

Research Article

System Evaluation and Management of College Students' Physical Exercise Behavior Stages Integrating Bayesian Association Rules Data Mining Algorithm

Juan Shi 

Xi'an Fanyi University, Xi'an 710105, China

Correspondence should be addressed to Juan Shi; shijuan518@xafy.edu.cn

Received 10 November 2021; Revised 7 December 2021; Accepted 20 December 2021; Published 28 May 2022

Academic Editor: Qiangyi Li

Copyright © 2022 Juan Shi. 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.

The form of association rules is simple, and it is efficient and convenient to apply. However, because association rules cannot express the connection between different rules, in some more complex application fields, when it is necessary to comprehensively consider the impact of multiple factors on the results, the application of rules is more difficult. In the process of reasoning about the node state, the influence of various factors (parent nodes) can be comprehensively considered. This paper proposes a Bayesian network-based association rule representation method. After mining the association rules from the data, through structural learning and conditional probability table learning, the original rules are finally used as Bayesian nets. This effectively expands the application of association rules. The experimental results show that after using MapReduce parallelization, the improved algorithm can not only process larger-scale data sets, but also save a lot of running time. The correlation between physical exercise behavior evaluation management, dimensional exercise motivation, exercise method, and college students' school adaptability reached a significance of 0.001, and the correlation between dimension exercise time and college students' school adaptability reached a significance of 0.05. The correlation between physical exercise behavior evaluation management and school adaptability in terms of interpersonal relationship adaptation, learning adaptation, campus adaptation, career adaptation, emotional adaptation, self-adjustment, and satisfaction reached significant levels.

1. Introduction

Pearl proposed Bayesian networks based on probability theory and graph theory and successfully applied to expert systems [1]. Because Bayesian networks have a solid theoretical foundation, rigorous inference calculation process, clear semantics, and other characteristics, and their intelligibility is very strong, it has naturally become a very active area for many researchers [2]. Significant results have been achieved in basic theory, standardization, reasoning learning and application areas. Because of the numerous knowledge fields involved and the wide range of applications, data mining has become one of the most focused areas of researchers and business organizations in the past ten years [3]. There are many data mining methods, among which the more typical ones are association analysis, sequence pattern analysis, classification analysis, cluster analysis, and so on. Data mining tools can perform

innovative analysis of historical information and predict future trends and behaviors so as to well support people's decision-making. Data mining technology has been continuously used in various new fields since its birth, and it has been widely used in commerce, banking, insurance, securities, telecommunications, biotechnology, and so on.

Mastering the physical exercise behavior of college students and the physical culture conditions of college campuses can explore the influence factors of college students' formation of final exercise habits and physical sports concepts [4]. This is conducive to the realization of the national policy implementation goals of national fitness and healthy China, and it is also conducive to the pertinence of colleges and universities to promote the development of college students' healthy behavior intervention measures; it is conducive to promoting the quality of life of college students, and it is also conducive to promoting college

students' development of comprehensive quality. The related exploration and analysis of college students' physical exercise behavior will help continue to enrich the domestic research results in related fields and provide a theoretical basis for college teaching reform and talent training [5]. It has long-term significance and reference for promoting the optimization and development of campus culture construction in the entire southwestern region. Physical exercise is an important way and means to promote physical and mental health in colleges. Through the study of college campus sports culture and student physical exercise behavior, the impact of campus sports culture on student sports behavior can be analyzed. The status quo of lifelong physical education, paying attention to the health of college students and physical exercise behavior, has important practical and theoretical significance.

Aiming at the problems in the mining of negative association rules, a mining algorithm that can simultaneously mine positive and negative association rules is proposed. The improved algorithm introduces negative items when mining frequent sets so as to ensure that infrequent sets that may imply negative association rules can be mined. In the generation stage of the association rules, the interest degree is used to cut down the mined association rules so as to ensure the validity of the obtained rules. Aiming at some problems that arise when association rules are applied in complex relationships, and at the same time to further expand the application of association rules, this paper proposes a method of expressing association rules with Bayesian nets. In order to simplify the calculation, the Bayesian network in the article adopts the noisy-OR model. When processing data sets whose scale reaches the maximum utilization rate of the computing platform, compared to algorithms that have not been parallelized, the improved Bayesian classification algorithm for massive data processing has higher operating efficiency, and the time saved varies with the data set. In terms of massive data processing, compared to the naive Bayes algorithm, although the running speed is slightly inferior, the accuracy rate has a higher improvement. Exercise time, exercise methods, and career adaptation do not have a significant correlation effect, exercise time and emotional adaptation do not have a significant correlation effect, exercise methods and self-adjustment do not have a significant correlation, and all others have a positive correlation effect. However, the correlation cannot explore whether there is a causal relationship between the physical exercise behavior assessment management and its various dimensions and the school adaptability and its various dimensions, and regression can investigate and verify the causal relationship between variables. Therefore, this study uses regression analysis to explore the causal relationship between physical exercise behavior assessment management and its dimensions and school adaptability and its dimensions.

2. Related Work

With the widespread application of association rule mining technology, many scholars and experts pay more and more

attention to the application of risk management [6]. Relevant scholars use association rule mining technology to conduct risk analysis on insurance business data [7]. Based on the insurance policy and claim information database established by insurance companies, they search for the characteristics of insurance applicants who have claimed and those who have not and find the riskier ones. Relevant scholars proposed a data mining-based customer credit risk rating system structure and analyzed the construction of a customer credit risk rating index system based on association rules and the construction of a refined visualization module based on a variety of data mining techniques [8].

Many scholars and experts have carried out extensive research on the application of association rule mining in various fields, such as commerce, banking, financial securities analysis, healthcare, engineering, insurance, and telecommunications, and discovered valuable knowledge through association rule mining [9]. Therefore, association rule mining has become the most mature, important, and active research content in data mining. After analyzing the existing research, association rule mining has been applied to risk management research, but it has not yet formed a systematic method to support the entire risk management process [10].

Since the concept of data mining technology was put forward, many experts and scholars have invested a lot of energy in research and made great contributions to the enrichment of data mining systems, which can produce many excellent research results [11]. The Stanford University in the United States has designed and implemented a set of data mining system with complete functions and excellent performance, called DM MNier [12]. The system has a wide range of applications, and there are many types of knowledge that can be mined, including association rules and time series patterns, etc. In addition, the system supports seamless connection with most of the current databases, so it is highly adaptable. IBM also invested a lot of money in the field of data mining and set up a QUEST project team, which analyzed and researched the current mainstream data mining technology, and designed and implemented DB2 Nitelligent MNier for data on this basis. The application is more extensive and has achieved better results.

At present, the number of domestic experts and scholars researching data mining technology is relatively small, most of them are concentrated in key domestic universities, and most of them focus on theoretical research, which can be roughly divided into two types [13, 14]. One is mainly engaged in the research of data mining algorithms, and the other is to transplant existing algorithms to new fields. There are relatively few daily application researches on data mining technology [15]. Therefore, there are few domestic independent research and development data mining products on the market. According to the domestic research situation, association rule mining is the most concerned. Its main function is to discover the association relationship between data items. The concept of association rules originated from the shopping basket problem. This method increases the sales of the supermarket and brings huge profits to the supermarket, which has attracted more and more people's

attention. At present, data mining methods are more and more widely used in China, especially with the advent of the era of big data, data analysis and rule grabbing all need the support of data mining technology [16, 17].

Since the association rule algorithm was proposed, it has received extensive attention and has produced many classic research results [18]. The most famous is the Apriori algorithm. The Apriori algorithm is the basis of the association rule algorithm. Other algorithms are more or less cited. It can be derived from the theory or improvement of it. The Apriori algorithm obtains frequent sets by cyclically scanning the database. The disadvantage of this is that it will generate a huge number of candidate sets, which requires frequent read-and-write operations, which increases the burden on the system. When the amount of data is very large, the computational efficiency of the Apriori algorithm will decrease. In response to the abovementioned shortcomings, many experts and scholars have proposed corresponding improvements [19]. For example, Cerna et al. introduces hashing technology in the Apriori algorithm so that it only needs to scan the database once, which can effectively improve efficiency [20]. Related scholars introduced a sampling method in the Apriori algorithm [21]. In this method, it divides the data into two parts: one is used to generate rules, and the other is used to verify rules. This kind of module can reduce the system burden and improve efficiency. Although the implementation of the above measures can improve the Apriori algorithm, it does not essentially solve the problem.

Relevant scholars pointed out that colleges and universities require college students to take physical exercises in order to satisfy the law of physical and mental development of college students and to carry out sports activities efficiently, consciously, and purposefully [22]. Obvious sports behaviors include obtaining sports information, establishing sports values and attitudes, meeting sports consumption needs, maintaining sports intensity, and organizing sports venues that specialize in activities. Invisible sports behaviors include sports attitudes, sports motivations, and sports' needs. Related scholars believe that sports can be divided into indirect sports behaviors and direct sports behaviors, both of which include sports values, sports' needs, and sports' motivations [23]. Sports behavior is a relatively broad concept; that is, all behavioral activities that are connected with sports can be called sports behaviors. These activities include not only the main manifestation of sports behavior-sports behavior, but also behavioral activities [24].

From the perspective of economic and technical background, the birth of the online sports platform benefits from people's growing demand for sports and the development of Internet technology and the development of the online sports platform have promoted the integration and innovation of Internet technology and sports. The online sports platform does not look at the value chain generation rules of the traditional sports industry. It is a typical bilateral market. In its development process, the platform must not only solve the problem between the growing user demand and the limited ability of its own, but also increase user viscosity, establish a good online reputation, and achieve good

communication through a series of online marketing plans. Combined with the definition of the network platform, the concept of the network sports platform can be defined as a platform that is supported by network technology and connected with sports-related affairs through the Internet and various terminal devices so as to serve network sports users and spread a certain sports culture.

3. Method

3.1. The Operating Mode of the College Students' Health Management Service System. The university's health management service system for college students is the main management department, which effectively combines the sports department, the school hospital, and the mental health consultation center to fully mobilize the advantages of all parties, break the traditional single-department independent operation model, and strengthen the relationship between departments and departments. In order to better ensure the operation of the service system and improve service quality, it is considered necessary for schools to set up corresponding supervision systems and incentive mechanisms. The supervision mechanism is mainly to supervise the authenticity of physical test results, the rationality of the questionnaire of the Mental Health Center, and the transparency of information. The incentive mechanism is mainly to set different material and spiritual rewards for departments and individual students. The main components are additional points rewards, certificate trophies and pennant rewards, and materials and economic rewards.

System management theory provides a solid theoretical basis for my idea. The basic characteristics of system theory are integrity, relevance, hierarchical structure, and so on. We have infiltrated and closely integrated system management theory with emerging disciplines such as information theory, electronic computer, and modern communication technology in the health management services for college students, making health management services more systematic, standardized, quantified, and individualized. Figure 1 shows the organizational structure of the university student health management service system under the Hadoop parallel processing platform.

The various departments of the school carry out their work in an orderly manner in accordance with the service content and operating mechanism of the university student management service system. All work processes are made and transparent, and they are implemented in every work after the freshmen enter the school so that students can more clearly understand the operation process of the entire system and truly feel the health protection that the service system brings to themselves.

The health management department makes full use of student health data. On the one hand, a universally effective health protection manual has been formulated for students whose health level is within a safe range; on the other hand, a targeted health intervention plan has been developed for every student whose health assessment does not meet the standards. The health management service system for college students is not an overnight process. The system needs to be constantly adjusted and perfected according to the changes in the

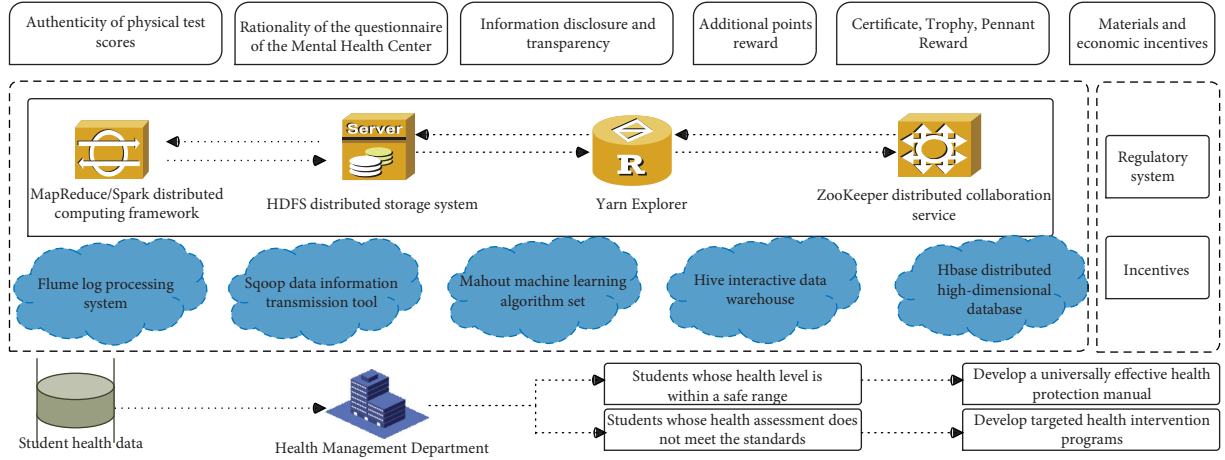


FIGURE 1: Organizational structure of university student health management service under Hadoop platform.

students' own health, so it is bound to be a process of continuous advancement and ascent. After collecting, analyzing, evaluating, guiding, and intervening students' personal health information for the first time, you perform the same process operation for the second time, and then analyze and compare the information with the first time, and conduct health model evaluation in time. According to the student's guidance and intervention effect, it analyzed and evaluated and then formulated an effective health intervention plan again.

3.2. The Basic Idea of Bayesian Network Representation of Association Rules. The form of association rules is very simple, and it is very convenient to apply. Since the association rules were put forward, a large number of mining algorithms have been produced. The speed of mining is getting faster and more accurate, but the association rules cannot express different rules. For example, there are multiple reasons for a result. Even assuming that there is an association rule between each pair of reasons and results, it is difficult for us to comprehensively consider all the reasons when analyzing the probability of the result. Therefore, there are difficulties in the application of association rules in some more complex application fields.

Bayesian network is a graphical model, which can be represented graphically on the same deterministic distribution between variable groups. Burgundy's network contains a set of structural models and related conditional authentication allocation functions. The structural model is a focused recursive graph. Nodes with random variables represent the specific characteristics of the entity, such as process, event, situation, and so on. The contour represents the deterministic ratio between variables. Each node in the graph has a positive conditional distribution function of the node specified by the parent node. In this way, Bayesian graphical network shows how to combine the conditional confirmation function of a group of nodes with the complete simultaneous confirmation distribution function. In the process of reasoning about the state of the nodes, considering the influence of various factors (parent nodes), we thought of using the Bayesian network to

represent the association rules. At the same time, because association rules and Bayesian networks are based on probability theory, our ideas are feasible.

3.3. Assumption of Cause Independence. One of the main problems in the application of Bayesian nets is the specification of conditional probability tables. To specify the conditional probability table of a node, it is necessary to consider all the value combinations of its parent node. If k is used to represent the state number of the parent node, its conditional probability table contains mk parameters. When m is large, the number of parameters that needs to be specified will increase sharply. In the Bayesian network, the parent node of a node can be regarded as the direct cause of the node, and according to people's thinking habits, the influence of a single cause on the result is easier to estimate. Therefore, if it can be ensured that the influence of each cause on the result is not interfered by other causes, the conditional probability table of the Bayesian network can be simplified. A simple formula can be used to calculate the "trust degree" of the node N_j taking the true value True; namely,

$$\text{Bel}(N_j = \text{True}) = 1 - \sum_I \left[c_{ij} \bullet \text{Bel}(N_j = \text{True}) \bullet \text{Bel}(N_{j+1} = \text{True}) \right]. \quad (1)$$

Among them, N_j is the j th Noisy-OR node in the network, N_i represents the i th direct precondition of the Noisy-OR node N_j , also known as the parent node, and Bel represents the degree of trust.

Under the assumption that all the evidences in the network are the ancestors of the node and the network is a pseudo-polytree, the "trust degree" of the node N_j taking the true value True can be calculated by the following formula:

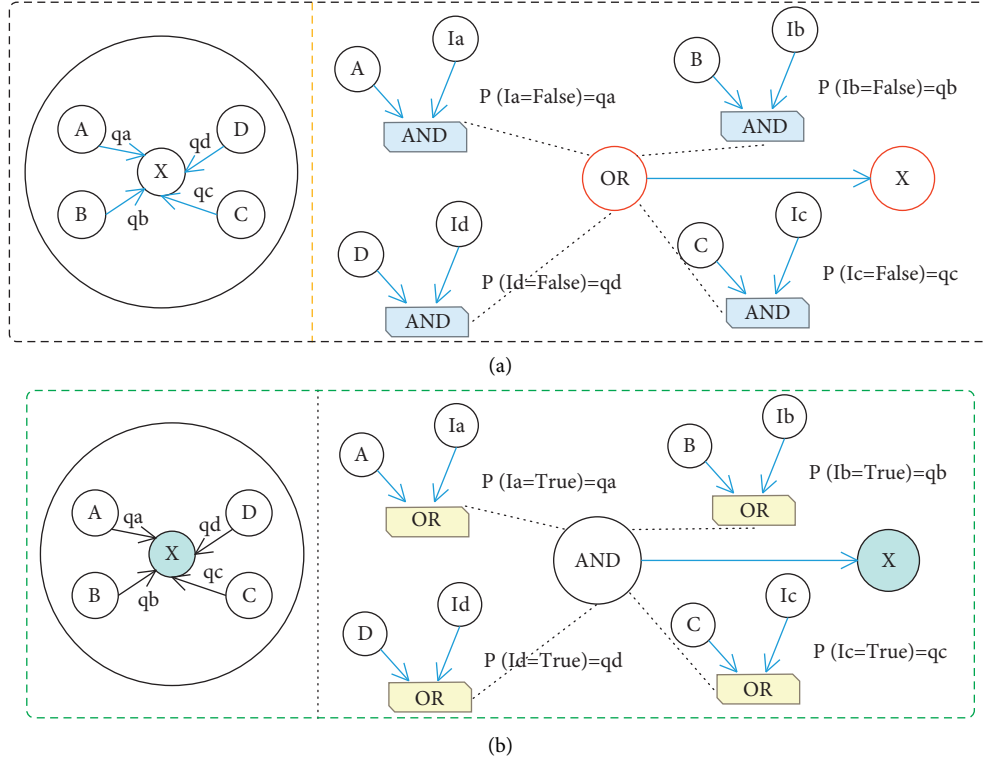


FIGURE 2: Comparison of the conceptual view of the Noisy-OR node and the conceptual view of the Noisy-AND node. (a) Conceptual view of Noisy-OR node. (b) Conceptual view of Noisy-AND node.

$$\text{Bel}(N_j = \text{True}) = \sum_I [1 - \text{Bel}(N_j = \text{True})] \cdot [1 - c_{ij} - \text{Bel}(N_j = \text{True})]. \quad (2)$$

The comparison of the conceptual view of the Noisy-OR node and the conceptual view of the Noisy-AND node is shown in Figure 2.

3.4. Bayesian Network Design of Association Rules. The structure of the Bayesian network is a directed acyclic graph. The directed edges in the graph represent the conditional (causal) dependence between variables (nodes). There is also a dependency relationship between the antecedent and subsequent parts of the association rules. Our idea is to express this dependency in the association rules with a Bayesian network structure. Apriori algorithm is called to use iteration to obtain frequent sets. Since only binomial association rules are needed, the Apriori algorithm is modified here to generate only binomial frequent sets.

A somewhat boundless graph is obtained, where the points are the points corresponding to all the items in the 1-item set in the frequent set L . What needs to be pointed out here is that the items in the association rules mentioned in the algorithm of this paper are different from the points in the Bayesian network. The points in the Bayesian network represent a variable. For example, the item xi in the association rule indicates the status of buying bread; then the point vi in the corresponding Bayesian network represents a

binary variable with the meaning of whether or not to buy bread.

Suppose Q is a node in the Bayesian network and $\text{Parent}(Q)$ is the set of Q 's parent nodes, then

$$P[\neg Q \longrightarrow \text{Parent}(Q)] = \sum P(\neg Q, P_i) \quad (3)$$

$$(P_i \longrightarrow \text{True}) \in \text{Parent}(Q).$$

Substituting the obtained association rule confidence into the above formula, we get

$$P[\neg Q \longrightarrow \text{Parent}(Q)] = \sum_i P(\neg Q, P_i) \cdot [1 - \text{confidence}(P_i|Q)],$$

$$P[\neg Q|\text{Parent}(Q)] = 1 - P[Q|\text{Parent}(Q)] \cdot \text{confidence}(P_i|Q). \quad (4)$$

According to the above formula, the conditional probability table can be constructed without additional calculations but only using the previous step data, which greatly reduces the complexity of calculations. After completing the learning of the conditional probability table, the entire conversion process from the association rule to the Bayesian network is completed.

3.5. Improved Apriori Algorithm for Mining Positive and Negative Association Rules. The support and confidence in negative association rules have the same meaning and form as positive association rules:

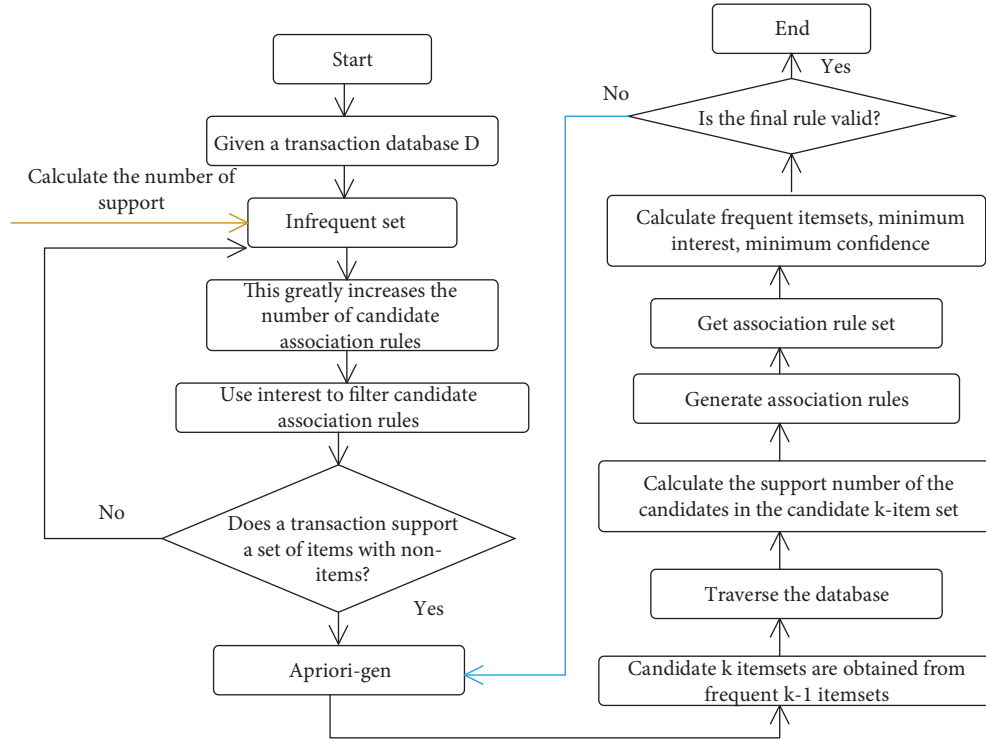


FIGURE 3: The flowchart of the improved Apriori algorithm.

$$\neg X \text{supp}(X) = \frac{[(X \subseteq T) \cap (D \subseteq T)]}{D} \quad (5)$$

The confidence of the negative rule $\neg X \Rightarrow Y$ is

$$\text{confidence}(\neg X \longrightarrow Y) = \frac{\text{supp}(X \cap \neg Y)}{\text{supp}(\neg X \cap Y)} \quad (6)$$

In order to eliminate the useless rules, the algorithm uses the degree of interest to filter the candidate association rules. The degree of interest is a measure that describes the closeness of the connection or influence between the antecedent and subsequent parts of an association rule.

This paper proposes an algorithm that can simultaneously mine positive and negative association rules, namely, the Ex-Apriori algorithm. This algorithm imitates the Apriori algorithm in the generation of frequent sets. The difference is that the state of the items is processed after the first-order frequent sets are generated, thereby ensuring that frequent sets including nonitems are obtained. Then, in the stage of generating association rules, the interest degree is used to screen the candidate rules so as to ensure the validity of the final result. Figure 3 shows the flowchart of the improved Apriori algorithm for mining positive and negative association rules.

In the process of mining association rules, due to the consideration of infrequent sets, the number of candidate frequent sets is greatly increased, resulting in a significant increase in the frequency of calculation of support numbers. In this regard, we proposed a novel support number calculation method, which converts the support judgment of

the transaction to the project into the OR operation of two binary numbers.

Because the Ex-Apriori algorithm is improved from Apriori, nonitems are added to the iteration in the process Generate All Frequent Items, which does not affect the theoretical basis of the Apriori algorithm, so after the call to GenerateAllFrequentItems, it is guaranteed to generate all frequent itemsets.

If the number of 1-item frequent sets is m , the maximum frequent set generated by the Apriori algorithm is k itemsets, and the maximum frequent item generated by the Ex-Apriori algorithm is j itemsets, then $k \leq j \leq m$. In this paper, the Ex-Apriori algorithm traverses the database at most m times (because an item and its nonitem cannot appear at the same time in the item set), so in the worst case, the time consumed by the Ex-Apriori algorithm and the Apriori algorithm is basically the same. Although the Ex-Apriori algorithm usually generates a higher-order frequent set than the Apriori algorithm, which causes an increase in the number of database traversals, we can control this overhead by adjusting the minimum support. Due to the introduction of nonitems, the number of items in the candidate set increases, and the number of calculations for the support number also increases. However, the algorithm in the article uses a novel support number calculation method, which greatly reduces this part of the time expenditure, so the Ex-Apriori algorithm is feasible in terms of efficiency.

At the same time, due to the increase in the number of generated frequent sets, the number of candidate association rules increases. In order to effectively eliminate redundant rules, the Ex-Apriori algorithm introduces interest in the Generate Rules process.

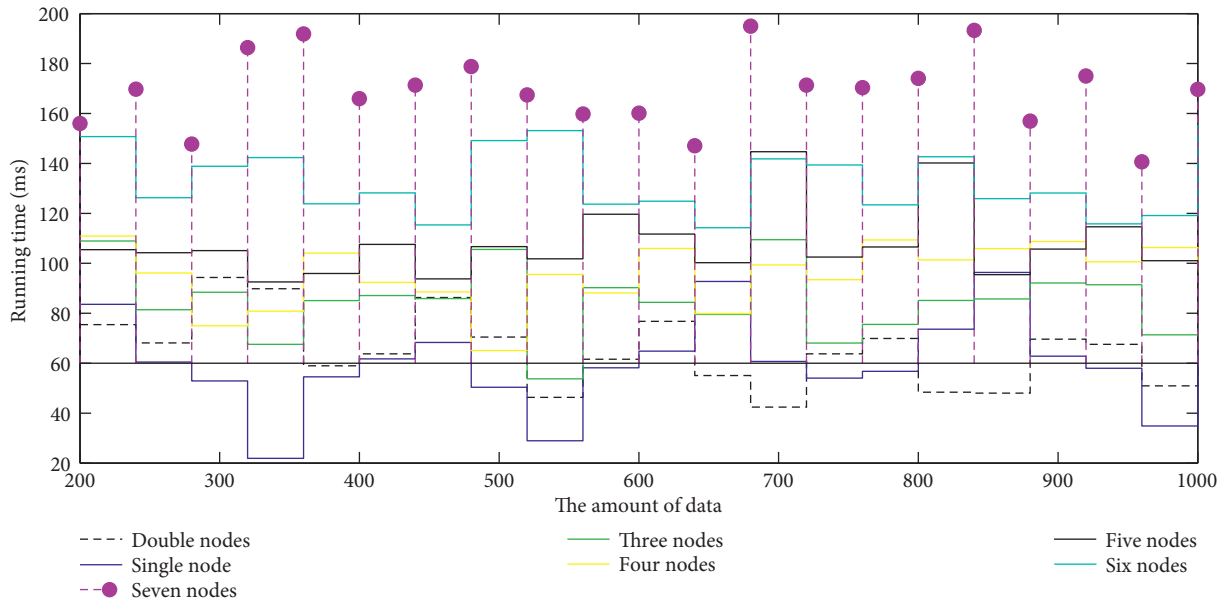


FIGURE 4: The running time of the improved algorithm for different data volumes.

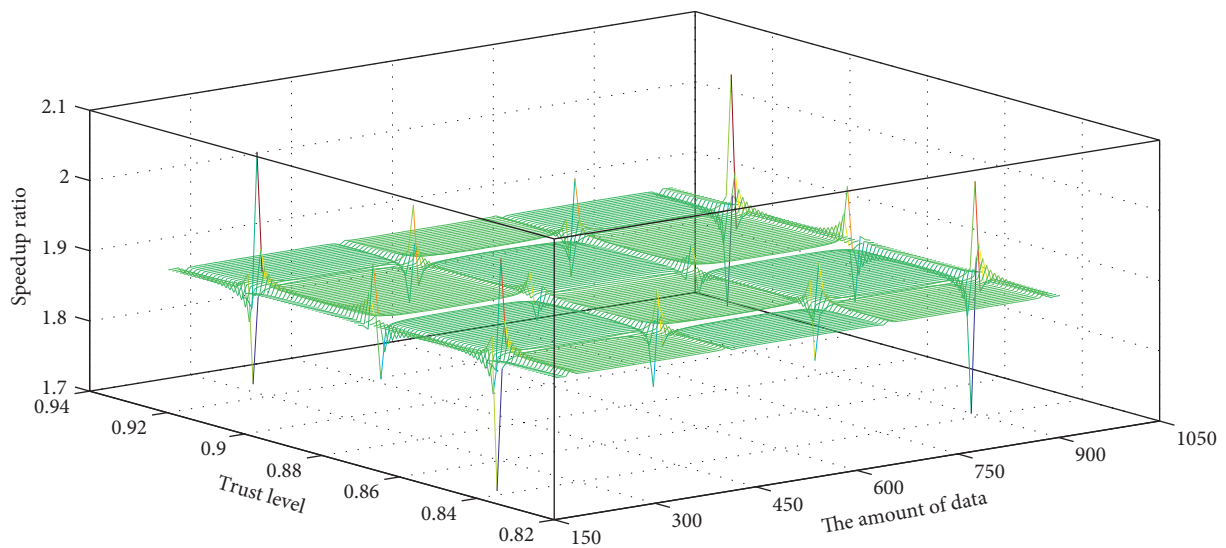


FIGURE 5: Speedup of the algorithm.

In summary, the Ex-Apriori algorithm reduces the complexity of calculations as much as possible on the premise of ensuring that the most association rules can be mined and effectively screens the mined association rules, so the algorithm is effectively feasible.

4. Results and Discussion

4.1. Comparison of Running Time of Different Numbers of Nodes. It can be seen from Figure 4 that with the expansion of the data set size, the parallelization of the 7-node MapReduce still requires less than 200 ms, which has a great advantage in the processing of massive data. The simulation of the speedup of the improved Apriori algorithm is shown in Figure 5.

It can be seen from Figures 4 and 5 that when processing a small-scale data set, the difference between single-node operation and full-node operation of the MapReduce platform is not very large, and the speed-up ratio (single-node running time divided by parallelization running time) is not very different. Although computing has certain advantages, it is not too obvious. However, as the scale of the data set expands, the computing advantages of MapReduce gradually come into play, and the proportion of time saved begins to increase with the increase of the data set. As the size of the data set continues to increase, the speedup ratio is also constantly fluctuating. This is because the MapReduce parallelization platform has a distributed storage process. This process takes a certain amount of time. In the processing of small-scale data sets, due to the small data set and short calculation run time, the distributed storage

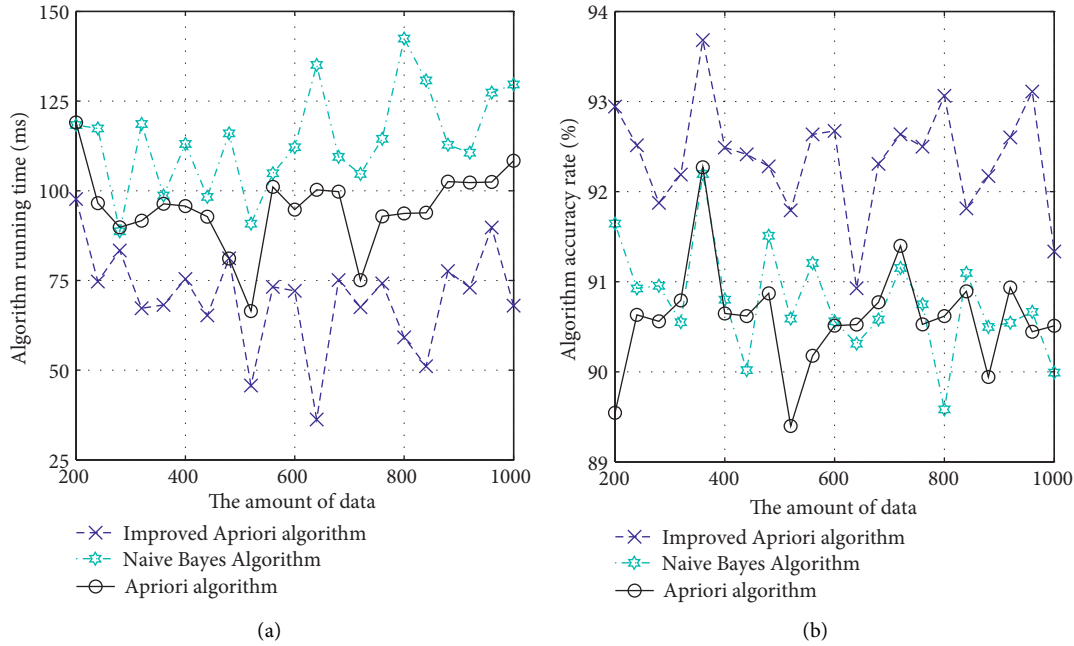


FIGURE 6: Comparison of the running time of the algorithm and the accuracy of the algorithm. (a) Comparison chart of the running time of the algorithm, (b) Comparison chart of the accuracy of the algorithm.

process will occupy a part of the total processing time, so the speedup ratio is low. When processing large-scale data sets, the processing running time is long, and the time occupied by the distributed storage process is only a very small part of the total processing time, so the speedup ratio will gradually increase. However, because the distributed storage process always exists, it is always impossible to form a super linear speedup.

4.2. Comparison of Full-Node Processing Results. The main purpose of this experiment is to test the speed and accuracy of the improved Apriori algorithm proposed in this article when processing massive amounts of data compared to other algorithms. In order to avoid the impact of insufficient data size and too few features on the parallelization algorithm, this experiment directly uses a larger-scale data set with a large number of features.

From the comparison in Figure 6, it can be seen that the improved Apriori algorithm maintains the highest running speed when dealing with large-scale data sets. This is because the improved Apriori algorithm has no clustering links, so the speed is faster. The naive Bayes algorithm is affected by the calculation method of the relationship between the two clusters and the four features (difference degree, feature coefficient, correlation degree, and result-based correlation degree), and the running speed of the naive Bayes algorithm is much slower. In addition, the improved Apriori algorithm has the highest accuracy rate, while also reducing the impact of a single threshold.

4.3. Regression Analysis of Physical Exercise Behavior Evaluation Management and School Adaptability Overall. In order to verify that physical exercise behavior evaluation

management has a significant positive effect on school adaptability, this research explores the relationship between physical exercise behavior assessment management and school adaptability through linear regression. The results are shown in Table 1.

It can be seen from Table 1 that the F -values in Model 1 and Model 2 are both significant at the P value of 0.001, indicating that both Model 1 and Model 2 are statistically significant. The difference between R^2 and adjusted R^2 in Model 1 and Model 2 is both 0.002, indicating that Model 1 and Model 2 have a good fit. The regression coefficients ($P < 0.000$) are significant, indicating that boys' school adaptability is higher than girls' school adaptability, and student leaders' school adaptability is higher. With the introduction of physical exercise behavior evaluation management, R^2 in Model 2 is 0.160, the change of R^2 is 0.098, and the explanatory variable has increased by 9.5%. The regression coefficient of physical exercise behavior evaluation management ($P < 0.000$) is significant, and the unstandardized coefficient of physical exercise behavior evaluation management is 0.426, which indicates that physical exercise behavior evaluation management has a significant positive impact on school adaptability.

The standardized coefficient of student leaders (0.137) is greater than that of gender (0.191), indicating that the impact of student leaders on school adaptability is greater than the impact of gender on school adaptability. It can be seen from Model 2 that the standardized coefficient of gender ($P > 0.05$) is not significant, indicating that with the entry of physical exercise behavior evaluation management, the impact of gender on school adaptability is not significant. The standardization coefficient (0.329) of physical exercise behavior evaluation management is greater than that of student leaders

TABLE 1: Regression analysis of physical exercise behavior evaluation management and school adaptability.

	Model 1 (improved Apriori algorithm)	Model 2 (naive Bayes algorithm)	Model 3 (Apriori algorithm)
Gender (male = 1, female = 0)	2.67	0.99	0.87
Student leaders (yes = 1, no = 0)	5.81	4.12	4.01
F-value	38.13	72.47	71.34
R ²	0.05	0.15	0.17
After R ² adjustment	0.04	0.14	0.16
R ² change	0.01	0.03	0.05

TABLE 2: Regression analysis of self-management and interpersonal relationship adjustment.

	Model 1 (improved Apriori algorithm)	Model 2 (naive Bayes algorithm)	Model 3 (Apriori algorithm)
Gender (male = 1, female = 0)	0.87	0.58	0.56
Student leaders (yes = 1, no = 0)	0.21	0.05	0.04
F-value	15.1	22.4	21.1
R ²	0.01	0.04	0.03
After R ² adjustment	0.02	0.05	0.04
R ² change	0.03	0.04	0.06

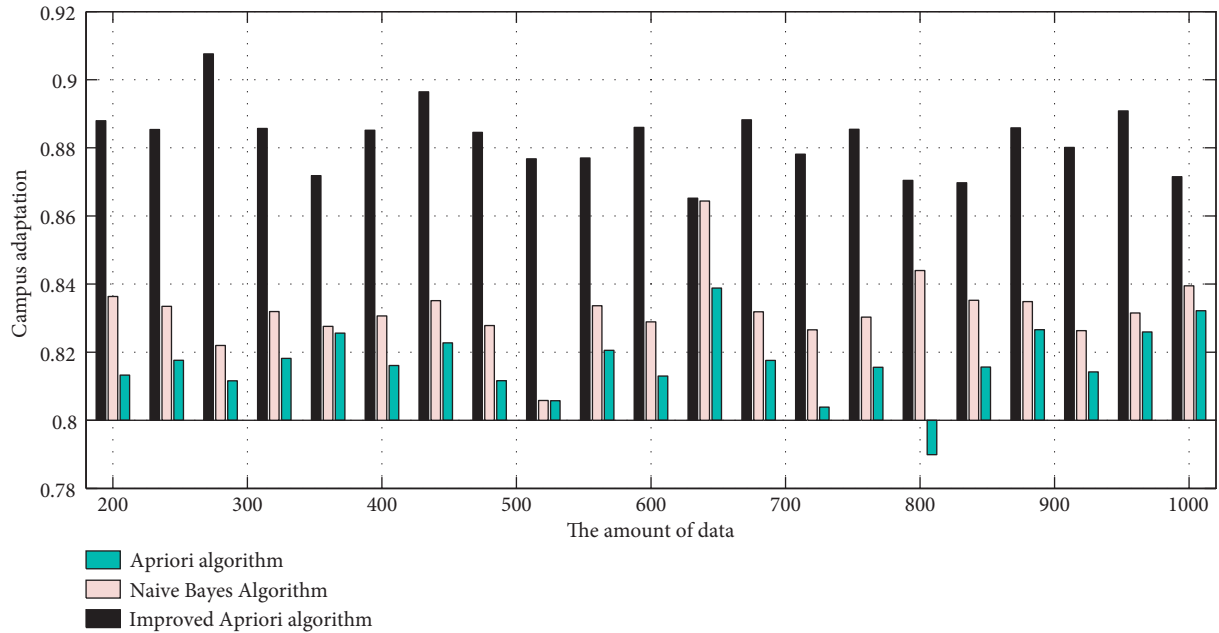


FIGURE 7: Regression results of campus adaptation.

(0.157), indicating that the impact of physical exercise behavior evaluation management on school adaptability is greater than that of student leaders on the impact of adaptability.

4.4. Regression Analysis of Physical Exercise Behavior Evaluation Management and Interpersonal Relationship Adaptation. In order to verify the hypothesis, physical exercise behavior evaluation management has a significant positive effect on the adaptation of interpersonal relationships. This study explored the relationship between physical exercise behavior evaluation management and interpersonal relationship adaptation through linear regression. The results are shown in Table 2.

It can be seen from Table 2 that the F-values in Model 1 and Model 2 are both significant at the P value of 0.001, indicating that both Model 1 and Model 2 are statistically significant. The difference between R² and adjusted R² in Model 1 and Model 2 is 0.002 and 0.003, respectively, indicating that Model 1 and Model 2 have a good fit. The gender unstandardized coefficient is 0.868, and the regression coefficient (P < 0.000) is significant, indicating that the interpersonal relationship adaptation of boys is higher than that of girls, and the unstandardized coefficient of student cadres (P > 0.05) is not significant. The regression coefficient of physical exercise behavior evaluation management (P < 0.000) is significant, and the unstandardized coefficient of physical exercise behavior evaluation

management is 0.071, which indicates that physical exercise behavior evaluation management has a significant positive impact on the adaptation of interpersonal relationships.

4.5. Regression Analysis of Physical Exercise Behavior Evaluation Management and Campus Adaptation. This section verifies that physical exercise behavior evaluation management has a significant positive impact on campus adaptation. This study explored the relationship between physical exercise behavior assessment management and campus adaptation through linear regression. The results are shown in Figure 7.

It can be seen from Figure 7 that the regressions of the campus adaptation of the three models are basically above 0.8, indicating that the three models are all statistically significant. However, it can be clearly seen that the regression results of the improved Apriori algorithm for campus adaptation are the most ideal.

In addition, the regression coefficient of physical exercise behavior evaluation management ($P < 0.000$) is significant, and the unstandardized coefficient of physical exercise behavior evaluation management is 0.071, which shows that physical exercise behavior evaluation management has a significant positive impact on campus adaptation.

5. Conclusion

The average school adaptability of college students is 171.97, and the average item is 2.867. This result shows that the level of college students' school adaptability is slightly lower than the middle level. The average value of job adaptation is 27.27, and the average value of items is 3.03, reaching a medium level. It can be seen that the overall school adaptability of college students is at the lower-middle level. Among all the dimensions, only the career adaptation has reached the middle level, and the other six dimensions are all at the lower-middle level. This shows that the adaptability of college students needs to be improved. Physical exercise behavior evaluation management has a significant positive effect on college students' school adaptability; physical exercise behavior evaluation management has a significant positive effect on school adaptability in terms of exercise motivation, exercise time, and exercise methods, and exercise motivation and exercise time. The order of the magnitude of the influence of exercise methods on college students' school adaptability is as follows: exercise motivation > exercise methods > exercise time; physical exercise behavior assessment management has the effect of physical exercise on college students' school adaptability. The mining algorithm proposed in this paper ensures the comprehensiveness of the mining rules as much as possible, but the experimental results show that the algorithm is slightly too time-consuming, so reducing the calculation time will be the focus of the next stage of work. At the same time, because there are many rules obtained from mining, how to propose a more reasonable measure to reduce the

association rules according to the characteristics of negative association rules is also the research focus of the next stage.

Data Availability

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

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

Acknowledgments

This study was sponsored by the "Theory and Practical Innovation Research on the Comprehensive Reform of College Physical Education Curriculum" team of Xi'an Institute of Translation and Interpretation (Team ID: XFU20KYTDC02).

References

- [1] A. Khan and S. K. Ghosh, "Student performance analysis and prediction in classroom learning: a review of educational data mining studies," *Education and Information Technologies*, vol. 26, no. 1, pp. 205–240, 2021.
- [2] F. Ouatik, M. Raoufi, F. Ouatik, and M. Skouri, "E-Learning & decision making system for automate students assessment using remote laboratory and machine learning," *Journal of E-Learning and Knowledge Society*, vol. 17, no. 1, pp. 90–100, 2021.
- [3] R. Kaviyarasi, T. Balasubramanian, and T. Balasubramanian, "Exploring the high potential factors that affects students' academic performance," *International Journal of Education and Management Engineering*, vol. 8, no. 6, pp. 15–23, 2018.
- [4] W. Xing, C. Li, G. Chen, and X. J. J. C. Huang, "Automatic assessment of students' engineering design performance using a bayesian network model," *Journal of Educational Computing Research*, vol. 59, no. 2, pp. 230–256, 2021.
- [5] S. A. Samerei, K. Aghabayk, A. Mohammadi, and N. Shiwakoti, "Data mining approach to model bus crash severity in Australia," *Journal of Safety Research*, vol. 76, pp. 73–82, 2021.
- [6] M. Qiu and Y. Yang, "Analysis of the current situation and characteristics of college student "online fraud cases"," *International Journal of Mobile Computing and Multimedia Communications*, vol. 12, no. 2, pp. 56–73, 2021.
- [7] G. Czibula, I. G. Czibula, D.-L. Miholca, and L. M. Crivei, "A novel concurrent relational association rule mining approach," *Expert Systems with Applications*, vol. 125, pp. 142–156, 2019.
- [8] S. Asadi, S. Jafari, and Z. Shokrollahi, "Developing a course recommender by combining clustering and fuzzy association rules," *Journal of AI and Data mining*, vol. 7, no. 2, pp. 249–262, 2019.
- [9] M. Ahamad and N. Ahmad, "Students' knowledge assessment using the ensemble methods," *International Journal of Information Technology*, vol. 13, no. 3, pp. 1025–1032, 2021.
- [10] Y. Zhu, H. Lu, P. Qiu, and K. J. Shi, "Heterogeneous teaching evaluation network based offline course recommendation

- with graph learning and tensor factorization,” *Neuro-computing*, vol. 415, pp. 84–95, 2020.
- [11] A. Khamparia and B. Pandey, “Association of learning styles with different e-learning problems: a systematic review and classification,” *Education and Information Technologies*, vol. 25, no. 2, pp. 1303–1331, 2020.
- [12] M. A. Job, “Data mining techniques applying on educational dataset to evaluate learner performance using cluster Analysis,” *European Journal of Engineering and Technology Research*, vol. 3, no. 11, pp. 25–31, 2018.
- [13] X. Du, Y. Pei, W. Duivesteijn, and M. Pechenizkiy, “Exceptional spatio-temporal behavior mining through Bayesian non-parametric modeling,” *Data Mining and Knowledge Discovery*, vol. 34, no. 5, pp. 1267–1290, 2020.
- [14] N. Ketui, W. Wisomka, and K. Homjun, “Association rule mining with permutation for estimating students performance and its smart education system,” *Journal of Computers*, vol. 30, no. 2, pp. 93–102, 2019.
- [15] S. M. Aslam, A. K. Jilani, J. Sultana, and L. Almutairi, “Feature evaluation of emerging E-learning systems using machine learning: an extensive survey,” *IEEE Access*, vol. 9, pp. 69573–69587, 2021.
- [16] F. Wu, Q. Zheng, F. Tian, and Z. Y. K.-M. M. N. J. F. Suo, “Supporting poverty-stricken college students in smart campus,” *Future Generation Computer Systems*, vol. 111, pp. 599–616, 2020.
- [17] B. Ke, M. Khandelwal, P. G. Asteris, and A. D. A. D. J. Skentou, “Rock-burst occurrence prediction based on optimized naïve Bayes models,” *IEEE Access*, vol. 9, pp. 91347–91360, 2021.
- [18] W. Puarungroj, P. Pongpatrakant, N. Boonsirisumpun, and S. Phromkhot, “Investigating factors affecting library visits by university students using data mining,” *LIBRES: Library and Information Science Research Electronic Journal*, vol. 28, no. 1, pp. 25–33, 2018.
- [19] M. A. Sulaiman, “Evaluating data mining classification methods performance in Internet of things applications,” *Journal of Soft Computing and Data Mining*, vol. 1, no. 2, pp. 11–25, 2020.
- [20] H. Gebregziabher, M. Meshasha, and P. Cerna, “Predicting the sero-prevalence of HBV, HCV, and HIV based on national blood of Addis Ababa Ethiopia using data mining technology’s,” *Am J Artif Intell*, vol. 1, no. 2, pp. 44–55, 2017.
- [21] C. K. Chiu and J. C. R. Tseng, “A bayesian classification network-based learning status management system in an intelligent classroom,” *Educational Technology & Society*, vol. 24, no. 3, pp. 256–267, 2021.
- [22] X. Weng, L. Qi, and P. Tang, “Research on college students’ physical exercise trend based on compartment model,” *Mathematics and Computers in Simulation*, vol. 180, pp. 24–42, 2021.
- [23] A. D. Herlambang, S. H. Wijoyo, and A. Rachmadi, “Intelligent computing system to predict vocational high school student learning achievement using Naïve Bayes algorithm,” *Journal of Information Technology and Computer Science*, vol. 4, no. 1, pp. 15–25, 2019.
- [24] S. Kausar, S. S. Oyelere, Y. K. Salal, and S. M. A. S. M. S. Z. X. Hussain, “Mining smart learning analytics data using ensemble classifiers,” *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, no. 12, pp. 81–102, 2020.