Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 1, pp. 241–253, DOI 10.12694/scpe.v25i1.2290

CLUSTERING ALGORITHM IN DIGITAL MANAGEMENT AND SUSTAINABLE SYSTEM CONSTRUCTION FOR URBAN RAIL TRANSPORTATION STUDENT EDUCATION

YIJIA LI*

Abstract. With the rapid growth of the national economy, people's demand for transportation is becoming increasingly strong. The rail transit business is booming in large and medium-sized cities, and the education management of urban rail transit students needs further reform. At the same time, digital information technology is widely used in various fields, and digital management of education has become one of the major development directions of education reform. The study proposes a specific construction path based on the analysis of the necessity of digital management of education for urban rail transportation majors, and then optimizes the K-medoids algorithm in the clustering algorithm and validates its education digital management effect. The outcomes show that the clustering precision of the upgraded K-medoids algorithm in the selected dataset is up to 92.68%, and the running time is all below 5s, with the lowest value being 3.9s; In the digital management of urban rail transit majors in universities, the precision obtained by the algorithm is all maintained at around 95%, and the satisfaction rate is all higher than 90%. The effectiveness of the proposed method has been verified, providing a new method for the management of digital education systems for urban rail transit students. It can better understand the needs and characteristics of students, help improve their learning effectiveness and educational quality, and achieve more targeted allocation of educational resources.

Key words: Internet of Things; path optimization; time windows; dynamic demand; Clustering; K-medoids algorithm; Rail transportation; Education; Digital management cost

1. Introduction. With the acceleration of urbanization and the rapid growth of population, the construction and management of urban rail transit system is facing more and more challenges [1]. To better meet the travel needs of urban residents, improve transportation efficiency, and enhance the quality of students' education, digital management and sustainable system construction have become one of the important directions of urban rail transit development. As an important data mining technology, clustering algorithm has increasingly attracted attention in its application in digital management and sustainable system construction of urban rail transit student education. As an efficient, fast and environmentally friendly transportation mode, urban rail transit is of great significance to the travel needs of urban residents [2]. With the rapid development of information technology, digital management has become a trend in modern management. Cluster algorithm, as a commonly used data mining technique, can classify and analyze data by dividing it into groups with similar characteristics. In the application of digital management and sustainable system construction in urban rail transit student education, efficient processing technologies and methods are urgently needed due to the complexity of teaching content. At the same time, with the continuous expansion of urban rail transit system, its management is facing more and more challenges [3]. Digital management of urban rail transit education resources is an important aspect of urban rail transit management. In addition, the K-medoids algorithm in clustering algorithms has a simple principle, but tends to fall into local optima. Further improvement is needed to meet the requirements of digital education management [4]. Therefore, this study optimizes the K-medoids algorithm by introducing the Artificial Bee Colony Algorithm (ABC) algorithm and applies it to the digital management of urban students' education. Through the method proposed in this article, it is expected to use clustering algorithms to rationally allocate educational resources, improve resource utilization efficiency, promote sustainable system construction, and promote the sustainable development of urban rail student education.

*College of Railway Transportation Management and Economics, Hunan Vocational College of Railway Technology, Zhuzhou 412000, China (yj543127986@126.com)

2. Related works. Digital management of education is an inevitable trend in the development of information technology, and in recent years, breakthroughs have been made in numerous studies. Raimundo's team summarized the application of blockchain technology in higher education management. Blockchain has become an important concept at the intersection of ICT and higher education. The results show that blockchain technology is widely used in education to optimize the efficiency of educational data management, and improve the effectiveness and security of the system. It also poses certain challenges for future research directions [5]. Scholars like Mohamed have developed a qualitative model for the digital transformation of higher education and management. A decision support system is used to effectively generate and manage an index of the importance of student experience and learning expectations. The study achieves effective student management by developing a transformation roadmap for the strategic management of universities and the changes in influencing factors [6]. Williamson B suggests that the management of higher education programs is beginning to move towards online platforms and towards digital and data. The process found that the Pearson method has begun to become the definitive approach to higher education platforms. Digital management is gradually linking higher education with the benefits of digitalization [3].

Shaturaev J. has developed a causal model to mitigate the potential impact on economic and educational management as a result of the Fourth Industrial Revolution. Operating under this model, the sustainable development of economic and educational management becomes more directional. It is also recommended that economic and educational organizations adopt the model in a rational way in order to remain competitive. A reference for the subsequent educational management and development of students is provided [8]. Barakina E Y and other scholars conducted research to address the use of intelligent robots in educational management. The study found a significant relationship between artificial intelligence and the sustainability of educational management. The results show that educational management can be effective in training students with the help of intelligent methods and using the talents developed for the research of new intelligent technologies.

The factors that hinder the implementation of technology are analyzed and appropriate recommendations for the education sector are developed [10]. Syahputra Y H et al. applied the K-means clustering algorithm to classify top students to further improve their performance, in response to the difficulty of searching and processing student data. The results indicate that this method can provide schools with the information and solutions needed to classify and determine advantageous classes, thereby improving the academic performance of students in school [11]. In order to conduct Big data analysis on innovative talent education mode, Luo Y's team designed an innovative talent education mode based on segmented information fusion regression statistical analysis, and integrated educational resources through data mining and information processing. The results show that the quantitative evaluation accuracy of this method is high and the convergence is better [11].

The application of clustering algorithms in digital education management has gradually become a research hotspot in this field. The improvement of the algorithm provides necessary methods and technical references for the information management and sustainable system construction of urban rail transit students. Hu's team proposed a fuzzy-based clustering algorithm (FCAN) in order to better apply the clustering algorithm. The process was followed by combining fast fuzzy and clustering algorithms with each other as Fast Fuzzy Clustering Algorithm (F2-CAN) to solve the shortcomings of slow convergence of FCAN. Using five data sets, F2-CAN was found to be more efficient in terms of both convergence speed and clustering accuracy. It provides a reference for future solutions for large-scale complex industrial networks [12]. Researchers such as Bindhu V propose to incorporate artificial intelligence techniques to reduce the response time and cost of the system. In the course of the experiments, the use of subspace clustering is proposed to handle the connectivity and sparsity between factors. A comparison is made between IoT and traditional image-based techniques, as well as between the proposed methods, to verify the effectiveness and widespread use. The outcomes demonstrate that in the IoT environment, the research proposed method is not only effective in dealing with noise, but also that subspace clustering helps to find the desired optimal strategy [13].

Hassan B A et al. propose an architecture that utilizes evolutionary clustering algorithms in order to reduce the ambiguity of frame form contexts in order to address the problem of time-consuming generation of formal contexts in educational data corpora. The outcomes demonstrate that the quality of the semantic concept hierarchy under the proposed method of the study can be maintained at a stable 89% compared to the traditional concept lattice, and some simplification of the data is achieved. And the clustering algorithm performs the concept lattice faster than other algorithms at different filling rates [14]. Chen et al. proposed a CNN-clustering algorithm-based model for image segmentation for license plate photo recognition to monitor the reasonable use of cars. The process also uses algorithms for localization and monitoring to optimize the detection accuracy. The process used the collected dataset for simulation experiments and the results show that the model proposed by the researcher outperformed all the traditional methods [15].

Shin D et al. experimentally applied the clustering algorithm to a mathematics education course to fully understand students' behavior and thought processes during the learning process and to automatically generate reports on students' performance on classroom assignments. The results show that the clustering algorithm performed well. And it is not limited to the field of mathematics education, largely provides a reference for future science education and indicates potential research directions [16].

In summary, the clustering algorithm has demonstrated good application in the processing of massive data, while most studies have upgraded it for its convergence speed problem, but there is less analysis on the stability and classification effect of this algorithm. The proposed K-medoids algorithm introduces the Gaussian similarity function, which increases the stability of the cluster center and makes it more robust, but it has the characteristic of strong randomness in selecting the centroid of the initial cluster. Therefore, the improved ABC algorithm is further introduced to achieve global point search, and the optimal solution is the initial centroid of the K-medoids algorithm. Meanwhile, the study uses the ABC algorithm to optimize the K-medoids algorithm in the clustering algorithm and validates it in the digital management of education for urban rail transportation majors, with a view to providing technical support for improving the learning effect of students in this major and improving the digital management system of education.

3. Digital management and sustainable system construction of urban rail transportation student education based on clustering algorithm. In this chapter, the digital management and sustainable system construction of urban rail transit student education are firstly carried out, and then the digital management of education is completed by improving the K-medoids algorithm.

3.1. Digital management and sustainable system construction for urban rail transportation student education. In mass transit education, the student is the main target and the focus is on learning. The digital management of education and the construction of sustainable systems must give priority to students [17]. The digital management of existing urban railway education is a necessity and a way to build a sustainable system. Firstly, digital technology has penetrated into many fields, including education. The educational management model of universities must comply with this progressive trend. In the process of improving overall level, providing higher quality management services can continuously meet one's own development needs [18]. Second, the need for high-quality development in universities has become increasingly important. As a transactional project, the education management of urban railway professionals, although centered on students, also affects all university staff. Taking advantage of digital technology, exploring and building advanced education management models and sustainable systems has also become one of the driving forces for sustainable school progress and students' professional empowerment. Finally, as urban railway students, they must have the ability to adapt to new technologies and create a convenient and effective learning environment with the help of digital media platforms for teaching and learning to achieve better learning outcomes. The necessity and impact relationship between digital management and sustainable system construction for urban railway students' education is shown in Figure 3.1.

In Figure 3.1, the necessity of digital management of urban rail transit student education and sustainable system construction is student-centered. Under the two dotted lines centered on students, there are mainly two objective requirements, namely the penetration of digital technology and the need for high-quality development of universities. Under the urgent needs of these two aspects, a relationship diagram within the dashed box in the figure has been formed. That is to say, the development of urban rail transit vocational education and the informatization of education management are two parts, which jointly respond to the development trend of digital education management. Sharing, openness and consistency are the three principles that must be followed in the student-centred digital management of education in the urban rail transport profession. Sharing requires that the digital education management system of this profession must take the whole school education management as the goal to achieve unified management and supervision. All functional modules in the system can provide and share information in real time to optimize the effectiveness and scientific nature of education

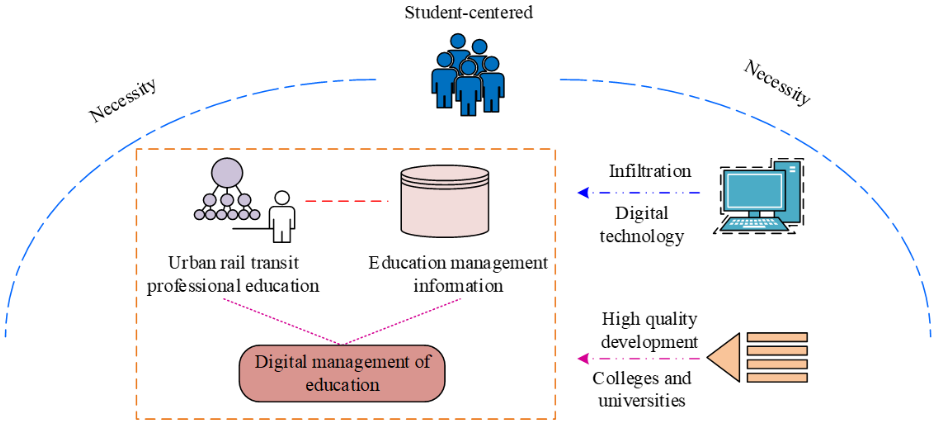


Fig. 3.1: The necessity and influence of digital management and sustainable system construction of urban rail transit students' education

management, and data security is guaranteed by identity verification and database encryption. To facilitate the construction of an education management platform and sustainable system, multiple open information modules are set up to achieve openness in a variety of integrated functions such as student outcomes enquiry, handling of various affairs and course scheduling, i.e. the principle of openness [19]. The principle of consistency requires standardized standards for the construction of educational digital platforms. By standardizing and ensuring the uniformity of educational management processes, and naming the information fields such as courses, teachers, and students in urban rail transit in a unique way, the relevance and accuracy of educational management are optimized. According to these three principles, to build a digital management and sustainable development system for urban rail transit professional education, it is first necessary to innovate digital technology in management methods. It includes scheduling professional courses through intelligent scheduling software, and integrated publishing of major events through WeChat official account. Secondly, to build a sharing platform for urban rail transportation teaching management, through digital teaching management, to achieve the sharing of quality education resources among teachers and students. The platform will also be maintained and upgraded to fully grasp the laws of teaching and coordinate the management of various departments, thus improving the decision-making level of education managers. In addition, the evaluation system of urban rail transit education management is further upgraded through multimedia technology. With the support of new media technology, education management evaluation can include multiple aspects of the subject, and the evaluation content is more social and diversified, providing a platform for urban rail transit students to showcase [20] . The path to build a digital management and sustainable system for urban rail transit student education is demonstrated in Figure 3.4.

3.2. Digital management of education based on the K-medoids algorithm. The K-medoids algorithm is based on the K-means algorithm, which uses random initialization to select the reference points for clustering [21]. Unlike the K-means algorithm, which uses the mean of the current cluster samples as the representative object, the K-medoids algorithm chooses the actual object as the representative of the cluster, thus reducing the degree to which the clustering effect is influenced by edge outliers. At the same time, the K-medoids algorithm optimizes the clustering effect by moving the samples between the division blocks, with objects contained within the same class cluster being highly similar and objects contained between different class clusters being significantly different. The K-medoids algorithm first selects a class cluster that represents the class of an actual object, and for the remaining objects, it divides them into the corresponding class clusters according to the Euclidean distance between sample points. Finally, clustering is achieved by iterating over other sample points instead of centroids, and the clustering results are compared for every two iterations,

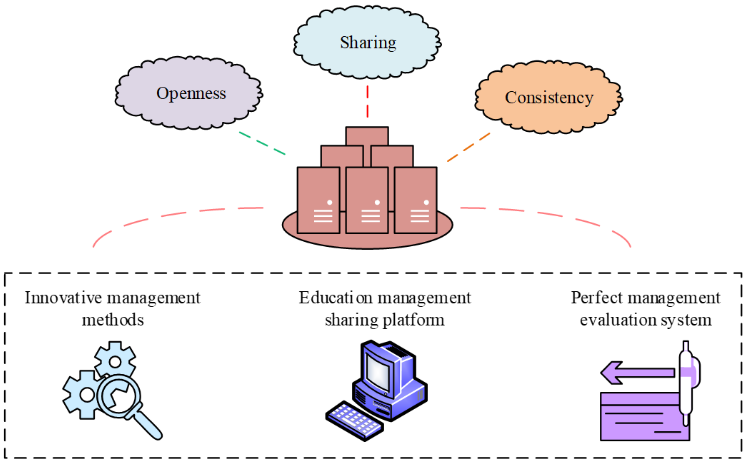


Fig. 3.2: Digital management and sustainable system construction path of urban rail transit students' education

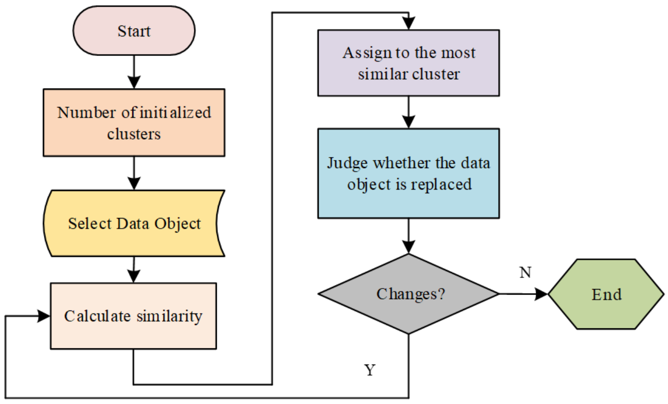


Fig. 3.3: Basic flow of K-medoids algorithm

keeping the best result as output [22]. The basic flow of the K-medoids algorithm is shown in Figure 3.3.

The spatially true distance of two sample points in the K-medoids algorithm is calculated as demonstrated in equation 3.1.

$$d(x_i, c_j) = \sqrt{\sum_{a=1}^{m} (x_{ai} - c_{aj})^2}, \quad j = 1, 2, 3...n; \quad i = 1, 2, 3, ...n \tag{3.1}$$

In equation 3.1, x_i and C_J are data objects, m represents the feature dimension, and ^{n} represents the first n data object. The mass of the cluster centroids is calculated by summing the squares of the errors, as demonstrated in equation 3.1.

$$E = \sum_{i=1}^{k} \sum_{p \epsilon c_j} d(p, c_j) \tag{3.2}$$

In equation 3.2, E denotes the data set,C_J denotes the class clusters.P is the object,and k denotes the number

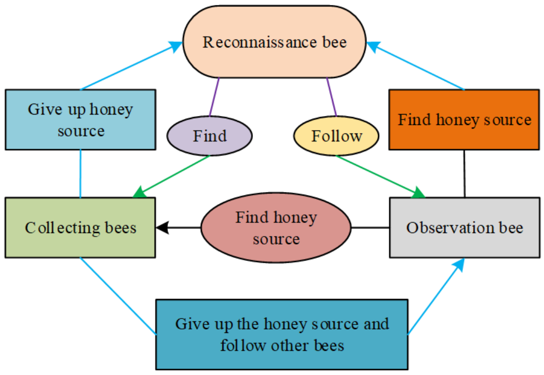


Fig. 3.4: Relationship diagram of bee species transformation of ABC algorithm

of class clusters into which the data set is divided. The K-medoids algorithm evaluates whether the quality of the clusters is optimized by a cost function as shown in equation 3.3.

$$E = E_2 - E_1 \tag{3.3}$$

In equation 3.3, E_2 is the sum of the absolute error values of all representative points in the replacement centroid data set and the new centroids of the class clusters. S is the total difference before and after replacement, and E_1 is the sum of the absolute error values of all representative points and the centroids of the class clusters before replacement is performed. If S is less than 0, the old centroid is replaced with the new centroid and the rest of the data samples are reclassified into the class cluster represented by the nearest centroid. If S is greater than or equal to 0, the current centroids are maintained. Although the K-medoids algorithm reduces the effect of noisy points and outliers on the clustering results, it still needs to mitigate the damage to the results caused by edge outliers [23]. The Gaussian similarity function can find the similarity of the data samples and increase the stability of the class cluster centers. Therefore, the study introduces Gaussian similarity to optimize the objective function of this algorithm, as shown in equation 3.4.

$$d(x_i, c_j) = e\frac{-\|x_i - x_j\|^2}{\sigma^2} \tag{3.4}$$

In equation 3.4, σ denotes the Gaussian kernel function parameter. After improving the objective function of the K-medoids algorithm by using Gaussian similarity, which is more robust in the face of isolated points, and considering its feature of random initial selection of initial clustering centroids, the study further optimizes it by means of the ABC algorithm. The ABC algorithm uses honey sources to represent the possible solutions in the sample solution space, and reflects the superiority of solutions by the good or bad honey sources, and in its description of the honey harvesting process, the bee species. The conversion is demonstrated in Figure 3.4.

In the starting phase of the ABC algorithm, the bees do not have any information about the sample, and a number of honey sources are generated by the formula demonstrated in equation 3.5.

$$X_ij = (X^j_{max}) - (X^j_{min})rand(0,1) + (X^j_{min}) \tag{3.5}$$

In equation 3.5, j represents a dimension and belongs to the D dimension.X^j_{max} is the upper limit of the searchable solution space.X^j_{min} is the lower limit, and X_ij represents the randomly obtained feasible solution space. The nectar volume of the nectar source is actually the relative size of the fitness value, which is calculated according to equation 3.6

$$fitness_i = \begin{Bmatrix} |fi|, f_1 < 0 \\ \frac{1}{f_i+1}, f_i \geq 0 \end{Bmatrix} \tag{3.6}$$

Once the bee species transformation is complete, a neighbourhood search is performed as shown in equation 3.7.

$$V_ij = rand()(X_{ij} - Xkj), i \neq k \tag{3.7}$$

In equation 3.7, V_ij represents the new nectar source found around the original source X_ij k=1,2......N_e, $j = \{1, 2, \ldots, D\}$, i and k are unequal and are obtained in a random way.rand() is a randomly selected value in [0,-1] and X_{kj} represents the nectar source location with the index value k . It is then evaluated whether the new nectar location is better than the old one, as shown in equation 3.8.

$$V_{ij}^{'} = \begin{Bmatrix} X_{ij}, f_v > f_x \\ V_{ij}, f_v \leq f_x \end{Bmatrix} \tag{3.8}$$

The observer bee decides whether to follow or not based on the fitness value passed by the nectar collecting bee, calculated by the probability obtained from equation 3.9.

$$p_i = \frac{fitness}{\sum_i fitness_i} \tag{3.9}$$

In equation 3.9, p_i represents the fitness value. In case the total number of neighborhood searches performed by any of the observation and honey collecting bees is greater than a limited number and the current nectar source location is satisfied, the honey collection at the current location is stopped and the original bee species is transformed into a scout bee, at which point the new nectar source location is found by equation 3.10.

$$X_i(n) = rand(0, 1)(X_{max} - X_{min})Basi \geq Limit + X_{min} \tag{3.10}$$

In equation 3.10,$Basi$ represents the total number of near-region searches, and $Limit$ is the finite number. The ABC algorithm runs to completion and outputs the global optimal solution when the swarm search combines other stopping criteria or reaches the maximum number of iterations. Although the ABC algorithm is simple and easy to implement with few parameters, it has obvious shortcomings such as easy premature termination and slow convergence in the later stages. Tent chaos mapping results in a flat and uniform distribution of mapping values, which can enhance swarm diversity. The IABC algorithm is based on the tent chaos mapping to obtain the initial nectar source, as shown in equation 3.11.

$$y_{i,j+1} = \begin{Bmatrix} \frac{1-x_{ij}}{1-u}, u \leq y_{ij} \leq 1 \\ \frac{y_{ij}}{u}, 0 \leq y_{i,j+1} < u \end{Bmatrix} \tag{3.11}$$

In equation 3.11, i represents the population size number.j represents the chaos number. u is the chaos parameter in the range of [0,2], and is the random number in [0,1]. The outcomeing population initialization formula for the IABC algorithm is demonstrated in equation (12).

$$X_{ij} = y_{i,j+1} \, (X_{max}^j - X_{min}^j) + X_{min}^j \tag{3.12}$$

In Eq. (12), X_{ij} denotes the randomly obtained feasible solution space. The IABC algorithm performs a binomial crossover of the global optimum with the new solution obtained from the honey bee neighbourhood search, as shown in Eq. 3.13

$$V_{ij} = \begin{Bmatrix} v_{ij,rand<cr} \\ x_j^{Global}, other \end{Bmatrix} \tag{3.13}$$

In equation 3.13, X_J^{Global} is the global optimum factor. The swarm is also constructed by adding a term to the equation, and the resulting position update equation is given by equation 3.14.

$$V_{ij} = \begin{Bmatrix} v_{ij,rand<cr} \\ x_j^{Global} + (x_j^{Global} - v_{ij}), other \end{Bmatrix} \tag{3.14}$$

Table 3.1: Software and Parameter Settings for test of IABC algorithm and ABC algorithm

Set Item	Specific Situation
Population Size	20
Maximum Number of Searches	100
Maximum Number of Iterations of Program	2000
Collecting Bees	10
Observation Bees	10
Experimental Software	MATLAB 2016

In equation 3.14, cr is the coefficient, whose main role is to coordinate the exploration and development capability of the algorithm. Finally, the K-medoids algorithm is fused with the IABC algorithm to obtain the final clustering algorithm as the IABCK-medoids algorithm. The IABCK-medoids algorithm first performs a global merit search on the dataset by utilizing the merit search advantage of the IABC algorithm, and the resulting optimal solution is the initial centroid of the K-medoids algorithm. The clustering results of the K-medoids algorithm are then passed to the IABC algorithm, which updates the swarm and iteratively updates it to achieve optimal clustering. The objective function of the algorithm is the standard absolute error formula, which is also used as a method to compute the fitness value, as shown in equation 3.15

$$fitness_i = E \tag{3.15}$$

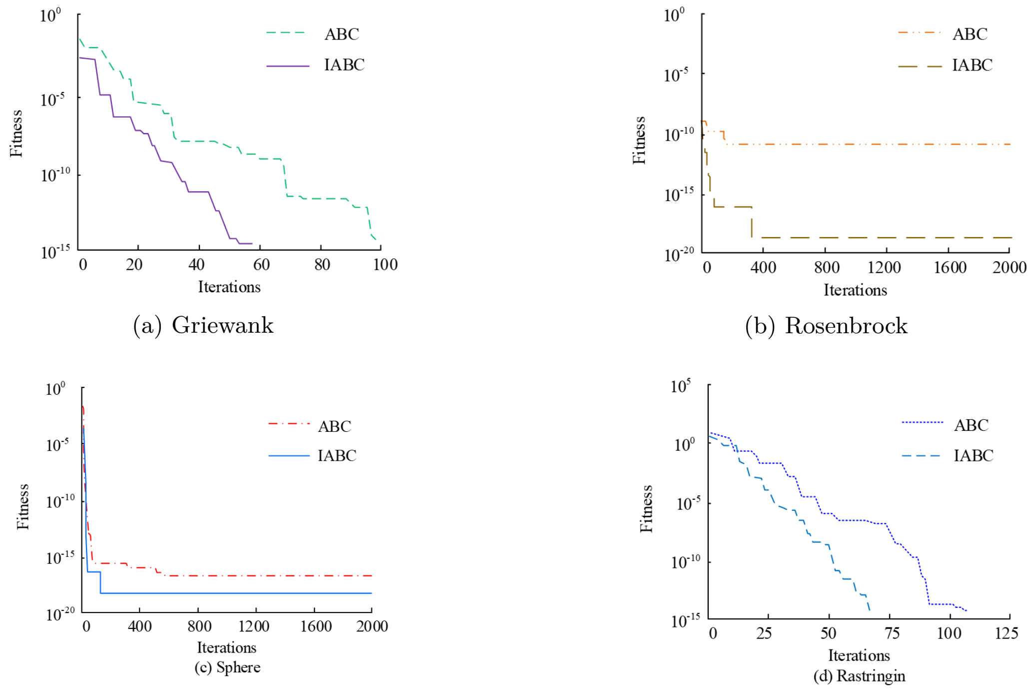
Finally, the IABCK-medoids algorithm is applied to the digital management of urban rail transportation student education, clustering student information and teaching evaluation, so as to achieve efficient management and sustainable system construction.

3.3. Improving the effectiveness of the K-medoids algorithm in the digital management of education. This chapter mainly tests the performance of the improved K-medoids algorithm, compares it with the traditional K-medoids algorithm and other algorithms, and finally verifies its practical application effect in the digital education management of urban rail transit students. The study uses the IABC algorithm to optimize the K-medoids algorithm to obtain the IABCK-medoids algorithm and apply it to the digital management and sustainable system construction of urban rail transit students' education. To analyze the application effect of this algorithm, the performance of the IABC algorithm is first verified. The IABC algorithm was compared with the ABC algorithm. To avoid overfitting and underfitting problems, the parameters were set with reference to relevant data and previous experience. The experimental software and associated parameters were set as shown in Table 3.1.

Four test functions, Sphere, Rastringin, Griewank and Rosenbrock, were selected to experiment with the IABC algorithm and the ABC algorithm, and the outcomes obtained are demonstrated in Figure 3.5. In Figure 3.5, subplots (a), (b), (c) and (d) correspond to the processing outcomes of the two algorithms in the Griewank, Rosenbrock, Sphere and Rastringin test functions, respectively. From Figure 3.5 (a) and (d), the IABC algorithm has fewer iterations and a better starting point for finding the optimum in both the Griewank and Rastringin test functions, and both achieve the optimum value. In Figure 3.5(b) and (c), for the Rosenbrock test function, which has a smaller fitness value, the convergence values of the ABC and IABC algorithms are 10-11 and 10-19, respectively; For the sphere test function, the IABC algorithm converges in about 130 iterations with a fitness value of about 10-18, while the ABC algorithm converges in about 540 iterations. Overall, the IABC algorithm is faster, can jump out of local extremes, and has a stronger search capability.

The performance of the IABCK-medoids algorithm is then verified by comparing it with the ABCK-medoids algorithm and the K-medoids algorithm. The experimental configuration is a Windows 10 64-bit operating system with a central processor of Intel Core i5-10600KF @ 4.10GHz and 8G RAM. The selected datasets were Iris, Wine, Glass, and Segmentation, and the details of the four datasets are shown in Table 3.2.

The clustering precision statistics for the three algorithms on the selected datasets are demonstrated in Figure 3.6. Figure 3.6 (a), (b) and (c) represent the clustering precision outcomes of the K-medoids, ABCK-medoids and IABCK-medoids algorithms in the four datasets, respectively. From Figure 3.6(a), the precision



(a) Griewank

(b) Rosenbrock

(c) Sphere

(d) Rastringin

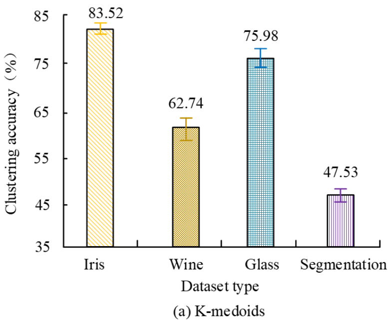
Fig. 3.5: Performance test outcomes of IABC algorithm and ABC algorithm in Sphere, Rastring, Griewink and Rosenbrook test functions

Table 3.2: Details of Iris, Wine, Glass, and Segmentation Datasets

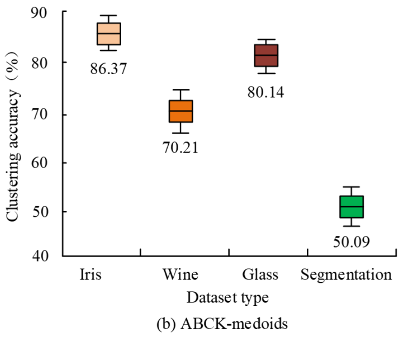
Dataset name	Number of clusters	Attribute dimension quantity	Number of objects
Iris	3	4	150
Segmentation	7	19	2310
Wine	3	13	178
Glass	6	9	214

of the K-medoids algorithm is 83.52%, 62.74%, 75.98% and 47.53% in the Iris, Wine, Glass and Segmentation datasets respectively, and the clustering precision in the dataset Segmentation is significantly lower. From Figure 3.6(b), the ABCK-medoids algorithm has the highest precision of 86.37% in the Iris dataset and the next highest precision of 80.14% in the Glass dataset. From Figure 3.6(c), the proposed IABCK-medoids algorithm achieves a clustering precision of 92.68% in the Iris dataset and still achieves a precision of over 70% in the Segmentation dataset. The comparison shows that the clustering precision of the IABCK-medoids algorithm in the Segmentation dataset is higher than the other two algorithms, with a maximum of 27.32%, and two datasets reach more than 90%, which is a clear advantage.

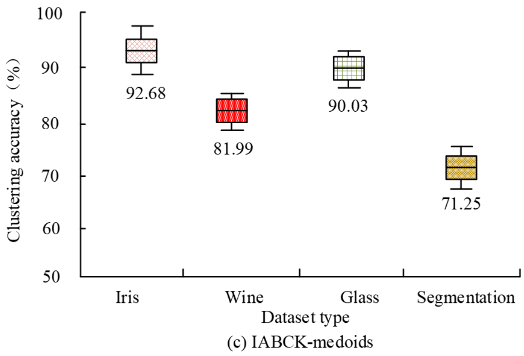
A comparison of the running times of the three algorithms on the selected datasets is shown in Figure 3.7. From Figure 3.7, the running time of the K-medoids algorithm is shorter in all four datasets, with the shortest in Iris at 1.4s and the highest in the Segmentation dataset at 2.4s. The running time of the ABCK-medoids



(a) K-medoids



(b) ABCK-medoids

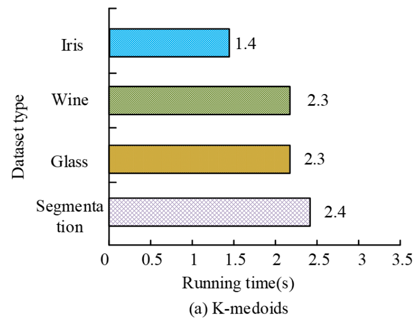


(c) IABCK-medoids

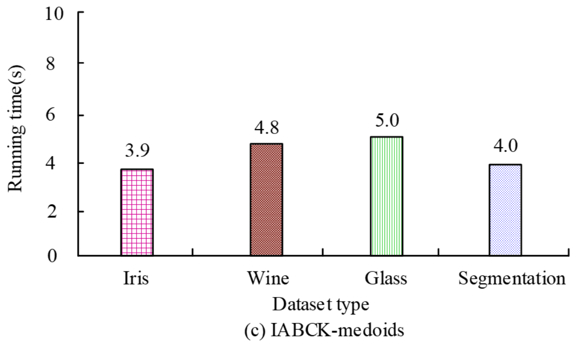
Fig. 3.6: Statistical outcomes of clustering precision of the three algorithms in the selected data set

algorithm is longer in all cases, with the highest at 9.1s and all over 7s. IABCK-medoids the IABCK-medoids algorithm has introduced the tent chaos mapping to improve the swarm diversity and the global optimality factor to achieve faster convergence of the algorithm, which inevitably prolongs the running time to some extent. However, it is still lower than the ABCK-medoids algorithm, and the accuracy is greatly improved compared to the K-medoids algorithm, with better clustering results and performance.

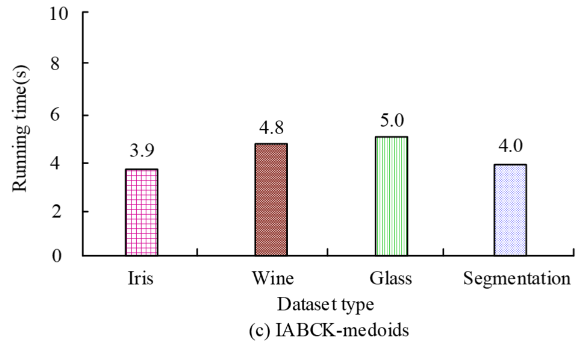
Finally, the method was applied to the digital management of urban rail transit students in a university, and the clustering effect of the three algorithms was evaluated in four aspects: student information, teaching data, course arrangement and grade management, and the usage evaluation of the four subjects: teachers, students, administrators and experts was statistically evaluated, and the obtained outcomes are demonstrated in Figure 3.8. Figure 3.8(a) demonstrates the comparison of the clustering precision of the three methods, and Figure 3.8(b) indicates the usage evaluation of the four subjects. From Figure 3.8(a), the precision of the K-medoids algorithm was generally in the range of 75% to 85%, with a minimum of 78% and a maximum of approximately 83%. The ABCK-medoids algorithm was generally in the range of 80% to 90%, with a maximum and minimum of 90% and 84%, respectively. The IABCK-medoids algorithm, on the other hand, fluctuated mostly around 95%, with a maximum of 97% and a minimum of over 90%. From Figure 3.8(b), the K-medoids algorithm is only slightly more satisfied than the AABCK-medoids algorithm in one of the four subject evaluations, while the other three are lower than the other two methods. The AABCK-medoids algorithm has a higher satisfaction rate, up to 90% in the teacher's evaluation, with an average satisfaction rate of around 85%. The IABCK-medoids algorithm, on the other hand, remained above 90%, with the highest rating of up to 96% for managers, having a better application.



(a) K-medoids

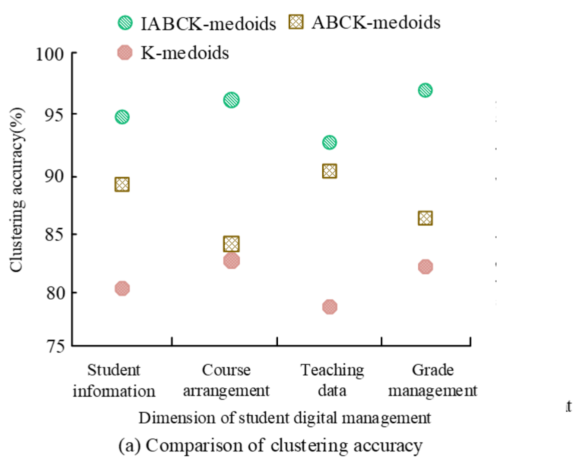


(b) ABCK-medoids

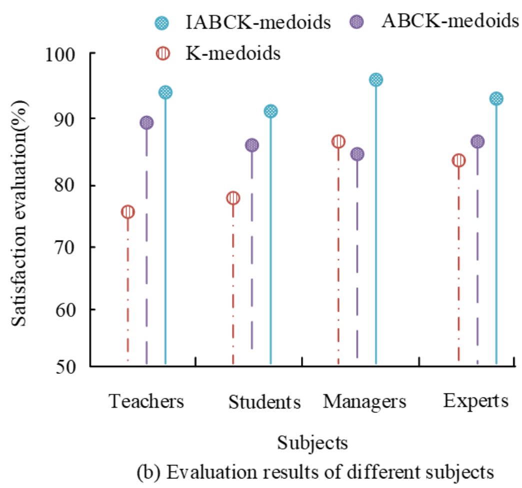


(c) IABCK-medoids

Fig. 3.7: Comparison of the running time of the three algorithms in the selected dataset



(a) Comparison of clustering accuracy



(b) Evaluation results of different subjects

Fig. 3.8: Three methods of educational digital clustering outcomes and satisfaction evaluation

4. Conclusion. The rapid development of urban rail transit scale has put forward higher requirements for the urban rail transit profession. The digital management of student education in this major has become one of the important means of talent cultivation in the rail transit industry. The study takes the urban railway students as an educational subject and uses information technology to achieve digital management; Then, the K-medoids algorithm is further optimized by improving the ABC algorithm to form the IABCK-medoids algorithm, and the algorithm is used to digitally manage education. The results show that the improved ABC algorithm converges to the Sphere test function at the 130th iteration, with an adaptive value of approximately 10-18, and converges to the Rosenbrook test function, demonstrating strong optimization ability; The IABCK medoids algorithm has a minimum accuracy of over 70% on Iris, Wine, Glass, and Segmentation datasets, which is 27.32% and 96% higher than the K-medoids and ABCK medoids algorithms. It can optimize the digital management of urban rail transit students. In conclusion, the method proposed in this study can help educational administrators to better understand the needs and characteristics of students in order to formulate targeted educational strategies. At the same time, it can help educational administrators to allocate resources rationally and improve the efficiency of resource utilization. In terms of theoretical significance, the proposed method can enrich the application scenarios of data mining in the field of student education and promote the development of data mining in the field of education. At the same time, it can help optimize the distribution and utilization of educational resources for urban rail transit students, thereby improving the sustainable development of the education system and providing theoretical support for the future development of urban rail transit student education. However, clustering algorithms need to rely on a large amount of student data and relevant feature information. If the amount of data is small or the feature information is not comprehensive, the effect of the clustering algorithm may be affected. At the same time, the clustering algorithm is sensitive to the selection of initial parameters, and different parameter settings may lead to different clustering results. Therefore, from the perspective of digital management of urban railway student education and building a sustainable system, further research will be conducted on how to combine other machine learning algorithms with clustering algorithms to improve the accuracy of student classification and resource allocation. At the same time, we can study how to use clustering algorithms to predict students' learning needs and behaviors to provide more personalized and accurate learning services. In addition, in the future, it is necessary to study how to apply clustering algorithms to optimize the operational efficiency of the urban railway student education digital management system, for example, by optimizing student grouping and resource allocation to improve the overall educational quality of schools.

5. Fundings. The research is supported by: Hunan University Student Innovation and Entrepreneurship Training Program Project; "Poetic Herbal" Humanities and Nature Notes, China (S202110541003X).

REFERENCES

[1] Decuypere, M., Grimaldi, E. & Introduction, L. Critical studies of digital education platforms. *Critical Studies In Education.* **62**, 1-16 (2021)

[2] Hakimi, L., Eynon, R. & Murphy, V. The ethics of using digital trace data in education: A thematic review of the research landscape. *Review Of Educational Research.* **91**, 671-717 (2021)

[3] Williamson, B. Education technology seizes a pandemic opening. *Current History.* **120**, 15-20 (2021)

[4] Ytre-Arne, B. & Moe, H. Folk theories of algorithms: Understanding digital irritation. *Media, Culture & Society.* **43**, 807-824 (2021)

[5] Raimundo, R. & Rosário, A. Blockchain system in the higher education. *European Journal Of Investigation In Health, Psychology And Education.* **11**, 276-293 (2021)

[6] Mohamed Hashim, M., Tlemsani, I. & Matthews, R. Higher education strategy in digital transformation. *Education And Information Technologies.* **27**, 3171-3195 (2022)

[7] Williamson, B. Making markets through digital platforms: Pearson, edu-business, and the (e) valuation of higher education. *Critical Studies In Education.* **62**, 50-66 (2021)

[8] And, S. and Management as A Result of The Fourth Industrial Revolution: An Education Perspective. *Indonesian Journal Of Educational Research And Technology.* **3**, 51-58 (2022)

[9] Barakina, E., Popova, A., Gorokhova, S., Technologies, V. & Education, A. European Journal of Contemporary Education. (2021)

[10] Syahputra, Y. & Hutagalung, J. Superior class to improve student achievement using the K-means algorithm. *Sinkron: Jurnal Dan Penelitian Teknik Informatika.* **7**, 891-899 (2022)

[11] Luo, Y. & An, Z. Research on self-learning system with "Internet+ Education" innovative talents education mode under big data background. *Computer Applications In Engineering Education*. **31**, 662-675 (2023)
[12] Hu, L., Pan, X., Tang, Z. & Luo, X. fast fuzzy clustering algorithm for complex networks via a generalized momentum method IEEE Transactions on Fuzzy Systems. (2021)
[13] Bindhu, V. & Ranganathan, G. Hyperspectral image processing in internet of things model using clustering algorithm. *Journal Of ISMAC*. **3** pp. 02 (2021)
[14] Hassan, B., Rashid, T. & Mirjalili, S. Formal context reduction in deriving concept hierarchies from corpora using adaptive evolutionary clustering algorithm star. *Complex & Intelligent Systems*. **7**, 2383-2398 (2021)
[15] Chen, J. & Zong, J. Automatic vehicle license plate detection using K-means clustering algorithm and CNN. *Journal Of Electrical Engineering And Automation*. **3**, 15-23 (2021)
[16] Shin, D. & Shim, J. systematic review on data mining for mathematics and science education. *International Journal Of Science And Mathematics Education*. **19**, 639-659 (2021)
[17] Gao, P., Li, J. & Liu, S. An introduction to key technology in artificial intelligence and big data driven e-learning and e-education. *Mobile Networks And Applications*. **26**, 2123-2126 (2021)
[18] Zaring, O., Gifford, E. & McKelvey, M. Strategic choices in the design of entrepreneurship education: an explorative study of Swedish higher education institutions. *Studies In Higher Education*. **46**, 343-358 (2021)
[19] Zhou, M., Dong, H., Zhao, Y., Ioannou, P. & Wang, F. Optimization of crowd evacuation with leaders in urban rail transit stations. *IEEE Transactions On Intelligent Transportation Systems*. **20**, 4476-4487 (2019)
[20] Yang, J. Research on Group Cooperative Learning Method Teaching Based on Urban Rail Transit Comprehensive Training Course. *International Journal Of Social Science And Education Research*. **4**, 202-206 (2021)
[21] Mashrabovich, M. The role of digital technologies in improving the quality of higher education. *ACADEMICIA: An International Multidisciplinary Research Journal*. **12**, 23-26 (2022)
[22] Dafir, Z., Lamari, Y. & Slaoui, S. survey on parallel clustering algorithms for big data. *Artificial Intelligence Review*. **54**, 2411-2443 (2021)
[23] Majhi, S. Fuzzy clustering algorithm based on modified whale optimization algorithm for automobile insurance fraud detection. *Evolutionary Intelligence*. **14**, 35-46 (2021)
[24] Bhattacharjee, P. & Mitra, P. survey of density based clustering algorithms. *Frontiers Of Computer Science*. **15**, 1-27 (2021)

Edited by: Mudasir Mohd
Special issue on: Scalable Computing in Online and Blended Learning Environments: Challenges and Solutions
Received: May 16, 2023
Accepted: Sep 1, 2023