Analysis of the Practice Path of the Flipped Classroom Model Assisted by Big Data in English Teaching

Zhan Du¹ and Jie Su²

1Quality Center, Shijiazhuang Institute of Railway Technology, Shijiazhuang 050041, China
2Department of Humanities and Social Sciences, Shijiazhuang Institute of Railway Technology, Shijiazhuang 050041, China

Correspondence should be addressed to Jie Su; sjnhelen@126.com

Received 11 October 2021; Revised 24 October 2021; Accepted 25 October 2021; Published 24 November 2021

Academic Editor: Tongguang Ni

Copyright © 2021 Zhan Du and Jie Su. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper makes a detailed analysis of the integrated mining algorithm, analyzes the characteristics of curriculum big data, and analyzes the existing problems of the current association rule mining algorithm, as well as the defects and deficiencies when applied to the curriculum data. Aiming at the problem of mining the entire data set by the mining algorithm, this topic proposes the idea of using the K-means algorithm for clustering processing and uses the Ball-tree structure on the basis of the original K-means algorithm to improve the efficiency of the algorithm. The data set is separated into several clusters of an appropriate number. In the flipped classroom, the basic knowledge is put before the class for learning, and the further deepening and practical application of language knowledge is completed in the class. Teachers can give timely guidance when encountering unsolvable difficulties so that students’ learning can be more effective. This new teaching model not only strengthens students’ confidence in learning and increases their interest in learning, but also increases the opportunities for students to interact with teachers and classmates in the classroom, allowing them to construct the meaning of knowledge in the fun of interactive communication. The classroom has become relaxed, lively, and attractive, and students’ sense of autonomy, self-learning ability, and collaborative inquiry ability have also been unknowingly improved. Among the main factors, the willingness to flip, emotional state, leadership role, and online learning input have a significant positive impact on collaborative learning performance, and the sense of competition has a partial negative impact on collaborative learning performance, of which positively affecting individual knowledge mastery. Among the nonmain factors, the degree of difficulty of the course, teacher-student interaction, teacher motivation, and evaluation mechanism have a significant positive impact on collaborative learning performance. Classroom assistive technology has a partial negative impact on collaborative learning performance. From the perspective of group performance, group-level performance considerations such as the quality of group conversations and the degree of group knowledge sharing in collaborative learning performance are more affected by nonsubject factors.

1. Introduction

With the integration and development of computer and communication technology, network infrastructure is becoming more and more perfect, equipment terminals are rich and diverse, and network applications are more extensive; the Internet has increasingly penetrated into people’s social life, opening a new era [1]. Due to the characteristics of the Internet and computer digital processing, all behaviors of people in the process of popularizing the use of the Internet are transformed into data information [2]. These data and information are collected and stored into digital records through terminals, networks, and servers, forming a big data environment in which everything is digital and everything can be quantified. The real material world in the Internet age is transformed into the real data world. The comprehensive campus network infrastructure, extensive application of information systems, and the popular use of new technologies, new equipment, and new applications by teachers and students have formed a mature campus big data environment with good conditions for big data applications [3]. In the context of the era of big data, flipped classrooms were born and developed rapidly as a new teaching model. In recent years, the research and application
of this model have been very hot in foreign countries [4]. Many domestic universities have applied this model to the teaching of different subjects and have gained successful experience [5]. The characteristic of flipped classroom is that it reverses the sequence of learning knowledge in class and completing homework after class in traditional teaching [6].

In traditional teaching, students acquire new knowledge in the classroom and consolidate knowledge by completing homework after class, while students in flipped classrooms acquire knowledge by watching videos made by the teacher before class and then resolve questions with teachers and classmates in class [2]. The introduction of flipped classrooms into college teaching will help promote the innovation and exploration of college teaching models and promote the transformation of college teaching models to informatization, high efficiency, individualization, and humanism. Related theories provide new ideas and theoretical references for the future application of flipped classrooms in college classroom teaching [7]. Flipped classroom teaching can fully stimulate students’ potential for independent and active learning. Before the class, students can learn independently by watching microvideos or other learning materials anytime and anywhere, and control the pace of learning by themselves. In this process, students will gradually cultivate their independent learning ability and sense of responsibility for learning. In the classroom, there are learning activities such as problem discussion, group cooperation, and results display. In this process, students' interpersonal communication and cooperation ability, high-order thinking ability and independent learning ability are gradually cultivated [8]. Through case studies on the use of flipped classrooms in colleges and universities, it analyzes and summarizes the teaching ideas, teaching strategies, and teaching principles implemented by flipped classrooms in colleges and universities. The construction of a classroom teaching environment that respects the individuality of students in teaching provides a reference for operation.

According to the existing association rule mining algorithm, its advantages and disadvantages are analyzed and ideas for improvement are put forward. We analyze the characteristics of the curriculum data set, and we analyze the reasons for the inapplicability of the current algorithm, propose an integrated idea of combining multiple algorithms, analyze its advantages, and verify the experiment according to the above idea as well as realize the proposed integrated algorithm. We compare the efficiency gap with the previous algorithm. In the flipped classroom, students start collaborative learning in the classroom based on the knowledge and background of the relevant courses. This is also a point that is special in traditional classrooms. The source of student knowledge and experience is the flipped teachers on the online learning platform of “Chinese College Students MOOC” online learning platform. The learning tasks and learning resources released by teachers in Yu Classroom before class, the weekly course online test, and the knowledge review link in the classroom all emphasize the importance of learners’ online self-study before class. At the same time, differences in the length and frequency of learners watching videos and completing exercises, as well as classroom test scores, also indicate differences in students’ online learning input. Students with a higher level of input are aware of the knowledge reserves and strategies required to complete the task, and it is relatively easy to make tacit knowledge explicit in each link of collaborative learning. Therefore, the degree of learner’s online learning input has a significant positive effect on both the individual learning performance of collaborative learning performance or the group learning performance.

2. Related Work

Flipped classrooms in Western countries (such as the United States, Australia, and the United Kingdom) have begun to be widely used in graduate and undergraduate teaching, and the teaching subjects and research fields involved include medical care, nursing, biochemistry, and materials science [9]. Their research on flipped classroom teaching methods is no longer limited to theoretical research and more turns to teaching practice. They have even used flipped classrooms to devote themselves to applications in other fields, such as using flipped classrooms to develop students’ learning skills and self-reflection abilities.

The American education consulting company Class Window released a survey report that revealed the application value of flipped classrooms [10]. A total of 450 teachers participated in the survey. Among them, 88% of teachers said that their satisfaction had improved, 46% of teachers reported a significant improvement, and 99% of teachers said that they will continue to use flipped classroom teaching next year. Columbia University conducted a survey of 203 students from the United States [11]. The survey found that 80% of the students said that they have more frequent and active interactions with teachers and peers in flipped classrooms, and there are more course materials they need to read in flipped classrooms. There are also more opportunities to learn and practice according to their own time, and more opportunities to demonstrate learning results. Students are more willing to regard learning as an active and active process. Speak Up Survey surveyed 43,000 K-12 students, parents, and administrators [12]. Through a survey of managers, it was found that 25% of managers indicated that flipped classrooms are the most important change in teaching, which surpassed educational games. It can be seen that the flipped classroom teaching method has generally been recognized by foreign students, parents, and managers and has high feasibility.

Domestic scholars mainly conduct research on the teaching strategy, teaching process, and subject and object of the flipped classroom [13]. Based on the flipped classroom under the conditions of informatization, the teaching strategy is more flexible, the teaching process focuses on problem-solving, and students have more decision-making power [14]. Although the research on flipped classroom content is relatively broad, the research depth is not enough. Furthermore, a mature theory has not yet been formed, and many schools, especially basic education schools, often just try to flip classroom teaching [15]. It is difficult to do in-depth research only by completing a few demonstration
classes on a certain topic. By trying several flipped classroom teachings with enthusiasm, the experience gained may not withstand the test of practice. Adhering to normalized flipped classroom teaching will greatly promote the research of flipped classroom teaching mode, and the practical experience gained is also worth learning. In the era of big data, the use of big data thinking to guide and optimize normalized flipped classroom teaching makes classroom teaching a new look. A middle school in Chengdu insists on normalized flipped classroom teaching, and it collects student learning process data, uses this data to guide and optimize flipped classroom teaching, and has gained some experience [16].

The University of Wisconsin-Madison used eTeach software to synthesize lecture videos and courseware simultaneously and successfully carried out the flipped classroom teaching reform [17]. Multimedia teaching technology creates necessary conditions for flipped learning and emphasizes the technical characteristics of flipped teaching mode [18–20]. In a teaching practice, related scholars helped students who were absent from the classroom make up lessons by recording videos of applying PPT in classroom teaching and then uploading them to the Internet for students to download and learn by themselves [21–23]. Through this teaching attempt, they found that the students’ understanding and internalization of the video requires the guidance and help of the teacher, so they reversed the order of the classroom. Through this teaching experiment, the idea of flipped classroom teaching has been further disseminated in American primary and secondary schools [24].

3. Method

3.1. Characteristic Analysis of Course Data. English flipped classroom teaching has provided a good path and opportunity for our educational scientific research work. It has constituted a cross-ethnic, cross-cultural, cross-professional, and cross-social communication platform. By studying these educational data, we can discover new information hidden in the data. The English cloud education platform network assisted by big data is shown in Figure 1.

Curriculum big data, as a relatively common type of data in education big data, has the following characteristics:

(1) Wide range of data sources: the big data of the course has a wide range of sources and a large scale. English flipped classroom teaching has many advantages, such as unique convenience and high efficiency. More and more students choose English flipped classroom teaching to learn and supplement their own knowledge reserves. There are also more and more educators and schools starting to use the English flipped classroom teaching platform for online teaching. More and more managers begin to use the English flipped classroom teaching management platform for online data management. Therefore, the English flipped classroom teaching platform has a large amount of student selection data, school curriculum data, educator teaching data, and a large number of comment data, etc. The source of curriculum big data is very wide, and it is no longer limited to schools or education bureaus.

(2) Higher potential value: curriculum big data does not only come from students, so the analysis and application of curriculum big data is no longer limited to students who need to learn but can be expanded to more lecturers and education managers. Therefore, curriculum big data has a higher potential value.

(3) More relevant: there may be certain relationships between data, and the values of certain variables may affect each other, but this correlation is generally weak and usually not causal. Compared with other types of education data, curriculum big data has greater opportunities for each feature variable to influence each other and stronger correlation. For example, people who choose to learn front-end technology may be more willing to learn front-end-related technology; students who learn testing techniques are more willing to learn such as automated testing, black-and-white box testing courses, and so on.

(4) Larger amount of data: generally speaking, curriculum big data has the characteristics of large quantity and high dimensionality. Some courses will have multiple characteristics, and some courses will also belong to multiple course categories at the same time. For example, in the course selection data of students, each course of selection data will have its own category. These categories are diverse, and each student’s course selection tendency is also different.

3.2. Problems in Association Rule Mining Algorithms and Improvement Ideas. Association rule mining algorithm is a mining algorithm that discovers frequent itemsets. It is an important subject in data mining. Its main function is to find hidden relationships between specific data sets and find valuable information from it. The process of association rule mining is roughly divided into two stages. The first part requires finding frequent itemsets in the original data set. The second part dig out association rules from these frequent itemsets through pruning and sorting.

Common association rule mining algorithms are mainly based on the corresponding association rules for the entire data set to mine. When the data set is very large, it will often produce massive amounts of redundant data and similar association rule mining results, which will also cause the final mining results of the association mining algorithm to be too accurate. For example, the Apriori algorithm finds frequent itemsets by frequently accessing the database. When faced with massive data sets, this frequent access to the database operation method will reduce the efficiency of the algorithm as the data set expands.

Although the FP-growth algorithm avoids a large number of database accesses by establishing the FP-tree, a large amount of data will also cause the FP-tree to be too
large. Therefore, this will greatly affect the overall efficiency of the algorithm and the accuracy of the mining results, which cannot be well adapted to the characteristics of curriculum big data.

Top-$K$ is an association rule mining algorithm designed by Fournier-Viger. Compared with the traditional association rule mining algorithm, the advantage of this algorithm is that the user can specify the number of rules to be generated $K$, avoid generating too many rules, so the efficiency is higher than the traditional association rule mining algorithm.

In this topic, improvements are made on the basis of Top-$K$ association rule mining algorithm, and correlation coefficients are added on the basis of Top-$K$ algorithm to improve the accuracy of the algorithm and further improve the efficiency of the algorithm. In view of the large number of data sets and high-feature dimensions, the $K$-means algorithm is improved to adapt to the situation of high-dimensional data. The $K$-means algorithm can refine the data set into multiple small data sets, and each data set represents a category. With the collection of the same feature vector, the accuracy and efficiency of the mining results are improved through the idea of classification and mining. However, the traditional $K$-means algorithm is not very adaptable to high-dimensional data, and the KNN algorithm uses KD-tree and Ball-tree to optimize the processing methods of high-dimensional data and compares the similarities between the $K$-means algorithm and the KNN algorithm. Therefore, this topic uses the Ball-tree structure to optimize.

### 3.3. Analysis and Improvement of $K$-Means Algorithm

Cluster analysis is to group according to the information involved in the data contained in the description of the objects or the relationship between them. There are many kinds of clustering algorithms, but not all algorithms are suitable for this experiment. Considering the efficiency of the algorithm, we prefer to choose the clustering algorithm with high efficiency. One of the biggest advantages of clustering methods based on definition partitions is their fast convergence rate and convergence speed, and for some large data sets, they are simple and efficient, with low time and complexity. The $K$-means algorithm is a typical clustering algorithm based on the distance between targets. The distance between targets is used as an evaluation index for the degree of similarity between targets. That is, if the distance between two targets is closer, the possibility of them being the same cluster is greater.

First, we need to determine the size $n$ of a known selected sample data set, and the feature dimension of each selected sample data set is $m$, recorded as

$$
\xi = \{ X_{i1}, X_{i2}, \ldots, X_{im} \}, \quad i = 0, 1, 2, \ldots, n - 1.
$$

![Figure 1: English cloud education platform network assisted by big data.](image-url)
Then, we need to set the required number of clusters $K$. And we need to randomly select the final cluster center point that is the same as the set number of clusters $k$ in a data set.

$$\zeta = \left[ \zeta_{ij} \ zeta_{2j} \ldots \ zeta_{mj} \right], \quad j = 0, 1, 2, \ldots, k - 1.$$  \hspace{1cm} (2)

For each point in the data set, we calculate the Euclidean distance between each centroid and each point. According to the criterion of the closest distance, they are divided into the corresponding classes closest to the cluster center (most similar), calculated as follows:

$$\xi(i, j) = \sqrt{\left( \zeta_{ij} - x_{11} \right)^2 + \left( \zeta_{ij} - x_{21} \right)^2 + \ldots + \left( \zeta_{ij} - x_{mj} \right)^2}.$$  \hspace{1cm} (3)

After putting all the data together, there will be $K$ sets. If the distance between the newly calculated centroid and the original centroid is much smaller than the expected threshold, then we can be sure that the sample clustering does not reach the expected classification, and the algorithm may be terminated. If the newly appearing center of mass differs greatly from the original center of mass, it will change, and the above steps must be repeated again.

It can be clearly seen that the $K$-means algorithm achieves the goal of sample clustering through multiple iterative calculations. In each iteration of the calculation process, it is necessary to reset and calculate the center point of each sample cluster, and replace the previous sample center point, which is used as the sample center point for the next iteration. The final sample center point can be considered as the final cluster center.

The $K$ value of the $K$-means algorithm affects the result of clustering. The original $K$-means algorithm uses a method of randomly selecting $K$ values. This method is extremely unstable and cannot get stable and good results. The larger the $K$ value of the cluster, the more detailed analysis and division of the data. When the $K$ value is less than the number of clusters, since the increase in the $K$ value will greatly increase the clustering degree of each cluster, the SSE (sum of square error) will also decrease. When the $K$ value is greater than the number of clusters, increasing the value of $K$ will increase the degree of polymerization, so the decline of SSE will also decrease sharply and then gradually flatten out. In other words, the $K$ value corresponding to the inflection point of the relationship graph between SSE and $K$ is the optimal number of clusters in the data set. This experiment uses SSE to determine the evaluation value. SSE is defined as follows:

$$\text{SSE} = \prod_{i=1}^{k} \prod_{p \in c_i} (m_i - p)^2.$$  \hspace{1cm} (4)

Among them, $C_i$ represents the $i$th cluster, $p$ represents the sample point in $C_i$, $m_i$ represents the center point of $C_i$, and SSE is the clustering error of all samples, representing the quality of the clustering effect.

The $K$ value determined by the elbow method can basically be determined as the optimal number of clusters in the current data set. This is more stable and accurate than the original $K$-means algorithm that randomly selects the $K$ value, which effectively avoids the randomness of the $K$-means algorithm itself and improves the stability of $K$-means.

The KD-tree structure and the Ball-tree structure are used to improve the efficiency of the KNN algorithm and the processing power on high-dimensional data, and the $K$-means algorithm and the KNN algorithm also have the same ideas. Therefore, the KD-tree and Ball-tree structures can also be used in the $K$-means algorithm to improve the efficiency of the $K$-means algorithm value and the processing capacity of high-dimensional data.

If the $K$-nearest neighbors in the data set of a sample belong to a certain category, the sample also belongs to this category. The basic implementation of KNN is linear scanning. At this time, the distance between the current instance of the input and each training instance is calculated. When a training set is very large, the calculation duration is very long and time-consuming, and the efficiency of this method is very low.

The KD-tree model is a high-dimensional indexed tree data structure, which is often used to search the data in space. However, the $K$-means algorithm is similar to the KNN algorithm and has similarities in the idea of the algorithm; that is, it is necessary to find a point on the data set through this point to find the closest point to it. Therefore, KD-tree can also be used in the $K$-means algorithm.

Ball-tree is analogous to KD-tree’s construction idea. KD-tree uses hyperrectangles for regional division, while Ball-tree uses hyperspheres for regional division. Although the construction process will take more time than KD-tree, performance on high-dimensional data sets will be more efficient. In the Ball-tree construction process, $T$ represents the selected data set, and $S$ represents the current data field. At the beginning of the creation process, the data set $T$ and the data domain $S$ need to be passed in. When there is only 1 data point in the data set, return to that node directly. Otherwise, first, we use the midpoint of all points as the center of the circle and use the distance from the midpoint to the farthest point as the radius to construct the initial circle, and we find the two points node 1 and node 2 with the furthest distance in the initial circle. If they are close to node 1, they are classified as clusters of node 1, and if they are close to node 2, they are classified as clusters of node 2, and then recursively construct small circles. Next, recursively we construct a smaller minimum circle from the solved current minimum circle, until there is only one node left, and it is stored as a leaf node. The structure and search process of Ball-tree are shown in Figure 2.

3.4. Rule Generation and Pruning. The original Top-$K$ association rule mining algorithm is based on the traditional “support-confidence” model, but the traditional “support-confidence” model can only find strong association rules in frequent item sets but cannot filter out strong associations. For rule $A \rightarrow B$ or $B \rightarrow A$, the correlation coefficient is defined as
The correlation coefficient \( \theta \) can reflect how much influence "the appearance of item set \( A \)" has on the appearance probability of item set \( B \). In practical applications, there are many problems in both positive and negative correlations that deserve our close attention. When \( \theta \) is greater than 1, it indicates that the former term has a positive promotion effect on the latter term, and the larger the value, the greater the promotion effect. On the contrary, it is an inhibitory effect. When \( \theta \) is 1, it can be seen from the formula that the occurrence of \( X \) has no effect on the occurrence of \( Y \), so we need not consider the case where \( \theta \) is 1.

\[
\theta(A, B) = \frac{p(A \cap B)}{p(A) \cdot p(B)}
\]

Therefore, on the basis of support and confidence, the correlation coefficient can only be used to determine the correlation between an item set. Correlation cannot be equal to causality, so the correlation coefficients of the two rules \( A \rightarrow B \) and \( B \rightarrow A \) are the same:

\[
\theta(A, B) = \begin{cases} 
> 0 & \text{positive correlation,} \\
= 0 & \text{independence,} \\
< 0 & \text{negative correlation.}
\end{cases}
\]

3.5. Construction of Teaching Mode. Based on the theory of constructivism and blended learning theory, combined with the subject characteristics of English courses and their own teaching experience, this paper proposes an English teaching model based on flipped classrooms, as shown in Figure 3.

Starting from the two dimensions of teacher activities and student activities, this model divides the teaching process into three stages in chronological order: pre-class, in-class, and after-class, showing the relationship between the various links in the teaching process. The pre-class stage of the flipped classroom is mainly to realize the transfer of knowledge, which is mainly composed of the production and preparation of the teacher’s teaching resources and the independent learning of the students. The in-class stage is mainly to realize the internalization of knowledge, and the task of English subjects is to complete the use of language. Therefore, in the classroom teaching stage, it aims to realize the application of language knowledge and divides teacher-student activities into answering questions and supplementing. Correction, creation of situations, group collaboration, display and communication, and evaluation feedback are four links, and students are included in the main body of evaluation. Teacher evaluation is combined with student self-evaluation and group evaluation to realize the diversification of evaluation subjects. After class, teachers should reflect on teaching and provide support and guidance for students’ review activities to help them complete knowledge remediation or knowledge expansion.

4. Results and Discussion

4.1. Questionnaire Design and Test Results. This research has compiled a “Questionnaire on Collaborative Learning of Students in Flipped Classrooms in Colleges and Universities.” The questionnaire is divided into two parts. The first part is mainly about the basic personal information of learners, including gender, grade, subject category, and the length of learning experience in the flipped classroom. It is not only to understand the relevant learning background of the learner, but also to prepare for the differences in these aspects of the influence of different factors in the subsequent data analysis. The second part is the main part of the questionnaire, which mainly reflects the collaborative learning performance of students in flipped classrooms and its influencing factors. Among them, the design of collaborative learning performance items is mainly based on the measurement indicators of collaborative learning performance in the flipped classroom in the previous chapter; the
item design of related influencing factors is mainly based on the 13 influencing factors determined by the learner as the main dividing standard. The questionnaire items are based on the existing research thoughts, written records of classroom observations with the group, and student interviews and exchanges as the standard, and they are discussed with the instructor of the flipped classroom in the class and conceived, designed, and modified. The items in the questionnaire are all single-choice questions. Among them, there are 4 items in the first part of the questionnaire and 35 items in the main part of the questionnaire.

Before forming a formal questionnaire, first, we conduct a trial test. The test subjects of this questionnaire are all students in the class, a total of 50 people. The questionnaire was distributed through the form of online questionnaires, and 50 questionnaires were actually returned. At the same time, the reliability of the test results is analyzed, and the commonly used measurement method is the Cronbach α coefficient method. If the α coefficient is above 0.9, the reliability is very good; if it is above 0.8, the reliability is good; if it is above 0.7, the reliability is good; if it is above 0.6, the reliability can barely be accepted; if it is above 0.5, it indicates that the scale needs major revisions; and if the α coefficient is below 0.5, this indicates that it should be abandoned. The internal consistency of α coefficient in the reliability test result is 0.956, indicating that the overall internal reliability of the questionnaire is ideal. Finally, the test result of the correlation coefficient between the item and the total score is shown in Figure 4.

4.2. Reliability and Validity Test Results of the Questionnaire. Most of the items in the questionnaire of this research are compiled based on literature review, classroom observations, and student interviews. There is no overlap between the formal questionnaire and the test questionnaire. The Cronbach alpha coefficient is 0.953 in total dimensions, indicating that the overall reliability is ideal. The test coefficient of influencing factors is shown in Figure 5. The lowest value of α is 0.65, which is above 0.6 and is acceptable, so the questionnaire reliability is good.

In questionnaire analysis, the validity test is also indispensable. The higher the validity, the more it can reflect the function and purpose of the question to be researched by the questionnaire. Questionnaire validity testing generally includes content validity analysis and structural validity analysis. From the perspective of content validity, the questionnaire in this study is the result of comprehensive review of relevant literature, classroom and group observations, and student interview investigation and analysis. Coordinating the frontline to flip the opinions of teachers, after trial testing and modification, based on the questionnaire formed by enriching the preparation work, it can be trusted in terms of content validity. The structural validity adopts the factor analysis method. First, the KMO and Bartlett sphere test of the data are carried out, and then according to the factor load and cumulative contribution rate, it is considered whether it is suitable for factor analysis. The closer the KMO value (0 < KMO < 1) to 1, the more common factors among variables and the more suitable for factor analysis. According to the data statistical analysis standard, the KMO value must be above 0.6 before the factor analysis can be done reluctantly, and it can be discarded if it is lower than 0.6. The KMO value is 0.822, which meets the index requirements. At the same time, the approximate chi-square of Bartlett’s sphere test is 589.323 ($P = 0.000 < 0.050$), reaching a significant level. There is a correlation between various influencing factors, which is suitable for factor analysis.
4.3. Descriptive Statistical Analysis of Data. In order to grasp the survey questionnaire data as a whole, descriptive statistics are made on the collaborative learning performance and influencing factor variables in the questionnaire. The basic personal information of the students was learned from 62 valid questionnaires. Among them, 54 boys and 8 girls were first-year university students. Investigators are the first time to come into contact with flipped classroom teaching. The questionnaire uses the Likert five-point scale, and 3 is divided into the middle value. As shown in Table 1, the average value of the observed indicators of collaborative learning performance is greater than 3, which indicates that in flipped classrooms in colleges and universities, the collaborative learning status of students is generally good. Among them, the average value of “individual learning attitude,” “group collaborative learning atmosphere,” and “group conversation quality” are above 4 points, which is also consistent with the data in the classroom observation record table. There are specific behavioral performances in other aspects, which also reflect the positive attitude of students to the assignment of tasks in collaborative learning.

4.4. Main Effect Test of Data. Using the method of multiple linear regression to determine the mathematical relationship between variables, we determine the path coefficients between variables, highlight which independent variables have significant effects and which ones are not significant, and analyze the effects of main factors on collaborative learning performance. The regression results of influencing factors on individual collaborative learning interest are shown in Figure 6. It can be seen that there is an extremely significant causal relationship between “emotional state,” “intention to flip” and individual collaborative learning interest. Students like to learn by video before class first, and then they can deepen and consolidate their knowledge points by solving problems in class. They recognize and accept the teaching form of flipped classroom, and naturally they will accept collaborative learning as a unique learning method in flipped classroom. Through group collaboration, it not only compares and learns other people’s problem-solving ideas, but also promotes the relationship between classmates, realizes the effect of group collaboration on course learning, and further stimulates students’ interest in collaborative learning.

As shown in Figure 7, with the individual learning attitude as the dependent variable, the influencing factors “emotional state,” “online learning input,” and “teacher-student interaction” enter the regression equation, and the regression results of the influencing factors on the individual learning attitude are obtained. The regression results show that there is a significant causal relationship between “emotional state,” “online learning input,” and “teacher-student interaction” and individual learning attitudes.

4.5. Data’s Moderating Effect Test. Through the main effect test on the data, the main factors affecting the performance of collaborative learning and their relationship are found, and the path coefficient is obtained. Some research hypotheses have also been verified. Among them, the nonsubject factors “grouping strategy” and “collaborative learning activity design” have no significant influence in the results of this research. Part of the reason is that the test subjects, whether it is a trial test or a formal survey, the students in these three classes only tried the grouping strategy of random grouping and lacked the feeling and evaluation of other grouping methods. Moreover, it was learned in the interviews that most students approve of this random grouping method, so there is no difference in opinions on the issue of “grouping strategy.” At the same time, the main effect test results did not indicate the significance level of the reliability test of the test questionnaire.

![Figure 4: Reliability test of the test questionnaire.](image)
Table 1: Descriptive statistical results of collaborative learning performance.

<table>
<thead>
<tr>
<th>Project</th>
<th>Standard deviation</th>
<th>Mean</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Team task completion quality</td>
<td>0.90</td>
<td>3.50</td>
<td>0.02</td>
</tr>
<tr>
<td>Individual performance</td>
<td>0.91</td>
<td>3.42</td>
<td>0.01</td>
</tr>
<tr>
<td>Group knowledge sharing</td>
<td>0.92</td>
<td>3.31</td>
<td>0.10</td>
</tr>
<tr>
<td>Group conversation quality</td>
<td>0.85</td>
<td>4.11</td>
<td>0.11</td>
</tr>
<tr>
<td>Individual task management ability</td>
<td>0.88</td>
<td>3.80</td>
<td>0.12</td>
</tr>
<tr>
<td>Individual cooperation ability</td>
<td>0.85</td>
<td>3.78</td>
<td>0.14</td>
</tr>
<tr>
<td>Individual expression ability</td>
<td>0.92</td>
<td>3.91</td>
<td>0.15</td>
</tr>
<tr>
<td>Individual knowledge</td>
<td>1.10</td>
<td>3.62</td>
<td>0.17</td>
</tr>
<tr>
<td>Individual interest in collaborative learning</td>
<td>1.01</td>
<td>3.55</td>
<td>0.10</td>
</tr>
<tr>
<td>Group collaborative learning atmosphere</td>
<td>0.92</td>
<td>4.00</td>
<td>0.16</td>
</tr>
<tr>
<td>Individual learning attitude</td>
<td>0.89</td>
<td>4.03</td>
<td>0.17</td>
</tr>
</tbody>
</table>

Figure 5: Reliability test of influencing factors.

Figure 6: The regression of influencing factors on individual collaborative learning interest.
“collaborative learning activity design.” It can only show that this factor does not directly affect the collaborative learning performance, but it may change the effect of other factors on the collaborative learning performance to a certain extent.

Moreover, different frontline flipped teachers have different considerations and ideas for the design of learning activities. Due to changes in the teaching objectives and content of each lesson, the design of collaborative learning activities in the classroom is also constantly being adjusted. Therefore, after the main effect analysis of the data, the adjustment effect is analyzed, and the influence of the “collaborative learning activity design” as the adjustment variable is tested, and the path model is improved.

The first step of the moderating effect test is to add the interaction term between the independent variable and the moderating variable as a new predictor variable and incorporate it into the subsequent moderating effect analysis. In order to accurately determine the interaction items, this research first analyzes the correlation between the factors. Here, we mainly examine the correlation between the “collaborative learning activity design” and the main factors. Using SPSS19.0 software to use Pearson’s correlation analysis on the data, the relevant results of the respective influencing factors of the main factors and the “collaborative learning activity design” are shown in Figure 8. The significance levels of “collaborative learning activity design”
and “flipping willingness,” “emotional state,” “leadership role,” “online learning input,” and “knowledge forgetting” are significantly correlated at the bilateral level. There is a significant positive correlation with turnover willingness, emotional state, leadership role, online learning input, and knowledge forgetting. The design of collaborative learning activities has an obvious moderating effect on collaborative learning performance.

The moderating effect shows that a suitable and high-quality collaborative learning activity design will “magnify” or “accelerate” the influence of the main factor on the performance of collaborative learning. Under the regulation of the collaborative learning activity design, the higher the learner’s willingness to flip, the stronger the individual task management ability and individual expression ability; and the higher the knowledge forgetting, the lower the memory retention, which will lead to lower individual expression ability. The group dynamics theory puts forward that the environment is one of the key group constituent elements, which refers to all the physical environment, social environment, and even individual psychological environment that may have an impact on the group. Therefore, in addition to the physical environment such as the flipped classroom, teachers should also help students construct problem situations through the design of collaborative learning activities, create a social environment that is conducive to learners’ collaboration and interaction, and promote effective learning of students. Therefore, the moderating effect of collaborative learning activity design is the influence of environment as an external factor on collaborative learning performance.

5. Conclusion

This article analyzes the characteristics of curriculum big data and finds that, as a common educational big data, it can better reflect the characteristics of education big data in the teaching work of students. And according to the characteristics of the big data of the course, the deficiencies of the current popular association rule algorithms are analyzed. Through the analysis, it is found that the current association rule mining algorithm has the problem of the target data set being too large and the accuracy of the mining results. Through the analysis, it is concluded that the data set can be clustered through the K-means algorithm to achieve the purpose of reducing the data set, and then the Ball-tree structure in the KNN algorithm is analyzed. After analysis, it is proposed that it can be used in the K-means algorithm Ball-tree structure to improve efficiency. Any high-quality classroom teaching will not lack teacher-student interaction, and as classroom teaching becomes more and more intelligent, teacher-student interaction is not only a simple human-to-human interaction, but also an interaction based on technical support. Regardless of its form, its purpose is to promote teaching and learning. Teachers’ questions in flipped classrooms are more complex and hierarchical. Students’ behaviors after asking questions last longer and the quality of learning is higher. Teacher-student interaction in flipped classrooms improves the quality of the classroom. At the same time, teachers’ speech and behavior feedback give students the feeling of being paid attention to. Appropriate participation of teachers in group collaboration promotes deeper discussions. The essence of teacher motivation is the concrete manifestation of teacher’s speech behavior in teacher-student interaction. Verbal motivation can positively strengthen learners’ learning motivation. At the same time, it can encourage and praise the quality of group results and students’ performance in answering questions, which is conducive to comparison and reflection between groups and students. Therefore, teacher-student interaction and timely and appropriate verbal stimulation are a powerful guarantee for good classroom collaborative learning performance. The starting point of evaluation is to test, and its purpose is to guide and motivate students through learning evaluation. The learning activities in the flipped classroom basically revolve around collaborative tasks, so the evaluation objects, methods, content, etc., are closely linked to each link of the learning activities. The group or individual score is a comprehensive result of objective score, teacher evaluation, and individual or group mutual evaluation of students. Therefore, members listen to the opinions of others and reflect on themselves, while also paying attention to the results of other groups and the performance of collaboration. If the teacher emphasizes the attitudes, methods, and skills of member collaboration in the evaluation content, it will be more conducive to improving the collaborative learning performance of the classroom.

Data Availability

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

Conflicts of Interest

All the authors declare that they do not have any possible conflicts of interest regarding the publication of this study.

References


