International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 4300-4311

https://doi.org/10.70517/ijhsa464278

Design of Athletes' Fitness Improvement Paths in the Framework of Intelligent Algorithm Optimization in New Era Sports Training

Yang Sun[1] and Jingsi Zhou[2,*]
[1] College of Physical Education and Health Science, Chongqing Normal University, Chongqing, 401331, China
[2] College of Physical Education, Wuhan Vocational College of Software and Engineering, Wuhan, Hubei, 430205, China
Corresponding authors: (e-mail: 13164694055@163.com).

Abstract This paper proposes an intelligent algorithm-supported design framework for athletes' physical fitness improvement paths, aiming to optimize the training program by grouping data mining techniques and clustering algorithms. Eight training groups of athletes from different regions of city A were selected as research objects, and the Apriori algorithm was used for in-depth correlation analysis of physical fitness data. K-means clustering algorithm was introduced to determine the grouping for the athletes' physical fitness data, and personalized training programs were formulated relying on the characteristics of each group. Ten athletes were randomly selected for investigation, the average number of times the heart rate of the ten athletes fell in the effective interval without program guidance was 8.2 times, and the average value of the number of times the heart rate of the ten athletes fell in the safe and effective interval under the personalized training program was 9.8 times, which is 19.51% more than that of the training without program guidance, which indicates that the intelligent algorithm provides a scientific and precise path design for the improvement of athletes' physical fitness.

Index Terms data mining, sports training optimization, K-means clustering, Apriori algorithm

I. Introduction

Physical training is the core part of competitive sports, and it is also an important guarantee for competitive sports to achieve excellent results. At present, the main way of athletes' physical training is specialized physical training, including strength training, speed training, endurance training and flexibility training [1]. Specialized physical training is an all-round practice of athletes' physical quality and sports skills under special sports load [2]. Although specialized physical training can help improve athletes' physical quality, it can also negatively affect athletes' physiological functions, thus reducing athletes' performance [3], [4]. Therefore, individualized programs need to be developed in conjunction with specialized physical training to improve athletic performance. However, although traditional physical training methods play an important role in improving athletes' physical performance, they are monotonous and repetitive, which can easily lead to athletes' boredom and affect the motivation and durability of participation in training [5]-[7]. And the lack of timely and effective feedback mechanism in the training process, resulting in students can not understand their training status and progress in a timely manner, and can not be adjusted according to the feedback of the training program and methods, also limiting the enhancement of the training effect of the space and the sustainability of the development of [8]-[10].

With the rapid development of information technology, the current application of information technology in athletes' physical training has become an inevitable trend, such as artificial intelligence, big data, the Internet of Things and cloud computing and other information technology in physical training has become possible. Information technology through algorithms, and relying on intelligent sports equipment real-time monitoring of the athlete's physiological and sports data, with the depth of the analysis of data information, accurate prediction of the future intensity of the exercise, to develop the optimal personalized program of the participants [11], [12]. On the one hand, it can provide accurate real-time monitoring and feedback to accurately obtain the required body indexes of the participants during training and improve the efficiency and effectiveness of training [13], [14]. On the other hand, the feedback data can be quantified, and based on the effective feedback data, the abnormalities of the body indexes during the training process can be detected in a timely manner, which ensures the safety of the participants and reduces the potential risks [15]-[17]. Under scientific guidance, students can gradually realize the physical fitness goals and achieve the best training effect.

First, information technology plays an important role in monitoring and assessing athletes' physical fitness. For example, heart rate monitoring equipment not only records data, but also captures subtle changes in heart rate in real time at various stages of training, which helps to reveal their cardiovascular function and endurance status. Schneider, C. et al. analyzed the response factors affecting heart rate measurements in athletes and never designed more effective heuristic heart rate measurements to provide scientific guidance and support for athletes' sport practices, including training [18]. Gajda, R. et al. explored the strengths and weaknesses of existing heart rate monitoring devices in assessing heart rate and rhythm in endurance athletes and incorporated expert recommendations to construct an improved heart rate monitoring system to provide athletes with the implementation of reliable exercise diagnostics [19]. Hernando, D. et al. assessed heart rate variability and heart rate intervals during physical activity in athletes using a wearable heart rate monitoring device to inform athlete training by calculating the correlation, reliability, and consistency between the two at different exercise intensities [20]. Gajda, R. et al. showed that heart rate monitors are a reliable tool for monitoring arrhythmia problems in endurance athletes, and that controlling the exercise intensity of athletes based on the results of ambulatory electrocardiograms reduces the problem of arrhythmia induced by the conductivity of exercise [21]. Muscle fatigue monitoring accurately assesses an athlete's adaptation to the current training intensity and identifies potential risks of fatigue accumulation or overtraining. Suchomel, T. J., et al. described numerous monitoring methods for measuring and monitoring muscle strength characteristics in physical trainers, which quantify athletes' training outcomes and give load-adjustment recommendations for athletes' training [22]. Saisho, O. et al. constructed a real-time visualization system of electromyography, which combined with the analysis of athletes' muscle applicability indexes to guide athletes' skill training and improve their physical training effect [23]. Zhang, X. et al. established characteristic functions for assessing muscle activity levels and analyzing muscle contraction processes, which, in combination with EMG monitoring methods, can provide an effective assessment of training levels and competitive status of athletes [24]. The above monitoring means can provide real-time feedback on the effectiveness of athletes' physical training, which can enable athletes to flexibly adjust their training intensity according to the detected situation and ensure reasonable and efficient training.

Secondly, information technology can provide a scientific basis for the development of appropriate training programs, thus ensuring that the training can effectively enhance physical fitness while avoiding over-exertion of bodily functions. Tan, L. and Ran, N. developed an athlete training process monitoring system given artificial intelligence terminal technology, which can not only comprehensively analyze the change rule of various indicators of athletes in the training state, but also use GPS to obtain real-time information about the location of the athlete to provide real-time guidance for the athlete's training program [25]. Wada, N. et al. emphasized that the optimization problem of health and fatigue deserves the focus of physically trained athletes by proposing probabilistic optimization models to adjust the training program to maximize fitness and minimize fatigue [26]. Wackerhage, H. and Schoenfeld, B. J. Integration of consideration of athletes' training data with individualized needs based on information technology-based training programs and exercise prescription to generate practice-appropriate training programs for athletes [27]. Kumyaito, N. et al. combined an adaptive particle swarm optimization method with constraints to develop an athlete training program that enhances the training effectiveness of the athlete while satisfying their physiological constraints, thus avoiding injuries and overtraining of the athlete [28]. Zhang, J. and Guo, J. proposed an improved ant colony optimization backpropagation method for evaluating the effectiveness of athletes' physical training while developing individualized training programs that maximize training benefits [29]. Baba, D. et al. constructed a scientific training guidance model based on athletes' physical training indexes, based on which the amount and intensity of physical training were strictly planned to safeguard athletes' health and optimize their sports performance [30]. The above dynamic adjustment mechanism based on intelligent information technology enables the training program to closely follow the actual state of the athletes, ensures that the training is always effective, further enhances the effect of physical training, prevents ineffective training and promotes efficient progress.

This paper firstly explores the path of big data processing to analyze the physical fitness to assist decision-making, using neural network model to realize the classification prediction of physical fitness state, and combining with data mining technology to correlate and analyze the physical fitness data. K-means clustering algorithm is selected to group the athletes and adjust the training plan and strategy for different groups. Take basketball program as an example to explore the digital application of intelligent algorithms in athletes' physical training. Combining the data collected from the physical fitness test and the Physical Fitness and Health Self-Assessment Form, the connection between the physical fitness information is mined by the Apriori algorithm. K-means algorithm is compared with FCM algorithm and hierarchical clustering algorithm to explore the effectiveness of clustering grouping in athletes' training. Introduce four-level classification method to verify the training grouping effect of K-means algorithm. The number of times that the heart rate reaches the safe and effective interval is used as the standard of evaluation to investigate the optimization effect of personalized training programs for athletes supported by intelligent algorithms.

II. Design of sports training programs for athletes supported by intelligent algorithms

Sports training is an important factor in improving athletes' special performance and competitive ability, and traditional physical training methods often rely on coaches' experience and intuition, lacking accurate data support and personalized guidance, which not only affects the effectiveness of training, but also may increase the risk of athletes' injuries. In the development process of contemporary competitive sports, relevant digital research has taken shape. Therefore, the purpose of this paper is to rely on data mining and analysis technology to analyze athletes' physical fitness data in depth, and use K-means clustering algorithm to realize the reasonable grouping of athletes' training, so as to formulate targeted training programs.

II. A. Big data-based fitness analysis

II. A. 1) Data acquisition

In order to increase the amount of data and make the results more accurate and representative, this paper adopts a multi-source method to collect data, including the actual test method and questionnaire method. First, the training groups of athletes were randomly selected, and n members of each training group were selected for physical fitness monitoring. The training groups were distributed in different areas, thus making the data sources more extensive. Secondly, the questionnaire was used to allow athletes to log in the webpage to access the results of the fitness test and fill in the questionnaire according to the results of the fitness test. In order to address the problem of possible erroneous data in it, it is also necessary to process the data results of the questionnaire as follows: in the results of the questionnaire, 20% of the data greater than the maximum value of the actual test method or 20% lower than the minimum value of the actual test method are considered as false positives and deleted from them.

II. A. 2) Data processing

Data cleansing is the process of cleaning the data, removing the useless components from it and leaving only the cleanest and most needed parts. Described in technical language, it is the use of technical means to obtain the needed parts of the data and eliminate the unneeded parts. The following 2 points should be noted in the process of data cleaning:

(1) To ensure the integrity of the data. The integrity of the data refers to the integrity of the core data, and any data cleansing must be based on the criterion of retaining the core data subject.

(2) To ensure the validity of the data. In the cleaning process, the core data should not be deformed and damaged due to cleaning, and the original state of the core data should be guaranteed. Based on the above 2 requirements, this paper summarizes the basic information of the collected data, comprehensive physical test data and single test data. Objects with incomplete information are eliminated and the corresponding physical test objects are screened. The big data are categorized according to the information, which can be divided into static information and dynamic information. Static information includes basic information, and basic information is information that does not change. Dynamic information, i.e., physical test data and program sports assessment results data, includes the basic information of physical fitness test subjects, comprehensive physical fitness test data, and individual program test data.

The neural network approach is based on the basic structure of neuronal organization in the human brain, constructed from mathematical and computer languages. The neural network is assumed to be a three-layer back-propagation algorithm, with the number of nodes in the input layer being N, the number of nodes in the output layer being M, and the number of nodes in the hidden layer being H. The range of values of H is dynamically adjusted according to the actual values of M and N until the model reaches the optimal result. Suppose the inputs are $x_1, x_2, \cdots, x_n$, the outputs are $y_1, y_2, \cdots, y_m$, the hidden layer nodes are $s_1, s_2, \cdots, s_h$, $x_1, x_2, \cdots, x_n$ represent the parameters of the training program such as lung capacity, grip strength, etc., and $y_1, y_2, \cdots, y_m$ represent the 3 states of good, fair and poor, then according to the forward transfer of information, the hidden node S_j can be obtained as shown in Equation (1).

$$S_j = f\left(\sum_{i=1}^{n} w_{ij} x_i + a_j\right) \tag{1}$$

where, f is the activation function from the input layer to the hidden layer; w_{ij} is the weight between the input layer node i and the hidden layer node j; x_i is the training item; a_j is the bias value from the input layer to the hidden layer; i is the input layer node number; n is the node serial number, and n is a natural number.

Through the hidden node S_j, the output y_k is calculated as shown in equation (2).

$$y_k = g\left(\sum_{j=1}^{h} w_{jk} s_j + b_k\right) \tag{2}$$

where, y_k is the network output value; g is the activation function from the hidden layer to the output layer; h is the serial number of the node in the hidden layer, and h is a natural number; w_{jk} is the weights between the node j in the hidden layer and the node k in the output layer; s_j is the node value of the hidden layer; b_k is the bias value from the hidden layer to the output layer; and j is the node number of the hidden layer.

In this paper, the S -type growth curve function $g(xi)$ is used as the activation function from the input layer to the hidden layer, as shown in Equation (3).

$$g(xi)=\frac{1}{1+e^{xi}} \tag{3}$$

where, xi is the sports item parameter; e is the mathematical constant Euler number.

The logistic regression function $g(z_i)$ is used as the activation function from the hidden layer to the output layer as shown in Equation (4).

$$g(z_i)=\frac{e^{z_i}}{\sum_{j=1}^{h}e^{z_i}} \tag{4}$$

where, z_i is the sequence of training iteration weights.

The data obtained from the actual test method was used as the sample set for model training, and model validation was carried out using the ten-fold cross-validation method, using the back-propagation algorithm neural network method, with the input layer selecting the parameters such as lung capacity and grip strength, and the output layer being three states of good, fair and poor.

II. A. 3) Data mining

An association rule is a strong relationship that may exist between two types of items, and a frequent itemset is a collection of items that often occur together. The algorithm for mining association rules is as follows: preparation of data. Athletes' fitness test data are collected and these data are discretized with continuous attributes to fragment the data and prevent data distortion due to intrinsic associations. Numerical fitness test data are discretized and distributed in several discrete intervals. For example, suppose that for the 50m sprint event, four discrete intervals are classified as excellent, good, qualified and unqualified, which are represented by the symbols I_1, I_2, I_3, and I_4, respectively; for the long term, four discrete intervals are classified as excellent, good, qualified and unqualified, which are represented by the symbols J_1, J_2, J_3, and J_4, respectively; and the seated body Forward bending is divided into 4 discrete intervals of excellent, good, qualified and unqualified, which are represented by the symbols H_1, H_2, H_3 and H_4, respectively. Then the record consisting of 50m sprint and solid throw can be expressed in the form of (I_k, J_m, H_n). The set of all frequent items is found by traversing the algorithm. Examples include sprint excellence, stride test excellence and 50m run excellence. Frequent item sets are the external performance characteristics of excellent individuals in physical training, and their logical relationship can be cited as a potential causal relationship, i.e., due to scientific training and reasonable planning, therefore the athletes are excellent in a single item of the test, and many excellent individual athletes make up the frequent item sets of each single item. According to this definition, strong association rules are generated from the frequent item set. For example, good BMI represents good pull-ups, which means that people with good BMI will also do good pull-ups. In this case, the process of mining frequent itemsets is as follows: in the first computation of the algorithm, each item is a member of the candidate 1-item set C_1, the algorithm scans all the records and counts the number of occurrences of each item, assuming that the minimum support is Min, and removes the members of the set C_1 whose occurrences are less than the minimum value, to obtain the set L_1 of frequent 1-item sets. Initialize $K=2$, i.e., first compute the frequent 2-item set. The set C_k of candidate K itemsets is generated by concatenating $LK-1$ with itself. The algorithm scans all records and counts the number of simultaneous occurrences of each item in C_k. Members of the set C_k with less than Min occurrences are deleted to obtain the set of frequent k itemsets $L_k\left(k=k+1\right)$.

II. B.Personnel training grouping based on K-means algorithm

In order to propose targeted training programs and provide intensive training for athletes, this paper introduces the K-means algorithm to realize a more scientific, efficient and flexible grouping of trainers. In the K-means algorithm, the elbow rule is usually used to determine the k value to minimize the sample and mass point squared error as the objective function, and the sum of the squared distance error between the mass point of each cluster and the sample point within the cluster is called the degree of distortion, for a cluster, the lower the degree of distortion, the tighter it represents the members of the cluster, and the higher the degree of distortion, the looser it represents the structure of the cluster.

In this paper, distance is used as a similarity metric and the center of each class is obtained from the mean of all values in the class and is described by each cluster center. For a given personnel dataset x, the Euclidean distance is chosen as the similarity metric, and the clustering objective is to minimize the sum of squared distances for each class, i.e., minimization:

$$L=\sum_{k=1}^{4}\sum_{i=1}^{n}(x_i-u_4)^2 \tag{5}$$

Combining the least squares method and Lagrange's principle, the clustering center is the corresponding average value. Meanwhile, in order to ensure the convergence of the algorithm, the final clustering center should be made as stable and unchanged as possible during the iteration process.

The flow of this algorithm is shown in Fig. 1, which is divided into the following four steps:

(1) Select the objects in the data space as the initial center, and each object represents a clustering center.

(2) For the data objects in the sample, according to their Euclidean distances from these clustering centers, they are classified into the class corresponding to the cluster center with the closest distance according to the closest distance criterion.

(3) Update the clustering centers by taking the mean value corresponding to all the objects in each class as the clustering center of that class and calculate the value of the objective function.

(4) Determine whether the values of the clustering centers and the objective function have changed, and output the results if they remain unchanged, or return to step (2) if they have changed.

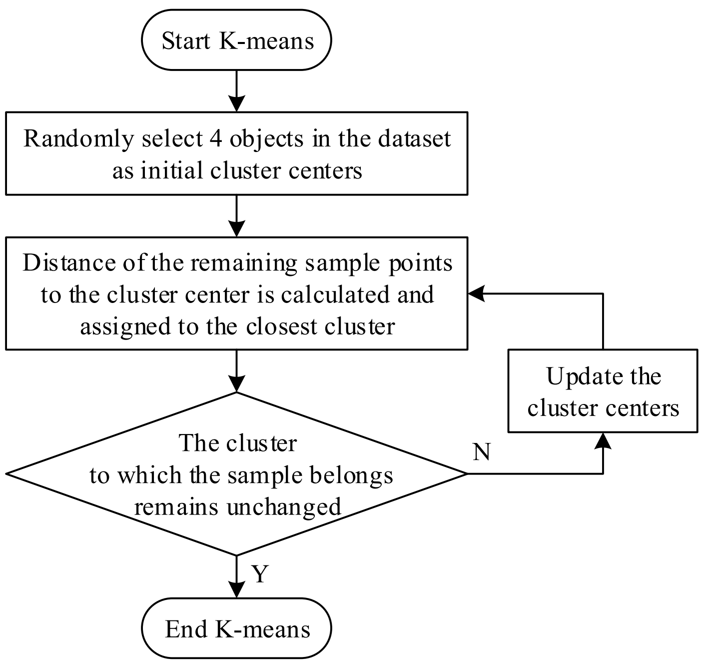

Figure 1: Flow of K-means algorithm

The contour coefficient is an evaluation index of the degree of denseness and dispersion of the class, for which, in the clustering task, it is desired to obtain the result that the clusters are as close as possible within the clusters and as far away as possible between the clusters, and the contour coefficient of the sample point X_A of the athlete A is expressed as follows:

$$S=\frac{b-a}{\max(a,b)} \tag{6}$$

where a denotes the average distance between the sample point X_A of the athlete A's performance and the samples of other athletes in the same cluster, called cohesion, and b denotes the average distance between X_A and all the samples in the nearest cluster, called separation. And the nearest cluster is:

$$C_A = \arg\min_{c_k} \frac{1}{n} \sum_{p \in c_k} \left| p - X_A \right|^2 \tag{7}$$

where p denotes the sample in cluster C_A, i.e., the closest cluster to the sample point X_A is selected as the nearest cluster after using the average distance of the athlete A performance sample X_A to all the samples of the cluster in question as a measure of the distance from that point to that cluster.

And then the Calinski-Harabasz criterion was utilized to determine the clustering evaluation indication:

$$VRC_k = \frac{SSB}{SSW} \cdot \left(\frac{N-K}{K-1} \right) \tag{8}$$

where SSB denotes the between-class variance, $SSB = \sum_{i=1}^{k} n_i \left\| m_A - m \right\|^2$, m denotes the centroid, and m_A denotes the center of the class in which the athlete A is located; SSW denotes the within-class variance, $SSW = \sum_{i=1}^{k} \sum \left\| m_A - m \right\|^2$; $(N-K)/(K-1)$ denotes the complexity; the larger the VRC_K ratio, the greater the data separation.

II. C.Development of Personalized Training Programs

After analyzing the athletes' physical fitness data, artificial intelligence algorithms can help coaches understand the athletes' sports conditions, so as to achieve scientific training and personalized teaching for athletes. Grouping athletes based on their physical fitness data to develop targeted training programs and strategies can help athletes improve their competitive level. Full data monitoring, feedback and optimization has become a new concept and new guideline for high-level competitive training.

III. Optimization effect of generating paths for athletes' training programs based on intelligent algorithms

III. A. Athlete Fitness Analysis

This paper takes the basketball program as an example to explore the digital application of intelligent algorithms in athletes' physical training. The content division of the traditional physical fitness test is relatively clear, and the items of the physical fitness test are divided into lung capacity, grip strength, sprinting, long-distance running and other detection indexes, which comprehensively assesses the physical fitness of athletes.

III. A. 1) Data acquisition

The data used in this paper comes from 8 athletes' training groups from different regions in City A. 10 athletes were selected from each training group, and the data include: athletes' physical fitness test data and data collected by the Physical Fitness and Health Self-Assessment Form, with a total of 1,858 entries in the dataset. The athletes' physical fitness data were taken to be stored and managed by SQL Server 2008. After deleting the relevant personal privacy information and irrelevant fields, the final sample information obtained included gender, height, weight, lung capacity, etc. The specific data are shown in Table 1.

Table 1: Physical Fitness Test Data of Athletes (Part)

Serial number	Gender	Height (cm)	Body weigh (kg)	Vital capacity (mL)	Holding strength (kg)	50m sprint (s)	800m long run (s)	1000m long run (s)	...
1	Male	186.35	88.36	4972	52.43	5.9	-	182	...
2	Female	184.44	79.28	3826	42.35	7.1	157	-	...
3	Male	185.61	86.32	4624	51.46	6.2	-	188	...
4	Male	192.49	93.14	4275	52.88	6.8	-	193	...
5	Male	191.36	92.48	4972	49.24	7.0	-	182	...
6	Male	189.43	87.44	3927	54.38	6.2	-	191	...
7	Female	183.29	86.03	3944	40.43	6.6	138	-	...
8	Female	179.35	82.49	4025	52.59	7.3	149	-	...

9	Male	188.92	85.34	4027	51.39	7.3	-	175	...
10	Female	185.31	81.33	4108	48.28	6.9	146	-	...
...	...	...	...	...	...	...	...	...	...

III. A. 2) Data mining

This paper combines the self-assessment information with physical test data, and uses Apriori algorithm to analyze the association rules of athletes' physical fitness data. According to the actual distribution of physical fitness indexes, the values of physical fitness test item scores are divided into three interval segments: 0~60, 60~80, 80~100. taking the minimum support min_sup=40%, the minimum confidence min sup=70%, the results of athletes' physical fitness data mining are shown in Table 2. From the data in the table, it can be analyzed that there is a correlation between the results of different test items, and when the results of some items are excellent, the results of some items related to them will also be better; while some items with poorer results, the results of some items related to them may also be worse.

As in Table 2, the aspects of athletes' athletic ability tested by the physical fitness programs involved in rules 2, 4, 7 and 8 are similar. From the test results, it is clear that when the athlete's sprinting level is high, the test scores in these programs are generally higher. When the coach carries out the development of the training program, he can try to design the training program for the improvement of the sprinting level to improve the athlete's performance in these items, instead of training to improve the performance of a certain item only, as in the traditional training program.

The 50-meter run and the interval step are mainly tests of the athletes' leg strength, and the results of the waist and abdominal strength test have a certain correlation with these two results. It proves that the strength of the athlete's sprinting ability is not only related to his own leg strength, but also has a great relationship with the waist and abdominal strength. According to the analysis results, it can be concluded that to strengthen the sprinting ability of athletes, in addition to strengthening the training of running, certain waist and abdominal strength training and leg strength training and guidance should be carried out.

Analyzing the test scores of several running events in Table 2 and comparing the athletes' bar pull-up scores can also find a certain relationship. Athletes with better running performance are generally weaker in upper body strength. Athletes with better running ability will focus more on running and lower limb strength training in their normal training, and less on upper limb strength training, which leads to the situation that the bar pull-up performance of athletes with good running performance in Table 2 is not satisfactory.

Table 2: Data mining results of athletes' physical fitness

	Support degree (%)	Confidence degree (%)	Rule
1	80.35	86.57	80 score<50m sprint<100 score⇒0 score<Horizontal bar pull-ups<60 score
2	82.40	83.05	80 score<50m sprint<100 score⇒80 score<Intermittent step<100 score
3	73.64	81.43	0 score<Horizontal bar pull-ups<60 score⇒60 score<Long run<80 score
4	64.47	80.35	60 score<Long run<80 score⇒80 score<50m sprint<100 score
5	62.45	79.53	80 score<50m sprint<100 score⇒80 score<Lumbar and abdominal strength<100 score
6	59.35	77.26	0 score<Long run<60 score⇒0 score<Horizontal bar pull-ups<60 score
7	42.68	75.31	60 score<Long run<80 score⇒80 score<Intermittent step<100 score
8	62.53	73.50	80 score<Long run<100 score⇒80 score<50m sprint<100 score
9	71.59	72.43	80 score<Lumbar and abdominal strength<100 score⇒80 score<50m sprint<100 score
10	44.63	71.39	0 score<Intermittent step<60 score⇒0 score<Long run<60 score
11	57.48	70.48	80 score<Lumbar and abdominal strength<100 score⇒80 score<Intermittent step<100 score

III. B. Athlete Training Groups

Apply K-means algorithm with FCM algorithm, hierarchical clustering algorithm two mainstream clustering algorithms to cluster the processed athletes' physical fitness test data, through the calculation of the distance cost function value to determine the optimal number of clusters k = 4, the three algorithms convergence curve of the objective function is shown in Figure 2. As can be seen from Figure 2, the K-means algorithm declines the fastest, and has converged when the number of iterations is 187, which is faster than the FCM algorithm and hierarchical clustering algorithm.

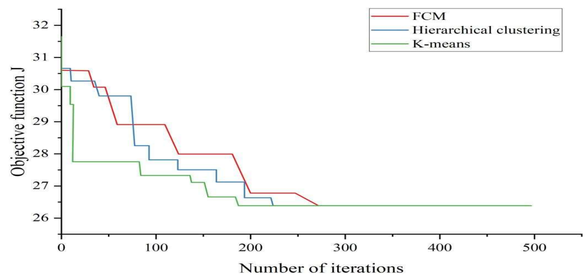

Figure 2: Comparison of convergence curves of objective functions

The location of the optimal particle of the population is coded as the four clustering center points, and the clustering results are shown in Table 3. As can be seen from Table 3, 80 athletes are divided into four categories, A, B, C and D. The center point of each cluster can reflect the overall situation of the physical fitness of the corresponding category of personnel. The group trainer can formulate a targeted training plan according to the average performance and number of personnel in each category to meet the training needs of different categories of personnel, and can strengthen the training for the short board to improve the training efficiency.

Table 3: Clustering results

Category	Number of people	Vital capacity	Holding strength	50m sprint	800m long run	1000m long run	...
A	17	72.85	86.87	76.85	77.95	78.88	...
B	36	80.76	82.66	80.98	82.36	81.93	...
C	19	88.45	77.82	85.22	87.95	85.69	...
D	8	93.52	72.85	91.98	92.38	90.82	...

In order to compare and analyze the K-means algorithm and the traditional four-level system classification method, and given that the four-level classification method usually only considers the performance of a single sport, the simulation experiment is conducted only with the 50m sprint as an example, and the four-level classification method is classified and divided according to the demarcation criteria of excellent, good, qualified and poor, and the classification results of the two methods are shown in Table 4. As can be seen from Table 4, the objective function J of the K-means algorithm is 5.281, which is smaller than the four-level classification method, i.e., the distance value between each sample point and the corresponding clustering center is smaller, so the classification effect is better.

Table 4: Comparison of classification results

Category	K-means		Four-level classification	
	Number of people	Central value	Number of people	Central value
A (Optimal)	17	21.53	12	21.44
B (Good)	36	22.62	38	22.58
C (Qualified)	19	23.09	21	22.91
D (Poor)	8	23.65	9	24.02
Objective function	5.281		6.923	

III. C. Training Program Generation

In order to verify the effectiveness of this paper in formulating personalized training programs for different groups of athletes, the personalized training process was compared with the training process without any program guidance, and the number of times the heart rate reached the safe and effective interval was used as the standard for evaluation.

Ten athletes were randomly selected for investigation, numbered X1~X10, and a long-distance running physical training was divided into 10 sampling cycles, and the number of times the corresponding heart rate reached the safe and effective interval in each athlete's training program was counted, and the experimental results were shown in Table 5, in which "*" indicates that it is not in the safe and effective interval of the heart rate, and EF denotes that the number of times the heart rate reached the effective interval during the training process.

Table 5: Results of personalized training experiment

Respondent number		The training process under the personalized training scheme										EF
		1	2	3	4	5	6	7	8	9	10	
X1	Speed (m/s)	9.24	8.71	7.42	6.85	6.64	6.42	5.31	6.17	6.07	6.01	10
	Heart rate (bmp)	136	138	135	128	142	141	139	144	135	143	
X2	Speed (m/s)	9.63	9.51	9.25	8.53	8.24	8.01	7.49	7.05	6.49	6.22	9
	Heart rate (bmp)	143	135	128	127	142	138	145	157*	148	151	
X3	Speed (m/s)	12.42	9.48	9.31	8.49	9.03	8.46	8.22	7.35	6.36	6.22	9
	Heart rate (bmp)	112*	144	142	139	127	128	136	141	146	137	
X4	Speed (m/s)	9.42	8.58	8.55	8.39	7.30	7.15	6.39	6.16	6.11	6.02	10
	Heart rate (bmp)	138	145	138	135	141	139	144	135	124	128	
X5	Speed (m/s)	9.85	9.34	9.17	9.03	8.49	8.34	8.12	8.02	7.43	6.38	10
	Heart rate (bmp)	127	128	136	136	138	135	128	136	141	124	
X6	Speed (m/s)	8.53	7.34	7.34	6.95	6.74	6.63	6.51	6.38	6.22	6.04	10
	Heart rate (bmp)	122	135	133	142	135	133	142	146	142	129	
X7	Speed (m/s)	9.33	8.94	8.77	8.53	8.35	8.21	8.16	7.35	7.11	6.38	10
	Heart rate (bmp)	142	135	133	128	136	136	127	142	138	145	
X8	Speed (m/s)	9.48	9.24	8.49	7.66	7.35	6.84	6.73	6.61	6.23	6.05	10
	Heart rate (bmp)	142	138	145	135	141	139	135	133	142	146	
X9	Speed (m/s)	9.34	8.56	7.34	7.46	7.43	6.85	6.62	6.37	6.15	6.02	10
	Heart rate (bmp)	126	124	139	133	142	151	147	142	138	145	
X10	Speed (m/s)	11.49	9.35	9.13	8.92	8.73	8.52	7.53	7.29	7.15	6.43	10
	Heart rate (bmp)	135	133	142	146	146	138	135	141	139	141	

In order to prove the effectiveness of the training program in this paper, the training process of ten athletes without the intervention of the training program is given, and the number of times that each athlete's heart rate falls in the safe and effective interval is calculated, and the experimental results are shown in Table 6, in which "*" indicates that it is not in the safe and effective interval of the heart rate, and EF indicates the number of times that the center of the heart rate during the exercise process reaches the effective interval.

Table 6: Results of the no-plan intervention experiment

Respondent number		The training process under the personalized training scheme										EF
		1	2	3	4	5	6	7	8	9	10	
X1	Speed (m/s)	9.84	9.63	8.45	8.38	8.27	8.16	7.47	7.36	7.42	6.31	7
	Heart rate (bmp)	159*	146	149	162*	161*	150	148	147	149	142	
X2	Speed (m/s)	9.39	8.35	7.43	7.01	6.95	6.72	6.61	6.42	6.25	6.03	10
	Heart rate (bmp)	147	141	132	136	139	124	127	128	133	142	
X3	Speed (m/s)	9.86	9.65	9.31	9.22	8.75	8.34	8.11	7.83	7.32	6.85	8
	Heart rate (bmp)	111*	126	128	131	135	138	142	149	161*	147	
X4	Speed (m/s)	9.74	9.48	8.35	8.01	7.45	7.21	7.03	6.38	6.11	6.02	7
	Heart rate (bmp)	112*	110*	125	129	137	145	143	168*	144	147	
X5	Speed (m/s)	9.94	9.82	9.64	9.42	9.31	8.85	8.71	8.62	8.33	8.06	9
	Heart rate (bmp)	159*	146	147	142	151	134	138	133	142	139	
X6	Speed (m/s)	8.82	8.53	8.11	8.02	7.86	7.71	7.52	7.33	7.15	7.06	7
	Heart rate (bmp)	150	151	148	162*	161*	159*	147	142	142	139	
X7	Speed (m/s)	8.42	8.42	7.93	7.81	7.66	7.23	7.05	6.93	6.71	6.64	9
	Heart rate (bmp)	146	135	132	131	129	126	112*	124	128	125	
X8	Speed (m/s)	9.48	9.22	8.38	7.39	6.94	6.55	6.20	6.11	6.04	6.01	9
	Heart rate (bmp)	135	146	142	138	131	130	127	122	115*	128	
X9	Speed (m/s)	10.42	9.59	9.65	8.35	8.31	8.02	7.88	7.73	7.46	7.21	8
	Heart rate (bmp)	133	133	142	136	144	149	150	158*	156*	142	
X10	Speed (m/s)	9.58	8.54	7.89	7.26	7.01	6.73	6.55	6.26	6.12	6.04	8
	Heart rate (bmp)	135	136	150	142	147	146	161*	159*	146	138	

The results of the two experimental groups showed that only one athlete's heart rate fell all the way within the safe and effective interval during the training sessions with no program guidance, while the training sessions of the other groups contained one or more heart rate values that were not within the safe and effective interval, and the average number of times the heart rates of the ten athletes fell within the effective interval was 8.2 times. The personalized training protocol corresponded to 98% of the heart rates falling within the safe and effective interval, and the mean value of the number of times the heart rates of the ten athletes fell within the safe and effective interval was 9.8, which is an increase of 19.51% compared to the training without the guidance of the protocol.

IV. Conclusion

In this paper, we designed a personalized physical training program for athletes supported by intelligent algorithms and explored its optimization effect through example analysis.

The athletes are clustered and grouped using K-means, and the K-means algorithm has converged when the number of iterations is 187, which is faster than the FCM algorithm and hierarchical clustering algorithm. Take 50m sprint as an example for simulation experiments, the objective function J of K-means algorithm is 5.281, which is smaller than the four-level classification method, i.e., the classification effect is better.

Ten athletes were randomly selected for investigation, the average number of times the heart rate of the ten athletes fell in the effective interval under no program guidance was 8.2 times, and the average value of the number of times the heart rate of the ten athletes fell in the safe and effective interval under the personalized training program was 9.8 times, which is 19.51% more than that of the training without program guidance. Thus, the validity of the training path supported by the intelligent algorithm was verified, i.e., the use of intelligent algorithms to target the development of exercise programs can help athletes to further improve the benefits of physical training, while avoiding the harm caused by over-training to their physical and mental health.

Funding

1.“Supported by the Science and Technology Research Program of Chongqing Municipal Education Commission (Grant No.KJQN202300566)”.

2. Supported by Chongqing Normal University Doctoral Program/ Talent Introduction Project "Research on the Cultural Orientation of the Behavioral Patterns of Chinese Professional Football Players" (Grant No.23XWB007).

3. Supported by Chongqing Social Science Planning General Project "Research on the Governance Model and Application of National Youth Football Characteristic Schools in Chongqing" (Grant No.2022NDYB197).

Reference

[1] Zhang, J. (2022). Special training for athletes to improve their physical and sportive capacity. Revista Brasileira de Medicina do Esporte, 29, e2022_0365.

[2] Osipov, A. Y., Nagovitsyn, R. S., Zekrin, F. H., Vladimirovna, F. T., Zubkov, D. A., & Zhavner, T. V. (2019). Crossfit training impact on the level of special physical fitness of young athletes practicing judo. Sport Mont, 17(3), 9-12.

[3] Kokarev, B., Kokareva, S., Atamanuk, S., Terehina, O., & Putrov, S. (2023). Effectiveness of innovative methods in improving the special physical fitness of qualified athletes in aerobic gymnastics. Journal of Physical Education and Sport, 23(3), 622-630.

[4] Sobko, I. M., Lyashenko, A., & Savchuk, A. G. (2020). Peculiarities of special physical training of young athletes in synchronized swimming. Health-saving technologies, rehabilitation and physical therapy, 1(1), 137-142.

[5] Campbell, B. I., Bove, D., Ward, P., Vargas, A., & Dolan, J. (2017). Quantification of training load and training response for improving athletic performance. Strength & Conditioning Journal, 39(5), 3-13.

[6] Brenner, J. S. (2016). Sports specialization and intensive training in young athletes. Pediatrics, 138(3).

[7] Fischetti, F., Alessio, V., Cataldi, S., & Greco, G. (2018). Effects of plyometric training program on speed and explosive strength of lower limbs in young athletes. Journal of physical education and sport, 18(4), 2476-2482.

[8] Huang, H., Huang, W. Y., & Wu, C. E. (2023). The effect of plyometric training on the speed, agility, and explosive strength performance in elite athletes. Applied Sciences, 13(6), 3605.

[9] Uthoff, A., Oliver, J., Cronin, J., Harrison, C., & Winwood, P. (2020). Sprint-specific training in youth: Backward running vs. forward running training on speed and power measures in adolescent male athletes. The Journal of Strength & Conditioning Research, 34(4), 1113-1122.

[10] Negra, Y., Chaabene, H., Hammami, M., Amara, S., Sammoud, S., Mkaouer, B., & Hachana, Y. (2017). Agility in young athletes: is it a different ability from speed and power?. The Journal of Strength & Conditioning Research, 31(3), 727-735.

[11] Xianguo, S., & Cong, W. (2021, April). Research on the Application of artificial intelligence technology in physical training. In 2021 2nd International Conference on Big Data and Informatization Education (ICBDIE) (pp. 261-264). IEEE.

[12] Li, X., Chen, X., Guo, L., & Rochester, C. A. (2022). Application of big data analysis techniques in sports training and physical fitness analysis. Wireless Communications and Mobile Computing, 2022(1), 3741087.

[13] Li, J. (2021). EVALUATION METHOD OF ATHLETES’SPECIAL PHYSICAL FITNESS BASED ON INTERNET OF THINGS. Revista Brasileira de Medicina do Esporte, 27, 62-65.

[14] Yao, W., & Zhihai, Z. (2022). Design of sports training data monitoring system based on wireless internet of things. Mobile information systems, 2022(1), 4162088.

[15] Zhang, B. (2021, October). Application of Smart Wearing Real-time Heart Rate Monitoring in Sports Training. In 2021 International Conference on Health Big Data and Smart Sports (HBDSS) (pp. 177-180). IEEE.

[16] Li, Z., & Song, J. (2024). Enhancing sports trainer behaviour monitoring through IoT information processing and advanced deep neural networks. International Journal of Embedded Systems, 17(1-2), 73-84.

[17] Zhang, Q. (2023, June). An Automatic Monitoring System for Sports Training Indicators Based on Deep Learning and Multiple Sensors. In 2023 International Conference on Applied Intelligence and Sustainable Computing (ICAISC) (pp. 1-6). IEEE.

[18] Schneider, C., Hanakam, F., Wiewelhove, T., Döweling, A., Kellmann, M., Meyer, T., ... & Ferrauti, A. (2018). Heart rate monitoring in team sports—a conceptual framework for contextualizing heart rate measures for training and recovery prescription. Frontiers in physiology, 9, 639.

[19] Gajda, R., Gajda, J., Czuba, M., Knechtle, B., & Drygas, W. (2024). Sports heart monitors as reliable diagnostic tools for training control and detecting arrhythmias in professional and leisure-time endurance athletes: an expert consensus statement. Sports medicine, 54(1), 1-21.

[20] Hernando, D., Garatachea, N., Almeida, R., Casajus, J. A., & Bailón, R. (2018). Validation of heart rate monitor Polar RS800 for heart rate variability analysis during exercise. The Journal of Strength & Conditioning Research, 32(3), 716-725.

[21] Gajda, R., Biernacka, E. K., & Drygas, W. (2018). Are heart rate monitors valuable tools for diagnosing arrhythmias in endurance athletes?. Scandinavian journal of medicine & science in sports, 28(2), 496-516.
[22] Suchomel, T. J., Nimphius, S., Bellon, C. R., Hornsby, W. G., & Stone, M. H. (2021). Training for muscular strength: Methods for monitoring and adjusting training intensity. Sports Medicine, 51(10), 2051-2066.
[23] Saisho, O., Tsukada, S., Nakashima, H., Imamura, H., & Takaori, K. (2019, September). Enhancing support for optimal muscle usage in sports: Coaching and skill-improvement tracking with sEMG. In Proceedings of the 2019 ACM International Symposium on Wearable Computers (pp. 206-210).
[24] Zhang, X., Li, S., Liu, H., & Cao, Z. (2022). Simulation study on changes of EMG and physiological parameters of athletes under training state based on nano biomechanics analysis. International Journal of Nanotechnology, 19(6-11), 1016-1033.
[25] Tan, L., & Ran, N. (2023). Applying artificial intelligence technology to analyze the athletes' training under sports training monitoring system. International journal of humanoid robotics, 20(06), 2250017.
[26] Wada, N., Ito, K., & Nakagawa, T. (2020). Optimal training plans on physical performance considering supercompensation. Communications in Statistics-Theory and Methods, 49(15), 3761-3771.
[27] Wackerhage, H., & Schoenfeld, B. J. (2021). Personalized, evidence-informed training plans and exercise prescriptions for performance, fitness and health. Sports Medicine, 51(9), 1805-1813.
[28] Kumyaito, N., Yupapin, P., & Tamee, K. (2018). Planning a sports training program using Adaptive Particle Swarm Optimization with emphasis on physiological constraints. BMC research notes, 11, 1-6.
[29] Zhang, J., & Guo, J. (2024, May). Application of Big Data Analysis in Sports Training: Personalized Training Plan Generation. In The World Conference on Intelligent and 3D Technologies (pp. 471-481). Singapore: Springer Nature Singapore.
[30] Baba, D., Mijaica, R., Nechita, F., & Balint, L. (2024). Evaluating the Effectiveness of the Annual Physical Training Plan for Masters+ 45 Women Half Marathon Athletes: A Guideline Model for Good Practices for Programming Effort Volume and Intensity. Sports, 12(9), 256.