Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 6875207, 13 pages
https://doi.org/10.1155/2022/6875207

Research Article

Construction of Data Mining Analysis Model in English Teaching Based on Apriori Association Rule Algorithm

Shufei Wang

Department of Foreign Language, Harbin University of Commerce, Harbin, Heilongjiang 150028, China

Correspondence should be addressed to Shufei Wang; 15104504977@163.com

Received 23 March 2022; Revised 19 April 2022; Accepted 27 April 2022; Published 15 June 2022

Academic Editor: Wei Liu

How to create an English data mining analysis model based on prior association strategy algorithm is learned. First, the basic principles of data mining and organizational strategy are investigated, and the cooperative strategy algorithm in data mining technology is studied. This document makes a comprehensive analysis of the classical Apriori algorithm and studies the extension of organizational strategy and the technique of deleting participation strategy after decision-making. Then, the application of Apriori algorithm in instruction management is analyzed. There is often a lot of information in command management. We need to collect the data, then sort, and review it. Finally, taking the two adults' classes, class (39) and class (40) as the research objects, of which class (39) has 41 people and class (40) has 43 people. The two classes were brought to the board of directors. The results showed that over the course of 4 months, the students scored better on the mixed tests in the lab than those in the control room. The average score of the subjects before the experiment was 17.6, the average score of the subjects after the experiment was 20.1, and the significant P value of the corresponding t-test was $p \leq 0.001$, less than 0.05, indicating significant differences between the two groups of data. After studying data using writing styles in class, students focused more on complex patterns, concepts, technical expressions, and line patterns. Dynamic assessment of student writing is an effort to improve student writing by using interactive assessment techniques.

1. Introduction

With the development of data information, database technology is more and more mature, and data are used more and more widely. The data information accumulated by the world is also developing at an exponential growth rate. In the 1990s, with the advent and rapid development of the Internet, the whole world was penetrated a small global village by the Internet. People in different places can exchange data and information between each other across time and space and provide help to each other through the Internet. After using the Internet, people are not only facing their own departments, their own units, or their own specific industry databases but also facing the whole ocean of information, covering all fields. From the appearance, there is not much correlation between these data, but if we carefully analyze them, we can find a large amount of technical knowledge information hidden in them. Until the mid-1980s, most databases can only simply perform data entry, data query, and data statistics. Data mining is the process of finding interesting knowledge from many data information stored in databases, data warehouses, or other information bases. The steps and process are shown in Figure 1. The application of this technology can not only enable people to experience and query the previously saved data but also find out the all-in-one relationship between these data, to speed up the transmission of information. In recent years, with the progress of science and technology, people gradually began to widely study rule association mining technology, which has also become an important topic in the direction of data mining. In data mining, association rule is one of the main technologies, which reflects the correlation and dependence between the events. It is also the most common form of mining local patterns in unsupervised learning system. If there is a correlation between two or more things, we can use other related things to predict the occurrence of something [1].

With the diversification of college English courses and methods, improving college English courses attracts more and more attention. Since 1990, English classes in many Chinese universities have completed certification procedures. Its main purpose is to improve the measurement of teaching and promote the improvement of teaching quality through regular monitoring of teaching process [2]. The measurement of good teaching is the important foundation of good teaching. Existing measuring instruments are often given a wide measuring range, which is usually based on multiple measuring levels. For example, the primary level measurement criteria include teacher teaching, syllabus, teaching, teaching results, etc., while the secondary level measurement criteria are the focus of the primary level. According to the role and direction of these indicators in good classroom teaching, different weights are used as a complete measurement standard. There are two main types of assessment: peer assessment and student assessment. This teaching method of quality evaluation has been widely used in colleges and universities and has achieved certain results. However, there are still some differences in the quality assessment process [3].

Based on this, this study argues that the objectives and findings of the key groups that affect English proficiency teaching in colleges and universities are the basis and basis for appropriate evaluation and improvement of teaching. This study presents a method to determine the best use of data mining techniques. Through the analysis of teaching materials, this study tries to find out the advantages of teaching strategies, situations, and environments, in order to provide new ideas and approaches for reviewing and improving teaching quality. Data mining (DM) is an approach to providing data that uses statistical techniques to find certain rules and identities of the data. It is a great way to get knowledge out of big data and is widely used in industrial decision-making and management practice. This study uses data mining to identify the characteristics of good teaching and acquire knowledge of teaching data storage [4].

2. Literature Review

Li and Yu believe that data mining can not only bring huge competitive advantages to enterprises and institutions but also create huge economic benefits for the society. Under such temptation, many famous companies in the world have also entered the ranks of data mining, using data mining systems to develop software and tools related to the company's future development [5]. Cui et al. argue that the composition of quality standards is largely based on the experience of manufacturers. There are a lot of details and attributes in the measurement, and the weight of the measurement is usually based on previous experience, or even the same. Many different disciplines accept the same teaching methods [6]. Wei believe that college English teaching has its own characteristics (such as teacher-student interaction and differences in teaching and events), and many factors that affect good teaching are often hidden. The curriculum, environment, and conditions were not easy to find. Only by demonstrating these situations can an effective teaching model be formed, which should not only be based on the characteristics of college English teaching but also study, appropriate, and exempt education [7]. Kong et al. conducted an empirical study on the characteristics affecting students' evaluation results and found that six characteristics can explain 25.8% of the variation of students' evaluation results. Students' evaluation of teachers' teaching has certain limitations, which can only be used as a reference for teachers to improve teaching [8]. On the other hand, the evaluation of students' performance can be greatly influenced by the way of Glanbock, but on the other hand, it has no influence on the evaluation of students' performance. Therefore, this quality evaluation method is a "result" evaluation method, which can only answer the question of "how" the teaching effect but cannot answer the questions of "why" and "how to improve" the teaching effect [9]. Tseng et al. use data mining technology to learn online English learning platform, college English listening, teaching, test scores, evaluation scores, and administration. For online English learning, assessment panels are used to group students. For English proficiency, the participatory algorithm is used to identify the relationship between output and input skills, and genetic algorithm is used to develop automated English language [10]. In terms of college English listening, Nes and See, based on the understanding of college English listening and research data on the Internet, used information such as gender, teaching, educational objectives, experiment, and language teaching, and found that the first influence was satisfactory in listening test, and the second influence was teaching [11]. Liu et al. believed that compared with other countries, in the field of Chinese companies and organizations, the use of data mining technology to help enterprises and organizations' commercial activities is still in its infancy. There are some good examples of data mining technology in Chinese industry and its application, so there is still a lot of room for improvement in data mining industry, mining technology, industry scientists, and project developers [12].

3. Method

3.1. Data Mining Foundation and Association Rules

3.1.1. Data Mining. With the development of computer technology, the database has evolved from the early file processing to a more powerful database system and further evolved into the current commonly used relational database system. Users can easily access the database through structural query language. With the increasing amount of information, some new problems arise. These fast-growing data are stored in many databases. The traditional database processing technology cannot effectively analyze and apply these massive data and provide users with valuable analysis results. In this context, data mining technology came into being. The operation of extracting useful, implicit, potential, and unknown information from a large amount of data with random interference, noise, errors, missing, and incomplete

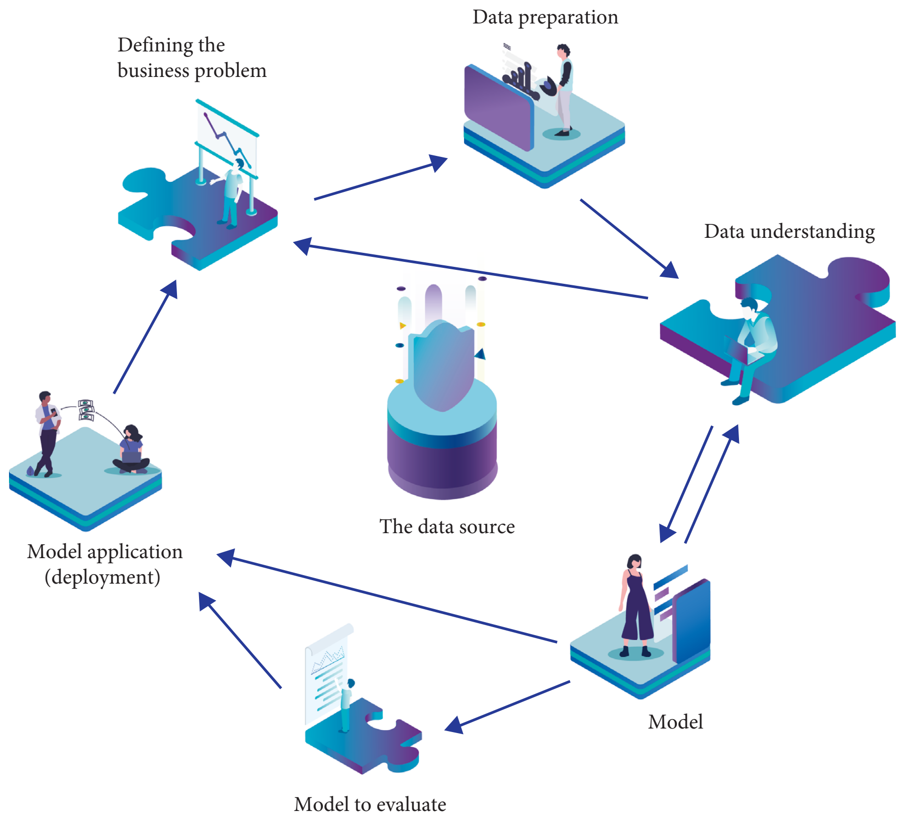

Figure 1: Data mining.

is called data mining. Based on potential knowledge and useful information, data mining technology extracts it from a large amount of data. It is the key link of knowledge discovery data (KDD). Its process is listed in Table 1 [13].

Typical data mining includes user interfaces, measurement models, mining engines, data warehouse servers, and data source text (such as databases, data warehouses, the World Wide Web, and other data). Its structure is shown in Figure 2. The knowledge discovery request put forward by the user is interpreted and converted into a specific pattern through the user interface, submitted to the data mining engine, and mined the cleaned data set through the database or data warehouse server [14]. The patterns found in the mining stage need to be further analyzed and evaluated. If there are redundant or irrelevant patterns, they need to be deleted to generate valuable knowledge and modify the knowledge base through the user. The data mining engine also updates accordingly with the update of the knowledge base.

The process of data mining can be described by three stages: data preparation, mining, and result description and evaluation. Constantly repeating the above three operations is the process of knowledge discovery data (KDD). The data mining process is shown in Figure 3.

(1) Data Preparation Stage. In the data mining process, the time-consuming stage is usually the data preparation stage, accounting for 50% of the whole process. The data preparation process can be divided into three parts: data selection, data preprocessing, and data transmission. The activities for each level are listed in Table 2.

(2) Data Mining Stage. The most concerned problem of scholars and experts in the field of data mining is the data mining stage. This stage really carries out the mining work, which shows the essence of data mining technology. The first is algorithm planning, such as data summary, classification, clustering, association rule discovery, or sequential pattern discovery, that is, to decide what type of data mining method to use. The second is to select an algorithm for the mining method. This directly affects the quality of the mining patterns. The application of data mining algorithm is carried out after the preparation is completed.

(3) Result Expression and Interpretation Stage. According to the different value of information, the extracted information is analyzed and studied, and the analysis results are described based on the user's decision-making objectives. The main tasks of this stage are as follows: deleting redundant and irrelevant information and analyzing and evaluating the data information found in the data mining stage [15]. If the results do not meet the needs of users, then it means that the mining results are inaccurate, and it is necessary to carry out mining calculation again, adjust the parameter value, change the data transformation mode, reselect the data, and even select a new data mining algorithm. In order to intuitively show the results of data mining to people, we should visually process the mining results, for example, "if...then..." rules describe the classification decision tree, which simplifies the difficulty of data analysis.

Table 1: Knowledge discovery process of database.

Step	Name	Operation
Step 1	Data cleaning	Clear noise or inconsistent data
Step 2	Data integration	Many files can be merged. Along with cleansing the data, this is considered the first step to storing the data in a data warehouse
Step 3	Data selection	Retrieve data related to the analysis task from the database
Step 4	Data transformation	Transform or share data through content or assembly into a format suitable for mining
Step 5	Data mining	Use intelligent methods to extract data patterns

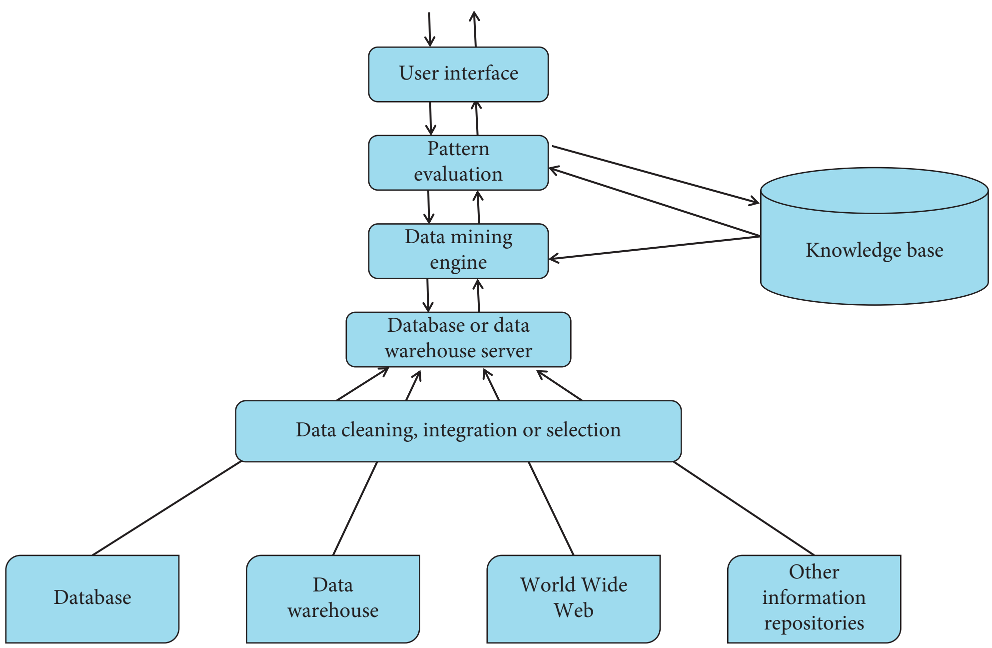

Figure 2: Structure of typical data mining system.

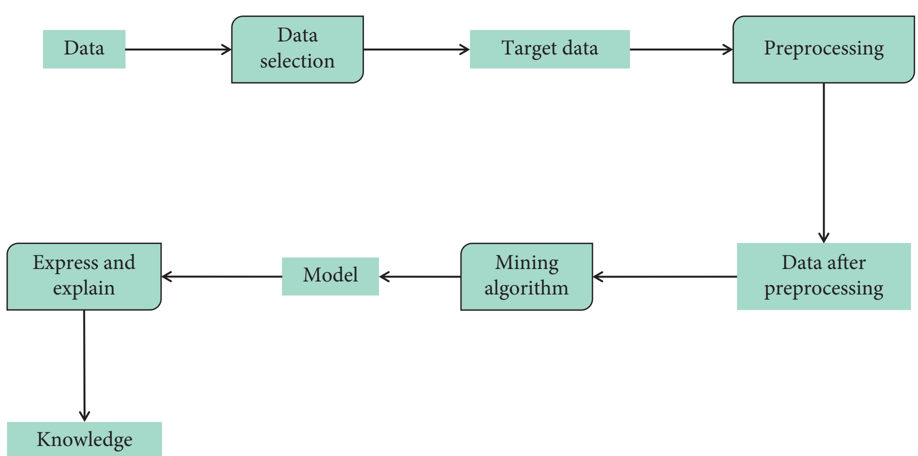

Figure 3: Data mining process.

3.1.2. Association Rules. The association discovery policy identifies associations interested in or affiliated with products in various forms. It is a major of data mining and has been widely studied in the industry in recent years.

At present, there are many different types of cultural organizations. Because of the grouping changes to the code, they can be divided into Boolean and numeric associations; according to the analysis of the data in the rule, it can be divided into a set of rules and a set of rules; depending on the size of the file involved in the code, it can also be divided into rule combinations and multiple consortions [16].

(1) Principle of Association Rule Algorithm. Association rules are ruling whose support and trust meet the threshold given

Table 2: Three substages of data preparation.

Data preparation	Explain
Data selection	Selecting a file usually involves deleting data affected by an existing file or database to create a target file
Data preprocessing	Data preprocessing process extracts data to meet the requirements of data mining
Data transformation	The main purpose of exchanging data is to reduce the width of the data, that is, to find the main features of the first feature in order to reduce the number of features or differences in data mining

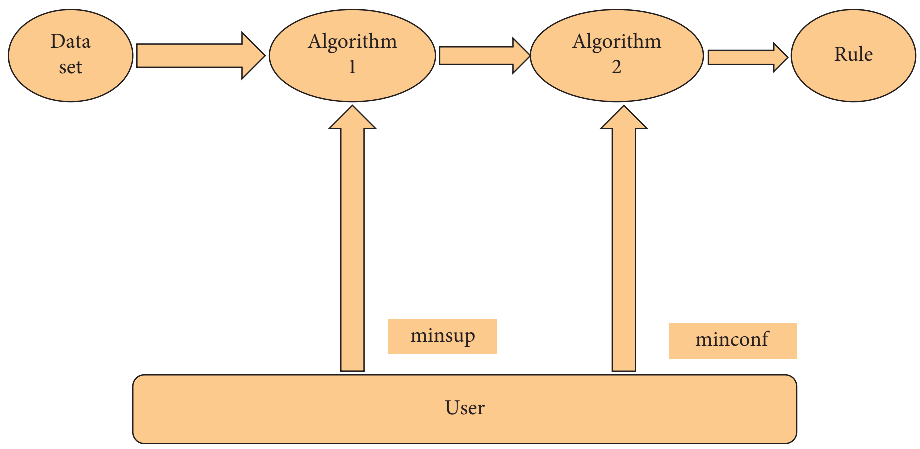

Figure 4: Basic model of association rule mining.

Table 3: Association rule mining steps.

Step	Operate
Find out the frequent item set	Find all frequent items, whose support is no less than a given threshold
Generate strong association rules	Find out the set of frequent items, namely, support and confidence, respectively

by users respectively, that is, strong rules. The data sample set corresponding to strong rules must appear frequently in the sample training set. We can deduce the confidence of the corresponding association rules according to the support of these frequent sample sets. In other words, the whole process of association rules is divided into two stages.

First, the frequent item set in the sample training set is determined by the preset minimum support threshold. This step is the central problem of the whole mining process, and it is also an important index to determine the efficiency of the algorithm. All association rules are generated according to the obtained frequent item set and the minimum confidence threshold. The implementation process of this process is as follows: for each element in the frequent item set, that is, each frequent item generates all nonempty subsets and calculates the confidence for each nonempty subset. If the confidence is greater than or equal to the minimum confidence, then the strong rule is output. The basic model is shown in Figure 4.

The basic steps of association rule mining are listed in Table 3.

In the process of mining association rules, the first step is the focus and key. The performance and effect of mining association rules directly depend on the execution results of the first step. Therefore, the first step is the main starting point of mining association rules algorithm. Compared with the first step, the second step is much simpler. All association rules of frequent itemsets are listed, and these association rules are evaluated and measured. Based on the confidence and support threshold, valuable association rules must meet the requirements of confidence and support threshold. In fact, all association rules generated by frequent itemsets must meet the requirements of support threshold. Therefore, there is no need to judge the support threshold of association rules, if the association rules are measured according to the requirements of confidence threshold, which simplifies the evaluation process.

Usually, users specify minimum support (min SUP) and minimum confidence (min CONF) according to mining needs. Support and confidence are two important concepts to describe association rules. The specific definitions are as follows:

(1) Support:

$$\text{support}\,(AB) = P\,(A \cup B) = \frac{NH_{A-B}}{N} \times 100\%. \tag{1}$$

That is, the probability that two itemsets A and B appear simultaneously in transaction set D, which is used to measure the statistical importance of association rules in the whole data set.

(2) Confidence:

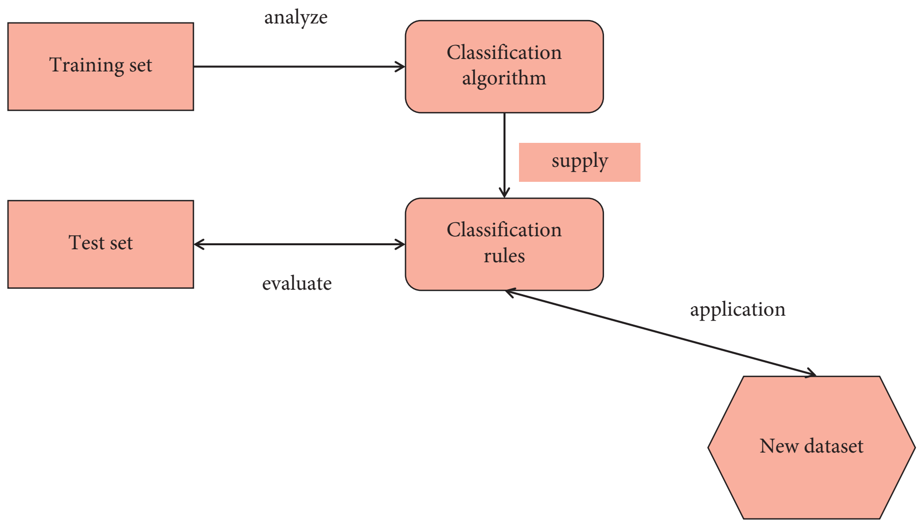

Figure 5: Data classification process.

$$\text{confidence}(AB) = P(A \cup B) = \frac{NH_{A-B}}{N} \times 100\%. \quad (2)$$

That is, in the transaction set D of itemset A, the probability that itemset B also appears at the same time [17].

The task of data analysis is classification and prediction. The classification task is to map each attribute set X to a predefined class label y by learning or obtaining an objective function F. Classification is mainly used to predict the category of data objects. First, according to the samples in the given training sample set, an appropriate classification model, that is, classifier, is constructed. In the process of classification, the classifier is used to specify the most appropriate class mark for the sample instances that have not determined the class mark in the data set.

The whole classification process can be divided into two stages: learning and classification. First, the classification algorithm is used to analyze the training data set and obtain the classification rules, which is the so-called learning stage; The second stage is the classification stage, which tests the accuracy of the classification rules obtained by the data set evaluation. If the accuracy is within the allowable range, the rule will be used in the new classification process. Otherwise, the classification rule will be abandoned. The whole process is shown in Figure 5.

We briefly describe the classification as follows: if a specific instance form can represent $(a_1, a_2, \ldots, a_n, c)$, where $a_i\,(i = 1, 2, \ldots, n)$ identifies the attribute field value and C represents the class mark. We briefly describe the classification as follows: given the training data sample set $D = \{x_1, x_2, \ldots, x_n\}$, the goal of the classification task is to analyze the data set D and determine a mapping function gate $f(A_1, A_2, \ldots, A_n) \longrightarrow C$, so that the instance $xi = (a_1, a_2, \ldots, a_n)$ of any unknown category can be identified by the appropriate class label C.

The amount of information can only reflect the uncertainty of the symbol, and the information entropy can be used to measure the overall uncertainty of the whole source X. Let something have n mutually independent possible results (or states), the probability of each result is $P(X_1), P(X_2), \ldots, P(X_n)$, and there are

$$\sum_{i=1}^{n} P(X_i) = 1. \quad (3)$$

Then, the uncertainty $H(X)$ of the thing is as follows:

$$H(X) = \sum_{i=1}^{n} P(X_i)\log_2 P(X_i). \quad (4)$$

Information gain is used to measure the expected reduction of the direct line, which can help us to ask the least questions when classifying the sample set. Suppose that the training set of tuples marked with class D has m different values of class label attribute [18], then it is recorded as $C_i\,(i = 1, 2, \ldots m)$. $|D|$ and $|C_{1,D}|$ respectively represent the number of tuples in D and the number of principles of class C_1 in D. If P_i is the probability value that a tuple in D belongs to class C_1, then the direct $Info(D)$ of D is described as the expected information for the classification of tuples in D, and the formula is written as follows:

$$\ln fo(D) = \sum_{i=1}^{m} \log_2(P_i). \quad (5)$$

Information gain is defined as the difference between the original classified information demand and the new classified information demand, which is as follows:

$$gain(D, A) = \ln fo(D) - \ln fo_A(D). \quad (6)$$

Therefore, the obtained information gain $gain(D, A)$ can also be defined as follows:

$$\text{gain}(D, A) = \ln fo(D) - \sum_{j=1}^{v} \frac{|D_v|}{|D|} \times nfo(D_j), \quad (7)$$

gain(D, A) indicates the expected compression of the lineage caused by clarifying attribute A. The larger gain(D, A)

is, the more information the test attribute A provides for classification. Therefore, for each attribute, it is sorted according to its information gain, and the attribute that obtains the maximum information gain is selected as the branch attribute. ID3 algorithm uses information gain as the splitting basis [19].

3.2. Apriori Algorithm. Apriori algorithm to realize association rule mining includes two basic steps: the first is to mine frequent itemsets from the transaction database and then generate association rules based on frequent itemsets.

Finding frequent itemsets layer by layer iteratively is the key and core of Apriori algorithm, which can be completed by pruning and connecting itemset. Generating a set C_k of candidate $k-$-item sets for the k-item frequent itemset L is the key to the connection step, which is realized by connecting the $k-$-item frequent itemset L_{k-1} with itself. The whole transaction data set is scanned and compared with the transaction items with the preset minimum support threshold, and the events with support lower than this threshold are deleted, which is called the pruning step.

Taking the generation of L_k through L_{k-1} as an example, this study explains how to use the Apriori algorithm when mining frequent items. The purpose of generating L_k through L_{k-1} can be realized by connecting and deleting.

3.2.1. Connection Steps. The two itemsets are connected in L_{k-1} and the candidate set C_k of L_k is calculated to find L_k Suppose L_{k-1} contains two itemsets l_i and l_j right, and the jth item in l_i and the penultimate item in l_i are represented by $l_i[j]$ and $l_i[k-2]$ respectively. In order to simplify the analysis process, it is assumed that the records in the database are stored in dictionary order. $L_{k-1} \oplus L_{k-1}$ is used to represent L_{k-1}-connection operation, which means that if l_1 and l_2 have the same first $(k-2)$ items, and l_1 and l_2 in L_{k-1} can be connected. In order to ensure the uniqueness of itemset, the constraint condition of $l_1[k-1] < l_2[k-1]$ must be satisfied.

3.3. Implementation of Apriori Algorithm in Teaching Management. The purpose of the joint venture is to find relationships between products in the database, which is a shopping basket analysis. The most famous example is the "Butt and Beer" story. The organizational policy has many uses. In sales, organizational rules can be applied to sales to make more money; as far as the insurance industry is concerned, if there is a difference in demand, then it could be fraudulent and needs further investigation. In terms of treatment, we can see combination therapy, talk about marketing, identify customers, and approve services of interest. The Apriori algorithm is one of the most classical algorithms in the custom organization [20].

Teaching management is very informative. This information provides flexibility for our ordinary business, but how does the information generated by these industries affect our management? Can we find some social organizations that can play a role in improving our day-to-day governance and decision-making? This concept is still rare in our approach to teaching management. The purpose of the system is to identify the relevance of the data generated during the course, draw final conclusions, and send them to management as a basis for supporting decision-making.

Command management often has a lot of data. We must first compile the data, then review and analyze the data, and store it in an archive. Then, the improved Apriori algorithm is used to eliminate the influence of these data in data mining technology, produce negative results, identify results, then make some adjustments, and finally produce the desired results from the support decision, presented in a computer. Therefore, the system can be divided into data acquisition data, collision data and analysis, operation data, collision, and presentation modules.

The design model of improved Apriori algorithm in instruction management is shown in Figure 6.

(1) Information Received and Received in Advance. This module is used to identify and compile all kinds of data in the management system, identify the data useful to the system, and extract and modify the format of these data for the system to use.

(2) Filter and Clean Files. Cleaning data in a data store is used to generate data. It can clean data, select subdevices, remove inaccurate data and duplicate data, and view important information in the process. Preliminary data can improve the quality of discovery data. After selecting the features of the former data, the research and exploration are completed by using analytical analysis, system analysis, classification, and clustering algorithms [21].

(3) The Improved Apriori Algorithm Was Used to Analyze the Data. The improved Apriori algorithm is used to create shared files that filter and clean up the data to get the results we used.

(4) Edit Information. Results may differ slightly due to different data storage and analysis algorithms. Therefore, the deviation case is processed and the second result is obtained. If the results are not satisfactory, then the archive is edited and finally, the results we want are obtained.

Main functional modules are as follows:

(1) *Data Received Before the Module Runs.* The module is designed by receiving data and predata. Data retrieval is typically based on data collected from each business report in the daily command management system. It is only extracted from the database. The business process data need to be analyzed into what we know and then converted into system files. Database processes are typically handled by SQL server: Windows databases.

(2) *Data Filtering Module.* This module uses data extraction to generate data for analysis. Due to the separation of different characteristics, the requirements for data are also different. Therefore, the main purpose of this module is to filter out the data that

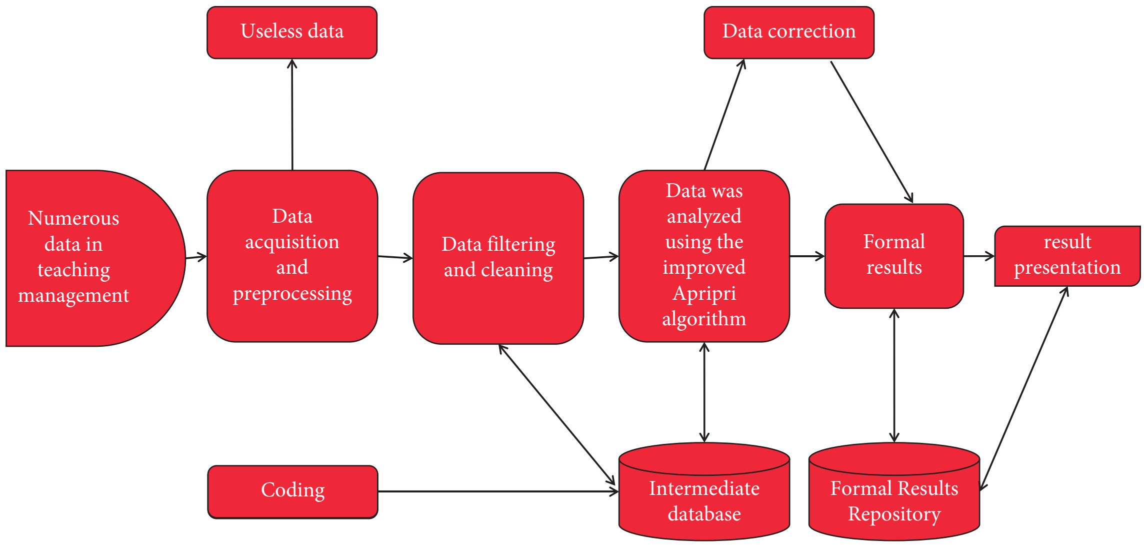

Figure 6: Application structure design of improved Apriori algorithm in teaching management.

does not conform to this specification, find the data we need, and remove the features and weak data recovered from data. The final analysis results are stored in the mean data, and these important data are separated to achieve the effect of data cleaning.

(3) *Data Analysis Module*. Data analysis is an important part of system operation. Its job is to use data from a central database as a data source. These data are then analyzed one or more times using the improved Apriori algorithm in order to get the relationship hidden in the data and make it useful to us [22].

(4) *Modification of Module Files*. Since the effectiveness of Apriori algorithm in data mining is unpredictable, we cannot be sure whether it will perform well after being tested. These models are used to edit and reevaluate the initial data of the results, resulting in multiple conclusions. If the conclusions are the same, then we consider the conclusions to be valid. The data editing process is shown in Figure 7.

(5) *Result Presentation Module*. The final result of data mining needs to be presented to the managers of the unit for decision analysis, so the first thing is to ensure the accuracy and reliability of the data. At the same time, it is also easy to understand the results, so the result presentation module generally adopts intuitive reports and icons, which are very common in common office software.

4. Experimental Analyses

4.1. Subjects. To answer the research questions, two adults from class 39 and class 40 were selected as research materials, with 41 students from class 39 as the laboratory and 43 students from class 40 as the administration room. Before the experiment, the composition of the final examination of the next semester of high school was selected as a pretest to determine the writing differences between the two classes. The reasons for choosing these two classes as experimental classes are as follows: first, they are taught by the author, which excludes the influence of teachers' characteristics on the research; second, the two classes use the same textbooks and the same writing style, eliminating the interference of different courses difficulty; finally, the average writing score was 17.6 in the test room and 18.8 in the control room, based on measurements obtained before the experiment.

4.2. Experimental Method

4.2.1. Questionnaire Survey Method. The questions were divided before and after the experiment. The main purpose of the questionnaire is to understand the current situation of students' writing, that is, students' English writing behavior, students' assessment of their writing ability, and the current situation of students' writing ability. The purpose of the post-test is to measure whether the improvement of DA standard and its application in English writing can improve students' writing skills and abilities [23].

4.2.2. Quantitative Analysis. This study uses the quantitative analysis method and SPSS 22.0 social science statistical software to statistically analyze the experimental data obtained in the process of teaching. The analysis and comparison of students' English composition scores before and after the experiment can be done by providing the data support; thus, the reliability of this study can be increased.

4.2.3. Interview Method. To better understand how students feel about the DA standard being used in high school English writing, the course also uses interviews. It is an additional source of information after the questionnaire and written test. This interview will help teachers better understand the changes in students' writing behavior after the experiment, whether their writing ability has improved, and whether they

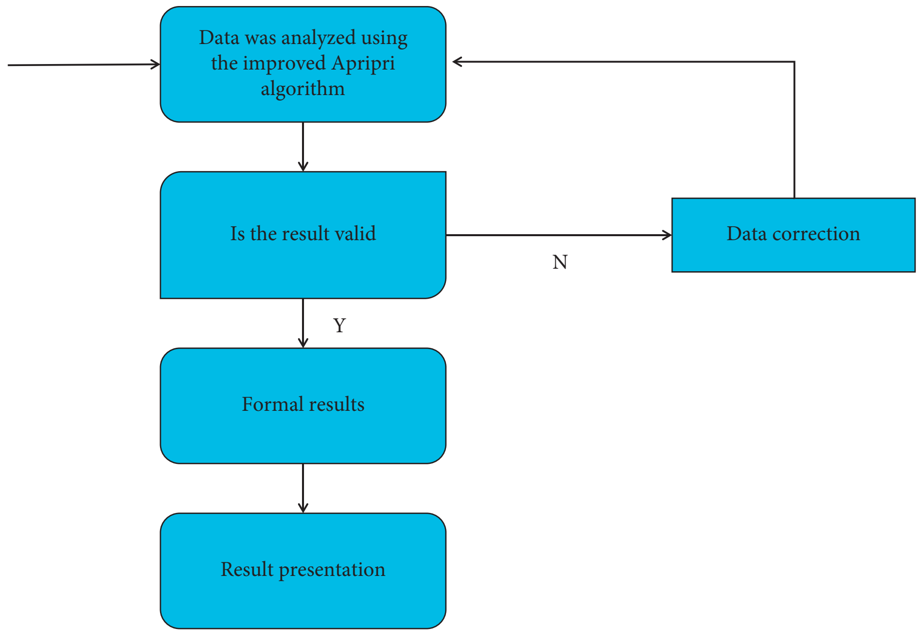

Figure 7: Flow chart of data correction.

are interested in the design and applied teaching of English writing classroom.

4.3. Research Process. This study consists of three stages: in the first stage, the students of the two classes fill in the preexperiment questionnaire in order to understand the students' current English writing attitude. When filling in the questionnaire, students should truthfully fill in the questionnaire in combination with their own personal experiences.

4.3.1. Construction of Writing Teaching Process Based on Dynamic Evaluation Model. In the experimental class, according to the characteristics of intrusive and interactive dynamic assessment introduced in Section 2, combined with the specific characteristics of English writing teaching, the author intends to apply the constructed.

DA model to the teaching practice of 4 months. The experimental process of DA mode is mainly divided into the following five stages: prewriting stage, mutual evaluation stage, modification stage, teacher evaluation stage, and final draft stage.

4.3.2. Specific Operation of Writing Teaching in Dynamic Evaluation Mode. Mainly through the use of the constructed DA model for teaching, one of the composition lessons is selected to show how the DA model is applied to the English writing class in senior high school [24].

4.3.3. Control Class Teaching Process. In the English writing class of the control class, the teachers still use the traditional teaching method; that is, according to the process of "teachers assign writing tasks, students write independently, then students hand in their compositions, teachers evaluate and teachers comment on the model composition." After the teacher assigns the writing task, he does not give any tips and help to the students. In the writing process, the middle school students are not allowed to use reference books or communicate with the students. They are handed in after completing the writing in class within the specified 30 minutes. The students' composition is a one-time manuscript, which is directly handed over to the teacher for evaluation and scoring without modifying the first and second drafts. Teachers only give a comprehensive score in the scoring link, rather than scoring separately. Students pay more attention to the scores assessed by teachers, not to the writing process and the help obtained in the process. Finally, the teacher leads the whole class to appreciate the standard model essay, analyzes the vocabulary and advanced sentence patterns used in the model essay, and then asks the students to imitate and apply them to the next composition writing.

4.4. Data Collection and Analysis. The data used in the experiment were one adult's final English test and two adults' writing scores in the final English test. The scores of the composition examinations are given independently. The concept of composition is close to college and students' life in the entrance examination, which is very difficult. Test scores are based on high school English standards. Therefore, the reliability of test data is very reliable. By observing the changes in test scores after the laboratory and control room, whether the standard of achievement is suitable for improving students' writing ability is determined. By observing changes in test scores before and after laboratory tests, this study determines whether Da applied standards in English teaching can support students' writing resources [25].

In order to understand the validity of this experiment, questions and interviews were also used in this study. The

Table 4: Questionnaire of students' writing status before the experiment.

Question number	Percentage (%)				
	①	②	③	④	⑤
1	0	2.4	7.3	23.3	58
2	7.2	14.6	30.5	24.4	16.3
3	12.4	19.5	34.2	23.2	10.3
4	8.2	27.2	8.2	35.4	17
5	6.1	15.9	47.5	21.6	8.9
6	6.1	17.1	58.5	15.3	3.7
7	19.8	24.4	42.7	7.6	2.8
8	28.1	30.5	32.9	7.3	0
9	12.3	28	39.1	14.5	2.5
10	0	1.2	7.8	22.1	69.5

questionnaire is divided into pretest questionnaire and post-test questionnaire. The pretest questionnaire consists of 10 questions and usually examines the current situation of English writing students. The post-test also contains 10 questions designed to explore the benefits of using the Da model for English teaching.

In this experiment, SPSS 22.0 social science statistical software is used to statistically analyze the experimental data.

4.5. Experimental Analysis. Before the experiment, 85 questionnaires were distributed to the laboratory and control room. The author asks the students of both classes to answer at the same time and brings them back to the scene. All 85 questionnaires were returned and all 85 questionnaires were valid. The pre-experiment questionnaire consisted of 11 small questions. The first two questions are designed to assess students' current writing behavior. Questions 3 to 6 are students' assessment of their own content, questions 7 to 8 are students' assessment of their current writing level, and the last two questions are students' assessment of the writing process. After collecting the data query, the author analyzed the research data from various angles. The evaluation results are listed in Table 4.

Questions 1–2 of this questionnaire are intended to investigate the current students' attitude towards writing. According to the data in the table above, 65% of the students in question 1 agree very much, 21.2% agree that English writing is very important, and no students disagree very much that English writing is very important. Therefore, up to 89.9% of the students agree on the importance of English writing. Question 2 talks about whether students like English writing class. About 30.5% of students have a vague attitude, 15.3% of students say they do not like English writing class, and 7.1% of students do not like English writing class very much, suggesting that students' current English writing attitude needs to be changed.

After the 4-month writing teaching experiment in the experimental class, the author issued a second questionnaire to the experimental class. This time, the author wants to investigate the effect of students' application of dynamic assessment model in English writing teaching after the experiment. There are 40 people in this experimental class.

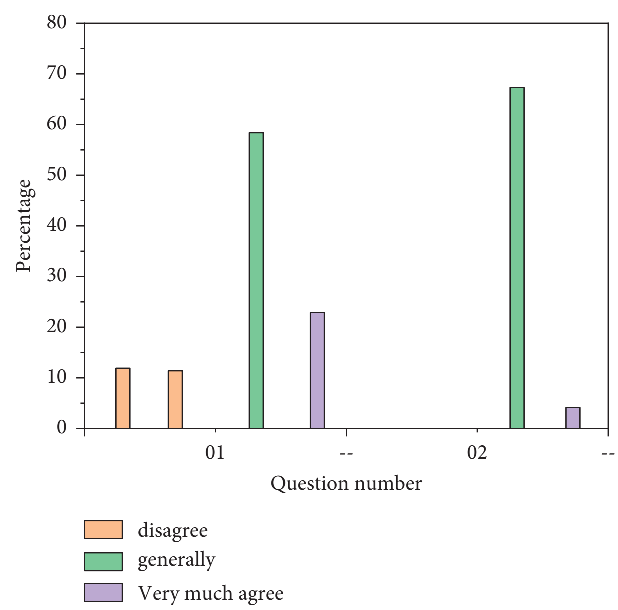

Figure 8: The influence of DA mode on students' writing confidence and habits.

About 40 questionnaires are distributed, 40 are recovered, and 40 valid questionnaires are available. Students are allowed to answer and hand them in within a unified time. The questionnaire contains 10 subquestions in total. Now, the percentage of people corresponding to each subquestion and each option is displayed in the form of column chart, as shown in Figure 8:

Questions 1 and 2 are to understand whether the students' writing confidence and writing habits have changed after the experiment compared with those before the experiment. After 4 months of dynamic evaluation teaching experiment, the author again investigated whether the students in the experimental class are confident in writing English compositions and whether they are used to making sketches and making outlines and revising them for many times. It can be clearly seen from Figure 8 that 57% of the students are confident in writing English compositions and 23% of the students are very confident in writing English compositions. Nearly half of the students in the first two classes had vague attitudes when asked whether they were confident of writing well, which shows that the dynamic assessment model is conducive to increasing students' confidence in English writing; as for question 2, the chart data show that 72% of the students have formed the habit of writing and revising their compositions for many times after the experiment, while the survey results of the same questions in the two classes before the experiment show that only 12% of the students will do so, which shows that the dynamic evaluation model has a good impact on students' writing habits.

In Figure 9, the author mainly investigates the students' recognition of DA mode and each process in the experimental class. Questions 3–6 of this questionnaire are to investigate the feelings of students in each link of DA mode writing. In question 3, none of the students in the class dislikes brainstorming very much. The proportion of students who like this link very much is as high as 70%. In the peer evaluation link of question 4, more than half of the students like this link. Students have slowly begun to adapt to the idea that teachers are no longer the only composition evaluators, and students can also help each other evaluate compositions [26].

Questions 7 and 8 in Figure 10 mainly investigate the impact of DA mode on students' writing interest and writing ability. As can be seen from question 7 in the figure, 55% of the students agree that DA mode improves students' interest in writing, and 20% of the students agree very much with the impact of DA mode on students' interest in writing. In question 8, none of the students strongly disagreed that DA mode can improve writing ability, and 65% of the students agreed that DA mode can improve writing ability. DA mode has a great impact on students' writing interest and writing ability.

To prove that there was no significant difference in English writing scores between the laboratory and the control room, the authors had students from both classes take the predictive test scores. The current SPSS 22.0 identification data are listed in Table 5:

Table 5 lists 40 students in the lab and 42 students in the control room. The standard deviation of the test class was about 1.37, and that of the control class was about 1.57. The average pretest score before the class was about 16.7, and that before the control class was about 18.9. There was a 0.2 difference between the pre-experimental class and the control class.

Four months later, in order to check the application of dynamic measurement standards in high school English classes and to see whether there is any difference in the scores of English writing test between the examination room and the control room, the city unified English test was taken. At the end of the last semester of the senior year, it is based on the follow-up exam section. The results of target analysis are listed in Table 6.

In Table 6, the scores between the test room and the control room in the later tests were higher than in the previous tests. The average score of the subjects in the pretest increased from 17.6 to 20.1, and the average score of the control class in the pretest increased from 18.9 to 20.6. After

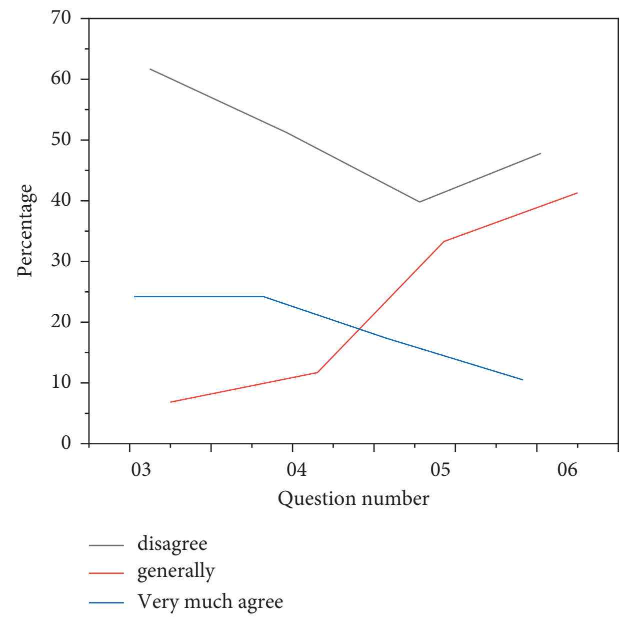

Figure 9: Students' recognition of each process of DA mode.

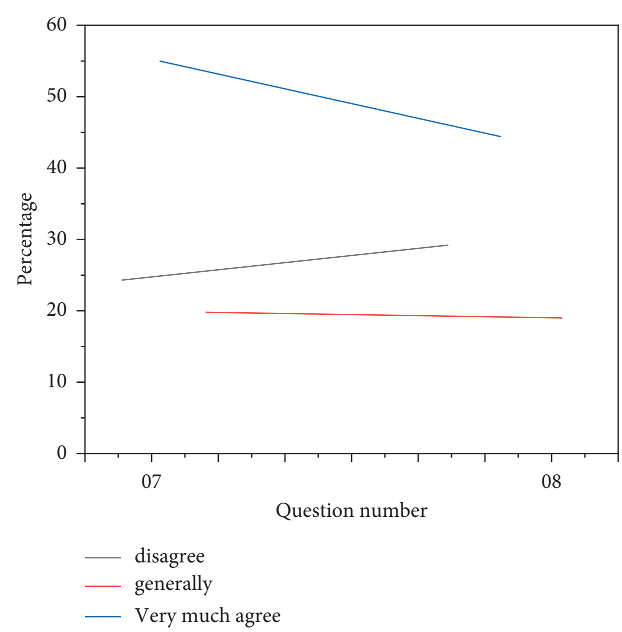

Figure 10: The influence of DA mode on students' writing.

4 months of English writing study, the result of the composition was good. Both classes have improved. The difference is that the results of the experimental class and the control class differ greatly in the post-test. At the same time, due to the use of variables, the experimental class achieved faster results, and the average score of the experimental class was 1.3 points higher than the control class.

Finally, students' scores on all types of questions have the greatest impact on students' passing exams and making responsible decisions, depending on where they study. The main model used in the experiment was students who had taken multiple English language courses or English tests.

Table 5: Description and statistics of pretest groups in experimental class and control class.

	Front side grouping	Figure	Average value (E)	Standard deviations	Standard error mean value
Class	Control class	41	17.8532	1.53742	24233
	Experimental class	43	17.5000	1.57165	25106

Table 6: Grouping description and statistics of post-test composition scores in experimental class and control class.

	Backside grouping	Figure	Average value (E)	Standard deviations	Standard error mean value
Backside	Control class	41	20.1340	1.19872	19543
	Experimental class	43	18.5554	1.30254	18792

When students have no special weakness in listening, speaking, reading, and writing, the probability of passing the exam is high; when students can complete the previous level through all types of questions, students are more likely to pass the final exam; in addition, students who pass CET 4 and 6 with high scores also have a higher chance of passing the exams. It can be seen from the data that the scores of multiple-choice questions, final level, CET 4, and CET 6 will all affect boys' basic English, and there are reasons for girls' learning. Therefore, students' basic English level is conducive to passing the exam.

5. Conclusion

To determine the teaching benefits of developing quality measurement standards and organizational policy algorithms and apply them to high school English teaching, the author conducted equivalent experiments in two classes and examined questionnaires and tests before and after exams. The material is discussed. The results show that the research problem has been successfully solved; that is, the use of dynamic measurement model in high school English class can improve students' satisfaction and grades. According to the characteristics of data processing and response behavior developed by the system, this study understands the theory and technology of data mining and machine learning. Using the logistic regression model, the improved decision tree model, and the model after integrating the logistic regression model and the improved decision tree model, the students' previous estimation can be completed. In classification problems, the logistic regression model is the most used algorithm in the industry, such as the media click pass value problem, while the improved logarithmic model decision is a powerful performance algorithm with multiple decision tree models. The prediction results are more accurate than the one-to-one decision model, which confirms the application of data mining technology in the problem classification of college English.

The systematic collection of body shape data is insufficient, especially the behavior data of students in English practice or test. All kinds of information are needed, both from the point of view of creating a student user profile and from the point of view of predicting a student's English proficiency in the past. There are many factors that can affect a student's ability to pass an exam later; for example, there are many factors that can affect a student's ability to pass an exam. In addition, there are many factors to consider in subsequent studies, such as not taking the test.

Data Availability

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

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This study was supported by 2021 Harbin University of Commerce Teacher's "Innovation" Project, The Practical Research of Cultivating the Compound Foreign Language Teachers in Heilongjiang Province from the Perspective View of ESP (20YYD218).

References

[1] P. H. Lu, J. L. Keng, F. M. Tsai, P. H. Lu, and C. Y. Kuo, "An apriori algorithm-based association rule analysis to identify acupoint combinations for treating diabetic gastroparesis," *Evidence-based Complementary and Alternative Medicine*, vol. 2021, no. 17, pp. 1–9, Article ID 6649331, 2021.

[2] G. Liu, "Information platform for classroom teaching quality evaluation and monitoring based on artificial intelligence technology," *Big Data Analytics for Cyber-Physical System in Smart City*, Springer, Singapore, pp. 149–155, 2020.

[3] T. S. Murthy, M. S. Roy, and M. K. Varma, "Improving the performance of association rules hiding using hybrid optimization algorithm," *Journal of Applied Security Research*, vol. 15, no. 3, pp. 423–437, 2020.

[4] W. S. Leite and R. C. D. Lamare, "List-based omp and an enhanced model for doa estimation with non-uniform arrays," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 6, pp. 4457–4464, 2021.

[5] J. Li and K. Yu, "An alternative to the bathe algorithm," *Applied Mathematical Modelling*, vol. 69, pp. 255–272, 2019.

[6] J. Cui, L. Kan, F. Cheng et al., "Construction of bifunctional electrochemical biosensors for the sensitive detection of the sars-cov-2 n-gene based on porphyrin porous organic polymers," *Dalton Transactions*, vol. 51, no. 5, pp. 2094–2104, 2022.

[7] L. Wei, "Study on the application of cloud computing and speech recognition technology in English teaching," *Cluster Computing*, vol. 22, no. S4, pp. 9241–9249, 2019.
[8] X. Kong, Y. M. Wang, X. Wang et al., "Monitoring damage evolution in a titanium matrix composite shaft under torsion loading using acoustic emission," *Acta Metallurgica Sinica*, vol. 32, no. 10, pp. 1244–1252, 2019.
[9] K. Glanbock, "Donaldo macedo (ed.) decolonizing foreign language education—the misteaching of English and other colonial languages," *Language Policy*, vol. 20, no. 1, pp. 135–137, 2021.
[10] J. J. Tseng, Y. S. Cheng, and H. N. Yeh, "How pre-service English teachers enact tpack in the context of web-conferencing teaching: a design thinking approach," *Computers & Education*, vol. 128, pp. 171–182, 2019.
[11] A. Nes and B. H. See, "Does explicit teaching of critical thinking improve critical thinking skills of English language learners in higher education? a critical review of causal evidence," *Studies In Educational Evaluation*, vol. 60, pp. 140–162, 2019.
[12] Y. Liu, Z. Yu, and Y. Yang, "Diabetes risk data mining method based on electronic medical record analysis," *Journal of Healthcare Engineering*, vol. 2021, no. 6, 11 pages, Article ID 6678526, 2021.
[13] R. Zhou, H. Chen, H. Chen, E. Liu, and S. Jiang, "Research on Traffic Situation Analysis for Urban Road Network through Spatiotemporal Data Mining: A Case Study of Xi'an, china," *IEEE Access*, vol. 9, no. 99, p. 1, 2021.
[14] N. A. Fauzi, N. Ali, P. J. Ker, V. A. Thiviyanathan, and L. H. Mun, "Fault Prediction for Power Transformer Using Optical Spectrum of Transformer Oil and Data Mining Analysis," *IEEE Access*, vol. 8, no. 99, p. 1, 2020.
[15] J. Hu, J. Fang, Y. Du, Z. Liu, and P. Ji, "Application of pls algorithm in discriminant analysis in multidimensional data mining," *The Journal of Supercomputing*, vol. 75, no. 9, pp. 6004–6020, 2019.
[16] J. Xu and Y. Liu, "Cet-4 score analysis based on data mining technology," *Cluster Computing*, vol. 22, no. S2, pp. 3583–3593, 2019.
[17] R. Aaij, S. Benson, M. D. Cian et al., "A comprehensive real-time analysis model at the lhcb experiment," *Journal of Instrumentation*, vol. 14, no. 04, 2019.
[18] X. Zhu and Y. Liu, "An efficient frequent pattern mining algorithm using a highly compressed prefix tree," *Intelligent Data Analysis*, vol. 23, pp. S153–S173, 2019.
[19] M. S. Akhtar, P. Sawant, S. Sen, A. Ekbal, and P. Bhattacharyya, "Improving word embedding coverage in less-resourced languages through multi-linguality and cross-linguality: a case study with aspect-based sentiment analysis," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 18, no. 2, pp. 15.1–15, 2019.
[20] U. Khadam, M. M. Iqbal, M. A. Azam, S. Khalid, and N. Chilamkurti, "Digital watermarking technique for text document protection using data mining analysis," *IEEE Access*, vol. 7, no. 99, p. 1, 2019.
[21] Y. Wang, H. Ye, T. Zhang, and H. Zhang, "A data mining method based on unsupervised learning and spatiotemporal analysis for sheath current monitoring," *Neurocomputing*, vol. 352, pp. 54–63, 2019.
[22] Y. Shi, T. Yu, Q. Liu, H. Zhu, and Y. Wu, "An approach of electrical load profile analysis based on time series data mining," *IEEE Access*, vol. 8, no. 99, p. 1, 2020.
[23] C. H. Chen, J. S. He, T. P. Hong, and S. R. Kannan, "Post-analysis framework for mining actionable patterns using clustering and genetic algorithms," *IEEE Access*, vol. 7, no. 99, p. 1, 2019.
[24] H. Zou, "Clustering algorithm and its application in data mining," *Wireless Personal Communications*, vol. 110, no. 1, pp. 21–30, 2020.
[25] N. Fatima, L. Li, S. Hong, and H. Ahmed, "Prediction of breast cancer, comparative review of machine learning techniques and their analysis," *IEEE Access*, vol. 8, no. 99, p. 1, 2020.
[26] J. Zhang, L. Chu, C. Guo, Z. Fu, and D. Zhao, "A novel energy management strategy design methodology of a phev based on data-driven approach and online signal analysis," *IEEE Access*, vol. 9, no. 99, p. 1, 2021.