WILEY | Hindawi

Research Article

Design of Online Education Information Management System Based on Data Mining Algorithm

Wei Chen, Daowen Ren [iD], and Meiyan Pang

Guangzhou College of Technology and Business, Guangzhou, 510850 Guangdong, China

Correspondence should be addressed to Daowen Ren; wei@gzgs.edu.cn

Online education is the core field in China now, and many online education institutions have reached thousands or even thousands of teachers. A large amount of information has been gathered in the corresponding personnel management system modules and talent databases of the institution itself, such as personnel student status query management and performance evaluation analysis and management. How to reuse these data to transform existing data management into usable knowledge has become a problem that organizations should not underestimate. The purpose of this paper is to study the theoretical knowledge related to data mining algorithms, put forward the process and classification of data mining, and focus on the mining process and common methods of data classification rules. Experimental data shows that 31% of parents chose online education in 2019, compared with only 14% in 2017. Among them, the natural lack of offline education resources in third- and fourth-tier cities online education makes up for the lack of offline resources. The results show that third- and fourth-tier cities will become the driving force for the development of online education in the future.

1. Introduction

At present, education management has gradually transformed into a core component of each institution of higher learning. With the gradual deepening of education system reform, the traditional teaching management system has been unable to keep up with the pace of modern education. As the educational measures of China's colleges and universities have been gradually implemented, the steady growth of the source of students in colleges and universities has brought a greater burden to the teaching management system. How to make computer technology and network technology more efficient and convenient to serve education in the field of education has become a brand new topic faced by higher vocational colleges. Since the emergence of data mining technology, there are many reasons to promote the development and diffusion of data mining technology. The emergence of data warehouse makes data mining technology develop and spread by leaps and bounds. In computer technology, especially network technology and parallel processing system, the development of computer system with fast speed and strong computing power provides a good environment for the implementation of data mining [1]. Data mining technology finds the meaning behind the data in a large number of irregular and noisy data.

The innovation of the article is firstly, it analyzes the data characteristics of the system and then selects the data analysis method according to the characteristics. Specifically, it should be carried out according to the whole process of data mining, data pre-preparation, data postprocessing, data transformation, data mining, and model evaluation.

2. Related Work

Data mining is a predictive analytics, Shousha et al. used data mining analysis to build a decision tree by reducing error (REP) technique and then used the Auto-WEKA tool to select the best classifier from 39 algorithms to predict advanced fibrosis [2]. However, the structure of his algorithm is more complicated, which makes the result

inaccurate. An emerging research topic that Xu L. has been researching in recent years is called Privacy Preserving Data Mining (PPDM), which has received extensive attention. In Xu et al.'s research, it can be known that privacy issues related to data mining from a broader perspective and investigate various methods that can help protect sensitive information [3]. Kavakiotis et al.'s research demonstrated that remarkable advances in biotechnology and health sciences have led to the generation of massive amounts of data, such as high-throughput genetic data and clinical information generated from large electronic health records (EHRs) [4]. In his research, Wu and Peng showed that in order to deal with the dynamics of training samples and improve the prediction accuracy, a data mining method consisting of K-means clustering and Bagging neural network (NN) was proposed for short-term WPF [5]. The purpose of Marozzo et al.'s research is to demonstrate how cloud software technologies can be integrated to achieve an efficient environment for designing and executing scalable data analysis workflows [6]. In Rupesh et al.'s research, a security system named Internal Intrusion Detection and Protection System (IIDPS) was proposed [7]. Sun et al. proposed an empirical method to efficiently preprocess and filter raw wind data using the total active power output of wind farms and the corresponding wind speed values, but these studies are not very practical [8]. However, the above methods only exist in the theoretical part, and their practicality is not strong enough.

3. Methods of Data Mining Algorithms

3.1. Data Mining Algorithms. With the development of multimedia and network, providing network learning environment and distance network teaching has become the research characteristic of computer in education application, and the intelligent teaching system based on Web emerges as the times require [9]. Rough set theory has many advantages: it does not need to provide additional information; the expression space is further compressed, reducing the number of input messages; it is easy to understand and easy to grasp [10]. An information table that approximates a two-dimensional association table is the ultimate purpose of rough set management. Its structure diagram is shown in Figure 1:

The figure just shows that under the action of the data mining system, the entire framework and process of online education institutions are very clear. Because the existence of the network can provide learners with a very rich network of educational materials, students' learning has become a resource-based learning to a certain extent [11]. The biggest advantage of resource-based learning is that students' individualized and autonomous performance can be fully utilized.

3.2. Naive Bayesian Methods. Naive Bayes classification algorithm is a simple algorithm. Due to the simplicity of its ideological basis, it first points out in solving the problem of the division of text objects. The corresponding relationship between every two words that appear in the bag of words

must be able to be independent of each other and related to each other. Each dimension existing in the object feature vector can also be related to each other independently of each other [12]. But at least we know we can now try to generalize it further to some more dimensional cases: a formal concept for Naive Bayesian classification includes:

(1) Set it as the item to be classified, and each element a is each attribute attribute of the element x

(2) There are category collections $C = \{y_1, y_2 \cdots y_n\}$

(3) Calculate $P(y_1|x), P(y_2|x), \cdots, P(yn|x)$

(4) If $P(y_k|x) = \max \{p(y_1|x)p(y_2|x), \cdots, p(y_n|x)\}$, then $x \in yk$

(1) Statistically obtains the conditional probability estimates of each feature attribute under each category, which is

$$P(a_1|y_1), P(a_2|y_1), \cdots, P(a_z|y_1), \tag{1}$$

$$P(a_1|y_2), P(a_2|y_2), \cdots, P(a_z|y_2), \tag{2}$$

$$P(a_1|y_n), P(a_2|y_n), \cdots, P(a_z|y_n). \tag{3}$$

(2) We can have the following derivation according to the formula Bess theorem:

$$P(y_k|x) = \frac{P(x|y_k)P(y_k)}{P(x)}. \tag{4}$$

Because the denominator can be a constant for all classes as long as it maximizes the numerator and because each feature attribute also has mutual constraints and independent connection mechanisms, there are

$$P(x|y_k)P(y_k) = P(a_1|y_k)P(a_2|y_k) \cdots P(a_m|y_k) = P(y_k)\prod_{(m]}^{m} P(a_j|y_k). \tag{5}$$

Classification is a fundamental problem in the field of data analysis and machine learning. Text classification has been widely used in many aspects such as network information filtering, information retrieval, and information recommendation. Data-driven classifier learning has been a hotspot in recent years, and there are many methods, such as neural networks, decision trees, support vector machines, and Naive Bayes. Compared with other well-designed and more complex classification algorithms, the Naive Bayes classification algorithm is one of the classifiers with better learning efficiency and classification effect.

3.3. Principle of Decision Tree. Decision tree rules can be easily understood and directly transformed into classification tree rules, which is another very vivid and intuitive
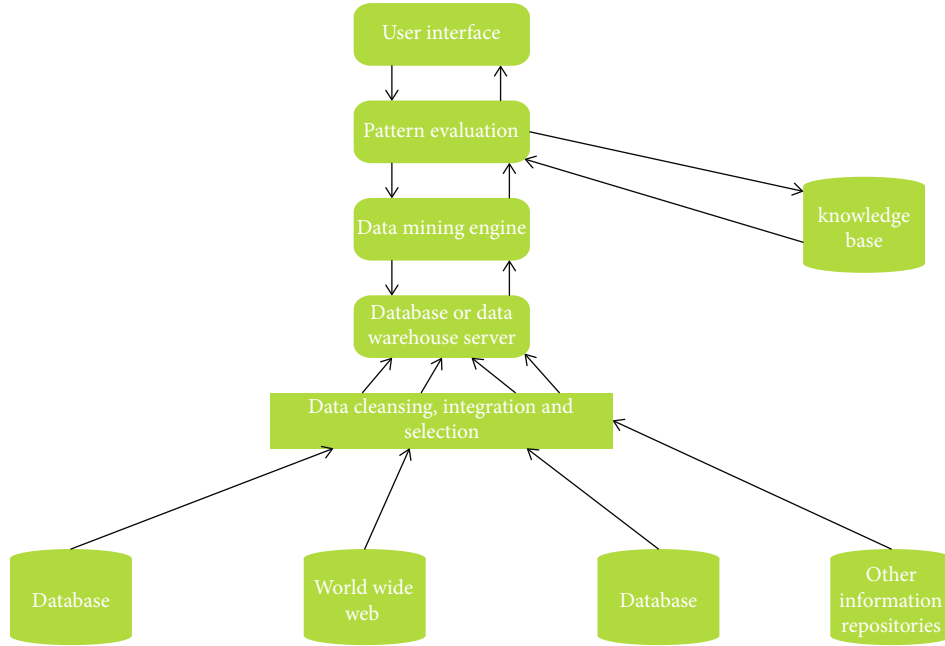
FIGURE 1: Data mining algorithm system structure diagram.

representation of the classification tree model. Classification tree (decision tree) is a very common classification method. It is a kind of supervised learning. The so-called supervised learning is given a bunch of samples; each sample has a set of attributes and a category, and these categories are determined in advance. Then a classifier is obtained by learning, which can give the correct classification of the newly appeared objects. Such machine learning is called supervised learning.

(1) Self-information volume. If the continuously sent signals are set to be $Y1, Y2, Y3 \cdots ..Yn$, the signals are sent until the Yi signal is accepted, and the uncertain signal is identified as $U(Yi)$, that is

$$U(Yi) = - \log_2 P(Yi). \qquad (6)$$

Among them, $P(Yi)$ represents the probability that the source sends out Yi.

(2) Information entropy. Another example is to measure the uncertainty of the signal source Y through the information entropy, namely:

$$H(X) = - \sum_{\lfloor m \rfloor}^{n} P(Yi) \log_2 P(Yi), \qquad (7)$$

where Y is the signal source and i is any possible number of symbols.

(3) Conditional entropy. Assuming that the signal sources X and Y are not independent of each other, the conditional entropy $H(X/Y)$ is used to measure

the overall uncertainty [13]. If the signal source corresponding to X is Xi, the signal source corresponding to Y is

$$H\left(\frac{X}{Y}\right) = - \sum_{|m|}^{n} \sum_{|n|}^{m} P(XiYj) \log_2 P\left(\frac{Xi}{Yj}\right). \qquad (8)$$

(4) Average mutual information. Correlation between signal sources U and V:

$$I(U, V) = H(U) - H\left(\frac{U}{V}\right). \qquad (9)$$

(5) According to the information theory, let the set S be the set containing the overall information of the whole training sample data, which contains the training instance data information set of n classes. If there are Si training instances in each class, the required amount of information I for classifying them can be expressed as

$$I(S_1, S_2 \cdots S_n) = - \sum_{[m]}^{n} P \log_2 (Pi). \qquad (10)$$

It can be clearly seen that the data samples we can get should be at least S or a data set containing more than N types of samples. In order to reduce the next task as much

as possible, we may need to select the feature with the highest gain coefficient in the information as a node for the next task and create a branch for feature division [14].

3.4. ID3 Algorithm. The ID3 algorithm was proposed in 1986. The algorithm is based on information theory and uses information entropy and information gain as measurement standards, so as to realize the induction and classification of data.

(1) Formal description

If it is known that there are the following n messages with the same maximum probability of the same information, then the high probability of each message with the same information is because x is $1/n$. The maximum amount of information that a message can pass out $-\log2(x) = \log2(n)$. If there is a given probability distribution, the total amount of information transmitted by the distribution is called the total entropy of the distribution x, that is:

$$U(x) = -(x_1 * \log_2(x_1) + x_2 * \log_2(x_2) + \cdots x_n * \log_2(x_n)). \tag{11}$$

(2) If we assume that we first divide the element class T into the following three sets $T\,1$, $T\,2$, $T\,n$ according to the value of the noncategory attribute V, respectively, then how can we determine the information content of any element class in the set T? We can first obtain [15] by determining the arithmetic mean of the arithmetic weights in the set $T\,i$, that is, the arithmetic mean of the arithmetic weights in INFO($T\,i$) is

$$\text{INFO}(V, T) = \sum_{i=1}^{n} \frac{|T_i|}{|T|} * \text{INFO}(T_i). \tag{12}$$

(3) Define the gain Gain (X, T) as

$$\text{Gain}(A, B) = \text{Info}(B) - \text{Info}(A, B). \tag{13}$$

Given that a set of noncategory attributes $C\,1$, $C\,2$, and $C\,N$ is constructed, the category attribute C and the training set T of records are constructed, and the specific calculation process of constructing a decision tree by the ID3 algorithm.

3.5. Feature Selection Strategy Based on Data Mining. In addition to the strategy method of feature node selection based on information gain, it also uses the strategy method that can automatically select the optimal split node according to the distribution of data categories of nodes, which is called Gini Index [16]. Assuming that the training set T contains more than n training samples, these training samples will be assigned to the Xth class, respectively. Among them,

the Ith class will be calculated as the proportion of time that occurs in the entire T set, and then the Gini Index of T is defined as

$$\text{Gini}(X) = 1 - \sum_{i=1}^{x} p_i^2. \tag{14}$$

Gini Index of this division is

$$\text{Gini}(B) = \sum_{i=1}^{m} \frac{n_i}{n} \text{Gini}(T_i). \tag{15}$$

3.6. BP Neural Network Algorithm. Given a node j in the output plane layer or hidden plane layer, the net input point I_j on node j is

$$I_j = \sum_i w_{ij} O_i + \theta_j. \tag{16}$$

Using the logistic function, given the net input of node j, then the output of node j is

$$O_j = \frac{1}{1 + e^{-I_j}}. \tag{17}$$

This formula is the extrusion formula of the BP network neural algorithm, which can map a larger input value range to a smaller range [17].

The error is propagated backwards by continuously updating the weights and biases representing the prediction error of the network [18]. For the output layer node v, the calculation formula of error Err_j is as follows:

$$\text{Err}_v = O_v(1 - O_v)(T_v - O_v), \tag{18}$$

where O_v is the actual output of node v and T_v is the known target value of node v based on a given training sample. And $O_v(1 - O_v)$ is actually the derivative of the logistic function.

By weighting the error of the nodes connected to v in the next layer, the error of the hidden layer node v can be quickly calculated. The error of the hidden layer node v is

$$\text{Err}_v = O_v(1 - O_v) \sum_k \text{Err}_k w_{kv}. \tag{19}$$

Among them, W_{kv} is the connection weight from node k to node v in the next higher layer, and Err_k is the error of node k.

The propagated error is reflected by updating the weights and biases. The formula for weight update is as follows, where Δw_{iv} is the change amount of w_{iv} weight.

$$\Delta W_{iv} = (1)\text{Err}_v O_i, \tag{20}$$

$$W_{iv} = W_{iv} + \Delta W_{iv}. \tag{21}$$

The number of variables l is a learning rate, usually used to refer to a set of constant values in the range between a

constant 0.0 and a constant 1.0 [19]. Back-propagation training uses stepped descent to continually find the most appropriate set of weights, which can be used to adjust the training data to minimize the squared error between the predicted neural network model values and the known true training sample values [20]. The formula can be updated by the following equation, where $\Delta\theta_v$ is the change in offset θ_V

$$\Delta\theta_v = (l)\text{Err}_v, \tag{22}$$

$$\theta_v = \theta_v + \Delta\theta_v. \tag{23}$$

The above method is an instance update method. The advantage of this BP neural network algorithm is that it can update the bias and weight after processing the training samples. This method is used more in practice, because instant update can often get more accurate results.

Based on the characteristics of students' online learning behavior data, this paper selects the ID algorithm of decision tree classification method for decision learning and rule extraction. Through the study of ID algorithm, it is found that if the correlation between attributes and the degree of association between attribute conditions expressed in the training set are considered and quantified for the reduction of rules, the computational speed and efficiency can be effectively improved while maintaining the classification accuracy. Based on this idea, an ID decision tree rule generation algorithm based on attribute correlation is proposed in this paper. The experimental data show that the rule generation and reduction of the full decision tree using ID algorithm is better than the original ID decision tree rule generation algorithm when constructing the online "learning behavior-effectiveness" model. The rule generation algorithm is more efficient than the original ID decision tree rule generation algorithm, especially when the training sample set size is further expanded.

4. Experiment Design of Online Education Platform

4.1. Design of Online Education Information Management System.
The system design should generally ensure that the B/S bus structure is used as much as possible. The three-tier structure in the B/S system architecture refers to the three-tier architecture with three basic subfunctions, such as database system, application server, and client browser, as shown in Figure 2:

The first layer is the client, which is an interface between the user terminal and the entire system. The client application is a web page. On the homepage of this website, users will be able to quickly and easily enter the website information you want in the information from template provided by this website anytime, anywhere. After the submission is successful, the user information is submitted to the server background, and the background automatically processes and automatically returns the user information to the server to the user. The background server here is the second-tier web server [21].

The second layer is a web server. The Web server will be used to quickly respond to user data input and requests received from the client browser and possibly edit the generated dynamic database into results and return the final data to the client browser. This system uses Java because Java is a relatively mature and reliable network development and design language with good compatibility. Some good system guarantees are provided to further improve the code and execution efficiency, maintenance-free performance, and system scalability of the researched and developed system. If these user data requests also include the browser's access to the data, then the web server will also be responsible for completing the data editing task together with the three-level dynamic database server. The third layer is the database server. The database server is mainly used to manage the database in a unified manner, and is responsible for coordinating and processing various database concurrent access service requests sent by different types of web server users.

The third layer is the database server. The database server is mainly used to manage the database in a unified manner and is responsible for coordinating and processing various database concurrent access service requests sent by different types of web server users.

4.2. Design of the Overall Function of the System.
Based on the review of the overall function analysis, requirements, and research work of the system, this chapter mainly focuses on the design and expounds the basic functions that constitute the application system. It is divided into several modules to describe the software operation and program, system function realization and design specification requirements, and other information. Each functional module runs stably under the control of the main control module. The system adopts the B/S mode and divides the system into three layers. The typical process of interaction between the layers of the system is shown in Figure 3:

The request sent by the system user on the page is judged by the controller, and the corresponding data information is retrieved from the background database.

4.3. System Function Design.
The system use case diagram can describe the operation functions of the system users well, but it cannot accurately represent the sequence relationship between the system functions and the sequence of operations. Therefore, the system module is redesigned with a flow chart here. The design process is as follows in Figure 4:

This system is to develop an online network remote training and resource distribution management system based on pure Java technology. Starting from an initial concept of system design, the purpose is only to develop a resource management platform system that can be used to build a platform for learning various online courses. At the same time, speaking from the principle of system technology, Java is used because Java is a relatively mature and reliable network development and design language with good compatibility. Some good system guarantees are provided to further improve the code and execution efficiency, maintenance-free performance, and system scalability of the researched and developed system. Several main
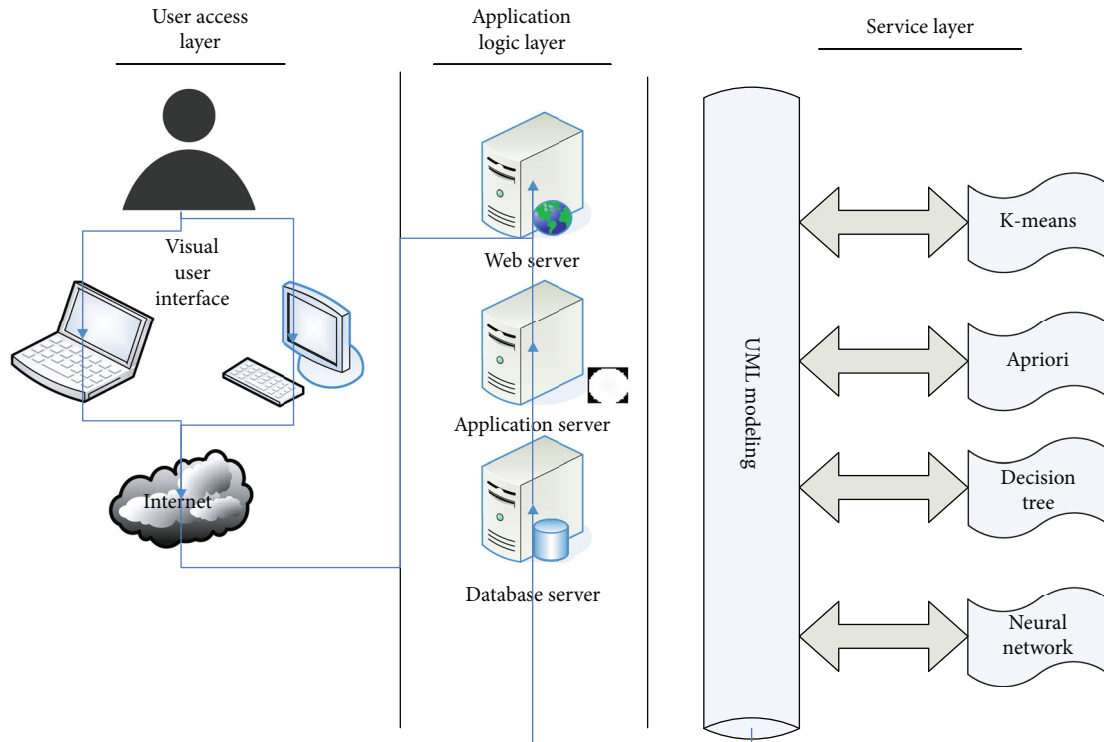
FIGURE 2: Online education information management system.

application functions provided by the system also include communication, student information management, assessment information management, work card management, and semester plan management. Because of the rapid progress in the construction of the network information environment, the network information technology education in primary and secondary schools in China has gradually become the target of everyone's criticism. Therefore, the school also needs us to follow the technological progress in various information system education and technology management education related to campus network technology education. It should be ready to adjust and update the thinking of teaching management methods and improve the concept of classroom teaching guidance in time according to the actual development and changes of current social information technology.

4.4. Detailed Design of the Database. From the design of the front-end web page to the design of the database running in the background, it all belongs to the development and design of the educational administration system [22]. One of the database designs is to use Microsoft SQL Server 2000 to establish a database. According to the functions that the website can achieve, a database named "ZhangQin" is established. In the database, a table is generally established for each module to store the information of each module. The basis of data classification research and analysis is data, and the types of data can be divided into continuous variables and categorical variables. In other words, the information of the same content, the same nature, and the

information that requires unified management are gathered together, and the dissimilar information and the information that need to be managed separately are distinguished. The relationships between the various sets are then determined to form an organized classification system. Here it will only introduced some of the most basic and commonly used table information in schools, such as administrators, teachers, students, courses, grades, and other information. In each user table here, the data type corresponding to each field that has been used has been redefined, and there is another user table specially reset for further convenience and control of user authority. The following is divided according to the system design goals and functional modules, as shown in Figures 5 and 6:

The main function provided by the learning area is mainly for users who are not using the system to learn. The main business includes user personal information login management system function and user account registration login information management query system module, real-time browsing and query management module of course content, online query of course resources, learning effect query management system module, personal online communication, interactive question and answer, learning record query management query system module, personal forum message and comment interaction record viewing management system function module, and famous teacher course resource screening system query management module and famous teacher article search function module.

The system includes user account and personal login and system login management functions, and the modules include
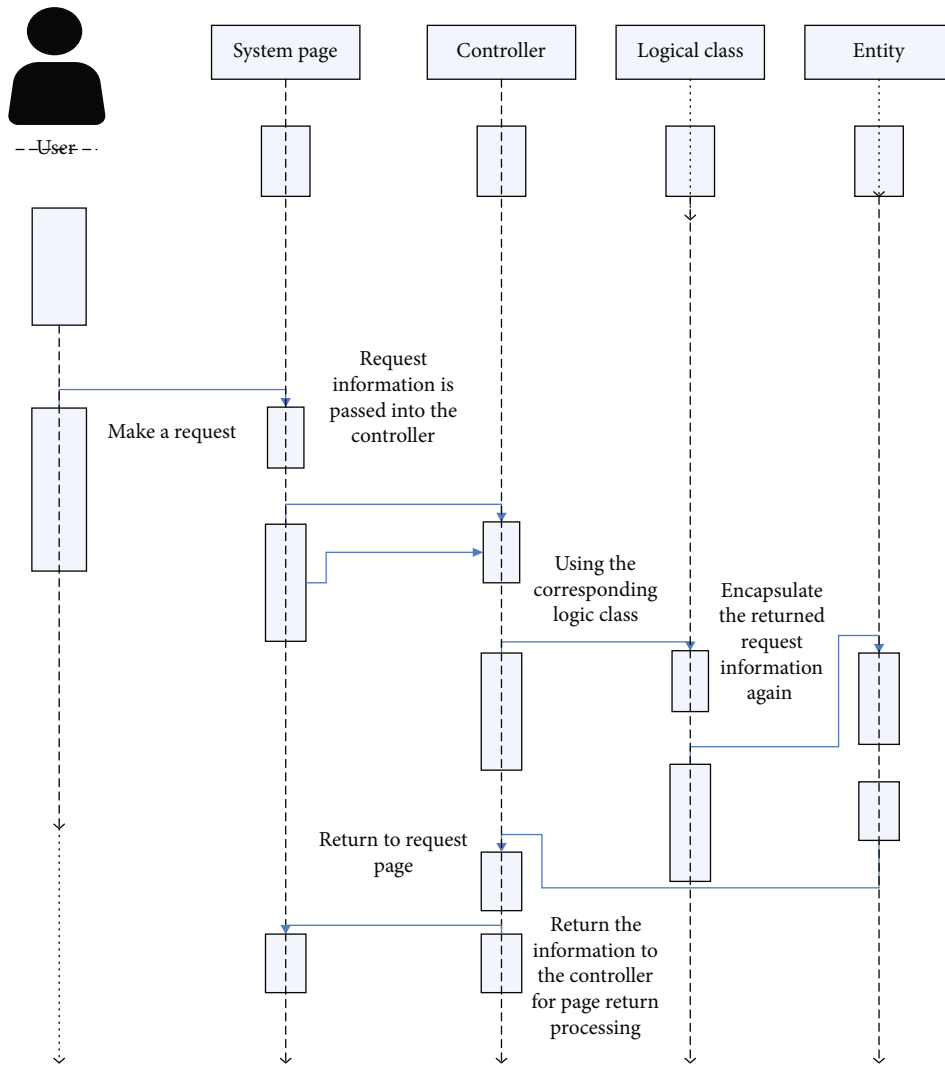
FIGURE 3: The overall interaction process of the system.

enterprise personal account registration functions and management functions. At the same time, it can provide individual account registration and management information query for corporate customers, in which account-related service information includes two main functions: query registration and management functions. These functions meet the needs of enterprise users' personal real-name registration and login, system query, customer modification, and record system. Browsing and management of course information, this module can also try to use various sorting or browsing methods to ensure that the target user is provided with a timely presentation of the course list. The user can simultaneously and directly browse and check that all the course lists meet the criteria for subject classification by subject teachers of the course or meet the criteria for teacher classification according to the teachers of the course.

A database is a "collection" of "data" organized according to certain rules and methods for a certain purpose. The database can be intuitively understood as a warehouse for storing data, but this warehouse is on the large-capacity memory of the computer, and the data must be stored in a certain format, because it not only needs to be stored, but also easy to find. The database contains as follows: (1) data sharing includes that all users can access the data in the database at the same time, and users can use the database in various ways through the interface and provide data sharing. (2) Compared with the file system, because the database realizes data sharing, it avoids users to create application files individually. A large amount of duplicate data is reduced, data redundancy is reduced, and data consistency is maintained. (3) The independence of data includes logical independence (the logical structure of the database and the application program in the database are independent of each other) and physical independence (changes in the physical structure of the data do not affect the logical structure of the data). (4) The data is in a scattered state, and there is no relationship between the files of different users or the same user in different processing. The database can be used to centrally
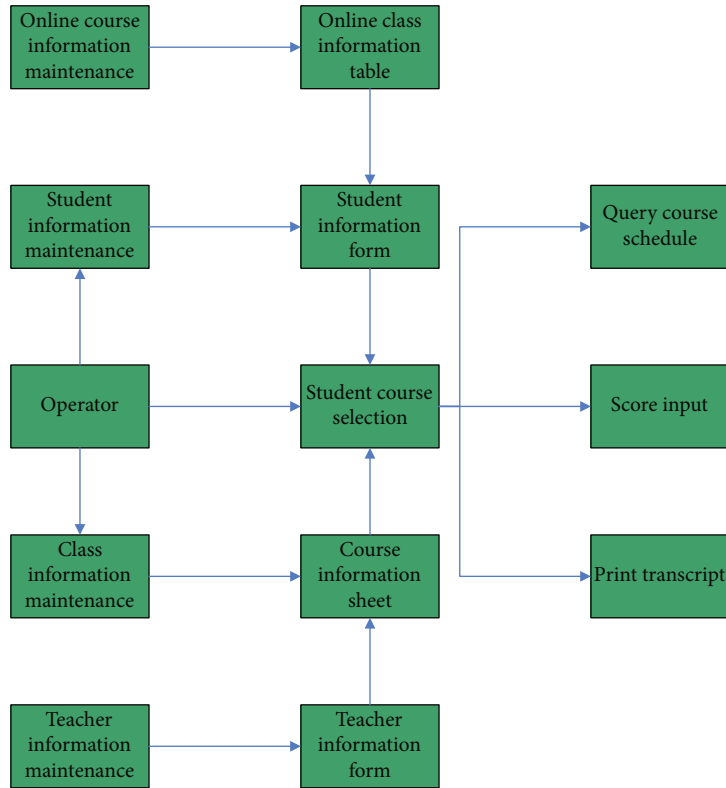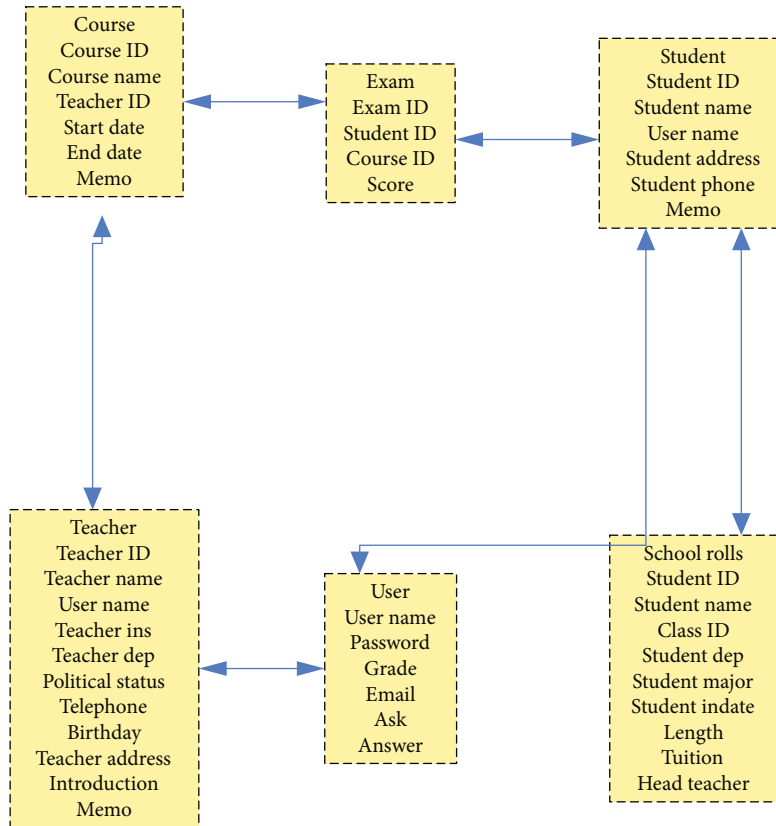
FIGURE 4: Teaching plan structure diagram.

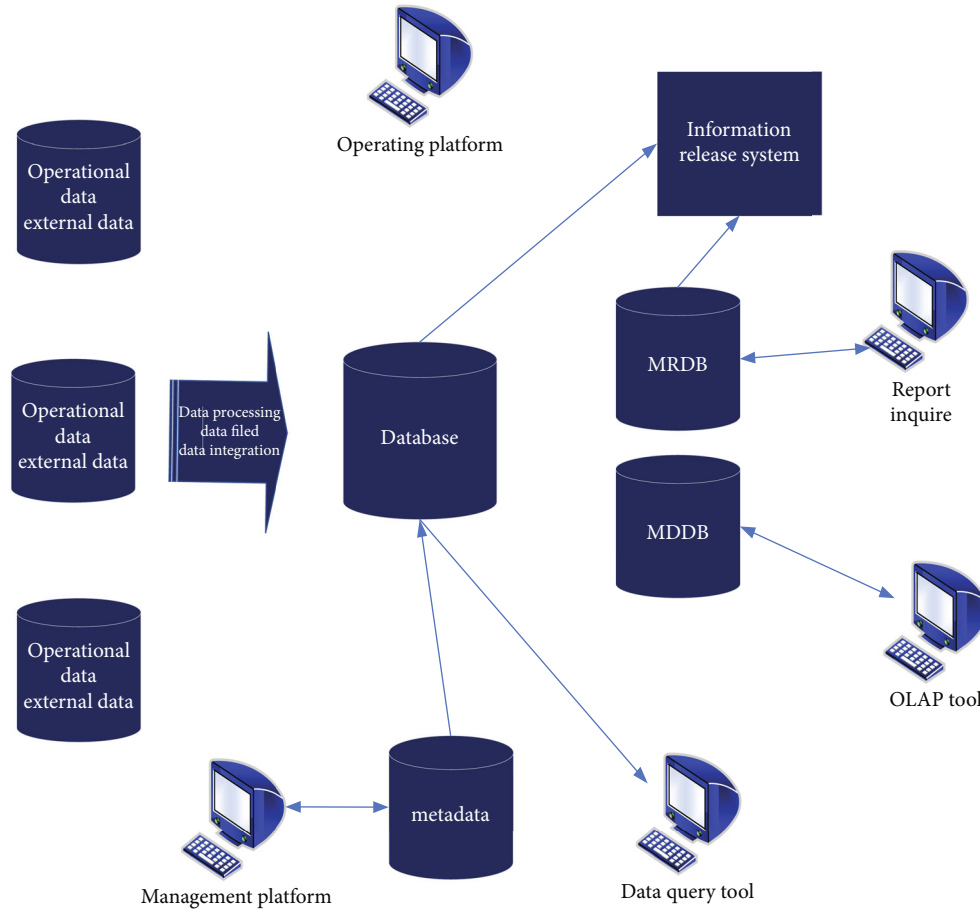FIGURE 5: System function structure diagram.

FIGURE 6: Data warehouse structure diagram.

control and manage the data, and the organization of various data and the connection between the data can be represented by the data model.

The online learning material management function module of course video mainly can realize a series of management functions such as providing online video course learning, real-time query management of learning material progress, and online download and viewing of learning materials related video materials. Users usually need to habitually use video course browsing or management software to conduct distance online learning when they have just officially entered the stage of video course learning and management. The management function module of the personal question and answer area mainly includes multiple functions such as question navigation in the personal question and answer area, searching for test questions, answering questions, viewing the list of all the questions I have asked, and viewing the list of all the answers to the questions that need my answer.

The data table is set up as follows:

(1) Administrator table

This table is used to store administrator and website initialization information, the structure is shown in Table 1:

(2) Teacher table

This table is used to store various basic information of teachers, and the structure is shown in Table 2:

(3) Student information sheet

This table is used to store various basic information of students, and the structure is shown in Table 3:

(4) System user table

This table simply records the information of system users, as shown in Table 4:

(5) Course information sheet

This table user stores course information, and the structure is shown in Table 5:

4.5. System Performance Test. Whether the system can really achieve various performance indicators close to or beyond the user's requirements is critical. And it takes this as a premise to find the system efficiency bottleneck that solves the possible problems in the system operation and finally

Table 1: Administrator table structure.

Listing	Data type	Length	Allow to be empty
admin	nvarchar	20	Yes
schoolName	nvarchar	20	Yes
bulletin	nvarchar	50	Yes
DAY	smalldatetime	4	Yes
pagesize	int	4	Yes
loginsToday	int	4	Yes
logins	int	4	Yes

Table 2: Teacher table structure.

Listing	Data type	Length	Allow to be empty
teacherID	int	4	No
teacherName	nvarchar	20	No
userName	nvarchar	20	No
teacherIns	nvarchar	50	Yes
teacherDep	nvarchar	50	Yes
politicalstatus	nvarchar	10	Yes
telephone	nvarchar	11	Yes
birthday	nvarchar	10	Yes
teacheraddress	nvarchar	50	Yes
introduction	ntext	16	Yes
Memo	nvarchar	200	Yes

Table 3: Student table structure.

Listing	Data type	Length	Allow to be empty
studentID	int	4	No
studentName	nvarchar	20	No
userName	nvarchar	20	No
studentAddress	nvarchar	50	Yes
studentPhone	nvarchar	11	Yes
Memo	nvarchar	200	Yes

Table 4: User information table.

Listing	Data type	Length	Allow to be empty
userName	nvarchar	20	No
password	nvarchar	20	No
grade	int	1	No
email	nvarchar	30	No
ask	nvarchar	50	No
answer	nvarchar	50	No

Table 5: Course information sheet.

Listing	Data type	Length	Allow to be empty
courseID	nvarchar	10	No
courseName	nvarchar	50	No
tearcherID	int	4	No
startDate	nvarchar	10	Yes
endDate	nvarchar	10	Yes
Memo	nvarchar	200	Yes

announcement by adding; finally, the user clicks to exit the system. A total of 150 registered and logged-in users need to be logged in. Initially, there can only be 5 registered and logged-in users, and then it will increase to 10 registered and registered users every two seconds. The general situation of the login results after the test is shown in Figure 7:

The statistical chart describes the test content and test results that will reflect the system operation in the test phase. It can be seen from the figure that the response time of the transaction under load reaches the highest peak average of 0.055 when the number of Vusers reaches 50. And it lists typical test scenarios for online training to confirm the system level and data mining algorithm development. The development of the system must first ensure economic feasibility and technical reliability. The use of computers, the design of an increasingly powerful network communication system, and the design of a perfect database have made the system gradually improved.

4.6. Development and Future Forecast of Online Education. With the rapid increase in the number of users of online video education in China in recent years, the overall scale will show a linear upward trend. As of December 2021, the total number of users of online video education websites in China has reached about 341.71 million, an increase of about 204.07 million compared with the same period in 2016; the website usage rate has reached about 34.6%, an increase of about 15.8 percentage points compared to the same period in 2016, as shown in Figure 8.

The year-on-year growth of online education is inseparable from the rapid development of the live broadcast market. At present, online education has become a new economic growth point, and the "live broadcast + education" model has begun to take shape. The government and local governments are also actively implementing policies to promote the healthy development of the "life + education" format. For example, the General Office of the State Council has previously issued an action plan document "Internet + Basic Education Promotion Work Plan" that has been implemented for many years. In the planning opinions, it is clearly stated that until 2022, the completion rate of the basic engineering education reform plan for famous teachers is basically to achieve 100%. It opened an online live room to explore the classroom form of online and offline integration. In the future, cities above the third and fourth tiers will further become a new powerful driving force for the healthy

achieves the goal of optimizing the overall operation efficiency of the system. It can be seen that the performance test of the system is very necessary [23]. For test case, the user logs in to the system first, uploads, and views a file; then the user viewed news, viewed a resource, and viewed an
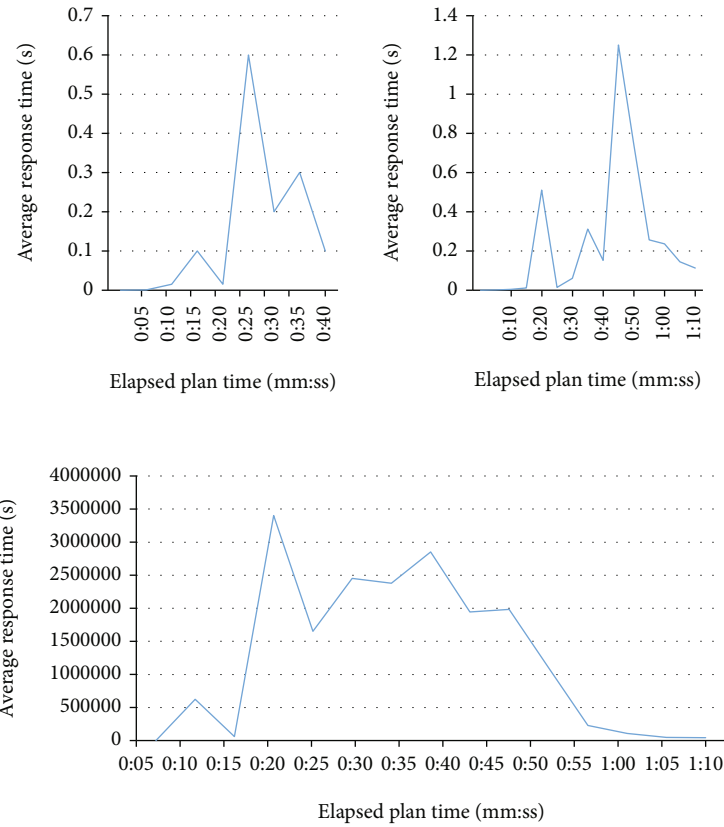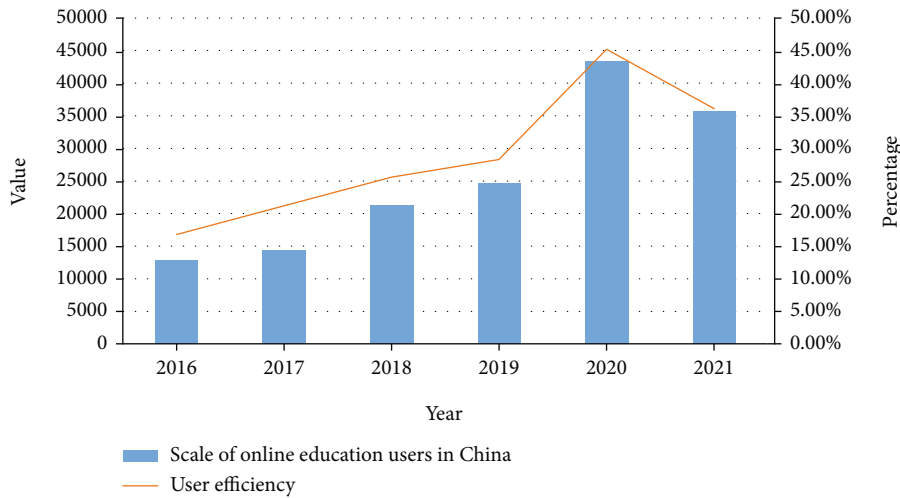
Figure 7: Throughput.

Figure 8: Changes in the scale and utilization rate of online education users in China from 2016 to 2021.

development of online training and education. With the rapid development, use, and further popularization of the new generation of Internet technology, the application penetration rate of online network education products will also be gradually increased. The survey shows that 31% of parents prefer online education in 2019.

The first- and second-tier markets of the online education industry are gradually maturing, while the third- and fourth-tier cities are still in the development stage and have

great potential for development. With the improvement of the economic level and the strengthening of education awareness, residents of third- and fourth-tier cities and townships will pay more attention to their spiritual needs. Online education can further exert its efforts to sink the market, deliver high-quality course resources to third- and fourth-tier cities, realize the sharing of high-quality educational resources, and meet the educational needs of residents in third- and fourth-tier cities.

5. Discussion

The system is essentially a network training resource management system developed based on data mining algorithms. The initial technical purpose of the development system should be to provide a management development system service that can be used to manage various online education application platform resources for development. In terms of current technical application, several main business functions of the development system service mainly include newsletter, student information management, evaluation information, management, work card management, and semester plan management [24]. With the increasing speed and development of China's information-based country construction, online distance education platforms are undoubtedly at the top of the field of social public education platforms in the future. Therefore, the education management of the online distance education network must also pay attention to keep pace with the times and adapt to the new needs of the rapid development of the educational society at any time. The background and significance of the research topic selection are guided by the ideological requirements of updating education management development and the concept of innovation and development of running a school. This paper summarized the current development of the online education platform, designed the organizational structure of this paper, and analyzed the related technologies involved in the system design and implementation in detail. It designs a layered system architecture according to the analysis of system requirements and the technical characteristics used and explains the framework technology used in the construction of each layer of the system in combination with the system architecture diagram. The design of the data model of the system includes the design of the system core class diagram, the design of the system entity object entity relationship diagram, and the application structure design of the database specific table.

6. Conclusions

This paper explores the shortcomings of traditional manual management training methods. On the basis of previous experience, it further summarizes the network and computer tool management network teaching platform, which replaces the modern management network training. Education is an important part of a country's development, and good education can improve the overall quality of the people and enhance the country's competitiveness. Using information technology to build an online education platform can break the time and space limitations of traditional offline education, allowing people to realize their desire to learn knowledge anytime and anywhere. The campus online education platform is mainly used to help students learn vocational courses in their spare time and can only learn in the classroom in the past. Students can share their questions and experiences about vocational courses in the question and answer area of the system. Through this system, teachers can broadcast lectures to students in the form of video or text. At the same time, the system also provides article pub-

lishing, making it easier for teachers and administrators to help students learn better.

Data Availability

No data were used to support this study.

Conflicts of Interest

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

Acknowledgments

References

[1] Q. Wang and P. Lu, "Research on application of artificial intelligence in computer network technology," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 33, no. 5, p. 1959015, 2019.

[2] H. I. Shousha, A. H. Awad, D. A. Omran, M. M. Elnegouly, and M. Mabrouk, "Data mining and machine learning algorithms using IL28B genotype and biochemical markers best predicted advanced liver fibrosis in chronic hepatitis C," *Japanese Journal of Infectious Diseases*, vol. 71, no. 1, pp. 51–57, 2018.

[3] L. Xu, C. Jiang, J. Wang, J. Yuan, and Y. Ren, "Information security in big data: privacy and data mining," *Ieee Access*, vol. 2, pp. 1149–1176, 2014.

[4] I. Kavakiotis, O. Tsave, A. Salifoglou, N. Maglaveras, I. Vlahavas, and I. Chouvarda, "Machine learning and data mining methods in diabetes research," *Computational and Structural Biotechnology Journal*, vol. 15, no. 5, pp. 104–116, 2017.

[5] W. Wu and M. Peng, "A data mining approach combining K-means clustering with bagging neural network for short-term wind power forecasting," *IEEE Internet of Things Journal*, vol. 4, no. 4, pp. 979–986, 2017.

[6] F. Marozzo, D. Talia, and P. Trunfio, "A workflow management system for scalable data mining on clouds," *IEEE Transactions on Services Computing*, vol. 11, no. 3, pp. 480–492, 2018.

[7] W. Rupesh, J. Sagar, and S. Rahul, "An internal intrusion detection and protection system by using data mining and forensic techniques," *IEEE Systems Journal*, vol. 11, no. 2, pp. 1–12, 2017.

[8] Q. Sun, L. Shi, Y. Ni, D. Si, and J. Zhu, "An enhanced cascading failure model integrating data mining technique," *Protection&Control of Modern Power Systems*, vol. 2, no. 1, pp. 33-34, 2017.

[9] S. Deng, Y. Du, and L. Qi, "A web service composition approach based on planning graph and propositional logic," *Journal of Organizational and End User Computing (JOEUC)*, vol. 31, no. 3, pp. 1–16, 2019.

[10] J. Zhang, S. O. Williams, and H. Wang, "Intelligent computing system based on pattern recognition and data mining algorithms," *Sustainable Computing*, pp. 192–202, 2018.

[11] Q. Wang, G. Z. Yao, G. M. Pan, J. Y. Huang, Y. P. An, and X. Zou, "Analysis of on medication rules for Qi-deficiency and blood-stasis syndrome of chronic heart failure based on data mining technology," *China journal of Chinese materia medica*, vol. 42, no. 1, pp. 182–186, 2017.

[12] E. B. Costa, B. Fonseca, M. A. Santana, F. F. de Araújo, and J. Rego, "Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses," *Computers in Human Behavior*, vol. 73, pp. 247–256, 2017.

[13] L. Zheng, W. Hu, and Y. Min, "Raw wind data preprocessing: a data-mining approach," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 1, pp. 11–19, 2014.

[14] L. I. U. XinYu, Q. Zhang, C. Zhang, Y. U. A. N. FangLin, and J. I. A. O. ShouTao, "Global major events in miocene and its significance: revelation from data mining," *Chinese science bulletin*, vol. 62, no. 15, pp. 1645–1654, 2017.

[15] W. Xia, Y. Liu, and D. Chen, "Design of public bicycle scheduling model based on data mining algorithm," *International journal of reasoning-based intelligent systems*, vol. 10, no. 1, pp. 68–76, 2018.

[16] M. Bakon, I. Oliveira, D. Perissin, J. J. Sousa, and J. Papco, "A data mining approach for multivariate outlier detection in postprocessing of multitemporal InSAR results," *IEEE Journal of Selected Topics in Applied Earth Observations&Remote Sensing*, vol. 10, no. 6, pp. 2791–2798, 2017.

[17] D. Caballero, A. Caro, M. del Mar Avila, P. G. Rodriguez, T. Antequera, and T. P. Palacios, "New fractal features and data mining to determine food quality based on MRI," *IEEE Latin America Transactions*, vol. 15, no. 9, pp. 1777–1784, 2017.

[18] H. H. Yang, M. L. Huang, C. M. Lai, and J. R. Jin, "An approach combining data mining and control charts-based model for fault detection in wind turbines," *Renewable energy*, vol. 115, pp. 808–816, 2018.

[19] J. R. Wang, W. F. Chen, Q. Zhang et al., "MORB data mining: reflection of basalt discrimination diagram," *Geotectonica et Metallogenia*, vol. 41, no. 2, pp. 420–431, 2017.

[20] C. J. Lin, C. Wu, and W. A. Chaovalitwongse, "Integrating human behavior modeling and data mining techniques to predict human errors in numerical typing," *IEEE Transactions on Human-Machine Systems*, vol. 45, no. 1, pp. 39–50, 2015.

[21] J. Zhang, S. O. Williams, and H. Wang, "Intelligent computing system based on pattern recognition and data mining algorithms," *Sustainable Computing: Informatics and Systems*, vol. 20, pp. 192–202, 2018.

[22] B. K. Blaylock, J. D. Horel, and S. T. Liston, "Cloud archiving and data mining of High-Resolution Rapid Refresh forecast model output," *Computers & Geosciences*, vol. 109, pp. 43–50, 2017.

[23] P. Vazan, D. Janikova, P. Tanuska, M. Kebisek, and Z. Cervenanska, "Using data mining methods for manufacturing process control," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 6178–6183, 2017.

[24] Y. Shi, D. Fu, X. Zhou et al., "Data mining to online galvanic current of zinc/copper Internet atmospheric corrosion monitor," *Corrosion science*, vol. 133, pp. 443–450, 2018.