An Algorithm for Data Management of Higher Education Based on Fuzzy Set Theory - Association Rule Mining Algorithm

Youmeng Guan
Xingtai Open University, Xingtai 054000, China
E-mail: guayou71546@163.com

Keywords: fuzzy set, higher education, association rule, data-based management

Received: September 22, 2023

Data management can enhance the efficiency of higher education management. When combined with data mining and other technologies, it can provide a sound foundation for making management decisions. This article combined the Apriori algorithm in association rules with fuzzy theory and optimized the FCM algorithm to mine fuzzy association rules for student course grades. The results indicated that the improved FCM algorithm demonstrated a more effective clustering effect on Iris, and the outcomes were closer to the actual values. Applying this method to fuzzy association rule mining could reveal the connections between students' course grades. For instance, when students achieved excellent grades in the course of computer application fundamentals, their performance in principles of computer composition was also good. Furthermore, if they obtained excellent grades in computer application fundamentals and good grades in principles of computer composition, their grades achieved in operating systems were also excellent. The experimental results validate the reliability of the fuzzy association rule mining algorithm, which allows for the discovery of associations between different courses. Consequently, it provides valuable support for education and teaching.

Povzetek: Članek obravnava izboljšanje upravljanja visokega šolstva z upravljanjem podatkov in rudarjenjem. Uporabljen je izboljšan algoritem FCM za odkrivanje povezav med ocenami študentov.

1 Introduction

Under the influence of technological development, the field of education and teaching is increasingly shifting towards digitization and data-driven approaches [1]. However, as colleges and universities expand and their systems continue to operate, the volume of data stored in the management systems grows annually, putting significant pressure on system performance. As a result, data mining techniques have been widely adopted in various aspects, such as predicting students' grades and evaluating the quality of teaching and learning [2]. Currently, research in this field has mainly focused on predicting student performance and grades, with relatively little analysis of the interrelationships between courses. However, course arrangement and design are crucial aspects of higher education teaching arrangements. In order to understand the reliability of data mining technique for course correlation analysis, this paper studied the application of association rule mining algorithms. To enhance mining efficiency, a fuzzy association rule combined with fuzzy set theory was designed to uncover correlations between students' grades in various courses. The goal of this study is to provide theoretical guidance for organizing educational and teaching work. The study results confirm that the proposed method effectively extracts useful data from education management systems, improves data-based management efficiency, enhances decision-making and management capabilities in colleges and universities, and contributes to the overall improvement of education and teaching quality. It is also beneficial for further application of data mining technology in educational work.

2 Related works

Current research on data mining techniques in educational teaching work is presented in Table 1.

Table 1: A summary table of related works

Literature	Method	Dataset	Result
Barua h et al. [3]	The MapReduce framework based on the proposed fractional competitive multi-verse optimization-based deep neuro-fuzzy network	Performance data of students	The mean squared error, root mean squared error, and mean absolute error are .0.3383, 0.5817, and 0.3915, respectively.
Sandoval et al. [4]	A prediction model based on low-cost variables and a sophisticated algorithm	Performance data of students	They improved the model by up to 12.28% in terms of root-mean-square

			error.
Joshi et al. [5]	CatBoost - an ensemble machine learning model	Performance data of students	The accuracy is 92.27%.
Fan et al. [6]	A deep learning method of recommending MOOCs to students based on a multi-attention mechanism comprising learning records attention, word-level review attention, sentence-level review attention and course description attention	Real-world data consisting of the learning records of 6,628 students for 1,789 courses and 65,155 reviews.	MOOC platforms must fully utilize the information implied in course reviews to extract personalized learning preferences.

3 Association rule mining algorithm

3.1 Association rules

Association rule mining algorithms were initially employed in supermarket shopping to analyze customers' purchasing patterns, enabling retailers to optimize product placement and increase sales [7]. Over time, the continuous advancement of association rule mining algorithms has led to their widespread adoption in data analysis across various domains, including healthcare [8] and engineering control [9].

Association rules can be written as $X \Rightarrow Y$, meaning that there is a high probability that a record containing X will include Y (X is called the former term, and Y is called the latter term). Taking supermarket shopping behaviors an an example, suppose:

of the customers who purchased a coke, 90% purchased potato chips (30% of all customers purchased both coke and chips).

In this example, "coke" is the former term and "potato chips" is the latter term, "90%" refers to the confidence level of the rule, while "30%" refers to its support level. According to this rule, coke and potato chips can be displayed in similar positions, thus increasing sales.

Suppose that in database D, the set of all items is: $I = \{I_1, I_2, \cdots, I_n\}$, and the set of some items is: $T = \{t_1, t_2, \cdots, t_n\}$. If there is A ⊂ I, B ⊂ I, and A ∩ B ≠ ∅, then an association rule is obtained: A ⇒ B. The following rules exist in the association rule.

(1) Support level: the proportion of records in D containing both A and B in all records, which is written as:

$$supp(A \Rightarrow B) = \frac{|\{T: A \cup B \subseteq T, T \in D\}|}{|D|}. \quad (1)$$

(2) Confidence level: the proportion of records in D containing both A and B in all records containing A only, which is written as:

$$conf(A \Rightarrow B) = \frac{|\{T: A \cup B \subseteq T, T \in D\}|}{|\{T: A \subseteq T, T \in D\}|}. \quad (2)$$

(3) Frequent term set: if there is $supp\{A\} \geq minsupp$, where $minsupp$ is the minimum support level, then $\{A\}$ is called the frequent itemset.

(4) Strong rule: If there is $supp(A \Rightarrow B) \geq minsupp$ and $conf(A \Rightarrow B) \geq minconf$, where $minconf$ is the minimum confidence level, then $A \Rightarrow B$ is called a strong rule.

3.2 Apriori algorithm

The Apriori algorithm is a classical association rule mining algorithm [10]. For database D, the sequence of steps involved in the mining process of the Apriori algorithm is outlined below.

(1) Database D is traversed to obtain a one-dimensional set of candidate items.

(2) Infrequent subsets in the one-dimensional candidate item set are pruned to obtain a one-dimensional frequent item set.

(3) The one-dimensional frequent item set is self-connected to obtain the two-dimensional candidate item set.

(4) Infrequent subsets in the two-dimensional candidate item set are pruned to obtain the two-dimensional frequent item set.

(5) The k-dimensional frequent term set is self-connected to obtain the k+1-dimensional candidate term set.

(6) Infrequent subsets in the k+1-dimensional candidate item set are pruned to obtain the k+1-dimensional frequent item set.

(7) There is no new set of frequent items, and mining ends.

The mining process of the Apriori algorithm with a minimum support threshold set at 2 can be illustrated in Figure 1, taking a simple dataset as an example.

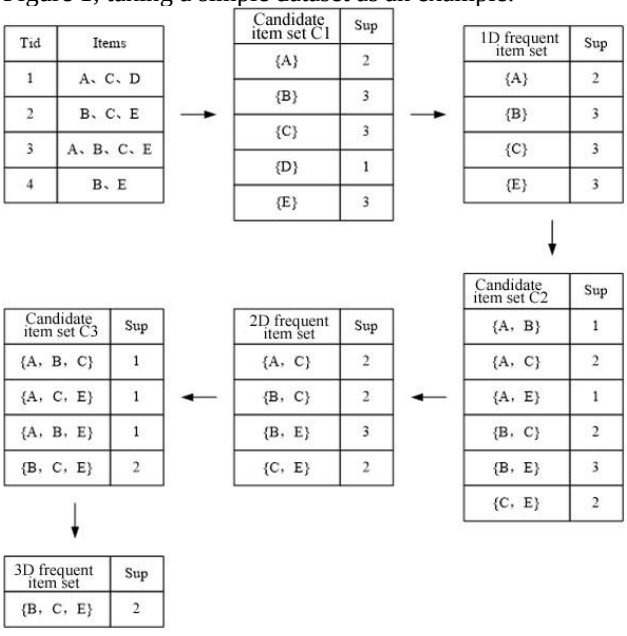

Figure 1: The mining process of the Apriori algorithm

As shown in Figure 1, the Apriori algorithm continuously scans and prunes the dataset, eventually obtaining a 3-dimensional frequent itemset {B, C, E}, at which point the algorithm terminates. However, the main drawback of the association rule in the case of the Apriori algorithm is that it can only deal with discrete data [11] and is easy to have the problem of overly hard boundary

division in the process of dividing continuous attributes into discrete intervals. For example, student grades are generally divided as follows.

0 < x < 60 = failed
60 ≤ x < 70 = pass
70 ≤ x < 80 = moderate
80 ≤ x < 90 = good
90 ≤ x ≤ 100 = excellent

According to this division, the difference between the students whose grades are 59 and 60 respectively is very small; however, they are placed in different sets. In addition, in real life, there are many non-numerical attributes that make this mining algorithm inapplicable. Therefore, to better apply association rules for mining higher education data, this paper combines them with fuzzy set theory.

4 Fuzzy association rules incorporating fuzzy set theory

4.1 Fuzzy set theory

In real life, there are many affairs without obvious boundaries, such as such as height or shortness, distance or proximity, goodness or badness, etc. Fuzzy set theory [12] introduces the concept of membership function and uses the notion of intermediate transition to realize the fuzzy processing of these affairs, which has made remarkable achievements in areas like fuzzy control [13], pattern recognition [14], and so on.

Suppose that A is a mapping from domain X to [0,1], written as:

$A: X \to [0,1]$,

then A is called the fuzzy set on X. $A(x)$ is the membership degree of the fuzzy set.

4.2 Fuzzy association rule algorithm

By combining fuzzy set theory with association rules, fuzzy association rules can be obtained [15]. Before mining the data, the attributes need to be discretized first, and the fuzzy C-mean (FCM) algorithm is used [16].

Suppose there is dataset $X = \{x_1, x_2, \cdots, x_n\}$, $j = 1,2,\cdots,n$, the purpose of the FCM algorithm is to divide X into c classes and get the clustering center set $V = \{v_1, v_2, \cdots, v_c\}$, $i = 1,2,\cdots,c$. Then, the membership degree of the j-th data belonging to the i-th class is written as u_{ij}, $u_{ij} \in [0,1]$, $\sum_{i=1}^{c} u_{ij} = 1$.

The objective function of the FCM algorithm can be written as:

$$J(U,V) = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^{m} \left\|x_j - v_i\right\|^2, \quad (3)$$

where $U = \left[u_{ij}\right]_{c\times n}$ is the membership matrix and m is the fuzzy factor.

The steps of the FCM algorithm are shown below.

(1) The clustering center is initialized.

(2) u_{ij} is calculated: $u_{ij} = \frac{1}{\sum_{r=1}^{c}(\|x_j - v_i\|/\|x_j - v_r\|)^{2/m-1}}$.

(3) The clustering center is updated: $v_i = \frac{\sum_{j=1}^{n}(u_{ij})^{m} x_j}{\sum_{j=1}^{n}(u_{ij})^{m}}$.

(4) Objective function J and the size of J in the last iteration are calculated. If J is less than or equal to termination condition ε or the specified number of iterations is reached, then it turns to (5), otherwise it returns to (3).

(5) Clustering result (V, U) is output.

After attribute discretization, assume that there is an arbitrary set of fuzzy attributes $X = \{x_1, x_2, \cdots, x_p\}$. The fuzzy support level of the i-th record in fuzzy database D_f for X is $Fsupp_i(X)$. Then, the fuzzy support level of X in D_f is:

$$Fsupp(X) = \frac{\sum_{i=1}^{n} Fsupp_i(X)}{|D_f|}. \quad (4)$$

The fuzzy association rule is written as $X_f \Rightarrow Y_f$, and its support level is written as:

$$Fsupp(X_f \Rightarrow Y_f) = \frac{\sum_{i=1}^{n} Fsupp_i(X_f \cup Y_f)}{|D_f|}. \quad (5)$$

Confidence level is written as:

$$Fconf(X_f \Rightarrow Y_f) = \frac{\sum_{i=1}^{n} Fsupp(X_f \cup Y_f)}{Fsupp(X_f)}. \quad (6)$$

Fuzzy association rules follow the same method to mine the data and get fuzzy association rules. The FCM algorithm has low complexity, is easy to implement, and is the most widely used fuzzy clustering method. However, the way of randomly determining the initial clustering center in the FCM algorithm may bring some negative effects on the results [17], making it difficult to guarantee their accuracy and determine whether the obtained optimal solution is globally optimal. In order to improve this problem and increase the reliability of fuzzy association rules, this paper uses density function to determine the initial clustering center. The FCM algorithm only considers distance measurements in its computation process. By incorporating a density function, it becomes possible to incorporate a measure of the spatial distribution of sample points. In a data, if a point is surrounded by many other data points, it means that it has a high density or is more likely to be the clustering center that reduces the effect of isolated noise point on clustering. The density function is:

$$M_j = \sum_{j=1, j\neq i}^{n} \frac{1}{\|x_i - x_j\|}, 1 \le i \le n, 1 \le j \le n. \quad (7)$$

The larger the value of M_j is, the more the data points distributed around is, and the larger the density is. The denser points are used as the initial clustering centers, while the region length is introduced to divide the cluster centers in order to avoid these points being in the same cluster:

$$D = \frac{D_{max}}{c}, \quad (8)$$

where D_{max} is the distance between the two farthest clustered samples in the dataset and c is the number of categories.

The procedure of determining the initial clustering centers by the improved FCM is shown below.

(1) The length of the region in dataset D is calculated, marking all sample points as searchable.

(2) Sample point x_i with the highest density is used as the initial clustering center.

(3) All sample points in the area where x_i locates are marked as unsearchable.

(4) It returns to step (2) for c times of iterations to find out the sample point with the highest density and use it as the initial clustering center.

By improving FCM, the influence of randomly generated initial cluster centers on clustering results can be avoided, thus preventing the selection of some isolated noise points as cluster centers and reducing result bias.

5 Student achievement analysis results

5.1 Experimental setup

The grades of 11 courses for 1,068 students from the Computer Science Department in the class of 2019 were exported from the student grade management system of Xingtai Open University for fuzzy association rule mining to eliminate incomplete values and abnormal values. Finally, the data of 1,052 students were retained, and some of them are shown in Table 2.

Table 2: Student course grades (unit: point)

Code	Course name	1	2	3	.....	1052
A	Computer application fundamentals	85	91	87	.....	65
B	Principles of computer composition	87	92	89	.....	64
C	Operating systems	77	93	86	.....	61
D	Data structure	76	89	87	.....	62
E	C programming	85	95	88	.....	69
F	Discrete number	77	85	81	.....	63
G	Software engineering	67	77	71	.....	49
H	Computer network	78	85	81	.....	55
I	Database application technology	67	78	73	.....	55
J	WEB development basics	65	77	74	.....	56
K	Computer network security technology	60	70	65	.....	48

The fuzzification results of students' course grades obtained using trapezoidal membership function are illustrated in Figure 1.

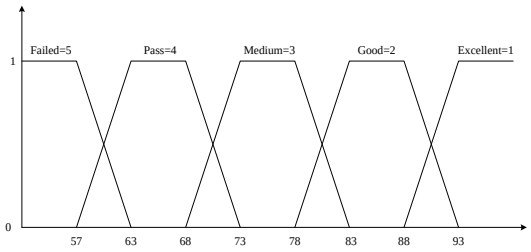

Figure 1: Fuzzification of students' course grades

The numbers 1-5 were used to assign codes to the students' course grades, corresponding to excellent-failure, respectively. After coding, the data were all written in the form of A1, B2, C3, etc., which facilitated subsequent mining of fuzzy association rules.

5.2 Clustering performance analysis

To determine the effectiveness of the improved FCM algorithm, experiments were conducted on the UCI's Iris dataset and Seeds dataset [18]. The UCI dataset is a commonly used standard benchmark for machine learning, and this article selected two subsets from it to validate the clustering performance of the improved FCM algorithm. The Iris dataset contains 150 samples. The samples were divided into three classes with four attributes, such as sepal length as displayed in Table 2. They were clustered using the traditional FCM and improved FCM algorithms, and the results were compared with the actual clustering centers given in the literature [19]. The comparison results are presented in Table 3.

Table 3: Comparison of clustering results for the Iris dataset

	Clustering	Sepal length /cm	Sepal width/ cm	Petal length /cm	Petal width/ cm
Literature [19]	1	5.00	3.42	1.46	0.24
	2	5.93	2.77	4.26	1.32
	3	6.58	2.97	5.55	2.02
Traditional FCM algorithm	1	5.37	3.46	1.54	0.33
	2	6.16	3.14	4.56	1.31
	3	7.03	3.16	5.88	1.98
Improved FCM algorithm	1	5.01	3.41	1.44	0.25
	2	5.99	2.81	4.31	1.29
	3	7.03	2.94	5.57	1.99

From Table 3, it can be found that the gap between the clustering results obtained by the traditional FCM algorithm and the actual results was large. For example, in cluster 1, the sepal length obtained by the FCM algorithm was 5.37 cm, while the actual result was 5.00 cm, with a difference of 0.37 cm. In contrast, the clustering results obtained by the improved FCM algorithm were much

closer to the actual results, with a smaller gap. The study on the Iris dataset showed that the improved FCM algorithm had better clustering results and can be applied in fuzzy association rules.

The Seeds dataset consists of 210 samples, which can be divided into three classes and seven attributes. The traditional FCM and improved FCM algorithms were used to cluster the dataset, and the results were compared with the actual clustering centers provided by the dataset. The results are presented in Table 4.

Table 4: Comparison of clustering results for the Seeds datasets

	Cluster	Area	Perimeter	Compactness	Length of kernel	Width of kernel	Asymmetry coefficient	Length of kernel groove
Actual value	1	0.76	0.79	0.69	0.73	0.77	0.37	0.76
	2	0.12	0.18	0.38	0.19	0.16	0.50	0.28
	3	0.38	0.42	0.67	0.36	0.47	0.26	0.32
Traditional FCM algorithm	1	0.71	0.69	0.72	0.75	0.71	0.39	0.67
	2	0.16	0.21	0.31	0.18	0.15	0.55	0.31
	3	0.31	0.45	0.68	0.37	0.51	0.24	0.29
Improved FCM algorithm	1	0.75	0.78	0.68	0.73	0.75	0.39	0.77
	2	0.13	0.18	0.37	0.21	0.17	0.55	0.28
	3	0.41	0.42	0.67	0.35	0.49	0.27	0.32

From Table 4, it can be observed that similar to the results of the Iris dataset, there was a significant discrepancy between the clustering results obtained by traditional FCM algorithm and the actual cluster centers. Taking Cluster 1 as an example, the traditional FCM algorithm yielded a length of kernel groove value of 0.67, which differed greatly from the actual value of 0.76. In comparison, the improved FCM algorithm achieved better alignment with the actual values in terms of clustering results. The testing conducted on two datasets demonstrated that improved FCM performed better in clustering tasks.

5.3 Fuzzy association rule analysis

Suppose min-supp = 0.3 and min-conf = 0.7. Fuzzy association rule mining was performed on students' course grades, and the partial results obtained are shown in Table 5.

Table 5: Fuzzy association rules

Former term	Latter term	Support level	Confidence level
A1	B1	0.35	0.97
J2	G2	0.36	0.95
K4	G4	0.32	0.88
A1, B2	C1	0.31	0.76
C2, D2	I1	0.33	0.85
E2, F2	I2	0.37	0.77
H1, J2	G3	0.32	0.91
I3, J3	K4	0.31	0.81
A1, C1	G1	0.33	0.84
A2, B2	K3	0.32	0.71

According to Table 5, the rules obtained are as follows.

(1) (Computer application fundamentals = excellent) ⇒ (principles of computer composition = excellent)

(2) (WEB development basics = good)⇒ (software engineering = good)

(3) (Computer network security technology = pass)⇒ (software engineering = pass)

(4) (Computer application fundamentals = excellent, computer application fundamentals = good)⇒ (operating systems = excellent)

(5) (Operating systems = good, data structures = good)⇒ (database application technology = excellent)

(6) (C programming = good, discrete mathematics = good)⇒ (database application technology = good)

(7) (Computer network = excellent, WEB development basics = good)⇒ (software engineering = moderate)

(8) (Database application technology = moderate, WEB development fundamentals = moderate) ⇒ (computer network security technology = pass)

(9) (Computer application fundamentals = excellent, operating systems = excellent)⇒ (software engineering = excellent)

(10) (Computer application fundamentals = good, principles of computer composition = good)⇒ (computer network security technology = moderate)

The above rules were analyzed. Taking rule 1 as an example, 35% of the students in the database satisfy this rule, i.e., i.e., performing excellently in both computer application fundamentals and principles of computer composition. Moreover, when a student is proficient in

computer application fundamentals, 97% of the students are also proficient in principles of computer composition. This shows that there is some commonality between the two courses and that students who are able to be proficient in computer application fundamentals are also able to be proficient in principles of computer composition.

Taking Rule 4 as an example, 31% of students in the database satisfy this rule, i.e., when students have excellent performance in computer application fundamentals and good performance in principles of computer composition, they can achieve excellent performance in operating systems. Additionally, 76% of students demonstrate excellent performance in operating systems when they have excellent performance in computer application fundamentals and good performance in principles of computer composition. This shows that computer application fundamentals and principles of computer composition serve as the foundation for understanding operating systems. If students have a good performance in these two courses, they are likely to achieve high grades in their operating systems course.

Take Rule 8 as an example, 31% of students in the database satisfy this rule, i.e., when students' performance in database application technology is moderate and their performance in WEB development fundamentals is also moderate, their performance in computer network security technology is qualified. When students have a moderate performance in database application technology and WEB development fundamentals, 81% of them demonstrate qualified performance in computer network security technology. This indicates that struggling with database application technology and WEB development fundamentals will make it more difficult for students to learn the course of computer network security technology.

Some correlations between courses can be identified according to the above rules to provide some guidance for the school's subsequent curriculum arrangement. For example, the study of computer application fundamentals and principles of computer composition should be preceded by courses in operating system and computer network security technology, in order to lay a good foundation. Similarly, courses on operating system and data structure should be arranged prior to database application technology to help students better understand the content of database application technology.

6 Discussion

The widespread application of data mining technology in educational teaching has provided scientific and reliable support for the management and decision-making of educational teaching work, effectively improving the efficiency of educational instruction. Considering the limitations of current data mining technology in course correlation analysis, this paper examined the usability of fuzzy association rule mining in higher education student course grade correlation analysis based on an association rule mining algorithm and proposed a method based on an improved FCM algorithm.

The improved FCM algorithm introduced in this paper performed better in clustering by incorporating density function design, as demonstrated through experiments on two standard test sets. The obtained clustering results were closer to the actual cluster centers of the data, indicating the reliability of the improved FCM algorithm in cluster analysis. Furthermore, when applied to fuzzy association rule mining for student course grades, this method identified certain associations between courses. For example, the performance in database application technology and web development basics will affect the grades in computer network security technology. Similarly, the performance in computer application fundamentals also has a certain impact on the grades in principles of computer organization. By analyzing these fuzzy association rules, we can understand which courses should be arranged as foundational courses earlier in the curriculum and which courses should be scheduled after completing foundational coursework. This provides theoretical support for course planners and can also be applied to student course selection systems to help students choose suitable courses for better learning outcomes, thereby avoiding a decline in interest and poor learning effectiveness caused by difficulty jumps during their studies. In general, using the fuzzy association rule mining algorithm to analyze the correlation between courses can help optimize teaching plans and improve the performance of course selection systems.

Due to the current application of data mining techniques mainly in predicting and evaluating students' performance, there has been limited research on course association rule mining. Consequently, valuable data in educational databases have not been fully utilized. This study serves as a starting point for applying fuzzy association rule mining in higher education data management, demonstrating the effectiveness of this method in analyzing course correlations. However, there is still room for improvement. For example, there are relatively few courses designed in the experiment, and a limited number of rules have been discovered. Additionally, other data mining algorithms that can be applied in this field need to be discussed. In future work, it is necessary to further improve the efficiency of mining and uncover more useful rules for education, providing guidance for educational teaching.

7 Conclusion

This paper primarily focuses on data-based management in higher education. Fuzzy association rule mining algorithms were utilized to analyze the association between students' course grades. The results indicated that the improved FCM algorithm exhibited a better clustering effect compared to the traditional FCM algorithm. Additionally, the fuzzy association rule mining based on the improved FCM algorithm yielded satisfactory outcomes. Moreover, the fuzzy association rule mining based on the improved FCM algorithm also achieved satisfactory results, revealing rules that were more meaningful and accurately reflecting the connections between courses. These findings provide a solid foundation for actual course arrangement.

References

[1] Xie B (2020). Construction of Teacher Culture in Applied Colleges under the Background of Educational Informationization. *Microprocessors and Microsystems*, 2020, pp. 103486. https://doi.org/10.1016/j.micpro.2020.103486

[2] Ma Y L, Cui C, Nie X, Yang G, Shaheed K, Yin Y (2019). Pre-course student performance prediction with multi-instance multi-label learnin. *Science China (Information Sciences)*, 62, pp. 200-205. https://doi.org/10.1007/s11432-017-9371-y

[3] Baruah A J, Baruah S (2021). Data Augmentation and Deep Neuro-fuzzy Network for Student Performance Prediction with MapReduce Framework. *International Journal of Automation and Computing*, 18, pp. 981-992. https://doi.org/10.1007/s11633-021-1312-1

[4] Sandoval A, Gonzalez C, Alarcon R, Pichara K, Montenegro M (2018). Centralized student performance prediction in large courses based on low-cost variables in an institutional context. *The Internet and Higher Education*, 37, pp. 76-89. https://doi.org/10.1016/j.iheduc.2018.02.002

[5] Joshi A, Saggar P, Jain R, Sharma M, Gupta D, Khanna A (2021). CatBoost - An Ensemble Machine Learning Model for Prediction and Classification of Student Academic Performance. *Advances in Data Science and Adaptive Analysis: Theory and Applications*, 13, pp. 1-28. https://doi.org/10.1142/S2424922X21410023

[6] Fan J, Jiang Y, Liu Y, Zhou Y (2022). Interpretable MOOC recommendation: a multi-attention network for personalized learning behavior analysis. *Internet Research: Electronic Networking Applications and Policy*, 32, pp. 588-605. https://doi.org/10.1108/INTR-08-2020-0477

[7] Prahartiwi L I, Dari W (2019). Algoritma Apriori untuk Pencarian Frequent itemset dalam Association Rule Mining. *PIKSEL Penelitian Ilmu Komputer Sistem Embedded and Logic*, 7, pp. 143-152. https://doi.org/10.33558/piksel.v7i2.1817

[8] Diamond B J, Happawana K A (2022). Association rule learning in neuropsychological data analysis for Alzheimer's disease. *Journal of Neuropsychology*, 16, pp. 116-130.

[9] Fu L, Wang X, Zhao H, Zhao H, Li M (2022). Interactions among safety risks in metro deep foundation pit projects: An association rule mining-based modeling framework. *Reliability Engineering & System Safety*, 221, pp. 1-16. https://doi.org/10.1016/j.ress.2022.108381

[10] Liu X, Sang X, Chang J, Zheng Y, Han Y (2021). The water supply association analysis method in Shenzhen based on kmeans clustering discretization and apriori algorithm. *PLoS ONE*, 16, pp. 1-21. https://doi.org/10.1371/journal.pone.0255684

[11] Erişti B, Yildirim O, Eristi H, Demir Y (2018). A New Embedded Power Quality Event Classification System Based on The Wavelet Transform. *International Transactions on Electrical Energy Systems*, 28, pp. e2597. https://doi.org/10.1002/etep.2597

[12] Zaki M J, Parthasarathy S, Li W, Ogihara M (1997). Evaluation of sampling for data mining of association rules. *Proceedings Seventh International Workshop on Research Issues in Data Engineering*, pp. 42-50.

[13] Zoghby H M E, Ramadan H S (2022). Enhanced dynamic performance of steam turbine driving synchronous generator emulator via adaptive fuzzy control. *Computers & Electrical Engineering*, 97, pp. 107666-. https://doi.org/10.1016/j.compeleceng.2021.107666

[14] Naranjo R, Santos M (2019). A fuzzy decision system for money investment in stock markets based on fuzzy candlesticks pattern recognition. *Expert Systems with Applications*, 133, pp. 34-48. https://doi.org/10.1016/j.eswa.2019.05.012

[15] Yavari A, Rajabzadeh A, Abdali-Mohammadi F (2021). Profile-Based Assessment of Diseases Affective Factors Using Fuzzy Association Rule Mining Approach: A Case Study in Heart Diseases. *Journal of Biomedical Informatics*, 116, pp. 103695. https://doi.org/10.1016/j.jbi.2021.103695

[16] Zare M, Koch M (2018). Groundwater level fluctuations simulation and prediction by ANFIS- and hybrid Wavelet-ANFIS/Fuzzy C-Means (FCM) clustering models: Application to the Miandarband plain. *Journal of Hydro-environment Research*, 18, pp. 63-76. https://doi.org/10.1016/j.jher.2017.11.004

[17] Siringoringo R, Jamaluddin J (2019). Initializing the Fuzzy C-Means Cluster Center With Particle Swarm Optimization for Sentiment Clustering. *IOP Publishing Ltd*, pp. 1-6.

[18] https://archive.ics.uci.edu/ml/datasets/Iris

[19] Bezdek J C, Hathaway R J, Sabin M J, Tucker W T (1987). Convergence theory for fuzzy c-means: Counterexamples and repairs. *IEEE Transactions on Systems Man & Cybernetics*, 17, pp. 873-877. https://doi.org/10.1109/TSMC.1987.6499296