Research Article

Markov Model-Based Curriculum Association Classification Model and Student Achievement Prediction

Hui Gu and XiuFen Deng

Guangzhou Huashang Vocational College, Guangzhou, Guangdong, 510000, China

Correspondence should be addressed to XiuFen Deng; xiufendeng@gzhsvc.edu.cn

Received 7 April 2022; Revised 4 May 2022; Accepted 10 May 2022; Published 23 May 2022

Academic Editor: Chia-Huei Wu

Copyright © 2022 Hui Gu and XiuFen Deng. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the increasing popularity of higher education and the increasing scale of students in schools, the educational administration management system of various colleges and universities has also accumulated a large amount of performance data. In the face of these massive data, many teaching managers still stay in the simple operation of adding, deleting, modifying, and checking the data and cannot effectively extract and analyze the useful knowledge and information hidden behind the data. Therefore, Markov model is proposed. The advantage of the Markov model is that it has better prediction effect on random series and data series with large volatility; that is, grey prediction model is used to reveal the overall trend of development and change of prediction data series. This paper studies the curriculum association classification model and student achievement prediction based on the Markov model. According to the student’s historical achievement, the average missed detection rate of future grade point is 48.65%, the average missed detection rate of Apriori algorithm is 35.5%, the average missed detection rate of FP growth algorithm is 43.2%, the average missed detection rate of this algorithm is 37.5%, and only 17 of the 40 grade point interval predictions match the actual interval. Make full use of the association rules between courses to provide students with early warning and teacher guidance for specific courses, which can effectively reduce the failure rate while reducing the academic burden. The Markov model can mine students’ data inside and outside the classroom, establish and improve college students’ achievement index system, then deeply analyze and discuss the development of college students’ achievement and ability, and finally give each student’s achievement results.

1. Introduction

Modern higher education from elite to mass, from unitary to pluralistic, leads to the rapid expansion of enrollment and the increasing shortage of teaching resources, which makes it difficult to teach students according to their aptitude. On the other hand, there are many kinds of courses. Taking communication engineering as an example, professional courses such as digital logic circuit, signal and system, and communication principle have potential related leading and follow-up courses, thus forming a complex tree structure [1]. When teaching, some teachers may not be able to make targeted teaching plans because they do not understand students’ learning of relevant courses in the early stage, which is easy to cause teaching disconnection or repeated teaching. The former increases the learning difficulty and hits the students’ learning enthusiasm, while the latter wastes teaching resources and reduces the learning efficiency. With the increasing popularity of higher education and the increasing scale of students in schools, the educational administration management system of various colleges and universities has also accumulated a large amount of performance data. In the face of these massive data, many teaching managers still stay in the simple operation of adding, deleting, modifying, and checking the data and cannot effectively extract and analyze the useful knowledge and information hidden behind the data [2, 3]. For students, they can more clearly understand their strengths and weaknesses in various qualities and abilities according to the evaluation results, so as to stimulate the enterprising power of their all-round development [4, 5]. For teachers, they can timely grasp the students’ academic status and comprehensive
quality development level from the evaluation results, so as to improve the teaching quality and promote the teaching reform. For education managers, the evaluation results can arouse their attention to the development of students’ abilities other than examination, so as to promote the reform of talent training methods and improve the quality of talent training.

The advantage of Markov model lies in its better prediction effect on random series and data series with high volatility; that is, the grey prediction model is used to reveal the general trend of the development and change of the predicted data series, while the Markov model prediction is used to determine the transition and change law of the data state, which just makes up for the limitations of the grey prediction model [1, 6]. The items in the data set have similar properties and functions, that is, the same or similar importance. The distribution of items in the database is uniform; that is, the frequency of occurrence is similar or similar. However, this is often not the case in the real database. In the eyes of users, the values and meanings of different items are different, and the frequency of different items appearing in the data set is also different. Because the transaction data set usually has the time characteristic, the support and confidence of the same rule are different in different time periods [7, 8]. Markov process predicts that a certain data system may appear in a certain state in the future according to the state of each data in the existing data series and the state transition rule constructed by the data series, thus providing theoretical support for relevant decision management.

The Markov model is a statistical method based on probability theory and mathematical statistics to establish a mathematical model in the random process of things’ development [9]. According to the nature of the Markov model, in a random process, the transition probability from one state to another is only related to the current state, and the impact on the future is all concentrated from one state to another is only related to the current state. Therefore, in order to establish a Markov model, under the guidance of the Markov model, the association patterns in the data set are mined more accurately and effectively.

The overall structure of this paper is divided into five parts.

Section 1 introduces the curriculum association classification model and the background and significance of student achievement prediction and then introduces the main work of this paper. Section 2 mainly introduces the curriculum association classification model and the related work of student achievement prediction at home and abroad. Section 3 introduces the principle and algorithm of the Markov model. Section 4 introduces the implementation of curriculum association classification model and student achievement prediction, as well as the analysis of the experimental part. Section 5 is the summary of the full text.

2. Related Work

2.1. Research Status at Home and Abroad. Fung et al. proposed three machine learning algorithms: linear regression, decision tree, and naive Bayes, combined with student feature screening to classify and predict students’ grades. Because the environment of traditional colleges and universities in China is quite different from that in foreign countries and the available data attributes are different, many domestic scholars use the behavior data in the teaching process of students, such as campus all-in-one card data and check-in data, to predict SAP [15] Ming et al. proposed the management data of about 700 courses in the University of Porto and studied the classification and regression...
found that the performance of decision tree and SVM algorithm was the best in the classification algorithm, and SVM, random forest, and AdaBoost in the regression prediction R2 have the best performance [16]. Zhao et al. proposed to use quantile regression model or decision tree, genetic algorithm, artificial neural network, and other technologies to find external factors related to students’ performance, but these factors only affect rather than determine the floating range of performance to a certain extent [17]. Lera et al. proposed the method of constructing the ability-oriented curriculum system, analyzed the importance of the curriculum according to the curriculum association rules generated by the Apriori algorithm, and then used the system clustering method to classify and process the association rules for the construction of the curriculum system [18]. Ahn and Trogdon put forward data mining methods, which take students' grades in various subjects as the research object, in order to find the hidden curriculum-related rules or patterns. Although these methods are conducive to optimizing the curriculum structure and improving the teaching quality, there are great limitations in the scope of curriculum analysis, and most of them are designed based on educational ideas. They do not fully excavate the useful information and teaching problems embodied in the curriculum association rules, so they cannot provide targeted guidance and early warning for students nor provide substantive teaching suggestions and decisions for the teaching development of colleges and universities [19]. Romanowski et al. put forward an algorithm model of personalized course selection recommendation and applied it to the data of computer courses. From the experimental results, we analyzed the computer courses with mutual promotion effect and then completed the course recommendation according to the needs of different students [20]. Parker et al. proposed that students’ current performance is not only influenced by the recent teacher’s effect but also restricted by various factors such as students’ learning foundation. If the overall foundation is good, their average performance will be high, which cannot be fully attributed to the teacher’s credit, so the current one or two students’ test scores cannot reflect the current teachers’ teaching level [21]. Bowcutt and Caulkins pointed out that foreign scholars had little research on the comprehensive quality evaluation of college students. Most of their research on student evaluation focuses on the evaluation of college students’ academic ability and learning effect. Academic ability refers to a person’s current level of learning and refers to the knowledge, abilities, and skills that students have achieved through learning, as well as the potential for subsequent learning at the existing level [22]. In the comprehensive quality evaluation of college students at the stage of the emergence of Tissenbaum, professional achievement is no longer the only indicator of overall evaluation of students, and researchers attach importance to all-round objective evaluation of students. Therefore, they have formulated corresponding comprehensive quality evaluation methods; adopted quantitative and qualitative methods, combined with computer science knowledge; and established a comprehensive, reasonable, visual, and comprehensive quality evaluation system for college students that meets the needs of today’s times [23]. Klapproth put forward that in the traditional evaluation of college students’ comprehensive quality in China, most researchers regard professional achievement as the only indicator to evaluate students’ comprehensive quality, ignoring the establishment of the concept of quality education, the scientific display of the relationship between courses, and the excavation of the internal relations of “morality, intelligence, physique, beauty and labor” [24].

2.2. Research Status of Curriculum Association Classification Model and Student Achievement Based on the Markov Model. Therefore, based on the Markov model, this paper studies the curriculum association classification model and student achievement prediction, studies the potential relationship between various courses in Colleges and universities, and uses this model and relevant learning prediction methods to mine relevant rules from students’ grades, so as to realize the prediction of related curriculum achievements between lower grades and higher grades, especially early warning for courses that may be difficult. It is suggested that teachers provide targeted guidance to students. Using the principal component analysis method, this paper puts forward nine dimensional core indexes such as comprehensive ability, language expression ability, logical thinking ability, physical quality, computer ability, practical operation ability, political quality, professional quality, and innovation ability and analyzes the performance of students’ individual comprehensive quality in the form of radar chart according to the quantitative value of each dimensional index. In the modeling process, combined with the student demographic data in the data set, the feature engineering is used to analyze the student performance, and the Markov model is used to establish the curriculum association classification model and student performance. Experiments based on the existing data show that the curriculum association classification model can effectively describe the characteristics of students. Based on the Markov model, this paper constructs the student group classification model and compares and analyzes the performance of the comprehensive quality of different student groups according to the classification results, so as to provide reasonable suggestions for promoting the all-round development of students, improving the teaching quality, and improving the quality of talent training.

3. Principle and Algorithm of the Markov Model

The Markov model is named after Russian mathematician A.A. Markov, and its essence is a kind of probability estimation. It regards the time series as a random process. By studying the initial probability of different states and the transition probability between states, it can predict the future state of things; that is, according to the current state of variables, it can predict the possible states in a certain period of time, providing some basis for decision makers. Therefore, it is not sufficient to use the unified minimum support to carry out the curriculum association classification.
model in different periods of the whole data set. Different sub-data sets are divided according to different time periods, and different support thresholds are set at different times to form support vectors. Markov process refers to the process that a system transfers from one state to another. This process has two important characteristics: no aftereffect and stationarity. For aftereffect, that is, in the development of the event, the second result state of the system is only related to the \( n - 1 \) time and has nothing to do with the previous state. Stability means that after a long time, the Markov process gradually tends to a stable state, regardless of the initial state. In our real world, there are many such phenomena; that is, in a certain system, under the known present conditions, the future situation is only related to the present conditions, while the future situation is not directly related to the past history. For example, let \( X_t \) be a random variable representing the system state at the discrete point of time \( t = 1, 2, \ldots \), and this random variable family \( \{X_t\} \) forms a random process. The state of a random process can be finite or infinite. A random process like this is a Markov process; that is, a certain time in the future is not related to any time before the present time but only related to the present time. The adaptive multi-Markov model adaptively determines the values of multiple minimum support and minimum confidence according to the different frequency characteristics of each item in the data set. The core steps of adaptive multi-minimum support association rule algorithm mining course relevance are shown in Figure 1.

One of the purposes of the curriculum association classification model is to find the trend of sequence changing with time and predict the possible value of rule support and confidence in the next time period through the analysis of the trend. Therefore, the analysis and prediction of support vector in the curriculum association classification model is to establish a Markov model and mine the association patterns in the data set more accurately and effectively under the guidance of the Markov model. Prediction is to understand the future with the help of the exploration of the past. Grey prediction is to make a scientific and quantitative prediction of the future state of the system through the processing of the original data sequence and the establishment, learning, discovery, and mastery of the development law of the system. It is mainly used for the fitting and prediction of the eigenvalue of a leading factor of the complex system, so as to reveal the change law of the leading factor and the future development trend. Through feature engineering, the effective features for students’ performance prediction are selected, the "student portrait" feature database is established, and the cart regression prediction based on "student portrait" is carried out, which proves the superiority of the model in predicting students’ performance. Model 2 only uses students’ historical score information. Based on the recommendation system algorithm, it tries not to use students' personal demographic data. According to several methods of Markov model fusion, because there are only two base models and they are regression models, they are more suitable for weighted average fusion. This chapter attempts to carry out weighted average fusion experiments on model 1 and model 2. The Markov model diagram is shown in Figure 2.

Markov model process is the basic model of Markov prediction method: \( x(k + 1) = x(k) \times P \), which predicts that a certain data system may appear in a certain state in the future according to the state of each data in the existing data series and the state transition rule constructed by the data series, and then provides theoretical support for relevant decision management. Among them, \( x(k) \) represents the state vector of the trend analysis and prediction future object in the data sequence at the moment, \( P \) represents the one-step transition probability matrix obtained from the original

---

**Figure 1: Flow chart of adaptive multimimum support Markov algorithm.**

- **Discrete score data set**
- **Calculate the support of each grade data item**
- **Find the value with the second derivative of 0 to determine the minimum support of multiple**
- **Under the constraint of multiple minimum support, the candidate set is generated recursively**
- **Compare the support of each candidate set with the maximum of the minimum support of the project**
- **Find the value of the second derivative of 0 to determine the minimum confidence**
- **Under the constraint of confidence lift mode, the markov model is obtained**

Wireless Communications and Mobile Computing
data sequence, and \( x(k+1) \) represents the state vector of the trend analysis and prediction future object in the data sequence at the moment \( t = k + 1 \).

**Support vector (SV):** the support vector of course association rule \( A \Rightarrow B \) or item set \( A \cup B \) has the following expression:

\[
SV = \left[ s_{(A \cup B)}_1, s_{(A \cup B)}_2, \cdots, s_{(A \cup B)}_n \right], \quad \text{s.t.} \quad s_{(A \cup B)} = f_{(A \cup B)} / M(n \in \{1, 2, \cdots, n\})
\]

(1)

\( f_{(A \cup B)} \) is the item set, \( A \cup B \) in the data subset \( D_i(i \in \{1, 2, \cdots, n\}) \) is the frequency of occurrence in the data subset \( D \) divided by time for the transaction data set, and \( M \) is the number of transactions in \( D \). If the support of item set \( A \cup B \) is \( s \), there is

\[
s = \sum_{i=1}^{n} s_{(A \cup B)}.
\]

(2)

Sometimes, it is more appropriate to use the frequency of item set to express the support. In this way, the support vector of item set is \( SV = [f_1, f_2, \cdots, f_n] \), and the corresponding support can be expressed as

\[
s = \sum_{i=1}^{n} f_i.
\]

(3)

**One-step transition probability matrix:** if the Markov state space \( I = \{1, 2, \cdots\} \), the matrix \( P \) composed of one-step transition probability \( p_{ij} \) is called the one-step transition probability matrix of system state.

\[
P = P(1) = \begin{bmatrix}
P_{11} & P_{12} & \cdots \\
P_{21} & P_{22} & \cdots \\
\vdots & \vdots & \ddots
\end{bmatrix}.
\]

(4)

Obviously, all elements of matrix \( P \) are nonnegative, and the sum of elements in each row is equal to 1.

In the Pullman-Kolmogorov equation, let \( \{X(n), n \in T_1\} \) be a homogeneous Markov chain; then, for any \( \{u, v \in T_1\} \), there is

\[
P_{ij}(u + v) = \sum_{k=1}^{\infty} P_{ij}(u)p_{kj}(v), \quad i, j = 1, 2, \cdots
\]

(5)

The matrix form of \( C - K \) equation is

\[
P(u + v) = P(u)P(v).
\]

(6)

The transition probability of \( n \) step is determined by \( C - K \) equation, and the recurrence relationship: \( P(n) = P(1)P(n-1) = PP(n-1) \), is obtained by making \( u = 1 \) and \( v = n - 1 \) in \( P(u + v) = P(u)P(v) \).

\[
P(n) = P^n.
\]

(7)

This shows that the \( n \) step transition probability of the Markov chain is \( n \) power of one-step transition probability, and the finite dimensional distribution of chain can be
Table 1: Comparison between Chinese women’s predicted results and actual results.

<table>
<thead>
<tr>
<th>Particular year</th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
<th>2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 km race walking</td>
<td>5274</td>
<td>8183</td>
<td>5522</td>
<td>5235</td>
</tr>
<tr>
<td>GM fitting value</td>
<td>5274</td>
<td>5281</td>
<td>5282</td>
<td>5283</td>
</tr>
<tr>
<td>Relative error</td>
<td>0.000</td>
<td>-0.0192</td>
<td>0.0428</td>
<td>-0.0091</td>
</tr>
</tbody>
</table>

Table 2: Comparison between predicted and actual results of Chinese women.

<table>
<thead>
<tr>
<th>Particular year</th>
<th>2018</th>
<th>2019</th>
<th>2020</th>
<th>2021</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 km race walking</td>
<td>5302</td>
<td>5304</td>
<td>5236</td>
<td>5290</td>
</tr>
<tr>
<td>GM fitting value</td>
<td>5286</td>
<td>5287</td>
<td>5288</td>
<td>5291</td>
</tr>
<tr>
<td>Relative error</td>
<td>0.0023</td>
<td>0.0031</td>
<td>-0.0101</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

4. The Realization of Curriculum Association Classification Model and Student Achievement Prediction

4.1. Systematic Design of Curriculum Classification and Student Achievement. The database of the school’s educational administration system stores students’ personal information and students’ achievements. According to these basic data, it can be screened and established to help analyze students’ achievement prediction. Based on the analysis of traditional university database data, this paper explores the relationship between students’ personal characteristics and achievements. Due to the complexity of factors that affect course achievements, some of them, such as students’ subjective learning attitude and objective learning environment, are difficult to control, which leads to the randomness and swing of exam results. The systematic prediction of achievement cannot be limited to the process of reasoning, but the information should be mined from the current academic achievement with the help of quantitative data. Therefore, it is difficult to establish a simple and practical prediction model of examination achievement by conventional methods. This paper attempts to accurately predict the students’ failing grades on the basis of reliable experimental data through the relevant analysis of courses and provide early warning of failing grades. The goal is to learn a linear model from the data set. There is a linear relationship between the predicted results and the attributes of the data set, and each attribute has a corresponding parameter weight. The larger the weight, the greater the influence of the parameter on the results. The modeling of the linear regression model focuses on finding the best parameter to fit the linear relationship. The reconstruction or revision of the curriculum system generally takes one academic year or longer as a cycle, with the participation of disciplines and specialties and key teachers. However, the problems that specific teachers are aware of cannot be solved in the new curriculum system in time and unimpeded, and the phenomenon that teachers both inside and outside the school share the teaching tasks often occurs. Especially, the participants in the planning of teachers outside the school and the teachers in the class are often not the same person, which also brings certain obstacles to the information communication in the curriculum setting.

The advantage of linear regression is fast and simple, which is suitable for the situation of small amount of data and uncomplicated data relationship. The disadvantage is that it is very sensitive to outliers and requires linear separability between attributes and prediction objectives. Due to the short course name, only the course name is segmented,
and the keyword repetition times are less. In this way, the calculated similarity is very low, and the correlation between the opening colleges of the course cannot be further excavated. In order to improve the reliability of course keyword similarity calculation after course name segmentation, this paper does not use direct course name segmentation when dealing with course word segmentation but connects the college name of students who have participated in this course with the course name, and the number of repetitions represents the number of students participating in this course, so as to form the course description text and then segment the word. Finally, the course description text of all courses forms the course description document corpus. Then, the course description text corpus is segmented. According to the characteristics of curriculum score system distribution, an appropriate score discretization method is proposed. Then, aiming at the limitation of single support in curriculum association rule classification and the difficulty of setting threshold parameters, an adaptive multiminimum support association rule classification model is proposed, and a statistical fitting method to adaptively determine the threshold of support and confidence is given. At the same time, valuable curriculum association rules are introduced. In order to eliminate the interference of irrelevant factors, first, make the following assumptions: (a) the proportion of usual grades in each semester remains unchanged, and the difficulty of test papers is roughly the same; (b) the learning ability and usual performance of each student are basically unchanged; (c) the review time that students devote to each course is basically the same; (d) each student is in the same examination environment; (e) students’ grades are real, and there is no influence of cheating and other problems; and (f) the criteria for teachers’ evaluation of papers remain basically unchanged.
However, there are some problems in this way: on the one hand, the artificially set threshold lacks the support of objective basis, and the value is subjective and arbitrary. On the other hand, the different values of support and confidence thresholds have a significant impact on the scale of candidate item sets and frequent item sets and the number of association rules generated in the mining process. In the direction of student achievement prediction, a lot of research shows that it is the most important factor for predicting student achievement. Secondly, the most commonly used attributes are subject achievement and demographic attributes, including gender, age, and family background. Research shows that girls’ learning style and attitude are more positive than boys’. In addition, comparing the attributes with Chinese traditional university characteristics is the attributes of students’ behavior stored in one-card data, so that they can learn more characteristics related to their grades. Many domestic studies have used one-card data to predict students’ grades. Therefore, this research adopts statistical fitting technology to realize the adaptive value of minimum support and minimum confidence, so as to improve the utilization rate of prior knowledge and the accuracy of mining results. Teaching is to instruct scholars to design different teaching methods for students of different levels according to their knowledge and ability level, so as to adapt to the characteristics of current students, and let each student learn something and improve together, thus narrowing the polarization. Therefore, it is suggested that the teaching staff of this major should implement stratified teaching in a targeted way and, at the same time, respect students’ personalities and hobbies and give full play to students’ main role in teaching, so that students can learn actively. This will not
only promote the improvement of students’ academic performance but also enable them to give full play to their individual strengths and finally achieve all-round development.

4.2. Experimental Results and Analysis. The overall trend of the annual best performance of Chinese women’s walking events shows a gradual upward and downward development trend. The specific data in each year increase and decrease. As a traditional prediction model, the original data sequence is required to show exponential regular changes, and its prediction value shows a single tone upward or monotonous downward trend. This prediction state is obviously inconsistent with the tortuous changes of sports performance. Therefore, the long-term performance prediction accuracy is not very high. In view of this, this experiment predicts and analyzes the annual best performance of Chinese women’s race walking events by the Markov model and then effectively solves this problem. Through the change law of its own time series, the performance of Chinese women’s race walking in the future can be predicted according to the principle of the model. The experimental results are shown in Tables 1 and 2.

From Tables 1 and 2, it can be seen that the prediction effect obtained by using the model directly is acceptable except that the prediction error of individual years is relatively large. The average relative error is 1.18% absolute value, and the maximum annual relative error is 4.32%. In order to improve the accuracy of prediction, Markov prediction model is introduced to process the results of the grey prediction model. In order to further improve the accuracy of the model, in the future prediction modeling, this paper first uses the metabolism method in the grey theoretical model to process the original data and then carries on the grey Markov modeling and prediction. If there are many states of data sequence division in the modeling process of the Markov prediction model, the middle value of grey element interval can be taken as the final prediction value to obtain higher prediction accuracy. If there are few divided states, a relatively conservative prediction principle can be adopted, and the lower value of the critical value of the grey element interval can be taken as the prediction value.

This paper takes 45 courses of 2021 students majoring in communication engineering in a college of telecommunications as the research object. First of all, the data is preprocessed, and the samples with missing grades may be deleted due to reasons such as repetition, drop-out, or absence from exams. The scores are divided into four grades, 90–100 corresponds to A, 80–90 corresponds to B, 70–80 corresponds to C, and below 70 corresponds to D. Set the minimum support supmin to 0.1 and the minimum confidence confmin to 0.1. In order to realize the failure warning, the failure score is increased to 70 points, and the top N courses with the highest support are selected as the prediction results. Three experiments are conducted, respectively, and the false alarm rates of Apriori, FP growth, and this algorithm are shown in Figures 3–5, when N changes.

It can be seen from Figures 3–5 that obvious rules can still be seen from the experimental results. When n increases, the false alarm rate increases, and the missed detection first decreases and then increases. The average false alarm rate of the Apriori algorithm is 38.3%, that of FP growth algorithm is 37.2%, and that of this algorithm is 40.5%. When n = 4, the false alarm rate is the lowest, so this value is selected as the number of courses for failing warning. It is assumed that all students’ grades are achieved in an ideal state, but this is not the case. Under the influence of various objective and subjective factors, especially when there is cheating or abnormal performance in the examination, the examination results often fluctuate to some extent, leading to false alarm or missed detection.

In order to realize the failure warning, the failure score line is increased to 80 points, and the first n courses with the highest support are selected as the prediction results. When N changes, the missed detection rates of Apriori and FP growth algorithms and the algorithm in this paper are shown in Figures 6 and 7.
As can be seen from Figures 6 and 7, according to students’ historical scores, the average missed detection rate of predicted future grade points is 48.65%, the average missed detection rate of the Apriori algorithm is 35.5%, and the average missed detection rate of the FP growth algorithm is 43.2%. The average missed detection rate of this algorithm is 37.5%, and only 17 of the 40 grade point interval predictions match the actual interval. In terms of prediction accuracy and practical significance, the method proposed in this paper not only achieves satisfactory results under the condition of higher prediction complexity and uncertainty but also makes full use of the association rules between courses to provide students with early warning and teacher guidance for specific courses, which can effectively reduce the failure rate while reducing the academic burden. The prediction of students’ performance is not the ultimate goal but to remind students of the final learning effect caused by their current learning state and performance and serve as an effective basis for guiding students to continue learning. For the subjects that have provided failure warning, students will certainly invest more time in learning and studying. Therefore, once the prediction model is established, it will be equivalent to failure and improve the false alarm rate to a certain extent.

5. Conclusions

On the basis of correlation analysis of students’ current grades, the research of course association classification model and student achievement prediction algorithm based on the Markov model proposed in this paper can accurately predict the subjects that students may fail in the future and combine spectral clustering algorithm to classify undergraduate courses to form a comprehensive and systematic course association system, which can provide guidance for university teaching management and students’ learning strategies. By using Markov models to give full play to their respective advantages, we can avoid considering the influence of various factors on our grades. At the same time, the accuracy of its performance prediction is higher, and its correlation degree, mean variance ratio, average relative error, and small error probability are all first-class indexes, which have stronger scientific theory support and better practical significance. According to the distribution characteristics of students’ achievement data, the course association classification model adopts the method of dividing the average score as the center and dividing the grade by standard deviation to complete the discrete processing of the achievement. Based on the Markov model, this paper studies the curriculum association classification model and students’ achievement prediction. The average missed rate of predicting future grades according to students’ historical achievements is 48.65%, that of the Apriori algorithm is 35.5%, and that of the FP growth algorithm is 43.2%. The average missed rate of this algorithm is 37.5%, and only 17 of the 40 grade point predictions match the actual interval. Make full use of the association rules between courses to provide students with early warning and teacher guidance for specific courses, which can effectively reduce the failure rate while reducing the academic burden. For example, recommending courses according to students’ interests, analyzing the related factors that affect their grades, and evaluating teachers’ teaching quality are all worthy of further exploration and research. How to reduce false alarm rate and missed detection rate and achieve more accurate course classification results through the Markov model and improving correlation analysis algorithm is the next research direction.

Data Availability

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

Conflicts of Interest

No competing interests exist concerning this study.

References


