Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 4, pp. 2476–2489, DOI 10.12694/scpe.v25i4.2872

TEACHING RESOURCE RECOMMENDATION OF ONLINE SPORTS COLLABORATIVE LEARNING PLATFORM BASED ON OPTIMIZED K-MEANS ALGORITHM

WEIGUO LI,* KE FENG,† TIANJIAO SHI,‡ AND JING HUA §

Abstract. The online collaborative learning platform for physical education is an interactive and open physical education teaching mode. To improve students' learning interest and efficiency, the online sports collaborative learning platform is designed. From the perspective of person-post matching, the role in the group is designed and the improved clustering algorithm is used to realize the grouping. The combination of the k-mean algorithm and the firefly algorithm is used to enhance the real-time and accuracy of learning resource recommendation. The outcomes demonstrated that the Firefly algorithm had obvious advantages in convergence speed and other aspects. Relative to the classical K-means algorithm and the Firefly algorithm, the average clustering accuracy of the presented algorithm was improved by 7.23 % as well as 2.18 %, and the average processing time was improved by 4.35 % and 2.26 %, respectively. In the dataset Iris, the average clustering accuracy and processing time were 91.29 and 8.65, respectively. The optimal, worst, and average values of the online collaborative learning platform on the ground of the firefly-optimized K-means algorithm were 0.3006, 3.2176, and 1.5234, respectively. The fusion algorithm proposed in this study can optimize the recommendation of teaching resources on sports online collaborative learning platforms, improve learners' learning passion, learning efficiency, and satisfaction, and relieve teachers' teaching pressure.

Key words: K-means algorithm; Firefly algorithm; Sports; Online collaborative learning platform; Cluster analysis

1. Introduction. Sports online collaborative learning is a new type of physical education teaching mode that uses the Internet as the carrier, computer network technology as the support, and integrates information technology with physical education courses. It is an open, personalized, and interactive online learning environment [1-2]. It can provide richer learning resources and more interactive methods, allowing learners to have more opportunities to communicate with other learners, share learning resources, learn experiences, and independently explore sports knowledge, skills, etc. With the continuous development of online collaborative learning platforms, a large amount of user data has been accumulated [3]. These data contain a lot of useful information, and how to discover potential knowledge and patterns from these data, to provide better and more effective online learning services for learners, is the focus of current research. With the continuous development of the internet, the application of online collaborative learning platforms in teaching is becoming increasingly widespread. Through these platforms, students can be provided with an open learning environment, enabling them to learn independently, collaborate, communicate, and learn. By mining and analyzing user data from online collaborative learning platforms, personalized learning guidance can be provided to learners, helping them better engage in sports online collaborative learning [4-5]. To increase students' interest in sports, a sports online collaborative learning platform model is designed and optimized by introducing the Firefly Algorithm (FA) to address the shortcomings of the K-means algorithm (KMA). The aim is to overcome the sensitivity of the KMA to initial centroid selection and ensure its clustering quality and accuracy.

2. Related works. In clustering analysis, the KMA is a fundamental partitioning method that can be applied to multiple data types, making it widely used in many fields. However, there are still problems such as being prone to falling into local optima. To accurately classify gas risks and improve the safety of coal mining operations, Huang et al. established a multi-factor coupling relationship analysis and warning model on the

*Basic Teaching Department, Hebei Women's Vocational College, Shijiazhuang, 050000, China
†Basic Teaching Department, Hebei Women's Vocational College, Shijiazhuang, 050000, China (Corresponding author, Ke_Feng23@outlook.com)
‡Basic Teaching Department, Hebei Women's Vocational College, Shijiazhuang, 050000, China
§Basic Teaching Department, Hebei Women's Vocational College, Shijiazhuang, 050000, China

ground of the KMA and Apriori algorithms. The model optimized the KMA through the initial clustering center, optimized the Apriori algorithm by filtering outliers in the collected dataset, and finally set the risk warning level using association rules for mining and analyzing outliers, the effectiveness of the model was verified [6]. To enhance the company's capability to match segmented customers in the market, the Li team used KMA and an adaptive particle swarm optimization algorithm to segment the customer group. Firstly, the inertia weight, learning factor, and position update method were redesigned to improve the particle swarm algorithm (PSA) to improve clustering accuracy. Then, the improved PSA was utilized for optimizing the K-means clustering center. After comparison, this method has certain effectiveness and practicality [7]. Gao and other scholars developed a multi-dimensional spatial feature vector expansion K-means model to address the problem of poor DoS detection methods and defense mechanisms. They optimized the weight of the K-means multi-dimensional feature vector through a genetic algorithm to improve the detection rate of DoS attacks. Simulation results showed that the model improved the accuracy to 96.88 % [8]. The Li team proposed a clustering model on the ground of the KMA and genetic algorithm to partition the electricity, heat, cooling, and gas loads of each building. Firstly, the KMA was utilized to uniformly analyze various indicators in the resource database, and then the genetic algorithm was used to optimize the configuration of each partition. Finally, the linear weighting method was used to sort and obtain the optimal partition configuration [9]. Wang Y and other researchers improved the KMA to address issues like long algorithm time in the clustering research of passenger hotspots. They first established a dynamically adjustable region, then used a Gaussian mixture model for data distribution statistics, and finally used the KMA for completing the clustering of various local regions. The results showcased that the algorithm could offer higher accuracy at the same time [10].

With the boost of the big data era, although there were some improvements in the personalized recommendation technology of academic resources, such as "cold start" and "data sparsity", there are still many urgent problems to be solved. Therefore, many scholars have conducted research on this. To enhance the learning performance and reduce resource consumption of online collaborative learning platforms, Han et al. developed a new effective edge learning framework for heterogeneous edges with resource constraints. By modeling the dynamic determination of collaboration strategies as an online optimization problem, they achieved a balance between performance and resource consumption. They also introduced synchronous and asynchronous learning modes to improve learning efficiency [11]. Tang J and other scholars proposed a system for optimizing English learning platforms on the ground of a collaborative filtering algorithm through in-depth analysis of recommendation algorithms for cognitive ability and difficulty, in response to the problem of poor recommendation performance on online English learning platforms. This method built a Spring Cloud platform, imported actual business data, and connected the recommendation system with the formal production system. After verification, the system design was reasonable [12]. To retain learning platform users and strengthen the competitiveness, researchers such as Xu H constructed a structural equation model of online learning platform user switching behavior on the ground of Push-Pull Mooring theory. After testing, information overload and dissatisfaction were the main influencing factors, while functional value and network externalities were secondary factors that influence user switching behavior [13].

Sun Z et al. developed an online English teaching assistance system on the ground of artificial intelligence education. This system combined deep learning with knowledge recommendation algorithms, and applied decision tree algorithms and neural networks to construct an implementation mode of applying decision tree technology in English teaching evaluation. Practice showed that this system can help students improve their learning efficiency [14].

To build an efficient sports network multimedia teaching platform, the Li W team used the black box test method to test the university sports management module and analyzed students' learning results from multiple perspectives. Finally, using the scoring method, a set of non-contradictory and non-repetitive indicators were selected. After verification, the system greatly improved the effectiveness of university sports teaching [15].

To sum up, although the existing resource recommendation technology has solved the phenomenon of academic information overload on the platform to a certain extent, it is not closely connected with the learning link, which makes it difficult for learners to actively learn with limited learning enthusiasm. To improve the cooperation efficiency of learners, meet their personalized needs in learning resources, and realize an online collaborative platform that can provide visual management and intelligent evaluation support for teachers and

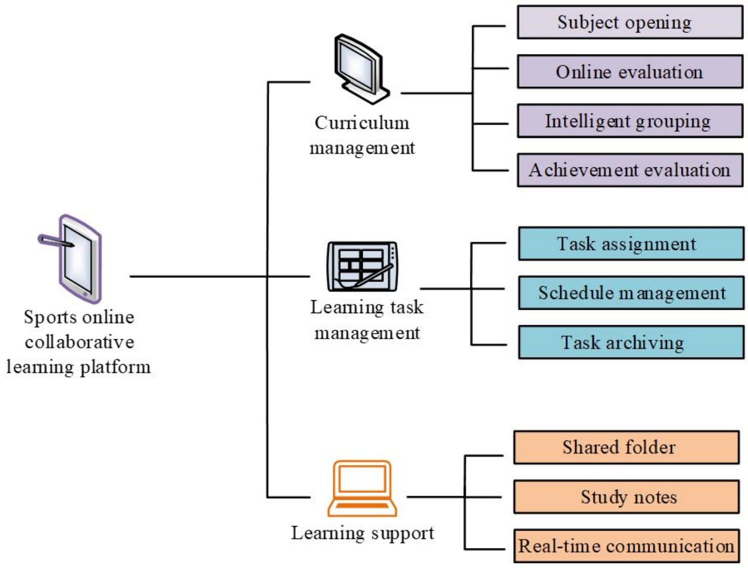


Fig. 3.1: Basic framework of sports online collaborative learning platform

learners, an online collaborative learning platform for sports is constructed. The recommendation of teaching resources is optimized by introducing the FA optimization k-mean algorithm to enhance learners' learning passion, learning efficiency, and satisfaction, and relieve teachers' teaching pressure.

3. Online collaborative learning platform construction based on FA optimized K-means algorithm. Through research and analysis of user needs, system business, and performance requirements, the overall function design of an online collaborative learning platform is completed. The main functions include learning project and task management, learning note recommendation, online assessment, and resource management, among others. Then, in response to the problem of traditional KMAs easily falling into local optima during the clustering process, the FA-optimized KMA is introduced. The optimized algorithm is used to select the initial clustering center, making the clustering results more accurate and optimizing the user experience of sports online collaborative learning platforms.

3.1. Online collaborative learning platform design for sports. The online collaborative learning platform for sports mainly utilizes methods such as course training, exam competitions, exam exercises, surveys, and training exchanges that students participate in online to track and manage their learning status and physical fitness throughout the entire process, thereby providing a comprehensive understanding of students' learning and training needs. It can also import the training, exams, and other content that students have participated in offline into the platform, forming a complete student learning and training file and providing strong data support for teachers to conduct fair evaluations and accurate analysis of student quality [16].

The research and design of a sports online collaborative learning platform mainly consists of three modules, including classroom management, learning task management, and learning support. The main functions of each module are shown in Figure 3.1. In the classroom management module, teachers can set topics, establish learning plans, provide physical education teaching guidance on the grounds of each student's learning process, and complete the writing and release of exams through the online evaluation module. After the physical education course is completed, students will also be graded. In this module, students can also see the arrangement of questions and projects and complete online exams, and at the end of the course, they can conduct self-evaluation and self-evaluation. They can also see the comprehensive scores of themselves and their team members.

The teacher arranges teaching tasks for each learning stage through the learning task management module. After the start of the physical education course, the team leader will assign learning tasks on the grounds of the

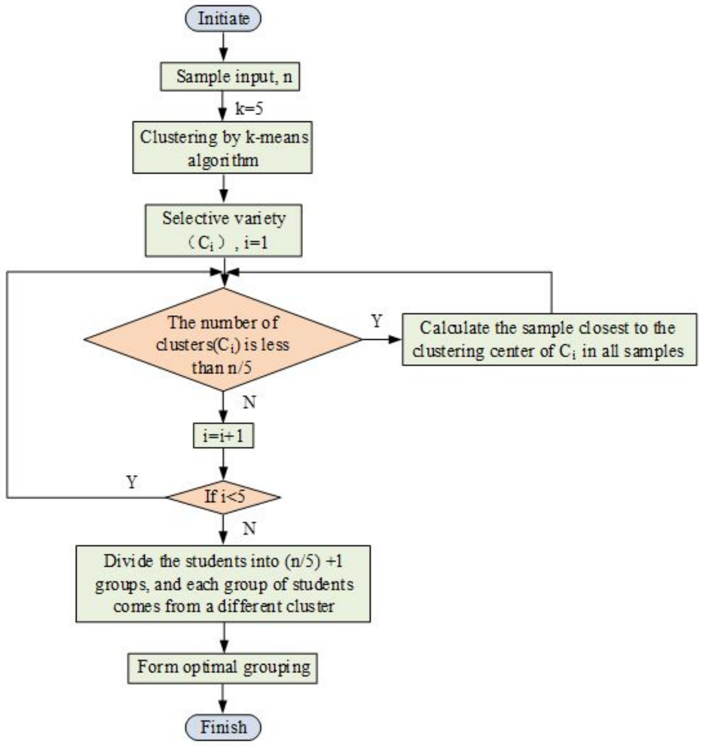


Fig. 3.2: Group working flow chart

learning abilities of the team members and set completion deadlines. Individual students can make appropriate adjustments accordingly. At the same time, teachers and students can see the learning progress of each group to enhance their positive competitive awareness and adjust their learning task arrangements promptly.

The recommendation of grouping strategies and learning resources has always been the most important issue in online collaborative learning platforms. Its main function is to help students find suitable learning companions and help them find their positioning in the learning group. On this basis, through reasonable division of labor and resource recommendations, individual learning motivation is stimulated, thereby improving overall learning efficiency and reducing the workload of teachers.

Figure 3.2 is the grouping workflow diagram of the research design. Students are classified using an improved K-means method; Then, on the grounds of the analysis of clustering results, the location of each type of student can be located. After the role type of each cluster is determined, to prevent uneven role distribution caused by large differences in the number of samples between clusters, it is necessary to recalculate the distance between all samples and the final cluster center, to adjust the samples' quantity within the cluster. Finally, according to the principle of "homogeneity between groups, heterogeneity within groups", students are randomly divided into groups, so that each student has a suitable positioning [17].

A 7-dimensional student feature model is constructed to group students, and the mathematical expression is shown in equation (3.1).

$$x = (A, B, C, D, E, F, G)\,, x \in X_m \tag{3.1}$$

In equation (3.1), A represents behavioral style, B represents knowledge acquisition ability, C represents knowledge application ability, D represents management ability, E represents interaction ability, F represents integration ability, and G represents learning level. A can be obtained through style testing, B, C, D, E, and F can be obtained by weighting and summing the dynamic data of students after logging in using the Analytic

Hierarchy Process. Firstly, five judgment matrices are established, and then feature vectors are calculated and normalized to obtain the weights of various indicators. The calculation method is showcased in equation (3.2).

$$x^* = \frac{x - \min}{\max - \min} \tag{3.2}$$

In equation (3.2), x^* represents the weight of each indicator. Then a consistency check is performed on the 5 judgment matrices. The relevant calculation is showcased in equation (3.3).

$$C.I. = \frac{\lambda_{\max} - n}{n - 1} \tag{3.3}$$

In equation (3.3), $C.I.$ represents the consistency index, $\lambda_{\max}$ represents the maximum eigenvector, and n represents the order of the matrix. Finally, the consistency ratio is calculated on the ground of the consistency index, as shown in equation (3.4).

$$C.R. = \frac{C.I.}{R.I.} \tag{3.4}$$

In equation (3.4), $C.R.$ serves as the consistency ratio and $R.I.$ represents the random consistency indicator. If $C.R.$ is below 0.1, the weight setting of the indicator is reasonable. G is obtained by weighting the static data obtained during registration with the dynamic data obtained after login, and the specific calculation method is shown in equation (3.5).

$$L = a \times \arg\_score + b \times s\_eval + c \times o\_eval + d \times score \tag{3.5}$$

In equation (3.5), $\arg\_score$ represents the average credit score, s_eval represents self-evaluation results, o_eval represents other evaluation results, $score$ represents the online setting results, a, b, c, and d are all weight coefficients.

3.2. K-means grouping and recommendation algorithm based on FA optimization. During the learning process, students' learning status reflects their characteristics. Grouping students with similar learning characteristics can better recommend suitable learning content for them. Research is on the grounds of improved clustering algorithms to group students in the classroom. By improving the accuracy of feature analysis, more optimized grouping results can be obtained. This method models students' ability attributes, locates their roles on the grounds of their characteristics, and finally groups them on the grounds of student evaluations to obtain classification results. The traditional K-means method clusters on the ground of the distance between samples, but due to its limitations in selecting the clustering center, the clustering results are often unsatisfactory, which in turn affects the grouping quality [18]. The Firefly Algorithm (FA) can greatly support the selection of cluster centers, so the research is on the ground of K-means and integrates FA.

The KMA is an unsupervised clustering algorithm that is easy to carry and possesses excellent clustering outcomes, making it extensively utilized [19]. The K-means clustering process is shown in Figure 3.3. Given the sample points T, U, V, W, X, Y, and Z, randomly set V and Z as the initial cluster centers. The clustering results after the first classification on the ground of Euclidean distance are: TVX, UXYZ. Then the above process is repeated until the cluster centers of each sample point do not change.

The KMA utilizes Euclidean distance for determining the similarity between samples [20]. Euclidean distance is the straight-line distance between two points in Euclidean space, and its calculation is shown in equation (3.6).

$$d(x_i, x_j) = \sqrt{\sum_{k=1}^{m} (x_i - x_j)^2} \tag{3.6}$$

In equation (3.6), $d(x_i, x_j)$ serves as the Euclidean distance between the sample x_i and the cluster center x_j of the cluster, m represents the spatial dimension, and k represents the cluster centers' quantity. For

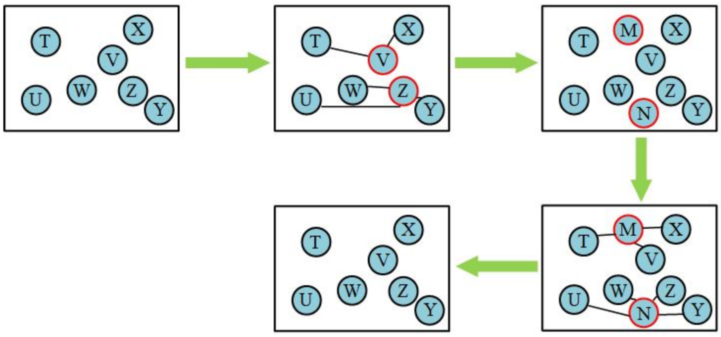


Fig. 3.3: K-means clustering process

student datasets $X = \{x_1, x_2, x_3 \ldots, x_m\}$, research randomly selects samples $u_1, u_2, \ldots, u_k$ as initial k clustering centers, divides sample points on the ground of the minimum Euclidean distance, and obtains the mathematical expression of the clustering center c^i closest to the data x_i, as shown in equation (3.7).

$$c^i = \arg\min_j \|x_i - u_j\|^2 \tag{3.7}$$

After dividing the sample points, the mean of the sample points in the k subcategory is recalculated according to equation (3.8), it is used as the new clustering center, and the above operation is repeated continuously until the clustering center no longer changes.

$$u_j = \frac{\sum_{i=1}^{n} \{c^i = j\} x^i}{\sum_{i=1}^{n} \{c^i = j\}} \tag{3.8}$$

The sample set $X = \{x_1, x_2, x_3 \ldots, x_m\}$ is divided into k clusters $\{C_1, C_2, \ldots C_k\}$ with a clustering center of $\{\mu_1, \mu_2, \ldots, \mu_k\}$, and its minimum objective function can be expressed as equation (3.9).

$$\begin{cases} J(x, \mu) = \sum_{i=1}^{5} \sum_{x_j \in C_i} d(x_j, \mu_i) \\ d(x_j, \mu_i) = \sqrt{\sum_{l=1}^{7} |x_{jl} - \mu_{il}|^2} \end{cases} \tag{3.9}$$

In the FA algorithm, there is a negative correlation between the brightness of fireflies and the value of the objective function. To decrease computational complexity and enhance the algorithm's time efficiency, this algorithm uses the reciprocal of the objective function $J(x, \mu)$ for representing the brightness of fireflies. The specific expression formula is shown in equation (3.10).

$$I(x_j) = \left(\sum_{i=1}^{5} \sum_{x_j \in C_i} \sqrt{\sum_{l=1}^{7} |x_{jl} - \mu_{il}|^2} \right)^{-1} \tag{3.10}$$

In equation (3.10), $I(x_j)$ represents the brightness of fireflies. To optimize the clustering effect, the formula for evaluating the brightness of a single firefly can be expressed as equation (3.11).

$$I(x_j) = \sum_{p \in x} dist(p, x_j)^2, j = 1, 2, \ldots, k \tag{3.11}$$

In equation (3.11), $dist(p, x_j)^2$ represents the sum of the squared distances of all data to x_j within the cluster with fireflies x_j as the cluster centroid. By studying the principle of equivalent infinitesimal substitution, the mutual attraction formula of FA has been improved to reduce computational complexity and improve computational speed. The improved mutual attraction formula can be expressed as equation (3.12).

$$\beta(d) = \frac{\beta_0}{1+\gamma d_{ij}^2} \tag{3.12}$$

In equation (3.12), β_0 serves as the maximum attraction, γ is a constant, representing the light intensity absorption factor, and d_{ij} represents the Cartesian distance from the firefly i to the firefly j. When updating the position of fireflies, inertia weights are added to the calculation formula of position updates to expand the search range of the firefly population and enhance the global optimization ability of the algorithm. The mathematical expression for updating the position of fireflies is shown in equation (3.13).

$$x_j(t+1) = \omega(t)\,x_j(t) + \beta_{ij}\left(x_i(t) - x_j(t)\right) + \alpha\,(rand - 0.5) \tag{3.13}$$

In equation (3.13), $x_j(t+1)$ represents the updated position of the firefly, t serves as the number of iterations, $rand()$ serves as the random value between [0,1], ω represents the inertia weight, α is a constant, and can generally be taken as any number within [0,1]. The calculation formula for the attraction β_{ij} of firefly i to j is equation (3.14).

$$\beta_{ij} = \beta_0 \exp\left(\gamma d_{ij}^2\right) \tag{3.14}$$

According to the characteristic of fast and then slow descent speed of the Firefly algorithm, when selecting inertia weights, a logarithmic descent strategy is used, and the current mathematical expression for weights ω_t is Equation (3.15).

$$\omega_t = \omega_{ls} - (\omega_{ls} - \omega_{le}) \log_{T_{\max}} t \tag{3.15}$$

In equation (3.15), ω_{ls} serves as the maximum value of inertia weight, ω_{le} serves as the minimum value of inertia weight, and ω_{le} serves as the maximum number of iterations.

The entire algorithm process is shown in Figure 3.4. Firstly, various parameters and the position of individual fireflies are initialized, and then the brightness of each firefly individual is counted on the ground of clustering partitioning criteria until every individual within the firefly population is selected. The brightness of each individual firefly is compared in the population. If the firefly $I(x_i)$ is larger than the firefly $I(x_j)$, the position of the firefly population is updated according to equation (3.13). Conversely, the firefly moves randomly. According to the classification principle of the nearest neighbor rule, the clustering result of the firefly is calculated, and then the new clustering center is recalculated on the ground of the current clustering result, and the position of the firefly is updated in the current population. Finally, the fitness values of fireflies are sorted, the optimal solution is found, and the optimal solution is output. Steps 3 to 6 are repeated for the firefly population that has completed a complete evolution until the stop condition is met or the maximum evolution algebra is reached.

4. Online collaborative learning platform performance verification . The study selected four different types of experimental datasets, comparing and analyzing the advantages and disadvantages of the improved KMA with other clustering algorithms. Then, through simulation analysis, the performance of the online collaborative learning platform was verified, and the grouping and recommendation effects of the platform were compared.

4.1. Simulation results of K-means algorithm based on FA optimization. This simulation was conducted on a PC with a Windows 64-bit operating system, with 4GB of physical memory and a CPU rate of 3.10GHz. The program was run using Matlab 2014b software. The maximum attraction $I(x_j)$ was set to 100, the absorption coefficient γ was set to 1, the step factor α was set to 0.06, the maximum number of iterations ω_{le} was set to 50, and the maximum fluorescence brightness I was set to 100. The four UCI datasets selected for the simulation experiment were standard experimental datasets commonly used in clustering algorithms, which were derived from the machine learning database established by Ervine at the University of California. The various information from these four datasets is demonstrated in Table 4.1.

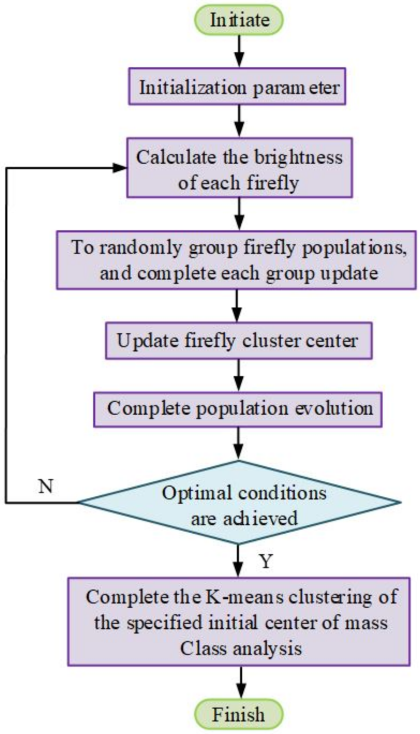


Fig. 3.4: Flowchart of K-means clustering algorithm on the ground of firefly algorithm

Table 4.1: Data set information

Data set	Number of data objects	Data dimension	Number of clusters
Iris	150	4	3
Wine	178	13	3
Hayes-Roth	162	5	3
Glass	214	9	6

Fig. 4.1: Sample contents of the four data set

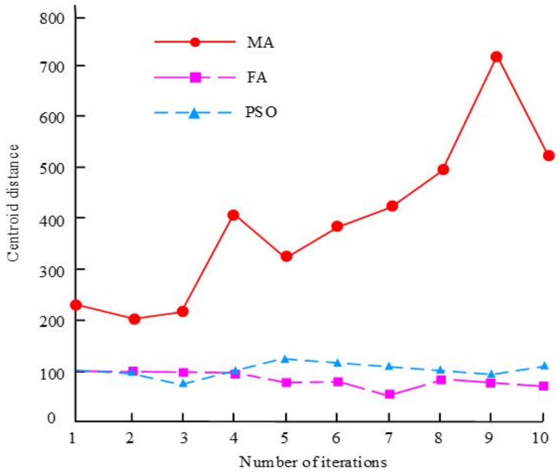


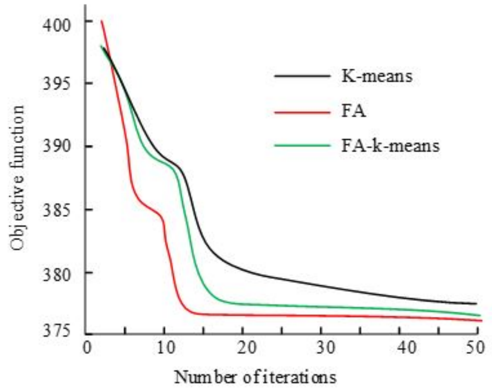
Fig. 4.2: Algorithm to actual clustering centroid distance

For verifying the possibility of the EFA algorithm breaking through local optima, the Monkey Algorithm (MA), FA algorithm, and Particle Swarm Optimization (PSO) algorithm were used in the experiment for calculating the distance between the actual clustering centroids of the Iris dataset. The outcomes are demonstrated in Figure 6. As shown in the figure, the MA algorithm had good optimization ability in the early stages of iteration. As the number of iterations increased, the distance between its centroids fluctuated significantly, indicating that the algorithm was prone to deviations during the solution process as the population evolved, leading to instability in the solution. Compared with FA, PSO had obvious advantages in convergence speed and had a gradually decreasing trend, indicating that FA can find global extremum from local extremum. The centroid distance fluctuation of PSO was greater than that of the FA algorithm, and although the overall curve was relatively stable, there was no obvious downward trend. Although PSO had a strong global search ability, it was still difficult to escape the dilemma of entering the local optimal solution.

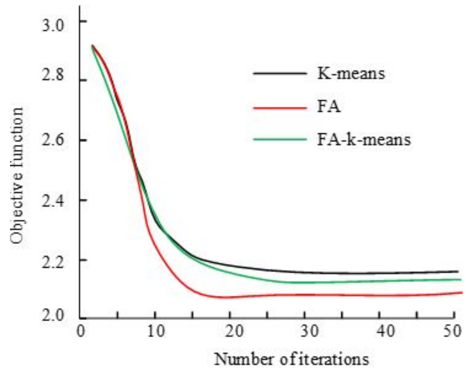
Figure 4.3 shows a comparison of the convergence curves in four datasets. The graph analysis demonstrates that when using the KMA, FA, and FA K-means selected centroids as the initial clustering centroids for analysis, the research found that although they both had certain convergence, they cannot accurately break through the local optimal dilemma. However, on the ground of the centroid selected by the FA-KMA, this clustering method not only had good stability but also had higher clustering accuracy and better convergence performance when performing clustering analysis on the specified initial centroid.

The study compared the clustering accuracy and processing time of FA KMA with KMA, and the specific outcomes are demonstrated in Table 2. The table analysis illustrates that the FA-KMA optimized the cluster center value k by using the maximum and minimum distance algorithms. Compared to the classical KMA and FA algorithms, the average clustering accuracy was improved by 7.23 % and 2.18 %, respectively, and the average processing time was improved by 4.35 % and 2.26 %, respectively. In the dataset Iris, the average clustering accuracy and processing time of the three algorithms were 91.29 and 8.65, which were better than other datasets due to their smaller dimensions. Overall, the average processing time and clustering accuracy of the FA KMA were superior to the FA algorithm and weighted KMA in all four datasets, demonstrating the effectiveness of the algorithm.

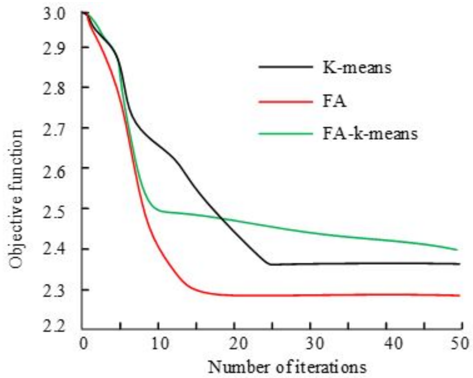
4.2. Platform performance verification and accuracy analysis of clustering results.. The study utilized Griewank, Alpine, and Salomon test functions to independently conduct 30 experiments using MA, FA, and FA K-means, respectively. Three test functions were selected with dimensions of 30 and 60, respectively. Figure 8 shows the optimization results of solving each test function once at 30 and 60 dimensions. From the graph, in the cases of 30 and 60 dimensions, FA-K-means performed better than MA and PSO for the three test



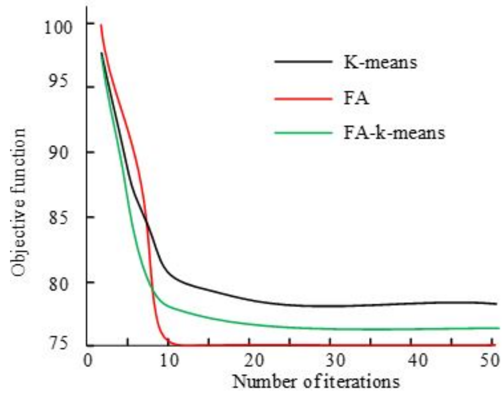
(a) Convergence Curve of the Glass data set



(b) Convergence Curve of the Hayes-Rothdata set



(c) Convergence Curve of the Wine data set



(d) Convergence Curve of the Iris data set

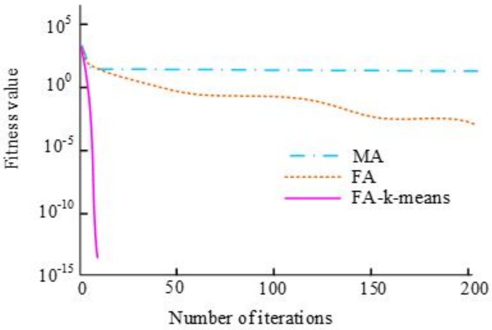
Fig. 4.3: Convergence curves of the three algorithms on six data sets

Table 4.2: Comparison of simulation results of different algorithms

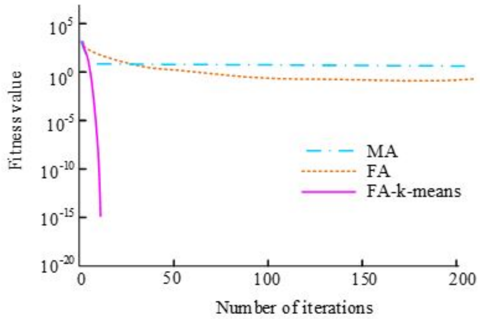
Data set		Glass	Hayes-Roth	Wind	Iris
K-means algorithms	Clustering accuracy (%)	60.72	81.47	70.44	91.50
	Processing time (s)	18.36	11.26	14.42	8.36
FA	Clustering accuracy (%)	62.47	81.27	70.70	91.20
	Processing time (s)	17.52	10.74	13.83	9.25
FA-K-means algorithms	Clustering accuracy (%)	63.14	82.36	72.15	92.15
	Processing time (s)	17.43	10.44	13.16	8.33

functions, and its convergence accuracy was significantly improved compared to other methods. Overall, the FA KMA can find the optimal solution with lower iterations and exhibit better stability for high-dimensional functions.

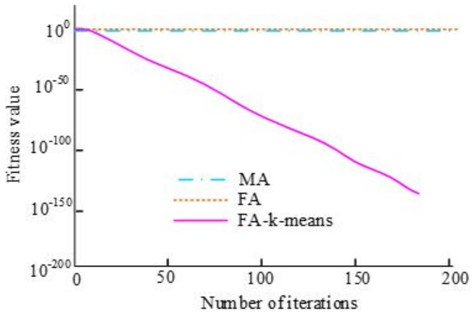
Table 4.3compares the optimization performance of the test function at 60 dimensions, using the optimal, worst, and average values of the global minimum to reflect the quality of the solution, and the standard deviation to reflect the stability of the optimal solution. The table showcases that the FA KMA had zero performance indicators in the Griewank test function, making it the best among several algorithms, while the MA algorithm



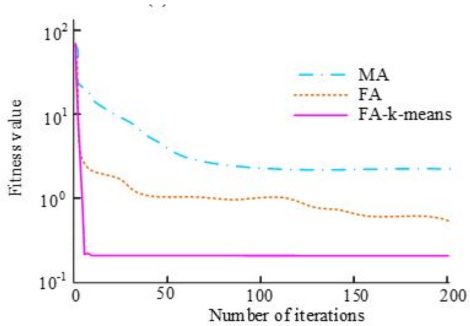
(a) 30 dimensional of Griewank function



(b) 60 dimensional of Griewank function



(c) 30 dimensional of salomon fuction



(d) 60 dimensional of salomon fuction

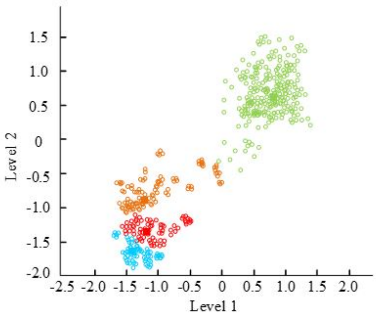
Fig. 4.4: Convergence diagram of two test functions

Table 4.3: The optimization performance comparison of the test function in 60 dimensions

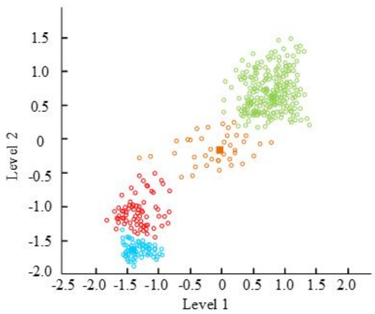
Function	Algorithm	Optimal value	Worst value	Average value	Standard deviation
Griewank	MA	24.1036	51.4285	37.4169	6.7895
	FA	0.0834	0.2215	0.1459	0.0346
	FA-K-means	0	0	0	0
Salomon	MA	0.4963	0.9045	0.7136	0.1026
	FA	2.3996	4.2015	3.2104	0.4789
	FA-K-means	0.0084	0.1995	0.1678	0.0645
Alpine	MA	13.1342	25.7154	19.8475	3.4982
	FA	72.1523	149.2457	103.9364	21.1978
	FA-K-means	0.3006	3.2176	1.5234	0.8692

had an average value of 37.4196. In the Salomon test function, the optimal, worst, and average values of the FA KMA were closest to the optimal values compared to the other three. In the Alpine test function, the optimal, worst, and average values of the FA-KMA were 0.3006, 3.2176, and 1.5234, respectively, which were the lowest relative to the two algorithms.

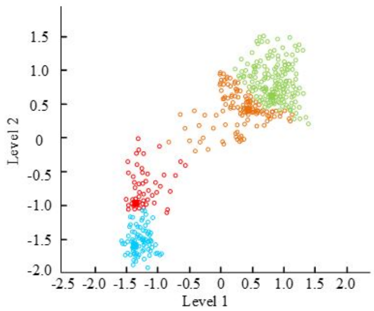
The study randomly selected 200 sample data and tested them using online collaborative learning platforms on the grounds of three different clustering algorithms. The outcomes are demonstrated in Figure 9, where the diamond denotes the cluster center, and the hollow point represents the sample points. In the figure, most of the cluster centers in the online collaborative learning platform on the ground of the KMA were relatively concentrated, but most of the sample points had a long Euclidean distance from the cluster center, indicating



(a) The clustering result Of K-means algorithm



(b) The clustering result Of FA algorithm



(c) The clustering result Of FA K-means algorithm

Fig. 4.5: Comparison of clustering results of the three algorithms

Table 4.4: Grouping and recommendation effect of online collaborative learning platform

Data set		Glass	Hayes-Roth	Wind	Iris
C-OCLP	Optimal value	729.65	101.33	162.63	80.67
	Worst value	1095.63	145.63	210.63	152.34
	Average value	912.64	123.48	186.63	116.51
	Standard deviation	81.52	21.58	23.45	17.64
P-OCLP	Optimal value	689.21	123.36	172.45	80.62
	Worst value	1140.36	168.77	250.36	153.23
	Average value	914.79	146.07	211.41	116.93
	Standard deviation	115.63	26.36	33.25	22.36
F-OCLP	Optimal value	600.68	136.55	196.24	80.06
	Worst value	854.23	186.41	263.47	152.12
	Average value	727.46	161.48	229.86	116.09
	Standard deviation	72.63	24.69	28.14	22.35

that the platform's recommendation accuracy for learning content needed to be improved. Compared with learning platforms on the ground of KMA, the recommendation accuracy of the FA method was improved, but the quality still needed further improvement; The learning platform on the ground of the FA-KMA had the minimum Euclidean distance between each sample point and its cluster center, which greatly improved the clustering results of the learning platform, thereby improving the efficiency of group clustering and the recommendation accuracy of learning content.

For testing the grouping and recommendation performance of the online collaborative learning platform proposed by the research institute, simulation analysis was conducted to compare the optimal, worst, average, and standard deviation of the learning platform on the ground of the FA KMA (F-OCLP), the learning platform on the ground of the Cuckoo KMA (C-OCLP), and the learning platform on the ground of the PSO-KMA (P-OCLP). The specific results are shown in Table 4.4. The data results in the table illustrate that the F-OCLP data results had better clustering quality compared to the other two learning platforms. In the Iris dataset, the optimal values for F-OCLP, C-OCLP, and P-OCLP were similar, with 80.67, 80.62, and 80.06, respectively, but F-OCLP was more stable. On the Wind dataset, the optimal value of F-OCLP was 263.47, and the optimization effect was significantly improved compared to the other three methods. In the Glass dataset, the optimal, worst, and average values of F-OCLP were significantly improved compared to the other three. In the Hayes-Roth dataset, the optimal, worst, and average values of F-OCLP were 136.55, 186.41, and 161.48, respectively, which was the highest compared to the other two learning platforms.

Table 4.5: Performance summary of FA-k-means algorithm

	Data set	Glass	Hayes-Roth	Wind	Iris
Algorithm performance	Clustering accuracy(%)	63.14	82.36	72.15	92.15
	Processing time(s)	17.43	10.44	13.16	8.33
Platform performance	Optimal value	600.68	136.55	196.24	80.06
	Worst value	854.23	186.41	263.47	152.12
	Average value	727.46	161.48	229.86	116.09
	Standard deviation	72.63	24.69	28.14	22.35

Table 4.5 summarizes the performance of the teaching resource recommendation algorithm of the sports online collaborative learning platform based on optimized k-mean. In the table, the research improved the traditional clustering algorithm by integrating the firefly algorithm and inertia weight and improved the clustering accuracy. Based on the analysis of clustering results, learners were automatically grouped and assigned roles. Fully combined with the real-time needs of learners and the characteristics of learning notes in the learning process of the sports online collaborative learning platform, learners' self-assessment and mutual assessment were added to the evaluation function, and learners' learning behavior data were obtained to automatically calculate the normal score, reduce the proportion of teachers' scoring, and make the results more objective and accurate by reducing the subjectivity of evaluation.

5. Conclusion. In response to the phenomenon of some students not acting and learning resource overload in the current online collaborative learning platform for sports, this study proposes a K-means grouping and recommendation algorithm on the ground of FA optimization, which uses random weight factors to influence the iteration of firefly positions and improves the random disturbance term to improve the diversity of the population. Last, simulation experiments verify the effectiveness. The results showed that FA had obvious advantages in convergence speed and a gradually decreasing trend, indicating that FA can find global extremum from local extremum. Compared with classical KMA and FA algorithms, the average clustering accuracy of FA-KMA was increased by 7.23 % and 2.18 %, respectively, and the average processing time was increased by 4.35 % and 2.26 %, respectively. In the dataset Iris, the average clustering accuracy and processing time were 91.29 and 8.65, while the FA-KMA had zero performance in several indicators of the Griewank test function. The average value of the MA algorithm was 37.4196. In the Alpine test function, the optimal, worst, and average values of the FA-KMA were 0.3006, 3.2176, and 1.5234, respectively. The FA K-means method minimized the Euclidean distance between each sample point and its cluster center. In the Iris dataset, the optimal values for F-OCLP, C-OCLP, and P-OCLP were 80.67, 80.62, and 80.06, respectively. On the Wind dataset, the optimal value for F-OCLP was 263.47, and on the Glass dataset, the optimal value for F-OCLP was 854.23. The optimal, worst, and average values of F-OCLP in the Hayes-Roth dataset were 136.55, 186.41, and 161.48, respectively. The application efficiency of the fusion algorithm proposed in this paper is low, and an adaptive mechanism will be introduced into the combination algorithm in the next step. In future work, research should be done on the implementation of diversified applications of the algorithm. To avoid the situation that the accuracy of the score results is affected by malicious bad reviews, the suspicious degree will be integrated into the next step, the evaluation with too large a score difference will be filtered out, and the accuracy of learning effect evaluation will be further improved.

REFERENCES

[1] Usman, A. & Abdullah, M. An Assessment of Building Energy Consumption Characteristics Using Analytical Energy and Carbon Footprint Assessment Model. *Green And Low-Carbon Economy.* **1**, 28-40 (2023)

[2] Ammari, A., Labidi, W., Mnif, F., Yuan, Z. & Sarrab, M. Firefly algorithm and learning-based geographical task scheduling for operational cost minimization in distributed green data centers. *Neurocomputing.* **490** pp. 146-162 (2022)

[3] Wang, Y., Liu, Y. & Feng, W. Waste Haven Transfer and Poverty-Environment Trap: Evidence from EU. *Green And Low-Carbon Economy.* **1**, 41-49 (2023)

[4] Gazman, V. A New Criterion for the ESG Model. *Green And Low-Carbon Economy.* **1**, 22-27 (2023)

[5] Choudhuri, S., Adeniye, S. & Sen, A. Distribution Alignment Using Complement Entropy Objective and Adaptive Consensus-Based Label Refinement for Partial Domain Adaptation. *Artificial Intelligence And Applications*. **1**, 43-51 (2023)

[6] Huang, Y., Fan, J., Yan, Z., Li, S. & Wang, Y. Research on Early Warning for Gas Risks at a Working Face Based on Association Rule Mining. *Energies*. **14**, 77-79 (2021)

[7] Li, Y., Qi, J., Chu, X., Tian, D., Feng, J. & Mu, W. Customer Segmentation Using K-means Clustering and the Hybrid Particle Swarm Optimization Algorithm. *The Computer Journal*. **113**, 136-158 (2022)

[8] Gao, L., Bian, Z. & Maode, M. Research on DoS Attacks Intrusion Detection Model Based on Multi-Dimensional Space Feature Vector Expansion K-means Algorithm. *IEICE Transactions On Communications*. **104**, 1377-1385 (2021)

[9] Li, Y., Liu, C., Zhang, L. & Sun, B. A partition optimization design method for a regional integrated energy system based on a clustering algorithm. *Energy*. **219**, 119562-119575 (2021)

[10] Wang, Y. & Ren, J. Taxi Passenger Hot Spot Mining Based on a Refined K-means++ Algorithm. *IEEE Access*. **9**, 66587-66598 (2021)

[11] Han, Q., Yang, S., Ren, X., Zhao, C. & Yang, X. OL4EL: Online Learning for Edge-Cloud Collaborative Learning on Heterogeneous Edges with Resource Constraints. *IEEE Communications Magazine*. **58**, 49-55 (2020)

[12] Tang, J. Optimization of English Learning Platform Based on a Collaborative Filtering Algorithm. *Complexity*. **6**, 45-59 (2021)

[13] Xu, H., Wang, J., Tai, Z. & Lin, H. Empirical Study on the Factors Affecting User Switching Behavior of Online Learning Platform Based on Push-Pull-Mooring Theory. *Sustainability*. **13**, 102-118 (2021)

[14] Sun, Z., Anbarasan, M. & Kumar, D. Design of online intelligent English teaching platform based on artificial intelligence techniques. *Computational Intelligence*. **37**, 1166-1180 (2020)

[15] Li, W. & Fan, X. Construction of Network Multimedia Teaching Platform System of College Sports. *Mathematical Problems In Engineering*. **15**, 703-714 (2021)

[16] Jain, K. & Saxena, A. Simulation on supplier side bidding strategy at day-ahead electricity market using ant lion optimizer. *Journal Of Computational And Cognitive Engineering*. **2**, 17-27 (2023)

[17] Muhiuddin, G., Mahboob, A. & Elnair, M. A new study based on fuzzy bi-Γ-ideals in ordered-Γ-semigroups. *Journal Of Computational And Cognitive Engineering*. **1**, 42-46 (2022)

[18] Yang, Y. & Song, X. Research on face intelligent perception technology integrating deep learning under different illumination intensities. *Journal Of Computational And Cognitive Engineering*. **1**, 32-36 (2022)

[19] Long, X., Chen, Y. & Zhou, J. Development of AR Experiment on Electric-Thermal Effect by Open Framework with Simulation-Based Asset and User-Defined Input. *Artificial Intelligence And Applications*. **1**, 52-57 (2023)

[20] Islam, A., Othman, F. & Sakib, N. Prevention of Shoulder-Surfing Attack Using Shifting Condition with the Digraph Substitution Rules. *Artificial Intelligence And Applications*. **1**, 58-68 (2023)

Edited by: Mudasir Mohd
Special issue on: Scalable Computing in Online and Blended Learning Environments: Challenges and Solutions
Received: Nov 17, 2023
Accepted: May 6, 2024