct cost control index set and the indirect cost control index set, of which the direct cost control index set is as follows:

$$U_1 = \{U_{11}, U_{12}, U_{13}\}, \tag{5}$$

Table 4: Importance scale and its meaning in scale method.

Serial number	Importance scale	Implication
1	1	U_i is just as important as U_j.
2	3	U_i is slightly more important than U_j.
3	5	U_i is obviously more important than U_j.
4	7	U_i is more important than U_j.
5	9	U_i is absolutely more important than U_j.
6	2, 4, 6, 8	The median value between two adjacent judgments.
7	Count backward	If the importance value of U_i and U_j is a_{ij}, then the importance ratio of the two is $1/a_{ij}$

where U_{11} to U_{13} represent the secondary indicators corresponding to the direct cost control indicators, respectively. The specific set of indirect cost control indicators is as follows:

$$U_2 = \{U_{21}, U_{22}\}, \tag{6}$$

where U_{21} to U_{22} represent the secondary indicator corresponding to the indirect cost control indicator.

The evaluation set is a set of grades for evaluating each medical accounting audit risk evaluation index. The constructed evaluation set is as follows:

$$V = \{V_1, V_2, V_3, V_4, V_5\}, \tag{7}$$

where V represents the constructed evaluation set; V_1 represents a high cost; V_2 represents a high level of cost; V_3 represents a medium level of cost; V_4 represents a low level of cost; V_5 represents a very low level of cost.

As the influencing factors of construction project cost control are mainly qualitative influencing factors, according to the principle of determining the membership function, this paper adopts the ridge distribution function. Considering that there is no numerical distinction between qualitative indicators, they cannot be quantified. It is necessary to first quantify these indicators and divide them into five grades: very poor, poor, medium, good, and excellent, with assigned grade values of 0.9, 0.7, 0.5, 0.3, and 0.1. Further, according to the principles of the determination of the affiliation function, the trapezoidal function is chosen here.

Aiming at the factor set U of construction cost control indicators, the risk assessment indicators at the same level are compared through the scaling method to obtain relative weights, so as to construct the judgment matrix. From Tables 1 and 2, it can be seen that the specific judgment is as follows:

$$A = [A_1, A_2], \tag{8}$$

where A represents the judgment matrix, and A_1, A_2 represent the judgment matrix elements corresponding to the material misstatement risk index and the inspection risk index. Finally, the eigenvectors of the construction cost control evaluation indicators of each construction project are calculated to obtain their weights, and the consistency test is carried out.

The membership relationship between the index set and the evaluation set is determined, and a fuzzy relationship matrix is constructed. The constructed matrix is as follows:

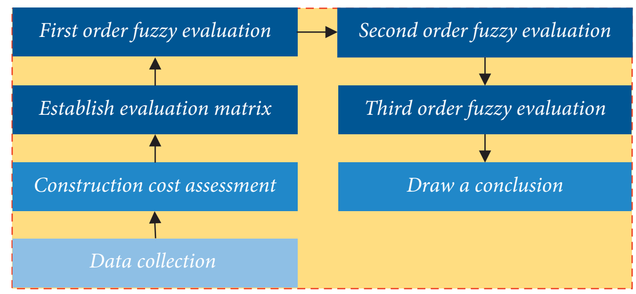


Figure 3: Basic flow of fuzzy hierarchical comprehensive evaluation.

$$R_i = \begin{bmatrix} r_{11} & \cdots & r_{13} \\ \vdots & \ddots & \vdots \\ r_{m1} & \cdots & r_{mn} \end{bmatrix}, \tag{9}$$

where R_i represents the fuzzy relationship matrix; r_{mn} represents the membership degree of the rating corresponding to the evaluation index.

According to the calculated weights and the constructed fuzzy relationship matrix, the medical accounting audit risk assessment results are obtained, as shown in the following formula:

$$B_i = W_i \times R_i, \tag{10}$$

where B_i represents the medical accounting audit risk assessment result; W_i represents the calculated index weight. The evaluation results are normalized, and finally, the results are analyzed and explained by fuzzy distribution method, weighted average method, and maximum membership degree method.

According to the data of construction engineering labor cost, mechanical cost, fuel power cost, and other data obtained based on big data theory, the construction cost of each part is analyzed, and then the fuzzy hierarchical comprehensive evaluation method is used to control and evaluate the construction cost of the entire construction project. The basic process is as follows: shown in Figure 3. Figure 3 shows the basic flow of fuzzy hierarchical comprehensive evaluation.

5. Risk Assessment Case Study

A company invests in the construction of apartment-style houses in a city to solve the living problems of residents in the city. The optimal design scheme is determined through the analytic hierarchy process based on the theory of big data, so as to control the cost. The specific construction plan is a 33-story frame shear wall structure with elevators inside. It is a unit apartment-style residence with three units per ladder. The area of the unit is 70–105 m^2. The exterior windows are made of aluminum alloy windows, and the exterior walls are made of beige paint. The interior is wall-mounted radiator, composite wood floor, etc. The surrounding greenery of the apartment is better, and there are cultural and entertainment facilities. Combined with the actual project, the optimal plan for the construction cost control of the project is obtained through big data analysis. At the same time, in order to highlight the superiority of big data theory in the analytic hierarchy process, the problem is further explained by comparing it with the traditional analytic hierarchy process. Test the accuracy data of construction cost control based on the big data analytic hierarchy process cost control evaluation method and the traditional analytic hierarchy process, and the obtained experimental results are shown in Figure 4. According to the experimental results of evaluation accuracy in Figure 4, the accuracy rate of construction cost control based on big data AHP is basically maintained at 90%, while the accuracy of construction cost control obtained without big data AHP is accurate. The rate remains at 80%. It can be seen that the construction cost control technology based on the big data theory is obviously better.

Figure 5 shows the weight distribution of the above construction cost control indicators in a pie chart. As can be seen from Figure 5, for the construction cost control of construction projects, the direct cost accounts for the largest weight, accounting for 74.3%, and the indirect cost accounts for the second weight, accounting for 25.7%. This shows that, for the construction of the project, both direct and indirect costs have a significant impact on the project construction cost, but direct cost control has a greater impact on project cost control. Therefore, in the actual construction process, more attention should be paid to the direct cost of the project control.

Figure 6 is a weight distribution diagram of direct and indirect cost control in the construction of construction projects. As can be seen from Figure 6(a), for direct cost control, the impact of labor costs has the largest weight, accounting for 63.9%, the mechanical cost has the lowest weight, accounting for 11.4%, and the material cost is in the second place. Between. In general, the weights of the direct cost control indicators of construction projects are ranked as follows: labor wages > material costs > machinery costs. This shows that in terms of direct cost control of construction projects, special attention should be paid to manual deployment during construction, and the work efficiency of personnel in various positions should be improved as much as possible to reduce unnecessary labor costs. As can be seen from Figure 6(b), special expenditures account for the largest proportion of indirect costs, and special attention should be paid to this impact when controlling construction costs. Special expenditures are more important than entrusted expenditures, which shows that when choosing construction cost control, the impact of special expenditures should be fully considered.

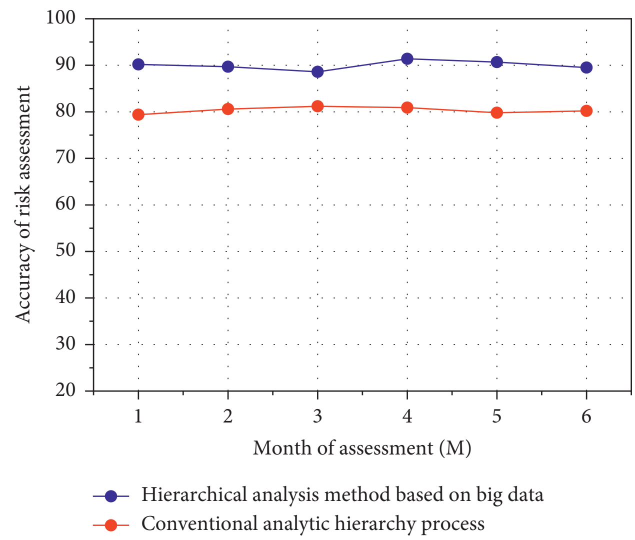


Figure 4: Accuracy test results of construction cost control in construction projects.

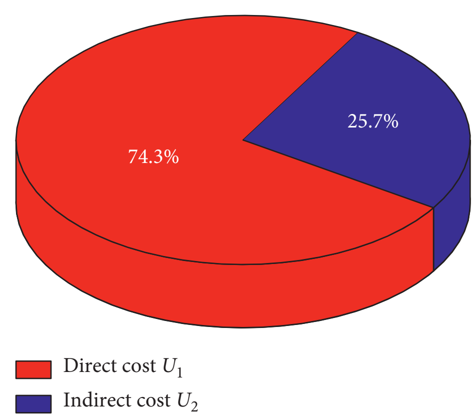


Figure 5: Weight distribution of primary indicators.

Table 5 shows the weight ranking of the three-level indicators of construction cost control on the target layer. It can be seen from the table that the consistency check meets the basic requirements. Among all the three-level indicators, the manual time fee occupies the largest weight, while the test and laboratory processing fee in the commission fee occupies the smallest weight. Therefore, for this project, among all possible construction costs in the comprehensive construction project, the construction cost control can mainly be controlled from labor costs, fuel power costs, etc. The above conclusion will also effectively guide the construction unit to control the construction cost of the project.

Based on the above analysis, direct cost is the main factor restricting the cost control of construction projects. After comprehensively analyzing the relationship between the actual situation of construction projects and construction costs, this paper puts forward the following suggestions for direct cost control of construction projects, in order to better guide project construction cost control.

First of all, according to the construction progress, at the beginning of the month, the labor quantity is reasonably

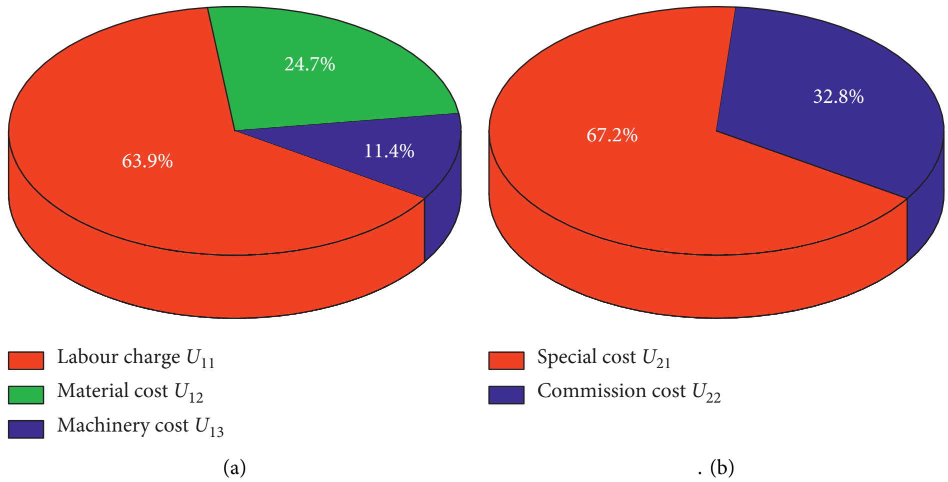


FIGURE 6: Weight distribution of direct and indirect cost control in construction projects. (a) U_1 weight distribution. (b) U_2 weight distribution.

TABLE 5: The weight ranking of the three-level indicators of construction cost control on the target layer.

Second indexes	Third indexes	Weight	Sort order	Consistency ratio
Labour charge U_{11}	U_{111}	0.207	1	0.0851 < 0.1
	U_{112}	0.135	2	
	U_{113}	0.091	5	
	U_{114}	0.041	10	
Material cost U_{12}	U_{121}	0.102	4	
	U_{122}	0.055	8	
	U_{123}	0.027	12	
Machinery cost U_{13}	U_{131}	0.056	7	
	U_{132}	0.029	11	
Special cost U_{21}	U_{211}	0.123	3	
	U_{212}	0.050	9	
Commission fee U_{22}	U_{221}	0.061	6	
	U_{222}	0.023	13	

calculated according to the process and the market labor unit price to calculate the control index for this month. Second, during the construction process, the daily labor quantity is continuously recorded according to the project subitem, and the same project is completed after a subitem is completed. The labor quantity in the quantity list quotation is compared to find out the existing problems, and the corresponding procedures are carried out to correct the control indicators. After completing several project subitems every month, the labor quantity in the engineering quantity list quotation of the respective divisions is compared to assess the completion of the control indicators. Saving labor through this control means reducing labor costs, that is, increasing corresponding benefits.

Secondly, the procurement department will issue the amount of materials required for the construction of this month to the procurement department. On the premise of ensuring the quality of materials, the check and storage link should be strictly control. Secondly, strictly implement the quality procedures and documents to ensure the reasonable stacking of materials and reduce secondary handling in the construction process. The specific operation is based on the progress of the project, and the quota is used. After completing a subitem, the control effect shall be assessed. At the end of the month, the control amount and price shall be compared with the actual quantity to assess the actual effect, whether there is a problem that the same developer calculates the price difference of materials, etc. The focus of materials cost management under the bill of quantities quotation is to keep abreast of market price changes throughout the construction period and to stock up on quality and inexpensive materials in a timely manner. In addition, the availability of materials during the construction process is enhanced by establishing its own material supply base to reduce the rate of material loss during the transportation of materials. The utilization rate extends the service life of the self-owned revolving materials, and the leased revolving materials are accurately calculated according to the construction period and the days of use are accurately returned to the rental unit when they are not needed to reduce the number of occupied days.

Finally, the expenditure of machinery fee is about its control index of the cost of construction products. The number of mechanical console shifts to be used is calculated according to the bill of quantities. Detailed records of the shifts are made every day during the construction process. If there is a power failure on-site that exceeds the time specified in the contract, the on-site visa record should be prepared with the contractor on the same day. The actual shifts will be used at the end of the month. Compare and analyze the reasons for the difference in quantity with the absolute number of control classes. Generally, a lease agreement is adopted for the price of machinery costs. The contract generally does not change during the settlement period. Controlling the actual consumption is the key. According to the on-site situation, the equipment should be arranged reasonably, and the time of entering and leaving the site should be reasonably arranged to make full use of the large equipment to reduce the cost.

6. Conclusion

Based on the existing fuzzy comprehensive evaluation method, aiming at the problem of construction cost control of construction projects, this paper uses the big data theory and introduces the analytic hierarchy process to comprehensively evaluate the construction cost control. The main conclusions are as follows:

(1) Based on the characteristics and principles of construction cost control, through the big data analysis theory, two main influencing factors are proposed when considering construction cost control, that is, direct cost and indirect cost, and various factors for construction cost control are further determined. On this basis, an evaluation index system for construction cost control of construction projects is established.

(2) Compared with the traditional analytic hierarchy process, the accuracy of construction cost control indicators obtained based on big data theory is higher.

(3) For the construction cost of construction projects, the proportion of direct costs is relatively large, and labor costs account for the largest proportion. On the contrary, indirect costs account for a relatively small proportion, and among them, the proportion of testing and laboratory processing fees is the smallest.

Data Availability

The experimental data 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 to report regarding the present study.

Acknowledgments

The study was supported by the project number 2022G03, which is a research on the top-level design of talent training in the context of Natural Resources for Engineering Construction Professional Clusters to Serve the Hunan "three highlands and four new mission" strategy; Innovation and Practice of Talent Training Model For Engineering Construction Specialty Cluster Under the Background of Natural Resources; the project number is GCZD01.

References

[1] R. Nunes Linhares, C. Mario Guimaraes Alcantara, E. Avila Gonçalves, F. Ramos, and M. Jose Loureiro, "Teaching evaluation by teachers from Brazil and Portugal: a comparative analysis," *American Journal of Educational Research*, vol. 5, no. 5, pp. 546–551, 2017.

[2] Z. M. Lv, *The Design of Mathematics Teaching Optimization Scheme Based on Data-Driven Decision-Making Technology Scientific Programming and Artificial Intelligence for Sensor Data Stream Analysis*, vol. 2021, Article ID 5377784, 12 pages, 2021.

[3] H. Derks, "Religion, capitalism and the rise of double-entry bookkeeping," *Accounting Business and Financial History*, vol. 18, no. 2, pp. 187–213, 2008.

[4] V. Antonelli, T. Boyns, and F. Cerbioni, "The development of accounting in europe in the era of scientific management: the Italian enginering conglomerate, ansaldo," *The Accounting Historians Journal*, vol. 35, no. 1, pp. 49–81, 2008.

[5] T. Mclean and T. Tyson, "Standard costs, standard costing and the introduction of scientific management and new technology into the post-second world war sunderland shipbuilding industry," *Accounting Business and Financial History*, vol. 16, no. 3, pp. 389–417, 2006.

[6] M. Matthews, T. Boyns, and J. R. Edwards, "Chandlerian image or mirror image? Managerial and accounting control in the chemical industry: the case of albright & wilson," *Business History*, vol. 45, no. 4, pp. 24–52, 2003.

[7] N. Robson, "Adapting not adopting: 1958 – 74. Accounting and managerial 'reform' in the early NHS," *Accounting Business and Financial History*, vol. 17, no. 3, pp. 445–467, 2007.

[8] R. K. Fleischman, T. Boyns, and T. N. Tyson, "The search for standard costing in the UnitedS tates and britain," *Abacus*, vol. 44, no. 4, pp. 341–376, 2008.

[9] D. P. Ritzwoller, J. G. Michael, and V. M. Michael, "Creating standard cost measures across integrated health caredelivery systems," *Journal of the National Cancer Institute*, vol. 97, pp. 80–87, 2005.

[10] G. Cokins, *Activity-based cost management: an executive's guide*, Wiley, Hoboken, 2007.

[11] S. Zhou, *Effective Cost Control Measures in the Construction Phase of Building*, Value Engineering, 2011.

[12] H. H. Feng and Y. Chen, "Analysis on the cost management of construction enterprises," *Applied Mechanics and Materials*, vol. 71-78, pp. 440–443, 2011.

[13] L. Liu, *Research on Knowledge Management of Construction Cost Control[C]//2015 International Conference on Management Science and Innovative Education*, 2015.

[14] Y. Zhang and S. Y. Zhuang, "Study on the cost control of construction project," *Applied Mechanics and Materials*, vol. 638-640, pp. 2351–2354, 2014.

[15] A. S. Oyegoke, R. Powell, S. Ajayi, G. A. G. R. Godawatte, and T Akenroye, "Factors affecting the selection of effective cost control techniques in the UK construction industry," *Journal of Financial Management of Property and Construction*, vol. 27, no. 2, pp. 141–160, 2021.

[16] N. N. Abbas and A. M. Burhan, "Investigating the causes of poor cost control in Iraqi construction projects," *Engineering,*

Technology & Applied Science Research, vol. 12, no. 1, pp. 8075–8079, 2022.

[17] J. Y. Dai and D. D. Ke, “Cost early-warning model system of large-scale construction project,” *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–10, Article ID 3541803, 2022.

[18] H. D. Xie, Y. G. Ge, and J. Y. Yi, “Cost Control Analysis of construction projects based on wireless communication and artificial intelligence decisions,” *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 8505922, 24 pages, 2022.

[19] D. Q. Zhang and X. B. Liu, “Dynamic control method and application of coastal engineering project construction cost based on rough sets,” *Journal of Coastal Research*, vol. 115, no. 1, pp. 7–9, 2020.

[20] P. M. Korke and G. P. Jarad, *Cost Control in Residential Construction Management*, 2016.

[21] X. Ai, *Analysis and Countermeasures of Project Cost Control in Construction Enterprises*, IEEE, 2011.

[22] T. L. Saaty, *The Analytic Hierarchy Process*, 2001.

[23] A J Berry, “Emerging themes in management control: A review of recent literature British Accounting Review,” *The British Accounting Review*, vol. 41, no. 1, pp. 2–20, 2009.

[24] R. Kaplan and S. cooper R, *cost & effect Using Integrated Cost Systems to Drive Profitability and Performance*, China Renmin University Press Co. LTD, Beijing, 2006.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 3384391, 14 pages
https://doi.org/10.1155/2022/3384391

Research Article

Research on Knowledge Management of Intangible Cultural Heritage Based on Linked Data

Zhaoyan Huang[1] and Tao Xu[2]

[1]*Guilin University of Technology at Nanning, Nanning 530001, Guangxi, China*
[2]*Zhejiang University of Science and Technology, Hangzhou 310023, China*

Correspondence should be addressed to Tao Xu; 114032@zust.edu.cn

Received 28 June 2022; Revised 26 July 2022; Accepted 2 August 2022; Published 27 August 2022

Academic Editor: R. Mo

At present, the protection of intangible cultural heritage has received more and more attention from all levels of society. Intangible cultural heritage is a treasure of national culture. It is an indispensable part of Chinese civilization, the crystallization of the wisdom of Chinese civilization, and represents the country's soft power. The effective organization and management of intangible cultural heritage knowledge is the premise and foundation for the protection, dissemination, and inheritance of intangible cultural heritage. Ontology and linked data technology provide a new method and realization path for the organization and management of intangible cultural heritage knowledge. In this paper, the intangible cultural heritage knowledge is organized reasonably semantically based on the method of linked data, and the purpose is to use the structure of linked data to express the resource data of different structures in a structured manner. This paper first introduces the meaning and background of the research and analyzes the relevant research at home and abroad. Second, it introduces the related knowledge of linked data, analyzes and sorts out the elements and semantic relationship of knowledge in the field of intangible cultural heritage, and designs and constructs the ontology model of intangible cultural heritage knowledge, Finally, based on linked data technology, the process of intangible cultural heritage knowledge organization and linked data set construction is studied, including key steps such as entity to RDF, entity association, linked data storage, and publication. The application of linked data technology in the field of intangible cultural heritage knowledge organization and management can promote the standardization and standardization of intangible cultural heritage knowledge management and is of great significance to the protection and inheritance of my country's intangible cultural heritage culture.

1. Introduction

With the development of society and the changes of the times, my country has made great achievements in both economic and technological development. In the long-term exchanges with countries around the world, the Chinese government has deeply realized the importance of cultural soft power in enhancing the overall national strength and has begun to strengthen the cultural identity of the Chinese people. As an important part of the excellent traditional culture of various ethnic groups, intangible cultural heritage shows the characteristics of the region, contains the wisdom of the group, and is a nonrenewable cultural wealth for every country. Intangible cultural heritage has various manifestations, covers a wide range of fields, and contains rich knowledge. Structural representation and semantic organization of knowledge in the field of intangible cultural heritage is an inevitable requirement for the protection of intangible cultural heritage and knowledge dissemination in the era of knowledge. However, in the management of intangible cultural heritage knowledge, the China Intangible Cultural Heritage website and the intangible cultural heritage dissemination platforms of various provinces and cities only display the information of intangible cultural heritage items and representative inheritors in the form of a list, lacking knowledge of projects and inheritors, projects and regions, and so on. The lack of effective organization and management of knowledge in the field of intangible cultural heritage results in the fragmentation of information, which greatly limits the integrity and dissemination of knowledge. At the same time, there is a lack of professional knowledge-based linked data sets in the vertical field of intangible cultural heritage [1], and a mature intangible cultural heritage encyclopedic knowledge base has not yet been formed. Although some achievements have been made in the research of Chinese encyclopedia knowledge map in the general field, in the field of Chinese knowledge encyclopedia, there is a lack of large-scale encyclopedic knowledge data sets such as DBpedia and Freebase. Since the data in the field of intangible cultural heritage are still dominated by semi-structured data and basic databases, there are many fragmented data, and there is a lack of high-quality open data sets for intangible cultural heritage; the single linear organization of intangible cultural heritage knowledge cannot reflect the diversity of intangible cultural heritage culture. characteristics, which cannot meet the needs of multidimensional disclosure and knowledge retrieval. This limits scientific research on intangible cultural heritage as well as the dissemination of intangible cultural heritage knowledge and the inheritance of intangible cultural heritage.

Foreign research on linked data is early. In January 2007, the Open Linked Data Project was launched, and the research on linked data began. In May 2010, the World Wide Web Consortium (W3C) established the Linked Data Incubation Group in the field of libraries. The main purpose is to help libraries use their data information to establish linked data, realize the effective organization and utilization of resources, and provide users with more comprehensive information. Intelligent retrieval services improve the relevance and interoperability of digital resources among libraries. In 2008, the Swedish National Library actively participated in the development of linked data, organized the Swedish Union Catalogue (UBRIS), and released it in the form of linked data. During the publishing process, the existing vocabulary was reused, which contains vocabularies of library domains such as DC, FOAF, and SKOS [2]. With the development of the Semantic Web, some domestic researchers have introduced ontology and linked data theories and methods into the field of intangible cultural heritage knowledge organization and knowledge management and have carried out a lot of research on ontology construction, resource integration, and resource organization. In terms of knowledge structure analysis and ontology construction research in the field of intangible cultural heritage, Lu et al. analyzed various elements of intangible cultural heritage from the perspective of systems theory and constructed a series of intangible cultural heritage items, people, things, events, documents, and so on. The ontology conceptual model is composed of core concepts and formulated metadata standards and specifications for intangible cultural heritage resources [3]. Huang et al. analyzed the difficulties in the construction of intangible cultural heritage knowledge ontology and took folk dance as an example to design an intangible cultural heritage knowledge ontology construction system for text and multimedia data [4]. Zhou et al. took the ontology method as a research perspective and designed a framework for the organization and retrieval of intangible cultural heritage information resources from the perspectives of ontology representation, semantic organization, and semantic retrieval and realized the ontology of intangible cultural heritage information resources by taking drama intangible cultural heritage projects as an example [5]. Ou and Tang proposed to divide natural language questions into simple sentences involving one data set and complex sentences involving multiple data sets in the study of automatic question-answering in the face of multiple RDF data sets and convert them into structured questions. The SPARQL query, and then integrate the answer, through experiments prove that the answer is more accurate [6]. Li and Zhao proposed an intangible cultural heritage archives resource development model based on linked data [7]. Dong used linked data technology to organize intangible cultural heritage knowledge semantically and realized the semantic disclosure and organization of intangible cultural heritage resources and their relationships [8].

The traditional knowledge management method of intangible cultural heritage projects is mainly classified and filed according to the subject, region, and level, unable to fully express and reveal the relationship between knowledge elements. With the deepening of knowledge organization and knowledge management research, the management of knowledge in the field of intangible cultural heritage has developed into new organizational methods such as topic maps, knowledge maps, and semantic organization, and more emphasis has been placed on the multidimensional disclosure of the relationship between domain knowledge [9]. In particular, the ontology and linked data technology in the Semantic Web technology stack provide new solutions for the semantic organization of knowledge in the field of intangible cultural heritage. Linked data is regarded as a lightweight implementation of the Semantic Web, which aggregates heterogeneous information resources in the form of linked open data, which can improve the visibility, sharing, and openness of resources and knowledge. The organization and management of intangible cultural heritage knowledge based on linked data are to express the knowledge elements and their attributes in the field of intangible cultural heritage in a structured and formalized manner under the guidance of a standardized and unified domain ontology model and to build the semantic relationship between the knowledge elements, then achieve the purpose of knowledge semantics and order, and provide open data acquisition and knowledge services.

Based on this background, because of the effective management and knowledge association of intangible cultural heritage knowledge, this paper proposes a method system for intangible cultural heritage knowledge organization and intangible cultural heritage knowledge association database construction based on linked data, including the design of intangible cultural heritage knowledge ontology model to the construction of linked data, the whole process of storage and publishing; research the knowledge association between the internally linked data sets in the field of intangible cultural heritage and the external knowledge databases of DBpedia and Geo Names; and establish an intangible cultural heritage knowledge related database with rich semantic relationships and knowledge interconnection and sharing, providing domain knowledge. Orderly organization and visualization of knowledge: It is helpful for the excavation and discovery of intangible cultural heritage knowledge as well as the international dissemination of my country's intangible cultural heritage culture.

2. Research Methods

2.1. Concepts and Principles of Linked Data. The concept of linked data comes from the "Father of the Internet," Tim Bemers Lee, who proposed the Web of Data when he analyzed the development and evolution of the Web in "Design Issues for the World Wide Web" published in 2006. In 2007, linked data was formally proposed in the Linked Open Data Project submitted by Chris Bizer and Richard Cyganiak to W3C. Different scholars have put forward their views on what is linked data. Tim Bemers Lee believes that linked data defines a URI specification that enables people to directly obtain digital information resources through HTTP and URI mechanisms. Wikipedia recommends linked data as a best practice for publishing, sharing, and linking all kinds of data, information, and knowledge on the Semantic Web using URI and RDF. It can be seen from this that linked data is a description specification, which describes data in the triple RDF format of "subject–predicate–object." URI determines the uniqueness and relevancy of resources in the data. Link association builds a new data network to realize intelligent applications. Because linked data solves the problem of linking distributed and heterogeneous data in the form of semantic association, it is recommended by W3C as the "Best Practice of the Semantic Web." Linked data has realized the extensive linking of data in different fields in the network and promoted the sharing and widespread use of various types of data. Users can quickly discover a lot of relevant data through data links. Figure 1 shows the situation of the Linking Open Data Cloud Diagram (LOD) as of March 2019. The data set already contains 1239 data sets and 16,147 links. Each node in the figure represents the data published by each data set. The lines between them represent the relationship between the data sets.

Tim Bemers Lee pointed out that the principles of linked data [10] mainly include four aspects: the first principle is to use URI as the name of the resource to identify things. All resources on the World Wide Web are assigned a unique URI as the resource identifier, which stipulates that everything in the World Wide Web uses a URI as a unique identifier, which is the primary condition of the Semantic Web. Using URI to standardize the name of things can avoid ambiguity and confusion; the second principle is to use HTTP URI to locate and find corresponding resources, that is, all URIs can be accessed and retrieved. It is stipulated to use HTTP URI as the link specification, which is conducive to discovering the required information resources according to the data link; the third principle is that when the URI of the resource is accessed, the information related to the resource is provided, and unified standard to represent this information, people can view data information and other classes and attributes through URI; the fourth principle is to provide relevant URI as much as possible to help people discover and obtain more information with potential use value and promote the sharing and utilization of information resources, truly realize the globalization of information. These four principles are important indicators of the semantic association of digital information resources on the World Wide Web and are where linked data fully reflects its ability to solve the problem of distributed and heterogeneous digital information resources on the Semantic Web. If these four principles are not fully satisfied, they may also play a certain role in data association, but it will greatly reduce the potential of data association and reduce the value-added benefits of data aggregation.

2.2. Related Technologies of Linked Data. Linked data can reflect the objective entities such as data and concepts and the relationship between them, which is an important factor for it as an effective way to achieve semantic relationships on the Semantic Web.

A URI is a string used to identify a resource, allowing users to interact with any resource through a specific protocol. Linked data is based on Web technology, mainly involving HTTP, HTML, and URI. HTTP is a hypertext transfer protocol, which is a standard for client and server requests and responses. HTML is a hypertext markup language, which is stored on the server. Web page files: HTML uses markup symbols to identify and standardize various parts of web page content; URI is a uniform resource locator, which realizes the unified positioning of network resources at the access address; and associated data further define and expand these three technologies, using URI at the same time. To solve the problem of naming and positioning [11], at the same time, the biggest feature of URI is that the identification is stable, the resource path is regarded as a part of the resource name, and it is not allowed to change at will so that the resource does not have information dislocation due to the change of the attribute, which is helpful to Stable Links to Achieve Semantic Associations in Dynamically Changing Network Environments. URI is the most critical technology for linked data. Using URI, anything in the World Wide Web can be given a unique identification name, and "anything" is different, mainly divided into information resources and noninformation resources. Information resources refer to the information resources on the Internet. Digital resources that can be found, such as pictures, videos, web pages, audio, and so on, exist in various physical objects outside the network, including nature, human society, and human meaning, such as mountains, rivers, people, and so on. The fourth principle of linked data refers to the use of relevant URI as much as possible to realize the RDF description of resources. It can be seen that URI can also be used for URI reuse while ensuring the uniqueness of resource identifiers. That is if a resource has been identified by authoritative URI and other data creators, the URI identifier of the resource can be used to ensure the uniqueness of the resource identifier. At present, institutions such as the Library of Congress and W3C provide terminology services and representative use cases based on linked data. At the same time, authoritative vocabulary sets such as

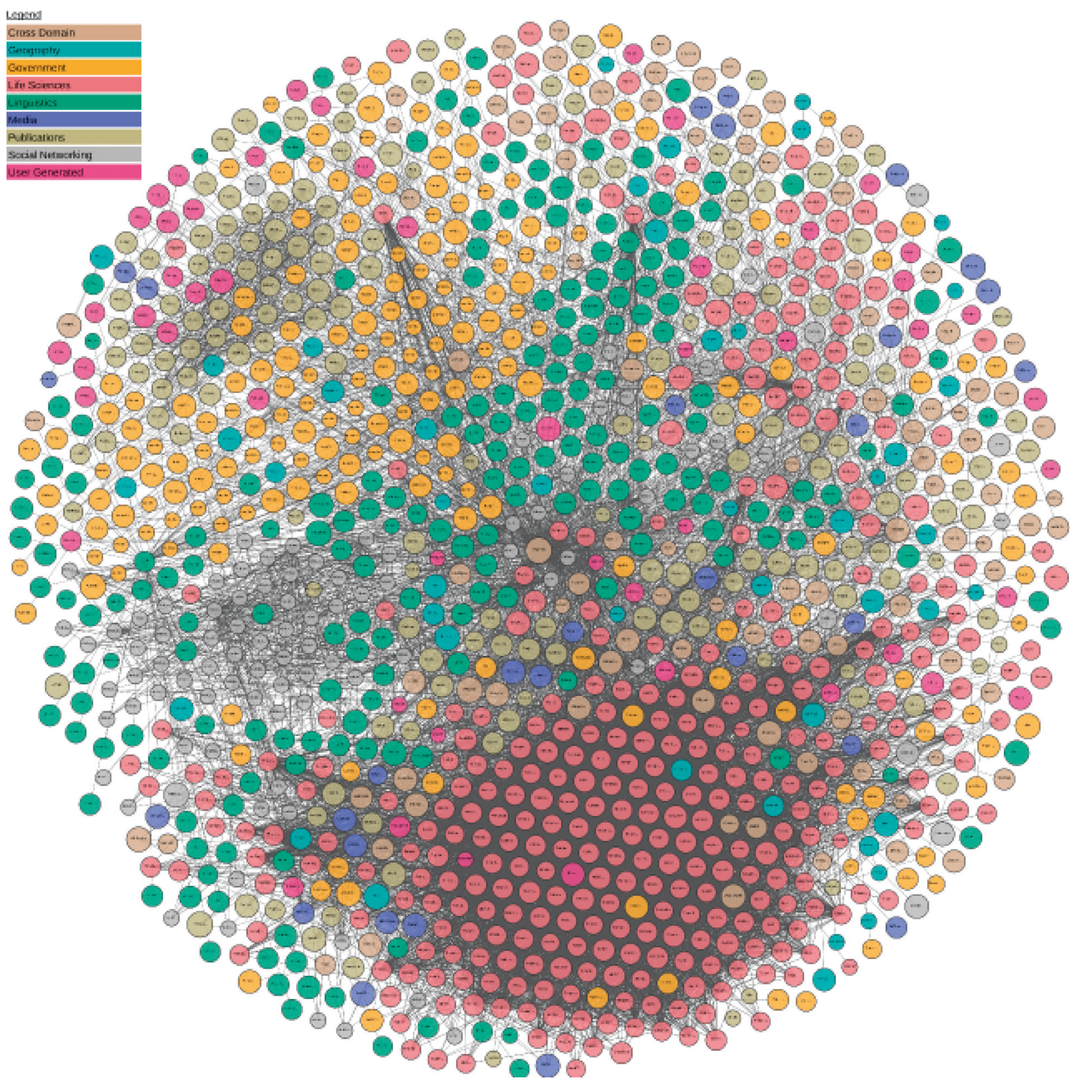


Figure 1: The linking open data cloud diagram.

DBpedia and FOAF also provide conceptual terms and associations for linked data. Therefore, the use of URI can effectively and uniformly identify resources, which not only has a clear indication function for resources but also ensures effective management of resources, which is conducive to identifying distributed and heterogeneous related resources in data association and realizing the spanning of linked data. Purpose of development: URI simply creates a unique identification name and ensures the stability of identification, thus enabling information resources to form a stable associative aggregation in the World Wide Web [12].

RDF is a resource description framework, which provides a way to flexibly describe diverse network resources. It is a markup language used to describe resources on the Web. It is a basic semantic format that can be understood by machines. RDF data can realize data are exchanged between computers with different types of operating systems and application languages so that each data in the network can be shared and utilized to the greatest extent. RDF is to describe resources by using the triple pattern of the "subject–predicate–object" structure under the condition of strictly following the network structure, in which the resource is the subject, the attribute of the resource is the predicate, and the attribute value is the object. The resource is defined by URI, and the predicate represents the relationship between subject and object, which is also represented by URI. Data can be clearly and accurately described through RDF. The resources in RDF contain links identified by URI, which can link to other related resources. At the same time, these links contain semantic relationships, indicating the relationship between resources. Data can be composed of many pieces of RDF, described by a resource identified by URI. The object of RDF can be a numerical attribute or an object attribute, that is, the resource identified by URI; the predicate between the subject and the object indicates the relationship between subject and object and can also be mark with URI. These URIs can come from normative vocabularies, such as FOAF, DC, SKOS, and so on.

SPARQL is a query language and data acquisition protocol especially developed for RDF. Using SPARQL, all information resources represented by RDF can be retrieved and queried on the network. The ultimate goal of SPARQL is to retrieve the Semantic Web in the same way that SQL retrieves relational databases. SPARQL can query RDF data between different data sources. The data source can be the data in the RDF format of the entity, or it can be the data in the virtual RDF format through the middleware. The SPARQL language can query based on graphs, and its query results can also be returned in the form of RDF graphs or data sets. An SPARQL query statement can be composed of five parts: statement, query form and result set, data set, graph schema, and result decoration. The relevant query statements are shown in Table 1.

2.3. Technical Feasibility Analysis. The core of linked data is to name data resources with URIs under the premise of following the HTTP protocol and organize and standardize data resources into the RDF format of the "subject–predicate–object" relationship to reflect the relationship between attributes and attribute values, and finally use HTTP URI to locate and query each data resource. The application development of linked data has promoted the generation of tools for constructing and publishing linked data, which mainly include three types.

2.3.1. Relational Database RDF Conversion Tool. The database is the main place for storing resources. The database is a data collection that organizes, stores, and manages data according to a certain data structure. The independence of data is conducive to the centralized control of data, provides convenience for massive resource storage, and promotes the combination of linked data and resources. Databases can be divided into hierarchical databases, network databases, and relational databases. The relational database is based on a relational model that reflects the relationship between entities and entities, which meets the requirements of relational data mining for associations between data resources. At present, the commonly used relational database management system is MySQL, which saves data in different data tables instead of storing all data in one database. Each data table has a primary key attribute that is not a null value. Both the key attribute and the primary key have a certain limited association relationship. Each data table associates the data attribute values in different data tables through the invocation of the primary key and the foreign key. Using MySQL can easily call the data of each database; even if each database is updated, it will not affect the values of other databases. Based on determining the relationship between data resources, data resources need to be processed, including the creation and publication of RDF [13]. D2R is a tool for publishing relational databases into linked data. It mainly includes three parts: D2RQ Mapping language, D2RQ Engine, and D2R Server. The main function of the D2RQ Mapping language is to convert relational data into mapping rules in RDF format. D2RQ Engine uses the D2RQ mapping file to convert the data in the relational database into RDF format. It does not convert the relational database into real RDF data but maps it into a virtual RDF format. This file is used to access the relational database. When generating data, the query language SPARQL of RDF data is converted into SQL language, and the SQL query result is converted into RDF format or SPARQL query result. D2R Server provides a query access interface for RDF data, which is convenient for SPARQL clients and traditional HTML browsers. The operation flow of D2R is shown in Figure 2.

2.3.2. Linked Data Tools That Directly Generate RDF Data. Such tools mainly include Virtuoso Universal Server and Sparq Plug. Virtuoso Universal Server is a commercial-grade linked data tool that enables XML, Web server, and network. The ideal carrier between them can use the SPARQL side to convert the data into RDF format. Sparq Plug uses SPARQL query language and HTML DOM to convert traditional HTML data into RDF form.

2.3.3. Other Linked Data Tools That Publish RDF Data. Pubby is a front-end linked data for SPARQL that provides a linked data interface for resources in RDF format. Pubby is used to connect the linked data interface and the SPARQL side. It can provide a linked data interface to connect to a local or remote SPARQL protocol server. At the same time, it can also provide simple HTML calls to the database to display available resources. Its working principle is shown in Figure 3.

In RDF, resources are identified by URI, but in most SPARQL data sets, URI cannot be directly referenced, that is, they cannot be accessed directly in a web browser. At this time, by establishing a Pubby server on the SPARQL side, the method of URI mapping is adopted to obtain the original URI information by connecting to the SPARQL terminal and returning the result to the client so that the user can obtain the URI information that can be used directly.

2.4. Design and Construction of Intangible Cultural Heritage Knowledge Ontology Model. Ontology is regarded as a clear formal specification of shared conceptual models. In the field of information science and computer, ontology can be regarded as a model, which is a formal expression of objectively existing objects or concepts, their attributes, and related relationships. For the effective organization and management of knowledge in the field of intangible cultural heritage, it is first necessary to clarify the structure of intangible cultural heritage knowledge, its constituent elements, and internal relationships. Then, based on referring to the international general ontology model, a knowledge ontology model in the field of intangible cultural heritage is established according to the knowledge characteristics in the field of intangible cultural heritage. Intangible cultural heritage knowledge ontology is a formal conceptual model formed by a highly abstract summary of the intangible cultural heritage connotation and its constituent elements [14]. Based on ontology theory and existing research results, according to the ideas and steps of domain ontology design and construction, and by analyzing the knowledge structure and constituent elements of intangible cultural heritage projects, this paper constructs an intangible cultural heritage

Table 1: SPARQL-related query statements.

Category	Check for phrases
For a known data set	@prefix dc: <http://purl.org/dc/elements/1.1> _:a dc : title "Linked Data" _:b dc : title "Resource Description Framework"
A SPARQL query to query all titles against a data set:	Statement : PREFIX dc:<http://purl.org/dc/elements/1.1> Query Mode : SELECT ?title Data set : FORM < http://example.org/knowledge/semantic_web> Graph Mode : WHERE{?x dc : title ?title} Result Modification : ORDER BY ?title

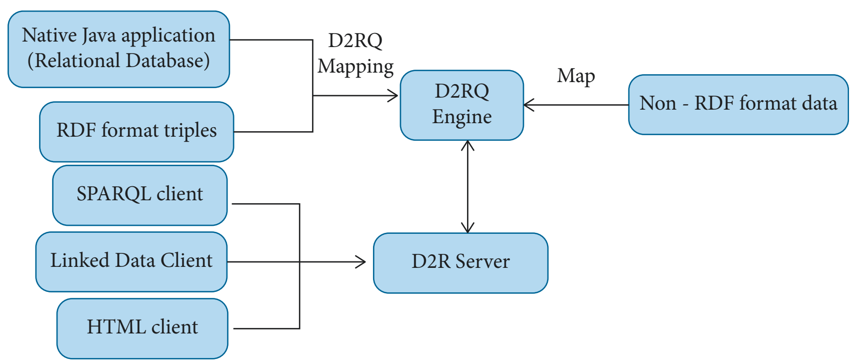


Figure 2: D2R operation flow chart.

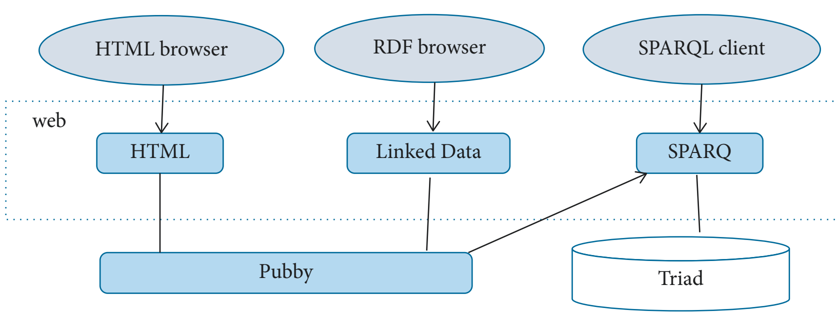


Figure 3: How Pubby works.

knowledge ontology model. Description and knowledge association provide a unified and standardized knowledge representation model and data model. At present, for the protection and inheritance of intangible cultural heritage culture in my country, a four-level intangible cultural heritage protection system of "national–province–city–county" has been formulated. Therefore, the basic knowledge about intangible cultural heritage can be regarded as composed of intangible cultural heritage items, inheritors, relevant institutions, project types, geographical locations, and other elements. [15]. Intangible cultural heritage project (ICH Project) is an abstraction of intangible cultural heritage projects, and its example refers to each specific project in the four-level list of intangible cultural heritage protection established by my country. An instance of an intangible cultural heritage item is a composite object, which not only has its connotative attributes but also includes related entities such as inheritors and regions.

Based on analyzing the basic structure and relationship of intangible cultural heritage knowledge, the ontology model of intangible cultural heritage knowledge designed and constructed in this paper is shown in Figure 4. The ontology model reference draws on the ontology models such as CIDOC CRM, FOAF, Geo Names, person relationship vocabulary (Relationship), and the Dublin Core (DC) metadata standard [16]. According to the core elements of knowledge in the field of intangible cultural heritage, the knowledge ontology in the field of intangible cultural heritage is abstracted into four core categories: intangible cultural heritage project (ICH project), representative inheritor (Person), geographic location (Place), and project type (Category). Each core class defines corresponding data properties, and the relationship between entities is described and revealed through Object Properties. The intangible cultural heritage knowledge ontology model provides a macrounderstanding of knowledge concepts and their relationships in the field of intangible cultural heritage,

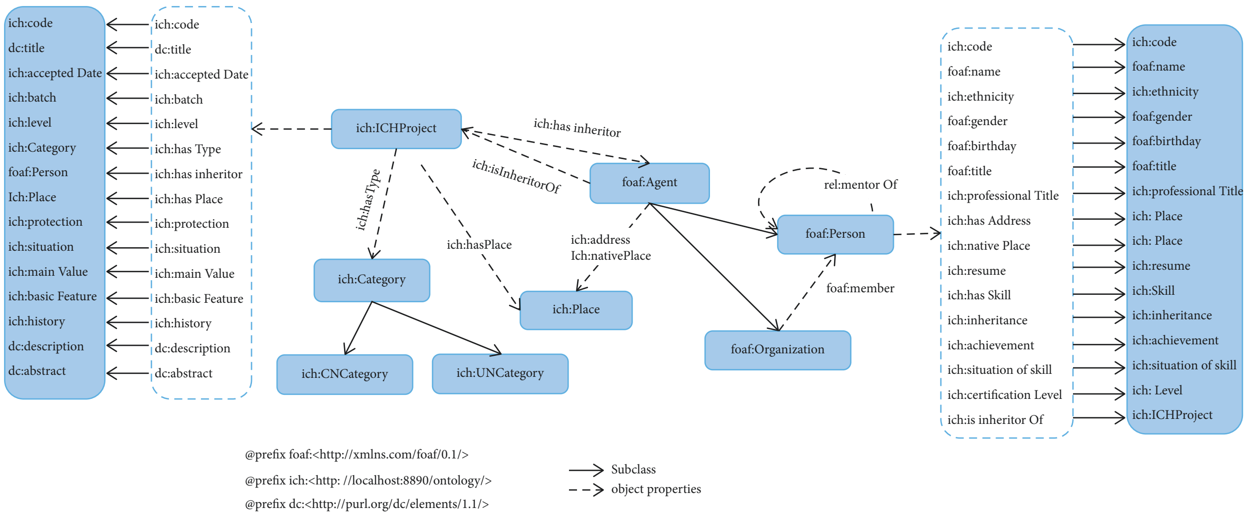


Figure 4: Intangible cultural heritage knowledge ontology model.

standardizes and unifies basic terms and relationships in the field of intangible cultural heritage, and accurately describes the internal relationship between knowledge concepts.

Item type category (Category) is used to construct the classification system of intangible cultural heritage items; by using the object attributes (skos: broader, skos : narrowe, etc.) that represent the relationship between the upper and lower levels of the concept, a multiperspective, multilevel, scalable intangible cultural heritage item classification system. Representative inheritors of intangible cultural heritage are an important part of intangible cultural heritage. The definition of the agent class (Agent) in the ontology model defined in this paper reuses the FOAF ontology model, and the agent class can be divided into two subcategories: "inheritance individual" and "organization and institution"; "individual" mainly refers to certified national. The representative inheritors at the provincial and municipal levels are individuals, and "organizations and institutions" refer to the declaration unit of some intangible cultural heritage items. For example, the declaration unit of the 24 solar terms in the lunar calendar is the China Agricultural Museum. For the representative inheritor of an intangible cultural heritage item, its attributes include the inheritor's number, name, gender, title, ethnicity, date of birth, place of origin, and other basic information, as well as the skills and skills they have acquired. In addition, the attributes in the relationship vocabulary (Relationship) are reused to better express the intricate inheritance relationship and inheritance lineage between inheritors. The geographic location class is defined in the ontology model. On the one hand, it expresses the geographical space of the distribution and circulation of intangible cultural heritage items, and on the other hand, it records information such as the place of residence and place of origin of the inheritor. And the geographic location class is given the data attributes of the administrative level of provinces, cities, counties, villages, and towns, which corresponds to the administrative divisions of our country. At the same time, associate each geographic location instance with the geographic database Geo Names to obtain information such as geographic location introduction, latitude, and longitude. Combined with geographic information system (GIS) technology, not only the spatial distribution of intangible cultural heritage can be expressed in the form of intuitive visualization of the spatial distribution of maps but also the hidden information of the spatial dimension in intangible cultural heritage can be excavated from a deep level. Using GIS spatial analysis techniques to analyze the spatial structure, evolution, and characteristics of intangible cultural heritage is of great significance for understanding the connotation of intangible cultural heritage [17].

The relationship between entities connects independent knowledge elements to form an intangible cultural heritage knowledge network, changing the single-clue model of traditional knowledge organization. In addition, based on these associations, by defining rules and relational reasoning, invisible knowledge can be inferred and discovered. Focusing on the organization and management of knowledge related to intangible cultural heritage projects, this article refers to the international common ontology models such as CIDOC CRM, FOAF, Relationship, and so on, and carries out the reuse and custom extension of an ontology according to the characteristics of intangible cultural heritage knowledge. The intangible cultural heritage knowledge ontology model is oriented to the organization and management of intangible cultural heritage knowledge, providing a formal representation of knowledge in the field of intangible cultural heritage, and meeting the knowledge management needs in the process of declaration and certification of intangible cultural heritage projects. The main function of building an ontology model of intangible cultural heritage is to standardize the description and formal expression of intangible cultural heritage knowledge, while the semantic transformation of data and the structured representation and storage of knowledge need to be implemented with the help of linked data technology.

3. Results and Discussion

3.1. Construction of Intangible Cultural Heritage Knowledge Association Data Set. The construction of the intangible cultural heritage knowledge association data set is a huge and systematic project. First, clarify the research scope and sort out the knowledge objects in the field of intangible cultural heritage; second, extract the domain entities, clarify the various attributes of the entities, build the domain ontology model, and form thesaurus and glossary; third, according to the ontology model, the entities are RDF attribute description, establish entity links, including entity links between internal entities and external open data; finally, select an appropriate data storage and publishing platform to provide access and data interfaces for humans and machines. Following the basic principles of linked data, the construction of a knowledge-linked database in the field of intangible cultural heritage can be divided into five key steps, namely data modeling, entity naming, entity RDFization, entity association, and entity publishing [18]. Among them, the data modeling process is the construction process of the knowledge ontology model in the field of intangible cultural heritage. This paper takes the Hubei Province intangible cultural heritage project as an example to explore the construction steps and specific implementation methods of the intangible cultural heritage knowledge association data set. The technical framework of the intangible cultural heritage knowledge association data set construction is shown in Figure 5.

3.2. Entity to RDF. Due to a large number of intangible cultural heritage projects in my country, and the dynamic changes in the identification of intangible cultural heritage projects and inheritors, the construction process of the entire knowledge base in the field of intangible cultural heritage needs to be carried out in layers and batches. There are two main sources of data. One is the data of intangible cultural heritage project declaration and representative inheritor certification application. The collected data are preprocessed such as data cleaning and stored in the database. Many traditional intangible cultural heritage information systems mainly use relational databases for data storage. This paper chooses to perform data semantic mapping based on D2RQ to convert the content in relational databases into linked data. Semantic mapping of data is to convert the two-dimensional table structure into associated data that is better at dealing with complex relationships and richer in semantic information; it specifically includes data table-to-class mapping, column-to-attribute mapping in data tables, and table-to-association mapping. In the mapping language, d2rq: Class Map is used to define the classes of the ontology model, which corresponds to the mapping of the data table, and d2rq: Property Bridge is used to define the attributes in the ontology model, which corresponds to the mapping of the columns and relational tables in the data table. The mapping of data in the relational database to RDF should follow the classes and attributes defined in the ontology model [19]. The five main data tables in the relational database are mapped to four core classes and their attributes, intangible cultural heritage item, inheritor, classification, and geographic location. The relational table is mapped to the "has Inheritor" object attribute, as shown in Figure 6. The fields of each data table are mapped to corresponding properties.

Finally, according to the data mapping file, use the dump-rdf tool of the D2RQ platform to convert the data in the relational database to generate an RDF/XML format file for use by other databases or third-party applications. Although D2RQ can also publish linked data, it is not flexible enough to update and manage data and has limited support for complex relationships and massive data. Therefore, this paper uses D2RQ to semantically transform the data into RDF/XML format files. Then, the data are stored in a special Triplestore database, and the server is configured to realize the associated publishing of the data and the data open interface. Table 2 corresponds to the mapping framework, which is the core statement for semantically mapping the item table, the inheritor, and the inheritance relationship table.

3.3. Entity Association. Entity association is based on the RDF description of entities, uses RDF links to describe the semantic relationship between different entity objects, and establishes associations with external data as much as possible to build a linked data network. Linking data to other open RDF data sets and vocabularies is a key step in enriching the semantics of linked data. Entity linking should more semantically link internal data with external open data sets and realize knowledge discovery through knowledge aggregation across domains, disciplines, and databases. In the Semantic Web environment, with the help of the standardization and open interconnectivity of linked data and the integration of multiple knowledge bases, the richness and breadth of knowledge in the field of intangible cultural heritage can be improved, providing scientific research and knowledge dissemination of intangible cultural heritage culture. Data Foundation and Knowledge Services [20]: To enrich the knowledge in the field of intangible cultural heritage, this paper chooses to match and associate data with DBpedia and Geo Names linked data projects. Through the association with DBpedia and its data source Wikipedia, the Chinese and English entries corresponding to the intangible cultural heritage items are obtained; through the association with the global geographic database (Geo Names), more information about the regions involved in the intangible cultural heritage field can be obtained. Use the OWL built-in attribute owl: same As to associate the internal knowledge entity with the entity in the external data set, indicating that the two linked entity objects are the same thing.

The DBpedia project is a large-scale knowledge data set based on linked data and established by extracting data from Wikipedia. It is the core of the linked open data cloud graph. The DBpedia data set contains a large amount of information about my country's intangible cultural heritage described in Chinese and English, and entity association with it can

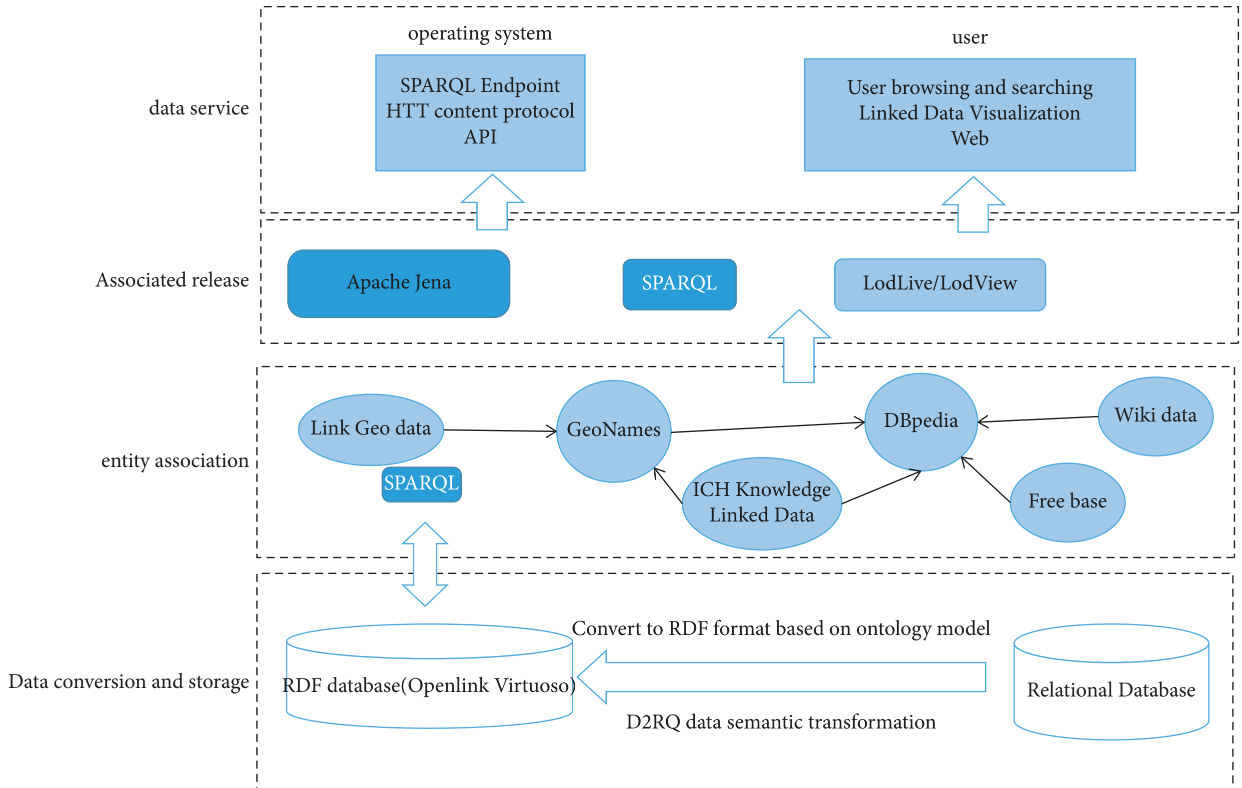


Figure 5: The technical framework for the construction of the intangible cultural heritage knowledge association data set.

enrich the knowledge in the field of intangible cultural heritage and improve the visibility and generality of intangible cultural heritage knowledge [21]. This paper adopts a combination of automatic retrieval and manual inspection. First, the resource items related to intangible cultural heritage items in DBpedia are retrieved online through SPARQL language, and then the retrieval results are screened and checked by manual inspection, and finally the retrieved resources are retrieved. The URI is associated with the internal intangible cultural heritage item entity through the owl: same As attribute. In the DBpedia ontology model, the db : abstract attribute represents the abstract of the resource, and the foaf: is Primary Topic of attribute links the Wikipedia page corresponding to the resource. The SPARQL Endpoint site of DBpedia is called online, and the related concepts of the intangible cultural heritage items are retrieved by constructing SPARQL sentences to retrieve the resource items containing the keyword. Due to the imperfections of DBpedia, for example, some intangible cultural heritage items lack entries, the Chinese information provided by DBpedia is incomplete, the resource entries related to intangible cultural heritage in Wikipedia are not included in DBpedia, and because the titles of intangible cultural heritage items are not uniform, through the above methods, only some entities are correctly matched, and some are not fully matched. Therefore, based on automatic retrieval and manual inspection, this paper associates as many internal intangible cultural heritage project entities with DBpedia-related resources as possible. In the end, about 1/3 of the intangible cultural heritage items are physically associated with DBpedia or Wikipedia, which also highlights the necessity of establishing an encyclopedic database of intangible cultural heritage in my country.

The Geo Names geographic database contains more than 10 million geographic names around the world and provides information such as geographic name alternatives, latitude and longitude, population, and Wikipedia. It adopts the principle of linked data to organize, defines a unique resource URI for each geographic name, and publishes the geospatial semantic information to the Internet. To obtain more information about geographic location, enrich domain knowledge, and provide a data basis for spatial analysis of intangible cultural heritage based on geographic information, this paper combines entities related to geographic location, such as intangible cultural heritage project application areas and inheritor's residence, with Geo Names database for the association. The Geo Names data can be obtained by calling its official API or by using the SPARQL endpoint provided by a third party [22]. To ensure the consistency with the DBpedia data association, this paper uses the SPARQL endpoint of the Factforge website to obtain Geo Names data. The core SPARQL query statements related to data matching and association are shown in Table 3.

The query sentence uses "Dark Biography" as the keyword to perform a full-text search on the Chinese tags of DBpedia and obtain the abstract information of the resource and the corresponding Wikipedia page. It is linked to the intangible cultural heritage item entities in the internal data set.

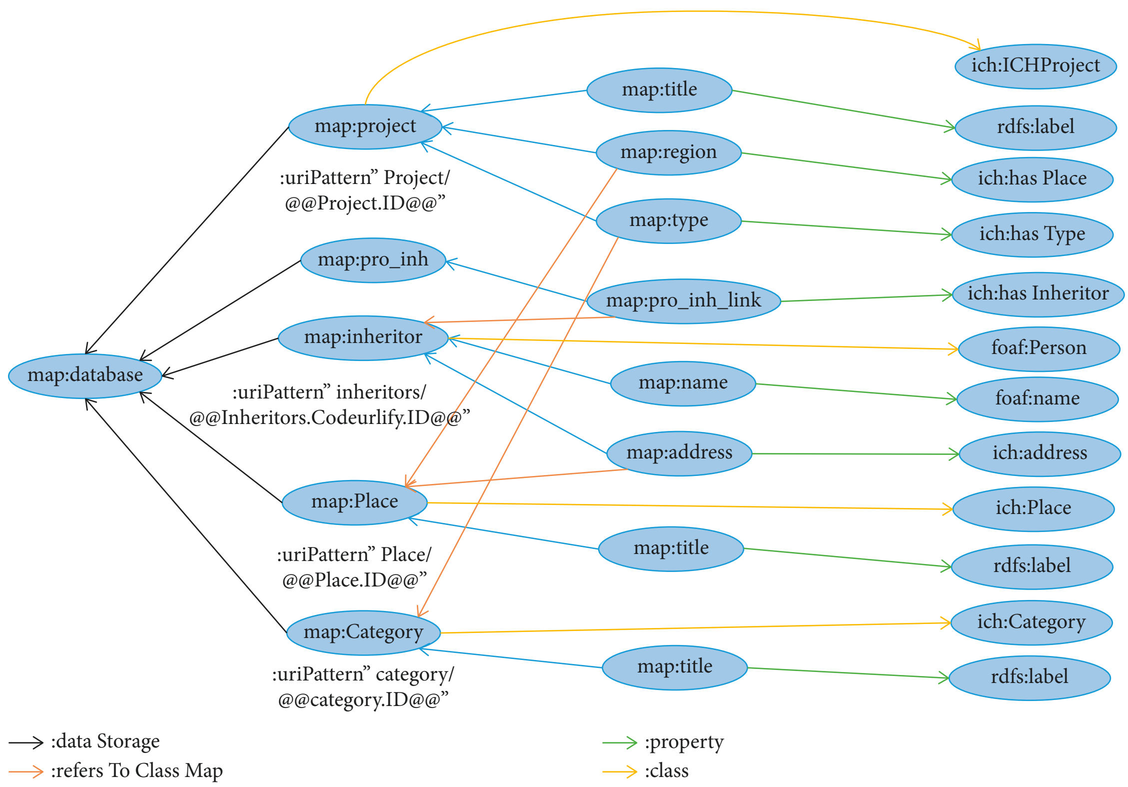


Figure 6: Mapping framework for classes and properties.

Retrieve the resource entry corresponding to "Enshi Prefecture" in the Geo Names database and limit the search scope to China (coded as CN) and the resource type to country and region (coded as A). The retrieval result will return the URI of the geolocation object, <http://sws.geonames.org/1811624/>, and you can further obtain information such as the latitude and longitude of the geographic location and the Wikipedia link. In the case of the same geographic names, disambiguation can be achieved by judging the municipal and provincial geographic names (parent ADM2, parent ADM1). In the above way, the association to the geographic location entity achieves a 100% match. Finally, based on constructing internal intangible cultural heritage domain knowledge-linked data, multiple categories of entities such as intangible cultural heritage items, people, and geographic locations are associated with external databases such as DBpedia and Geo Names. The number of entities and triples in the finally constructed intangible cultural heritage domain knowledge association database are shown in Table 4. Among them, there are more than 10,050 triples, involving multiple categories of objects such as intangible cultural heritage items, inheritors, institutions, geography, and types. There are 505 entities associated with DBpedia and 295 entities associated with Geo Names. By linking DBpedia, it can be more convenient to associate with Wikipedia, WIKIDATA, YAGO, and other resources. Moreover, the constructed intangible cultural heritage knowledge-linked data are completely open according to the W3C standard, and data services can be obtained through online retrieval, SPARQL Endpoint, and other data cabling methods.

3.4. Storage and Release of Intangible Cultural Heritage Knowledge-Related Data. After converting all kinds of data and knowledge in the field of intangible cultural heritage into the form of linked data, it needs to be persistently stored and published. The storage and publishing of linked data directly affect the sharing and reuse of data [23]. There are many ways to store and publish linked data. This article adopts the native method for storage, configures the server for linked data publishing, and provides open data services and knowledge services. Open Link Virtuoso is used as the storage and management database of RDF data, and Lod View is used for users to provide data browsing of the intangible cultural heritage knowledge base and use Lod Live to provide a visual representation of the intangible cultural heritage knowledge association graph.

The storage scheme of linked data is roughly divided into relational data-based storage, NoSQL database storage, and Tri-plestore database storage. Among them, the Triplestore database is specially developed for the characteristics of the RDF data structure and has efficient data storage, query, and reasoning mechanisms; at the same time, because it adopts a unified data model, it can achieve efficient interaction between data. Considering the expansion of knowledge in the field of intangible cultural heritage and the growth of data in the future, this paper chooses Open Link Virtuoso, which is

Table 2: Statements of data semantic mapping (part).

Category	Check for phrases
#Intangible cultural heritage project table/mapping to ich : ICHProject class	map: project a d2rq: class map; d2rq: data storage map: database; d2rq: uri pattern "project/@@project.ID@@"; d2rq: class ich: ICHProject; d2rq: class definition label "project"; .
#Inheritor information table/mapping to foaf : Person class	map: inheritor a d2rq: class map; d2rq: data Storage map: database; d2rq: uri pattern "inheritor/@@inheritor.id@@"; d2rq: Class foaf: person; d2rq: class definition label"inheritor";
#The relationship table between inheritors and intangible cultural heritage items/mapping to ich : has inheritor object attributes	map: pro_inh_link a d2rq: property bridge; d2rq: Belongs to class map map: project; d2rq: property ich: has inheritor; d2rq: Refers to class map map: inheritor; d2rq: join "pro_inh . Inheritorid => inheritor . urlcode"; d2rq: join "pro_inh . Projectid => project . urlcode"; .

Table 3: Data matching and associated core SPARQL query statements.

Category	Check for phrases
Intangible cultural heritage item class entities are matched against DBpedia	PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#> PREFIX dbo: < . Org/ontology/" title = "http://dbpedia . Org/ontology/ ">http://dbpedia . Org/ontology/> SELECT ? res ? title ? Wikipedia ? abstract FROM < http://dbpedia . Org> Where{? Res rdfs: label? Title . ? title bif: contains""Dark biography." ? res foaf: is primary topic of ? Wikipedia . ? res dbo: abstract ? abstract.filter(lang(? abstract) = "zh").}
Geolocation entities are matched against Geo names	PREFIX gn: < http://www . Geonames.org/ontology# > SELECT ? gnfeature WHERE { ? gnfeature gn: country Code'CN '. ? gnfeature gn: alternate Name'Enshi Tujia and Miao Autonomous Prefecture'@ zh. ? gnfeature gn: feature Classgn: A . } LIMIT100

widely used, like the database management software. The virtuoso database is a cross-platform scalable high-performance database management software that provides SQL, XML, RDF database management functions and supports the storage and management of billions of scale triples. The virtuoso database provides multiple mechanisms such as WEB pages or ISQL commands for data import. The original data related to intangible cultural heritage is semantically described or semantically mapped to generate RDF format data and then imported into the Virtuoso data storage, and the IRI (Internationalized Resource Identifiers) of the named graph to which the data belong is specified.

The intangible cultural heritage domain knowledge RDF data constructed above needs to be published through linked data to realize the utilization and sharing of data; the commonly used publishing methods of linked data are based on static RDF/XML files, based on relational data, and based on RDF data repository, using RDFa formula and other ways. Based on the Virtuoso database platform, this paper follows W3C's four principles of linked data publishing, configures a linked data publishing server, and provides services such as RDF data management, linked data browsing, and SPARQL endpoints and content negotiation. To fully and comprehensively display the attributes and relationships of each entity in the intangible cultural heritage-related data set, the Lod View tool is used to provide users with a browse of the intangible cultural heritage-related data. Lod View is a Web application developed based on Jena and Spring framework. It supports the parsing of International Resource Identifiers (IRI) conforming to the W3C standard and is a tool for converting RDF data format to HTML; configuring Lod View's SPARQL site and multimedia display and Attributes such as latitude and longitude; and returning the correct RDF data and web page description according to the content negotiation mechanism. The entities of the intangible cultural heritage knowledge association data set are distinguished according to the types of entities, including intangible cultural heritage items, inheritors, place names, and project types. The URI format of entity naming is http://base URI]/[entity Type Name]/[entity ID], where [entity Type Name] corresponds to the class in the ontology model; such as http://localhost:

Table 4: Statistics of knowledge-related data in the field of intangible cultural heritage.

Category	Number
Triad	10050
entity	1200
Intangible cultural heritage project entity	580
Heir entity	530
Associate with DBpedia	505
Associate with Geo names	295

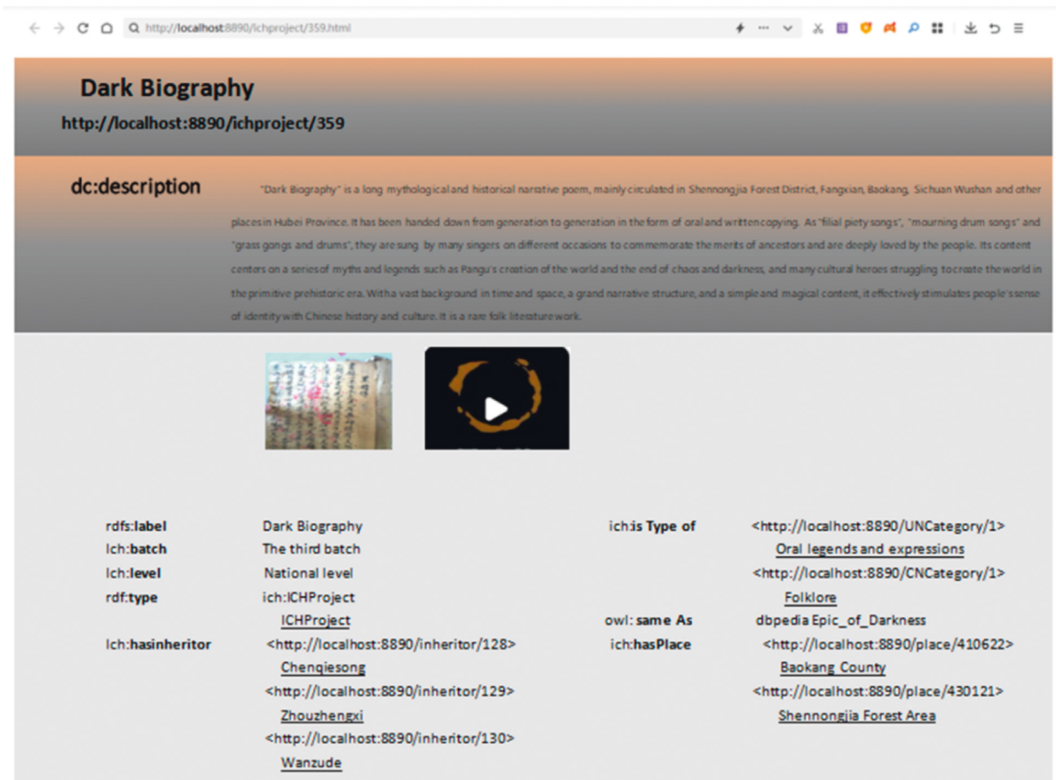


Figure 7: Intangible cultural heritage item object page.

8890/resource/ichproject/359 represents the instance of "Intangible Cultural Heritage Project Class" "Dark Biography"; the page that uses the LodView tool to display the details of this instance is shown in Figure 7.

This page shows all the properties of the "Dark Biography" instance of the intangible cultural heritage item. The upper part of the page is a multimedia display part, which visually displays multimedia resources such as pictures and videos related to it. The page lists the representative inheritors, types, distribution areas and other related objects of the project. For example, link the representative inheritor of the project through the ich: has inheritor attribute and click the hyperlink to jump to the detailed page of the inheritor object. When the user accesses, it provides an intuitive HTML page; when the application accesses, it returns the data in the corresponding format such as RDF/XML, RDF/Turtle, JSON, and so on, according to different content requests. In the intangible cultural heritage knowledge association data set, geographic location objects are associated with Geo Names. When the user accesses the information page of the instance of the geographic location class "Shennongjia Forest Area," the location in the map is displayed according to the latitude and longitude information in the form of an online map. The middle of the page displays the specific attributes of the geographic location class and is associated with DBpedia, Geo Names and Wikipedia; at the bottom of the page, through the inversion of the ich : has Place attribute, the intangible cultural heritage items, and inheritor objects owned in the region are retrieved in reverse.

4. Conclusions

The development of the Semantic Web and linked data has provided new ideas and methods for the organization and sharing of knowledge in the field of intangible cultural heritage and has changed the representation and expression of intangible cultural heritage knowledge. To realize the effective organization and management of intangible cultural heritage knowledge, based on processing and sorting out the knowledge in the field of intangible cultural heritage, this paper sorts out the elements and semantic relations of intangible cultural heritage knowledge and constructs the ontology model of intangible cultural heritage knowledge to reveal the field of intangible cultural heritage. A wealth of knowledge and the interconnectedness of knowledge: the intangible cultural heritage knowledge-linked data set constructed based on linked data technology provides data consumption and sharing in the form of linked open data and provides knowledge services such as domain knowledge related to display and knowledge visualization based on linked data sets. The ideas and technologies of ontology and linked data have great advantages in the organization of knowledge in the field of intangible cultural heritage. Using linked data technology to build a knowledge base in the field of intangible cultural heritage in my country and physically linking it with international open data sets can improve knowledge in the field of intangible cultural heritage, relevance, sharing and openness, and influence.

The construction of high-quality intangible cultural heritage knowledge open data sets and knowledge service platforms, on the one hand, provides a scientific data set for deeper analysis and research on intangible cultural heritage culture and on the other hand promotes intangible cultural heritage knowledge to play a greater role The dissemination and inheritance of intangible cultural heritage are of great significance. The intangible cultural heritage contains a wealth of knowledge. This article only studies and sorts out the basic knowledge elements such as intangible cultural heritage items, inheritors, regions, and inheritance relationships. The granularity of knowledge organization needs to be further refined. Intangible cultural heritage knowledge ontology model still needs to be further expanded and enriched according to actual needs. The follow-up research will expand the data sources and research objects. In the big data environment, the organization and management of the massive heterogeneous knowledge of the intangible cultural heritage will be studied.

Data Availability

The experimental data 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 to report regarding the present study.

Acknowledgments

This study was supported by Education Sciences Planning 13th Five-Year Key topics of the Ministry of Education in 2020 "Research on the practice mode and performance evaluation system of vocational education supporting the industrialization of intangible cultural heritage poverty alleviation" (Item No: DJA200312); Guangxi Philosophy and Social Sciences Planning Research Project in 2021 "Collection and research of oral archives of intangible cultural heritage inheritors of traditional skills in Guangxi" (Item No: 21BMZ011); Special Topic of Guangxi Ethnic Education Research in 2019 in the 13th Five Year Plan of Guangxi Educational Science "Research on Guangxi Zhuang cultural inheritance and education under the guidance of socialist core values" (Item No: 2019ZJY085) Phased research results; and Philosophy and Social Science Planning Project in Inner Mongolia Autonomous Region in 2021: Research on the Protection and Inheritance of Intangible Cultural Heritage "Wushen Horse Race" (project no. 2021NDB165).

References

[1] Q. Jia and P. Wang, "Research on the aggregation of historical archives resources based on linked data," *Library and Information Work*, vol. 65, no. 10, pp. 105–112, 2021.

[2] S. Jong-Soo and I. J. Chung, "Dynamic FOAF management method for social networks in the social web environment," *The Journal of Supercomputing*, vol. 66, no. 2, pp. 633–648, 2013.

[3] C. Lu, Y. Xiong, and C. Liu, "Construction of Knowledge Organization System of Intangible Cultural Heritage Resources Based on Ontology and Metadata," *Library Theory and Practice*, no. 3, pp. 39–43, 2016.

[4] Y. Huang, W. Lu, Q. Cheng, and S. Deng, "The design and implementation of intangible cultural heritage knowledge ontology construction system——taking tibet's "guozhuang" and "duixie" as examples," *Journal of Tibet University for Nationalities (Philosophy Social Science Edition)*, vol. 37, no. 1, pp. 20–26+154, 2016.

[5] Y. Zhou, Y. Zhao, and J. Sun, "The research path of intangible cultural heritage information resources organization and retrieval—inspection and design based on ontology method," *Journal of Information*, vol. 36, no. 8, pp. 166–174, 2017.

[6] S. Ou and Z. Tang, "Research on automatic question answering technology for library linked data," *Chinese Library Journal*, vol. 41, no. 6, pp. 44–60, 2015.

[7] S. Li and Y. Zhao, "Development of intangible cultural heritage archive resources based on linked data," *China Archives*, no. 6, pp. 71–73, 2016.

[8] K. Dong, "Research on semantic organization of intangible cultural heritage based on linked data," *Modern Information*, vol. 35, no. 2, pp. 12–17, 2015.

[9] Y. Chen, Z. Chen, and Y. Qiu, "Research on the development strategy of enterprise archives informatization based on knowledge management," *Management and Technology of Small and Medium Enterprises*, no. 2, pp. 10–12, 2022.

[10] Y. Lu, "Research on Digital Archives Resource Sharing Based on Linked Data," Central China Normal University, Wuhan, China, Master, 2017.

[11] M. Huang, "Design of teaching resource library system based on java web technology," *Electronic Technology and Software Engineering*, no. 1, pp. 229–232, 2022.

[12] Y. Cao, Z. Gong, D. Chen, and Y. Kang, "Linked data technology and its research status," *Library Theory and Practice*, no. 11, pp. 42–45, 2014.

[13] W. Liu, "Linked data: concept, technology and application prospects," *Journal of University Libraries*, vol. 29, no. 2, pp. 5–12, 2011.

[14] Q. Nian, L. Li, S. Zhao, and H. Wu, "Knowledge modeling in the field of ontology intangible cultural heritage," *Journal of*

Fujian Normal University of Technology, vol. 40, no. 2, pp. 119–126, 2022.
[15] S. Ye and L. Pan, "The formation, expression and transformation of the inheritance and development of intangible cultural heritage: based on the perspective of cultural capital," *Jinyang Journal*, no. 3, pp. 109–114, 2022.
[16] M. Fan, F. Zhu, and Y. Wu, "A web data sharing technology based on metadata," *Journal of Mianyang Normal University*, vol. 23, no. 2, pp. 34–40, 2004.
[17] X. Zhang, "Research on the distribution characteristics of rare tree species in Pingnan County based on GIS spatial analysis," *Fujian Forestry*, no. 6, pp. 42–45, 2021.
[18] P. Akeem, P. Anhtuan, N. P. Thanh, and P. H. Chien, "Data-driven construction safety information sharing system based on linked data, otologies, and kowledge gaph tchnologies," *International Journal of Environmental Research and Public Health*, vol. 19, no. 2, 2022.
[19] J. Lu and Y. Li, "Mapping method of object relational database to RDF(S)," *Computer Science*, vol. 48, no. 10, pp. 145–151, 2021.
[20] H. Gao, "Research on ontology mapping based on similarity calculation in semantic web," *Network Security Technology and Application*, no. 2, pp. 41-42, 2022.
[21] J. Shi, "Research on Entity Linking Technology Based on DBpedia Knowledge Base," Southeast University, Dhaka, Bangladesh, Master, 2018.
[22] W. He, "SPARQL Joint Query and its Application," Dalian Maritime University, Dalian, China, 2014.
[23] X. Zhu and R. Wang, "Research on ontology construction of intangible cultural heritage image semantic information and its associated data storage and release," *Modern Intelligence*, vol. 41, no. 6, pp. 54–63, 2021.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9890363, 1 page
https://doi.org/10.1155/2023/9890363

Retraction

Retracted: Study on Employee Performance Evaluation Based on Adaptive Feature Selection Fuzzy Algorithm

Mobile Information Systems

Received 13 September 2023; Accepted 13 September 2023; Published 14 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] Y. Qian and J. Yin, "Study on Employee Performance Evaluation Based on Adaptive Feature Selection Fuzzy Algorithm," *Mobile Information Systems*, vol. 2022, Article ID 6985508, 13 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 6985508, 13 pages
https://doi.org/10.1155/2022/6985508

Research Article

Study on Employee Performance Evaluation Based on Adaptive Feature Selection Fuzzy Algorithm

Yuan Qian [1] and Jiemin Yin [2]

[1] *Nanchang Institute of Science&Technology, Nanchang 330108, Jiangxi, China*
[2] *School of Law, Shanghai University of Finance and Economics, Shanghai 200433, China*

Correspondence should be addressed to Yuan Qian; qianyuan@ncpu.edu.cn

Received 22 June 2022; Revised 20 July 2022; Accepted 29 July 2022; Published 26 August 2022

Academic Editor: R. Mo

In our study, in terms of performance evaluation methods, the performance evaluation algorithm based on genetic algorithm and fuzzy comprehensive performance evaluation algorithm is introduced, their advantages and disadvantages are compared and analyzed, and the design idea of fuzzy performance evaluation algorithm based on compound elements is proposed. It can be divided into seven steps: first, clarify the evaluation purpose and object; second, select the optimal evaluation mode and method; third, compile the evaluation index system; fourth, it is to collect information extensively; fifth, the evaluation adopts a variety of methods, multiple angles, and multiple sides to collect materials so that the conclusion of the evaluation has sufficient factual basis; the sixth is to process the information and make a comprehensive evaluation; the seventh is to analyze the results and write an evaluation report. Based on the existing algorithms, the fuzzy performance evaluation algorithm based on compound elements is studied, and the detailed design of the algorithm is presented. Through the comparison and analysis of the performance evaluation algorithm based on compound elements and the performance evaluation algorithm of specific elements, the superiority of the performance evaluation algorithm based on compound elements is experimentally verified by comparing the operation time and classification accuracy. The performance evaluation system based on the performance evaluation algorithm designed in this study, combined with the official business of the performance evaluation system, modularizes the administrative management activities of the enterprise, digitizes the electronic office information, and conducts in-depth exploration of the unstructured natural language. A real-time performance evaluation system based on the arrangement of corporate administrative activities has been established. By designing and implementing a performance-assisted analysis system based on text content analysis, which is suitable for performance evaluation systems, it solves the imperfect problem of performance evaluation based on electronic enterprise administrative management.

1. Introduction

In the transition period from demographic dividend to talent dividend, the performance management of Chinese enterprises at this stage is particularly important [1]. It is of great significance to the company's prospects and the ability development of employees, so improving performance management means improving the core competitiveness of the company and employees themselves [2]. In most Chinese enterprises, performance management has been using the most traditional performance management method. However, with the advent of the market economy, more and more state-owned enterprises have found that the traditional performance appraisal methods can no longer meet the needs of the survival and development of enterprises in the process of using them in recent years [3]. Therefore, some enterprises have begun to seek a more scientific and reasonable performance appraisal system that is more meaningful to the development of the enterprise, and the changes in state-owned enterprises are particularly obvious.

At present, the more popular methods are KPI (Key Performance Indicator), forced distribution method, pairwise comparison method, balanced scorecard, 360-degree evaluation method, and so on [4]. However, due to the influence of China's national conditions, this method originated in the United Kingdom and will always encounter some small problems in the implementation process in China [5]. In the current application of KPIs, many enterprises do not follow the above points to determine indicators, and even some indicators are formulated very arbitrarily, which is one of the common drawbacks of KPI application in Chinese enterprises [6]. When the KPI indicators are not specified scientifically and perfectly, the role of the KPI will be greatly reduced, thus forming a certain negative effect on the strategic management of the enterprise [7]. There is no further decomposition of performance goals to the basic management and operating personnel of the enterprise. KPIs fail to provide a complete set of indicators framework systems with specific guiding significance for operations.

Schain et al. believed that in the process of modern social development, information technology has become one of the indispensable and important means, and is widely used in all aspects of production and life [2]. However, with the advancement of reform, many state-owned enterprises have also begun to change their performance management models [8]. Their performance management is abandoning the original assessment method and changing to a more scientific and reasonable assessment method. For example, goal-oriented KPIs have become analytical performance management [9]. The main means and basis of indicators. This not only makes the company's strategic goals and employees more closely linked but also can make the staff's behavioral standards to follow. However, there are still some difficulties in the progress of the new performance management module of many companies, for example, there are many problems in the implementation of KPI [10].

The public welfare nature of non-profit organizations has aroused the necessary attention of society to their business practices. The performance level of non-profit organizations has become the target of the government, enterprises, and the public. Many scholars have also begun to study the performance evaluation of non-profit organizations [11]. The performance evaluation of non-profit organizations is an emerging topic. Compared with the performance evaluation of enterprises, the performance evaluation of non-profit organizations has the characteristics of the particularity, complexity, diversity, openness, and public welfare of the evaluation objects [12]. The performance evaluation has brought problems such as subjective one-sidedness of evaluation indicators and difficult quantification of evaluation indicators. Since the fields and industries involved in non-profit organizations are very wide, and there are certain differences between them, in order to fully explain the performance evaluation of non-profit organizations, the author selects large public enterprises as the research object [13]. According to the current situation of large enterprises in my country, follow the principle of purpose, the principle of comprehensiveness and simplicity, the principle of scientificity, the principle of measurability, and the principle of representativeness [14].

Most of the existing fuzzy clustering algorithms are based on the fuzzy comprehensive evaluation algorithm of specific elements [15]. The main problems of such fuzzy performance evaluation algorithms are: employee performance evaluation is a multi-element, multi-level evaluation system, and the fuzzy comprehensive evaluation algorithm of specific elements lacks the responsibility for the overall study of performance evaluation, and the algorithm lacks hierarchy, so the above two types of algorithms are not suitable for data mining of complex employee performance evaluation systems [16]. Enterprise engineering project management is a complex nonlinear system operation process, and its performance evaluation has the characteristics of multi-objective, multi-index, and multi-stage, among which progress, cost, quality, and safety are the key indicators of evaluation [17]. The key to performance evaluation is to approximately restore the original system by establishing a nonlinear mapping, and the nonlinear mapping is the algorithm model to be constructed [18]. The neural network is a highly nonlinear dynamic system, and its strong nonlinear fitting ability can well find nonlinear functions to effectively approximate the mapping relationship of the attractor trajectory state in the embedded space, making it useful in performance evaluation [19].

The so-called fuzzy comprehensive evaluation algorithm refers to a systematic and comprehensive evaluation of the evaluation factors of all the intervention objects at the same level so that a complete process can be regarded as the end of an evaluation [20]. Specifically, a fuzzy comprehensive evaluation is based on fuzzy mathematics, applying the principle of fuzzy relationship synthesis, quantifying some factors with unclear boundaries and difficult to quantify, and comprehensively evaluating the membership level of the evaluated object from multiple factors [21]. First, determine the factor (indicator) set evaluation (level) set of the object to be evaluated; then determine the weights of each factor and their membership degree vectors respectively, and obtain the fuzzy evaluation matrix; finally, the fuzzy evaluation matrix and the weight vector of the factors are subjected to fuzzy operations and combined [22]. Perform normalization to get fuzzy comprehensive evaluation results. Its characteristic is that the evaluation is carried out on an object-by-object basis, and it has a unique evaluation value for the evaluated object and is not affected by the object set where the evaluated object is located [23]. The purpose of a comprehensive evaluation is to select the winning object from the object set, so it is also necessary to sort the comprehensive evaluation results of all objects. Fuzzy evaluation deals with the fuzzy evaluation objects by precise digital means and can make a scientific, reasonable and realistic quantitative evaluation for the data that contains ambiguity [24].

The types of performance evaluation include benefit index, cost index, intermediate index, and interval index. Benefit-type indicators require that the larger the index value, the better, such as output value and interest rate value [25]. Cost-based indicators require the smaller the value, the better, such as energy consumption and expenses [26]. The centered index requires the value to be in the center, the better, and the interval index requires the best value to fall within a certain interval. When calculating the performance evaluation of different types of indicators, it is necessary to carry out type conversion [27]. In the performance evaluation, the cost-type indicators, intermediate-type indicators, and interval-type indicators are all converted into benefit-type indicators. The method of converting cost-type indicators into benefit-type indicators is $x^* = M\text{-}x$, where M is a maximum upper limit value of indicator x. The method of converting the middle-type index to the benefit-type index is

$$x^* = \begin{cases} 2(x-m), \leq x \geq \dfrac{M+m}{2}, \\ \\ 2(M-m), \dfrac{M+m}{2} \leq x \leq M, \end{cases} \tag{1}$$

where m is a low limit value of x. The method of converting an interval-type indicator to a benefit-type indicator is:

$$x^* = \begin{cases} 1 - \dfrac{q_1 - x}{\max(q_1 - m, M - q_2)}, x < q_1, \\ \\ 1, q_1 \leq x \leq q_2, \\ \\ 1 - \dfrac{x - q_2}{\max(q_1 - m, M - q_2)}, x > q_2, \end{cases} \tag{2}$$

$q_1 \leq x \leq q_2$is the most reasonable range for x. To determine the weight of the evaluation index, first, determine the weight coefficient. The performance evaluation weight coefficient can be determined by the subjective weighting method according to the importance of the evaluation index, or it can be weighted according to the degree of change of the evaluation index value based on the objective weighting method, and the two can also be combined to form a combined weighting method [28]. Evaluation index aggregation is a comprehensive evaluation value that combines the values of multiple evaluation indicators into a whole through a mathematical model. There are many ways to assemble evaluation indicators, some are simple and some are more complex. Simple aggregation methods such as averaging the evaluation indicators, the weighted average of the evaluation indicators, max operator and min operator, etc.

This study develops an employee performance evaluation system based on the company's employee performance evaluation needs. The system mainly from the perspective of the evaluator's activities establishes the performance of six modules of employee information management, attendance management, salary management, performance management, system management, and report management. In terms of performance evaluation methods, the performance evaluation algorithm based on genetic algorithm and fuzzy comprehensive performance evaluation algorithm is introduced, their advantages and disadvantages are compared and analyzed, and the design idea of fuzzy performance evaluation algorithm based on compound elements is proposed. Specifically expounds on the process of employee performance evaluation, which can be divided into the following seven steps: one is to clarify the evaluation purpose and object; method; the third is to compile the evaluation index system; the fourth is to collect information extensively, and the evaluation uses a variety of methods, multiple angles, and multiple aspects to collect materials so that the conclusion of the evaluation has sufficient factual basis; the sixth is to process information and comprehensively evaluate; Analyze the results and write an evaluation report.

This study first analyzes the improved algorithm of performance evaluation and designs a fuzzy performance evaluation algorithm based on compound elements. After functional testing and performance testing, the expected research and development goals have been achieved, and finally, a performance evaluation system has been implemented.

2. Algorithm and Research Object

2.1. Algorithm Design. The existing fuzzy clustering algorithms are mostly fuzzy comprehensive evaluation algorithms based on specific elements, but the main problem of this type of fuzzy performance evaluation algorithm is that employee performance evaluation is a multi-element, multi-level evaluation system, and the specific element fuzzy comprehensive evaluation algorithm [29]. There is a lack of overall research on responsible performance evaluation, and the algorithm lacks hierarchy, so the above two types of algorithms are not suitable for data mining of complex employee performance evaluation systems [30].

A fuzzy comprehensive weighted evaluation algorithm for compound elements is dived into four steps [31]. (1) first, the factors that affect the evaluation are classified according to their attributes, and different levels of factors are correspondingly different. After classification, various factors will be reduced. The advantage is that the weight is easy to allocate reasonably, and the weight coefficient will not be too small, so there will be no situation where the fuzzy evaluation "data" of specific elements is submerged; (2) the fuzzy comprehensive evaluation of specific elements is carried out gradually from the bottom-level factors to the high-level factors. After the evaluation factors are divided into multiple levels, the factors at the lowest level are more specific, and the evaluation results tend to be more reasonable; (3) comprehensive evaluation is carried out layer by layer according to various factors, and the second, third, ..., n-level fuzzy comprehensive evaluation is obtained at one time; (4) in the performance evaluation system of company employees, the design idea of "weighting" is introduced because it needs to be designed according to different types of evaluators such as customers, company employees, inspection teams, etc. [32] These evaluators with different identities will evaluate the results have different weight effects, so these weights must be added first when calculating the membership degree, which can improve the accuracy of the evaluation. In the system design, a concept of weighting evaluation results is defined. The purpose is to avoid the same membership degree in the evaluation results so that the evaluation results can determine specific quantitative numbers, which is convenient for ranking the employees of the company under the same level.

Determine the evaluation factor set Assuming that the evaluation of things, there are evaluation sets composed of n evaluation factors $U = \{u_1, u_2, u_3, \ldots, u_n\}$, where

u_i, $i = 1, 2, \ldots, n$, is ith factors affecting evaluation. Assuming that there are m levels of evaluation results:

$$V = \{v_1, \ldots, v_n\}, \tag{3}$$

where v_j, $j = 1, 2, \ldots, n$ is the jth level.

Usually, the results are also consistent with the actual situation. The weight set of the first-level factor relative to the evaluation object is:

$$W = \{w_1, w_2, \ldots, w_n\}, \tag{4}$$

where $0 \le w_i \le 1$, $\sum_{i=1}^{n} w_i = 1$.

The weight set of the second-level factor relative to the first-level factor is:

$$\begin{cases} W_1 = \{w_1^1, w_1^2, \ldots, w_1^n\}, \sum_{i=1}^{n} w_1^i = 1, \\ W_n = \{w_n^1, w_n^2, \ldots, w_n^n\}, \sum_{i=1}^{n} w_n^i = 1. \end{cases} \tag{5}$$

The algorithm flow chart is shown in Figure 1.

From Figure 1, in the fuzzy evaluation system, the weight is very important, it reflects the role of different evaluation factors in the comprehensive decision-making process, and will directly affect the result of the comprehensive decision-making. To a certain extent, the weight can reflect the actual situation and can be given by experience.

Establish a weighted specific element evaluation matrix: first determine the degree of membership of a single factor to each evaluation level, the result of factor u_i is a fuzzy subset r_i, if there are k members forming an evaluation group, each member will evaluate each factor with a level, if the evaluation factor u_i level is:

$$R_i = \left(\frac{k_{i1}}{k}, \frac{k_{i2}}{k}, \ldots, \frac{k_{\mathfrak{J}}}{k}\right) = r_{i1}, \ldots, r_{\mathfrak{J}}. \tag{6}$$

When the members of the evaluation group are different (for example, the system includes customers, employees of the same level company, and inspection team evaluation), if the evaluation consists of k people, they are divided into P categories, and the number of people in the t category is k, then:

$$k = k_1 + k_2 + \ldots + k_n. \tag{7}$$

The weight is:

$$\beta = \beta_1 + \ldots + \beta_n. \tag{8}$$

If the evaluation factor is u_i, the various types of people at the level v_i are:

$$k_{ij1}, \ldots, k_{ijp}. \tag{9}$$

Thus, we can obtain:

$$R_i = \left(\frac{\sum_{t=1}^{P} \beta_t k_{ijt}}{\sum_{t=1}^{P} \beta_t k_t}, \ldots, \frac{\sum_{t=1}^{P} \beta_t k_{imt}}{\sum_{t=1}^{P} \beta_t k_t}\right) = (r_{i1}, r_{i2}, \ldots, r_{\mathfrak{J}}). \tag{10}$$

If each factor can be represented by a fuzzy quantity R_i, a fuzzy matrix can be used to represent the evaluation results of n factors:

$$R = (R_1, R_2, \ldots, R_n) = \begin{vmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{n3} \end{vmatrix}, \tag{11}$$

where $i = 1, 2, \ldots, m$.

The comprehensive performance evaluation formula is based on composite elements:

$$\begin{aligned} B = W \cdot R = (w_1, w_2, .., w_n) \cdot \begin{vmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{n3} \end{vmatrix} \\ = (b_1, b_2, \ldots, b_m), \end{aligned} \tag{12}$$

where $\cdot$ is composition operation, in our study, $M(\wedge, \vee)$ model is used.

$$b_j = \overset{i=1}{\underset{n}{\vee}} \left(w_i \wedge r_{ij}\right), j = 1, 2, \ldots, m, \tag{13}$$

where B is the comprehensive evaluation result and b_j represents the membership degree of the evaluation object to the jth level in the evaluation set. At the same time, it can also be seen that the difference weighted average method in the same grade is designed as follows:

$B = (b_1, b_2, \ldots, b_m)$, the evaluation result set can be normalized as:

$$B = (j_1, j_2, \ldots, j_3), j_i = \frac{b_i}{\sum_{j=1}^{m} b_j}, i = 1, 2, \ldots, m. \tag{14}$$

The evaluation results is:

$$V = \sum_{i=1}^{m} j_i v_i. \tag{15}$$

2.2. Performance Evaluation Algorithm Mining Objects. According to the needs of the system, the company's employee performance evaluation database is designed as Tables 1–4.

The table field design of the rater is shown in Table 1.

The field design of the company employee information table is shown in Table 2.

The fields of the company's employee performance evaluation information statistics table are shown in Table 3.

The fields of the company's employee performance evaluation information table are shown in Table 4.

The fuzzy comprehensive evaluation method for specific elements is very simple, and a fair and reasonable evaluation can be obtained under the condition of a few factors. However, when the evaluation is more complex and considers many complex factors, this method will lead to the

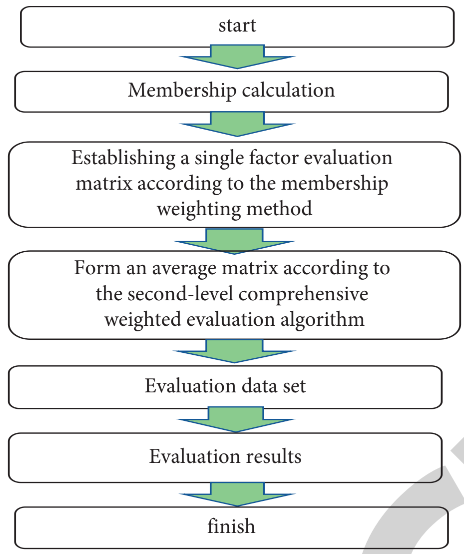


Figure 1: The algorithm flow chart.

Table 1: Rater table fields.

Names	Data types	Null or not
User name	Varchar ()	Not null
Password	Varchar ()	Not null
Permission	Varchar ()	Not null

Table 2: Division employee information form fields.

Names	Data types	Null or not
Employee code	Varchar ()	Not null
Employee name	Varchar ()	Not null
Employee information	Demo	Not null
Employee property	Varchar ()	Not null
Results	Int (30)	Not null

Table 3: Company employee performance evaluation information statistics table fields.

Names	Data types	Null or not	Illustration
Sn	Varchar ()	Not null	Employee code
Nam	Varchar ()	Not null	Employee names
Stu	Int (30)	Not null	Employee numbers
Tea	Int (30)	Not null	Evaluated employee numbers
Dep	Int (30)	Not null	Evaluated the number of department inspectors
Uni	Int (30)	Not null	Number of school inspectors evaluated
Ping	Varchar ()	Not null	Whether the evaluation conditions are met

Table 4: Company employee performance evaluation information table fields.

Names	Data types	Null or not	Illustration
Teacher	Varchar ()	Not null	Employee code
Superior	Int (30)	Not null	Number of times rated as superior
Good	Int (30)	Not null	Number of times rated as good
Median	Int (30)	Not null	Number of times rated as median
Bad	Int (30)	Not null	Number of times rated as bad

limitation of evaluation and assessment. Subjectively, people cannot determine the degree of the impression of various factors on the evaluation object, resulting in the inability to reasonably distribute the weights of various factors.

The weight satisfies the condition $\sum_{i=1}^{n} w_i = 1$. It can be seen that the fuzzy comprehensive evaluation algorithm of the specific element is used. Because the evaluation result is obtained according to the multiplication of the fuzzy matrix, the weight assigned to each factor is very small. Therefore, with smaller weights, a lot of information will be "overwhelmed" after the operation, and it is basically difficult to come up with any reasonable results. For example, when there are n kinds of employee evaluation factors, it will appear: $B = W \cdot R(\vee w_1, \vee w_2, \ldots, \vee w_n)$.

That is, if all the elements in the B set are the same and the largest element in W, then the information in the evaluation set of a specific element in R will be "submerged", and the evaluation result will be meaningless. When the evaluation system is quite complex, there are many types of evaluation object factors involved, and each factor has different levels. In the fuzzy comprehensive evaluation of specific elements, the level of the factors cannot be reflected.

3. Results and Discussion

3.1. Comparison and Analysis of Performance Evaluation Algorithm Based on Compound Elements and Performance Evaluation Algorithms of Specific Elements. The time consumption of this method is mainly divided into three parts, the time for calculating the attribute kernel C_O, the time for calculating the reduced classification and the time for selecting the non-kernel attribute, the complexity is: $t_c = O(|C||U|^2), t_{\text{non-core}} = O(|C||U|^2)$, $t_{c\_m} = O(|C||C-1||U|^2)$. The fuzzy performance evaluation algorithm based on specific elements and the performance evaluation algorithm based on genetic algorithm in most literature have the same complexity, and they are both attribute reduction algorithms based on information entropy, which are compared here. Its time consumption consists of two parts, including $t_c = O(|C||U|^2)$ and $t_{\text{non-core}} = O(|U|^3)$. Obviously, due to the need to additionally calculate the complementarity between attributes, the method in this paper has no advantage in time consumption. The same problem exists in space consumption.

The weighted specific element evaluation matrix constructed in the improved algorithm must have the same evaluation factors for evaluators with different identities before they can evaluate. Because of this system, the evaluation factors of each company employee can be the same for different evaluators, so the algorithm is more accurate for the performance evaluation of the company's employees. However, when this algorithm encounters a particularly complex evaluation system, the demand for data will be very large, and multiple matrix operations will be required, and the efficiency of the program will be significantly reduced. However, with the continuous development of computer hardware, the efficiency of software execution will also improve.

As can be seen from Table 5, for most data sets, the results of the algorithm in this paper are consistent with the results of the CEBARKCC algorithm, except for the third data set "SPECT" and the first data set "Australian Credit", but they finally get the reduced set, also have the same number of attributes. The time consumption of the algorithm is not given in the table, and only two of the datasets are time tested in this experiment. The algorithm is implemented under VC++6.0, and the test computer is configured as: a 2.4 GHz processor and 384M memory. For the third dataset "SPECT", the time consumption of the algorithm CEBARKCC is 803 ms, and the time consumption of the algorithm in this paper is 2967 ms. This is because the dataset has 22 conditional attributes, so calculating the complementarity between two attributes takes a lot of time, about 2215 ms. For the second dataset "IRIS", the time consumption of both algorithms is the same, 15 ms. The experimental results are the same as the previous analysis, and the algorithm has no time advantage.

For the data sets "Australian Credit" and "SPECT", the reduction results obtained by the two methods are different. Using the CART and J48 classifiers in the data mining software WEKA to perform 10-fold cross-validation experiments on the reduction results, two methods based on the classification accuracy of the performance evaluation attribute reduction set of the fuzzy algorithm are shown in Table 6. The experimental results show that the reduced attribute set obtained by this method has high classification accuracy.

This experiment selects nine datasets from the UCI database for testing. For the continuous variables in the data set, discretization preprocessing using the method provided by the ROSETTA software was used in the experiment, and the incomplete cases and inconsistent data of the data set were also removed to ensure the completeness and consistency of the data set. The algorithm CEBARKCC is an attribute reduction algorithm based on conditional information entropy proposed by Wang Guoyin et al. The reduction results are also given in Table 6.

3.2. Realization of the Function of Performance Evaluation System. The system first modifies the operation stage of the company's employee statistics table; second, extracts the first record data currently submitted as needed, and modifies it on the view formed by the retrieval; and third, after the process is over, the first factor data is stored directly. In addition, other factor data can be modified according to the simple IF statement for the retrieval attempt; finally, after evaluating the employees of the company, you can return to continue to evaluate other employees of company, but the evaluator cannot evaluate the employees already Evaluated employees of the company.

When making statistics on the stag table, the system first operates in the process of submitting data, and stores the data of the first factor directly in the database in time for reuse in this process; second, simply modify the second one, to achieve the purpose of colleges and universities using the

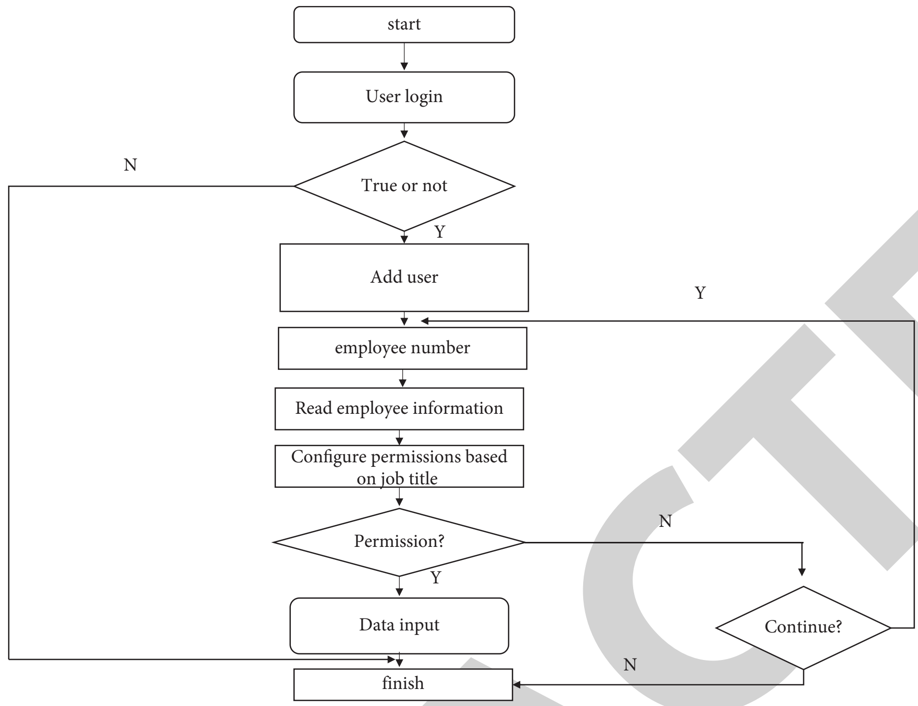


Figure 2: Add user information flowchart.

Table 5: Operation time comparison (milliseconds).

	Performance evaluation algorithm based on genetic algorithm (ms)	Fuzzy performance evaluation algorithm based on specific elements (ms)	Method in this study
Australian credit	240.6	1366.2	152.5
SPECT	366.1	1684.2	194.2

Table 6: Classification accuracy comparison.

		Fuzzy performance evaluation algorithm based on specific elements (%)	Algorithm in our study (%)
Australian Credit	J48	85.4	87.2
	CART	84.4	86.23
SPECT	J48	72.2	74.6
	CART	70.1	72.2

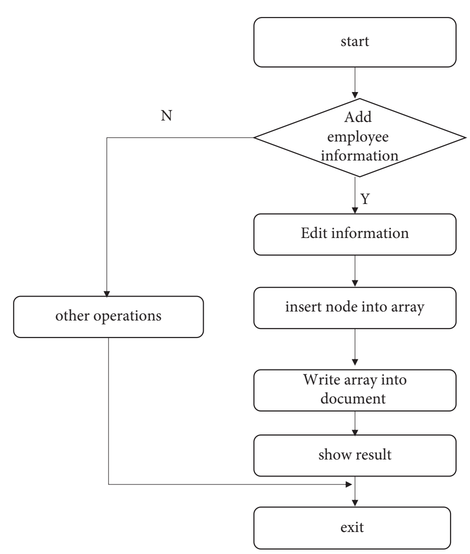


Figure 3: Add employee information flow chart.

views that have been retrieved. The evaluation factor set includes professional level, work discipline, work style and collaboration, work content, professional title level, academic research, work attitude, work method, work effect, etc. This study assumes that the weights of various evaluation factors are the same (Figure 2).

After the evaluation, the data is directly stored in the employee information table of the company, and when viewing the evaluation results, you can see the results of all the employees of the evaluated company. In this functional module, the design of the multi-level fuzzy comprehensive weighted evaluation algorithm and the realization of the program about the company's employees (Figure 3).

When designing a multi-level fuzzy comprehensive weighted evaluation algorithm in this paper, it is required that the number of customers who must be evaluated must be more than 10 employees, more than 5 employees of the same level company, and more than 5 experts from the

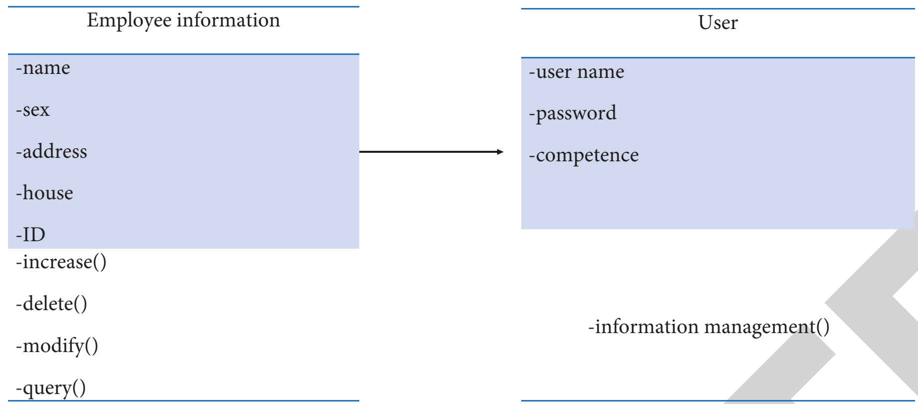


Figure 4: User management module class diagram.

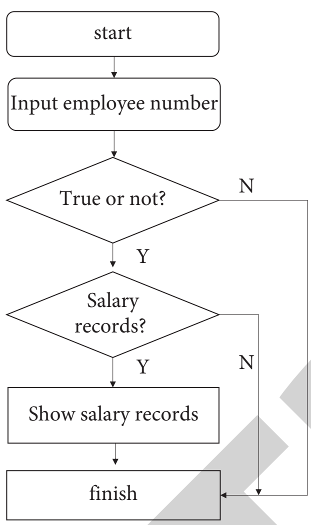


Figure 5: Query employee salary flow chart.

inspection team. Only in this way can the objective evaluation be ensured (Figure 4).

The evaluation result set including the evaluation result set of professional level, work discipline, work style, and cooperation is divided into very good, good, average, and poor; the evaluation result set of work content, academic research, work attitude, work method, work effect is divided into excellent, good, medium, and poor.

The process of querying employee salary includes the following contents. First, enter the employee number, determine whether the number is correct through CA certification, and determine whether the employee number exists in the company's performance evaluation system through a database keyword query. If only one of the two judgments has problems, then directly end the task of the salary management module; otherwise, display the salary record of the employee (Figure 5). From a practical point of view, the company's employee performance evaluation activities have an operational process, which is also the evaluation procedure mentioned in this article. This evaluation procedure is mainly discussed from the point of view of the activities of the evaluator. Specifically, look at a company's employee performance evaluation process.

The performance-assisted analysis system based on text analysis studied in this article is oriented to the existing e-government system. It builds a performance evaluation system by tracking the business office realization process and data flow analysis and realizes the quantitative and qualitative analysis of the total performance of government personnel. From a practical point of view, the company's employee performance evaluation activities have an operational process, which is also the evaluation procedure mentioned in this article. This evaluation procedure is mainly discussed from the point of view of the activities of the evaluator. Specifically, look at a company's employee performance evaluation process.

3.3. Performance Evaluation System and Framework Needs Analysis. The performance evaluation system based on text analysis proposed in this project is oriented to the existing company performance evaluation system, comprehensively compares the current relatively complete performance evaluation methods and systems, combines the data characteristics in the research field, and is based on the original company performance evaluation system, implementing a performance-assisted analysis function subsystem [33]. By analyzing the structured and unstructured related data of related modules such as enterprise administrative management activities, employee performance management file information processing, employee performance management file information flow records, etc., analyze the roles, tasks, and task completion of enterprise administrators in related matters On this basis, a quantitative model of the workload of enterprise administrators is established, so as to establish a performance evaluation system to count the quantitative workload of individuals and departments within a given time range.

The data storage involved in the company performance evaluation system contains both structured data of database table fields and unstructured natural language that does not depend on table structure. There is a clear hierarchical structure between structured information data and data, and there are certain operating specifications, which are easy to analyze but have limitations in content; unstructured information realizes human-computer interaction to a certain extent, which is convenient for users to use, but the

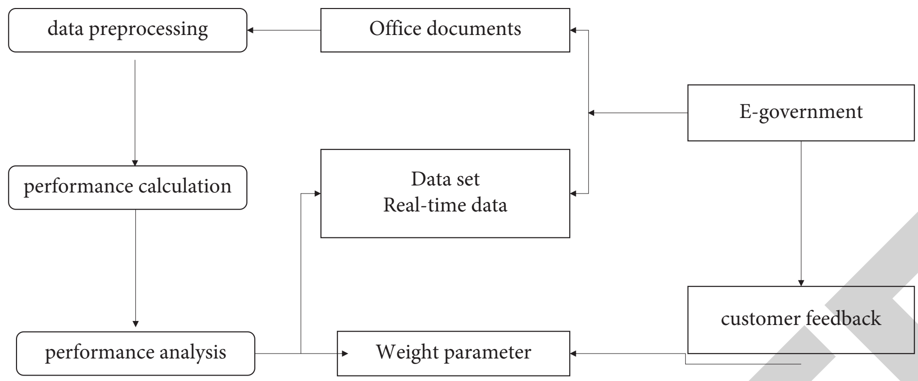


Figure 6: System design framework diagram.

disadvantage is that its scalability is low and the analysis cost is high. The performance-assisted analysis system based on enterprise administrative management content combines structured and unstructured language processing. Based on the analysis of the modules and functions of the existing company performance evaluation system, a comprehensive, fair, and efficient performance evaluation system is formulated to objectively evaluate the contribution of employees to the performance of the enterprise, and also provides a basis for the salary level of employees [34].

The performance-assisted analysis system based on text analysis is based on the company's performance evaluation system (automated office system), with the help of support platform technology, combined with the business management and data flow in the original performance evaluation system, to develop a new performance evaluation function module and add it to the original performance evaluation system. In the evaluation system platform. Comprehensive analysis of the official process of the performance evaluation system, enterprise administrative management activities, and employee performance management file information processing are connected to the activities of various departments and are closely related to the office of each department. The electronic enterprise administrative management office mainly involves corporate administrative management activities [35]. Employee performance management file information management, and enterprise administrative management are the main tasks of the three modules. The statistics of these work contents are directly related to the calculation of individual and departmental work performance.

In the integration with the original performance evaluation system, the overall structure of the system is defined in the structure diagram of the interface between the performance evaluation system and the original performance evaluation system, data flow, source data reading method, result data storage method, and parameter adjustment method (Figure 6).

The performance analysis tasks of the performance evaluation system mainly include. First, run the system at a fixed time every day to calculate the real-time data of the performance statistics of the day; in the database and each table of the database, the performance evaluation system based on text analysis can read the relevant database field information in the original performance evaluation system from the database, so as to analyze the relevant data of the original performance evaluation system; third, for a large number of documents in the original performance evaluation system, the performance evaluation system first reads the document storage path from the relevant fields of the database, and then searches for the required documents in the document storage space according to this path, avoiding the tediousness of obtaining a large number of document sets at one time [36].

System data is structured record data stored in an SQL database. The offline processing stage requires the use of all records of the relevant field in the relevant table. In the online processing stage, the data is transmitted to the system by querying the database in real-time, and the system obtains the corresponding individual and partial performance values by reading the corresponding fields of the database [37]. The input data mainly includes the enterprise administrative management activity module in the automated office system, the relevant records involved in the enterprise administrative management module, the relevant records of the employee performance management file information statistics (to-do, in-progress, and processed) and employee performance management. Office documents involved in the file information management module.

For the input of the enterprise administrative management activity module, the input contents include the enterprise administrative management number (ID) in the enterprise administrative management activity table, the enterprise administrative management initiator (ITEM_ REGISTER), and the enterprise administrative management participation (ITEM_ COMMISSIONLEADER), etc. The main input data of the management module is the related fields involved in the process of employee performance management file information processing and circulation, such as current processing personnel (ACTORID), processing time (PROCESSTIME), circulation status (FLOSTATE_ ID), processing opinions (ATTITUDE) Wait. For the acquisition of Office documents, we first obtain the absolute path (SUB_ DOC_ CONTENT) stored in each document from the database field, find the two documents to be calculated according to the personal modification record and calculate the edit distance; the employee performance management file information statistics module input. The data are valid fields related to the processing of pending documents, documents in progress, documents to be executed, and identification [38].

The output of the system is the quantified value of the workload of each person and each department in a given time period and its specific components. For the offline processing stage, the output data is saved in a structured text file (recommended CSV file structure). In the online processing stage, the system sets up scheduled tasks, runs the system regularly every day without affecting the operating efficiency of the original system, and returns the results to the automated office system in real-time.

The system output data is stored in eight tables in the database pasboc, each table stores the field information corresponding to different implementation functions, such as employee ID (USER_ID), department (CDEFAULTDEPARTMENT), role type (DOCACT_CREATOR), role weights (DOCACT_CREATO_R), delay time (OP_RE_INTERVAL), processing time (CPR_OP_INTERVAL), employee performance management file information execution process map parameters (MAP_OPTION).

3.4. Algorithm Function Module Analysis. In the performance evaluation system, the processing process of employee performance management file information includes the circulation process of drafting, issuing, receiving, forwarding, etc., and in different stages of circulation and between different nodes in the same stage, the processing content and content of employee performance management file information by different handlers. The processing methods are different. How to effectively identify the work content and work performance of different enterprise administrators in the information flow of employee performance management files is a key link in the construction of the performance evaluation system. Different from the traditional performance evaluation model, the performance evaluation model proposed in this paper combines the unstructured natural language in the performance evaluation system, that is, the "opinion" short text information in the management of employee performance management file information, combined with the flow of employee performance management file information before and after. It uses machine learning, natural language processing, and other related technologies to mine effective information in the text and analyze the association rules between different information, and identify the work of different personnel in the process of employee performance management file information flow [39]. Types handle "roles". At the same time, an execution state diagram model of employee performance management file information flow is established, each key node in the employee performance management file information flow is tracked in real-time, and the relationship between each node is identified. Based on the business relationship and data flow analysis of the electronic enterprise administrative management office system, the employee performance management file information handling roles designed in the system are divided into five categories: creator, approver, reviewer, executor, and forwarder.

The enterprise administrative management arrangement module of the company's performance evaluation system involves multiple participants in the activity, and in different enterprise administrative management activities and different stages of the same enterprise administrative management activity, each participant is responsible for different activity tasks, ranging from complex ones. Effectively identifying the division of duties in the process of enterprise administrative management arrangement can be used as an important reference for performance evaluation. The data involved in the enterprise administrative management arrangement is semi-structured text data. By comprehensively analyzing the workflow and data storage characteristics of the enterprise administrative management arrangement, different job roles and the mutual working relationship between different roles can be extracted [40]. The role types in the management arrangement are divided into the following two categories: the initiator of the enterprise administrative management arrangement and the participant of the enterprise administrative management arrangement.

The processing of employee performance management file information in enterprise administration revolves around the creation of employee performance management file information, the processing and forwarding of employee performance management file information, and the processing of employee performance management file information including the addition and deletion of employee performance management file information documents., modification, in each node of employee performance management file information flow, it may involve the processing of employee performance management file information by enterprise administrators. The workload in employee performance management file information processing.

Taking the amount of editing of the employee performance management file information document by each person in the process of employee performance management file information processing as a rigid indicator, it is calculated that each office employee receives the employee performance management file information from the employee performance management file information in the process of employee performance management file information flow [41, 42]. The amount of editing between two employee performance management file information versions (that is, the number of modified characters in the document) for information forwarding the statistics of the editing volume of performance management file information is used as a quantitative reference index in the quantitative calculation of performance evaluation [43, 44].

The performance-assisted analysis system based on the electronic office system combines the business flow, office process, structured and unstructured enterprise administrative management data flow process, and interrelationship in the performance evaluation system, and modularizes the business office. Analyzing and processing structured and unstructured data, mining key information related to the work content and work type of enterprise administrative personnel in text data, and establishing a role type based on enterprise administrative management arrangement work relationship identification, employee performance management file information management. The comprehensive performance real-time evaluation model of statistics, document editing and processing statistics, the realization of each model.

4. Conclusions

The performance evaluation system based on the composite element performance evaluation algorithm proposed in this study is oriented to the existing company performance evaluation system, comprehensively compares the current relatively perfect performance evaluation methods and systems, and combines the data characteristics of the research field. Based on the system, a performance auxiliary analysis function subsystem is realized. By analyzing the structured and unstructured related data of related modules such as enterprise administrative management activities, employee performance management file information processing, employee performance management file information flow records, etc., analyze the roles, tasks, and task completion of enterprise administrators in related matters. On this basis, a quantitative model of the workload of enterprise administrators is established, so as to establish a performance evaluation system to count the quantitative workload of individuals and departments within a given time range.

Facing the existing company performance evaluation system, it builds a performance evaluation system by tracking the business office implementation process and data flow analysis and realizes the quantitative and qualitative analysis of the performance of the corporate administrative personnel. Taking the performance evaluation system as a reference, different from the traditional performance evaluation method, we track and analyze the structured and unstructured data in the performance evaluation system, and use machine learning, natural language processing, and other technologies to analyze the text information. At the same time, a performance evaluation model based on each work module is established to quantitatively calculate and count employee work performance. To sum up, this study mainly studies the following aspects of the system design.

First, the performance evaluation algorithm based on a genetic algorithm and fuzzy comprehensive performance evaluation algorithm is introduced, and then the advantages and disadvantages of the performance evaluation algorithm are compared. Through the company's employee performance evaluation and assessment database to study the performance evaluation algorithm mining objects, make a performance evaluation system and framework demand analysis, including: performance evaluation system demand analysis, system performance evaluation algorithm overall demand analysis, system input and output data format. The algorithm is applied to performance evaluation, and the algorithm requirements of employee management module, system management module, and performance management module are analyzed respectively.

Based on the existing algorithms, the fuzzy performance evaluation algorithm based on compound elements is studied, and the detailed design of the algorithm is given. Through the comparison and analysis of the performance evaluation algorithm based on compound elements and the performance evaluation algorithm of specific elements, the superiority of the performance evaluation algorithm based on compound elements is experimentally verified by comparing the operation time and classification accuracy.

The performance evaluation algorithm performance evaluation system based on composite elements designed in this paper, combined with the office business, office process, structure, and unstructured enterprise administrative management data transfer process and interrelationship of the performance evaluation system, modularized the enterprise administrative management activities. Digitize the electronic office information, dig deep into the unstructured natural language, and establish a real-time performance evaluation system based on the arrangement of enterprise administrative management activities, employee performance management file information processing, document content processing, and employee performance management file information statistics. By designing and implementing a performance-assisted analysis system based on text content analysis, which is suitable for performance evaluation system, it solves the imperfect problem of performance evaluation based on electronic enterprise administrative management.

The data storage involved in the company performance evaluation system includes both structured data in the form of database table fields and unstructured natural language that does not depend on the table structure. There is a clear hierarchical structure between structured information data and data, and there are certain operating specifications. It is easy to analyze but has limitations in content. Unstructured information achieves human-computer interaction to a certain extent and is convenient for users to use, but the disadvantage is that It has low scalability and high analysis cost.

The performance-assisted analysis system based on the performance evaluation algorithm of composite elements is based on the electronic office cloud platform of the Science and Technology Innovation Committee—performance evaluation system (automated office system), with the help of the support platform technology, combined with the business management in the original performance evaluation system And data flow, develop new performance evaluation function modules to add to the original performance evaluation system platform. Comprehensive analysis of the office process of the performance evaluation system, enterprise administrative management activities, and employee performance management file information processing are connected to the activities of various departments and are closely related to the office of each department. The electronic enterprise administrative management office mainly involves corporate administrative management activities. Employee performance management file information management and enterprise administrative management are the main tasks of the three modules. The statistics of these work contents are directly related to the calculation of individual and departmental work performance.

On the basis of focusing on the analysis of business process and data flow in enterprise administrative management activities and employee performance management file information processing, this study analyzes the unstructured natural language and relevant information mining, and establishes a performance evaluation model based on various evaluation indicators, through the assignment of weights to different job roles and real-time data analysis, to establish a performance evaluation system that can quantitatively and qualitatively analyze the work of employees, and evaluate the results objectively and efficiently, to achieve a comprehensive evaluation of individual and departmental performance in the enterprise.

Data Availability

The experimental data 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 to report regarding the present study.

References

[1] G. D. Ilyushin and V. A. Blatov, “Cluster self-organization of intermet allic systems: Nanoclu sterprecursors and self-assembly of crystalstructures,” *Russian Journal of Inorganic Chemistry*, vol. 56, no. 2, pp. 12-13, 2011.

[2] M. Schain, M. Tóth, Z. Cselényi et al., “Quantification of serotonin transporter availability with [11C]MADAM--a comparison between the ECAT HRRT and HR systems,” *NeuroImage*, vol. 60, no. 3, pp. 601-602, 2012.

[3] L. Bass, P. Kazman, and P. Clements, *Software architecture in Practice*, Addison-Wesley Educational Publishers Inc, 3rd, pp. 1–355, 3rd edition, 2014.

[4] M. L. Jenkins and G. Lloyd, “Howcorporate philosophy and strategy . Shape theuse of HR information systems,” *Personnel*, vol. 62, no. 5, pp. 28–38, 2011.

[5] S. B. Roach, “Future system to streamline Army HR and Soldierpay,” *EN*, vol. 61, no. 3, pp. 122–124, 2012.

[6] A. Nagendra and M. Deshpande, “Human Resource Information systems (HRIS) in HRP lanning and development in Midto large Sized organizations,” *Procedia-Social and Behavioral Sciences*, vol. 133, no. 133, p. 55, 2014.

[7] R. David, “Gettingon schedule. Hospitalsare dumping manual scheduling for automate dsystems thattie intoother HR software,” *Healthcare Informatics*, vol. 24, no. 12, pp. 55-56, 2011.

[8] B. Kapoor and J. S. Sherif, “Global human resources (HR) information Systems,” *Kybernetes*, vol. 41, p. 41, 2012.

[9] T. Rollins, “Pay for performance: Is it worth the trouble?” *Personnel Administrator*, vol. 33, no. 5, pp. 42–46, 1988.

[10] Advisory, *Committee to the National Science Foundation Shaping the Future*, National Press Club, no. 10, p. 36, 1996.

[11] E. Rasmussen, “Games and information: introduce to Game theory,” *Blackwell Personnel*, vol. 1, no. 1, pp. 35–39, 1989.

[12] R. S. Kaplan and D. P. Norton, “The balanced score card measures that drive performance,” *Harvard Business Review*, vol. 69, no. l, pp. 71–79, 1992.

[13] J. C. Paradi, S. Smith, and C. Schaffnit-Chatterjee, “Knowledge worker performance analysis using DEA: an application to engineering design teams at bell Canada,” *IEEE Transactions on Engineering Management*, vol. 49, no. 2, pp. 161–172, 2002.

[14] T. del Rosario, “Report on staff performance assessment and enhancement systems of the agency for agricultural research and development,” *Asian Institute of Management*, vol. 6, pp. 65–72, 2003.

[15] K. Moser, H. Schuler, and U. Funke, “The moderating effect of raters’ Opportunities to Observe Ratees’ job performance on the Validity of an assessment Centre,” *International Journal of Selection and Assessment*, vol. 7, no. 3, pp. 133–141, 1999.

[16] C. E. Miller and C. L. Thornton, “How accurate are Your performance appraisals?” *Public Personnel Management*, vol. 35, no. 2, pp. 153–162, 2006.

[17] M. Armstrong and A. Baronl, *Performance Management [M]*, The Cromwell Press, London, 1998.

[18] D. H. Flint, “The role of organizational justice in multi-source performance appraisal: theory-based applications and directions for research,” *Human Resource Management Review*, vol. 9, no. 1, pp. 1–20, 1999.

[19] W. J. Smith, K. V. Harrington, and J. D. Houghton, “Predictors of performance appraisal Discomfort,” *Public Personnel Management*, vol. 29, no. 1, pp. 21–32, 2000.

[20] G. Blua, “Testing the longitudinal impact of work variables and performance appraisal satisfaction subsequent overall job satisfaction,” *Human Relations*, vol. 52, no. 8, pp. 1099–1113, 1999.

[21] K. Kuvaas, “Performance appraisal satisfaction and employee outcomes: mediating and moderating roles of work motivation,” *International Journal of Human Resource Management*, vol. 17, no. 3, pp. 504–522, 2006.

[22] S. Lyubomirsky, L. King, and E. Diener, “The benefits of Frequent Positive affect: does Happiness lead to Success?” *Psychological Bulletin*, vol. 131, no. 6, pp. 803–855, 2005.

[23] T. A. Wright and S. E. Hobfoll, “Commitment, psychological well-Being and job performance: an examination of conservation of resources (COR) theory and job burnout,” *Journal of Business and Management*, vol. 9, no. 4, pp. 389–406, 2004.

[24] R. Abraham, “Emotional intelligence in organizations: a conceptualization,” *Genetic, Social, and General Psychology Monographs*, vol. 125, no. 2, pp. 209–224, 1999.

[25] F. L. Sehnlidt, “The role of general cognitive ability and job performance: why there cannot be a debate,” *Human Performance*, vol. 15, pp. 187–210, 2002.

[26] A. L. Kristof, “Person-organization fit: an integrative review of its conceptualizations, measurement, and implications,” *Personnel Psychology*, vol. 49, no. 1, pp. 1–49, 1996.

[27] R. Libby, “Availability and the generation of hypotheses in analytical review,” *Journal of Accounting Research*, vol. 23, no. 2, pp. 648–667, 1985.

[28] C. P. Parker, B. Baltes, S. A. Young et al., “Relationships between psychological climate perceptions and work outcomes: a meta-analytic review,” *Journal of Organizational Behavior*, vol. 24, no. 4, pp. 389–416, 2003.

[29] J. Pfeffer, “Six dangerous myths about pay,” *Harvard Business Review*, vol. 5, pp. 109–119, 1998.

[30] L. Darling-Hammond, S. P. Newton, and R. C. Wei, “Developing and assessing beginning teacher effectiveness: the potential of performance assessments,” *Educational Assessment, Evaluation and Accountability*, vol. 25, no. 3, pp. 179–204, 2013.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9787416, 1 page
https://doi.org/10.1155/2023/9787416

Retraction

Retracted: Quality Evaluation of College Employment Based on Fuzzy Comprehensive Evaluation and Immersive Virtual Realization Technology

Mobile Information Systems

Received 13 September 2023; Accepted 13 September 2023; Published 14 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.

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] Y. Ji and Y. Luo, "Quality Evaluation of College Employment Based on Fuzzy Comprehensive Evaluation and Immersive Virtual Realization Technology," *Mobile Information Systems*, vol. 2022, Article ID 6285591, 13 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 6285591, 13 pages
https://doi.org/10.1155/2022/6285591

Research Article

Quality Evaluation of College Employment Based on Fuzzy Comprehensive Evaluation and Immersive Virtual Realization Technology

Yahui Ji and Ye Luo

Civil Aviation Flight University of China, Deyang, Sichuan, China

Correspondence should be addressed to Yahui Ji; jiyh@cafuc.edu.cn

Received 19 June 2022; Revised 20 July 2022; Accepted 22 July 2022; Published 23 August 2022

Academic Editor: R. Mo

In order to further understand the employment forms, employment dynamics, and employment competitiveness of college students, this paper puts forward an employment quality evaluation model based on fuzzy comprehensive evaluation and immersive virtual realization technology in combination with the employment needs and environment of college students. Through horizontal and vertical comparison, field research, and employment quality evaluation model, this paper deeply understands the employment level of college students. From the perspective of the evaluation results of the employment-related work of the 2020 college students, the excellent employment rate has reached 86.25%. From the perspective of the employment quality trend of graduates, the graduate employment quality score is 77.08, and the undergraduate and junior college students' scores are 74.97 and 77.66, respectively. From the perspective of the excellent employment rate, the employment quality evaluation model proposed in this paper is helpful for the employment of college students.

1. Introduction

Immersive virtual reality is a simulation that generates an interactive and immersive real world on a computer by using graphic systems and various control interface devices; that is, a three-dimensional virtual world is generated by using computer simulation to provide users with a real-world simulation of visual, auditory, tactile, and other senses in the virtual environment, as shown in Figure 1 [1]. Traditionally, previous studies on employment usually focused on the amount of employment, such as the employment rate and unemployment rate. However, they seldom pay too much attention to the "quality" of employment, so in the long run, the employment rate has naturally become a crucial standard in the research of traditional employment [2]. The employment rate has even become an important indicator to reflect the employment situation of a country and region. However, previous studies on the employment rate only reflected the quantity of employment and the number of undertakings through the use of simple labor statistical indicators and neglected the quality. This not only will seriously ignore the working environment of the labor process but also will not even consider factors such as working hours. Therefore, only focusing on the employment rate and unemployment rate cannot meet the needs of employment research [3]. Therefore, the aim of this paper is to focus on the quality of employment work on the basis of previous studies and to better improve the employment rate through the evaluation of the quality of employment work.

2. Literature Review

Some scholars have proposed that the quality of employment is a comprehensive category of good or bad conditions in which the means of production and workers are combined to obtain returns or income in the whole process of employment [4]. From the microscopic analysis, the quality of employment is actually centered on the workers and all the elements associated with it. From a macroperspective, the quality of employment is the different degrees of good and

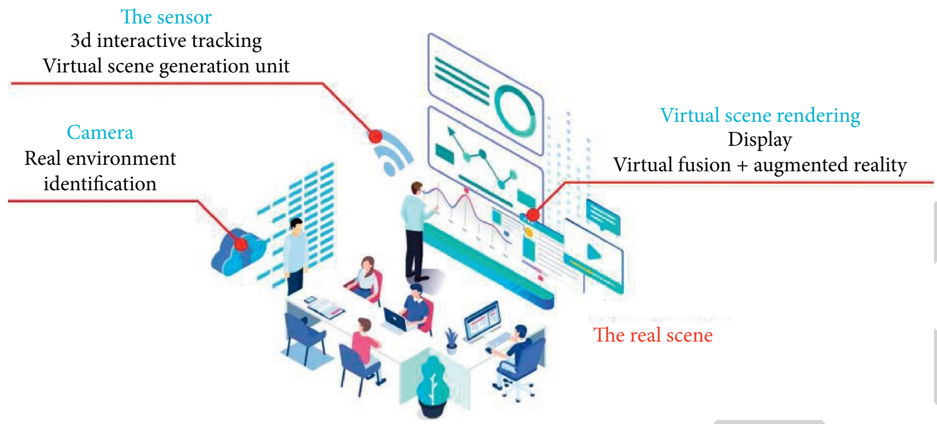


Figure 1: Immersive virtual reality.

bad that all workers in a region, industry, or country react to when they work [5]. In addition, it is believed that decent work covers two levels of employment quality and employment quantity in connotation, the so-called decent work is actually the good or bad level of employment quality, and the production work with high employment quality corresponds to decent work [6]. Some scholars argue that the quality of employment is actually a measure of the comprehensive employment status of workers in the process of completing their employment behavior [7]. The quality of employment includes nine aspects, namely, working conditions, health and welfare, equal opportunities, whether to enjoy social security, the nature of work, training and career prospects, labor relations, personal dignity, and safety. We think the employment quality not only refers to the state in the employment process but also includes the whole employment process itself. It argues that the evaluation of employment quality should come from two aspects: objective conditions and subjective satisfaction. The objective conditions refer to job stability, job opportunities, social reputation, development space of the industry, salary and welfare, and so on [8]. Subjective satisfaction refers to whether the job itself is satisfied, whether the workplace is satisfied, whether the conditions of the work unit are satisfied, and whether the salary is satisfied. Throughout many studies on employment quality by foreign scholars, almost all of them started from the research on the related concepts of employment quality, and many of them were inspired by the research on "decent work" [9]. There are many similarities between "decent work" and "quality of employment," but there are also many differences. Scholars have always stressed that "quality of employment" and "decent work" should not be equated [10].

3. Immersive Virtual Reality Technology Evaluation System

3.1. Determination of Evaluation Object. The immersive product virtual evaluation system is guided by the theory of perceptual engineering and evaluates products from the perceptual point of view of users. Therefore, it is more suitable for daily life products rather than functional technical products. Such products widely exist in daily life and are often observed and exposed by the public [11]. The specific product range includes household appliances, digital products, furniture lighting, storage ornaments, kitchen and bathroom supplies, and personal care, as shown in Figure 2.

3.2. Establishment of Evaluation Index System

3.2.1. Collection of Evaluation Indicators. The collection of evaluation indicators is carried out with the help of the perceptual vocabulary classification method, which is to use the hierarchical inference method to establish a tree-like analysis chart and gradually refine the details [12]. The implementation procedure is shown in Figure 3.

First, 0 times of perceptual vocabulary of the product shall be determined according to the market positioning and research of the product, and then multilayer perceptual vocabulary shall be gradually deduced to complete the index collection. The specific process is shown in Figure 4 [13]. In the process of collecting indicators, we should make the vocabulary more representative and comprehensive and not miss the important perceptual vocabulary.

3.2.2. Screening and Classification of Evaluation Indicators. After the indicators are collected, it is necessary to simplify and screen the collected indicators, eliminate unnecessary indicators, classify and sort the indicators, and establish an evaluation indicator model [14]. The following points should be done when screening indicators: First, the number of perceptual words to be screened should be controlled within the critical value. Second, the problem design should focus on the expression of the evaluator's perceptual will, rather than the understanding of the design elements. The problem should be simple and easy to understand without consuming the evaluator's energy. Third, the selection of evaluators should focus on the people who have demand for products and have close contact with products, subdivide them, and do a good job in user positioning [15]. The importance of the collected perceptual indicators is scored by 40 evaluators using the Likert scale method. There are five grades of 1–5, corresponding to "completely unimportant," "unimportant,"

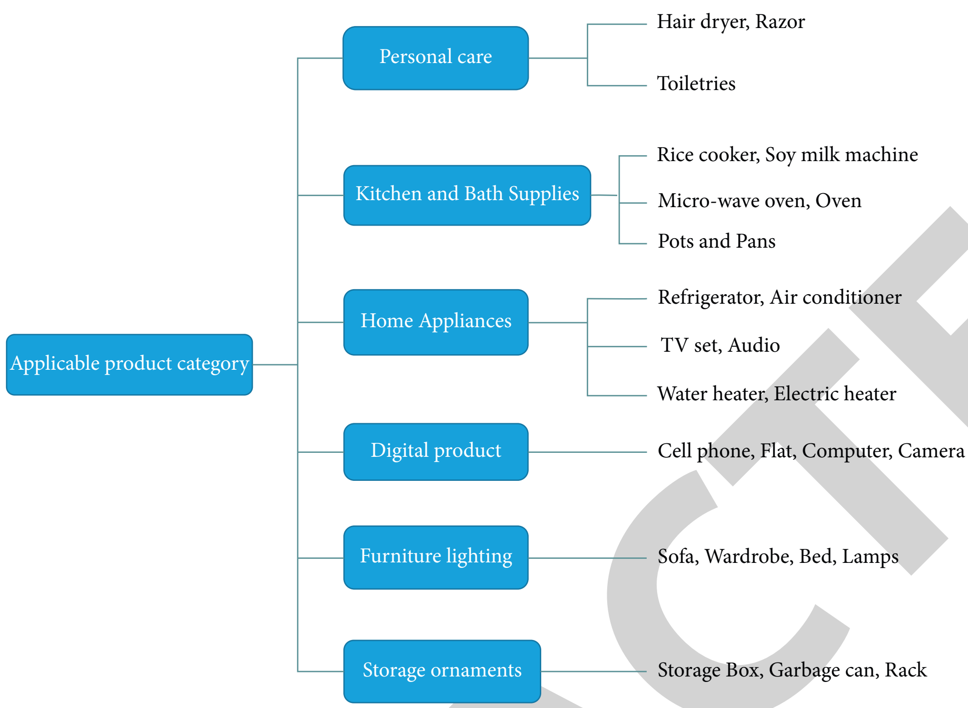


Figure 2: Applicable product categories.

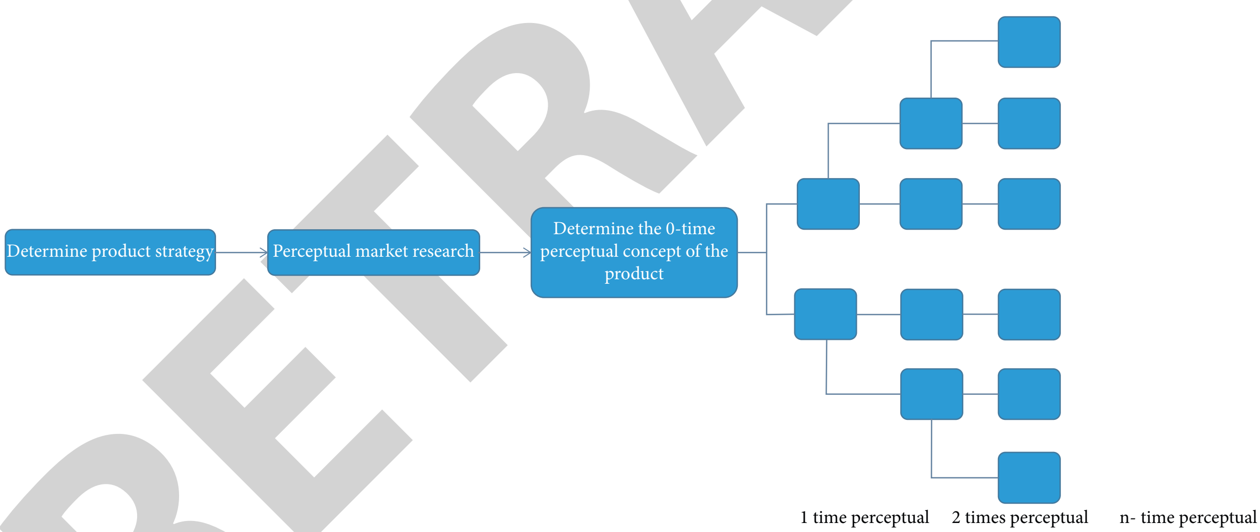


Figure 3: Implementation diagram of perceptual vocabulary classification.

"general," "important," and "very important." The specific form is shown in Table 1.

Import the survey data into SPSS software for descriptive analysis, calculate the total score and average value, and get the statistical analysis table of perceptual indicators. The emotional indicators are arranged by the mean value of scores, and a critical value is selected to divide the top n items with high average scores as the emotional evaluation indicators; that is, the evaluators believe that these indicators are very important or more important factors for the design evaluation of such products [16]. Using the Likert scale method, the correlation degree between the perceptual indicators and the evaluation angle is divided into five grades: "very relevant," "relatively relevant," "general," "relatively unrelated," and "completely unrelated," and the corresponding score is 5–1 points. Take the survey form of product appearance as an example, and the specific form is shown in Table 2.

Summarize the survey results, input them into SPSS software for data processing, obtain the average value of the correlation coefficient between each evaluation index and each evaluation angle, and generate the average score table of the correlation degree of the evaluation index, as shown in Table 3.

Use SPSS software to perform cluster analysis on the data in the table, check and draw the tree view, and get the relevant summary table, evaluation index cluster table, and

Figure 4: Flowchart of indicator collection.

Table 1: Index screening questionnaire.

Perceptual index	Completely unimportant (1 point)	Unimportant (2 points)	General (3 points)	Important (4 points)	Very important (5 points)
Simple and generous					
		. . .			
Fashion trends					

Table 2: Evaluation index correlation questionnaire.

Angle	Correlation degree				
	Completely unrelated (1 point)	Less relevant (2 points)	Average (3 points)	Relatively relevant (4 points)	Very relevant (5 points)
Indicator 1					
. . .					
Indicator n					

Table 3: Average data of correlation degree of evaluation indicators.

Evaluation index	Appearance modeling angle	Color material angle	Usability perspective	Human factors engineering
Indicator 1	Average value	. . .	. . .	Average value
		. . .		
Indicator n	Average value	. . .	. . .	Average value

tree view. In the tree view, different evaluation indicators will be classified into a certain evaluation angle category [17]. Sort out the evaluation indicators according to categories to get the evaluation indicator model, as shown in Figure 5.

3.2.3. Calculate Evaluation Index Weight. The weight of indicators is calculated by quantitative statistical method. The importance of the selected evaluation indicators is investigated and divided into five levels: unimportant, less important, important, more important, and very important. 67% of the three levels of "important," "very important," and "very important" are selected as the proportion of the limit [18]. The following formula is used:

$$\delta_i = \frac{a_i}{\sum_{i=1}^{n} a_i}, \qquad a_i = \sum_{j=1}^{3} L_j C_{ij}, \tag{1}$$

where j stands for "important, relatively important, and very important," $j = 3$;

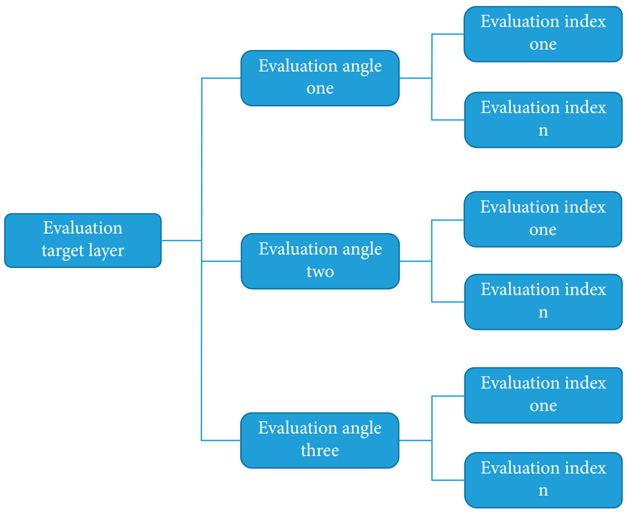


FIGURE 5: Evaluation index model.

$$L_1 = \frac{3}{3+4+5} = 0.25, \quad (2)$$

$$L_2 = \frac{4}{3+4+5} = 0.33, \quad (3)$$

$$L_3 = \frac{5}{3+4+5} = 0.42. \quad (4)$$

Equation (2) represents the proportion of general important categories; equation (3) represents the proportion of more important categories; equation (4) represents the proportion of very important categories; a_i, the weight value δ_i of index i, is obtained by normalization, where

$$\sum_{i=1}^{n} \delta_i = 1, 0 \le \delta_i \le 1\, (i = 1, 2, \ldots, n). \quad (5)$$

Calculate the weight of each evaluation index of the product, and adjust it appropriately according to the design focus of the product to obtain the final weight of each evaluation index [19]. The weight of each evaluation angle is the sum of the weights of each evaluation index under it, and the weight value of each evaluation angle is added to 1.

4. Fuzzy Comprehensive Evaluation of the Quality of College Students' Employment

4.1. Construction of Quality Evaluation Index System for College Students' Employment. According to the ABC method, after referring to relevant literature, the indicators are divided into three groups to form an indicator system for the quality of college students' employment. The weights of each indicator are assigned by using the expert scoring method, as shown in Table 4.

4.2. Introduction to Fuzzy Comprehensive Evaluation Model. Fuzzy comprehensive evaluation can be divided into single level and multipole fuzzy comprehensive evaluation according to the division of factor set. In order to illustrate the problem considering the complexity and operability of operation, this paper uses two-level partition and fuzzy comprehensive evaluation. According to the fuzzy comprehensive evaluation and the above-established evaluation index system for the employment quality of college graduates, an evaluation model is established [20].

(1) Determine the evaluation index set. According to the above evaluation index system for the employment quality of college graduates, we can see the following:

Primary indicators include

$$X = (x_1, x_2, x_3). \quad (6)$$

Secondary indicators include

$$x_1 = (x_{11}, x_{12}, x_{13}, x_{14}, x_{15}, x_{16}, x_{17}, x_{18}, x_{19}). \quad (7)$$

Similarly, we can get

$$\begin{aligned} x_2 &= (x_{21}, x_{22}, x_{23}, x_{24}, x_{25}, x_{26}, x_{27}, x_{28}, x_{29}), \\ x_3 &= (x_{31}, x_{32}, x_{33}, x_{34}, x_{35}, x_{36}, x_{37}, x_{38}). \end{aligned} \quad (8)$$

(2) Determine the weight of each indicator layer. It can be seen from the above that the weight of each indicator is

$$\begin{aligned} a &= (a_1, a_2, a_3) = (0.2, 0.5, 0.3), \\ a_1 &= (a_{11}, a_{12}, a_{13}, a_{14}, a_{15}, a_{16}, a_{17}, a_{18}, a_{19}) \\ &= (0.1, 0.1, 0.05, 0.1, 0.2, 0.05, 0.2, 0.05, 0.15), \\ a_2 &= (a_{21}, a_{22}, a_{23}, a_{24}, a_{25}, a_{26}, a_{27}, a_{28}, a_{29}) \\ &= (0.2, 0.1, 0.1, 0.05, 0.05, 0.1, 0.15, 0.05, 0.2), \\ a_3 &= (a_{31}, a_{32}, a_{33}, a_{34}, a_{35}, a_{36}, a_{37}, a_{38}) \\ &= (0.2, 0.1, 0.1, 0.2, 0.05, 0.1, 0.05, 0.2). \end{aligned} \quad (9)$$

(3) Determine the decision set. The decision set of the previous indicators is expressed as

$$\begin{aligned} V &= (V_1, V_2, V_3, V_4, V_5) \\ &= (\text{very good, good, generally, poor}). \end{aligned} \quad (10)$$

Four-level rating system, through questionnaire survey, is carried out to clarify the distribution of index values and then determine $R_i, i = 1, 2, 3$.

$$R_1 = \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ r_{21} & r_{22} & r_{23} & r_{24} \\ \cdots & \cdots & \cdots & \cdots \\ r_{91} & r_{92} & r_{93} & r_{94} \end{bmatrix}. \quad (11)$$

The above formula is the evaluation membership matrix of the employer on the school evaluation.

$$R_2 = \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{13} \\ r_{21} & r_{22} & r_{23} & r_{24} \\ \cdots & \cdots & \cdots & \cdots \\ r_{91} & r_{92} & r_{93} & r_{94} \end{bmatrix}. \quad (12)$$

Table 4: Evaluation index system of college students' employment quality.

Evaluation index system of college students' employment quality	Evaluation of the employer on the school (x_1) (0.2)	Staff service attitude (x_{11}) (0.1)
		Style of study and school spirit (x_{12}) (0.1)
		School hardware (x_{13}) (0.05)
		Communication and cooperation with units (x_{14}) (0.1)
		Teaching practice link (x_{15}) (0.2)
		School popularity (x_{16}) (0.05)
		Teaching and management level (x_{17}) (0.2)
		Satisfaction with employment services (x_{18}) (0.05)
		Satisfaction with specialty setting and curriculum system (x_{19}) (0.15)
	Evaluation of employers on graduates (x_2) (0.5)	Executive force (x_{21}) (0.2)
		Professional foundation (x_{22}) (0.1)
		Development potential (x_{23}) (0.1)
		Adaptability (x_{24}) (0.05)
		Comprehensive quality (x_{25}) (0.05)
		Work style and professionalism (x_{26}) (0.1)
		Practical ability (x_{27}) (0.15)
		Innovation ability (x_{28}) (0.05)
		Organization, communication, and coordination (x_{29}) (0.2)
	Graduates' evaluation on school employment (x_3) (0.3)	Help of employment information to employment (x_{31}) (0.2)
		Hardware facilities and logistics services (x_{32}) (0.1)
		Study style construction (x_{33}) (0.1)
		Teaching management level (x_{34}) (0.2)
		Vocational guidance (x_{35}) (0.05)
		Curriculum (x_{36}) (0.1)
		Daily management and service (x_{37}) (0.05)
		Employment assistance satisfaction (x_{38}) (0.2)

The above formula is the evaluation membership matrix of the employer on the school graduates.

$$R_3 = \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ r_{21} & r_{22} & r_{23} & r_{24} \\ \cdots & \cdots & \cdots & \cdots \\ r_{81} & r_{82} & r_{83} & r_{84} \end{bmatrix}. \tag{13}$$

The above formula is the membership matrix of graduates' comments on the employment work of the school.

(4) Conduct fuzzy comprehensive evaluation.

The fuzzy comprehensive evaluation of secondary indicators shall be carried out first, and the evaluation method is

$$B_i = a_i * R_i, \tag{14}$$

where B_{ii} is the primary indicator and $i = 1, 2, 3$. There are three primary indicators. The composition matrix B is as follows:

$$B = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} & b_{15} \\ b_{21} & b_{22} & b_{23} & b_{24} & b_{25} \\ b_{31} & b_{32} & b_{33} & b_{34} & b_{35} \end{bmatrix}. \tag{15}$$

The first level evaluation of the index is carried out, and the matrix E is obtained. The method is

$$E = a * B. \tag{16}$$

(5) Conduct comprehensive evaluation.

According to the results of B, take the excellent value, that is, the sum of the first and second items. According to the principle of maximum membership, the greater the value, the better the quality of employment in colleges and universities and the higher the degree of recognition from all aspects [21].

4.3. Example Analysis of College Students' Employment Quality Evaluation

4.3.1. Sample Selection. The research example in this section is a certain electric power university. The questionnaire was sent through e-mail, WeChat, and so on using the existing electronic equipment and network channels. A total of 400 questionnaires were distributed, of which 394 were valid. The basic data of the respondents involved in this questionnaire, such as gender, graduation type, and registered residence before enrollment, are shown in Tables 5 and 6, respectively.

Table 5: Gender composition in the sample.

Gender	Number (person)	Percentage (%)
Male	205	52.0
Female	189	47.9
Total	394	100

Table 6: Statistics of sample graduation types.

Educational background	Number (person)	Percentage (%)
Graduate student	123	31.2
Undergraduate	156	39.5
Junior college students	115	29.1
Total	394	100

By analyzing the data collected in the above sample tables, we can see the following:

(1) There is little difference between the number of men and women in the sample, and the data obtained will not have gender tendency, which is more objective and reliable.

(2) The graduate groups involved in this survey mainly include graduate students, undergraduates, and junior college students, which are relatively uniform in terms of sample distribution.

There are 30 questions in this questionnaire, all of which are in the form of multiple-choice questions. The specific content is set according to all the indicators in the graduate employment quality evaluation system designed above [22]. At the same time, the following questionnaire result calculation tables are designed. When making questionnaire statistics, the questionnaire data are quantified according to the tables. As for the score setting of options, if the question has 4 options, the score distribution is 100%, 75%, 50%, and 25%, respectively. If the question has only 3 options, the score distribution is 100%, 66.667%, and 33.333%, respectively. As for the question on the way and way of employment in question 22, considering the small difference between "others' recommendation" and "others," the score of this question is allocated as 100%, 33.333%, and 33.333%.

4.3.2. Research Process. The questionnaire results of all primary indicators are summarized, and the secondary indicators and employment quality scores of graduate students, undergraduate students, and junior college students are analyzed by SPSS software. The results are shown in Table 7.

The questionnaire results of all primary indicators are summarized, and SPSS software is used to analyze the correlation between secondary indicators and employment quality scores of graduate students, undergraduate students, and junior college students. The results are shown in Table 8.

By analyzing the data at graduate level, the correlation between vocational skill training indicators and employment path indicators and employment quality is less than 0.3, and the significance is greater than 0.05, indicating that vocational skill training indicators and employment path indicators are not related to the employment quality of graduate level graduates; that is, the employment quality of graduate level graduates has nothing to do with vocational skill training and employment path [23]. The correlation values of occupational health indicators, working hours indicators, interpersonal relationship indicators, gender equity indicators, and work nature indicators are greater than 0.3 and less than 0.5, indicating that the above five indicators are weakly correlated with the employment quality of graduates at the graduate level; that is, the above five indicators have an impact on the employment quality of graduates at the graduate level, but the impact is weak. The correlation values of employment service quality index, promotion opportunity index, job stability index, and career belonging index are greater than 0.5 and less than 0.8, indicating that the employment quality of graduates at graduate level is positively correlated with the above indicators; that is, the employment quality of graduates at graduate level is affected by employment service quality, job promotion opportunity, job stability, and career belonging. Among all the secondary indicators of graduate level graduates, only the correlation value of employment location indicator, enterprise credit tolerance indicator, and salary reward indicator is greater than 0.8, indicating that the employment quality of graduate level graduates is highly related to the three indicators; that is, the employment quality of graduate level graduates is mainly affected by the employment location, the social reputation of the enterprise, and the salary [24].

By analyzing the data at the undergraduate level, the correlation value of employment path indicators is 0.067, less than 0.3, with a significance of 0.671, greater than 0.05, indicating that the employment path indicators are not related to the employment quality of graduates at the undergraduate level; that is, the employment quality of graduates at the undergraduate level has nothing to do with the way they obtain employment opportunities. The correlation values of employment service quality indicators and occupational health indicators are greater than 0.3 and less than 0.5, indicating that the above two indicators are weakly correlated with the employment quality of graduates at the undergraduate level. The above employment service quality and occupational health have an impact on the employment quality of graduates at the undergraduate level, but the impact is small. The correlation values of working time index, vocational skill training index, promotion opportunity index, employment location index, enterprise credit tolerance index, interpersonal relationship index, gender equity index, occupational belonging index, and work nature index are greater than 0.5 and less than 0.8. It shows that the above nine indicators are positively correlated with the employment quality of undergraduates. The above nine factors have a certain impact on the employment quality of graduates at the undergraduate level. By analyzing the data of junior college students, the correlation values of employment service quality indicators and employment path indicators are less than 0.3, and the significance values are greater than 0.05, indicating that the employment service quality indicators and employment path indicators have nothing to do with the employment quality of junior college students. The employment quality of college graduates has

Table 7: Descriptive analysis of secondary indicators and employment quality.

Arrangement	Primary index	Secondary index	Mean value	Standard deviation	N
Graduate level	Employment opportunities	Comprehensive employment rate	—	—	123
		Employment service quality	5.1641	0.67085	123
	Working conditions	Occupational health	2.5742	0.37279	123
		Working hours	2.1094	0.38593	123
		Vocational skills training	2.4531	0.29401	123
		Promotion opportunities	7.1352	1.70563	123
		Place of employment	2.2448	0.61585	123
	Labor relations	Enterprise trust	4.0820	0.74434	123
		Interpersonal relationship	5.9424	0.99904	123
		Working stability	4.4219	0.81922	123
	Employment equity	Employment path	1.7500	0.98374	123
		Gender equity	2.2656	0.55335	123
	Work value	Salary and remuneration	5.2734	1.54680	123
		Sense of professional belonging	4.8750	1.84308	123
		Nature of work	7.4609	1.38125	123
Undergraduate level	Employment opportunities	Comprehensive employment rate	—	—	127
		Employment service quality	1.8191	0.42334	127
	Working conditions	Occupational health	2.1037	0.45841	127
		Working hours	4.2819	0.67489	127
		Vocational skills training	3.5287	0.63166	127
		Promotion opportunities	5.5103	0.90060	127
		Place of employment	3.5079	0.46215	127
	Labor relations	Enterprise trust	3.5287	0.77052	127
		Interpersonal relationship	5.0634	0.74099	127
		Working stability	3.7945	0.59235	127
	Employment equity	Employment path	2.2910	0.67233	127
		Gender equity	2.0106	0.41032	127
	Work value	Salary and remuneration	4.2340	1.33860	127
		Sense of professional belonging	5.2483	1.67768	127
		Nature of work	4.3006	0.94089	127
Junior college level	Employment opportunities	Comprehensive employment rate	—	—	114
		Employment service quality	2.4107	0.46120	114
	Working conditions	Occupational health	5.7918	1.37279	114
		Working hours	2.9894	0.59409	114
		Vocational skills training	9.4531	1.29401	114
		Promotion opportunities	1.1352	0.36204	114
		Place of employment	2.2448	0.61585	114
	Labor relations	Enterprise trust	3.0820	0.54434	114
		Interpersonal relationship	1.9424	0.29904	114
		Working stability	2.4219	0.41922	114
	Employment equity	Employment path	3.7500	1.98374	114
		Gender equity	4.2656	0.85335	114
	Work value	Salary and remuneration	5.2734	1.54680	114
		Sense of professional belonging	3.8750	0.84308	114
		Nature of work	2.2609	0.38125	114

nothing to do with the quality of employment service and the way to obtain employment opportunities.

4.3.3. Evaluation and Analysis of Sample Employment Quality. According to the results of the questionnaire and the employment quality evaluation system established above, the employment quality scores of the graduates in the sample in 2020 can be obtained. The statistics are shown in Table 9 and Figure 6.

As can be seen from Figure 6, the employment quality of graduates at the graduate level in 2015 was the highest, 77.08 points. There is a small difference in the employment quality between graduates of undergraduate level and graduates of junior college level, which are 74.97 and 74.66 points, respectively, and there is a certain gap with graduates of graduate level.

4.3.4. Evaluation of the Employer on the Electric Power University. 150 units were contacted through questionnaires and other means in this survey. A total of 87 employers finally participated in this survey, mainly those who have come to the school for recruitment. They are generally familiar with the school, with a wide range of distribution, wide industry coverage, and strong representativeness. Investigate the employer according to the requirements of the index system. The survey data are shown in Table 10.

Table 8: Correlation analysis between secondary indicators and employment quality.

Arrangement	Primary index	Secondary index	Relevance	Significance	N
Graduate level	Employment opportunities	Comprehensive employment rate	—	—	123
		Employment service quality	0.512*	0.013	123
	Working conditions	Occupational health	0.402*	0.038	123
		Working hours	0.424*	0.023	123
		Vocational skills training	0.178	0.329	123
		Promotion opportunities	0.771**	0.002	123
	Labor relations	Place of employment	0.949*	0.012	123
		Enterprise trust	0.857*	0.033	123
		Interpersonal relationship	0.475**	0.006	123
		Working stability	0.662**	0.000	123
	Employment equity	Employment path	0.185	0.311	123
		Gender equity	0.358*	0.044	123
	Work value	Salary and remuneration	0.918*	0.016	123
		Sense of professional belonging	0.572**	0.000	123
		Nature of work	0.460**	0.001	123
Undergraduate level	Employment opportunities	Comprehensive employment rate	—	—	127
		Employment service quality	0.302*	0.034	127
	Working conditions	Occupational health	0.465*	0.019	127
		Working hours	0.526*	0.025	127
		Vocational skills training	0.508*	0.026	127
		Promotion opportunities	0.625*	0.026	127
	Labor relations	Place of employment	0.507*	0.020	127
		Enterprise trust	0.724**	0.002	127
		Interpersonal relationship	0.519**	0.000	127
		Working stability	0.874**	0.010	127
	Employment equity	Employment path	0.067	0.671	127
		Gender equity	0.612*	0.011	127
	Work value	Salary and remuneration	0.845**	0.002	127
		Sense of professional belonging	0.508**	0.000	127
		Nature of work	0.611**	0.000	127
Junior college level	Employment opportunities	Comprehensive employment rate	—	—	114
		Employment service quality	0.139	0.386	114
	Working conditions	Occupational health	0.842**	0.001	114
		Working hours	0.636*	0.016	114
		Vocational skills training	0.815**	0.001	114
		Promotion opportunities	0.487*	0.027	114
	Labor relations	Place of employment	0.521**	0.002	114
		Enterprise trust	0.702*	0.011	114
		Interpersonal relationship	0.481**	0.020	114
		Working stability	0.438**	0.000	114
	Employment equity	Employment path	0.276	0.078	114
		Gender equity	0.587**	0.007	114
	Work value	Salary and remuneration	0.718**	0.006	114
		Sense of professional belonging	0.600**	0.000	114
		Nature of work	0.422**	0.010	114

N represents the number of samples.

Table 9: Comprehensive evaluation of employment quality.

Educational background	Employment quality (points)
Graduate student	77.08
Undergraduate	74.97
Junior college students	74.66

4.3.5. Evaluation of the Employer on the Graduates of the Electric Power University. In this survey, 150 units were contacted by issuing questionnaires and other means. A total of 87 employers finally participated in this survey, all of which are units from all walks of life that receive a large number of graduates from our university. They have a good understanding of the overall situation of our graduates. Investigate the employer according to the requirements of the index system. The survey data are shown in Table 11.

4.3.6. Evaluation of Graduates of the Electric Power University on Employment. The survey contacted 500 graduates of the year 2020 through questionnaires and other means. Finally, 442 graduates of the year 2020 participated in the survey, including 321 boys and 121 girls, including graduate and undergraduate students, covering all majors. The survey data are shown in Table 12.

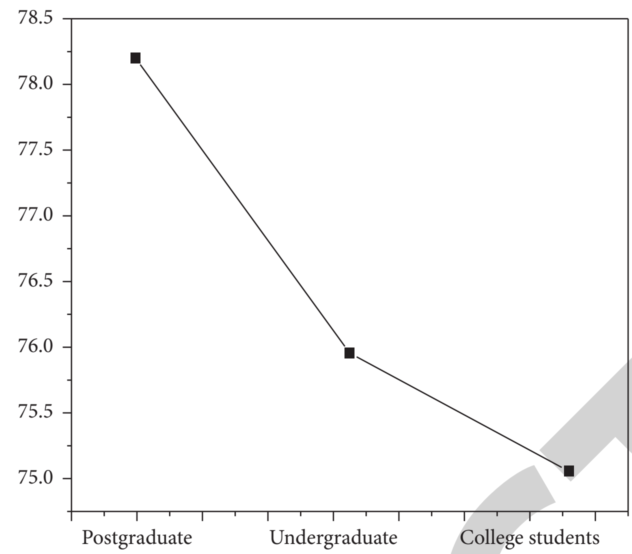


Figure 6: Employment quality trend of graduates.

Table 10: Evaluation of the employer on the electric power university

Evaluating indicator	Very good (%)	Good (%)	General (%)	Difference (%)	Total (%)
Staff service attitude	47.50	48.00	4.50	0.00	100.00
Style of study and school spirit	35.00	52.50	12.50	0.00	100.00
School hardware facilities	25.00	67.50	7.50	0.00	100.00
Communication and cooperation with units	17.50	65.00	17.50	0.00	100.00
Teaching practice	35.00	50.00	15.00	0.00	100.00
School popularity	42.50	55.00	2.50	0.00	100.00
Teaching and management level	27.50	65.00	7.50	0.00	100.00
Satisfaction with employment services	15.00	85.00	0.00	0.00	100.00
Satisfaction with specialty setting and curriculum system	57.50	42.50	0.00	0.00	100.00
Satisfaction with recruitment site, facilities, and equipment	37.50	60.00	2.50	0.00	100.00

Table 11: Evaluation of the employer on the graduates of the electric power university.

Evaluating indicator	Very good (%)	Good (%)	General (%)	Difference (%)	Total (%)
Executive power	21.43	78.57	0.00	0.00	100.00
Professional foundation	33.33	67.50	2.80	0.00	100.00
Development potential	26.19	71.43	2.50	0.00	100.00
Adaptability	21.43	73.81	4.76	0.00	100.00
Comprehensive quality	33.33	61.90	4.76	0.00	100.00
Work style and professionalism	27.50	69.05	7.50	0.00	100.00
Practical ability	15.00	71.00	10.22	1.68	100.00
Innovation ability	9.25	70.00	16.43	2.60	100.00
Organization, communication, and coordination skills	9.25	70.00	20.05	0.00	100.00

4.4. Analysis of Evaluation Results

(1) *Overall Evaluation*. From the above evaluation results, it can be seen that the excellent rate of graduates' employment in this electric power university is $0.2925 + 0.6084 = 0.9009$. It shows that the quality of the graduates' employment work of the electric power university is very high and has been basically satisfied in all aspects.

(2) *Specific Evaluation of All Aspects of Work*.

① According to the evaluation of the employer on the electric power university, employers have a high degree of recognition for the work related to school employment, and the excellent rate has reached 91.62% ($0.3229 + 0.5933$). According to the detailed data of each evaluation, there are 3 items whose evaluation indicators are satisfactory, good, high, and above, and the proportion is 100%, there are 4 items whose proportion exceeds 90%, and the proportion of the lowest 3 items exceeds 80%. It can be seen from the statistical chart that the employers are most satisfied with our school in three aspects, 100% of them think that the service attitude of the staff is good or above, the popularity of the school is high or

Table 12: Graduates' evaluation on employment.

Evaluating indicator	Very good (%)	Good (%)	General (%)	Difference (%)	Total (%)
Help of employment information to employment	45.43	48.62	5.78	0.36	100.00
Hardware facilities and logistics services	35.33	55.63	4.85	3.54	100.00
Study style construction	46.19	44.83	8.41	0.52	100.00
Teaching management level	42.43	47.39	6.95	2.82	100.00
Vocational guidance	58.33	31.68	6.84	3.15	100.00
Curriculum	40.50	49.43	9.88	0.48	100.00
Daily management and service	55.20	34.24	6.92	3.89	100.00
Employment assistance satisfaction	31.91	34.24	16.43	15.69	100.00

above, and the style of study and school spirit are good or above. Secondly, more than 90% of the students are satisfied with the employment service or above, the specialty setting and curriculum system are reasonable or above, the hardware facilities of the school are good or above, and the teaching and management level is high or above. The relatively low proportion is 87.5%, 85%, and 82.5%, respectively, which are satisfied with or above the recruitment site, facilities, and equipment, good at or above the communication and cooperation with the unit, and good at or above the teaching practice. In addition, the main factors considered by employers for recruitment to schools are the quality of students (77.5%), school running history (40%), and traffic factors (25%).

From the perspective of unit evaluation, first, the one-stop high-quality service of "site service, information release, image publicity, interview and negotiation, and contract signing guidance" that the school's employment guidance center has always adhered to has been widely recognized by the unit, effectively deepened the friendship with the majority of employers, and consolidated the employment market, and we need to continue to maintain and work hard. Second, in recent years, the popularity of our school has been significantly improved and has been recognized by many units, especially nonpower system units [25]. Thirdly, the excellent school spirit and learning style formed by the long-term school running history of the university are the highlights of our school running, and also one of the important factors that employers value and consider as recruitment factors.

② According to the overall evaluation [26–33] of the graduates of the electric power university by the employer, the excellent rate reached 93.63% $(0.2048 + 0.7315)$. This strongly proves that the educational idea formed by the school over the years is correct. The vast majority of the trained talents can meet the needs of the society and are widely recognized by the unit. This will undoubtedly provide a useful reference for the school's educational idea and development direction in the future. According to the statistical data of various evaluation indicators, most of the graduates have outstanding qualities, which are welcomed and affirmed by the units, such as executive ability, professionalism, work style, and professional knowledge. This is also a vivid reflection of our rigorous school running as a science and engineering college, which is worthy of our continued maintenance and further development.

The indicators highly recognized by the employer are executive ability, professional foundation, development potential, adaptability, comprehensive quality, work style, and professionalism. The proportion of "good + very good" is more than 95%, of which executive ability is 100%. The three indicators of hands-on practice ability, innovation ability, organization, communication, and coordination ability are relatively low, but the proportion is also more than 80%. With regard to "the aspects that graduates need to be improved and improved," the employers believe that the most important things that our graduates need to improve are organization, communication and coordination ability (69.05%), innovation ability (45.24%), and hands-on practice ability (35.71%). This also just confirms the three indicators with a low proportion of "good + very good." As for the "main reasons for accepting graduates," the most important factors considered by employers are professional counterparts (90.48%), solid professional foundation (64.29%), and high comprehensive quality (59.52%).

③ According to the evaluation of the graduates of 2020 on employment-related work, our employment work has made some achievements, and the excellent rate has reached 86.25% $(0.4191 + 0.4434)$. At the same time, we need to continue to improve and improve. The 2020 graduates have a relatively positive evaluation on employment and related work, in which they think that the employment information is of great help to employment, they are satisfied with or above the hardware facilities and logistics services, they have a good style of study, they have a high level of teaching management, and they are satisfied with or above the vocational guidance, all of which are more than 90%. The satisfaction with employment assistance was the lowest, only 66.15%.

At present, there are three main problems in the employment work: first, the overall expectation of fresh graduates is high, but the employment gap between graduates of various majors is large, and some fresh graduates of some majors think that the employment work of the university has not been fully taken into account; second, most students do not understand the employment responsibilities and relevant employment policies of the employment guidance center, colleges, and departments, which is not conducive to the development of employment management and service work; third, the school's assistance measures for the special employment groups and vulnerable groups need to be understood and felt by more students. Therefore, we need to attach great importance to the problems reflected in the evaluation of fresh graduates, strengthen publicity, narrow the employment gap between majors, improve the assistance system, and provide better services for graduates.

5. Conclusions and Suggestions

This research aims at the problem that the evaluation subject of the employment quality evaluation system of college graduates in China is single; that is, an evaluation system is only applicable to a certain part of graduates. This study proposes to divide the college graduates into three levels, namely, graduate students, undergraduate students, and junior college students, and independently design the employment quality evaluation system for college graduates at all levels. This study evaluates the employment quality of college graduates from five aspects: employment opportunities, working conditions, labor relations, employment equity, and work value. In addition, 15 secondary indicators have been set, including comprehensive employment rate, employment service quality, occupational health, working hours, vocational skills training, promotion opportunities, employment location, enterprise trust, interpersonal relationship, job stability, employment path, gender equity, salary, sense of occupational belonging and nature of work. This paper analyzes the objective facts and subjective feelings and constructs a three-level and five dimension evaluation system for the employment quality of college graduates. From the perspective of the evaluation results of the employment-related work of the 2020 college students, the excellent employment rate has reached 86.25%. From the perspective of the employment quality trend of graduates, the graduate employment quality score is 77.08, and the undergraduate and junior college students' scores are 74.97 and 77.66, respectively. From the perspective of the excellent employment rate, the employment quality evaluation model proposed in this paper is helpful for the employment of college students. From the evaluation results, the evaluation of employment quality in colleges and universities should pay attention to the combination of various evaluation methods. Nowadays, the commonly used evaluation method for the employment quality of college graduates is through questionnaires; that is, the graduates fill in the questionnaires themselves, hand them over to the school, and then hand them over to the relevant education authorities after the school makes preliminary statistics. Although this method has high operability, because there is only one evaluator, and it is the evaluator himself, the subjectivity of the results is serious, and it cannot fully reflect the employment quality of college graduates. When evaluating the employment quality of college graduates, we can adopt a combination of various evaluation methods. First of all, collecting data not only is limited to the questionnaire but also can adopt anonymous telephone communication to eliminate unnecessary concerns of graduates and understand the real situation of graduates' employment quality from the perspective of graduates themselves, independently design the employment scheme, so as to obtain the attention of graduates to various employment contents, and set a more scientific and reasonable evaluation weight for the employment quality of college graduates and Internet data collection, through a wider range of relevant data collection on the Internet, to strengthen the authority of the evaluation of the employment quality of college graduates, and other ways, to ensure that the data obtained are persuasive. As the terminal of employment behavior, the employment unit should also evaluate the employment quality of college graduates. The employment quality is the reflection of behavior in a period of time, not the immediate reflection of behavior. After the employment of graduates, the employment unit, through a period of contact and understanding, evaluates the results of its own employment of the graduates and timely provides the relevant information of the graduates after employment so that the relevant departments can obtain the required data more accurately and quickly when evaluating the employment quality of college graduates.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] B. Peng and X. X. Zheng, "Exploring the training mode of engineering disciplines in Chinese colleges and universities to meet the needs of an innovative society," *International Core Journal of Engineering*, vol. 6, no. 5, pp. 149–151, 2020.

[2] Y. Feng, "An evaluation method of pe classroom teaching quality in colleges and universities based on grey system theory," *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 6, pp. 6911–6915, 2020.

[3] D. Yan, "Analysis of the policy text of entrepreneurship education in Chinese colleges and universities (1998–2019)," *Open Journal of Social Sciences*, vol. 8, no. 7, pp. 218–231, 2020.

[4] D. Patimo, "Faculty performance evaluation system of state universities and colleges in the philippine eastern visayas region," *JISAE: Journal of Indonesian Student Assessment and Evaluation*, vol. 6, no. 2, pp. 159–167, 2020.

[5] W. Yu and X. Liu, "Behavioral risky multiple attribute decision making with interval type-2 fuzzy ranking method and topsis method," *International Journal of Information Technology and Decision Making*, vol. 21, no. 2, pp. 665–705, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 8404931, 9 pages
https://doi.org/10.1155/2022/8404931

Research Article

Machine Learning-Based Classification and Evaluation of Regional Ethnic Traditional Sports Tourism Resources

Cuihan Li and Shaojun Lyu

College of P.E. and Sports, Beijing Normal University, Beijing 100875, China

Correspondence should be addressed to Cuihan Li; 11132020301@bnu.edu.cn

Received 24 June 2022; Revised 21 July 2022; Accepted 28 July 2022; Published 21 August 2022

Academic Editor: R. Mo

Traditional sports in ethnic minority regions are a valuable cultural heritage. Regional ethnic traditional sports are not only a sports business but also a tourism resource. The construction of a reasonable regional sports tourism resource classification model is fundamental to the development of sports tourism resources. However, the existing sports tourism resources classification is mostly constructed manually based on the national standard tourism resources classification system. The efficiency and accuracy of the traditional manual classification are poor and cannot reflect the characteristics of regional ethnic traditional sports tourism. In order to solve the above problems, a machine learning-based classification method for regional ethnic traditional sports tourism resources is proposed. Firstly, the relevant concepts and characteristics of traditional sports tourism resources are introduced. Then, taking the development of traditional sports of ethnic minorities in Yunnan Province as the research object, SWOT analysis, literature, interview, questionnaire, and mathematical statistics are used to investigate and analyse the overall status of the development of regional ethnic traditional sports. Secondly, a classification evaluation method based on an optimised back-propagation (BP) neural network is proposed. Finally, the optimised BP neural network model is applied to the classification of traditional sports tourism resources. The experimental results show that the optimised BP model performs well in the classification of traditional sports tourism resources, verifying its effectiveness.

1. Introduction

Traditional sports in China's 55 ethnic minority regions are a valuable cultural heritage. Nowadays, with the rapid development of rural tourism, it is important to analyse and study traditional sports from the perspective of tourism resources to promote the development of regional ethnic traditional sports. China is rich in ethnic cultures and characteristic landscapes. Intangible cultural heritage has been well protected in China and has become an important tourism resource. However, the traditional sports of ethnic minorities contain rich ethnic cultural connotations and are also an important tourism resource.

For an ethnic minority, regional ethnic traditional sports contain not only the exercise of labour skills and combat skills, but also many elements of ethnic minority culture, art and religion. Traditional ethnic minority sports have a history of several hundred or even thousands of years. In recent years, the construction of rural tourism in China has been developing at a rapid pace. According to statistics, more than 85,000 villages with tourism as a development component have been built in China. Traditional sports of ethnic minorities have received great attention as an important tourism resource in cultural and ecological villages.

Sports tourism originated in the UK and has been widely concerned with the vigorous development of sports [1–5]. The global sports tourism industry is growing at a rapid rate of 14% per year and is expected to reach the USA $412.7 billion by 2022, with nearly USA $180 billion in the Asia-Pacific region. In China, there is a growing demand for diversified sports and tourism and leisure. Sports tourism has become an important embodiment of a healthy lifestyle [6, 7]. In recent years, China has issued a series of important documents and policies to actively promote the development of the sports tourism industry. It is expected that by 2022, the total number of sports tourism visits will reach 1 billion, accounting for 15%

of the total number of tourism visits. The total consumption scale of sports tourism exceeds 1 trillion yuan.

As the sports tourism industry continues to develop, with the resulting growing economic and social impact, academic research related to sports tourism is gradually gaining attention [8–11]. At present, research topics on sports tourism show a diversified trend, including sports tourism characteristics and types, the economic impact of sports tourism, the characteristics of sports tourists and sports tourism destination planning. As the material basis of sports tourism activities, the importance of sports tourism resources is indisputable, and it concerns the development of the entire sports tourism industry. Most studies have mainly focused on areas such as the characteristics of sports tourism and the development of products, while insufficient attention has been paid to sports tourism resources [12–14]. In practice, the inaccuracy of resource classification has caused the problem of low efficiency in the actual use of sports tourism resources, limiting the development of sports tourism products and making the structure of sports tourism products single, thus hindering the development of sports tourism.

The rational classification of sports tourism resources is an important condition for the study of sports tourism resources. However, most of the existing research focuses on the development and evaluation of sports tourism resources, so there is a lack of specific research on the classification of sports tourism resources. The classification of sports tourism resources has not yet formed a unified standard. Compared to traditional tourism resources that focus on enjoyment, regional ethnic traditional sports tourism resources place more emphasis on the participatory and experiential aspects of the activities. The existing national standard tourism resource classification system does not reflect the characteristics of sports tourism resources, and therefore cannot be applied to the task of classifying regional ethnic traditional sports tourism resources. Therefore, it is necessary to establish a regional ethnic traditional sports tourism resources classification method with unique sports characteristics.

The classification of regional ethnic sports tourism resources facilitates census statistics and helps to capture the quality of the region's distinctive sports tourism resources. The tourism industry has a long history of development compared to the new sports tourism industry. Research on tourism is also relatively well developed. The dichotomous approach is the most widely used in the classification of tourism resources [15–17]. The dichotomous approach classifies tourism resources into two broad categories, natural and artificial, based on the different attributes of the resources. Although sport tourism resources are a subset of tourism resources, they are somewhat different from traditional tourism resources that focus on landscape. There is no uniform standard for the classification of sports tourism resources. A few common approaches to classifying sports tourism resources are shown in Table 1. Most existing sports tourism resource classifications are manually constructed based on the national standard tourism resource classification system. The efficiency and accuracy of the traditional manual classification is poor and does not reflect the characteristics of regional ethnic traditional sports tourism.

Table 1: Methodology for classifying sports tourism resources.

Basis of classification	Type	
Properties	Natural resources	Human culture resources
Purpose	Participatory	Ornamental
Space	Land	Water/air
Function	Leisure	Stimulation

As a fundamental artificial intelligence technique, machine learning algorithms have been widely used in various fields, especially for classification tasks of complex data. Currently, the main machine learning algorithms include multivariate discriminant analysis, logistic regression, decision tree classification, neural networks, genetic algorithms, support vector machines and cluster analysis, among others [18–21]. Srihadi et al. [22] proposed a method for classifying tourism markets based on cluster analysis. The method reveals the behavioural models of four categories of foreign tourists by clustering and analysing the differences in tourists' lifestyles. At this stage, as the research on artificial intelligence continues to advance, neural networks are developing very rapidly. Neural networks can operate intelligently by mimicking neurons in the human brain. This method is usually used to find the optimal solution, and we usually call it a network that can be fed back. Qin et al. [23] proposed a recursive neural network-based method for classifying tourism resources. Afzaal et al. [24] proposed a machine learning-based method for classifying the sentiment of tourism reviews. It can be seen that machine learning techniques have been used in many applications in the traditional tourism industry. However, after extensive investigation, it was found that no research has been conducted to apply machine learning techniques to the field of sports tourism resource classification.

In summary, the classification of regional ethnic traditional sports tourism resources is a complex issue. The efficiency and accuracy of traditional manual classification are poor and cannot reflect the characteristics of regional ethnic traditional sports tourism. Therefore, this study proposes a machine learning-based method for classifying regional ethnic traditional sports tourism resources. The aim of the research is to use the strong learning and self-adaptive ability of BP neural networks to improve the accuracy and efficiency of the classification of regional ethnic traditional sports tourism resources.

The main innovations and contributions of this study include:

(1) Taking the development of traditional sports of ethnic minorities in Yunnan Province as the research object, the SWOT analysis, literature, interview, questionnaire, and mathematical and statistical methods were used to investigate and analyse the overall status of the development of traditional ethnic sports in the region.

(2) Based on the analysis of classification principles, a Quantum Genetic Algorithm BP (QGA-BP) neural network model was proposed by introducing a

probabilistic evolutionary mechanism and applied to the classification of regional ethnic traditional sports tourism resources.

The rest of the study is organized as follows: In Section 2, the research objects and methods are studied in detail, while Section 3 provides the classification method of sports tourism resources based on the QGA-BP neural network model. Finally, the study is concluded in Section 5.

2. Research Subjects and Methods

2.1. Research Objects. This study is based on traditional sports of ethnic minorities in Yunnan Province, China, and surveys and interviews were conducted with local villagers and tourists between May 2018 and October 2019. This study takes sports tourism resources as the research object and completes the task of classifying regional ethnic traditional sports tourism resources based on the principles of classification.

2.2. Research Methodology. The aim of the study was to collect literature on traditional ethnic sports according to the purpose and content of the study, summarise the concepts related to sport tourism resources by analysing the valid information in the literature and establish a classification of ethnic traditional sports tourism resources by combining the existing research results.

Based on expert interviews, the questionnaire was designed according to the principles of the social survey method. For traditional sports in ethnic minority areas, we designed two levels of research questionnaires. The sample distribution of the questionnaire is shown in Table 2. The questionnaires were targeted at local villagers and tourists. A total of 350 questionnaires were distributed, of which 150 were distributed to villagers (Questionnaire A) and 146 were returned, accounting for 97.33%. 200 questionnaires were distributed to tourists (Questionnaire B) and 197 were returned (98.00%).

A split-half reliability test was conducted on the villagers' questionnaire and the tourists' questionnaire, and the correlation coefficients $r1 = 0.87$ and $r2 = 0.85$ were obtained. Both exceeded 0.85, indicating good reliability and that the questionnaire met the survey requirements. The calibration process for the split-half reliability test used the Spearman–Brown formula [25–27]:

$$r_{SB} = \frac{2r_{hh}}{1 + r_{hh}}, \tag{1}$$

where r_{SB} indicates the reliability indicator of the whole test after correction and r_{hh} indicates the split-half reliability.

SWOT analysis was introduced by Steiner, an American management professor, in the early 1980s and is one of the common methods of competitive intelligence analysis [28–30]. Through investigation and analysis, the SWOT analysis method can derive Strengths, Weaknesses, Opportunities, and Threats that are closely related to the object of study, and arrange them in a matrix in a certain order, as shown in Figure 1.

Table 2: Distribution of samples.

Object	Number	Number of recoveries	Effective number	Efficient (%)
Villagers	150	146	146	97.33
Tourists	200	197	196	98.00
Total	350	393	342	97.71

Then, using the idea of system analysis, the various factors are matched with each other and a series of corresponding conclusions are drawn through the analysis. The SWOT analysis method consists of the following steps: analysing environmental factors, constructing a SWOT matrix and formulating action strategies. External environmental factors include opportunity factors and threat factors, which are objective factors. Internal environmental factors include strengths and weaknesses, which are subjective factors. The factors are ranked according to their degree of influence and a SWOT matrix is constructed.

In the mathematical and statistical process, SPSS 19.0 statistical software was used to conduct relevant statistics and analysis of the returned questionnaires. The rationale framework for the classification of regional ethnic traditional sports tourism resources is shown in Figure 2.

3. Classification of Sports Tourism Resources Based on QGA-BP Neural Network Model

3.1. Principles of Classification. Sport tourism resources are a resource system consisting of various types of individual resources within a certain area, containing multiple single elements or individual resources, making the types of sport tourism resources more complex. It is important to have a full understanding of the various types of resources as a whole in order to maximise the range of sport tourism resources and thus ensure that the classification system encompasses all categories of sport tourism resources. When classifying sports tourism resources, it is important to ensure that resources in the same category meet the same classification criteria. In classifying, we accurately capture the characteristics of each type of resource so that resources with the same attributes are grouped together.

3.2. Analysis of Classification Indicators. This study uses two levels to build a classification system for sports tourism resources. The first level contains 2 types. Based on the meaning and characteristics of sports tourism resources, they are divided into two broad categories: participative and ornamental. The second level contains 6 categories. Based on the spatial distribution of the resources, the participative sports tourism resources are classified into the categories of geography, water, competition and leisure. Based on the classification of sports tourism, the spectator

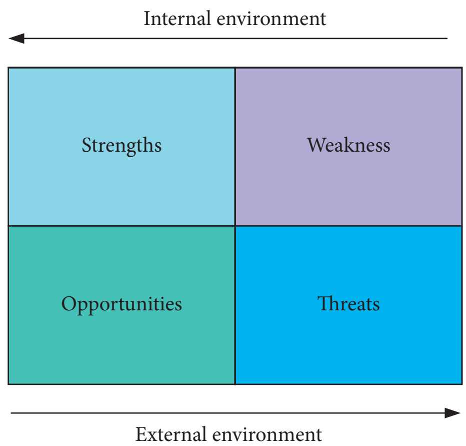


Figure 1: Traditional matrix for SWOT analysis.

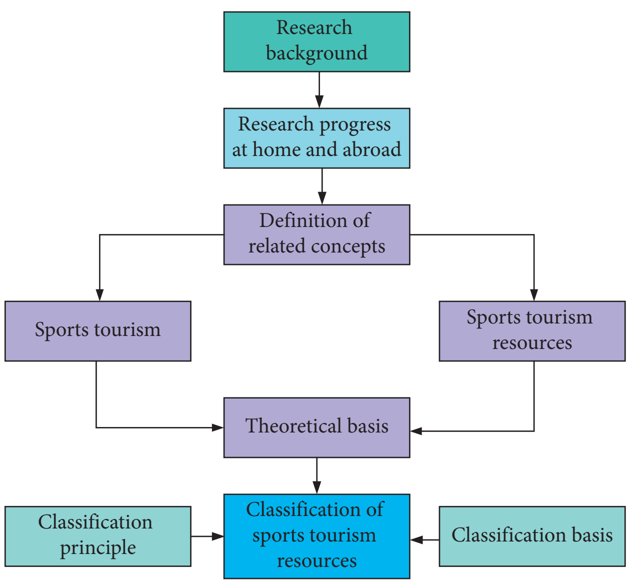


Figure 2: Rationale framework for the classification of tourism resources.

sports tourism resources are divided into performance and event categories.

3.3. Quantum Genetic Algorithm QGA. Genetic Algorithm (GA) is a randomised search algorithm [31–33] that seeks optimal solutions by simulating the behaviour of a population of organisms. GA is often used to optimise the weights and thresholds of neural networks. However, in practice, like similar heuristic search algorithms, GA suffers from long search time and low solution accuracy.

QGA is a new optimisation algorithm based on the GA algorithm [34]. Unlike the GA algorithm, QGA introduces the concept of quantum computing. First, the quantum state is considered as a primitive block of information. Then, the quantum states are superimposed. The problem of uncertain polynomial complexity is well solved by the quantum juxtaposition calculation method. QGA performs the following procedure [35].

Step 1: initialize the quantum population $Q(t_0)$.

Step 2: using quantum collapse operation to deal with population problems. Then, record each solution after the corresponding measurement.

Step 3: a fitness function is used to complete the evolution of the population.

Step 4: if the stop condition is satisfied, the iteration can be exited immediately. Otherwise, the next step will be performed.

Step 5: record each solution of the population $Q(t_0)$ after the iteration.

Step 6: analyse each solution using a fitness function and update the population size using a quantum rotation gate.

Step 7: continue with the iterative process and return to Step 4.

When measuring all individuals in the initial population and selecting the output value according to the probability amplitude of qubits, the quantum state of the QGA is irreversible and therefore the quantum bit encoding reduces the generation of computational errors. In addition, the optimisation of the quantum circuit reduces the negative effects of noise.

3.3.1. Quantum Bit Coding. Each gene in the encoding pattern of quantum bits has n parameters.

$$q_j^t = \left[\begin{array}{c|c|c|c|c|c|c|c|c|c|c|c} \alpha_{11}^t & \alpha_{12}^t & \cdots & \alpha_{1k}^t & \alpha_{21}^t & \alpha_{22}^t & \cdots & \alpha_{2k}^t & \alpha_{n1}^t & \alpha_{n2}^t & \cdots & \alpha_{nk}^t \\ \beta_{11}^t & \beta_{12}^t & \cdots & \beta_{1k}^t & \beta_{21}^t & \beta_{22}^t & \cdots & \beta_{2k}^t & \beta_{n1}^t & \beta_{n2}^t & \cdots & \beta_{nk}^t \end{array}\right], \tag{2}$$

where q_j^t denotes the chromosome of individual j in generation t, k denotes the number of quantum bits in a single gene, and n denotes the total number of chromosomes. All quantum bits code $(\alpha, \beta)^T$ of a population individual is initialized to $(1/\sqrt{2}, 1/\sqrt{2})$ such that each chromosome is expressed with equal probability:

$$\varphi_{q_i^t} = \sum_{k=1}^{2^m} \frac{1}{\sqrt{2}^m} |S_k\rangle, \tag{3}$$

where $|S_k\rangle$ is the quantum rotation gate at different quantum bits in a single gene and x_i has a value of 0 or 1.

3.3.2. Quantum Revolving Doors. The adjustment of the quantum revolving door operates is shown as follows [36]:

$$U(\theta_i) = \begin{bmatrix} \cos(\theta_i) & -\sin(\theta_i) \\ \sin(\theta_i) & \cos(\theta_i) \end{bmatrix}. \tag{4}$$

The update process is shown as follows:

$$\begin{bmatrix} \alpha_i' \\ \beta_i' \end{bmatrix} = U(\theta_i) \times \begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix}, \tag{5}$$

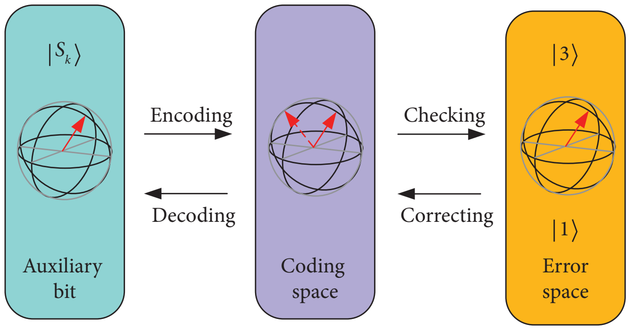


Figure 3: Quantum manipulation methods.

where $(\alpha_i, \beta_i)^T$ is the probability amplitude, $(\alpha_i', \beta_i')^T$ is the adjusted probability amplitude, and θ_i is the rotation angle. The quantum manipulation method is shown in Figure 3.

3.4. Optimising BP Neural Networks by QGA. Traditional BP neural networks have the disadvantages of slow convergence and falling into local optima. In order to overcome these problems, this study implements sports tourism resource classification through QGA-BP neural networks. Firstly, we optimise the BP network by QGA and set its weights and thresholds. Secondly, in order to make the output value of the optimised BP network model as small as possible close to the solution of the objective function, we search for a better search domain in a specific solution space. The QGA-BP algorithm operates as follows:

Step 1: initialize the population to $Q(t_0) = \{q_1^t, q_2^t, q_3^t, \cdots\cdots, q_n^t\}$.

Step 2: use the operation of quantum collapse for the initial population. Denote the record of measurements as $P(t_0) = \{p_1^t, p_2^t, p_3^t, \cdots\cdots, p_n^t\}$, p_i^t denoting the tth measurement for the tth generation of individuals.

Step 3: set the fitness function to $f = 1/E$. $E = \sum_{i=1}^{n} e_i^2$ is the sum of squares of the errors.

Step 4: keep a record of the fitness value and the best one for each individual in $Q(t_0)$.

Step 5: update the individuals by dynamically adjusting the strategy if the iteration stopping condition is not met.

Step 6: continue with the iterative process, and then jump to Step 2.

3.5. Classification of Sports Tourism Resources. As shown in Figure 4, the sports tourism resource classification method in this study is divided into two modules. The first part is the data pre-processing module, which is responsible for normalising the collected sports tourism resources data. The second part is the QGA-BP module. This module firstly inputs the pre-processed data into the input layer of the neural network and obtains the data features through convolutional processing. The data features are then compared with the desired value. If they do not match the desired result, the output layer will output 0. Otherwise, the network model will continue to be trained until the output layer outputs a result of 1.

4. Experimental Results and Analysis

4.1. Model Setup. A Sigmoid function is used for the transfer function of the QGA-BP neural network model. Meanwhile, the negative gradient function was used for the training function. The data samples we used originated from minority regions in Yunnan Province. Firstly, the dataset was divided equally into 100 groups, each group containing 50 pieces of data. Then, the dataset was treated as a non-linear function. The number of hidden layer neurons in the QGA-BP neural network model was determined by an empirical formula:

$$M = \sqrt{I + O} + a, \tag{6}$$

where a is a constant between 1 and 10. After 200 training sessions, the target error of the three-layer structure of the neural network is 10^{-10}. The length of the chromosomes of the QGA is $5 \times 10 + 10 \times 2 + 10 + 2 = 82$, i.e. 82 variable parameters. The bit number for each variable parameter of the QGA is 20, and the range is $[-1, 1]$. The total number of populations of the QGA is 20.

4.2. Model Performance Validation. The simulation was completed in a laptop computer with a central processor of Intel(R) i5-9300H CPU@2.40 GHz 2.40 GHz and a graphics card of NVIDIA GeForce GTX 1050. The simulation software is MATLAB 2019b. After 1000 iterations, the output results of the two models are shown in Figures 5 and 6 respectively.

The experimental results show that the QGA-BP model has the advantages of fast convergence and being less likely to fall into a local optimum solution. After 4 s of time, the fitness value of the individuals gradually increased. After 6 s, the individual fitness values continue to be optimised. After 9 s, the fitness value of the individual basically stabilises. It can be seen that the optimal fitness value of the QGA-BP model is significantly better than that of the traditional GA-BP model, and is more suitable for application in the classification of sports tourism resources.

The root mean square errors of the two models over 16 iterations are shown in Figures 7 and 8, respectively.

We can see that the training errors of both models show a linear decline from the beginning, but the decline is faster for QGA-BP. When the number of iterations was 15, the training error was consistent with the target error. This is because the GA-BP neural network model fluctuates greatly in the training process, so it cannot reach the required accuracy quickly.

4.3. Practical Case Studies. The GA-BP model and QGA-BP model were used to classify 30 sets of samples in the test dataset respectively. The classification errors of the traditional sports tourism resources with 10 sets of sample data are shown in Table 3.

Overall, the QGA-BP model was able to achieve the predicted results when applied to the traditional sports tourism resource classification, overcoming the problem of

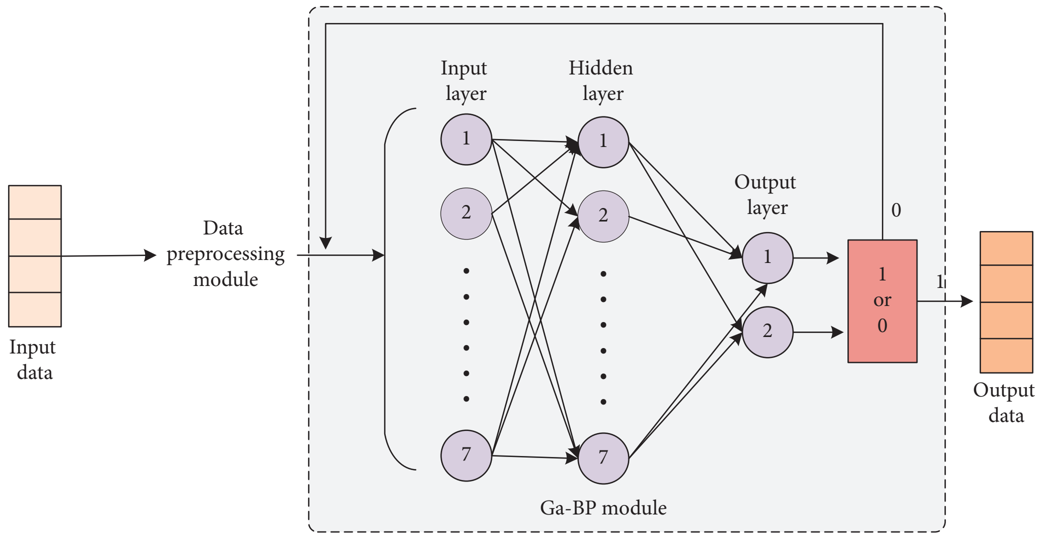


Figure 4: Traditional sports tourism resource classification based on QGA-BP neural network.

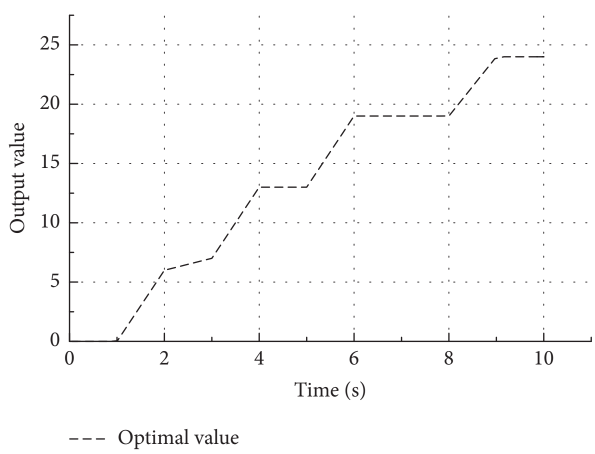


Figure 5: Output values of the QGA-BP neural network model.

Figure 7: Root mean square error of QGA-BP.

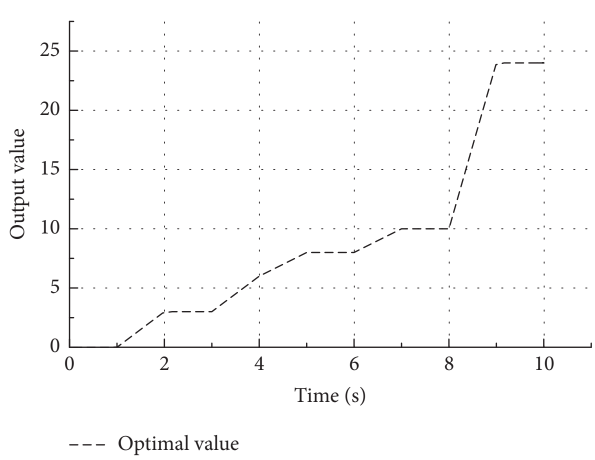


Figure 6: Output values of the GA-BP neural network model.

Figure 8: Root mean square error of GA-BP.

Table 3: Classification errors of traditional sports tourism resources.

Data	GA-BP	QGA-BP
Group 1	0.38273833	0.23973934
Group 2	0.38497236	0.2382949
Group 3	0.98736224	0.87469329
Group 4	0.82923224	0.57392474
Group 5	0.84739427	0.7329424
Group 6	0.04874424	0.0328294
Group 7	0.49482922	0.3928329
Group 8	0.12334422	0.0928372
Group 9	0.22455993	0.1039482
Group 10	0.39283244	0.2828492

Table 4: Classification of traditional ethnic sports tourism resources in Yunnan province.

Main category	Subcategory	Test sample of sports tourism resources	Number
Participative	Geography	Group 1, group 5, group 7	3
	Water	Group 6, group 8	2
	Competition	Group 2, group 3, group 4, group 10, group 11, group 12, group 14, group 15, group 16, group 20, group 22, group 23, group 25	13
	Leisure	Group 26, group 27, group 28, group 30	4
Ornamental	Performance	Group 9, group 13, group 17, group 18, group 21	5
	Event	Group 19, group 24, group 29	3

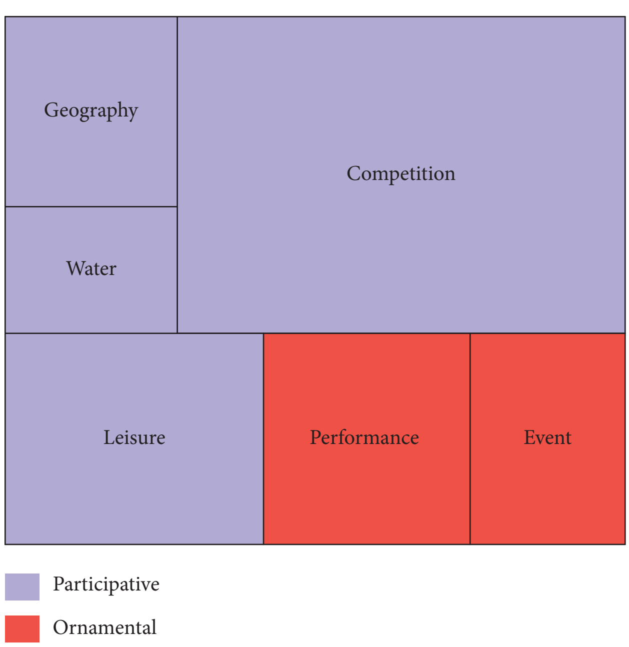


Figure 9: Tree diagram of classification results.

Table 5: Effect of the number of neurons in the hidden layer on the network model.

Number of neurons in the hidden layer	Classification errors
120	0.48736224
240	0.19036288
480	0.05672872
600	0.09726538
1024	0.12937344

the GA-BP model where large fluctuations in training error can occur.

The representative minority traditional sports tourism resources in Yunnan Province were collected and the collected resources were classified according to the proposed sports tourism resource classification method, and the statistical results are shown in Table 4 and Figure 9.

4.4. Neuronal Ablation Experiments. Because there are a large number of neurons in the hidden layer, a reasonable design of the number of neurons can not only reduce the amount of computation, but also help to improve the accuracy of sports tourism resource classification. Therefore, a large number of ablation experiments were done for how to set the number of neurons, and the experimental results are shown in Table 5.

In order to investigate the effect of the number of neurons per layer on the QGA-BP model, different numbers of neurons were chosen: 120, 240, 480, 600 and 1024. The results showed that when the number of neural units was too small, the network did not have enough fitting ability, which led to too large a classification error. As the number of hidden layers continues to increase, the classification error gradually decreases. However, an increase in the number of neurons leads to an increase in the time required for training. Therefore, in the QGA-BP neural network, we set the number of units per layer to 480.

5. Conclusion

This study uses QGA to optimise the BP network and applies it to the classification of traditional sports tourism resources. Taking the development of traditional sports of ethnic minorities in Yunnan Province as the research object, the SWOT analysis method, literature method, interview method, questionnaire method and mathematical statistics method are used to investigate and analyse the overall status of the development of regional ethnic traditional sports. The validity of the proposed QGA-BP model was verified through the sports tourism data obtained from the minority regions in Yunnan Province. The QGA-BP model effectively improves the accuracy of the classification of traditional sports tourism resources. Subsequent attempts will be made to apply the QGA-BP model to areas such as corporate finance and disaster risk assessment, so as to further verify its applicability and robustness.

Data Availability

The experimental data 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 to report regarding the present study.

Acknowledgments

The study was supported by "the National Social Science Fund of China, China (Grant no. 2020CTY002)" and "Fundamental Research Funds for the Central Universities, China (Grant no. 2020NTSS08)."

References

[1] L. E. Pedauga, A. Pardo-Fanjul, J. C. Redondo, and J. M. Izquierdo, "Assessing the economic contribution of sports tourism events: a regional social accounting matrix analysis approach:," *Tourism Economics*, vol. 28, no. 3, pp. 599–620, 2022.

[2] C. C. Yang, C. C. Shen, Y. S. Lin, H. W. Lo, and J. Z. Wu, "Sustainable sports tourism performance assessment using grey-based hybrid model," *Sustainability*, vol. 13, no. 8, p. 4214, 2021.

[3] R. W. Pouder, J. D. Clark, and G. G. Fenich, "An exploratory study of how destination marketing organizations pursue the sports tourism market," *Journal of Destination Marketing & Management*, vol. 9, pp. 184–193, 2018.

[4] Y. Chen and Y. Chi, "Harnessing structures in big data via guaranteed low-rank matrix estimation: recent theory and fast algorithms via convex and nonconvex optimization," *IEEE Signal Processing Magazine*, vol. 35, no. 4, pp. 14–31, 2018.

[5] G. R. Burmester, "Rheumatology 4.0: big data, wearables and diagnosis by computer," *Annals of the Rheumatic Diseases*, vol. 77, no. 7, pp. 963–965, 2018.

[6] E. Koc, "Risk and safety management in the leisure, events, tourism and sports industries," *Tourism Management*, vol. 54, pp. 296-297, 2016.

[7] Y. Zhu and X. Liu, "Big data visualization of the quantification of influencing factors and key monitoring indicators in the refined oil products market based on fuzzy mathematics," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 42021, pp. 6219–6229, 2021.

[8] X. -L. Meng, "Statistical paradises and paradoxes in big data (I): law of large populations, big data paradox, and the 2016 US presidential election," *Annals of Applied Statistics*, vol. 12, no. 2, pp. 685–726, 2018.

[9] G. Tsekouropoulos, O. Gkouna, D. Theocharis, and A. Gounas, "Innovative sustainable tourism development and entrepreneurship through sports events," *Sustainability*, vol. 14, no. 8, Article ID 4379, 2022.

[10] S. Shin, W. Chiu, and H. W. Lee, "For a better campus sporting experience: scale development and validation of the collegiate sportscape scale," *Journal of Hospitality, Leisure, Sports and Tourism Education*, vol. 22, pp. 22–30, 2018.

[11] J. J. Yang, H. W. Lo, C. S. Chao, C. C. Shen, and C. C Yang, "Establishing a sustainable sports tourism evaluation framework with a hybrid multi-criteria decision-making model to explore potential sports tourism attractions in taiwan," *Sustainability*, vol. 12, no. 4, Article ID 1673, 2020.

[12] E. Vayena and A. Blasimme, "Health research with big data: time for systemic oversight," *Journal of Law Medicine & Ethics*, vol. 46, no. 1, pp. 119–129, 2018.

[13] S. Nadal, O. Romero, A. Abelló, P. Vassiliadis, and S Vansummeren, "An integration-oriented ontology to govern evolution in Big Data ecosystems," *Information Systems*, vol. 79, no. 9, pp. 3–19, 2019.

[14] Y. Cao, H. Song, O. Kaiwartya et al., "Mobile edge computing for big-data-enabled electric vehicle charging," *IEEE Communications Magazine*, vol. 56, no. 3, pp. 150–156, 2018.

[15] L. Laing, M. Piekarz, I. Jenkins, and P. Mills, "Risk and safety management in the leisure, events, tourism and sports industries," *Sport Management Review*, vol. 22, no. 2, pp. 288-289, 2019.

[16] P. Cichosz, "Unsupervised modeling anomaly detection in discussion forums posts using global vectors for text representation," *Natural Language Engineering*, vol. 26, no. 5, pp. 551–578, 2020.

[17] Y. Bao, Z. Tang, H. Li, and Y. Zhang, "Computer vision and deep learning–based data anomaly detection method for structural health monitoring," *Structural Health Monitoring*, vol. 18, no. 2, pp. 401–421, 2019.

[18] L. Shi, Y. Zhu, Y. Zhang, and Z. Su, "Fault diagnosis of signal equipment on the lanzhou-xinjiang high-speed railway using machine learning for natural language processing," *Complexity*, vol. 2021, no. 8, Article ID 9126745, 13 pages, 2021.

[19] S. Talukdar, M. W. Naikoo, J. Mallick et al., "Coupling geographic information system integrated fuzzy logic-analytical hierarchy process with global and machine learning based sensitivity analysis for agricultural suitability mapping," *Agricultural Systems*, vol. 196, pp. 103343–103351, 2022.

[20] V. Landa, Y. Shapira, M. David et al., "Accurate classification of fresh and charred grape seeds to the varietal level, using machine learning based classification method," *Scientific Reports*, vol. 11, no. 1, pp. 13577–13615, 2021.

[21] M. Tamal, M. Alshammari, M. Alabdullah, R. Hourani, H. A. Alola, and T. M Hegazi, "An integrated framework with machine learning and radiomics for accurate and rapid early diagnosis of COVID-19 from Chest X-ray," *Expert Systems with Applications*, vol. 180, Article ID 115152, 2021.

[22] T. F. Srihadi, S. D. Hartoyo, D. Sukandar, and A. W. Soehadi, "Segmentation of the tourism market for Jakarta: classification of foreign visitors' lifestyle typologies," *Tourism Management Perspectives*, vol. 19, pp. 32–39, 2016.

[23] Q. Li, S. Li, J. Hu, S. Zhang, and J. Hu, "Tourism review sentiment classification using a bidirectional recurrent neural network with an attention mechanism and topic- enriched word vectors," *Sustainability*, vol. 10, no. 9, Article ID 3313, 2018.

[24] M. Afzaal, M. Usman, and A. Fong, "Tourism mobile app with aspect-based sentiment classification framework for tourist reviews," *IEEE Transactions on Consumer Electronics*, vol. 65, no. 2, pp. 233–242, 2019.

[25] I. Van Heerden and A. Bas, "Viewpoint: AI as author - bridging the gap between machine learning and literary theory," *Journal of Artificial Intelligence Research*, vol. 71, pp. 175–189, 2021.

[26] B. D. M. Lopes, L. C. B. Silva, I. M. Blanquet, P. Georgieva, and C. A. F Marques, "Prediction of fish mortality based on a probabilistic anomaly detection approach for recirculating aquaculture system facilities," *Review of Scientific Instruments*, vol. 92, no. 2, Article ID 025119, 2021.

[27] J. Kadupitiya, F. Sun, G. C. Fox, and V Jadhao, "Machine learning surrogates for molecular dynamics simulations of soft materials," *Journal of Computational Science*, vol. 42, Article ID 101107, 2020.
[28] L. Silva, M. Coutinho, and C. Santos, "Hardware architecture proposal for TEDA algorithm to data streaming anomaly detection," *IEEE Access*, vol. 9, pp. 103141–103152, 2021.
[29] K. S. Rallis, C. R. Hillyar, M. Sideris, and J. K Davies, "T-cell-based immunotherapies for haematological cancers, Part A: a SWOT analysis of immune checkpoint inhibitors (ICIs) and bispecific T-cell engagers (BiTEs)," *Anticancer Research*, vol. 41, no. 3, pp. 1123–1141, 2021.
[30] K. Almutairi, S. J. Hosseini Dehshiri, S. S. Hosseini Dehshiri, A. Mostafaeipour, A. X. Hoa, and K Techato, "Determination of optimal renewable energy growth strategies using SWOT analysis, hybrid MCDM methods, and game theory: a case study," *International Journal of Energy Research*, vol. 46, no. 5, pp. 6766–6789, 2022.
[31] I. Apostol, M. Preda, C. Nila, and I Bica, "IoT botnet anomaly detection using unsupervised deep learning," *Electronics*, vol. 10, no. 16, p. 1876, 2021.
[32] 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.
[33] A. Y. Prasad and B. Rayanki, "A generic algorithmic protocol approaches to improve network life time and energy efficient using combined genetic algorithm with simulated annealing in MANET," *International Journal of Intelligent Unmanned Systems*, vol. 8, no. 1, pp. 23–42, 2020.
[34] L. Xu, S. Zhao, N. Li, Q. Gao, T. Wang, and W. Xue, "Application of QGA-BP for fault detection of liquid rocket engines," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 5, pp. 2464–2472, 2019.
[35] S. Klembt, "Light detection nears its quantum limit," *Nature*, vol. 597, no. 7877, pp. 483-484, 2021.
[36] H. Bao and W. Wang, "Canonical bases arising from quantum symmetric pairs of Kac–Moody type," *Compositio Mathematica*, vol. 157, no. 7, pp. 1507–1537, 2021.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 7932684, 12 pages
https://doi.org/10.1155/2022/7932684

Research Article

Study on the Sustainable Development Strategy of School Soccer Based on the Background of Big Data Era

Jun Zhang, Dong Liang, and Zhang Sun

Guangzhou Xinhua University, Guangzhou, Guangdong 510520, China

Correspondence should be addressed to Dong Liang; bcld@xhsysu.edu.cn

Received 11 July 2022; Revised 27 July 2022; Accepted 4 August 2022; Published 20 August 2022

Academic Editor: R. Mo

Big Data is the most popular concept in this era, which is the massive amount of information and related technology generated by the information explosion in the era of "Internet+." Big Data is the most popular concept of our time. With the most advanced technology to collect, analyze, organize, and store data, Big Data can effectively handle all kinds of complex information. Because of this, big data is widely favored by all walks of life. In China's sports industry, the use of big data has become mature and has shown its unique advantages. With the development of campus soccer in China in the past decade, how to use big data to promote the sustainable development of campus soccer in China has become a key issue for sports workers to consider today. Based on the above background, this paper proposes a system combining data mining and personalized data recommendation to collect and analyze the information of campus soccer to promote the sustainable development of campus soccer. First, we propose a data mining method based on deep learning data mining network model combined with migration learning to address the data mining problem. The method uses the knowledge of historical model parameters and applies them to new tasks, thus solving the problem of network training when samples are lacking and improving data utilization and data mining effects. Then, for the data recommendation problem, a new deep learning method is proposed, which performs effective intelligent recommendation by pretraining. In the initial phase, the corresponding low-dimensional embedding vectors are learned, which capture information reflecting the relevance of students to soccer sports. During the prediction phase, a feed-forward neural network is used to model the interaction of student and soccer sport information, where the corresponding pretrained representative vectors are used as inputs to the neural network. Finally, it is experimentally verified that the data mining method proposed in this paper can effectively improve the data mining performance and efficiency, and the proposed data recommendation method possesses better accuracy than the traditional methods. The use of this system can effectively collect and analyze campus soccer information, which helps to develop campus soccer and promote the sustainable development of campus soccer.

1. Introduction

Soccer, as the world's number one ball, has been attracting attention all over the world since its birth. China, as the largest developing country in the world, has also been committed to the development of soccer. After continuous learning and exploration, it embarked on the road of professional development of soccer in the early 1990s. However, more than two decades of learning and efforts have not resulted in rapid progress, and the results of Chinese soccer are still unsatisfactory. To this end, it is particularly important to explore a sustainable path for the development of Chinese soccer, and the government has specifically formulated a strategy for the development of school soccer, placing its hopes on the youth and introducing a series of policies to promote the popularity of school soccer, which has also ushered in unprecedented development opportunities for youth soccer [1]. Of course, opportunities and challenges coexist, and some problems inevitably arise in the process of school soccer development.

Through the survey, it was found that in the institutions visited, there was no long-term development plan for campus soccer, but only regular teaching documents such as training plan, syllabus, and teaching schedule of comprehensive nature. Some institutions have set up school soccer teams, but they are prone to "discontinuity," and there are no incentives for students to participate in the games due to their personal hobbies, so it is difficult to guarantee the training level and game performance. The current problems facing the development of campus soccer are shown in Figure 1, which can be summarized into four main aspects.

(1) Lack of scientific theoretical guidance for sports selection. Sports selection is the beginning of competitive sports, and a good start is the foundation of success. So in a sense, the scientific level of soccer player selection directly affects the development level of soccer in a country [2]. The research of youth selection in China has been lagging behind for a long time, especially some coaches at the grassroots level still uphold the traditional selection concept, relying on subjective experience to select, resulting in many excellent sports talents being buried. The selection of school soccer players also inevitably faces these problems, so scientific theoretical guidance in sports selection is particularly important.

(2) Sports training is difficult to complete with quality and quantity. Athletic training is an essential part of improving athletic performance, and without quality training, it is difficult to achieve the expected results. In the process of training campus soccer players, you can often find a variety of problems, some from the teachers, some from the students' parents, and some from the students themselves, how to balance the relationship between various aspects, and deal with the contradiction between student learning and training, which is also a subject that requires in-depth research. Cultivating an excellent athlete is a long-term process that requires perseverance and persistence, so sports training is the guarantee of athletes' success.

(3) Sports competition fails to form a long-term mechanism. Soccer league is an important part of campus soccer activities and is the main way to evaluate the effect of campus soccer development, and the quality of soccer league development also directly reflects the development of campus soccer [3]. According to the National School Football Competition Program, efforts are made to build a school soccer competition system based on inter-school competitions, interschool competitions and regional competitions, and to improve the four-level league mechanism of school soccer. In recent years, various types of campus soccer competitions have increased, but still no long-term mechanism or scale effect has been formed. Therefore, it is imperative to reform campus soccer, normalize, and scale campus soccer competitions, and establish a campus soccer competition system in line with regional characteristics.

(4) The cultural atmosphere of campus soccer is lacking. Looking at the traditional soccer teams in Europe and South America, it is not difficult to find that they all have a strong soccer atmosphere, and soccer seems to be an indispensable part of their lives, while the soccer culture in China is relatively absent, especially in the campus [4]. Most people look more at your cultural achievements, and under the guidance of such social values and recognition, many young people with athletic talent eventually did not embark on the development of soccer, and the construction of campus soccer culture has become an urgent problem.

In professional soccer, a large amount of data involving tactical behavior analysis is collected, and cooperation with computer science can lead to more rapid development of soccer techniques and tactics because big data presents new ideas on data management and analysis methods commonly used in sports. In recent years, the equipment for data collection on soccer performance has changed day by day and the quality and quantity of data have grown rapidly, resulting in teams having large amounts of data to process on a daily basis [5]. The human form of statistics and analysis has long been unable to match the rapid development of today's soccer game, and the efficiency and speed of big data processing is evident. With the application of big data, technical and tactical performance, which are more abstract and usually evaluated qualitatively, can be quantified and analyzed to provide more scientific suggestions for players' daily training and games.

In tactical performance analysis, passing, as one of the most common and frequent elements, deserves the attention of coaches and players. How to define and distinguish the difference between effective and ineffective passes cannot be determined entirely by subjective judgment, and data-oriented content is needed to support this. Some researchers have assessed the effectiveness or ineffectiveness of passing by developing a model that combines passing effectiveness with consistent offensive performance, rather than relying on the occurrence of infrequent probability events. The model is built to help assess any position, individual tactics, efficiency, player comparisons, and team capabilities, while the method is applied to the game and can also help teams identify important team roles. Tracking data analytics applied to soccer mechanics and tactics demonstrates the efficiency and speed of big data. Whereas previously technical and tactical analysis was only superficially useful for game performance, advances in the data analysis process have now greatly improved the rationalization of athletic performance [6]. Addressing the aggregation regarding performance feature construction, space, and time through multidisciplinary collaboration is critical to unlocking the potential of soccer position tracking data.

Currently, the expected product dissemination of soccer events is more popular. The application of Internet and big data technologies can collect a large amount of data from individual teams and players and display and summarize them through different indicators to predict the performance of teams and players in the game, presenting more intuitive data for fans and optimizing their experience. Many event communication companies will use the information

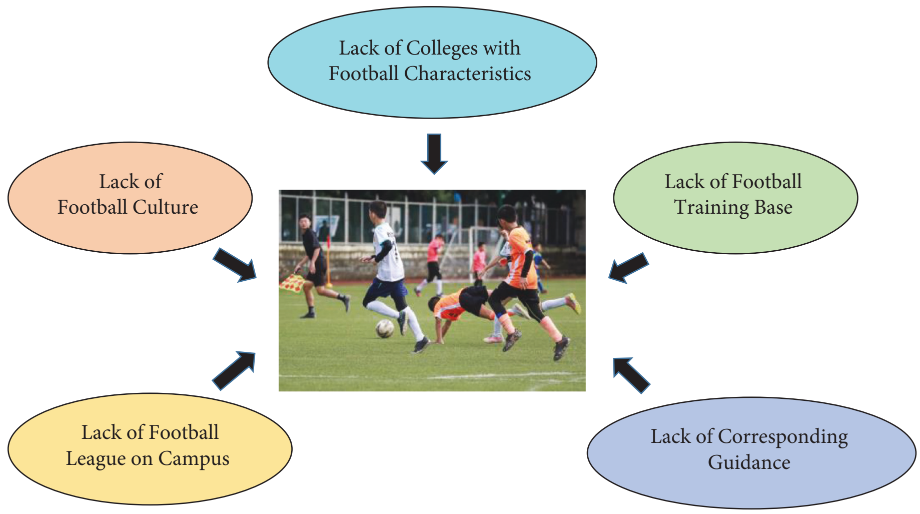


Figure 1: Problems faced by the development of campus football.

technology interaction mode to realize the interaction between commentary and fans, fan-to-fan interaction, etc., to enrich the audience's viewing experience, while the background can make timely adjustments through the collection of user feedback information [7]. The application of big data is also reflected in sports betting. By collecting massive amounts of data, tournament communication companies can provide reference results to fans, who can analyze the results of matches based on the data and attract more people to watch soccer matches. The tournament distribution company can increase the amount of users and collect more data to achieve a virtuous cycle.

Introduce the advanced concept of Internet and big data into the practice of campus soccer and combine big data with student physical fitness test, physical education entrance examination, and four levels of campus soccer league to make technological selection of campus soccer players based on objective data. This can prevent defects such as the static and one-sided nature of the data of the selection index in the past, and also provide certain reference and reference for campus soccer teaching, training, and competition. At the same time, a wide area network for athlete selection should be established nationwide so that the whole country can be played as a single game and resources can be shared.

The main contributions of this paper are as follows: in order to better promote the sustainable development of campus soccer, this paper proposes a combined data mining and data recommendation approach for the collection and analysis of campus soccer sports information. First, a data-mining model based on migration learning, and MMD is proposed for the data-mining task, and MMD is used to measure the distribution differences between source and target data, so that the network model can be adjusted accordingly according to the size of MMD. The model is pretrained using the source domain data and retrained and fine-tuned using the target domain data, thus improving the accuracy and efficiency of data mining. Then, a collaborative filtering recommendation system based on deep neural networks for explicit feedback data is constructed for the data recommendation task, which uses feedforward neural networks to establish the connection between student and soccer information. Finally, the proposed method is experimentally verified to be effective in analyzing campus soccer data, which are conducive to promoting the sustainable development of campus soccer in the era of big data.

2. Related Works

2.1. The Current Situation of Campus Soccer Development. Before exploring school soccer, we must first understand what "school soccer" is and only by understanding its nature and characteristics can the developed intelligent system be close to the actual application needs and serve school soccer to the greatest extent and solve the related problems. According to some researchers, school soccer is a general term for all kinds of soccer activities that focus on cultivating young people's interest in soccer, in the form of competitive matches or games. The purpose of school soccer activities is to improve students' physical fitness, cultivate a spirit of hard work and a sense of cooperation, and its ultimate goal is not just to choose and compete [3]. With the development of large-scale campus soccer activities and the construction and improvement of the four-level soccer league system not only can we popularize soccer knowledge and skills for young students, so that more students understand soccer and are willing to join the sport but also increase the soccer population, discover and cultivate soccer reserve talents from it, and lay a good foundation for the development of soccer in China.

School soccer activities have been carried out for nearly 11 years, and in response to the current situation of school soccer, some researchers point out that certain schools carry out soccer training activities superficially, focusing only on the performance form of efforts, and do not go deep into the organization and promotion of soccer activities. Some schools, due to limited funds, invest most of their funds in cultural learning programs and are unable to provide comprehensive financial support for soccer training programs, resulting in the failure of soccer training programs [8]. The lack of professional soccer teaching staff is also an important reason why school soccer activities cannot be carried out. Most physical education teachers only have basic knowledge of soccer and cannot teach students soccer skills and tactics in depth, which is a great obstacle to the cultivation of soccer talents. In addition, some students' parents only focus on cultural learning and worry that soccer will affect their students' performance in cultural classes, so they do not support students to play soccer.

The aim of school soccer is to make more students participate in soccer, feel the fun of soccer, experience the culture and spirit of soccer, and guide the development of students in all aspects of body and mind. Combining literature and personal experience, the authors conclude that the development of campus soccer faces the following problems in practice: (1) sports safety; (2) a large lack of professional soccer coaches/teachers; (3) inability to objectively and comprehensively reflect players' performance; (4) lack of scientific talent training model; (5) lack of efficient and professional competitive training assistance; (6) lack of "intelligent" management method. The above problems largely restrict the vigorous development of school soccer. Solving the above problems is conducive to the further promotion of school soccer, enhancing the physical fitness of youth and the spirit of teamwork, and promoting China's progress toward a strong sports nation.

Internet+ is a new industry of Internet development under Innovation 2.0, which organically integrates the Internet with various traditional industries to create a new development ecology through cross-border integration, innovation-drive, and structural reshaping. Studying the deep integration of campus soccer and the Internet is conducive to improving the development of intelligent systems for campus soccer [9]. Unlike the inefficient traditional methods, a smart campus soccer system must be combined with the Internet to create a scientific, rational, and efficient campus soccer management and training model.

With the increasing maturity of "Internet+" technology, the emergence of intelligent soccer fields has emerged. Parents only need to connect to the network through 4G or Wi-fi, and they can view students' performance on the smart soccer field in real time with the help of cell phones and other communication devices, using Internet streaming media and other technologies, parents can make real-time scrolling caption comments, interact with other parents, and support one-click forwarding of the game video to the WeChat circle of friends. In addition, through the use of big data and cloud computing technology, ordinary users can have real-time location near the location of the stadium, real-time understanding of the use of the stadium, using cell phones and other intelligent devices to make reservations for the use of the stadium and other operations to enhance the use of soccer fields, improve the rate of reasonable allocation of soccer fields, and greatly solve the problem of soccer field shortage [10]. The combination of Internet technology and campus soccer can realize the function of "intelligent coach" and provide innovative new thinking for campus soccer training.

The combination of Internet and campus soccer not only improves the intelligent teaching level of campus soccer schools but also greatly saves the teaching cost and human cost in the teaching process, and gives full play to the convenience and wisdom of the Internet. For example, in the school soccer work, with the help of Internet technology, the performance data of students at each stage can be recorded, so that a more comprehensive and detailed understanding of the overall level of students [11]. Teachers no longer rely on personal intuition and usual observation to provide targeted training for students, but only need to check the changes in students' performance data to scientifically and reasonably arrange targeted training for students, and in the training process, the intelligent system of campus soccer can also give suggestions on training directions and training programs, greatly reducing teachers' labor intensity. In addition, the intelligent human ball data interaction can also reduce the subjective bias brought by human operation and ensure the objectivity and accuracy of the data.

With the increasing prevalence of big data, soccer, the world's number one sport, needs to progress with the times and keep up with the pace of development. The application of big data in soccer can not only greatly improve the accuracy and speed of information acquisition but also monitor in real time whether the various physiological references in athletes' sports are within a reasonable level range, effectively enhancing the safety and protection of athletes. Only by scientifically using big data technology and adopting an innovative education model can we accelerate the development of Chinese soccer faster.

2.2. Current Status of Data Mining Research. Before data mining, we should clarify our data orientation, determine the direction and scope of data mining, and then implement data mining to avoid data redundancy, data bias and other problems, and avoid blind mining. Data mining is the substantive mining of data, and then select the suitable algorithm for the data research according to the theme, and then implement the data mining work, this link is the core link of data mining work. The main methods of data mining include four categories, as shown in Figure 2.

The decision tree generated in different scenarios will be different, so the decision tree will also be called classification tree, regression tree, etc. The classical algorithms of decision tree data mining methods are the following: (1) Cart algorithm—it is a simple binary tree algorithm, which is often used in simple data to generate a simple binary tree structure. (2) ID3 algorithm [12]—ID3 algorithm is a relatively early algorithm in the decision tree algorithm, and it is based on data information through a series of rules to find the attributes represented by each node in the tree, the entropy of the algorithm as the basis for classification, the data will eventually generate the form of decision trees. (3) C4.5 algorithms [13]—this algorithm uses information gain or entropy to optimize the process of decision tree node classification and improve the decision tree to make it more friendly.

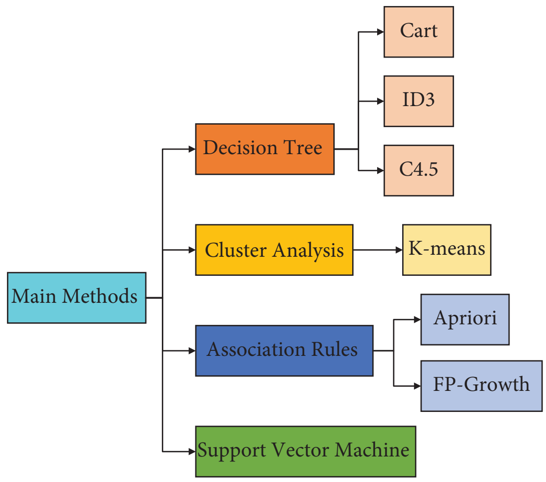


Figure 2: Main methods of data mining.

Clustering analysis is essentially to find out the classification basis of data according to the research topic and according to this basis to classify the data, and refine the data into different types of data sets, and ensure that the data in each set has similarity, and there are differences between different sets, and then use the data visualization technology to represent them, and friendly to the user, that is called cluster analysis. The main algorithm is K-means algorithm, the outstanding advantage of this algorithm is the principle of simple, efficient application, very suitable for the processing of large-scale data.

Association rule analysis is one of the more commonly used methods in data mining work. Association rules refer to the hidden rules of relationship between things, and association rule analysis refers to the process of finding and analyzing the information between things and association rules with set values [14]. The Apriori algorithm can solve the analysis of association rules of corresponding data, but it has some shortcomings, so some researchers proposed the FP-Growth algorithm to make up for the shortcomings of Apriori algorithm in generating candidate item sets.

Support vector machine (SVM), a binary classification model, is defined as a linear classifier with maximum interval in the feature space, which is different from the perceptron when the interval is maximum. The core concept is that the support vector samples play a key role in the recognition problem, and the support vector is the nearest sample point to the classification hyperplane, which is the support vector classifier, and the classification hyperplane is used to divide the sample data into two.

In terms of data mining objects, data mining will mostly favor multimodal data mining in the later stage [15]. At the present stage, most of the data mining is based on the corresponding algorithm, and the algorithm process is not easy to be understood by users, so the data mining visualization research has certain research significance.

In financial investment, it is possible to predict the future trend of stocks through data mining algorithms based on historical stock trading data, thus promoting profitability [16]. In terms of fraud information identification, the identification of fraud information is mainly through data mining methods to obtain certain typical characteristics of fraudulent behavior, and when the relevant business processing or behavior is highly similar to or matches these typical characteristics, it is possible to issue warnings to relevant personnel through some interactive techniques. In addition, data mining has very great potential for application in bank lending, process optimization, oil reserve prediction, drug synthesis and development, chemistry, chemical industry, etc.

2.3. Current Status of Personalized Data Recommendation Research. Recommender system is a new research field combining various disciplines such as data mining, prediction algorithms, and machine learning. In the definition of recommendation system, it is pointed out that in daily life, whether it is an understood event or an unknown event, people always need to make decisions. When facing familiar things, people can often rely on past experiences to make reasonable decisions; however, when facing unknown things, people need others' verbal suggestions, book reviews, movie reviews, recommendations [17]. To make judgments—recommender systems, which match different users with items from a large number of items that match their interest preferences but are not observed by the users, are considered to be becoming an important business with significant economic impact.

The core idea of the content filtering-based recommendation technique is that the user's historical selection record or preference record is used as the reference recommendation, and the items with high correlation with the reference recommendation in other unknown records are mined as the content recommended by the system [18]. Then calculate the similarity between the user's preferences and the recommendation object to be tested in terms of content; finally, rank the similarity between the recommendation object to be tested and the user's preferences, so as to select the recommendation object that matches the user's interest preferences.

The content-based recommendation technology often reduces the timeliness of information due to being time-consuming when dealing with large-scale information content; the collaborative filtering technology is prone to cold start problem when facing new items; and the hybrid recommendation technology is a recommendation method that retains the advantages of different recommendation technologies and avoids their disadvantages by incorporating different algorithms into the recommendation system. The current hybrid recommendation is mainly divided into prefusion, postfusion, and midfusion: (1) Prefusion—refers to the fusion of multiple recommendation algorithms into one model, such as in the process of product recommendation. (2) Midfusion—this hybrid recommendation technology generally [19]. (3) Postfusion—this method places great importance on the recommendation results, mainly by comparing the recommendation effects of different recommendation algorithms to get a more reliable sequence of recommended objects, and finally recommending this sequence to users.

Deep learning algorithms are powerful because they can learn and deal with complex problems such as human beings, analyze and calculate linear or nonlinear feature sequences from multiple dimensions in the face of complex scale data, and automatically learn features that meet users' needs from massive data [20]. Deep learning techniques can not only discover the hidden potential features of user behavior records, but also capture the interaction features of user-user, user-item, and item-item nonlinear relationships, which brings more opportunities for system performance improvement and can overcome some obstacles encountered in traditional recommendation techniques to achieve more accurate recommendations.

3. Algorithm Design

3.1. Data Mining Model Design. Transfer learning is the mainstream machine learning method to address this problem, which transfers the knowledge learned from the previous task to the new task, with the aim of achieving better learning results in the new task. As the driving force of future deep learning, transfer learning can effectively solve the above problems. Therefore, migratory learning is introduced in data mining to make efficient use of big data and explore the commonality of different data sets.

First, the source and target domains are preprocessed accordingly. Then, the original deep neural network model is trained with the source domain data, or the trained network structure and parameters are already available as the source domain classification or prediction model [21]. The difference of data distribution between the source domain and the target domain is analyzed by MMD to obtain the distribution distance. The original network structure is adjusted according to the MMD to obtain a new target domain network, and the parameters obtained from the original network training are selectively migrated. Generally, the source domain and the target domain distribution will have some differences, that is, the MMD is higher than the set threshold, then the network needs to add or replace new hidden layers for the target domain model to learn new knowledge.

It is based on the principle of finding the mean of samples with different distributions by finding a continuous function on the sample space, and then finding the difference between these two means as the difference in means corresponding to these two distributions [22]. The maximum value of this difference is the MMD of the two distributions:

$$MMD[F, p, q] = \sup_{f \in F}\Big(E_p[f(x)] - E_q[f(y)]\Big). \tag{1}$$

As the size of the observed data set increases, constraints are needed to speed up the convergence of the empirical estimates of MMD. Using the complete inner product space becomes the Hilbert space, and the regenerative kernel Hilbert space can be expressed as the dot product in the space:

$$f(x) = \langle f, \phi(x) \rangle_H. \tag{2}$$

The inner product can be replaced by a kernel function, and for mappings in higher-dimensional spaces, the radial basis kernel function is usually used:

$$k(x, x') = e^{-\|x-x'\|^2/2\sigma^2}. \tag{3}$$

In general, MMD can be regarded as the distance between two points in the regenerative kernel Hilbert space, which can be used to measure the distance of two distributions.

According to the above migration steps, the network is not adjusted for the MMD below the set threshold, and the network is directly trained unsupervised with the target domain data based on the original parameters and model, as shown in Figure 3.

The parameter fixing is canceled, and the whole network is fine-tuned using the target domain data to obtain the final target domain network model.

3.2. Personalized Data Recommendation Algorithm Design. The goal of a recommendation system is to recommend content or products that users like or are interested in. In simple terms, a user will give a high rating to content or products that he is interested in or likes. Therefore, it is a simple and effective strategy to recommend a product to a user that he may give a high rating to. In summary, recommendation systems can be broadly viewed as a problem for predicting the user's rating prediction for a product.

A variety of recommendation system models have emerged for the rating prediction problem. In general, most of the current recommendation system models can be classified into two categories: (1) content-based recommendation systems, and (2) collaborative filtering-based recommendation systems. Among them, content-based recommendation systems make recommendations by extracting features from user or product content, such as user profile content, product description content [23]. The collaborative filtering-based model, on the other hand, obtains the available recommendation system by deriving from the user's historical interaction behavior. The user's historical behavior can be the user's click record, purchase record, or rating record of a product on a specific website. Among them, the user's rating records are again the most commonly used. Recommendation systems based on collaborative filtering models are currently gaining more attention because of their higher recommendation accuracy.

Based on the machine-learning algorithm, training a usable recommendation system with the user's historical interaction data is the most commonly used collaborative filtering method. Here, the commonly used machine-learning model is based on matrix decomposition model (MF). However, artificial neural network (ANN)-based models have also started to gain attention in recent years [24]. The MF-based approach can better model simple user-product interactions, but more complex interactions cannot be handled effectively due to the low complexity of the MF model.

Therefore, the model can be divided into two main phases, as shown in Figure 4: (1) feature learning phase: the feature learning model generates the corresponding low-dimensional user and product feature vectors based on the user-user and product-product co-occurrence relationships

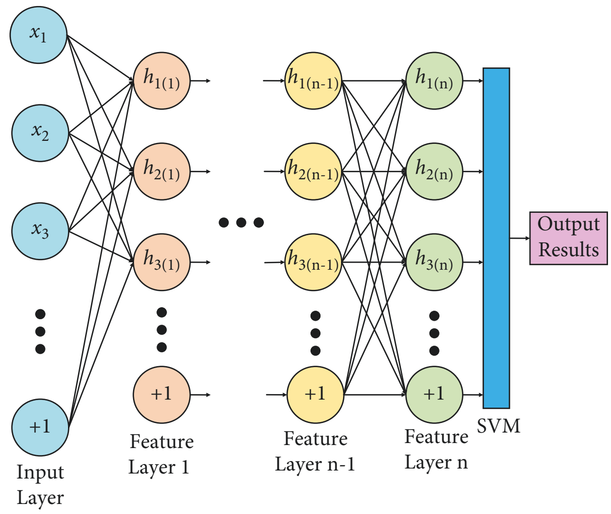


FIGURE 3: Network adjustment process based on transfer learning.

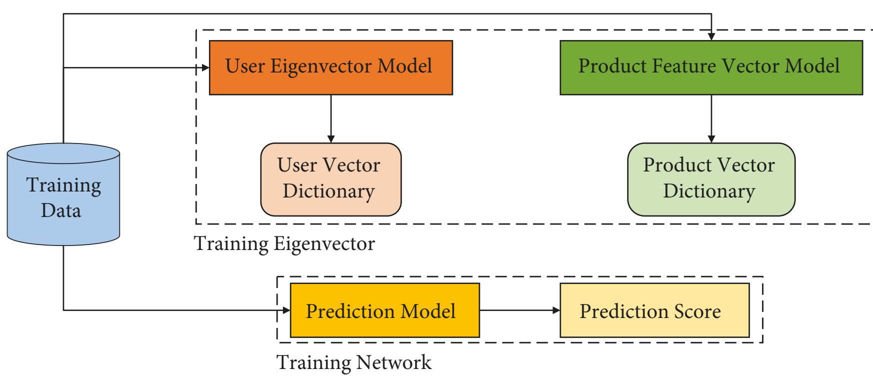


FIGURE 4: Model two-stage learning framework.

through the user's rating matrix. (2) Neural network training phase—the final score is predicted by the rating prediction neural network by using the user's product's CM or RIM feature vector as input, which is calculated by layer-by-layer operations in the network.

Compared to the previous models, the proposed model is able to utilize both co-occurrence and interaction. This is because the pretrained feature vectors can obtain user and product feature vectors that contain co-occurrence features, while the predictive neural network can simulate interaction relationships. Previous models can either handle co-occurrence but not interaction, or the opposite. The pretrained feature vectors can be easily used in different branches of the model. Training the feature vectors and neural networks separately can achieve higher prediction accuracy than joint training.

The product and user feature vectors obtained by the feature learning model reveal the co-occurrence characteristics of users and products, but these feature vectors do not directly yield the final rating prediction results. Therefore, additional components are needed to estimate the user's rating of the product. Since artificial neural networks are able to efficiently extract features and model complex objective functions, as well as fuse input features from multiple perspectives [25]. Therefore, neural networks are an ideal model for rating prediction. In this section, a detailed description of how the neural network is used to generate the final predicted scores from the obtained feature vectors is presented.

The goal of rating prediction is to output a real number of predicted values based on the features of a given user and product as input, and this predicted value represents an estimate of the score given to the product by the user. Thus, the rating prediction problem can be viewed as a regression problem. Therefore, the most straightforward way to use neural networks is to directly input the CM or RIM features of users and products obtained in advance into a feedforward neural network to obtain the prediction values. However, the prediction accuracy of this approach is not high enough [26–29].

This history-based perspective differs from the basic perspective that utilizes only current users and products; it will represent users or products by their historical records. The general architecture of the network is shown in Figure 5. In which, the network takes CM and RIM features as input. The symbols "FC" and "CONV" in the figure indicate the use of fully connected layer and convolutional layer, respectively. In general, the network can be divided into two main parts: (1) multiview feature extraction and (2) integrated prediction.

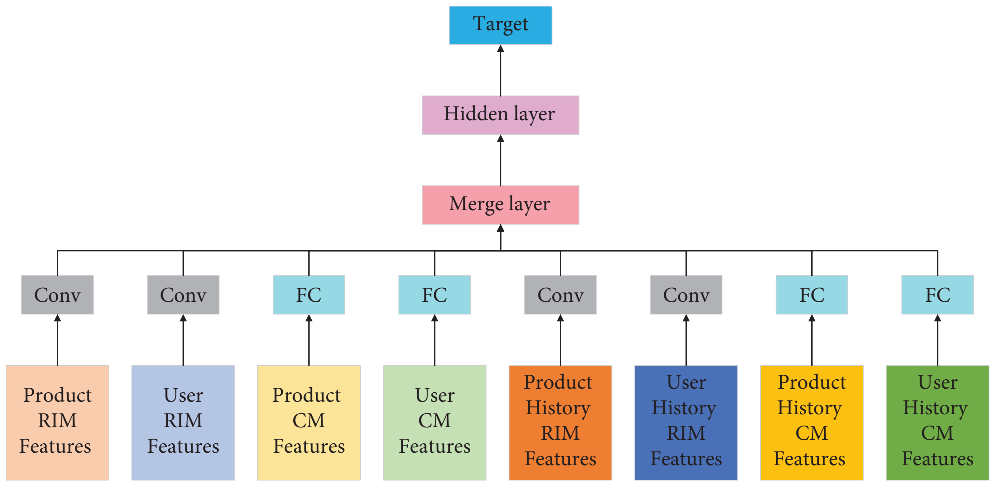


FIGURE 5: Overall structure of multiview prediction neural network.

As shown in Figure 5, in order to obtain the prediction of the product given by the user, two feature inputs are required in the feature extraction phase under the perspective of current perspective and historical perspective. For the current perspective, the corresponding feature vectors are input. The CM and RIM features of the user from the current and historical perspectives are entered on the left side of the network.

The CM features of users and products exist under different spaces, since they are trained independently of each other. To extract their features, two different fully connected neural networks are introduced to transform the CM features of users and products, respectively. The fully connected neural network is used because the CM feature vectors are holistic and each CM feature vector uniquely represents a user or a product. Then, the corresponding fully connected feature extraction layer is represented as follows:

$$\begin{aligned} \alpha_t(tj) &= g\left(W^e e_j\right), \\ \alpha_u(ui) &= g\left(W^r r_i\right). \end{aligned} \tag{4}$$

In contrast to CM, users and products, given K different rating categories, will each have K RIM feature vectors representing the different ratings. In order to use a unique feature to represent users and products, K different RIM feature vectors connecting users and products are used to represent them:

$$\begin{aligned} v_i(ui) &= \left[r_i^1, r_i^2, \ldots, r_i^k\right], \\ \mu_j(tj) &= \left[e_j^1, e_j^2, \ldots, e_j^k\right]. \end{aligned} \tag{5}$$

In order to efficiently process the historical information effectively, it is first necessary to perform further feature extraction from the set of users' evaluated products or the feature vector corresponding to the set of users who have evaluated products, and then use the extracted features for further prediction by neural networks.

The neural network is trained by the stochastic gradient descent method, and the evaluation results of the network on the validation set are used as a condition to end the training early. If the results do not improve within a few cycles on the validation set, then the training is ended. In addition, if the improvement on the validation set is small in one training cycle, the learning rate of the training will be halved. It is worth noting that any emergent CM or RIM feature vectors are not updated during the network training. This will destroy the semantic meaning of the CM or RIM feature vectors on the corresponding space. In addition, the structure of the multiview network results in the possibility of a feature vector appearing in different branches of the network, which makes training more difficult.

4. Experiments

4.1. Data Mining Experiments. In the experimental data, the ratio of the number of source domain data to target domain data is set to 7 : 3, and several target domains with different distributions are collected. The simulated scenario is a situation where it is difficult to obtain enough target domain data. The variation curves of predicted MAPE with MMD for different models are shown in Figure 6. From the curves, it can be found that the effect of migration learning is correlated with the maximum mean difference MMD between source and target domains, and if the MMD is small enough, it means that the target domain data are highly consistent with the distribution of source domain data, and the migration learning-based model works better than the original model. In this case, the source domain and the target domain can be regarded as in the same domain, then they can be used as a training sample of the network at the same time, if adding new layers may lead to an increase in network parameters, resulting in overfitting, which brings undesirable effects instead. Therefore, we take the measure of fine-tuning the whole original network directly using the target domain data.

The results of representative specific prediction experiments are shown in Table 1. From Table 1, we can see that there is a distance between the source domain and the target domain, that is, the value of MMD is larger, and our model works better. The model not only can effectively migrate the knowledge from the source domain but also allows the network to use the target domain data to learn new

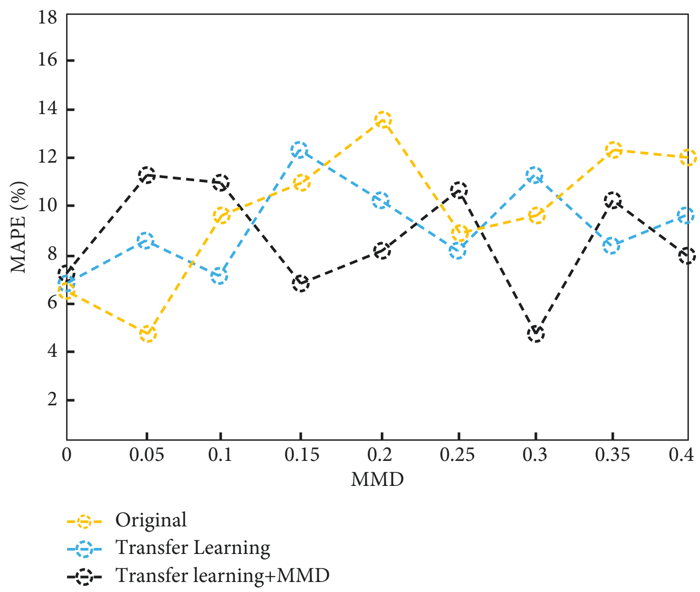


Figure 6: Change curve of MAPE predicted by different models with MMD.

Table 1: Representative prediction experiment results.

MMD	Original (%)	Transfer learning (%)	Transfer learning + MMD
0.045	11.26	9.64	12.92%
0.096	9.74	11.37	10.20%
0.158	19.37	8.96	11.83%
0.213	7.81	12.13	13.56%
0.279	8.35	10.42	15.49%
0.343	12.64	11.58	14.37%

knowledge by adding new layers to achieve the combination of the two domains' knowledge. However, if the MMD is too large, which means that the distribution of the source and target domains is too different, then negative migration will occur. No matter how you adjust the network result, the final migration result is worse than the original network trained with only the target domain data. In this case, the knowledge of the source domain is counterproductive.

A comparison of the prediction curves under different MMD-based models is shown in Figure 7. From Figure 7, it can be seen that our model fits better. In contrast, the original model deviates from the actual curve in several places, and the results are relatively poor. Specifically, the model performs best when the MMD belongs to the interval (0, 0.17), that is, only the target domain data need to be used to fine-tune the trained model for the source domain data. If the MMD is larger and falls in the interval (0.17, 0.35), the model needs to be retrained using the target domain for adding new layers, and finally, the learning rate is reduced and the full network is fine-tuned. When the MMD is larger than 0.35, the migration learning has a negative effect, and it is necessary to find another source domain dataset with stronger relevance.

4.2. Personalized Data Recommendation Experiments. In this section, we test the effectiveness of the proposed model applied to the Top-N recommendation problem. MovieLens-1M is used as the dataset, and the model parameters introduced previously are used. First, the evaluation criteria are introduced, and then the comparison method is presented, and finally, the experimental results and analysis are given. The same consistent test conditions are used. The detailed experimental setup is as follows: with 90% of the data selected as the training set and the rest as the test set. In particular, in the test set, only the samples with a rating of 5 are kept, and the rest are discarded directly. Despite the Top-N problem, the training set is used to train a model for score prediction during the training phase. Therefore, all training and model details follow the previous model. Recall is used as the test criterion to evaluate the performance of Top-N recommendations. The comparison methods are the following: (1) Movie Average—the N products with the highest average scores are recommended to users. (2) Bias Matrix Factorization (BMF)—a MF model that incorporates user and product biases. (3) SVD++—a hybrid model of MF method and neighbor-based method.

Figure 8 shows the recall of different methods at different N. It can be seen from the figure that the proposed method can achieve a recall rate of 0.32 for $N = 10$, and overall, the proposed method is significantly higher than the other models under different N. The experimental results indicate that the proposed method not only has good accuracy in score prediction but also performs well on Top-N. However, this does not mean that a method with high RMSE will necessarily guarantee good Top-N performance. For example, although BMF can achieve an RMSE of 0.85, it has a recall of 0.17 at $N = 10$.

In order to verify the rationality of the multiview structure, the performance under different viewpoint subnetworks and different combinations of subnetworks was tested, and the results are shown in Table 2. In the table, different subnetworks are classified into separate viewpoints and multiviewpoints. Under the current perspective, the performance with CM or RIM feature vectors alone is tested

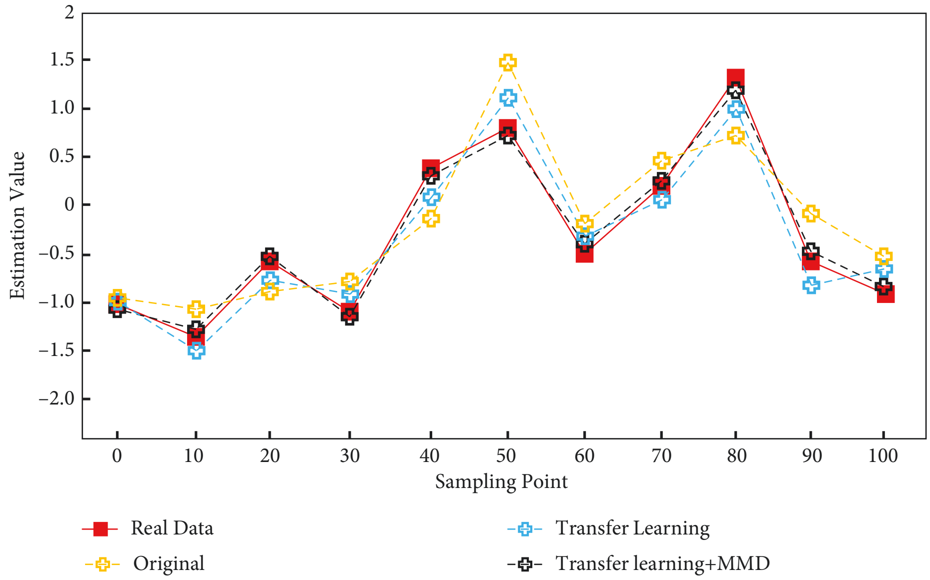


Figure 7: Experimental results of prediction curves of different models.

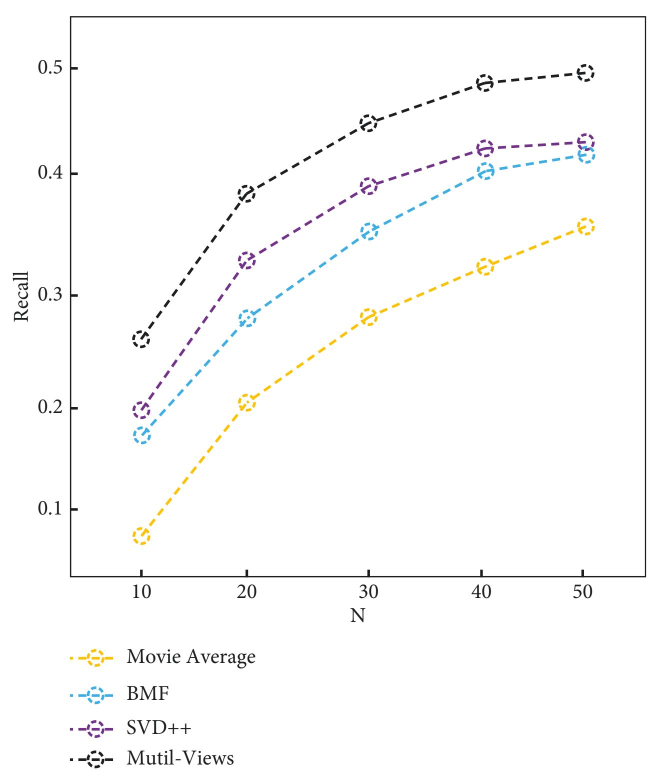


Figure 8: Recall rate of different N values.

separately. Similarly, two separate subnetworks using CM or RIM feature vectors are also tested under the historical view. Note that the use of both CM and RIM features in the current view is also considered as multiview here. Among all the single-view models, using the RIM feature in the historical view is the best performer, reaching 0.834, which is close to the final model's 0.821. This result indicates that the RIM feature in the historical view plays the most critical role in the final prediction of the multiview model. When both RIM and CM are used in the current perspective, the corresponding performance is better than the 0.837 and 0.838 of the CM and RIM single-view models, while the performance of using both RIM and CM in the historical perspective is only equal to the 0.852 of using RIM alone in the historical perspective. However, if all views are used simultaneously, the performance is still further improved to 0.834. Therefore, in general, the use of multiple views is better than the single-view model.

Table 2: RMSE experimental results from different perspectives.

View type		Vector model	RMSE
Single view	Current users and products	CM	0.837
	Current users and products	RIM	0.838
	Historical users and products	CM	0.853
	Historical users and products	RIM	0.852
Multiple view	Current users and products	CM + RIM	0.843
	Historical users and products	CM + RIM	0.834
	Historical and current users and products	CM + RIM	0.821

5. Conclusion

The advent of the era of big data has provided an opportunity for the development of soccer in China, especially for the selection and education of school soccer. It is especially important to deeply integrate big data with school soccer and give full play to the role of "big data" in the selection, teaching, training, and competition of school soccer. Based on the existing big data technology, a method combining data mining and data recommendation is proposed to promote the sustainable development of campus soccer. First, a data mining model based on migration learning and MMD is proposed for the verse mining problem, which improves the data utilization and data mining performance and efficiency. In which, MMD is introduced as a measure of the difference of data distribution between the source domain and the target domain, and the migration learning method and the adjustment network model are designed according to the size of MMD; finally, the data mining model based on migration learning and MMD is established. Then, for the personalized data recommendation problem, a deep collaborative filtering model that does not depend on content features using deep neural networks is proposed. The model learns the feature vectors of users and products, respectively, and the interaction between students and soccer sports information extracted by neural networks using the previously obtained features, and a multiview prediction model with two different feature vectors is applied. Finally, the proposed data-mining model is experimentally verified that the proposed method can not only improve the accuracy of data mining but also improve the efficiency of network learning; the accuracy of the proposed data recommendation algorithm prediction also exceeds the traditional method and the previous deep model. This paper is an attempt for the sustainable development of school soccer in the era of big data, and we hope to provide reference for other researchers. In the future, we plan to conduct research on sustainable development strategies for school soccer based on convolutional neural networks.

Data Availability

The datasets used during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This paper is one of the research results of the Higher Education Teaching Reform Project "Academic Achievement Research Based on Learning Participation" in Guangzhou Xinhua University, item number: 2020J016; and The Key Platforms and Major Scientific Research Projects of Education Bureau of Guangdong Province-Young Innovative Talents (Humanities and Social Sciences), China (Grant no. 2015WQNCX187).

References

[1] R. Millington, A. R. Giles, N. van Luijk, and L. M. C. Hayhurst, "Sport for sustainability? The extractives industry, sport, and sustainable development," *Journal of Sport & Social Issues*, vol. 46, no. 3, pp. 293–317, 2022.

[2] S. C. Kong, "Partnership among schools in e-Learning implementation: implications on elements for sustainable development," *Journal of Educational Technology & Society*, vol. 22, no. 1, pp. 28–43, 2019.

[3] B. Lusheng, M. S. Aman, and L. B. Hooi, "Revolutionary of school youth soccer management development in," *Journal of Positive School Psychology*, vol. 6, pp. 4487–4503, 2022.

[4] S. Loland, "Olympic sport and the ideal of sustainable development," *Journal of the Philosophy of Sport*, vol. 33, no. 2, pp. 144–156, 2006.

[5] D. Matarrita-Cascante, M. A. Brennan, and A. E. Luloff, "Community agency and sustainable tourism development: the case of La Fortuna, Costa Rica," *Journal of Sustainable Tourism*, vol. 18, no. 6, pp. 735–756, 2010.

[6] K. Readman, "The soccer finals as an analogy for sustainable futures in assessment," *Primary and Middle Years Educator*, vol. 7, no. 2, pp. 10–15, 2009.

[7] M. O. A. González, J. S. Gonçalves, and R. M. Vasconcelos, "Sustainable development: case study in the implementation of renewable energy in Brazil," *Journal of Cleaner Production*, vol. 142, pp. 461–475, 2017.

[8] B. M. McNoe and D. J. Chalmers, "Injury in community-level soccer: development of an injury surveillance system," *The American Journal of Sports Medicine*, vol. 38, no. 12, pp. 2542–2551, 2010.

[9] C. M. P. Meylan, J. B. Cronin, J. L. Oliver, M. G. Hughes, and S. Manson, "An evidence-based model of power development in youth soccer," *International Journal of Sports Science & Coaching*, vol. 9, no. 5, pp. 1241–1264, 2014.

[10] Z. M. Darus, A. K. A. Rashid, N. A. Hashim, Z. Omar, M. Saruwono, and N. Mohammad, "Development of sustainable campus: universiti Kebangsaan Malaysia planning and strategy," *WSEAS Transactions on Environment and Development*, vol. 5, no. 3, pp. 273–282, 2009.

[11] R. Agricola, J. H. J. M. Bessems, A. Z. Ginai et al., "The development of cam-type deformity in adolescent and young male soccer players," *The American Journal of Sports Medicine*, vol. 40, no. 5, pp. 1099–1106, 2012.

[12] C. Romero and S. Ventura, “Data mining in education,” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 3, no. 1, pp. 12–27, 2013.

[13] R. Baker, “Data mining for education,” *International encyclopedia of education*, vol. 7, no. 3, pp. 112–118, 2010.

[14] S. Olafsson, X. Li, and S. Wu, “Operations research and data mining,” *European Journal of Operational Research*, vol. 187, no. 3, pp. 1429–1448, 2008.

[15] S. Zhang, C. Zhang, and Q. Yang, “Data preparation for data mining,” *Applied Artificial Intelligence*, vol. 17, no. 5-6, pp. 375–381, 2003.

[16] B. Thuraisingham, “A primer for understanding and applying data mining,” *It Professional*, vol. 2, no. 1, pp. 28–31, 2000.

[17] C. Wang, Y. Zheng, J. Jiang, and K. Ren, “Toward privacy-preserving personalized recommendation services,” *Engineering*, vol. 4, no. 1, pp. 21–28, 2018.

[18] C. H. Chen and C. Toumazou, “Personalized expert recommendation systems for optimized nutrition,” *Trends in Personalized Nutrition*, vol. 12, pp. 309–338, 2019.

[19] K. Kesorn, W. Juraphanthong, and A. Salaiwarakul, “Personalized attraction recommendation system for tourists through check-in data,” *IEEE Access*, vol. 5, pp. 26703–26721, 2017.

[20] D. Geiger and M. Schader, “Personalized task recommendation in crowdsourcing information systems — current state of the art,” *Decision Support Systems*, vol. 65, pp. 3–16, 2014.

[21] K. L. O’Halloran, S. Tan, D. S. Pham, J. Bateman, and A. Vande Moere, “A digital mixed methods research design: integrating multimodal analysis with data mining and information visualization for big data analytics,” *Journal of Mixed Methods Research*, vol. 12, no. 1, pp. 11–30, 2018.

[22] C. H. Wu, “Data mining applied to material acquisition budget allocation for libraries: design and development,” *Expert Systems with Applications*, vol. 25, no. 3, pp. 401–411, 2003.

[23] Z. Huang, D. Zeng, and H. Chen, “A comparison of collaborative-filtering recommendation algorithms for E-commerce,” *IEEE Intelligent Systems*, vol. 22, no. 5, pp. 68–78, 2007.

[24] L. Xie, W. Zhou, and Y. Li, “Application of improved recommendation system based on spark platform in big data analysis,” *Cybernetics and Information Technologies*, vol. 16, no. 6, pp. 245–255, 2016.

[25] Z. Liu, Y. Ma, H. Zheng, D. Liu, and J. Liu, “Human resource recommendation algorithm based on improved frequent itemset mining,” *Future Generation Computer Systems*, vol. 126, pp. 284–288, 2022.

[26] W. Gang, A. Gunasekaran, E. Ngai, and T. Papadopoulos, “Big data analytics in logistics and supply chain management: certain Investigations for research and applications,” *International Journal of Production Economics*, vol. 176, pp. 98–110, 2016.

[27] D. Opresnik and M. Taisch, “The value of Big Data in servitization,” *International Journal of Production Economics*, vol. 165, pp. 174–184, 2015.

[28] Y. Sun, H. Song, A. J. Jara, and R Bie, “Internet of things and big data analytics for smart and connected communities,” *IEEE Access*, vol. 4, pp. 766–773, 2016.

[29] M. Hong, M. Razaviyayn, Z. Q. Luo, and J. S. Pang, “A unified algorithmic framework for block-structured optimization involving big data: with applications in machine learning and signal processing,” *IEEE Signal Processing Magazine*, vol. 33, no. 1, pp. 57–77, 2016.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 2752498, 10 pages
https://doi.org/10.1155/2022/2752498

Research Article

The Role Model of Inclusive Finance in Regional Economic Development Based on Carbon Neutrality Theory

Zeyi Yu

School of Management, Fudan University, Shanghai 200433, China

Correspondence should be addressed to Zeyi Yu; yu_zeyi@fudan.edu.cn

Received 21 June 2022; Revised 4 July 2022; Accepted 8 July 2022; Published 30 July 2022

Academic Editor: R. Mo

The most significant global challenge is global climate change; carbon emission reduction and carbon neutralization are the last way to solve global climate change. Achieving carbon neutrality will significantly impact the global economy and the national economy. The "dual carbon goals" proposed by China will also substantially impact the global economy and regional economic development. Focusing on regional economic development, achieving carbon neutrality is not only an opportunity for regional economic development but also a challenge. To achieve the goal of carbon neutrality, there must be institutional innovation and management innovation, which will provide new opportunities for balanced development among regions. With the development of inclusive finance, financial services for many disadvantaged groups and small and medium-sized enterprises have been further developed, thereby promoting economic development. With the development of Internet finance, big data, cloud computing, and other technologies, the development of inclusive finance, etc., the concept has also emerged. However, due to the short time of digital financial inclusion, the imperfect regulatory system, and the low degree of public participation, it is worth investigating whether the development of digital financial inclusion has a catalytic effect on the improvement of economic growth. Using the data of 31 provinces to build a panel regression model, this paper studies the market development, the development degree of digital financial inclusion, and the impact of digital financial inclusion on the relationship between factor markets and product markets and regions from the perspective of factor markets and product markets.

1. Research Background

The economic development of a country or region must adapt to broader development trends to obtain comparative advantages and effects. All countries that can seize this opportunity have achieved great success in the evolution of human society. In a country's development process, any region that can grasp the opportunities contained in the country's development trend will gain certain development advantages in the country's development model. From the point of view of the nation, grasping the trend of world development is the nation's priority; at the regional level, grasping the trend of national development is the focus of regional development. Climate change is the biggest challenge to world development today. The global response to climate change is an important issue to be solved urgently. However, due to the increase in extreme weather phenomena caused by global climate change, the increasing awareness of climate change among scientists, and the increasing importance of governments around the world about its impact. As an active party to the United Nations Framework Convention on Climate Change, China has not only participated in a series of agreements on temperature control advocated by the Framework Convention but also actively took measures to achieve higher temperature control indicators [1–5]. At the ninth meeting of the National Economic Committee, Xi Jinping conducted in-depth discussions on achieving carbon peaking and carbon neutrality. Carbon neutrality is the ultimate goal we must achieve in these three areas and the only way to combat global climate change.

2. Literature Review

Since the concept of financial inclusion was put forward, its policy effects have received widespread attention from the academic community. Scholars have researched the impact of inclusive finance on regional innovation, increasing residents' consumption, alleviating the imbalance of urban and rural economic development, improving poverty status, and promoting economic growth. Scholars at home and abroad have various definitions of inclusive finance. Wang defined financial development as the ratio of financial scale to economic scale in their paper on Establishing and measuring the evaluation index system for high-quality development of 15 sub-provincial cities in China in the new era [1]. Tian and Gao used these data to benchmark financial inclusion (the proportion of the population using formal financial services) in the Research on the driving mechanism of technological innovation [2]. Shi clearly defined the three dimensions of the development level of financial inclusion. The purpose of financial inclusion development is to enable all members of the economic society to obtain and use formal financial resources. It facilitates the efficient allocation of productive resources, potentially reducing capital costs [3]. Wei and Zhao believe that inclusive finance is an expansion of microfinance with marginalized and small-scale characteristics. A systematic and formal financial service system has been formed based on microfinance, and its core concept is to serve the low- and middle-income groups and the poor [4]. Qi believe that the practical significance of inclusive finance is to allow all social strata to have equal access to financial services so that financial resources can better support economic development, and areas with better economic development can Use scientific and technological means to promote the development of inclusive finance [5]. The related Research by Qi and Yuan et al. mentioned that digital finance has the advantages of convenient service, accurate matching, and low cost. This innovative digital finance can overcome the dependence of traditional finance on physical outlets, shorten the business chain, and improve the financial business process system [6, 7]. With the widespread application of Internet technology in modern medical care, e-commerce, remote communication, and other daily life, its application value in the financial field has also received widespread attention. With the digital characteristics of Internet technology, the advantages of digital finance have become more and more obvious. People began to think about using Internet technology to make financial inclusion play a greater role. In their Research, Xiao pointed out that the development of the Internet is a necessary condition for the rise of innovative financial services. New Internet financial products such as Yu'ebao and Alipay can only provide financial services such as payment and lending with the help of rich Internet platforms [8]. Digital Finance has broken through the limitations of time and space with its technological advantages to carrying out various financial services. It is a new financial service model that integrates artificial intelligence, mobile payment, and digital technology to realize fast payment, lending, investment, financing, etc. [9]. Sun's Research pointed out that the focus of financial innovation at this stage is the integration of finance and digital technology, which has changed the development mode of inclusive finance and gave birth to a digital inclusive finance development model. Digital financial inclusion is based on the remote transmission function of Internet technology, which expands the scope of financial services [10].

Since its introduction in China in 2006, inclusive finance has been highly valued by all walks of life. The existing ResearchResearch provides a good reference for this paper. From the review and analysis of existing literature, because the concept of high-quality economic development has been proposed relatively quickly, there are currently few articles on the relationship between inclusive finance, especially digital inclusive finance, and high-quality economic development. Some studies on financial inclusion, especially digital financial inclusion, are mostly based on theoretical discussions. However, the environment facing our country's financial development is complex, and its effects on the macroeconomy may deviate from micro theories. Therefore, based on theoretical analysis, we should also fully consider our country's current economic development status to conduct empirical Research. In addition, it will deeply analyze whether the Internet penetration rate will impact the implementation of digital inclusive finance policies to provide ideas and directions for the development of inclusive digital finance, as well as practical guidance for the realization of innovation-driven innovation and high-quality economic development.

3. Carbon Neutrality

Carbon neutrality has become the common goal of the global response to climate change. However, to achieve this goal, we need to face the following challenges: CO_2 emissions are enormous, and the average concentration of CO_2 in the air continues to rise; India, Russia, and other world Powerful countries have not yet committed to the time to achieve carbon neutrality [11]; The energy consumption structure is still dominated by fossil energy, with coal, oil, natural gas, and new energy being "four parts of the world." The proportion of new energy is relatively low. Energy transformation is facing severe challenges. The global distribution of new energy sources such as solar and wind energy has intermittent and spatial distribution differences, which brings severe challenges to the large-scale development of new energy sources; Hydrogen energy, CCUS, energy storage, etc. The application cost of technology is very high, and there is no large-scale commercial promotion and application. Our country's energy consumption and CO_2 emissions can be roughly divided into three stages: slow development (gentle slope area), rapid development (steep slope area), and stable development (platform area). From 1980 to 2001, our country's energy consumption was in a gentle slope area. The annual increase in energy consumption was 0.43×108 tons, and the CO_2 emission increased to 0.93×108 tons [12]. The annual increase in CO_2 and the energy consumption ratio reached 2.2; China's energy consumption showed a steep slope from 2002 to 2013, with an annual increase in energy consumption of 2.06×108 tons, an annual increase in CO_2 emissions of 4.50×108 tons, and the proportion of the

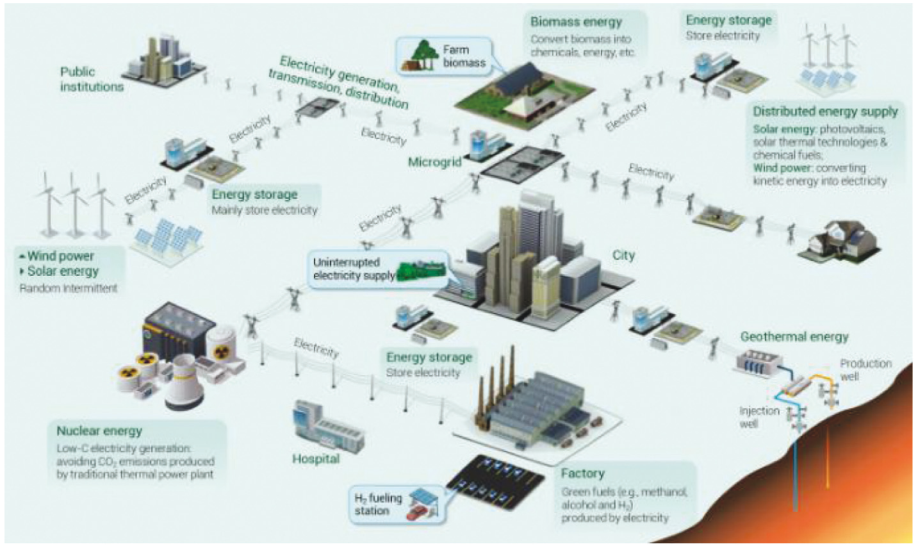


Figure 1: Carbon neutralization theory technology.

annual increase in CO_2 emissions in annual energy consumption was 2.2; the platform area was 2014. For the energy consumption in the period of ~2020, the annual new energy consumption will be 1.12×108 tons of standard coal, the carbon dioxide emission will increase by 0.81×108 tons per year, and the annual increase in CO_2 energy consumption ratio will be 0.7. Our country's energy consumption structure is mainly coal, oil and gas resources are relatively light, and new energy resources are relatively scarce [13]. Figure 1 shows the related carbon neutralization theory technology. In order to achieve carbon neutrality, China must transform the "big three, small" energy structure into a "three small and one big" new energy structure.

4. The Connotation of Carbon Neutrality

"Carbon Neutrality" aims at the dynamic balance between carbon dioxide emissions caused by human activities and the Earth's carbon cycle and uses carbon-free new energy to replace fossil energy in an orderly manner. The study of minimizing the impact of human activity footprint on the natural environment is an interdisciplinary subject between energy science and social science [14, 15].

"Carbon Neutrality" is a discipline of energy science and energy research idea centered on the interaction and coordinated development of the Earth, energy, and human beings. "Carbon Neutrality" includes three main research contents: (1) In the Earth system, the impact of energy generation and consumption on the Earth's climate and environment, reflecting the interconnection between the Earth and energy, (2) The Earth's environment is the evolution of human beings. The product is the impact of human activities on the environment and the interaction between man and nature; (3) The human beings use science and technology to develop energy and develop energy as the driving force, which is the interactive relationship between man and energy.

Establishing a carbon-neutral global energy system is integral to "carbon neutrality" research [16]. The earth's energy originated from the prosperity, extinction, burial, and evolution of the sun. With the evolution of the Earth system, it finally became an "artificial sun" (controllable nuclear fusion), thus realizing "from the sun, and then back to the sun." The energy cycle [17]. Under carbon neutrality, a "new energy" + "smart source" system will be formed with clean, carbon-free, intelligent, and efficient energy development goals at the core. Energy technology will be transformed from a resource-based type based on resource advantages to a technology-leading type [18]; our country's energy structure will be transformed from the direct consumption of single energy to the consumption of secondary energy of electricity from centralized use to intelligent, balanced energy management. The purpose of carbon neutrality is consistent with the purpose of energy research [19]. Figure 2 shows the financial analysis under carbon neutrality theory. The energy perspective reveals the symbiotic distribution of fossil energy and non-fossil energy in the earth system, the orderly replacement and transformation of carbon energy and hydrogen energy, and the harmony between the energy system and green earth—the laws of natural change in development.

5. Empirical Analysis of the Impact of Digital Inclusive Finance on Regional Economic Development Competitiveness Based on Carbon Neutrality

Assuming that there are n decision-making units DMU in this model, each DMU has m types of input items and s output items [20]. The corresponding weight vectors are recorded as $A = (a_1, a_2, \ldots, a_m)^T$ and $B = (b_1, b_2, \ldots, b_s)^T$ respectively. Among them, the input item vector of the jth is denoted as $X_j = (x_{1j}, x_{2j}, \ldots, x_{mj})^T$, $Y_j = (y_{1j}, y_{2j}, \ldots, y_{sj})^T$, $j = 1, 2, \ldots, n$. Among them: x_{ij} is the total input of

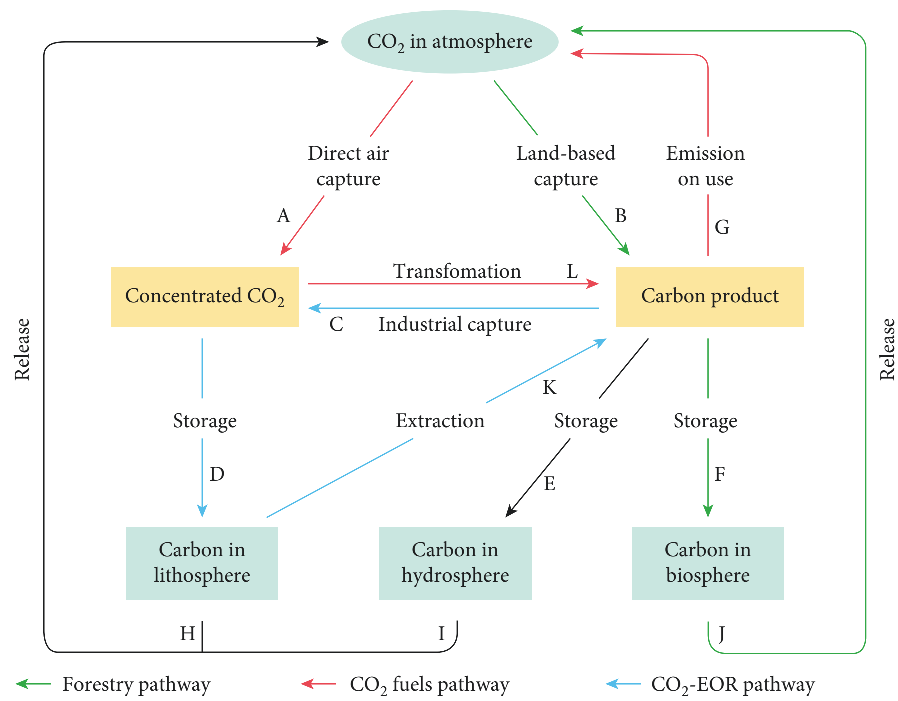


FIGURE 2: Financial analysis under carbon neutrality theory.

the jth decision-making unit to the ith type of element, y_{ij} is the total output of the jth decision-making unit to the rth type of product, $x_{ij} * y_{ij} > 0$; a_i is the input item of the ith type Weight coefficient, b_r is the weight coefficient of the rth output item, $a_i * b_r > 0$. Therefore, the relative efficiency evaluation index of the jth DMU factor input and product output is obtained as follows:

$$D_j = \frac{\sum_{r=1}^{s} b_r y_{rj}}{\sum_{i=1}^{m} a_i x_{ij}}. \quad (1)$$

Select suitable input weight A and output weight B to make the jth DMU relative efficiency evaluation index $D_j \le 1$. Perform performance evaluation on the kth DMU, denoted as DMU_k, and take the efficiency index of the kth DMU as the goal to obtain a general DEA optimization model:

$$\mathrm{MAX} = \frac{B^T Y_j}{A^T X_j}. \quad (2)$$

In order to simplify the operation of this model, it is transformed into an equivalent linear programming model, and the Archimedes infinitesimal variable ε is introduced. ε represents a positive actual number greater than 0 but less than any positive number, which constitutes the C^2R model of DEA. The specific dual linear programming model is as follows:

$$D(\varepsilon) = \min\left[\theta - \varepsilon\left(e^m\right)S^- + e^s S^+\right)],$$
$$\mathrm{S.t.}\left\{ \sum_{j=1}^{n} X_j\lambda_j + S^- = \theta K_K, \sum_{j=1}^{n} Y_j\lambda_j - S^+ = Y_k. \right. \quad (3)$$

Among them, $\lambda_j = 1, 2, \ldots, n, \lambda_j, 0$ is the dual variable; e^m is the m-dimensional unit vector, that is, e^s is the s-dimensional unit vector, that is $e^s = (1, 2, \ldots, 1) \in E_s$, S^+ and S^- are slack variables, $X_j = (x_{1j}, x_{2j}, \ldots, x_{mj})^T$, $Y_j = (y_{1j}, y_{2j}, \ldots, y_{sj})^T$, $X_k = (x_{1k}, x_{2k}, \ldots, x_{mk})^T$, $Y_k = (y_{1k}, k, \ldots, y_{sk})^T$, $\lambda_j \ge 0, j = 1, 2, \ldots, n$, $S^+ = (S_1^+, S_2^+, \ldots, S_s^+) \ge 0$, $S^- = (S_1^-, S_2^-, \ldots, S_s^-) \ge 0$.

Model C^2R was initially only set as the constant return to scale of the project production technology. After the continuous improvement of the model by researchers, the assumption of the project production technology was developed into a situation of diminishing returns to scale. If the constant return to scale is assumed (C^2R) Changed to the variable returns to scale assumption (VRS), the DEA model is perfected as:

$$V(\varepsilon) = \min\left[\theta - \varepsilon\left(e^m S^- + e^s S^+\right)\right],$$
$$\mathrm{s.t.}\left\{ \begin{array}{l} \sum_{j=1}^{n} X_j\lambda_j + S^- = \theta x_K, \\ \sum_{j=1}^{n} Y_j\lambda_j - S^+ = Y_K, \\ \sum_{j=1}^{n} \lambda_J = 1. \end{array} \right. \quad (4)$$

The relative efficiency calculated by the model under the assumption of VRS is pure technical efficiency. The relative efficiency calculated under the assumption of the C^2R condition is technical efficiency or total efficiency, which includes scale efficiency and pure technical efficiency.

The C^2R model shows that when the output Y_K of the kth DMU remains unchanged, the input factor X_k should be kept as low as possible in the same proportion. If C^2R obtains the best solution, and if $\theta_k = 1$, $S^{k-} = 0$, and $S^{k+} = 0$ are relatively effective at this time, the evaluated DMU is both technically practical and effective in scale; if $\theta_k = 1$, $S^{k-} = 0$, and $S^{k+} = 0$ are not zero vectors at the same time, the evaluated DMU is called weak DEA Effective, the decision-making unit does not satisfy both scale effectiveness and technical effectiveness, solve its VRS model, if $\theta_k < 1$, it is said that the evaluated DMU is not DEA effective.

6. The Role of Inclusive Finance in Promoting High-Quality Development

The role of inclusive finance in economic development has two aspects: First, SMEs are an essential link in our country's economic development. The "China Economic Census Yearbook" shows that by the end of 2019, the scale of small and medium-sized enterprises in China has reached about 70% [21]. In essence, inclusive digital finance is an organic integration of inclusive finance and digital finance, which can accurately mine and analyze user data, provide more accurate financial services, and improve service quality and efficiency. It provides powerful help to solve small and medium-sized enterprises' complex financing and financing difficulties. At the same time, inclusive development has also provided more financial services for people in remote and backward areas and promoted regional economic development. Essential financial services in China have covered 99% of the population, and the coverage rate of outlets in rural areas has reached 96% [22, 23]. Digital inclusive finance expands the coverage of financial services, lowers the threshold for marginal customers to obtain financial services, broadens the investment and wealth management channels for marginal customers, and increases the channels for marginal customers to obtain benefits. From the perspective of Keynesian economics, the endogenous growth of income can promote residents' consumption, expand domestic demand, and promote the high-quality development of the region.

6.1. Inclusive Finance Has Played an Essential Role in Promoting the High-Quality Development of the Regional Economy. The factor market is an essential factor in promoting high-quality economic development. In real life, factor markets, labor, capital, and other factor markets and product markets are optimally allocated according to the principle of efficiency to maximize the productivity of the entire society [24]. On this basis, accelerating the construction of the factor market system has great practical significance for promoting the high-quality development of our country's economy. The balanced development of the product market is conducive to realizing the rational allocation of the industry and the promotion of fair competition in the market, which is an essential factor in promoting the sustainable development of the regional economy and society. The development of the capital factor market can promote the sustainable development of the economy from the aspects of resource allocation, productive investment, and technological progress. At present, China's capital factor market is operating well. The yield of 10-year treasury bonds is around 3.2%, and good results have been achieved in resource allocation, capital accumulation, and technological progress [25]. Further development of the market is conducive to promoting the high-quality development of our country's economy. Figure 3 shows the theoretical analysis of financial inclusion under carbon neutrality theory.

Table 1 displays the high-quality economic development index of different provincial-level regions in China from 2011 to 2018. It can be seen from the table that the average value of the high-quality development index of the digital economy every year is close to the median value. Moreover, with time, the values of the two are getting closer and closer, which indicates that the development of digital financial inclusion in various regions has gradually tended to be balanced. At the same time, it can be found that the high-quality development index of China's digital economy in 2018 has increased by more than seven times that of 2011, with an average annual growth of 33.37% [26]. That shows the rapid growth trend of digital financial inclusion in China. At the same time, according to the different growth rates, it can be divided into three development stages, of which the first 5 years are the accumulation stage of development. At this stage, thanks to the breakthrough development of information technology, the development of inclusive digital finance is also moving toward a big step forward. After this stage, the development growth rate slowed down, and inclusive digital finance encountered a development bottleneck. Most provinces have a small growth rate at this stage, and even Ningxia has declined; 2016–2018 is a stage of stable development. With the introduction of the normative documents for developing inclusive digital finance, China's digital inclusive finance regulatory system is becoming more and more perfect. The corresponding development of inclusive digital finance has also entered a sustainable and stable development stage [27].

6.2. The Development Level of the Factor and Product Markets Plays an Essential Role in Improving the High-Quality Development of the Regional Economy. The combination of inclusiveness and digital finance will provide new impetus for developing the regional economy. First, promote the development of inclusive finance, promote the improvement of urbanization level, and infrastructure construction, maintain the supply of essential elements, and promote the development of emerging elements. Secondly, the development of inclusive finance can effectively guide the flow of funds to high value-added industries and increase the output of high-value-added products, thereby promoting the high-quality development of the regional economy; Financial markets are increasingly required to provide adequate support for the real economy, and at the same time, there are more and more problems caused by economic development and income disparities between regions [28]. In our country's middle- and low-income areas, the market for

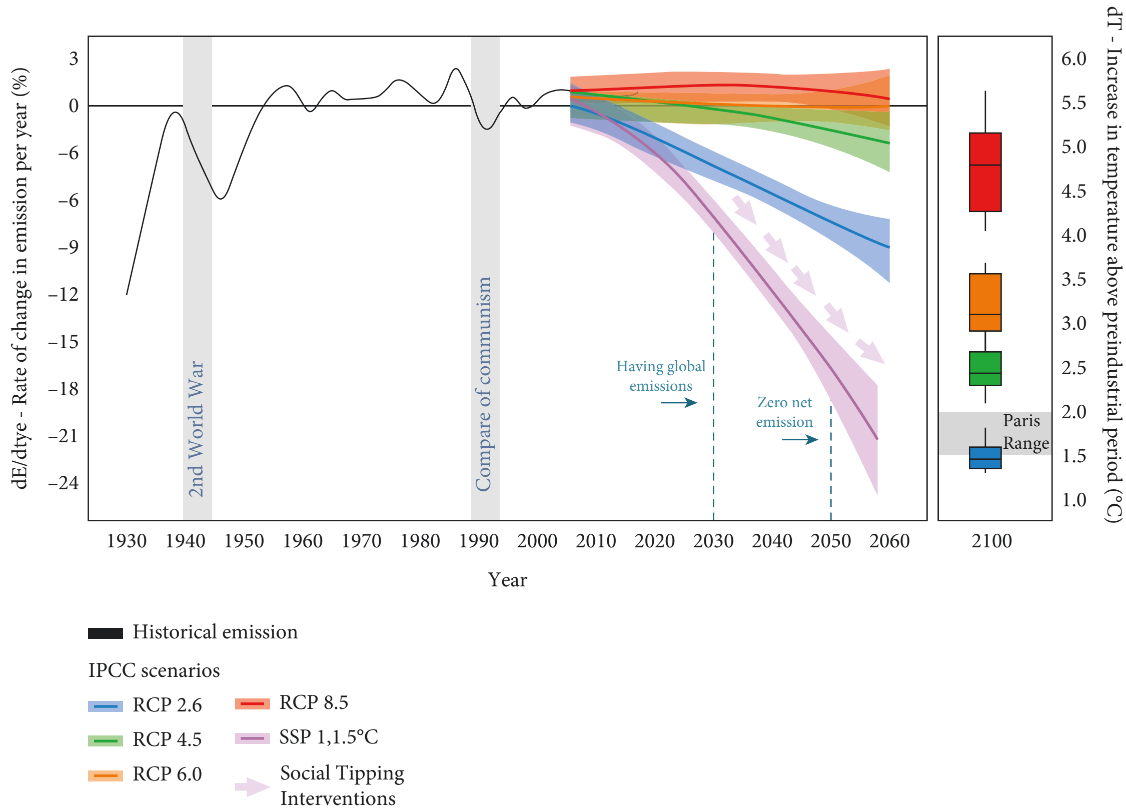


Figure 3: Theoretical analysis of financial inclusion under carbon neutrality theory.

Table 1: High-quality economic development index.

Area	Province	2011	2012	2013	2014	2015	2016	2017	2018
Central	Beijing	0.4174	0.4294	0.4393	0.4381	0.4552	0.4797	0.5241	0.5216
	Tianjin	0.2759	0.2918	0.2716	0.2815	0.2718	0.2215	0.2421	0.2198
	Hebei	0.1504	0.1501	0.1571	0.1601	0.1647	0.1719	0.1889	0.2173
	Liaoning	0.2538	0.2681	0.2551	0.2414	0.1912	0.1918	0.2161	0.2212
	Shanghai	0.3525	0.3569	0.3491	0.3535	0.3458	0.3453	0.3534	0.3611
	Jiangsu	0.3166	0.3462	0.3228	0.2928	0.3228	0.3122	0.3125	0.3629
	Zhejiang	0.2979	0.3300	0.3417	0.3415	0.3616	0.3614	0.3616	0.4130
	Fujian	0.2315	0.2524	0.2543	0.2424	0.2640	0.2967	0.2725	0.2817
	Shandong	0.1944	0.2492	0.2134	0.2254	0.2243	0.2346	0.2348	0.2643
	Guangdong	0.3531	0.3656	0.3832	0.3837	0.4002	0.4057	0.4371	0.5043
	Hainan	0.2014	0.2106	0.2007	0.2006	0.2024	0.2065	0.2179	0.2278
Central	Shanxi	0.1423	0.1666	0.1659	0.1646	0.1641	0.1896	0.1645	0.1831
	Jilin	0.1522	0.1590	0.1608	0.1639	0.1642	0.1702	0.1783	0.1979
	Heilongjiang	0.1792	0.1911	0.1980	0.1877	0.1838	0.1874	0.1885	0.2047
	Anhui	0.1644	0.1850	0.2002	0.1985	0.2099	0.2161	0.2213	0.2331
	Jiangxi	0.2007	0.2119	0.2094	0.2143	0.2224	0.2350	0.2428	0.2553
	Henan	0.1247	0.1393	0.1487	0.1543	0.1634	0.1724	0.1861	0.2027
	Hubei	0.1586	0.1674	0.1756	0.1837	0.1916	0.2057	0.2119	0.2253
	Hunan	0.1911	0.1945	0.1895	0.1894	0.1994	0.1949	0.2143	0.2194
West	Chongqing	0.2368	0.2463	0.2467	0.2547	0.2602	0.2507	0.2551	0.2872
	Sichuan	0.1720	0.1838	0.1893	0.1935	0.1954	0.1971	0.2341	0.2257
	Guizhou	0.1275	0.1529	0.1526	0.1625	0.1528	0.1633	0.1828	0.2525
	Yunnan	0.1308	0.1423	0.1567	0.1574	0.1575	0.1522	0.1596	0.1784
	Tibet	0.2486	0.2390	0.2715	0.2596	0.2545	0.2548	0.2729	0.2762
	Shaanxi	0.1583	0.1693	0.1836	0.1939	0.2337	0.2366	0.2163	0.2239
	Gansu	0.1036	0.1379	0.1339	0.1335	0.1330	0.1430	0.1338	0.1348

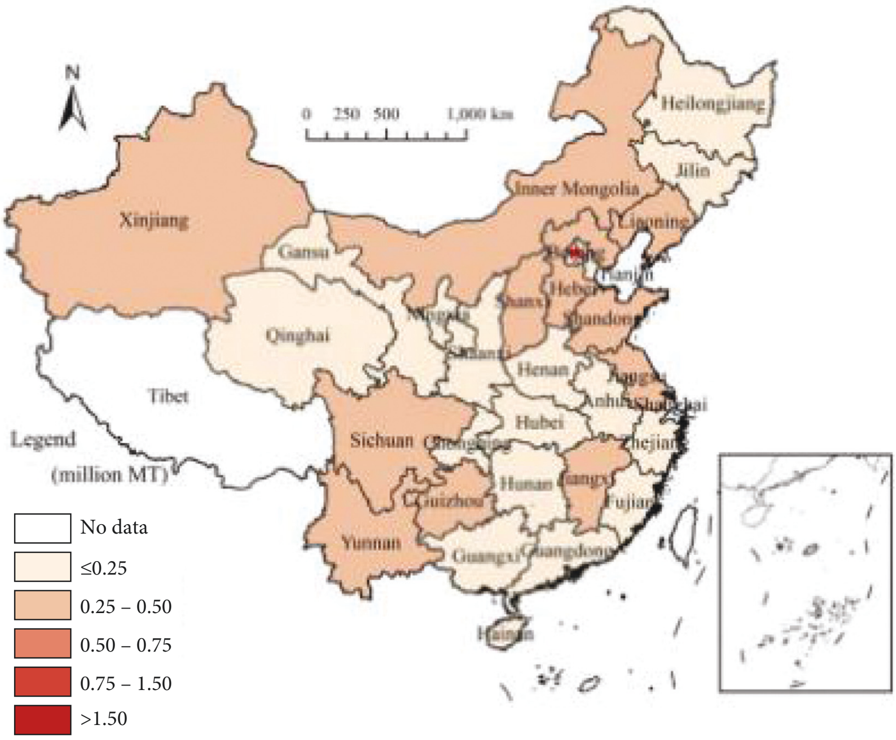


Figure 4: Data analysis of financial inclusion theory under China's carbon neutrality theory in 2020.

capital factors is relatively low-developed. The number of financial institutions is small, the level of financial services is low, and the people's financial awareness is poor. The popularization of inclusive finance in our country will provide more financial support for moderately underdeveloped regions and expand the coverage of financial services in poor regions, thereby enhancing the role of the financial market in promoting high-quality regional economic development [29]. Figure 4 shows the data analysis of financial inclusion theory under China's carbon neutrality theory in 2020.

7. Application Models of Inclusive Finance in Promoting High-Quality Development

7.1. Indicator Selection and Data Description. This paper takes 31 provinces and autonomous regions from 2011 to 2020 as samples, selects the following indices as research variables and uses the Wind database and the yearbooks of various provinces and cities for analysis.

7.1.1. High-Quality Economic Development. This paper firmly grasps the connection between the five development concepts and economic growth, and according to the research ideas of Cheng Xiang et al., shares the indicators of five dimensions, and uses the entropy weight method to establish the evaluation index of the high-quality development of the provincial economy. Among them, the main influencing factor of technological progress is technological progress, and the selected indicators include R&D investment intensity, technology market turnover, the number of invention patents per 10,000 people, and the full-time equivalent of R&D personnel in industrial enterprises above designated size; among them, urban-rural income ratio [30]. The urban–rural consumption ratio, the ratio of the output value of the secondary and tertiary industries to GDP are the coordination dimensions; the green dimension selection indicators include: forest coverage rate, green coverage rate in built-up areas, per capita park green space area, domestic waste harmless treatment rate, environmental protection Expenditure accounts for the proportion of fiscal expenditure, the number of days with air quality reaching or good; the indicators for selecting the open dimension are: trade surplus, the number of foreign-funded enterprises; the number of public library collections in public libraries, the per capita public library collections, and the number of public transport vehicles owned by 10,000 people, the number of urban workers who participate in the endowment insurance and the basic medical insurance for urban workers [31].

7.1.2. Inclusive Finance. The development of provincial-level inclusive finance was evaluated with the indicators of inclusive finance represented by the Institute of Digital Finance of Peking University as an indicator.

7.1.3. Market Development Level. China's provincial marketization index compiled by experts such as Fan Gang takes the product market development level as an indicator. The factor market contains more content, including financial marketization, human resource supply, marketization of technological achievements, etc. Therefore, we separate the financial market-oriented financial products from the development index of the factor market as a financial market. The development index and the excluded financial marketization index are used as the factor market development score (Factor) [32–49]. In addition to the development level of financial inclusion and the development degree of each market, this paper also introduces some control variables to reduce the estimation error caused by ignoring variables: trade openness (open): the scale of foreign trade affects economic growth. A decisive role. Fiscal Expenditure: As the primary measure for local governments to ensure economic development and maintain economic stability, fiscal expenditure plays an essential role in regional economic development. The article analyzes our country's fiscal expenditure with the ratio of provincial fiscal expenditure to GDP as an indicator [33]. Urbanization level: In recent years, with the rapid development of urbanization, the role of the urbanization level in the region cannot be ignored. The article selects the urban population ratio of each province as the evaluation index [34].

7.2. Mode Settings. This study aims to explore the impact of different markets and inclusive finance on high-quality development and further examine the impact of the development level of inclusive finance on our country's high-quality economic development.

In the actual economic operation, the economic activities of any region cannot exist independently, and there will be a certain degree of connection and interaction between economic units in different regions. Existing studies have shown obvious spatial correlations between Paving the way for low-carbon development globally and along the "Belt and Road" report released (Zai Xie) [35] and Supporting green and low-carbon development of economy and society with an interdisciplinary model system (Qu Shen) [36] in different regions. Therefore, this paper uses a spatial econometric model to empirically test the impact of digital finance on high-quality economic development and its spatial spillover effect. Spatial econometric models mainly include Spatial Error Model (SEM), Spatial Autoregressive Model (SAR), and Spatial Durbin Model (SDM). The general expression of the model is shown in formula (5):

$$Y_{it} = p\sum_{j=1}^{n} W_{ij}Y_{jt} + \beta X_{it} + \gamma\sum_{j=0}^{n} W_{ij}X_{jt} + \mu_i + \xi_i + u_{it}, u_{it} = \lambda\sum_{i=1}^{n} W_{ij}u_{jt} + \varepsilon_{it}. \tag{5}$$

Among them: Y_{jt} is the high-quality economic development index of the ith region in year t; p is the spatial autocorrelation coefficient of the explained variable; X_{it} is the set of all explanatory variables in the ith region in year t; β is the estimate of the corresponding explanatory variable coefficient; y is the spatial autocorrelation coefficient of each explanatory variable; W_{ij} is the spatial weight matrix element of the ith and jth regions; μ_i and ξ_i are the spatial and temporal fixed effects, respectively, u_{it} is the spatial error term; λ is each disturbance term of the spatial autocorrelation coefficient. When $p = 0$, $y = 0$ is the SEM model; $\lambda = 0$, $y = 0$ is the SAR model; $A = 0$ is the SDM model.

Table 2: Global index results.

Years	Digital finance		High-quality economic development	
	Index	P	Index	P
2011	0.0383	0.001	0.022	0.028
2012	0.0447	0.000	0.216	0.030
2013	0.0384	0.000	0.021	0.042
2014	0.0414	0.000	0.021	0.044
2015	0.0415	0.000	0.176	0.083
2016	0.0315	0.000	0.145	0.124
2017	0.0235	0.003	0.095	0.293
2018	0.0145	0.009	0.031	0.481

The results are given in Table 2. It can be seen that the indices of China's digital finance from 2011 to 2018 were all positive numbers and were statistically significant at the 5% level; the indices of high-quality economic development were also cheerful and were at the 10% level from 2011 to 2015.

The above is significant, indicating a certain spatial autocorrelation between China's digital finance and high-quality economic development, which is manifested in positive spatial aggregation. Therefore, it is more reasonable to use the spatial econometric model to explore the impact of digital finance on high-quality economic development.

8. Conclusions and Recommendations

From the perspective of factor markets and product markets, this paper uses data from 31 provinces to construct a panel regression model to study market development, the degree of development of digital financial inclusion, and the impact of digital financial inclusion in various factors markets and product markets—relationship with the region. According to our country's high-quality development level and inclusive finance development level, this paper studies the heterogeneous effects of different development levels in our country. The study found that the development of digital financial inclusion and the development of financial markets will both play a positive role in promoting the high-quality development of the regional economy, and financial inclusion can promote high-quality regional development by optimizing the development of factors, products, and financial markets. Develop. In addition, from the regression analysis of small samples, it can be seen that in areas with high levels of development, the regulation effect of inclusive finance on various markets and high-quality development levels is the most obvious; In contrast, the level of inclusive development in inclusive digital finance continues to improve, Its function of regulating various markets and economic development is also gradually weakening.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] S. Wang, "Efficiency analysis of Chinese science and technology finance based on the DEA-malmquist model," in *Proceedings of the 2nd International Conference on Economic Development and Innovation (EDI 2021)*, Hong Kong, December 11-12, 2021.

[2] L. Tian and J. Gao, "Research on financing modes of catastrophic risks in China:Cross-sectional or inter-temporal," in *Proceedings of the 6th International Conference on Innovation and Management*, London Uk, 26th July 2009.

[3] D. Shi, "Input and output analysis of financial innovation system in China," in *Proceedings of the 6th International Conference on Innovation and Management*, Changsha, China, May 17 to 18, 2009.

[4] W. Jingqiang and Y. Zhao, "A research on Chinese financial sustainable development," in *Proceedings of the 7th International Conference on Innovation and Management*, 7-7 July 2010.

[5] Qi Huang, "Study on rural finance against the background of Internet finance in China," in *Proceedings of the Second International Conference on management, economics and law*, China, 11 September 2021, https://doi.org/10.2991/aebmr.k.210909.015.

[6] H. Qi, "Impact of digital finance on corporate efficient: evidence from empirical research in China," in *Proceedings of the 2nd International Symposium on Economics, Management, and Sustainable Development (EMSD 2021)*, Hangzhou City, China, July 15-16, 2022.

[7] B. Yuan, K. Zhao, and X. Shen, "Digital finance and corporate value: an empirical evidence from China," in *Proceedings of the 2nd International Symposium on Economics, Management, and Sustainable Development (EMSD 2021)*, Hangzhou City, China, July 15-16, 2022.

[8] Ke Xiao, "The contribution of policy-oriented finance to independent innovation," in *Proceedings of the 6th International Conference on Innovation and Management*, Changsha, China, May 17 to 18, 2019.

[9] M. Yu, "Research on relative efficiency of input and output of sci-tech finance based on DEA model:A case study in shanghai," in *Proceedings of the Fifth International Conference on Economic and Business Management (FEBM 2020)*, 14 December 2020.

[10] S. Sun, "Development strategy of technology finance industry based on Internet thinking," in *Proceedings of the 2021 International Conference on Intelligent Transportation,Big Data & Smart City(ICITBS 2021)Part*, Xi'an, China, 27-28 March 2021.

[11] J. Peng, "Research on poverty reduction effect of inclusive finance in China ethnic areas," in *Proceedings of 4th International Conference on Economy, Judicature, Administration and Humanitarian Projects(JAHP 2019)(Advances in Economics, Business and Management Research*, vol. 94, Hong Kong, September 2019.

[12] M. Jiang, H. Yu, Guangxi inclusive finance research from the perspective of rural rejuvenation," in *Proceedings of the 2nd International Symposium on Social Science and Management Innovation(SSMI 2019)*, Xi'an City, China, November 29-30, 2019.

[13] J. Wang, "Digital inclusive financial and total factor productivity," in *Proceedings of the 2nd International Symposium on Frontiers of Economics and Management Science (FEMS 2021)*, Dalian City, China, May 15-16, 2021.

[14] Y. Zhang, "Inclusive financing and labor migration——evidence from China," in *Proceedings of the 2nd International Symposium on Frontiers of Economics and Management Science (FEMS 2021)*, Dalian City, China, May 15-16, 2021.

[15] J. Wang, G. Hu, and Y. Lei, "Research on inclusive financial system in China: from the perspective of financial support for disadvantaged fields," in *Proceedings of the 9th International Conference on Innovation and Management*, June 25, 2012.

[16] Y. Li and X. Wei, "The development of inclusive finance in chongqing," in *Proceedings of 4th International Conference on Economy, Judicature, Administration and Humanitarian Projects (JAHP 2019) (Advances in Economics, Business and Management Research,VOL.94)*, September 2019.

[17] B. Wu, J. Yang, T. Duan, and W. Miao, "An analysis of credit risk of middle and small-sized enterprises in the inclusive financial system:an example from China," in *Proceedings of The 2019 World Congress on Computational Intelligence, Engineering and Information Technology (WCEIT 2019)*, San Francisco, USA, October 22 - 24, 2019.

[18] C. Chang and C. Fang-Tzu, "Applying data mining in order to motivate inclusive finance," in *Proceedings of 4th International Conference on Humanities Science, Management and Education Technology (HSMET 2019)(Advances in Social Science, Education and Humanities Research,VOL.334)*, Kunming, China, September 29-30, 2016.

[19] Y. Wang, "Government behavior in the development of inclusive finance in China from the perspective of historical institutionalism: regulation or boost? Based on international case comparison," in *Proceedings of 2019 3rd International Seminar on Education, Management and Social Sciences (ISEMSS 2019) (Advances in Social Science, Education and Humanities Research, VOL.345)*, Zhuhai, China, January 14-15, 2017.

[20] Y. Wang, "Empirical research of poverty alleviation effects of digital inclusive finance from the perspective of internet——based on 2011-2015 provincial panel data," in *Proceedings of the 2018 3rd International Conference on Society Science and Economics Development(ICSSED 2018)*, Zhuhai, China, October 22-23, 2016.

[21] C. Zhang and F. Guo, "Financing path of small and micro enterprises of science and technology in China based on Internet finance," in *Proceedings of the International Conference on Communication, Innovation and Economic Management in 2020 (ICCIEM)*, Samuel Hall, December 26, 2020.

[22] Z. Fu and J. Jian, "Analysis of the relevancy between the Internet finance and small-and-medium-sized enterprises financing: based on the survey of wuhan E," in *Proceedings of the 14th International Conference on Innovation and Management*, 2017.

[23] Z. Zhou, "Analysis of marketing strategies for related products in China's financial market from the Internet perspective——taking ant credit pay as an example," in *Proceedings of 2nd International Symposium on Social Science and Management Innovation(SSMI 2019)*, Xi'an City, China, November 29-30, 2019.

[24] D. I. Shi, "Output analysis of financial innovation system in China," in *Proceedings of the 6th International Conference on Innovation and Management*, Changsha, China, May 17 to 18, 2019.

[25] H. Qi, "Impact of digital finance on corporate efficient: evidence from empirical research in China," in *Proceedings of the 2nd International Symposium on Economics, Management, and Sustainable Development (EMSD 2021)*, Hangzhou City, China, November27-28, 2021.

[26] B. Yuan, K. Zhao, and X. Shen, "Digital finance and corporate value: an empirical evidence from China," in *Proceedings of the 2nd International Symposium on Economics, Management, and Sustainable Development (EMSD 2021)*, Hangzhou City, China, November27-28, 2021.

[27] M. Cui, "Research on bank supply chain financial risk management based on game analysis," *Economic Times*, vol. 1744, no. 2, p. 022027, 2020.

[28] H. Li, "Internet financial risk and prevention analysis," in *Proceedings of the 3rd Guangzhou International Forum on Finance (GZIFF 2020)*, Guangzhou City, China, November 19-22, 2020.

[29] J. Han, "Risk prevention and countermeasures of Internet finance in China," in *Proceedings of 4th International Conference on Economics and Management, Education, Humanities and Social Sciences (EMEHSS 2020)*, Suzhou City, China, December 5-6, 2020.

[30] S. H. An, E. Ban, I. Y. Chung, Y. H. Cho, and A. Kim, "Antimicrobial activities of pp based topical gels," *Pharmaceutics FEMS*, vol. 13, no. 12, p. 2021, 2021.

[31] Yu Liu, "Vigorously promoting green and low-carbon development," *China Industrial Economics*, vol. 2021, no. 23, p. 223, 2021.

[32] H. Zhou, "The key to achieving the "dual carbon" goal is green and low-carbon energy development," *Forum of Chinese Party and Government Cadres*, vol. 2021, no. 11, p. 554, 2021.

[33] C. Chan, "New situation, new opportunities, and new challenges for Chengdu to promote urban green and low-carbon development," *Low Carbon Economy*, vol. 9, no. 02, pp. 45–61, 2022.

[34] Na Wen, "Effects, problems, and coping strategies of low-carbon development in Shenyang. Brand and Standardization," *Effects*, vol. 2021, no. 03, p. 212, 2021.

[35] Z. Xie, "Paving the way for low-carbon development globally and along the "Belt and Road" report released," *Comprehensive utilization of resources in China*, vol. 2021, no. 05, 2021.

[36] Qu Shen, "Supporting green and low-carbon development of economy and society with an interdisciplinary model system," *Environmental Management in China*, vol. 2021, no. 03, pp. 77–89, 2021.

[37] S. Jiang, "Countermeasures to promote ecological priority green and low-carbon development in our country," *Journal of Shandong Agricultural Engineering College*, vol. 21, no. 07, 2021.

[38] L. Jin and T. Che, "Research on the green and low-carbon development of Liaoning under the background of carbon neutralization," *Liaoning Economics*, vol. 2021, no. 03, p. 27, 2021.

[39] S. Wang and H. Wang, "Strategy research on the healthy development of China's digital economy under the background of dual carbon goals," *Contemporary economic management*, vol. 22, 2018.

[40] L. Li, W. Chen, and H. Wang, "Application research of digital RMB in green finance," *Rural Finance Research*, vol. 2021, no. 06, 2021.

[41] Z. Liu, Y. Zhang, and H. Li, "Digital inclusive finance, multidimensional education, and farmers' entrepreneurial behavior," *Mathematical Problems in Engineering*, vol. 2021, p. 6541437, 2021.

[42] L. Li and C. Mao, "Big data supported PSS evaluation decision in service-oriented manufacturing," *IEEE Access*, vol. 8, no. 99, pp. 154663–154670, 2022.

[43] L. Zhou and H. Wang, "An approach to study the poverty reduction effect of digital inclusive finance from a multidimensional perspective based on clustering algorithms," *Scientific Programming*, vol. 2021, p. 4645596, 2021.

[44] L. Li, C. Mao, H. Sun, Y. Yuan, and B. Lei, "Digital twin driven green performance evaluation methodology of intelligent manufacturing: hybrid model based on fuzzy rough-sets AHP, multistage weight synthesis, and PROMETHEE II," *Complexity*, vol. 2020, no. 6, pp. 1–24, 2020.

[45] T. Wang and C. Ji, "How does digital finance affect people's income: evidence from China," *Security and Communication Networks*, vol. 2022, p. 7025375, 2022.

[46] L. Li, B. Lei, and C. Mao, "Digital twin in smart manufacturing," *Journal of Industrial Information Integration*, vol. 26, no. 9, p. 100289, 2022.

[47] X. Liu, J. Zhu, J. Guo, and C. Cui, "Changnan. Spatial association and explanation of China's digital financial inclusion development based on the network analysis method," *Complexity*, vol. 2021, p. 6649894, 2021.

[48] L. Li, T. Qu, Y. Liu et al., "Sustainability assessment of intelligent manufacturing supported by digital twin," *IEEE Access*, vol. 8, pp. 174988–175008, 2020.

[49] Y. Li, H. Long, and J. Ouyang, "Digital financial inclusion, spatial spillover, and household consumption: evidence from China," *Complexity*, vol. 2022, p. 8240806, 2022.

Hindawi
Mobile Information Systems
Volume 2023, Article ID 9898413, 1 page
https://doi.org/10.1155/2023/9898413

Retraction

Retracted: The Relationship between Machine Translation and Human Translation under the Influence of Artificial Intelligence Machine Translation

Mobile Information Systems

Received 13 September 2023; Accepted 13 September 2023; Published 14 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] Z. Lihua, "The Relationship between Machine Translation and Human Translation under the Influence of Artificial Intelligence Machine Translation," *Mobile Information Systems*, vol. 2022, Article ID 9121636, 8 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 9121636, 8 pages
https://doi.org/10.1155/2022/9121636

Research Article

The Relationship between Machine Translation and Human Translation under the Influence of Artificial Intelligence Machine Translation

Zhao Lihua

School of Foreign Languages, University of Sanya, Sanya, Hainan 572022, China

Correspondence should be addressed to Zhao Lihua; zhaolihua@inu.edu.vn

Received 10 June 2022; Revised 23 June 2022; Accepted 28 June 2022; Published 13 July 2022

Academic Editor: R. Mo

Machine translation based on artificial intelligence has many commercial applications, such as Google translation, Baidu translation, and Youdao translation. More artificial intelligence and its translation are still used in all aspects of life. Therefore, we should reexamine its impact on the relationship between human translation and machine translation. Therefore, based on this background, this paper discusses the impact of the development of artificial intelligence machine translation on the relationship between human translation and machine translation. Although the translation accuracy and overall situation of machine translation based on artificial intelligence are similar to that of human translation, the basic algorithm of machine translation is still a program that judges right and wrong through computer code. It cannot simulate the "faithfulness, expressiveness, and elegance" of human translation in combination with social background and human culture. However, for some mechanical operations, such as business translation, scenes with low requirements, such as common vocabulary in daily tourism, can still meet the needs. Therefore, under the influence of artificial intelligence machine translation, the relationship between the two is that machine translation can replace human translation in some aspects, but it cannot replace human translation.

1. Introduction

Machine translation refers to the mutual transformation between two languages with the help of computers. Its fundamental algorithm is to integrate the corresponding relationship between the two languages into the word database in advance. When the language to be translated is input, the translated sentences will be split according to the word structure, and then the split required translation words will be transformed into the words corresponding to the translation language according to the corresponding relationship in the word database, Finally, it is integrated into the translated sentence meaning according to the part of speech structure of the target language [1]. Human translation means that the translator first masters and masters the two languages to be translated and then translates the language object to be translated into the translated language content after understanding the meaning. The two methods have their own advantages and disadvantages. In terms of accuracy, the accuracy of human translation is usually greater than that of machine translation [2]. However, in terms of the portability of translation, the convenience of machine translation is far more than that of human translation. This is because machine translation can get rid of the limitation of translators. It carries out machine translation according to the predetermined correspondence of database words, and there is no need for the object who puts forward the translation demand to master the translated language [3]. As far as the shortcomings of the two are concerned, the biggest disadvantage of human translation is that the quality of translation results greatly depends on the translator's translation skills. At the same time, because the translator digests the language to be translated first and then transforms it into another language he is familiar with, the meaning tendency of the translated content is easily affected by the translator [4]. In the age when the computer has not been invented, the communication between different languages mainly depends on translators to translate works such as characters so that people can understand the local customs and national culture outside different languages. Yan Fu, a famous translator in early China, once proposed that translation should follow the principle of "faithfulness, expressiveness, and elegance." Among them, "letter" mainly refers to the accuracy of translation. The translated content should correspond to the meaning of the original text and should not violate the meaning of the original text [5]. "Da" refers to the fluency of translation. Due to the different semantic structures and speaking methods between different languages, if the translation is carried out word by word, the content as a whole may not be smooth or reversed, making people unable to understand the meaning of the original text. In the process of translation, translators should not stick to the format of the original text, reintegrate according to the translated language, and make the translated content smooth and clear without violating the meaning of the original text [6]. "Elegance" refers to the fit between the translation content and the language. In the process of translation, for example, the sentence patterns and artistic conception of the two languages are different when the ancient text is translated into modern vernacular. If the ancient text is simply translated into modern vernacular, it will lose a bit of charm, so the translator should keep the original text as quaint as possible in the process of translation. In China's early translation works, translators followed the above principles for manual translation [7]. However, it seems unrealistic to require machine translation according to this standard. Machine translation mainly depends on computer language, and its underlying logic is binary language that can only distinguish 01. Therefore, computer translation cannot think about the smoothness and elegance of translation like human beings. Therefore, with the popularization of artificial intelligence machine translation, the problems of machine translation accuracy are well solved. In this context, we should also reconsider the relationship between machine translation and human translation. Therefore, this paper focuses on the principle of artificial intelligence machine translation and its impact on the relationship between machine translation and human translation.

2. Related Work

According to the time when the artificial intelligence algorithm appears, machine translation is mainly divided into two stages—traditional machine translation and neural machine translation, that is, what we call artificial intelligence machine translation. The origin of machine translation comes from the pursuit of hegemony between the United States and the Soviet Union. During the cold war, the United States had a demand for translation in order to learn more about the relevant information of the Soviet Union [8]. IBM was the first company to use computers to complete translation tasks. They used computers to translate Russian into English for the first time. This is the first time that human beings use computer to translate. In fact, the function of the translation system is very simple, which is very similar to today's common language manual. Nevertheless, IBM has also opened the competition for the development of machine translation among countries. However, in the 40 years since the rise of computers, machine translation has not made greater progress and remains at the stage of traditional machine translation [9]. In 1966, the American automatic processing Advisory Committee even declared that machine translation was useless, saying that it was not necessary to continue its research. Until the 1990s, with the further development of computer network technology, people began to carry out mathematical modeling and analysis through computer technology. At this time, it is found that a large number of parallel and research can be carried out on a language through statistical knowledge, and the statistical analysis results can be imported into the database to establish the best mathematical model, and finally machine translation can be carried out through the seen mathematical model [10]. This discovery once again made machine translation a hot research field and then achieved a lot of results. For example, Li explored the shortcomings of traditional machine translation in one translation according to the word comparison relationship, proposed to use the mathematical model to make entropy statistics on the translated content, and differentiated training on the statistical results according to the maximum entropy. The best training result is the translation priority and the highest content [11]. This method greatly improves the performance of statistics-based machine translation mechanization and greatly improves the translation results of words with multiple meanings. And this method is far more innovative than other studies in the same period. Then, in 2006, Google created Google translation platform, which is one of the largest translation platforms on the Internet. The underlying machine translation algorithm is still statistical machine translation, which means that statistical machine translation method has become the mainstream method in practical application. In the following time, most translation platforms adopt statistical machine translation method, but the biggest problem of this method is that the translation results mainly rely on the statistical results obtained after the computer analyzes a large number of parallel sentences [12]. However, insufficient attention is paid to the sentence meaning and word meaning of the translated object itself. In the process of translation, grammar and other problems often occur. Therefore, statistical machine translation algorithms are still under continuous research, but due to the limitations of the development of the times, these shortcomings of statistical machine translation have not been well solved. Until 2013, nal developed a decoder architecture suitable for machine translation based on neural network algorithm under artificial intelligence. The architecture uses the cyclic neural network for autonomous learning and converts the learning results into a continuous vector. Through the continuous training of the vector, the translation results of the target language are finally obtained [13]. This research is the first time to combine artificial intelligence with machine translation. Later, with the development of artificial intelligence neural network, it is found that artificial intelligence has the ability of autonomous learning, which can carry out autonomous learning on the existing results of statistical machine translation and obtain the most appropriate translation results again. And experiments have proved that the performance of artificial intelligence machine translation model is much better than that of statistical machine translation model when the learning corpus is sufficient. Therefore, more and more researchers began to explore machine translation under artificial intelligence [14]. In 2015, Red combined with the principle of psychological attention mechanism and introduced it into the artificial intelligence algorithm, aiming at the problem that artificial intelligence could not focus on the fixed continuous vector in the training process of translating long sentences, resulting in a large loss of source language information in the training process, which significantly improved the accuracy of artificial intelligence translation model [15].

3. Method

The accuracy of human translation can approach 100%, but it also depends on the level of translators, the level of expression of the original text, the industry, the time of submission, and other factors. The accuracy of machine translation depends on the language, industry, the quality of the original text, the training corpus, the training model, and other factors. From the perspective of translation fluency, human translation stresses "faithfulness, expressiveness, and elegance," but it will not be fully reflected in actual commercial translation. The development of mathematical logic has experienced continuous improvement. From the initial assumption of universal language to the establishment of logical typology. In essence, the establishment of mathematical logic is to symbolize the human thinking process, so a set of highly formal symbolic languages is constructed. To further explore the relationship between human translation and machine translation, we must first analyze the two translation methods and explore the relevance of the results from the method. The methods and meanings of human translation have been described in detail above, so this section only focuses on the traditional machine translation and artificial intelligence machine translation methods. In this part, we first introduce the most commonly used method based on editing distance in traditional machine translation methods as an example. On this basis, we further introduce the optimization direction of artificial intelligence machine translation with transformer model and compare their translation results with human translation results to find the similarities and differences. The transformer algorithm model based on artificial intelligence machine translation is shown in Figure 1.

Among the traditional machine translation methods, the machine translation method based on similarity calculation is one of the earliest and its translation methods. It is the first time that people use computers to translate between two languages. First, multilingual machine translation adopts a unique model framework, which can reduce some deployment or training costs. Unified training of a model will bring about some knowledge sharing. Some rich languages can transfer some knowledge to some low-resource languages, which can improve the translation effect of low-resource language pairs. At the same time, due to multilanguage mixing, some low-resource language pairs can see some inputs that were not seen before, which can improve the generalization ability of low-resource languages to the model. It has a certain creativity in model architecture and algorithm ideas. Therefore, this paper takes this as an example to explore the tradition and its translation algorithms. Computer translation algorithm based on similarity mainly comes from the field of speech recognition. It looks for the best translation result by converting the editing distance between the two words to be translated. The editing distance here refers to finding the best translation result of the content to be translated with the help of operations, such as insertion, replacement, and deletion between two words. Among them, different operation modes represent different translation costs. Therefore, we introduce the concept of word error rate Bleu. Generally, the algorithm definition expression of word error rate is shown in the following formulas.

$$\mathrm{BLEU}-n=\mathrm{BP}\times\ \exp\sum_{i-1}^{n}\lambda_t\log\ \mathrm{precision}_i, \tag{1}$$

$$\mathrm{BP}=\min\left(1,\frac{\mathrm{output}-\mathrm{length}}{\mathrm{reference}-\mathrm{length}}\right). \tag{2}$$

However, because the algorithm is converted from the field of speech recognition, the way of machine recognition is still quite different from that of speech recognition in the actual process of machine translation. For some sounds or word positions, the laws followed are not completely consistent, so there may be errors in the position of words or the meaning of translated words in actual translation. On this basis, some scholars have proposed an improved algorithm based on the position of words. The improved algorithm is less sensitive to the position of words in the translation process and abandons the past way of translating a paragraph of translated content according to a specific word order for the first time in the translation process but adopts an independent and continuous translation method. That is, when translating a paragraph of content, the translation order of each word does not need to follow its order in the whole translation content. In this way, the degree of freedom of word translation is higher, the whole translation result is more flexible, and the accuracy rate has increased. The specific improved algorithm expression is shown in the following formula.

$$\mathrm{BLEU}-4=\min\left(1,\frac{\mathrm{output}-\mathrm{length}}{\mathrm{reference}-\mathrm{length}}\right)\prod_{i=1}^{4}\mathrm{precision}_i. \tag{3}$$

When measured based on n-gram parameters, the formula can be transformed into

$$\mathrm{BLEU}-N=\frac{\sum_{s\in\{\mathrm{ReferenceSunnaries}\}}\sum_{\mathrm{gram}_n\in S}\mathrm{Count}_{\mathrm{match}}\left(\mathrm{gram}_n\right)}{\sum_{s\in\{\mathrm{ReferenceSunnaries}\}}\sum_{\mathrm{gram}_n\in S}\mathrm{Count}\left(\mathrm{gram}_n\right)}. \tag{4}$$

The change of the accuracy of the improved algorithm is shown in Figure 2. From the perspective of error accuracy, we divide the translation results into three categories:

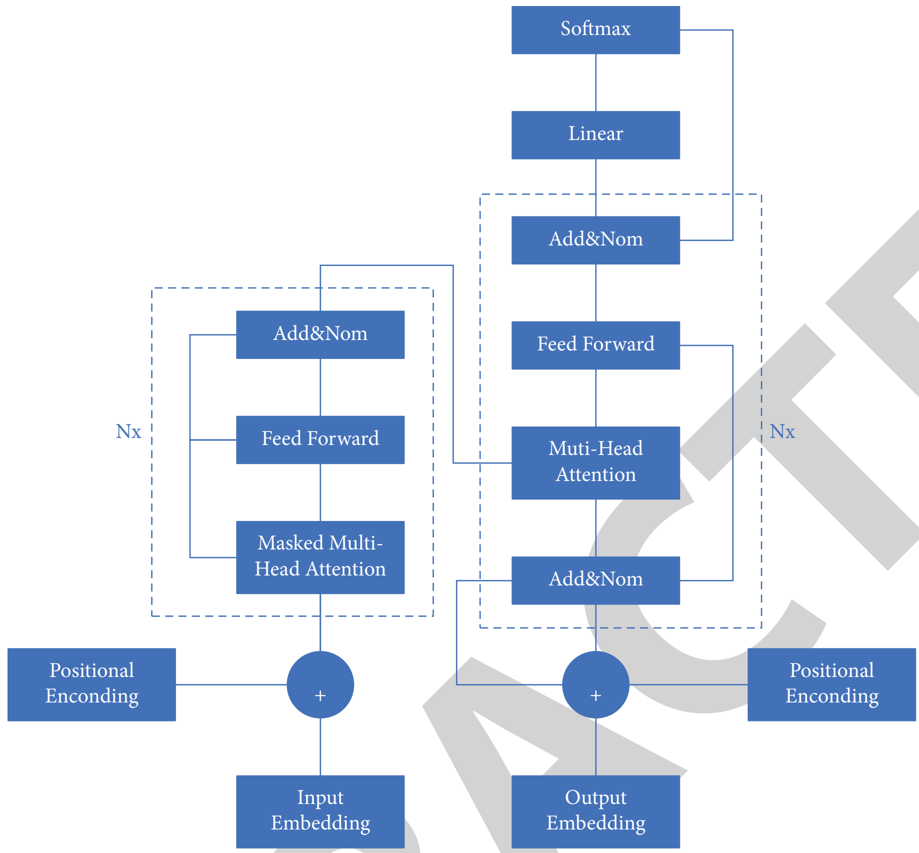


Figure 1: Diagram of transformer model based on artificial intelligence.

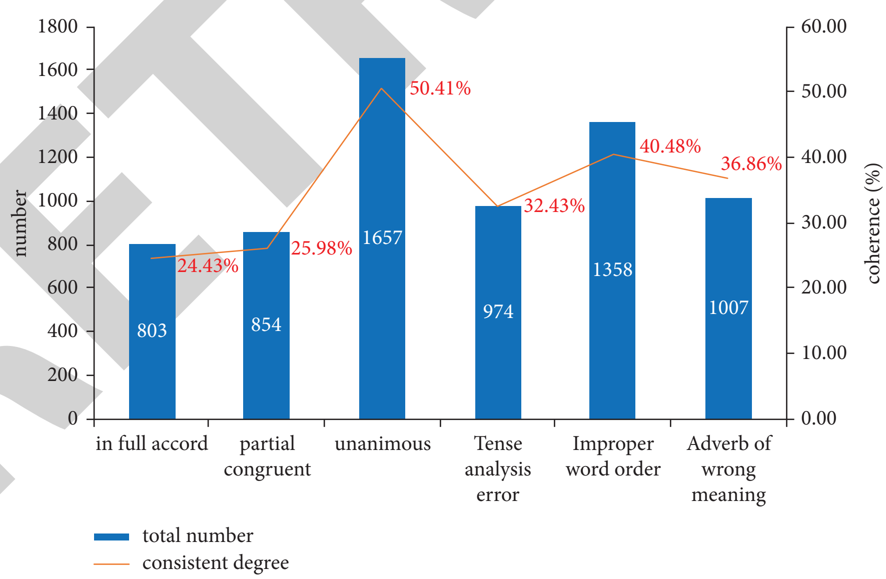


Figure 2: The number and consistency of errors in improved corpus translation.

complete consistency, partial consistency, and consistency. In addition, for some translation errors, we divide them into semantic errors, word order errors, part of speech errors. and so on. The measurement of accuracy is mainly based on the conclusion of the evaluators. For example, if the evaluators think that the translation result is inaccurate, the accuracy of the translation result is completely consistent. If some evaluators think that there is an error in a certain segment of the translation content, but other translators think that the segment is used alone, but there is an error when combined with other segments, we call it partial consistency error. On the other hand, the translator thinks that there is a semantic error in the translation; that is, if there is a semantic error in the translation, he thinks that there is no semantic error in the translation.

The artificial intelligence machine translation model is mainly based on the neural network algorithm, and its underlying logic is mainly completed by relying on the end-to-end algorithm principle. Compared with the traditional neural network, which only analyzes the best answer through statistical information, artificial intelligence machine translation can use the self-learning function of neural network to continuously iterate and update in the model training and finally find the best result. However, the algorithm as a whole is more complex than machine translation and has a certain premise of use. Generally, we believe that there is a certain space for the probability change when the set X composed of the input language to be translated and the set Y composed of the target language to be translated are closed sets. See formulas (5) to (7) for specific set expression and probability evaluation expression.

$$X = \{x_1, x_2, x_3, \ldots, x_m\}, \tag{5}$$

$$Y = \{y_1, y_2, y_3, \ldots, y_m\}, \tag{6}$$

$$p(y|x;\theta) = \prod_{j=1}^{n} p\left(y_j|y_{<j}, x; \theta\right). \tag{7}$$

After determining the applicable premise of artificial intelligence machine translation, the long-term and short-term memory network architecture is mainly used in the specific training of the model. The long-term and short-term memory network architecture usually takes the input layer as the coding segment and the output layer as the decoding end. In this way, we can train each translation result as many times as possible, effectively avoid the complex calculation process caused by artificial intelligence algorithm, and effectively improve the calculation efficiency. The calculation formula of long-term and short-term memory network is shown in the following formula.

$$i_\tau = \sigma\left(W_{ix}x_t + W_{ih}h_{t-1} + b_i\right). \tag{8}$$

However, in the process of practical application, we found that the value trained according to the formula cannot transfer all the training information to the next neuron when it outputs the final value through the sigmoid activation function, which means that the current training information cannot be transmitted evenly, which will eventually lead to inaccurate results. Therefore, we decided to choose tanh function with more convergent derivative as the activation function, which will produce more new information in transmission than sigmoid.

After different training iterations according to this algorithm, the error value and efficiency change of artificial intelligence machine translation results are shown in Figure 3. It can be seen that with the increase of training times, each training iteration increases the accuracy of the model, and its error value will decrease. That is, there is a training growth period during the operation of artificial intelligence

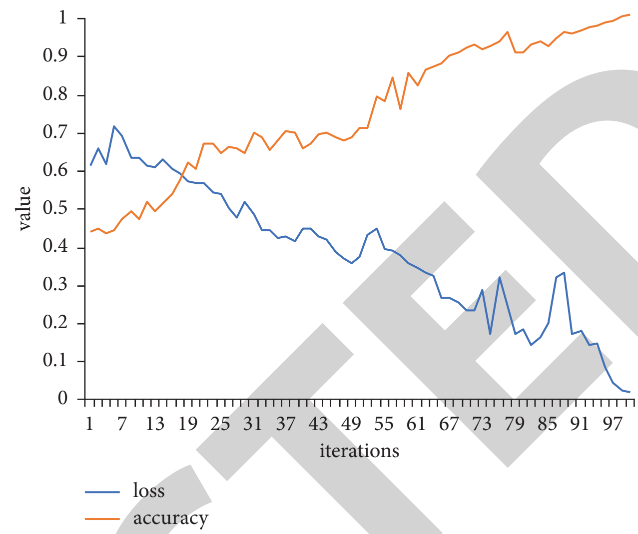


Figure 3: Error value and accuracy change of machine translation based on artificial intelligence algorithm.

machine translation. With the iteration of the model, the translation results will be better.

Finally, we evaluate the results of traditional machine translation and artificial intelligence machine translation. In order to better reflect the relationship between robot translation and artificial translation, here we use the artificial way to evaluate the results of traditional machine translation and artificial machine translation so as to pave the way for the follow-up exploration of the relationship between the two and artificial translation. We take the correlation coefficient as the evaluation method.

After determining the evaluation method, we evaluate the translation of 37 languages randomly selected from the commonly used English-Chinese human translation corpus, and take the recall rate and accuracy rate as the data indicators for the actual test.

Finally, we will sort out the evaluation results of traditional machine translation and artificial intelligence machine translation algorithms from the perspective of algorithm and model, as shown in Figure 4. Small probability events caused by random events should be avoided, which will affect the accuracy of the evaluation results. Small-probability events, that is, events with very small probability of occurrence (usually P ≤ or), have important applications in statistics. That is, such events can occur in theory, but the probability of occurrence is small, and the probability of occurrence in this test is almost zero. For example, winning the lottery is a typical small-probability event. There may be a grand prize in each issue (the probability is very low), but for a lottery winner, there is almost no possibility that he will win the grand prize by buying a bet (the probability that a small probability event will occur in a test). In fact, this is an important theoretical basis for the application of small-probability events in statistics—the principle of small probability. That is, the probability of a small-probability event occurring in a test is very small. If it does happen, statistics doubts its authenticity. The conclusion of statistics based on the principle of

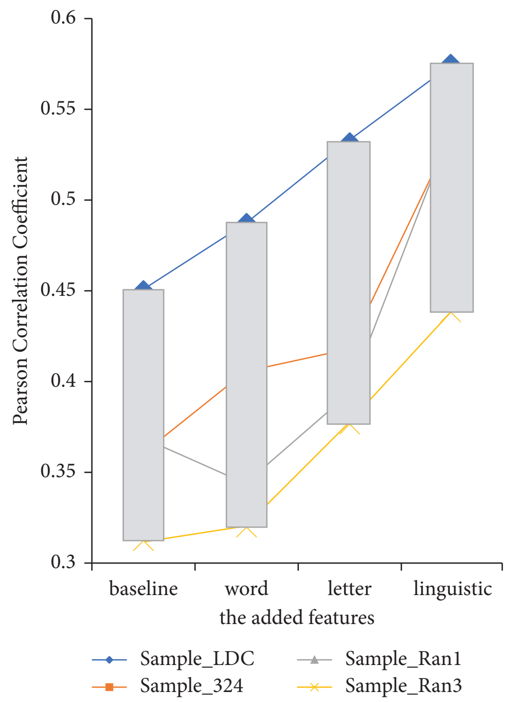


Figure 4: Comparison of model evaluation between traditional and artificial intelligence machine translation algorithms.

small probability is very correct, but there is also the risk of making mistakes. We selected four groups of samples from the English-Chinese human translation corpus and tested them on different evaluation dimensions. The results show that compared with the traditional robot translation, the machine based on artificial intelligence performs better in the accuracy of translation answers. In addition, it performs better in part of speech and word order.

4. Result Analysis and Discussion

Through the above exploration, we have a preliminary understanding of the meaning of human translation and machine translation, as well as the underlying algorithm logic of traditional machine translation and artificial intelligence machine translation. In order to further explore the relationship between human translation and machine translation, we also compare and evaluate the differences, advantages, and disadvantages between traditional machine translation and artificial intelligence machine translation. The advantages of traditional human translation include the following: the translator can interpret the context and convey the same meaning, rather than direct literal translation. The translator can understand the creative use of language, such as puns, metaphors, slogans, and so on. Machine translation features short processing time and faster processing speed than human translation. One tool can complete the translation of multiple languages and master many more languages than human translation. But the accuracy is not guaranteed. After mastering the relevant basic knowledge, we will discuss the relationship between human translation and machine translation in this chapter

Figure 5: Traditional machine translation system and practical operation mode.

and talk about the changes of the relationship between them in combination with the background of artificial intelligence machine translation.

The relationship between machine translation and human translation began with the birth of machine translation. It has been 63 years since the birth of machine translation in 1949. The development of machine translation is inseparable from the development of computer industry. Therefore, if you want to further explore the relationship between other human translation, you should analyze it in combination with the trend background of computer development. When computers were not widely used, machine translation was just a written theory. Warren Weaver, as a consulting theory researcher, first proposed the concept of machine translation in 1949, but it was not well applied. At this time, the relationship between machine translation and human translation began to appear, but the relationship is weak. The main translation method in the translation industry is still human translation. Until 1954, IBM and Georgetown University in the United States first integrated the translation function into the computer and invented the world's first IBM 701. Limited by the development of computers at that time, the translation machine was almost as large as the computer, with complex and clumsy operation. The specific model is shown in Figure 5.

However, in the following decades, computer translation was still used, only affected by the rigid algorithm, multiple semantics, and word order errors. At this time, the traditional computer translation was not widely used by the people. Therefore, the translation work was still more manual translation. After that, with the development of computer artificial intelligence algorithms, machine translation has also ushered in great changes. With the complexity and accuracy of algorithms and the wide promotion of computer applications, artificial intelligence machine translation is being used more and more in the translation industry. See Figure 6 for details.

It can be seen that machine translation is applied in all walks of life, among which the health care industry accounts for the highest proportion. This is because with the development of globalization and the development of national economy, more and more foreign products are imported

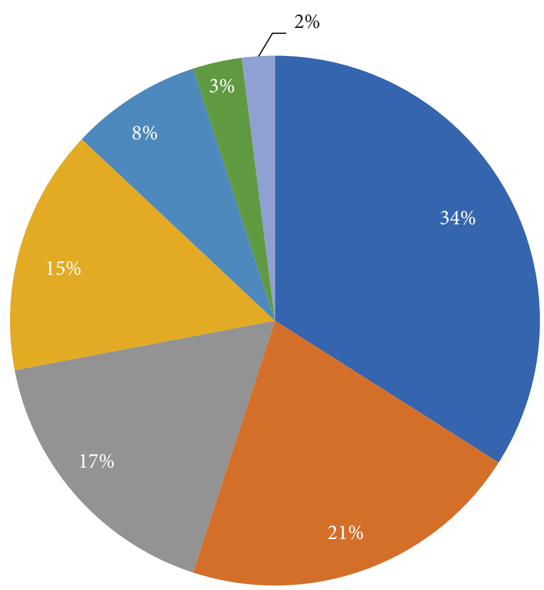


Figure 6: Proportion of machine translation applications in the industry.

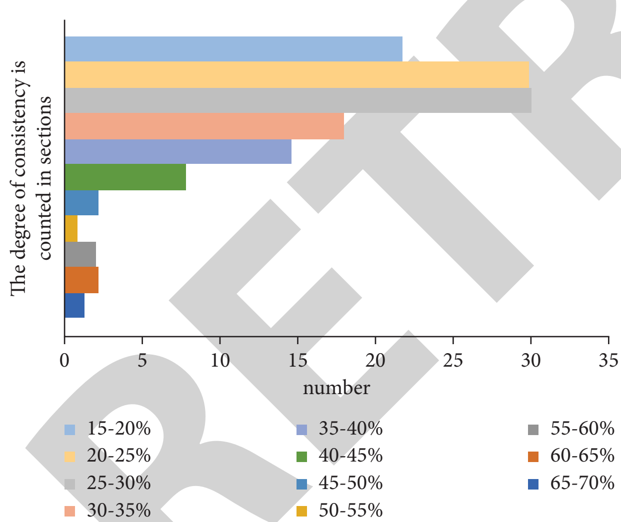


Figure 7: Comparison of consistency between artificial translation and artificial intelligence machine translation.

into China. The translation of products is relatively fixed, so brands usually choose machine translation with high efficiency and low cost. On the whole, at the present stage, with the development of artificial intelligence translation, the relationship between artificial translation and machine translation has gradually changed from human-oriented to machine-oriented to human-oriented. In order to further determine the reliability of this relationship, we compared the accuracy of human translation and machine translation based on artificial intelligence. The results show that there are few inconsistencies between artificial translation results and artificial intelligence machine translation results, and the inconsistencies are mainly maintained at 15–20%. It shows that although the translation accuracy and overall situation of machine translation based on artificial intelligence are not much different from that of human translation, the fundamental algorithm of machine translation is still a program for judging right and wrong through computer code, which cannot simulate the degree of "faithfulness, expressiveness and elegance" of human translation in combination with social background and humanistic culture, but for some mechanical operations such as business translation, scenes with low requirements such as common vocabulary in daily tourism can still meet the needs well. Therefore, under the influence of artificial intelligence machine translation, the relationship between the two is that machine translation can replace human translation in some aspects, but it cannot replace human translation (Figure 7).

5. Conclusion

This paper discusses the influence of the development of artificial intelligence machine translation on the relationship between human translation and machine translation. The advantages and disadvantages of traditional human translation include that the translator can interpret the context and convey the same meaning instead of direct translation. The translator can understand the creative use of language, such as puns, metaphors, slogans, and so on. Compared with human translation, machine translation has the characteristics of short processing time and fast processing speed. One tool can complete the translation of multiple languages and master many more languages than human translation. However, accuracy cannot be guaranteed. Therefore, this paper reexamines the impact of artificial intelligence on the relationship between human translation and machine translation. Although the translation accuracy and overall situation of machine translation based on artificial intelligence are similar to those of human translation, the basic algorithm of machine translation is still a program that judges right and wrong through computer code. It cannot be combined with social background and human culture to simulate human translation.

However, there are some limitations in this research. With the application of artificial intelligence technology in different fields, its potential risks have also caused ethical dilemmas at different levels. This will cause severe unemployment problems for the wide application of artificial intelligence in the future.

Data Availability

The experimental 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 regarding the publication of this work.