

Retraction

Retracted: Construction of Innovative Working Mode of College Counselors Based on Data Mining Technology

Wireless Communications and Mobile Computing

Received 29 August 2023; Accepted 29 August 2023; Published 30 August 2023

Copyright © 2023 Wireless Communications and Mobile Computing. 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This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] L. Wang and J. Xu, "Construction of Innovative Working Mode of College Counselors Based on Data Mining Technology," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 8094732, 9 pages, 2022.

Research Article

Construction of Innovative Working Mode of College Counselors Based on Data Mining Technology

Lei Wang¹ and Jian Xu ²

¹Security Division, Jiangsu University of Science and Technology, Zhenjiang, Jiangsu 212028, China

²Youth League Committee, Jiangsu University of Science and Technology, Zhenjiang, Jiangsu 212028, China

Correspondence should be addressed to Jian Xu; xujian@just.edu.cn

Received 25 May 2022; Revised 8 July 2022; Accepted 12 July 2022; Published 16 August 2022

Academic Editor: Kalidoss Rajakani

Copyright © 2022 Lei Wang and Jian Xu. 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.

Data mining technology can analyze and mine university management data, provide more data support for university management, and play an important role in optimizing teaching quality, but these technologies are rarely applied to the work of university counselors. Based on this, this paper studies the construction and research of college counselors' innovative work mode based on data mining technology. Based on the simple analysis of the impact of big data on the traditional working mode of counselors, this paper introduces the application of data mining commonly used in colleges and universities and puts forward the existing shortcomings. The innovative working mode of college counselors is designed. In the analysis of students' daily behavior, the advantages of cluster analysis and support vector machine are used to analyze students' consumption behavior. The Apriori algorithm is applied to the student achievement early warning management to improve the Apriori algorithm. Simulation results show that the proposed algorithm can shorten the running time, reduce the number of frequent item sets, and improve the classification accuracy.

1. Introduction

The development of Internet technology provides effective means for the management of college counselors and the communication between students and makes up for the shortcomings of the traditional counselor working mode. At present, big data in colleges and universities can promote full coverage of work, grasp the status of students at any time, and provide a lot of help for enriching educational means. The network itself also provides more educational ways for counselors and has a certain effect on cultivating students' independent personality. At present, the data in the daily management of college counselors basically come from the campus network, and the working mode is constantly improving, but these data belong to shallow data analysis, and the association and information between data have not been mined and used. Apriori algorithm is the first association rule mining algorithm and the most classic algorithm. It uses the iterative method of layer by layer search to find the relationship between item sets in the database to

form rules. The process includes connection (matrix-like operation) and pruning (deleting those unnecessary intermediate results). The concept of item set in this algorithm is item set. A set containing K items is a set of K items. The frequency of the item set is the number of transactions that contain the item set, which is called the frequency of the item set. If the item set meets the minimum support, it is called frequent item set [1].

Based on this background, this paper studies the construction and research of college counselors' innovative work mode based on data mining technology, which is mainly divided into four chapters [2]. The first chapter briefly introduces the evaluation of college counselors' work mode and the chapter of this study; Chapter 2 introduces the application of data mining technology in colleges and universities at home and abroad and summarizes the shortcomings of the current research [3]. The third chapter constructs the counselor working mode based on data mining technology, analyzes it from the perspectives of student behavior analysis and achievement early warning, and applies the data

mining algorithm. The fourth chapter simulates the behavior analysis and achievement early warning of the college counselor working mode constructed in this paper. The experimental results show that compared with the traditional algorithm, the improved algorithm proposed in this paper can improve the classification accuracy, shorten the running time, and provide more intuitive data for counselors [4].

The innovation of this paper is to combine the characteristics of cluster analysis and support vector machine, put forward a new clustering method to cluster analyze students' daily consumption, and apply it to the support vector machine model. In the performance early warning design, we improve the shortcomings of the traditional Apriori algorithm, improve the pruning strategy in the optimization operation of the Apriori algorithm, reduce the number of candidate item sets, divide the large database into independent and irrelevant small databases, and shorten the operation time [5].

2. State of the Art

At present, colleges and universities at home and abroad have basically established their own campus network and began to be applied to teaching management, which provides a lot of convenience for teachers and students. How to use these campus data has become one of the hot topics. In the analysis of students' learning behavior, Pang C et al. combined with the traditional clustering analysis algorithm and random forest algorithm to improve the traditional algorithm, combined with the human skeleton model to identify students' classroom behavior in real time, and constructed the network topology model [6]. Bechter BE et al. used potential contour analysis methods to analyze students' motivation in their research [7]. Yang Q et al. proposed a personalized course recommendation algorithm based on hybrid recommendation. The improved new Apriori algorithm is used to realize association rule recommendation. The user-based collaborative filtering algorithm is the main part of the algorithm, which solves the problems of data sparsity and cold start [8]. Cantabella M and others use the big data framework to implement statistics and association rule technology, so as to speed up the statistical analysis of data. The results obtained were demonstrated and evaluated using visual analysis technology to find the trends and defects of students' use of LMS [9]. Liu N et al. in the study of College English score analysis, the College English test system software based on data mining mainly realizes the automatic generation of test papers by computer programs, sets the test time, automatically judges the test scores of candidates, and gives the scores on the spot [10]. Pan T uses Apriori algorithm to mine the hidden correlation between physical fitness indicators from college students' sports data and identify the indicators closely related to college students' physical fitness [11]. In the analysis of students' psychological problems, Liu J and others used data mining technology to realize the dynamic management of psychological early warning data, monitor the psychology of high-risk groups in real time, and improve the accuracy and effectiveness of early identification and early warning of students' psycho-

logical crisis [12]. Yang Zhao adopted the management measures of "standardization of data collection, diversification of data application, and institutionalization of data management" to improve the quality of education and teaching [13].

To sum up, we can see that there are many researches on data mining and university management at home and abroad. Almost all universities have realized information management in management, and the data needed for counselor work is also very sufficient, but these data often cannot provide more data support for management work. In data mining, clustering analysis and association rule-based algorithms related to data mining are emerging and improving. However, the application of data mining algorithm in colleges and universities is mostly reflected in the establishment of database, and the application of counselor working mode is obviously insufficient. Therefore, it is of great practical significance to study the construction of innovative work mode of college counselors based on data mining technology.

3. Methodology

3.1. Constructing the Innovative Working Mode of College Counselors. At present, information technology has been realized in the work of college counselors, which can record all the data of students in school. These data have the characteristics of multiple isomerization and a large number. In such an environment, the management of counselors and students' behavior will be stored. These data can be mined and used to improve the traditional work model. The counselor's work model covers many contents, such as work content and work means. This paper improves the work model from the technical point of view.

The improvement of counselor's working mode is based on big data mining, which analyzes and processes students' data, looks for key information, and converts this information into usable knowledge. At present, the amount of data provided by college counselors is huge and complex, so it is necessary to extract and filter the data to find association rules, as shown in Figure 1. In the design, it is also necessary to provide friendly interactive functions and interfaces for easy understanding.

Exchange and cooperate with colleges and universities in data collection. On the basis of ensuring data privacy, open the data structure. The collected information covers all-in-one card data, access control data, educational administration data, library data, etc. Most of these data are stored in CSV format. In data mining analysis, we need to rely on the fusion of multisource data, so we also need to purify when accessing. If the data of students' economic status belongs to discrete data, other data can be imported directly. When the information of multiple databases is fused, the fused data will be affected by other factors, such as noise data and incomplete data. Only by ensuring the quality of data can we ensure the results of data mining and analysis. Therefore, the data needs to be processed to some extent. For wrong data, redundant data, and missing data, interpolation method and mean value method are used to ensure the integrity of the data. If there is a logical error in the data, it is directly treated as missing value.

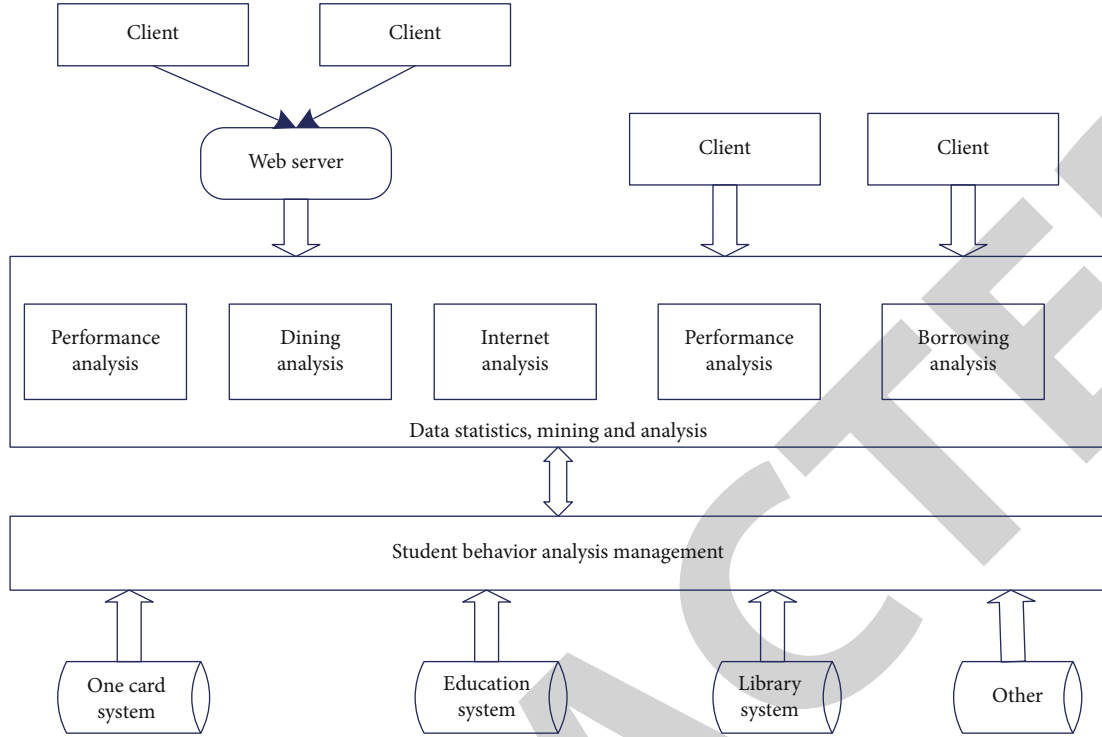


FIGURE 1: Innovative design of the work mode of college counselors.

3.2. Cluster Analysis of Student Behavior. The data sources used in the work of college counselors are diverse. These big data basically cover all the data of students in school. Only by carrying out targeted analysis on these data and extracting the effective characteristics between the data can we carry out the work better. The analysis of data changes covers static data and dynamic data. In addition to some structured data, these data also have a lot of unstructured processing. In terms of content, it includes students' learning data and daily life data. Classifying data can provide important technical support for establishing innovation work model.

In the analysis of students' daily living conditions, we use the short text data mining method to mine students' family conditions and later text data, use the text information as keywords for analysis, find keywords with high correlation with economic grade, and then use machine learning algorithm for data mining analysis [14]. For language, vocabulary is an independent and meaningful element. At present, most word segmentation algorithms are based on word frequency statistics, which can meet the processing of natural language and analyze a large number of texts according to this method [15]. In the analysis of the importance of vocabulary, TF-IDF is used to evaluate, check the frequency and frequency of occurrence, and reflect the importance of vocabulary to the document. The formula is expressed as follows:

$$w = \frac{n_{i,j}}{\sum n_{k,j}} \times \log \frac{|D|}{1 + |\{j : t_i \in d_j\}|}, \quad (1)$$

where n represents the number of words and D represents the number of files. This algorithm can filter out common words

and retain important information. Chi square test is used to test the independence of independent words. After selecting a specific text, in order to ensure mutual independence, chi square test is carried out. The formula is:

$$x^2 = \frac{N(ad - bc)^2}{(a + c)(a + b)(b + d)(c + d)}, \quad (2)$$

where N represents the number of documents. The larger the value, the more obvious the text belongs to this category. When analyzing the trajectory of students' behavior, we need to consider not only the behavior of students in a certain period of time but also the number of times students appear in a certain place. If there are many times in a canteen, it shows that this position is very important for students. The behavior trajectory belongs to spatial change. The closer the time period is, the more it can reflect the behavior trajectory of students. In order to reflect the trajectory of students at different time points, we need to consider the difference of occurrence times, cross time period, and boundary value. In the calculation of student similarity, the cosine similarity calculation method is used, and the formula is expressed as:

$$e_{s,t,p} = 1 \sum_n^{N_{stp}} e_{st} f_{stp}(t), \quad (3)$$

where f represents the marginal effect function of students, f represents the number of occurrences, and the formula is:

$$f_{stp}(t) = e^{1/H|d-d_0|}, \quad (4)$$

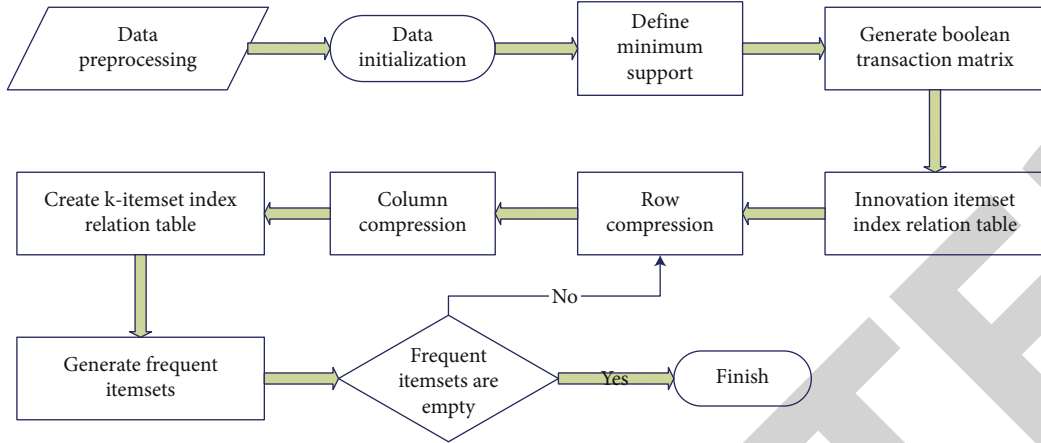


FIGURE 2: Early warning process based on association rules.

where d represents the time and H represents the absolute maximum value. Further calculate the similarity of students, and the calculation formula is:

$$\text{sim}(s_1, s_2) = \frac{\sum_{t \in T} \sum_{p \in P} e_{s_1, tp}^d e_{s_2, tp}^d}{\sqrt{\sum_{t \in T} \sum_{p \in P} e_{s_1, tp}^d} \sqrt{\sum_{t \in T} \sum_{p \in P} e_{s_2, tp}^d}}, \quad (5)$$

where s represents the student, T represents the time period, and p represents the probability of occurrence. The range of similarity is 0~1. If the trajectories are completely different, the value is zero. If they are completely similar, the value is 1. The greater the similarity value, the more identical the trajectories of the two students. Suppose that the training sample set is represented by L and I , respectively, belonging to the first type, the expected value output is positive, and the others are negative. Construct decision discriminant function to ensure that all data are correctly classified. Assuming that the samples are linearly separable, the training set can be separated by the hyperplane, and the classification interval is the largest; this surface is the optimal classification variable. At this time, the discriminant function is expressed as:

$$f(x) = \text{sgn} \{ (w^* \cdot x) + b^* \}, \quad (6)$$

where b^* represents the classification threshold. Calculate the maximum value of the objective function with the formula:

$$w(\delta) = \sum_{i=1}^n \delta_i - \frac{1}{2} \sum_{i,j} \delta_i \delta_j y_i y_j (y_i \cdot y_j). \quad (7)$$

If the sample cannot be linearly divided, it is calculated by nonlinear function, and the kernel function is selected. The formula is expressed as:

$$k(x, y) = \exp \{ -|x - y|^2 / \delta^2 \}. \quad (8)$$

The kernel function maps the data to the high-dimensional space, solves the optimal classification hyper-

plane in the high-dimensional space, and obtains the objective function and classification function. The formula is expressed as:

$$w(\delta) = \sum_{i=1}^n \delta_i - \frac{1}{2} \sum_{i,j} \delta_i \delta_j y_i y_j (x_i \cdot x_j), \quad (9)$$

$$f(x) = \text{sgn} \left\{ \sum_{i=1}^n \delta_i y_i (x_i \cdot x) + b^* \right\}. \quad (10)$$

3.3. Early Warning Model Based on Association Rule Algorithm. In the innovation of efficient counselor working mode, student behavior analysis and achievement early warning are necessary modules. In functional design, data storage and mining and achievement early warning need to be realized. According to the demand, the big data of colleges and universities can send data to the file system regularly and update it every other period of time. Therefore, spark platform can be used in data preprocessing and statistical analysis. After the data preprocessing is completed, part of the basic data is stored in the database, and the other part of the data is used to analyze the learning behavior of students. This part of the data is used as the initial data, then association rule analysis is carried out, and the analysis results are stored in the database. See Figure 2 for details.

The university database sharing platform collects students' information, including all-in-one card data, educational administration data, library data, and other data. At present, this work can be carried out regularly and updated regularly. The amount of all-in-one card data and library storage data is relatively large, and the data of educational administration system and library system will be relatively small. If the data is not updated within a period of time, the updated data will not appear on the corresponding platform. The algorithm based on association rules is used in student learning early warning mining. The data to be analyzed mainly comes from the student learning behavior data in the data, and then the results are fed back to the system database.

At present, there are many association rule algorithms, such as multicycle mining algorithm, constraint-based

mining algorithm, and parallel algorithm, which are classical algorithms [16]. When selecting association rule algorithm, it needs to be analyzed in combination with the characteristics of the course. In the research, this paper mainly analyzes the situation of students' learning early warning, and the data involved are related to their grades, including students' class status, book borrowing, course grades, and other information. These data cover many aspects and have different attributes. Therefore, multidimensional association rules need to be used in data mining [17]. Among the discrete multidimensional association rule algorithms, Apriori algorithm is a classical algorithm. This paper is also based on this algorithm when mining student information.

Among many association rule mining algorithms, Apriori algorithm is one of the classical algorithms. The algorithm takes advantage of the properties of two frequent item sets. It is assumed that there are n frequent item sets and any other one is infrequent item set. In the mining process, it is necessary to find all frequent item sets and generate association rules according to this [18]. Compared with ALS algorithm, the most obvious disadvantage of Apriori algorithm is that the candidate set may grow exponentially in data mining. Apriori algorithm has strong scalability and applicability, but it also has its own defects. In the improvement of Apriori algorithm, the method based on matrix compression is used to assume that the frequent item sets have synchronization, and the properties can be inversely inferred [19]. The transaction matrix is formed by the transformation of the database, and the item sets are sorted according to the size. If the item sets are in the transaction, it is considered that D is 1; otherwise, it is 0. Such a transaction matrix is the transaction database, that is, the Boolean transaction matrix. After conversion, it is expressed as:

$$M = I \begin{bmatrix} d_{11}, d_{12}, \dots, d_{1n} \\ d_{21}, d_{22}, \dots, d_{2n} \\ \dots \\ d_{m1}, d_{m2}, \dots, d_{mn} \end{bmatrix}. \quad (11)$$

From this matrix, we can see that the item sets corresponding to m are I_m , respectively. Any k rows in the transaction matrix M form an item set. The formula is:

$$\text{support}(C_k) = \sum_{j=1}^n (d_{i1j} \wedge d_{i2j} \wedge \dots \wedge d_{ikj}). \quad (12)$$

When the K value is 1, the calculation formula of item set support is expressed as:

$$\text{support}(C_1) = \sum_{j=1}^n (d_{ij}) (i = 1, 2, \dots, m). \quad (13)$$

In the calculation process, candidate item sets will appear in each iteration, and the index relationship table is used to express the support. If the data set appears 4 times in the Boolean transaction matrix, the corresponding sup-

port is 4. Recalculate the pruned item set using Boolean transaction matrix [20]. With the increase of the amount of data, the number of item sets is large, which increases the difficulty of calculation. In order to reduce the computational complexity, the parallel processing of multithreaded operation is used. Divide the data based on certain standards to get different partitioned databases. This improved method can carry out parallel data mining and improve the mining efficiency. Divide and conquer the database to obtain an independent and irrelevant partitioned database, which is expressed as:

$$D = \{D_1, D_2, \dots, D_n\}, \quad (14)$$

where D_i stands for partitioned database, and all independent databases form a transaction database. For each partitioned database, the improved Apriori algorithm is used to analyze students' behavior. The results of frequent item sets can be obtained, which are represented by L_k . The final results are combined and expressed as:

$$L = L_1 \cup L_2 \cup \dots \cup L_n. \quad (15)$$

4. Result Analysis and Discussion

4.1. Simulation Analysis of Student Behavior. Different index data have different dimensions and need to be standardized. This paper uses the method of normalization to zero mean for processing. The daily consumption data of students are analyzed by clustering algorithm. The clustering results are used to filter the original data and discrete data, and the consumption data to be analyzed are selected. There are some occasional data in students' daily consumption, which are not universal, and some data cannot reflect individual consumption, such as electricity and books. Therefore, select the data closely related to students in the student all-in-one card consumption data, such as supermarket, catering, and fruit consumption.

Cluster the students' behavior data, establish the sample set of consumption data, calculate the sample similarity matrix and Laplace matrix, and calculate the eigenvalue. According to the consumption situation, the number of clusters is set to 4. Select the financial status from the existing list of students with financial difficulties for investigation. In the simulation analysis of the algorithm, the model based on support vector machine is affected by the selection of kernel function and has some changes. In order to more comprehensively analyze the impact of different characteristics on simulation, different types of data are selected for analysis, and different kernel function training models are adopted. The results are shown in Figures 3 and 4.

It can be seen from the data in the figure that clustering analysis using students' daily consumption data can accurately analyze students with different economic conditions, and the accuracy is high. The choice of different kernel functions has a certain impact on the accuracy, but the overall difference is small. The prediction accuracy of RBF kernel function is the highest.

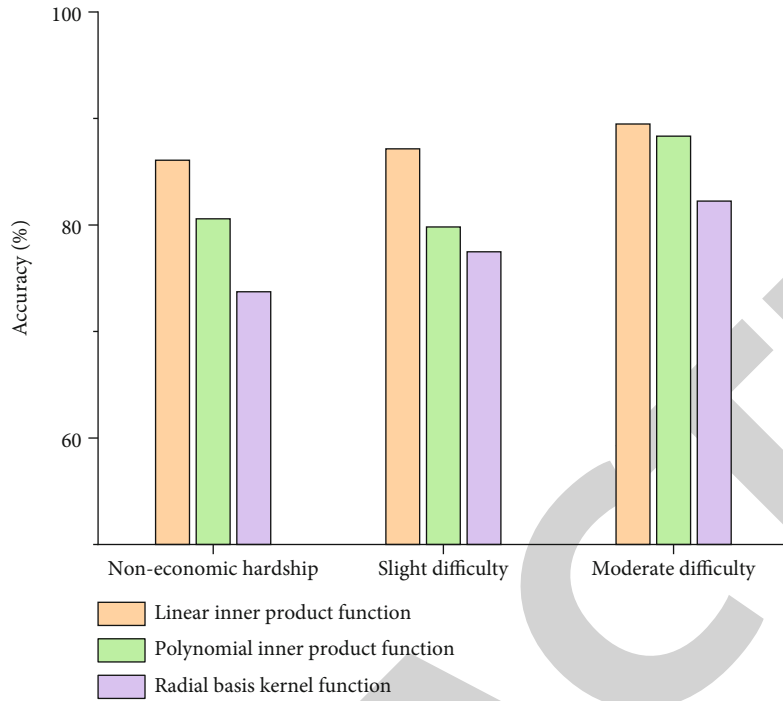


FIGURE 3: Accuracy rate of difficulty in economic certification of students in Group 1.

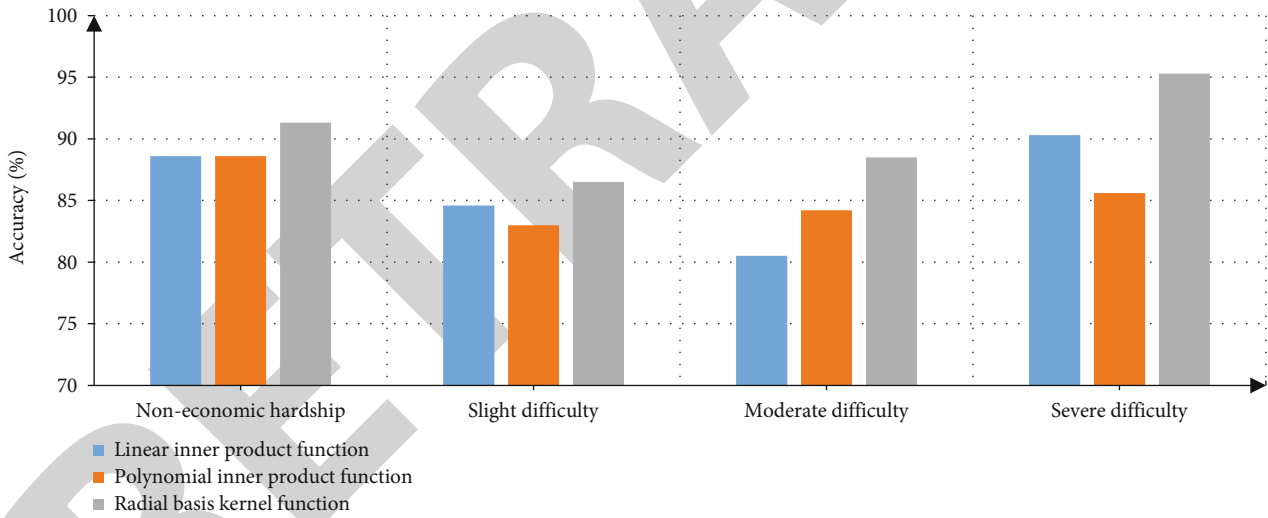


FIGURE 4: Accuracy rate of difficulty in economic certification of students in Group 2.

4.2. *Simulation Analysis of Performance Early Warning.* The effectiveness of this algorithm is verified before the simulation analysis of performance early warning. The same software and hardware platform is used for simulation analysis. The test data comes from the student achievement database, including 500 records, and each record contains multiple fields. In the test, two methods of fixed minimum support and transformed minimum support are used for the test. When the minimum support is set to 0.2, the number of candidate item set C1 is 37, the number of candidate item set C2 is 478, the number of candidate item set C3 is 1268, the number of candidate item set C4 is 684, the number of candidate item set C5 is 65, and the number of candi-

date item set C65 is 35. You can see that the candidate item set rises first and then decreases.

Test the running time of each algorithm under the same candidate item set. It can be seen from the figure that the running time of the traditional Apriori algorithm and the improved Apriori algorithm on the candidate item set C1 is basically the same. For other candidate item sets, the running time of the improved Apriori algorithm is significantly shorter than that of the traditional Apriori algorithm, which shows that the improved Apriori algorithm used in this paper can effectively shorten the running time of the item set (Figure 5).

In order to further analyze the performance of the improved algorithm, test the running time of each algorithm

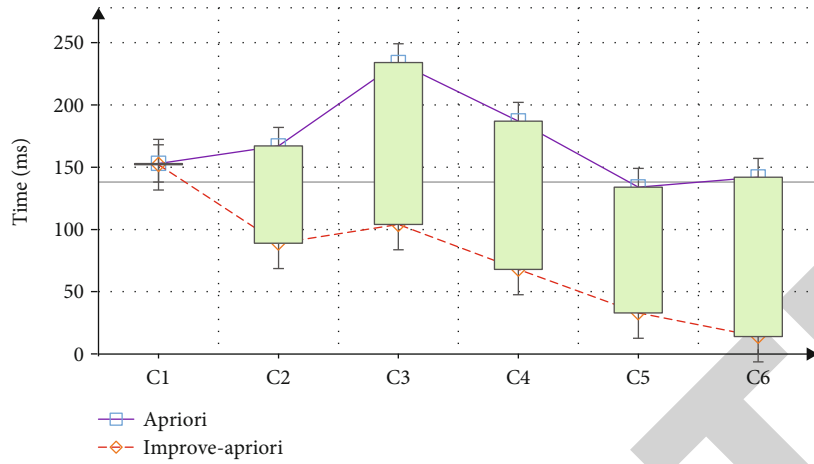


FIGURE 5: Time to generate candidate item sets under different algorithms.

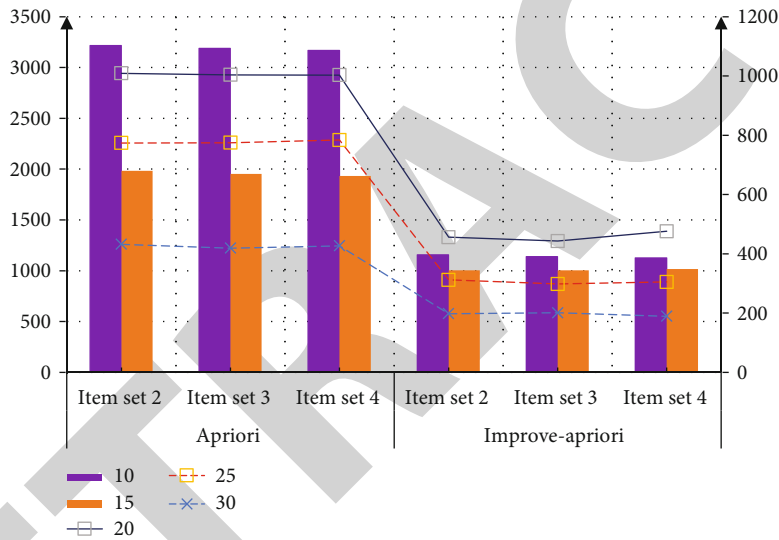


FIGURE 6: Support running time under different algorithms.

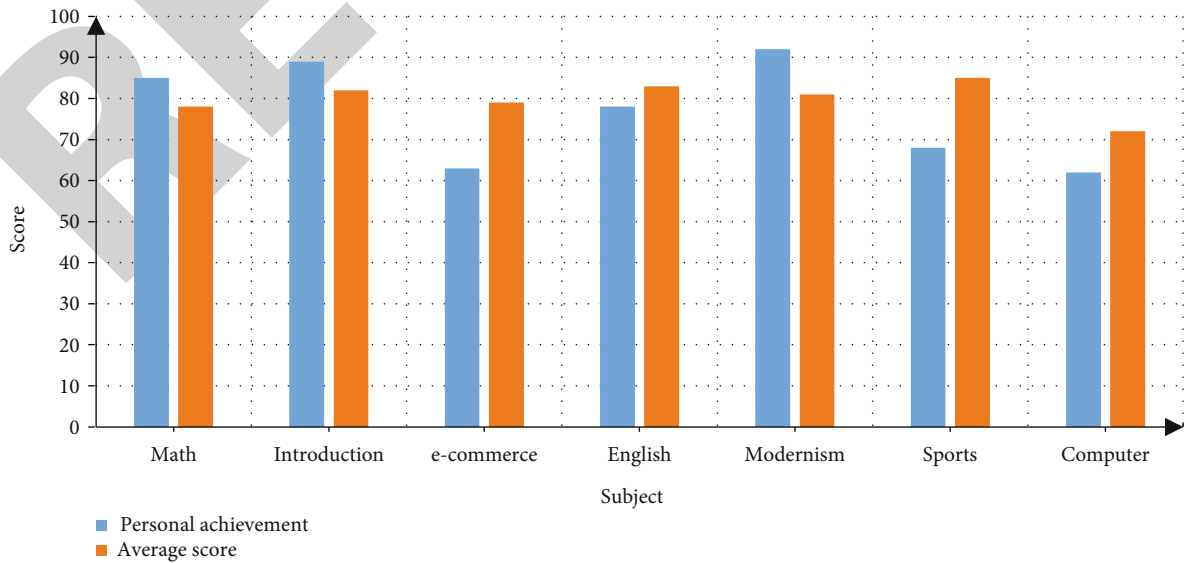


FIGURE 7: Early warning of students' academic performance.

under different support. All algorithms run for 5 times and take the average value. The test results are shown in Figure 6. It can be seen from the data in the figure that the improved algorithm proposed in this paper can show advantages under different support degrees, indicating that the operation efficiency of the improved algorithm is significantly improved, and this advantage is more obvious with the reduction of support degrees.

In the application of the algorithm, the characteristic data is discretized and discretized according to the properties, value range, and characteristics of the characteristic attributes. Visual technology is used to analyze students' grades and provide intuitive understanding for counselors and students. From Figure 7, we can see that students' grades in all subjects are more than 60 points, and there is a significant gap between personal achievement and average score. Personal achievements in learning, introduction, e-commerce, and modernism are higher than average; personal achievements in English, e-commerce, sports, and computer are higher than average. The algorithm can accurately predict students' academic performance.

As shown in Figure 7, the student's performance is compared with the average level of the major. You can clearly see the comparison between the student's performance and the average level of each subject, the counselor can clearly see the student's learning status, and you can give early warning for backward subjects.

5. Conclusion

- (1) It is processed by normalizing to zero mean. Clustering algorithm is used to analyze students' daily consumption data. In view of the shortcomings of data mining algorithm, an optimization and improvement strategy is proposed, and the Apriori algorithm is optimized by combining clustering analysis with support vector machine
- (2) Use the same software and hardware platform for simulation analysis. The test data comes from the student achievement database, including 500 records, each of which contains multiple fields. In the test, the fixed minimum support and the transformed minimum support are used to test. The running time of the traditional Apriori algorithm is basically the same as that of the improved Apriori algorithm on candidate item set C1. For other candidate item sets, the running time of the improved Apriori algorithm is significantly shorter than that of the traditional Apriori algorithm. The improved algorithm proposed in this paper can show advantages under different degrees of support, indicating that the operation efficiency of the improved algorithm has been significantly improved, and this advantage is more obvious with the reduction of support
- (3) In the application of the method, the characteristic data is discretized and discretized according to the

nature, value range, and characteristics of the characteristic attributes. Visual technology is used to analyze students' grades and provide intuitive understanding for instructors and students. In data mining, the data capacity is huge, and the results are also very complex. Some noisy data may appear, which will affect the mining results. This aspect needs improvement

Data Availability

The figures used to support the findings of this study are included in the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

Thanks are due to those techniques who have contributed to this research. This work was supported by the Research on the Working Path of Ideological and Political Education in Universities Based on the Integration of Second Classroom and Ideological and Political Education, Special Project of Ideological and Political Education in Universities, Jiangsu Social Science Application Research Boutique Project, 2021 (Code: 21SZB-040); and Research on the Construction of College Students' Entrepreneurial Ability Evaluation Model from the Perspective of Ideological and Political Education, Special Project of Ideological and Political Work, Jiangsu University Philosophy and Social Sciences Research Project, 2018 (Code: 2018SJSZ336).

References

- [1] T. Wu, S. Chen, Y. Tian, and P. Wu, "A feature optimized deep learning model for clinical data mining," *Chinese Journal of Electronics*, vol. 29, no. 3, pp. 476–481, 2020.
- [2] L. Wang and S. J. Chung, "Sustainable development of college and university education by use of data mining methods," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 16, no. 5, p. 102, 2021.
- [3] X. Jie and L. Yang, "CET-4 score analysis based on data mining technology," *Cluster Computing*, vol. 5, pp. 1–11, 2018.
- [4] Z. He and C. Yu, "Clustering stability-based evolutionary K-means," *Soft Computing*, vol. 23, no. 1, pp. 305–321, 2019.
- [5] F. Liu, D. Yang, Y. Liu et al., "Use of latent profile analysis and k-means clustering to identify student anxiety profiles," *BMC Psychiatry*, vol. 22, no. 1, pp. 1–11, 2022.
- [6] C. Pang, "Simulation of student classroom behavior recognition based on cluster analysis and random forest algorithm," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, pp. 2421–2431, 2021.
- [7] B. E. Bechter, J. A. Dimmock, J. L. Howard, P. R. Whipp, and B. Jackson, "Student motivation in high school physical education: a latent profile analysis approach," *Journal of Sport & Exercise Psychology*, vol. 40, no. 4, pp. 206–216, 2018.
- [8] Q. Yang, P. Yuan, and X. Zhu, "Research of personalized course recommended algorithm based on the hybrid

- recommendation,” *MATEC Web of Conferences*, vol. 173, no. 3, p. 03067, 2018.
- [9] M. Cantabella, R. Martínez-España, B. Ayuso, J. A. Yáñez, and A. Muñoz, “Analysis of student behavior in learning management systems through a Big Data framework,” *Future Generation Computer Systems*, vol. 90, pp. 262–272, 2019.
- [10] N. Liu and D. Li, “Framework and performance analysis of college English testing system based on data mining technology,” *Journal of Intelligent and Fuzzy Systems*, vol. 3, pp. 1–11, 2021.
- [11] T. Pan, “An improved Apriori algorithm for association mining between physical fitness indices of college students,” *International Journal of Emerging Technologies in Learning (iJET)*, vol. 16, no. 9, p. 235, 2021.
- [12] J. Liu, G. Shi, J. Zhou, and Q. Yao, “Prediction of college students’ psychological crisis based on data mining,” *Mobile Information Systems*, vol. 2021, 7 pages, 2021.
- [13] Y. Zhao, “Research on the application of university teaching management evaluation system based on Apriori algorithm,” *Journal of Physics: Conference Series*, vol. 1883, no. 1, pp. 012033–012033, 2021.
- [14] M. Ali, S. Khalid, and M. H. Aslam, “Pattern based comprehensive Urdu stemmer and short text classification,” *IEEE Access*, vol. 6, pp. 7374–7389, 2018.
- [15] A. Tommasel and D. Godoy, “A social-aware online short-text feature selection technique for social media,” *Information Fusion*, vol. 40, pp. 1–17, 2018.
- [16] Y. Guo, M. Wang, and X. Li, “Application of an improved Apriori algorithm in a mobile e-commerce recommendation system,” *Industrial Management & Data Systems*, vol. 117, no. 2, pp. 287–303, 2017.
- [17] Y. Djenouri and M. Comuzzi, “Combining Apriori heuristic and bio-inspired algorithms for solving the frequent itemsets mining problem,” *Information Sciences*, vol. 420, pp. 1–15, 2017.
- [18] S. Zhao, B. Huang, and L. Fei, “Linear optimal unbiased filter for time-variant systems without Apriori information on initial conditions,” *IEEE Transactions on Automatic Control*, vol. 62, no. 2, pp. 882–887, 2017.
- [19] P. H. Lu, J. L. Keng, F. M. Tsai, P. H. Lu, and C. Y. Kuo, “An Apriori algorithm-based association rule analysis to identify acupoint combinations for treating diabetic gastroparesis,” *Evidence-based Complementary and Alternative Medicine*, vol. 2021, 9 pages, 2021.
- [20] M. F. Javed, W. Nawaz, and K. U. Khan, “HOVA-FPPM: flexible periodic pattern mining in time series databases using hashed occurrence vectors and Apriori approach,” *Scientific Programming*, vol. 2021, 14 pages, 2021.