Retraction

Retracted: Application of the Data Mining Model in Smart Mobile Education

Mobile Information Systems

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1].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] G. Wang and Z. Ding, "Application of the Data Mining Model in Smart Mobile Education," *Mobile Information Systems*, vol. 2022, Article ID 3111581, 7 pages, 2022.

Hindawi

Research Article

Application of the Data Mining Model in Smart Mobile Education

Gang Wang [1] **and Zuqin Ding** [2]

^{1}*Institute of Physical Education, Soochow University, Suzhou, Jiangsu 215021, China*
^{2}*School of Mathematical Science, Huaiyin Normal University, Huaian, Jiangsu 223300, China*

Correspondence should be addressed to Gang Wang; 20148439@stu.sicau.edu.cn

In order to solve the problem that education departments urgently need intelligent and efficient information technology to deal with massive data, so as to mine valuable information for management decision-making, an overall scheme framework of education informatization is proposed. The framework takes data mining as a tool, combined with the theoretical knowledge of cloud computing, and takes the student data of a school as an example to verify the practicability of the framework system. The results are as follows: the framework clusters the students of a school. The overall level of students in cluster 1 is high, accounting for 45.07% of the total number. These kinds of students have solid basic theory, strong logical thinking ability, and excellent professional knowledge learning. The overall level of students in cluster 2 is average, accounting for 38.03% of the total number. This kind of students should pay attention to the study of professional courses and the cultivation of professional skills. The rest are students in cluster 3. These kinds of students have weak basic professional knowledge and poor thinking and logic ability. Through the accuracy before and after pruning output from the model, the accuracy before pruning is 87.32% and the accuracy after pruning is 97.37%. The noise data are eliminated in the process of pruning. The framework established in this study provides a certain decision-making basis for education and teaching and explains the feasibility and effectiveness of the application of data mining technology in educational informatization.

1. Introduction

As human society enters the information age, education also enters the information age with human society, knocking at the information age of education [1]. In the information age, the introduction of new educational technology promotes the rapid growth of educational informatization, and massive data are generated and accumulated in the database. However, understanding these data has far exceeded people's ability. In the end, a large amount of data cannot be effectively used, forming the phenomenon of information island, data explosion, but poor knowledge [2]. Facing this challenge, data mining technology came into being and showed strong vitality. It can find hidden and neglected laws and patterns from the vast ocean of data, so as to better support decision-making [3].

The mining of massive data is like a "treasure" to be developed. How to effectively integrate and manage these data, mine the laws and patterns useful for decision-making, and convert the existing management data into knowledge that can be used by managers, so as to facilitate the managers of relevant education departments to make decisions, improve the management level and school quality, and finally realize the education information reform. It is now the focus of research and discussion in the education industry, as shown in Figure 1.

2. Literature Review

Based on the background of the high level of social informatization development, in order to meet the needs of new skills training in radio and television, education reform should be carried out. Because the concept of education has always existed in the deep background of people's hearts, in order to change this concept and improve the quality of education today, we should deeply reform and cultivate

modern high-quality new knowledge, and education is a suitable method [4]. However, in the context of the rapid development of information technology, the data stored in the school media are increasing day by day, which makes the leaders of the education industry unable to obtain good and sound information when choosing resources, resulting in some good data not being installed and used, resulting in "data islands." In the modern life with the rapid development of industry, the demand for information technology is far from enough. Data applications only effectively solve the above problems [5, 6].

The purpose of using data mining in education and teaching is to find useful information from the big data collected by e-learning. The ultimate goal is to benefit all participants in the learning process, provide a basis for partial and nonpartisan decision-making in learning, and facilitate improved instructional procedures [7]. Traditional data mining is used for solutions, but rarely from the user's point of view, focusing more on technical issues such as algorithms and design models in data mining systems. Such procedures are highly demanding for professionals and usually only apply to professionals and do not involve professionals. Therefore, many companies require additional development costs for data mining and support for expertise and content [8, 9]. With the emergence and development of cloud computing, cloud platforms can start from the perspective of serving users. The concept of serving users provides a good solution for data mining, therefore, designing a data mining application platform based on cloud computing service type and using it in research will make it easier for academic leaders to use data mining to assist schools teaching and teaching management [10, 11].

The research of the data mining system in China's cloud environment begins with China mobile's data mining based on cloud computing, that is, the construction of a "big cloud" cloud computing platform [12]. At the end of 2008, China Mobile and the Institute of Technology of the Chinese Academy of Sciences jointly developed PDMiner, a data mining software based on cloud computing, which can solve many cloud computing problems. As a result, data mining applications in the cloud computing mode began to appear. Fan et al. proposed to build a powerful and scalable big data processing platform based on cloud computing through open-source technologies Hadoop and Xen and provide a practical and high-quality open source EMR platform and Haizui text data processing case [13]. Wu et al. proposed to use the Hadoop open-source platform to develop an integrated code mining algorithm based on the Apriori algorithm and finally determined the performance of the data mining platform in the cloud environment around [14]. Bardak et al. developed a data mining service architecture based on cloud computing and provided a set of detailed data mining service models in cloud environment, which laid the foundation for the design and implementation of data mining technology in cloud mode [15]. Attari et al. published a data mining service architecture based on cloud computing and provided a set of detailed data mining service models in the cloud environment, which laid the foundation
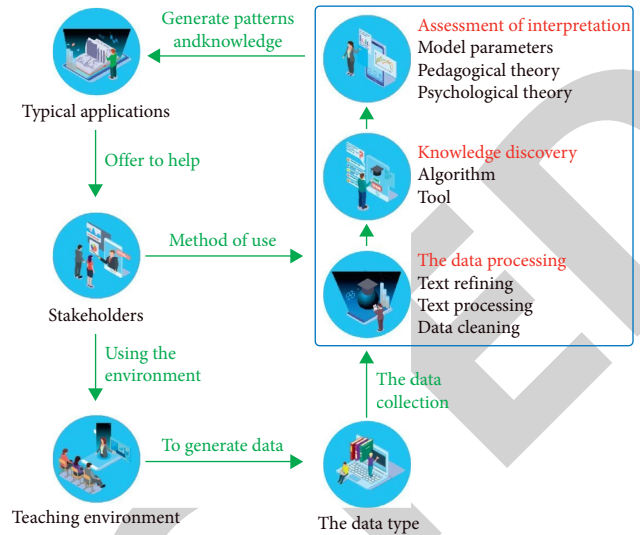

Figure 1: Data mining model.

for the design and field data mining technology in the cloud model [16].

Due to the ever-increasing demand for mining equipment, the need to develop data mining application platforms in the cloud environment is getting faster and faster. The next step in my country's research on the data mining support system in the cloud environment will focus on improving the data mining architecture and mechanism algorithm in the cloud environment [17].

Completing the application of data mining technology based on cloud computing in urban education, a large amount of important data can be found, which can not only promote the success, revision, and development of education but also provide principles. This is necessary to support the development of education and health for a variety of decision-making issues in school management. It can be seen that the content of this study is the application of data mining in teaching information. These studies have had a significant impact on increasing the use of curriculum in regular teaching at level II and in improving grade levels and achievement.

3. Research Methods

3.1. Data Mining.
Data mining refers to a complete process, which is to mine effective, unknown, and practical information from the database. Use this information to provide a certain basis for decision makers and enrich knowledge. The basic process of data mining is shown in Figure 2 [18]:

3.2. Student Characteristic Analysis Module.
The student characteristic analysis module is mainly based on the basic information and achievements of students. By analyzing the basic characteristics of students' learning, learning preferences, learning history, and professional knowledge structure, it forms a learning characteristic analysis model, classifies students' characteristics, and provides guidance for the learning of different types of students [19]. The student feature analysis module can be summarized as a clustering
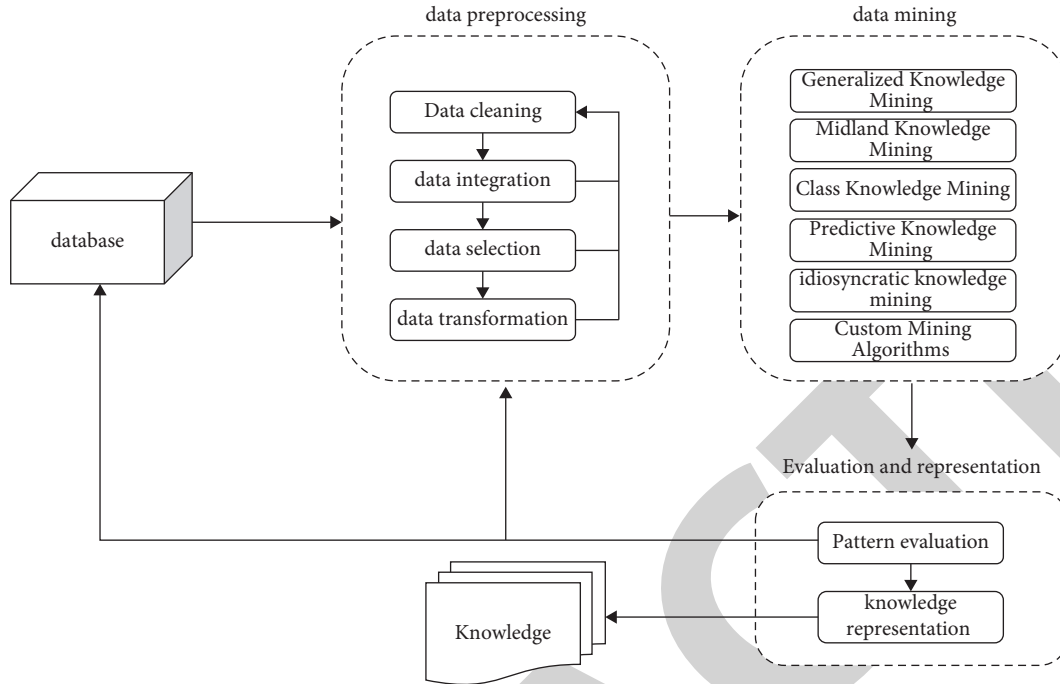
Figure 2: Data mining process.

problem. The clustering algorithm can be used to classify students and summarize the characteristics of each discipline. Because the K-means algorithm is simple and fast, it does not need to mark a large number of training tuple sets or patterns. It can adapt to changes and distinguish the useful features of different groups [20]. Therefore, the module uses the K-means algorithm and SPSS Clementine to build the model.

K-means clustering, also known as fast clustering, belongs to the partition clustering method. In the clustering results, each sample point belongs to only one class, and the clustering variables are numerical [21].

There are many methods with data sample set distance in the clustering algorithm. Because the objects processed by the K-means clustering algorithm are numerical, Euclidean distance is used to measure the difference between data samples. The Euclidean distance between data points x and y is the square root of the square sum of the difference between the values of two variables of two points [22]. The definition is as follows:

$$\text{EUCLID}(x, y) = \sqrt{\sum_{1=1}^{n} (x_i - y_i)^2}, \tag{1}$$

where x_i is the ith variable value of point x and y_i is the ith variable value of point y.

3.3. Establishment of the Student Characteristic Data Model

3.3.1. Data Import. The module is modeled by the SPSS Clementine data mining software, and the data are imported from the original set of Excel data [23].

3.3.2. Parameter Setting. According to the characteristics of students' learning courses, the K-means algorithm is used to analyze and study students' characteristics. The parameter setting adopts the default setting of the software, the maximum number of iterations is 20, and the set coding value is 0.70711, which can meet the needs of original dataset processing.

3.3.3. Determination of Data Flow. According to the K-means algorithm flow, determine the data flow, including data source, type selection, data audit, K-means model, and table output. The data flow is shown in Figure 3 [24]. Through data review, abnormal processing and missing value processing are carried out for the data, and scatter diagram and histogram output are carried out for the courses in clustering through graph nodes [25].

3.3.4. Output of Data. After using the data flow to build the model, the clustering results are output through the analysis of the SPSS Clementine software, including the proportion of various samples, the total square of samples, various variances, various mean values, and clustering results after clustering.

3.4. Establishment of the Employment Factor Data Model

3.4.1. Data Import. The results of data preprocessing are introduced into the C5.0 decision tree algorithm model of SPSS Clementine for the analysis and prediction of students' employment factors.
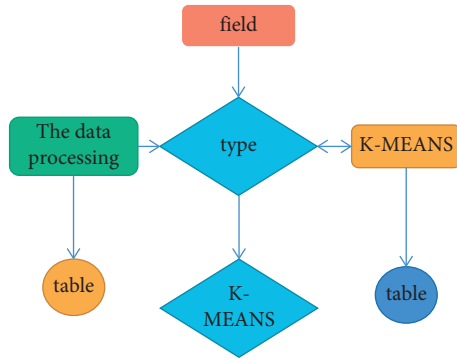
Figure 3: Data flow diagram.

3.4.2. Setting of Relevant Parameters

(1) In this model, all courses are selected as input variables and signing units are selected as output variables. Signing units are classified by course scores.

(2) In order to better analyze the employment factors of each course, the model sets the decision tree before and after pruning, and the pruning setting of the decision tree is 75%.

(3) The output results can be displayed in two forms: rule set and decision tree.

(4) In order to better evaluate the quality of the employment factor data mining model and test the accuracy of the model, the analysis field is added to the data flow.

After the above settings, the data flow of the C5.0 algorithm is realized, as shown in Figure 4:

3.5. Student Education Evaluation Module. The simulation process of this module is as follows:

(1) Index selection

This module mainly uses the scores of grade 12 e-commerce students in a school to comprehensively evaluate and analyze all students, so the indicators used are all courses of e-commerce.

(2) Data standardization

Input the data with the SPSS software and standardize the 15 course indicators of the course. The data standardization is automatically executed by the factor process of the SPSS software (the correlation judgment between indicators is omitted).

(3) Determine the number of principal components

From the correlation coefficient obtained in step 2, we know the characteristic root and variance contribution rate of the matrix. Since the contribution rate of the first five principal components is 72.825%, which can well reflect the overall index, the number of extracted principal components is 5.

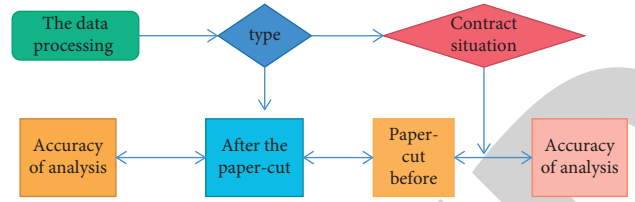(4) Expression for determining principal component function

Figure 4: C5.0 decision tree algorithm data flow.

After the principal component coefficient vector is obtained, the principal component function is calculated.

(5) Calculate the comprehensive principal component value and evaluate and analyze it

The five principal component functions calculated from the above matrix reflect different course index information, respectively, and finally obtain the following comprehensive principal component formula:

$$F = 0.0.46F_1 + 0.08F_2 + 0.068F_3 + 0.063F_4 + 0.057F_5. \quad (2)$$

4. Result Analysis

4.1. Student Characteristic Analysis Module Simulation. Figure 5 shows the course mean curve, including the average value of each course in each category. It can be seen from the course curve, the proportion and standard deviation of various samples that all kinds have been well distinguished. Only the scores of computer operation, cognition practice, ideological theory, and computer foundation of college students tend to be the same. The analysis shows that the characteristics of these courses are the obvious characteristics of short-term training, the required basic and comprehensive quality is low, and there are more subjective components in the results.

To sum up, the student characteristic analysis module obtains the following results through *K*-means analysis:

(1) The overall level of students in cluster 1 is high, accounting for 45.07% of the total number. The average scores of all subjects are more than 70. The students have solid basic theory, strong logical thinking ability, and excellent professional knowledge learning.

(2) The overall level of students in cluster 2 is average, accounting for 38.03% of the total number. The scores of each subject fluctuate up and down in 70 points, and the gap is not very obvious. The approximate curve direction is consistent with cluster 1. The personality characteristics of such students are not obvious, and the lower change is little, which is higher than cluster 3 and lower than cluster 1. Therefore, these students should pay more attention to the study of professional courses and the cultivation of professional skills. Through the study and cultivation of professional knowledge, we can better cultivate students' white confidence, strive to get
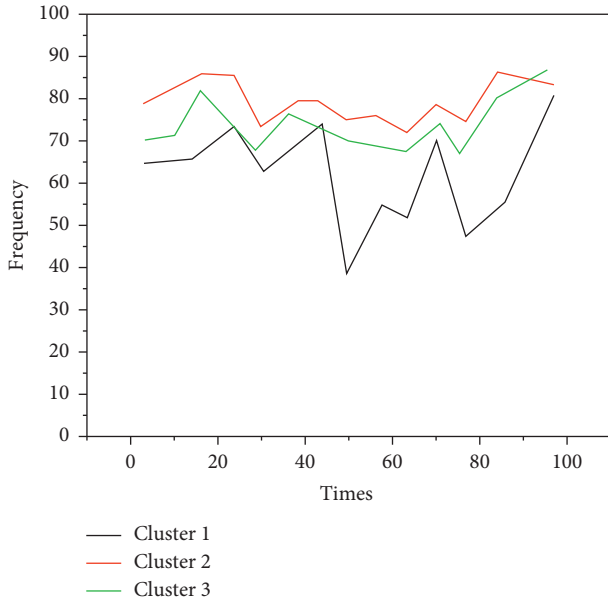
Figure 5: Various mean curves.

close to the score of cluster 1, and develop in the direction of professional technicians.

(3) The overall level of students in cluster 3 is relatively poor, accounting for 16.9% of the total number. The downward fluctuation of the curve is obvious, indicating that this kind of students have weak professional basic knowledge, poor thinking logic ability, and obvious backward academic performance. Therefore, the teaching focus of this kind of students lies in the learning of basic knowledge, the cultivation of basic professional skills, and the construction of the basic learning system.

4.2. Simulation of Employment Factor Analysis Module. Through the construction of the employment factor analysis module, the application of the employment factor data mining model is realized, which is divided into three layers, and two layers are trimmed compared with the decision tree before pruning.

Table 1 provides the accuracy before and after pruning output through the model. The accuracy before pruning is 87.32%, and the accuracy after pruning is 97.37%. The noise data are eliminated in the process of pruning.

4.3. Simulation of Student Education Evaluation Module. Combined with the standard data, we can calculate the comprehensive value of principal components of e-commerce grade 12 students in a school and sort the comprehensive principal component values. Some results are given in Table 2. It can be seen that the students with student numbers 4, 5, and 62 have higher comprehensive principal components, indicating that the comprehensive performance evaluation of these three students is high. At the same time, we can see the five principal component values affecting each student's comprehensive evaluation, analyze

Table 1: Model accuracy before and after pruning.

	Correct	Error	Accuracy (%)	Total
Before pruning	62	9	87.31	71
After pruning	67	4	97.31	

Table 2: Comprehensive evaluation of principal component analysis.

Student number	F_1	F_2	F_3	F_4	F_5	F
4	5.15	0.41	0.83	0.95	−0.13	2.51
5	4.32	0.14	−0.12	1.49	0.20	2.10
62	4.08	0.24	0.75	0.21	−0.16	1.95

different principal components, excavate the specific influencing factors of students' comprehensive evaluation, and put forward relevant solutions to specific problems, so as to provide basis for improving the overall quality level of students.

To sum up, a few independent new comprehensive evaluation indexes can be used to represent the original index variables with a large number and mutual connections, which not only avoids the mutual interference and overlap between the evaluation information but also reflects the amount of information contained in the previous indexes as much as possible, which provides a guiding basis for teaching research and management and students' comprehensive evaluation.

5. Conclusion

The rapid development of information technology has had a great impact on the development of educational informatization, which makes the relevant departments of educational institutions produce more data and promotes the continuous growth of the amount of educational information data, so that the data in the database continue to accumulate and cannot be fully utilized over time. In this context, based on the theory of data mining and cloud computing, this study puts forward the education informatization framework, instantiates some functions of the framework, realizes the application of the data mining application platform based on cloud computing service mode in education, provides a scientific decision-making basis for the education department, and becomes an indispensable part of the management decision support system.

The specific work contents are as follows:

(1) This study introduces the relevant basic theoretical knowledge of data mining, including the definition, task function, data mining technology, and data mining process of data mining, and introduces in detail the data mining algorithms used in this study, including clustering, decision tree, association rules, and principal component analysis.

(2) This study introduces the concept, service mode, and related applications of cloud computing in the field of teaching in detail. On this basis, combined with

data mining, this study constructs the framework of data mining application platform based on cloud computing service mode.

(3) According to the actual situation of educational informatization, this study puts forward the educational informatization framework, which includes the knowledge of cloud computing and data mining, and explains the design of the main function in the educational informatization framework.

(4) SPSS Clementine software is used to instantiate and simulate some functions in the educational informatization framework, including student characteristic module, school rescue management module, employment factor module, and student education evaluation module, which provides a certain decision-making basis for education and teaching and explains the feasibility and effectiveness of the application of data mining technology in educational informatization.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] Y. Jin, Y. Yang, B. Yang, and Y. Zhang, "Evaluation model of educational curriculum in higher schools based on deep neural networks," *Mobile Information Systems*, vol. 2021, no. 3, Article ID 6275096, 1–8 pages, 2021.

[2] Z. Shao, H. Sun, X. Wang, and Z. Sun, "An Optimization Mining Algorithm for Analyzing Students' Learning Degree Based on Dynamic Data," *IEEE Access*, vol. 8, no. 99, pp. 113543–113556, 2020.

[3] X. Zhang, "Evaluating the quality of internet-based education in colleges using the regression algorithm," *Mobile Information Systems*, vol. 2021, no. 1, Article ID 7055114, 9 pages, 2021.

[4] L. Liu, Y. Wang, and C. Ma, "The cultivating strategies of pre-service teachers' informatization teaching ability oriented to wisdom generation," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 16, no. 6, p. 57, 2021.

[5] A. Farhadi Esmaeily and M. Najafzadeh, "Developing a decision tree based on data mining method for detecting the influential parameters on the power of flood destruction," *Amirkabir (Journal of Science and Technology)*, vol. 53, no. 5, p. 5, 2021.

[6] H. Asami, M. Golabi, and M. Albaji, "Simulation of the biochemical and chemical oxygen demand and total suspended solids in wastewater treatment plants: data-mining approach," *Journal of Cleaner Production*, vol. 296, no. 2-4, Article ID 126533, 2021.

[7] R. Zhou, H. Chen, H. Chen, E. Liu, and S. Jiang, "Research on Traffic Situation Analysis for Urban Road Network through Spatiotemporal Data Mining: A Case Study of Xi'an, china," *IEEE Access*, vol. 9, pp. 75553–75567, 2021.

[8] H. Moayedi, M. a. M. Abdullahi, H. Nguyen, and A. S. A. Rashid, "Comparison of dragonfly algorithm and Harris hawks optimization evolutionary data mining techniques for the assessment of bearing capacity of footings over two-layer foundation soils," *Engineering with Computers*, vol. 37, no. 1, pp. 437–447, 2021.

[9] T. Yang, L. Zhang, T. Kim, Y. Hong, D. Zhang, and Q. Peng, "A large-scale comparison of artificial intelligence and data mining (ai&dm) techniques in simulating reservoir releases over the upper Colorado region," *Journal of Hydrology*, vol. 602, no. 6, Article ID 126723, 2021.

[10] E. P. Booker and G. E. Jabbour, "Antiviral nanoparticle ligands identified with datamining and high-throughput virtual screening," *RSC Advances*, vol. 11, no. 37, pp. 23136–23143, 2021.

[11] Y. Li, R. K. Shyamasundar, and X. Wang, "Special issue on computational intelligence for social media data mining and knowledge discovery," *Computational Intelligence*, vol. 37, no. 2, pp. 658-659, 2021.

[12] Y. Liu, Z. Yu, and Y. Yang, "Diabetes risk data mining method based on electronic medical record analysis," *Journal of Healthcare Engineering*, vol. 2021, no. 6, Article ID6678526, 11 pages, 2021.

[13] J. Fan, M. Zhang, A. Sharma, and A. Kukkar, "Data mining applications in university information management system development," *Journal of Intelligent Systems*, vol. 31, no. 1, pp. 207–220, 2022.

[14] B. Wu, D. Qin, J. Hu, and Y. Liu, "Experimental data mining research on factors influencing friction coefficient of wet clutch," *Journal of Tribology*, vol. 143, no. 12, pp. 1–14, 2021.

[15] S. Bardak, T. Bardak, H. Peker, E. Sözen, and Y. Çabuk, "Predicting effects of selected impregnation processes on the observed bending strength of wood, with use of data mining models," *Bioresources*, vol. 16, no. 3, pp. 4891–4904, 2021.

[16] M. Y. N. Attari, B. Ejlaly, H. Heidarpour, and A. Ala, "Application of data mining techniques for the investigation of factors affecting transportation enterprises," *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, no. 99, pp. 1–16, 2021.

[17] R. Wang, "Exploration of data mining algorithms of an online learning behaviour log based on cloud computing," *International Journal of Continuing Engineering Education and Life Long Learning*, vol. 31, no. 3, p. 371, 2021.

[18] S.-H. Liao, R. Widowati, and P. Puttong, "Data mining analytics investigate facebook live stream users' behaviors and business models: the evidence from Thailand," *Entertainment Computing*, vol. 41, no. 1, Article ID 100478, 2022.

[19] I. Parvez, J. Shen, I. Hassan, and N. Zhang, "Generation of hydro energy by using data mining algorithm for cascaded hydropower plant," *Energies*, vol. 14, no. 2, p. 298, 2021.

[20] J. Li, "Analysis of the mental health of urban migrant children based on cloud computing and data mining algorithm models," *Scientific Programming*, vol. 2021, no. 5, Article ID 7615227, 7 pages, 2021.

[21] T. Li and C. Zhang, "Research on the application of multimedia entropy method in data mining of retail business," *Scientific Programming*, vol. 2022, no. 2, Article ID2520087, 13 pages, 2022.

[22] N. Suo and Z. Zhou, "Computer assistance analysis of power grid relay protection based on data mining," *Computer-Aided Design and Applications*, vol. 18, no. 4, pp. 61–71, 2021.

[23] X. Xu, L. Li, and A. Sharma, "Controlling messy errors in virtual reconstruction of random sports image capture points

for complex systems," *International Journal of Systems Assurance Engineering and Management*, vol. 1, no. 1, 2021.

[24] D. Selva, B. Nagaraj, D. Pelusi, R. Arunkumar, and A. Nair, "Intelligent network intrusion prevention feature collection and classification algorithms," *Algorithms*, vol. 14, no. 8, p. 224, 2021.

[25] R. Huang, P. Yan, and X. Yang, "Knowledge map visualization of technology hotspots and development trends in China's textile manufacturing industry," *IET Collaborative Intelligent Manufacturing*, vol. 3, no. 3, pp. 243–251, 2021.