

Retraction

Retracted: An Empirical Study on Big Data Model and Visualization of Internet+ Teaching

Mathematical Problems in Engineering

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

Copyright © 2023 Mathematical Problems in Engineering. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1].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

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

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

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

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

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

References

- [1] H. Zhang and S. Tsai, "An Empirical Study on Big Data Model and Visualization of Internet+ Teaching," *Mathematical Problems in Engineering*, vol. 2021, Article ID 9974891, 10 pages, 2021.

Research Article

An Empirical Study on Big Data Model and Visualization of Internet+ Teaching

Hua Zhang ¹ and Sang-Bing Tsai ²

¹*School of Foreign Languages, Weifang University, Weifang 261061, China*

²*Regional Green Economy Development Research Center, School of Business, Wuyi University, Nanping, China*

Correspondence should be addressed to Hua Zhang; zhanghua198316@126.com and Sang-Bing Tsai; sangbing@hotmail.com

Received 4 April 2021; Accepted 29 April 2021; Published 11 May 2021

Academic Editor: Chenxi Huang

Copyright © 2021 Hua Zhang and Sang-Bing Tsai. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we conduct an in-depth study and analysis of “Internet+ Business English” teaching through teaching big data model and visualization and elaborate on the change of educational objectives in the context of Internet education. From the perspective of individual value coordinates, the goal of education is to adhere to the people-oriented, personalized, and comprehensive free development of human beings; from the perspective of social value coordinates, the goal of education is to cultivate innovative talents for social innovation and development; through the multiple perspectives of the curriculum in the context of Internet education, we analyze the goal orientation and value reshaping of the curriculum in the context of educational change. The curriculum of the future will develop toward an intelligent curriculum and introduces the curriculum design and the form of curriculum organization in the context of Internet education. A comparison of the constructed regression and classification of a total of 27 data models reveals that the model constructed based on all data in the integrated education system is the most effective. The multiple linear regression model explained up to 66.5% of the variance in student academic performance; the explanatory power of the social and demographic characteristics dimension variables ranged from approximately 13% to 18%, the personal characteristics dimension variables ranged from 7% to 20%, and the student input dimension variables ranged from 10% to 17%. The highest correct prediction rate of the binary logistic regression model was 69%.

1. Introduction

Since entering the information age, the rapid development and widespread popularity of computer technology and Internet technology have made the Internet everywhere and indispensable, gradually penetrating from life and entertainment to social and economic fields [1]. The deep integration of information technology and education has profoundly changed the status quo of the education industry, with the emergence of intelligent course selection systems, library lending systems, study room seat reservation systems, dormitory access control systems, etc., generating a large amount of student-related data [2]. The modernization of higher education governance has put forward new requirements and challenges to the informatization of education management, and the

informationization of education management is the only way to realize the modernization of higher education governance, and the teaching management information system is the concrete realization of the information of education management [3]. This is conducive to controlling student development trends and making macroeconomic adjustments. Allocate teacher resources according to the comprehensive quality and ability assessment of the student body to maximize the quality improvement of all students. According to objective evaluation, one of the paths to improve the educational governance ability of colleges and universities in the research on educational data includes the establishment of educational data collection, conversion and transmission standards, the improvement of educational data sharing mechanism, the exploration of educational data mining and analysis, the discovery of

educational laws behind educational data, the promotion of precise teaching management and scientific-educational decision-making, the better service to the improvement of teaching quality, the transformation of management mode, and the sustainable development of students in colleges and universities. This will help to improve the quality of teaching and learning, change the management mode, and develop the students.

With the advent of the cloud era, marked by cloud storage and cloud computing, how learners learn online is becoming increasingly diverse. The data from many cloud learning platforms provide strong support for personalized learning research and prediction [4]. However, since the existing classroom environment is still the primary environment for learning, it has richer data on learning behaviors. The cube is divided in 3D space, and the equivalence surface is generated based on the features of the vertices of each cube, and the class division is performed by considering the features of the values on the 8 vertices at the same time. The 3D data field is varied and coloured to produce an image on the screen (it penetrates the whole space like an X-ray and represents the distribution of matter and internal structure and information in a semitransparent way, simulating the propagation of light). It is extremely difficult for educators to design instructional processes based on the many learning behaviors of all learners in the classroom, and teachers commonly intervene only to manage discipline and participation in the classroom, without providing personalized assistance. Moreover, this traditional intervention is often not targeted and timely, making it ineffective in improving learners' effectiveness [5]. When this state persists, it may even cause learners' target learning outcomes to drift further and further away from their learning goals. Therefore, research related to teacher interventions using analytic techniques is indispensable.

Through the integration, management, and control of educational big data, modern management of higher education is realized from the perspective of data and relying on data evidence, which provides more scientific, objective, comprehensive, and intelligent management and decision-making for sustainable development of universities and makes higher education governance and educational research gradually scientific. From the perspective of pedagogy, evaluation models are needed to cross the gap between qualitative and quantitative evaluation, and they no longer rely on manual grading and expert assessment but rely on the power of data to truly present reality. It can comprehensively grasp the characteristics of teacher-training students, depict their growth trajectory, evaluate their academic level, and predict the development trend of learning, to achieve the goals of process-oriented, comprehensive, intelligent, personalized, and refined learning, management, and guidance of teacher-training students. At the same time, it constructs a trinity evaluation and observation system for students, teachers, and schools, providing an effective and feasible tool for reforming and exploring the model of teacher-training excellence.

2. Current Status of Research

Weber et al. proposed the ELM-ART intelligent interactive web-based teaching system, which provides interactive online learning resources, knowledge navigation, question answering, and analysis [6]. By analyzing the behavioural characteristics of Twitter users, Ma et al. found that highly extroverted users tend to relieve their emotions through social tools more than other users, linking the emotions of highly extroverted users to their behaviors and providing a new entry point for analyzing user behavior: user emotions [7]. Based on the results, Ma and Tsai further analyzed user behavior on the Twitter platform and concluded that the user's personality and individuality could be analyzed and predicted by the user's online behavior patterns [8]. In studying the characteristics and changes of the user's emotional state, Polyakova et al. found that the user's emotional state changed over time [9]. Dahdouh et al. conducted a study in the direction of emotion modeling [10]. They designed and implemented the AutoTutor system by acquiring the user's emotional state information through sensing devices, using the student's expressions, body postures, and conversational language as cues to build the student's emotional portrait [11].

Template-based methods are methods for extracting entity relationships from text using predefined templates and consist of two main types. One is based on trigger words or strings, and the other is based on dependency syntax, which can usually build rules starting from verbs and qualifying the lexical and edge dependency relationships on nodes [12]. The advantages of pattern matching classifier include being simple, having high accuracy, being easy to implement on small-scale datasets, and being simple to build [13]. The disadvantages are low recall of extraction results, domain-specific templates need to be constructed by experts, it is difficult to consider all patterns, the pattern needs to be defined for each relation, poor scalability, and difficulty to maintain. The supervised learning approach treats relationship recognition as a relationship classification task. In this case, relationship extraction is essentially a multi-classification problem, which requires first defining the relationship type and then labeling the data [14]. Faced with the emergent problem of too large a labor cost, many research efforts have increasingly focused on semisupervised or unsupervised learning methods. Currently, semi-supervised mostly uses distance-supervised methods to obtain a large amount of automatically annotated corpora to improve efficiency [15]. Unsupervised learning methods learn models mainly through clustering in a large corpus and then perform relationship recognition on new data. This method does not require a large amount of manually labeled corpora, but this weakly supervised learning method requires high data quality with the corpus, and noisy data will accumulate when automatically labeling relations, which affects the effectiveness of the model. The current research work focuses on both corpus quality and algorithmic models to improve the performance.

Construct a comprehensive, objective, quantitative, and data-driven index system for comprehensive quality

assessment of teacher-training students. Based on this index system, the data-driven representation of each index is studied to ensure the accuracy and usability of the comprehensive quality assessment model for teacher-training students. Considering the specificity of data in this study, when constructing the regression model of academic performance, based on the difference of data scope and data volume, the data model is constructed from the steps of pure objective data-combination of subjective and objective data, respectively, from the teaching management information system data, objective data in the integrated education data system, and all the subjective and objective data in the integrated education data system, respectively; the data volume of these three aspects are from large to small, but the number of study variables is from small to large.

3. Visual Analysis of Internet+ Business English Teaching Big Data

3.1. Internet+ Business English Big Data Model Design. The Bayesian school is different from the frequency school in judging some random distribution parameters, which is a random variable rather than a fixed value. If a person is lazy and unmotivated, the probability of failure will be higher. The posterior distribution obtained by Bayesian networks can be applied to some sample observations, and the estimates are more reasonable. The emergence of the Bayesian network is in line with the law of people's thinking in daily life and is very valuable for research. One is based on trigger words or strings, and the other is based on dependency syntax, which can usually build rules starting from verbs and qualifying lexical and edge dependency relationships on nodes.

The learning environment in the context of Internet education is a network environment with interactive capabilities based on computer networks. In a learning environment supported by computer network technology, learners can take advantage of a huge library of educational resources (including e-books, digital libraries, museums, and learning resource databases), powerful functions provided by computers (expensive devices that are normally inaccessible for use provided by network device sharing, cognitive functions provided by cognitive tools, guided learning functions provided by intelligent agents, etc.), and the network provided by communication capabilities (such as communication with learning partners, teachers, volunteers, and parents); this learning environment is ideal and can provide learners with the ability to expand their learning space in a nearly unlimited way, making possible the free realization of learners' desire for personal development [16]. As a new form of education that scientifically combines modern information technology and educational teaching practices, Internet education not only changes traditional education in terms of means and forms, but also gives a new meaning to education in terms of concepts, processes, methods, and the roles of teachers and students at many deep levels.

Both oral fluency of the teacher and English-only instruction bring auditory stimulation to the students and are

analyzed together here [17]. A fluent foreign language teacher can model the language well to the students and implicitly influence their oral expression. In the course, the teacher speaks fluently, slowly, and clearly, which is a positive model for the students' oral learning. At the bottom of the video, there is a "speed adjustment" function, so students can adjust the teacher's speech speed according to their needs, which is a kind of personalized setting [18]. Also, the intonation of the teacher is lighter and heavier, so that students can pay attention to the points when they need to emphasize them.

The deduction is a method of reasoning from the general to the individual by using known general principles to observe and study objects and deduce conclusions about things. The deduction is an important research method in scientific research, which can not only expand and deepen existing knowledge but also foresee the future scientifically and provide forward-looking clues for discoveries. For example, the theoretical deductive method is used in this paper, which discusses the coming of the era of the Internet of Everything and Teaching Internet. At the same time, the framework concept of the Internet of Things is deduced based on the IoT. This is because the Internet of Teaching is a specific application of the Internet of Things in the education industry. Teaching networking is the advanced form and inevitable trend of education in the context of Internet education [19]. The paper further deduces the framework of IET, including learner service, teacher and teaching service, education resource, and education management service. The paper also proposes to strengthen the exploration of the "education" level and return to the essence of education.

Principal component analysis (PCA) is a commonly used method for dimensionality reduction. High-dimensional data are generally very sparse, and in a high-dimensional hypercube, most of the data are at the boundaries. This makes analyses such as clustering and outlier analysis, which focus on the distance between points and the degree of aggregation, meaningless. And there is also data redundancy in high-dimensional data, which reduces the speed of processing data later and takes up storage resources. Therefore, it is necessary to represent high-dimensional data with low-dimensional data without losing information as much as possible. PCA style is the simplest method to reduce the dimensionality of random variables that are originally componential correlated:

$$X = (x_1, x_2, \dots, x_n)^2. \quad (1)$$

Converted to a new random vector with linear independence between components by orthogonal transformation.

$$Y = (y_1, y_2, \dots, y_n)^2. \quad (2)$$

PCA is a dimensionality reduction method based on linear transformation and projection. For example, if all the data points in 3D space are gathered around a certain plane, the projection of the data on this plane can be used to replace the original data and reduce the 3D data to 2D data without losing much information, such as the distance between two points. For methods using nonadaptive learning rate

transformations, the absolute value of the learning rate has a large impact on the performance of the model, and researchers often use the step change strategy. Some of the latest research methods, such as cyclical learning rate, are also being explored in academia. The problem that PCA solves is how to find the optimal one so that the projection of the original data on the plane is as dispersed as possible [20]. Mathematically, the dispersion of the projection can be expressed as the variance of the projection of the data on the axes. Therefore, it is necessary to calculate a new set of bases so that the variance of the original data is as large as possible under the representation of the new bases. And it is required that this set of bases is two-by-two uncorrelated, i.e., the covariance is 0. With two features a , b , the original data matrix after 0-averaging is

$$X = \begin{pmatrix} a_1 & b_1 & a_2 & b_2 \\ a_3 & b_3 & a_4 & b_4 \end{pmatrix},$$

$$C = \begin{pmatrix} \frac{1}{m} \sum_{i=1}^n a_i^2 & \frac{1}{m} \sum_{i=1}^n a_i^2 b_i^2 \\ \frac{1}{m} \sum_{i=1}^n a_i b_i & \frac{1}{m} \sum_{i=1}^n b_i^2 \end{pmatrix}. \quad (3)$$

This is a symmetric matrix. On the diagonal are the variances of features a , b , respectively, and off the diagonal are the covariances between the corresponding two features. Our original data is represented by a new basis that meets the requirements, then $Y = PX$, where P is the matrix of basic transformations from X to Y . The covariance matrix of D is

$$D = \frac{1}{m} (PX)(PX)^T. \quad (4)$$

In the model building and evaluation phase, the data set created in the data preparation phase is analyzed through a variety of modeling methods. Some modeling methods have specific requirements on the form of the data, and sometimes it is necessary to go back to the data preparation phase for some data collection and data cleaning. Multiple modeling approaches can be used for the same problem, and comparisons between multiple modeling approaches are required to determine the optimal approach. The result of the model building phase is a set of mathematical models or rule-based models. These models need to be evaluated by certain methods to determine their efficiency and quality. The modeling task of this study is relatively clear, which is to build a predictive model of the academic performance of students in University A. There can be multiple modeling methods to accomplish this task [21]. This study proposes to use traditional regression analysis modeling methods and data mining classification modeling methods to build academic performance prediction models, respectively, and it may also be necessary to compare the effectiveness and applicability of different algorithms in the classification modeling methods. In this study, both regression analysis methods and data mining methods in a general sense are referred to as educational data mining methods in a broad sense, as shown in Figure 1.

From the functional dimension, the education data of universities include basic data, status data, resource data, and behavior data. The basic data includes all kinds of basic information to guarantee education activities, the status data includes the operation status of education equipment, education environment, and education business, the resource data includes the teaching resources generated by the construction in the education process, and the behavior data includes the teaching behavior data of teachers and the learning behavior data of students in teaching activities, etc. The advantage of the Internet is that you can learn anytime and anywhere. Discrete data is numerical data with countable values between any two values; continuous data is numerical data or temporal data with infinite values between any two values; categorical data is data representing a finite number of categories or distinguishable groups stored with numerical data or text-based data.

Visualization technology uses computer graphics and image processing-related methods to convert abstract multidimensional data into human-eye recognizable graphic images displayed on a computer screen and provides a human-computer interaction interface so that people can manipulate and change the graphics through the interface. Visualization is an intersection of multiple theories, methods, and technologies, and visualization in computational science is an important branch of it. The main applications are line graphs, bar graphs, contour (surface) drawing, body drawing, etc.

In recent years, with the development of information technology and e-commerce platforms, the data stored in the database has become more diversified and unstructured, which makes the visualization of data begin to cover more domain contents. Data visualization requires one to be able to reorganize the picture from different dimensions to represent the relationship between data and data and to gain insight into the rules that follow the data being. The main methods used are a parallel coordinate method, pixel-oriented method, tree diagram, spider web diagram, etc.

$$P(\text{grab}|\text{correct}) = P(\text{grab}) \frac{P(\text{correct}|\text{grab})^2}{P(\text{correct})^2}. \quad (5)$$

The equation predicts the probability of passing the user's coursework corresponding to the user's corresponding answer profile based on historical a priori data for all users on the system, which can be used as a criterion for intercourse evaluation. After a user has taken any test paper question from the interactive aptitude test component, the system can evaluate the user's individual comprehensive coursework pass rate based on the latest probability combined with the user's answer rate PCCPA, which can be formalized as

$$\text{PCCPA} = \frac{\text{PECS}}{\text{CCPA}}. \quad (6)$$

From a modern educational and psychological point of view, a person's interest in learning is a positive psychological state characteristic of his knowledge and a force that can push him to continue his research. For example, if a

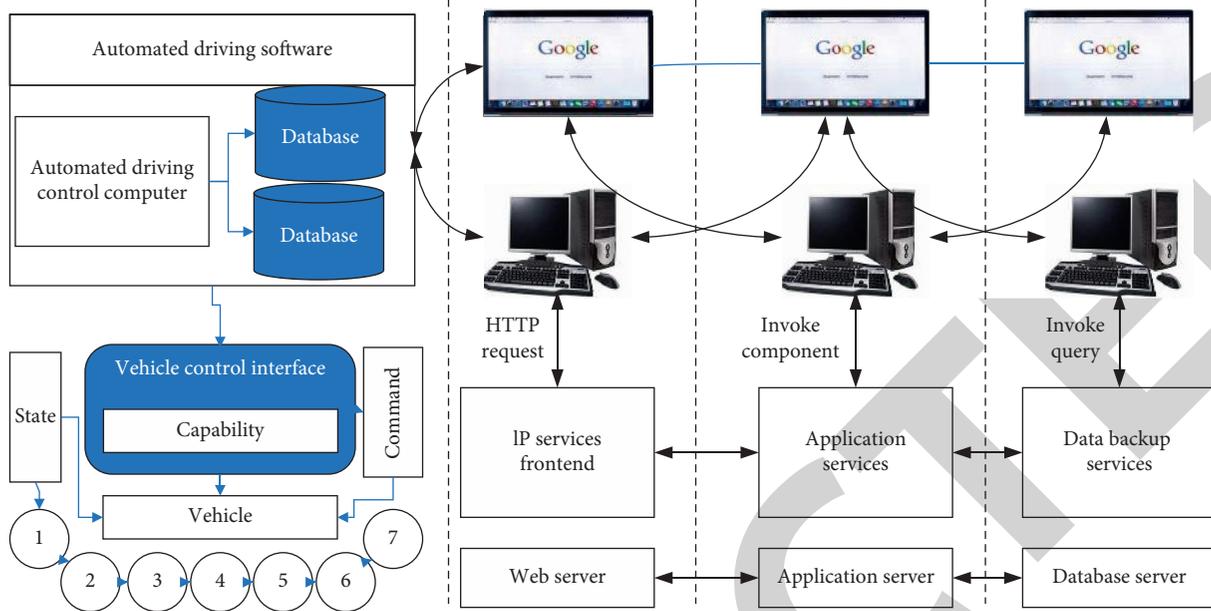


FIGURE 1: Internet+ Business English big data model design.

student has a strong interest in a subject, then he will spare no effort to study it further, and in addition to showing an active state in class, he will also concentrate on studying it outside of class. According to the research, this type of interest can be broken down into several categories. The first is direct interest and indirect interest. Direct interest means that the thing that people are interested in has a strong attraction that makes individuals feel good about something and study it spontaneously; indirect interest does not mean that an individual is interested in persevering with it under the influence of some medium, but that an individual learns it spontaneously, mainly because he or she realizes that he or she needs this kind of knowledge very much, for example, when it is related to personal or collective interests. There are also some examples, for example, when an individual wants to be recognized by the people next to him, he will try to learn to express himself and slowly develop his interest.

3.2. *Visualization Analysis Design.* The pixel-oriented method is a visualization technique based on scatter plots and computer pixel displays. Its approach is to map each data to a color space, find a color for the data values, and represent it as a two-dimensional, colored pixel point. By tracing many data points in two-dimensional space, relationships between data can be explored from the color distribution. This method requires the highest possible resolution of the computer display device to be able to represent larger amounts of data and discover more precise patterns. For example, when depicting canteen consumption data, each consumption bill is abstracted as a point, marked on a geographic location, and mapped to a color interval plotted with the number of consumption amounts, so that an ethical pixel map will visually reflect the consumption intensity and dollar amount of each stall in the canteen.

The spider web diagram is an icon-based visualization method that presents n -dimensional data in a two-dimensional chart. The spider web diagram method first sets up multidimensional attributes as axes starting from the same point and firing in different directions, with the axes scaled evenly from the minimum to the maximum value. Then, by tracing points on each dimension and connecting the lines, a graph like a spider web picture is derived. The method is like the parallel coordinate method, but the visual effect is very different. It can also be used to demonstrate the strengths and weaknesses of students in various areas.

The teaching goal includes the goal of teaching as well as the goal of learning. It is a kind of goal that is finally achieved through teachers' teaching and students' learning. It is a very important element in instructional design. Teachers should carefully design the teaching methods they choose when preparing lessons. The design of teaching methods is to determine the use of teaching methods: what content to use what method. Note that there is often more than one method of teaching in a lesson, and it is better to design the method specifically for each subtopic or each knowledge point. All assessments in quality education should be a motivator to move students into subsequent learning with the good heart of a successful person and never to find loopholes in students. To this end, with the purpose of motivation, we should give different award levels in the multidimensional evaluation of teachers' assessment and self-assessment, mutual assessment, and parents' assessment, regardless of which activities are outstanding or have made progress, so as to minimize the psychological burden of horizontal comparison among classmates, so that students can see their progress and achievements, and everyone can experience the joy of success and continuously enhance their self-confidence.

A tree diagram is a visualization method that shows the structural hierarchy between data. The idea is to divide the n -dimensional data space into different levels, with each level containing the next level under it. The tree diagram divides the screen into subrectangles of different sizes according to the data hierarchy, and the subrectangles recursively divide the hierarchy below them [22]. According to the number of word nodes m , the rectangle is divided into m subrectangles, either horizontally or vertically, and the two ways are crossed to distinguish the sibling and parent relationships between different rectangles. The size of the delimited rectangle depends on the data size of each child attribute. The attribute data size at a higher level is the sum of the attribute data sizes at lower levels. The tree diagram containing the hierarchy is well suited for displaying the hierarchical structure of the teacher education comprehensive quality assessment model, and the rectangular division of the tree diagram is well suited for displaying the different content cards in chunks on the screen.

As shown in Figure 2, the comprehensive recommendation strategy combined with interactive guidance gives two types of recommendations: recommendation index and comfort zone guidance. The recommendation index is the application of the comprehensive passing rate of all users, which shows users the recommendation rate of each courseware by sorting and uses the courseware with the largest recommendation index as the blackout recommendation by default; the comfort zone guidance will show the courseware suitable for learning within the user's comfort zone, giving users a feeling of easy access and familiarity and playing the role of guiding individual learning interests. At the same time, in the process of learning, visualization of the user's learning growth process can increase the interest in contextual learning, making users feel the positive response of the system to their individual and understand the role of the course in improving their knowledge mastery level. The visualization of the user's progress helps to increase the user's confidence in learning and thus facilitates the process. At the bottom of the video, there is a "speed adjustment" function, so that students can adjust the speed according to their needs. The teacher's tone of voice is also light and heavy, so that when the emphasis is on the key points, the pronunciation is heavier to draw the students' attention to them. The visualization of the learning process includes the visualization of the user's proficiency level and the status and gaps of the target level, and the visualization of the user's learning ability assessment. The visual display of the user's proficiency level of the course is based on the user's chapter assessment method as the data statistics, and the average of the user's quiz score for the course material is summarized by chapter dimension. The target level is a preset target value set at the time of course creation, set separately by chapter, giving an assessment criterion, and the user's chapter assessment method target is the preset target value. The user's learning ability is used in this recommended strategy as a demonstration of learning ability in the form of the individual's composite course pass rate PCCPA.

According to the study, this type of interest can be specifically subdivided into several categories. First, there are

direct and indirect interests [23]. Direct interest means that the thing that people are interested in has a strong attraction that makes individuals feel good about something and spontaneously go to study it and learn it; indirect interest does not mean that the individual is interested under the compulsive influence of some medium, but that the individual spontaneously goes to learn, mainly because he or she realizes that he or she needs this kind of knowledge very much, for example, when it concerns personal or collective interests. There are also some examples; for example, when an individual wants to be recognized by the people next to him, he will try to learn to express himself and slowly develop his interest.

4. Analysis of Results

4.1. Parameter Comparison Experiment and Analysis.

There are two important parameters in the model training process: learning rate and dropout value. If the learning rate is too large, the model will converge too fast and may exceed the optimal value, while if the learning rate is too small, the model will converge too slowly and affect the training efficiency and may even cause the model to fail to converge. Therefore, the learning rate and the dropout are added in this paper, and the model effect is compared with the learning rate of 0.01 and 0.001, respectively. Figure 3 shows the comparison of model recognition when two learning rates are used.

At the same time, dropout can be used to avoid the overfitting phenomenon in model training. Dropout is to let a neuron stop working temporarily with a certain probability P during the forward propagation process, which makes the model generalize more and does not rely too much on some local features and achieves the effect of regularization to some extent. PCA solves the problem of finding the optimal plane so that the projection of the original data on the plane is as dispersed as possible. Figure 4 shows the comparison of the model with the learning rate of 0.01 and the dropout set to 0.1 and 0.5, which shows that the difference between them is not large. Figure 4 shows that the loss value gradually decreases with the increase of the number of iterations of epoch, and the loss decreases to a smaller value and is more stable after 10 iterations. By comparing the above parameters, the model with a learning rate of 0.01 and a dropout value of 0.5 is finally selected.

In contrast to traditional forms of school organization, virtualization will be a common form of management and teaching organization in the schools and classrooms of the future. In the form, there are no fixed geographical space and time constraints, and members of the organization work together to achieve organizational goals through a high degree of self-regulation and a high degree of value orientation. In the teaching scenario, virtual mainly refers to the teaching activities based on the network environment, including non-face-to-face teaching implementation, teaching, and research activities, fictitious teaching and research experts, intelligent teaching, and research resources equipped. However, it needs to be emphasized that the virtual type

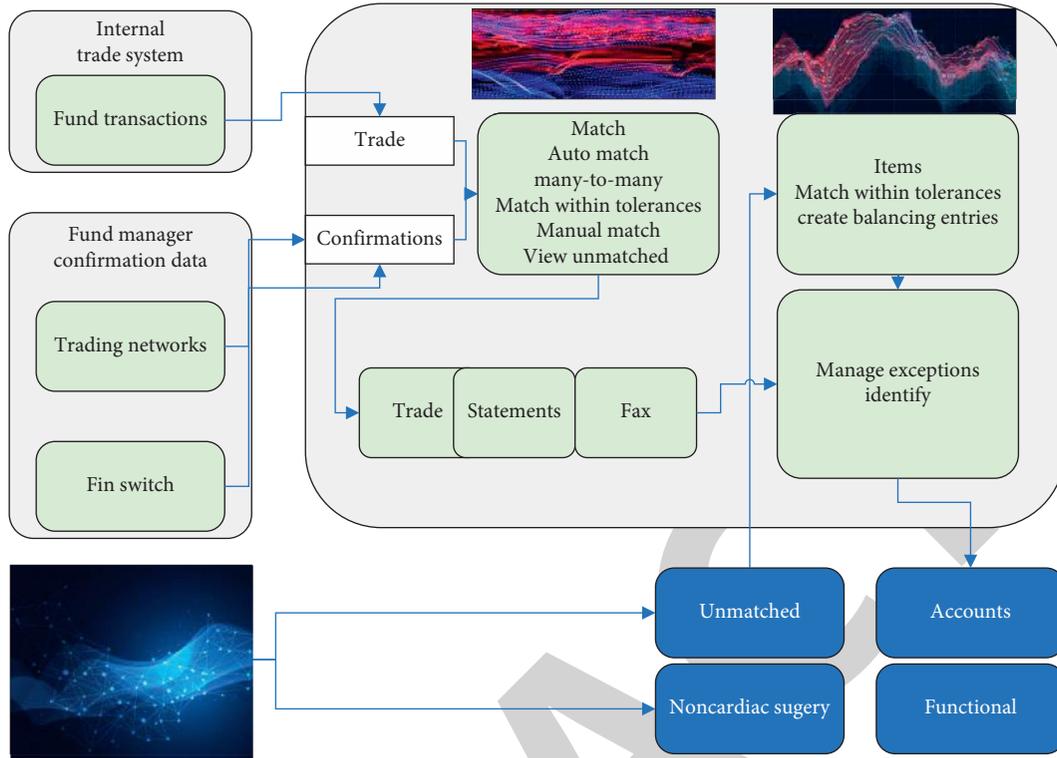


FIGURE 2: Visual analysis design.

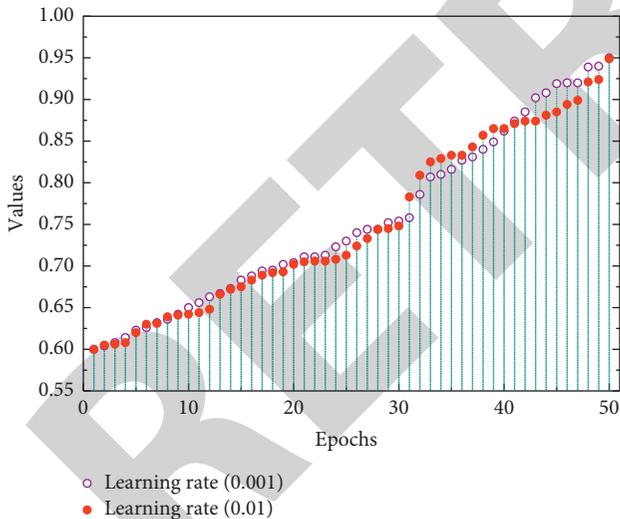


FIGURE 3: Model effect for learning rate of 0.001 and 0.01.

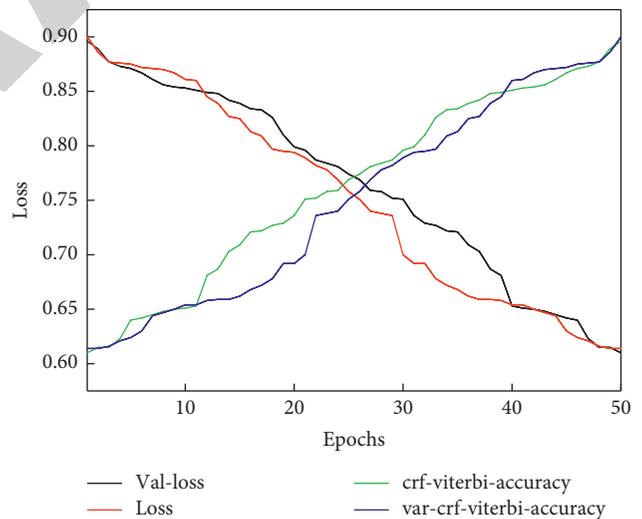


FIGURE 4: Effect of loss with the number of iterations.

must have a real aspect in its essence, as shown in Figure 5. The traditional forms of in-depth classroom and subject group research are still indispensable. The technological means of virtualization can span time and space, providing easier, faster, and even smarter tools for teaching and learning to be more effective. Overall, in the school of the future, close integration of virtual and real is the complete and scientific connotation.

From the experimental results, we can see that each model can achieve certain entity recognition results.

Among them, the accuracy and recall rate of the CNN + BiLSTM-CRF model improve 0.3% and 3.6%, respectively, based on the BiLSTM-CRF model, which shows that the Dict + CNN + BiLSTM-CRF model in this paper is better than the BiLSTM-CRF model. It indicates that the combination of CNN and BiLSTM can effectively use the local and global information of text to obtain more complete text features, which makes the model more accurate and can achieve better-named entity recognition results. And after adding the domain lexicon, the F1 value is

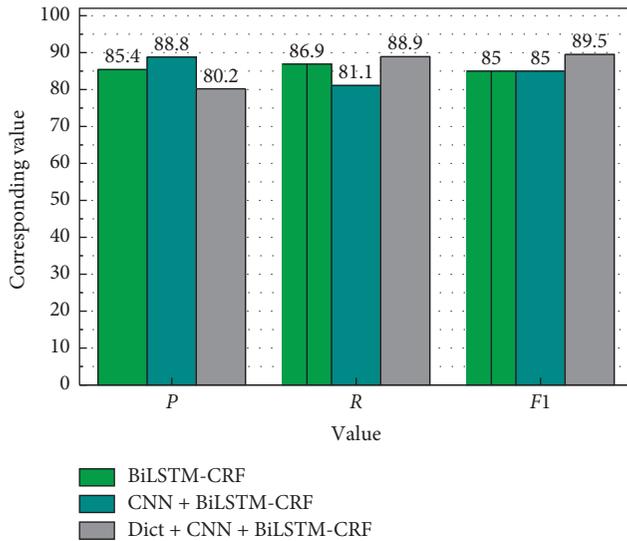


FIGURE 5: Comparison of experimental results between this model and other models.

improved by 2.6% on top of the CNN + BiLSTM-CRF model, which shows that the lexicon can help to further improve the model effect. In summary, the CNN + BiLSTM-CRF model proposed in this paper can achieve better results in named entity recognition work and can significantly improve the entity recognition effect to some extent after the introduction of the domain lexicon. This paper also uses the model to complete entity prediction for unlabeled text data and ensure the quality of entity data of the knowledge graph.

4.2. Visual Design Analysis. Inside the module, the visualization focuses on how to represent the module information. For modules with subnodes below, the detailed information of each subnode needs to be hidden when it is not clicked on, and the general information of all subnodes is displayed. For example, in Figure 6, the total score of students' overall quality should be the total score of each ability level, not the individual score of each level. This total score ranges from 0 to 100 and is divided into five bands; the higher the band is, the closer the student is to "teacher excellence." Regarding student's score, the highest, lowest, and average scores of all students are plotted in a bar graph to show the student's position in the student population. The overall evaluation screen has a time selection function that displays the results of the evaluation period selected by the user. The sub-modules that are expanded by clicking on the total evaluation are also based on the time information of the parent module and display the details of the evaluation results for the selected period.

The purpose of this is to provide a holistic picture of the student's overall situation. In addition to the total score, the trend of the student's score within each evaluation period and the total score for each subarea should be displayed. For the scoring trend, a line graph is used, with the user dynamically entering the starting time as an interactive excuse

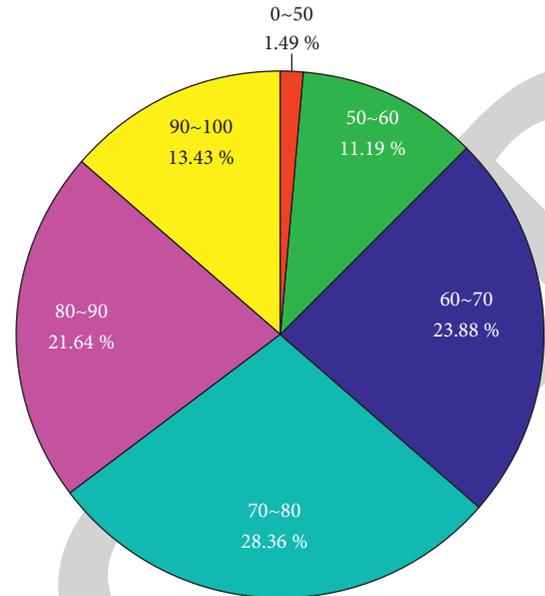


FIGURE 6: Comprehensive quality evaluation total score segment commentary table.

to return the time and score of each evaluation within this period, depicting the line graph with the time as the horizontal axis and the score as the vertical axis. The trend of the overall quality of the students can be seen in the line graph, whether it is a gradual improvement or a linear increase, which can be used as simple feedback of the effect of the reform of teaching measures. For the subspect scores, a spider web diagram is first used to draw the proportion of each subspect score to the total score. The purpose of this diagram is to show the students' strengths and weaknesses, as well as the degree of balance in their development. Although there is some overlap with the tree diagram structure that uses the area to represent the proportion to split the parent nodes, the spider web diagram is more suitable for expressing the ability composition, and the spider web diagram appears at the upper level and the split tree diagram appears at the lower level, which does not conflict with each other. Moreover, the split tree diagram can only show the distribution of evaluation results in the current evaluation cycle when clicked on, while the spider web diagram can be used in the upper layer to view the evaluation spider web diagram for each past evaluation cycle by adjusting the time selection, as shown in Figure 7.

In the leaf node module, the overview information is already available to the user and detailed information needs to be displayed. First, the submodule should still show the overall rating score for this module. Use a bar chart to draw the score for this student, the highest score, the lowest score, and the average score among students in this area. Use a line graph to derive the trend of the student's score under this module by user input time. Then, list all the details in the nondata form under this module such as participation, organization, and awards for club activities, competitions, and lectures.

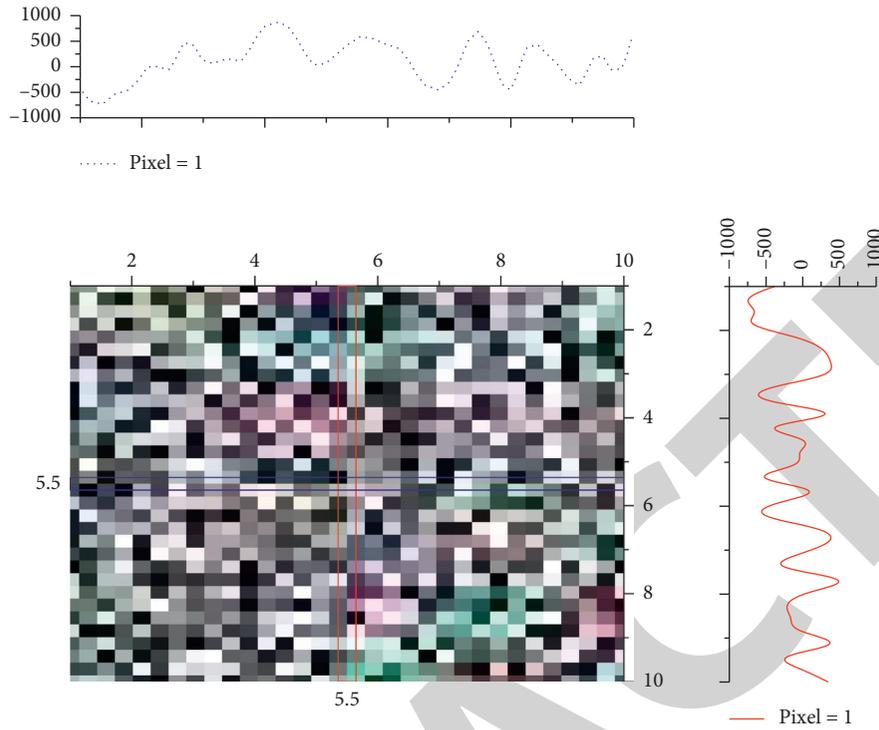


FIGURE 7: Evaluation of spider web diagram.

From the user’s learning data, it can be observed that the user had a general understanding of the course in advance and chose to learn it spontaneously during the learning process. Although the user did not choose to skip Technique 1, the subsequent posttest results also showed the user’s proficiency in Technique 1 and passed the test. Through Figure 8, we can observe that the system constantly adjusts the recommendation index according to the different choices and assessment results of the user, and at the same time, the learning power of the user in the learning process increases from 3500 to 5273, and the lower limit of the learning zone gradually stabilizes at 0.20, which means that all the lessons of this course are not very difficult for the user and are suitable for direct learning. In response to the shortcomings of the original static and subjective evaluation methods, we constructed a business English method based on big data, built multilevel evaluation indexes from six aspects, and established the correspondence between the indexes and campus life data to build a data-based process evaluation method and studied the changes of Internet education and education.

A recommendation method is proposed based on the Bayesian formula for calculating the efficiency of courseware, which will give different recommendation feedback to different users after they choose different courses. This chapter also proposes a way to calculate the learning comfort zone based on Bayes’ theorem to provide different users with learning content suitable for themselves. Finally, based on the proposed method and the basic learning model, a simple courseware learning system is designed for practical testing. The simulation test is conducted based on the reference to the actual test, and the final results show that the method

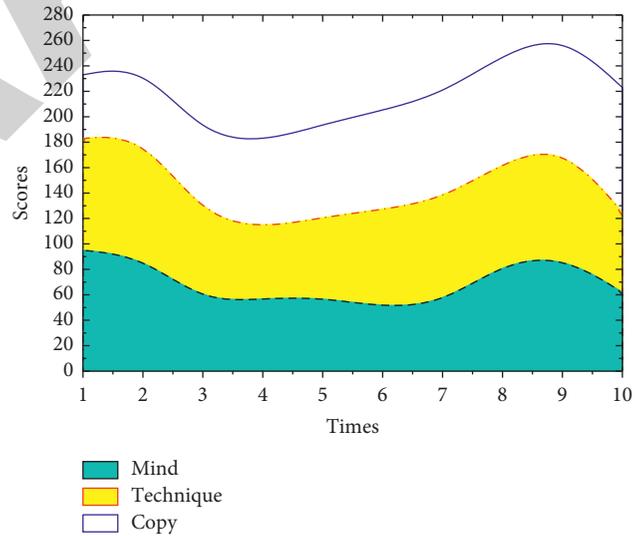


FIGURE 8: Score trend graph.

proposed in the thesis can not only be used as an effective recommendation method but also can provide users with more basis for learning growth and improve their learning efficiency.

5. Conclusion

Internet education covers the contents of two different fields represented by the Internet information technology and education, which belong to two different fields of human practice and value, and there are both correlations and

certain conflicts between them. The new technologies represented by the Internet, big data, and cloud computing should essentially emphasize the practicality of the means and pursue the objectivity, accuracy, and maximum efficacy of the operation in the education process. In response to the shortcomings of the original static and subjective evaluation methods, a business English method based on big data is constructed, multilevel evaluation indicators are built from six aspects, and the correspondence between the indicators and campus life data is established to build a data-based process evaluation method; to study the change of Internet education and education, the new information technology represented by the Internet and education must be closely integrated with both, starting from the technology. In the study of Internet education and educational change, it is necessary to study not only the promotion and profound impact of new technologies such as the Internet, big data, and cloud computing on education, but also the impact of new technologies on a series of educational concepts, ideas, and models. In the study of Internet education and educational change, we should study not only the promotion and profound influence of new technologies such as the Internet, big data, and cloud computing on education but also the impact and influence of new technologies on educational concepts and ideas and models.

Data Availability

All the information used to support the findings of this study is presented within this paper.

Conflicts of Interest

The authors declare that no conflicts of interest exist concerning this study.

References

- [1] P. Lade, R. Ghosh, and S. Srinivasan, "Manufacturing analytics and industrial Internet of things," *IEEE Intelligent Systems*, vol. 32, no. 3, pp. 74–79, 2017.
- [2] Z. Lv, X. Li, and H. Lv, "BIM big data storage in WebVRGIS," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 4, pp. 2566–2573, 2019.
- [3] F. J. Ai, X. Z. Yin, R. C. Hu, H. L. Ma, and W. Liu, "Research into the super-absorbent polymers on agricultural water," *Agricultural Water Management*, vol. 10, no. 1, pp. 1–10, 2021.
- [4] W. Liu, H. Ma, and A. Walsh, "Advance in photonic crystal solar cells," *Renewable and Sustainable Energy Reviews*, vol. 116, Article ID 109436, 2019.
- [5] X. Zhang, C. Zang, H. Ma, and Z. Wang, "Study on removing calcium carbonate plug from near wellbore by high-power ultrasonic treatment," *Ultrasonics Sonochemistry*, vol. 62, Article ID 104515, 2020.
- [6] R. H. Hariri, E. M. Fredericks, and K. M. Bowers, "Uncertainty in big data analytics: survey, opportunities, and challenges," *Journal of Big Data*, vol. 6, no. 1, pp. 1–16, 2019.
- [7] H. Ma, X. Zhang, F. Ju, and S.-B. Tsai, "A study on curing kinetics of nano-phase modified epoxy resin," *Scientific Reports*, vol. 8, no. 1, p. 3045, 2018.
- [8] H. Ma and S.-B. Tsai, "Design of research on performance of a new iridium coordination compound for the detection of Hg^{2+} ," *International Journal of Environmental Research and Public Health*, vol. 14, no. 10, p. 1232, 2017.
- [9] A. G. Polyakova, M. P. Loginov, and E. V. Strelnikov, "Managerial decision support algorithm based on network analysis and big data," *International Journal of Civil Engineering and Technology*, vol. 10, no. 2, pp. 291–300, 2019.
- [10] L. Cao, "Data science," *ACM Computing Surveys*, vol. 50, no. 3, pp. 1–42, 2017.
- [11] K. Dahdouh, A. Dakkak, L. Oughdir, and F. Messaoudi, "Big data for online learning systems," *Education and Information Technologies*, vol. 23, no. 6, pp. 2783–2800, 2018.
- [12] M. D. Lytras, V. Raghavan, and E. Damiani, "Big data and data analytics research," *International Journal on Semantic Web and Information Systems*, vol. 13, no. 1, pp. 1–10, 2017.
- [13] R. Godwin-Jones, "Scaling up and zooming in: big data and personalization in language learning," *Language Learning & Technology*, vol. 21, no. 1, pp. 4–15, 2017.
- [14] C. Yang, Q. Huang, Z. Li, K. Liu, and F. Hu, "Big data and cloud computing: innovation opportunities and challenges," *International Journal of Digital Earth*, vol. 10, no. 1, pp. 13–53, 2017.
- [15] J. Zetino and N. Mendoza, "Big data and its utility in social work: learning from the big data revolution in business and healthcare," *Social Work in Public Health*, vol. 34, no. 5, pp. 409–417, 2019.
- [16] M. K. Saggi and S. Jain, "A survey towards an integration of big data analytics to big insights for value-creation," *Information Processing & Management*, vol. 54, no. 5, pp. 758–790, 2018.
- [17] L. Mo, W. Sun, S. Jiang et al., "Removal of colloidal precipitation plugging with high-power ultrasound," *Ultrasonics Sonochemistry*, vol. 69, Article ID 105259, 2020.
- [18] D. Gao, Y. Liu, and Z. Guo, "A study on optimization of CBM water drainage by well-test deconvolution in the early development stage," *Water*, vol. 10, no. 7, 2018.
- [19] X. Jiang, A. Lin, and H. Ma, "Minimizing the thermal bridge through the columns in a refrigeration room," *Applied Thermal Engineering*, vol. 165, Article ID 114565, 2019.
- [20] A. Mohamed, M. K. Najafabadi, Y. B. Wah, E. A. K. Zaman, and R. Maskat, "The state of the art and taxonomy of big data analytics: view from new big data framework," *Artificial Intelligence Review*, vol. 53, no. 2, pp. 989–1037, 2020.
- [21] S. B. Tsai and H. Ma, "A research on preparation and application of the monolithic catalyst with interconnecting pore structure," *Scientific Reports*, vol. 8, no. 1, 2018.
- [22] T. R. Rao, P. Mitra, R. Bhatt, and A. Goswami, "The big data system, components, tools, and technologies: a survey," *Knowledge and Information Systems*, vol. 60, no. 3, pp. 1165–1245, 2019.
- [23] W. Liu, H. Shi, H. Ma, and S.-B. Tsai, "Improving renewable energy source in automotive applications," in *Green Production Strategies for Sustainability*, S. Tsai, B. Liu, and Y. Li, Eds., pp. 210–237, IGI Global, Harrisburg, PA, USA, 2018.