Personalized Motion Scheme Generation System Design for Motion Software Based on Cloud Computing

Jinkai Duan
Department of Public Courses, Inner Mongolia Technical College of Construction, Hohhot, 010070, China

***Abstract*—The increase of national attention has also promoted the growth of the scale of sports health industry. However, for ordinary people who lack professional knowledge, intuitive data cannot make correct sports planning. Therefore, aiming at the problem that it is difficult for ordinary people to make correct exercise plan according to intuitive data, a personalized exercise plan generation system based on cloud computing is proposed. By analyzing the user's movement and physical data, the system uses cloud computing resources and machine learning algorithms to provide customized exercise recommendations for users. The key innovation of the research is the combination of improved random forest algorithm and reinforcement learning, while improving the performance of the algorithm on unbalanced sample sets. The results indicated that the accuracy of the improved random forest was 0.985 higher than that of the precision weighted random forest. The research algorithm was 9.04% higher on average than the original random forest algorithm and 2.71% higher than the accuracy weighted random forest algorithm. In terms of the accuracy of personalized motion scheme generation of motion software, the improved algorithm reached 95.05% at most, and its recall rate reached 83.46% at most. Compared with the existing sports software solutions, the research system can generate personalized sports programs more accurately, promote the development of the sports health industry and improve the national physical health level. The system can provide users with personalized sports suggestions, and utilize the powerful computing power of cloud computing to realize real-time processing and analysis of large-scale user data, providing users with timely sports feedback and suggestions.**

Keywords—Cloud computing; sports; random forest algorithm; personalization; system

I. Introduction

Nowadays, China's economic level is constantly improving, but the increasing cost of living has also had a negative impact on the physical health of its citizens [1]. Physical exercise can maintain the daily health of human body. Studies have shown that taking appropriate exercise methods and maintaining appropriate amount of exercise can improve health level [2]. The increase of people's attention has also promoted the growth of the scale of the sports health industry. At present, the domestic outdoor sports industry market has great room for growth. At the same time, with the deep popularization of the Internet, China's online sports health market is also growing. The huge user base brings a massive amount of information data. Therefore, how to mine the user characteristic information hidden in these data, analyze the user's body development trend, and provide more reasonable suggestions is the key research direction of sports health application in the intelligent era. Zheng W et al. proposed a programming practice recommendation algorithm based on knowledge structure tree (KSTER). This algorithm was predicated on the quantification of students' cognitive level and knowledge needs, the construction of a KSTER, and the combination of a learning goal prediction method to realize personalized programming practice recommendation. Experiments demonstrated the superiority of this algorithm in terms of precision and recall ratio, as well as its effective improvement of students' learning efficiency [3]. Netz Y et al. evaluated a new tool that uses smartphone sensors to remotely assess and deliver personalized exercise plans. Fifty-two healthy older volunteers participated in the test, which showed significant improvements in strength, flexibility, and balance. Research confirmed the potential of this tool to deliver personalized exercise programs in a home setting [4]. Due to the difference in human body constitution, everyone has a large difference in the exercise methods and amount of exercise. Similar to medical consultation, customized exercise recommendations based on the user's physical characteristics have formed the concept of exercise prescription. Cloud computing integrates computing and storage into a scalable resource pool to improve resource utilization [5]. This study aims to overcome the shortcomings of the random forest (RF) algorithm and improve the classification performance and accuracy of unbalanced sample sets by introducing weighting factors and optimizing the voting mechanism based on the traditional RF algorithm. By collecting users' movement and physical sign data, cloud computing and machine learning algorithms are used to generate a sports health cloud platform for sports programs, and promote the concept of sports health. The innovation of research methods is mainly reflected in five aspects. First, cloud computing integration. The cloud computing platform used integrates computing and storage resources, improving the processing capability and real-time feedback effect of large-scale data. Second, the RF is optimized. The traditional RF algorithm is improved by adding weight factors to improve the classification performance on unbalanced data sets. Third, reinforcement learning application. Reinforcement learning strategies are used to optimize exercise plan generation, dynamically adjust plans based on user feedback, and improve user experience. Fourth, it provides a personalized exercise plan based on the user's physique and goals, and realizes an innovative breakthrough in exercise software. Fifth, compared with traditional algorithms, the advantages of the improved algorithm in accuracy rate, recall rate and F1 score are verified.

The research mainly includes six sections. Section I introduces the background and significance of the research on cloud platform and intelligent motion system. Section II summarizes the intelligent sports software, mainly for the detailed analysis of the achievements of the current domestic and foreign experts and scholars in the direction of sports health assistance system design. Section III is the research method, which is mainly divided into two sections. In subsection one, the research proposes the software scheme system design. In subsection two, the research proposes a personalized motion scheme generation system for motion software based on cloud computing and improved RF algorithm. Section IV is the performance analysis of the personalized motion scheme generation system of sports software. Discussion is given in Section V. Section VI summarizes the research methods and results analysis. At the same time, the shortcomings of research methods and future research directions are proposed.

II. Related Work

Sports can maintain the daily health of human body. A large number of studies have shown that taking appropriate exercise can improve people's health level. Yang C and others proposed to monitor sports energy consumption through cloud services and the Internet of Things and establish an effective data model for difficult operation and implementation of the method of detecting sports energy consumption. Research showed that the detection error of this method for sports energy consumption was less than 2% [6]. In order to promote the clinical application of metabolomics, Castelli F. et al. addressed problems such as insufficient standardization of data generation, low metabolite recognition rate, and complex data processing. These challenges affected the interoperability and reusability of data. In addition, metabolomics outputs were complex and expensive, and features needed to be simplified for on-site applications [7]. Xie D et al. designed a wearable energy-saving fitness tracking system for health monitoring of athletes based on the integrated of deep learning. The algorithm used smart phone applications to track the steps of using Internet of Things technology. Research showed that energy efficiency system performance was improved [8]. Virtual reality technology integrated digital image processing, computer, artificial intelligence, multimedia, sensors, and other information technologies to provide support for the creation and experience of virtual worlds. Its combination with cultural relic protection created a new field. Ling Z et al. explored the personalized health care pavilion display system based on VR and deep learning. A new computational model was designed to improve the system performance. Its effect was verified through application scenario testing [9].

To understand the cognition of high school physical education teachers to the rules of basketball competition, Saputro A et al. used data collection technology and descriptive research methods to investigate the high school physical education teachers in their local area. The results showed that most physical education teachers had a higher cognition of basketball game rules [10]. Men Y et al. proposed a motion recognition algorithm to improve prediction efficiency. A database model was constructed to analyze sports training and competitions. The results showed that this research improved the prediction performance of vicious energy motion [11]. Wang L et al. established the SVM model. Based on texture feature extraction method and in-depth analysis of non-downsampling contour wave transformation, an intelligent motion feature recognition system was constructed. The research showed that the system effectively and accurately extracted and recognized motion features [12]. Sun C M D et al. introduced time control factor to design classification strategy when using support vector machine for classification. Experiments showed that the research method achieved fast target discrimination when the ambient brightness changed, and improved recognition accuracy of complex human posture [13]. Wu B et al. proposed a motion assistant system to identify, query and analyze motion features through deep learning, aiming at the low detection performance of target detection technology in complex background. The experiment showed that the system had good performance in recognizing and classifying human motion behavior characteristics [14].

To sum up, the current research and market application of sports health applications lack effective user data. The application research of data algorithm in sports health utilizes the data uploaded by the sports health monitoring system. However, with the popularity of intelligent sports devices such as smart bracelets, the scale of users' motion data is growing. Cloud computing has powerful computing power, storage capacity and other advantages. Therefore, this study combines cloud computing with machine learning to build a cloud platform. The system uses the improved RF algorithm to solve the unbalance of sample data. Second, the powerful processing power of cloud computing is used to analyze the user's motion data in real time and provide personalized suggestions through machine learning algorithms to ensure the accuracy of the motion scheme.

III. Personalized Motion Scheme Generation System Design and Research for Motion Software Based on Cloud Computing

In this study, an improved RF algorithm is used to generate a personalized motion scheme based on the user's motion and physical data. It comprehensively considers individual differences such as the user's age, health status and medical background. The machine learning classification algorithm is used to accurately identify individual characteristics through a large number of physical test data. The improved algorithm uses a weighted voting mechanism to optimize the classification effect of the decision tree, thus improving the performance of the model when dealing with unbalanced sample sets. The system has adaptive capabilities and can process user data in real time. The high-performance computing capabilities of cloud computing platforms can also be used to dynamically adjust exercise suggestions to ensure personalized and adaptive solutions. Through intelligent analysis and real-time feedback, the system is able to provide users with highly personalized exercise plans that take into account individual differences.

A. Software System Scheme Design

In traditional enterprise software development, computer hardware infrastructure is an important part. Some small and medium-sized enterprises need to purchase physical machine servers with high cost from server manufacturers [15]. However, large software enterprises or server manufacturers have the ability to manufacture large-scale physical machine servers. Therefore, computing and storage devices will be redundant or idle [16]. Cloud computing provides computing resources over the Internet, giving users remote access to data storage, computing power, and applications. Cloud service types typically include infrastructure as a service (IaaS), platform as a Service (PaaS), and software as a service (SaaS). IaaS provides virtualized hardware resources. PaaS supports application development and management. SaaS provides standardized application services. Major cloud providers such as AWS and Microsoft Azure offer a wide range of services for applications such as artificial intelligence, the Internet of Things. To ensure data security and privacy, these services use measures such as data backup, encrypted transmission, and access control, and comply with regulations such as GDPR. Cloud computing not only improves data analysis capabilities, but also provides personalized solutions while ensuring compliance. In terms of usage mode, the cloud computing service users use is to obtain the service provider's storage hardware, underlying network, storage space, and other resources through remote technology. Fig. 1 shows the conceptual model of cloud computing [17].

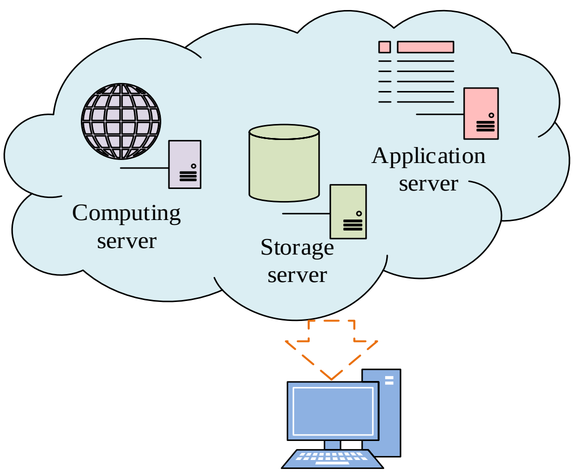

Fig. 1. Conceptual model of cloud computing.

In Fig. 1, cloud computing includes two parts, cloud and terminal. Cloud refers to the collection of network, storage, service and security devices. Cloud boundary is determined by the size of various hardware devices. Terminals refer to terminal devices that can connect to the cloud and help users access and use cloud computing services, as well as various intelligent devices with networking capabilities. The architecture of cloud computing system has three layers, namely core service, service management and user access [18]. The role of the core service is to virtualize the software and hardware so that user services have better performance in terms of reliability, availability, and other aspects to meet the diverse high-quality service needs of users. The service management layer guarantees the stability, reliability and high quality of the cloud platform during operation. The user access function is to provide the user access interface. The cloud platform mainly includes cloud server, database and other components. The main function of the server is to communicate with the client and perform business functions. The research platform is to accurately generate motion schemes for users, and there are many commonly used classification algorithms. RF is is mainly used for regression and classification. The algorithm can freely choose the type of decision tree, and investigates the use of classification and regression tree (CART) as the generation model of a single RF classifier. CART uses GINI as feature selection index, as shown in Formula (1) [19].

$$GINI(P) = \sum_{i=1}^{k} p_i (1 - p_i) \tag{1}$$

In Formula (1), k represents the number of samples classified. p_i is the probability that the sample belongs to this category. For a given sample set D, the expression is shown in Formula (2).

$$GINI(D) = 1 - \sum_{k=1}^{k} \left(\frac{|d_k|}{D} \right)^2 \tag{2}$$

In Formula (2), d_k represents the classification of samples in D. CART is a binary tree. If the training set is divided into two categories because of A attributes, the expression of GINI is shown in Formula (3).

$$GINI(D,A) = \frac{D_1}{D} GINI(D_1) + \frac{D_2}{D} GINI(D_2) \tag{3}$$

The classification result of a RF is determined by the classification pattern of all decision trees, as shown in Formula (4).

$$I(x) = \arg\max_{Y} \sum_{i=1}^{k} I(h_i(x) = y) \tag{4}$$

In Formula (4), y represents the classification category. $h_i(x)$ is the tree of decision tree. Y represents the classification result of a single decision tree. The cloud computing platform is a SaaS service. Its main function is to obtain, store and process data such as web pages and intelligent terminals through servers in Alibaba Cloud [20]. Cloud platform system architecture includes user, data transmission, control, business logic, and data. The user layer is a channel for human-computer interaction and data upload. App and web pages can also be used as human-computer interaction terminals. The app can accept the data uploaded by the body feature sensor, while the web can allow users to upload manually entered data. Data transmission transmits data through the Internet. Control uses Nginx as the front-end server for handling high load access. The business logic is the core platform structure layer, which is realized by running the code of each business function module through the Tomcat server cluster. Data function response operation requirements, mainly through the Redis database and MySQL database, storing physique data, motion data, and scheme library. The hybrid replication is used to achieve data synchronization in the master and slave databases, as shown in Fig. 2.

In Fig. 2, first of all, the main node writes the corresponding SQL statement to the binary text before each transaction completes data update. After the transaction is completed, the main node submits the transaction through the database storage engine. Then, the slave node is connected to the master node by the I/O process, and the slave node determines the initial location of the required binary code. After the master node accepts the request, it parses the request from the slave node. Then, the log message after reading the binary file at that location according to the bit given in the request is published in the I/O process of the node. The main contents of the log information are the file name and the end position of the file. This position is defaulted to the start position of the next request. After the master node returns the data to the slave node, the slave node adds the content to the end of the relay file in order. Meanwhile, the returned file description information is added to the master info file. The slave node detects the newly added files through the SQL thread, and processes the SQL statements of the corresponding nodes, thus realizing the master slave data replication.

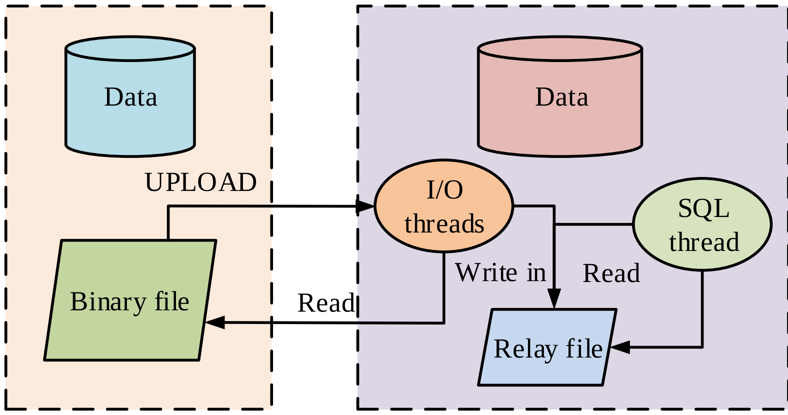

Fig. 2. Master / Slave synchronization execution process.

B. *Design of Personalized Motion Scheme Generation System for Motion Software Based on Cloud Computing and Improved Random Forest Algorithm*

Exercise is an effective way to keep healthy. However, exercise method and amount suitable for everyone are different. Therefore, the cloud computing platform provides users with reasonable online training programs and plays an auxiliary role in sports and fitness. Although the RF classification effect and generalization ability are good, the algorithm also shows an over-fitting phenomenon with too much noise. In addition, the performance of the algorithm under unbalanced sample sets needs to be improved. Therefore, the research improves RF shortcomings, and uses the improved algorithm as the core function of the cloud platform motion scheme generation algorithm. The research mainly improves RF from two aspects: improving the voting mechanism and improving the performance of the RF algorithm in unbalanced sample set. For RF, the input value is set to X and the output value is set to Y. Then, there is a definable interval function $mg(X,Y)$, as shown in Formula (5) [21].

$$mg(X,Y) = P_\theta(h(X,\theta) = Y) - \max_{j \neq Y} P_\theta(h(X,\theta) = j) \tag{5}$$

In formula (5), θ represents the random input value formed in the random selection of training sets from a single subtree in the forest. $h(X,\theta)$ represents the output value. RF classification intensity is the expected value of the interval function. The calculation expression is shown in Formula (6).

$$S = E_{X,Y}(mg(X,Y)) \tag{6}$$

The generalization error of RF is shown in Formula (7).

$$PE^* \le \overline{\rho}(1 - S^2)/S^2 \tag{7}$$

In formula (7), S indicates the classification strength. $\overline{\rho}$ represents the average similarity between decision trees. If the performance of RF is to be improved, its error should be reduced. The classification result function of RF is improved by introducing weight factor, as shown in Formula (8).

$$I(x) = \arg\max_Y \sum_{i=1}^{k} W_i * I(h_i(x) = y) \tag{8}$$

The probability of RF classification for any result α is shown in Formula (9).

$$P_\theta(h(X,\theta) = \alpha) = \frac{W_\alpha}{W_t} \tag{9}$$

In Formula (9), W_α is the total weight of the α decision tree. W_t represents the total weight of all decision trees in the forest. The interval function of RF can be obtained, as shown in Formula (10).

$$mg(X,Y) = \frac{W_Y}{W_t} - \frac{\arg\max_{j \neq Y} W_j}{W_t} = \frac{W_Y}{W_t} - \frac{W_\beta}{W_t} \quad (10)$$

It is assumed that the tree weights of the same classification result are the same, then Formula (11) can be obtained.

$$w_Y * N_Y > w_\beta * N_\beta \quad (11)$$

In Formula (11), N_Y represents the number of decision trees with classification results. Y represents the decision tree classified as the weight of other results and the maximum result. If the formula does not hold, formula (12) is obtained.

$$\frac{N_Y}{N_\beta} > \frac{w_\beta}{w_Y} \quad (12)$$

When the voting mode of RF is equal voting, there is $w_Y = w_\beta$, as shown in Formula (13).

$$\frac{N_Y}{N_\beta} > 1 \quad (13)$$

When voting with equal rights, $N_Y > N_\beta$ can get the correct classification results. When different weights are used, if the weight of the incorrectly classified tree is less than that of the correctly classified tree, Formula (12) exists.

$$\frac{N_Y}{N_\beta} > \lambda, \lambda < 1 \quad (14)$$

In Formula (12), in classification, the number of trees with correct classification results can be less than the number of trees with wrong classification results. Meanwhile, compared with equal voting, when correct and wrong trees in the classification result remains the same, the weight of the correct tree in RF classification is higher, indicating that the weight of the incorrect tree is smaller. Good classification performance indicates smaller generalization error. Fig. 3 shows the steps of the improved RF algorithm after adding weight factors.

In Fig. 3, the improved RF algorithm uses test set pairs to test after the forest is constructed according to the original algorithm flow. The AUC value of each decision tree in RF is obtained and used as the corresponding weight value of classification result of each subtree. Finally, the classification results are obtained by weighted voting. Formula (15) displays the classification results of the improved RF.

$$I(x) = \arg\max_{Y} \sum_{i=1}^{k} weight_{AUC}(i) * I(h_i(x) = y) \quad (15)$$

To satisfy high traffic, stability and other access, a sports health cloud platform is built on Alibaba Cloud using frameworks and component technologies such as Springboot, React, Nginx, MySQL, and Redis. The database module of the platform is very critical, mainly composed of Redis database and MySQL database. Its main function is to provide data storage, query, modification and other operations. Fig. 4 displays the workflow of the database module.

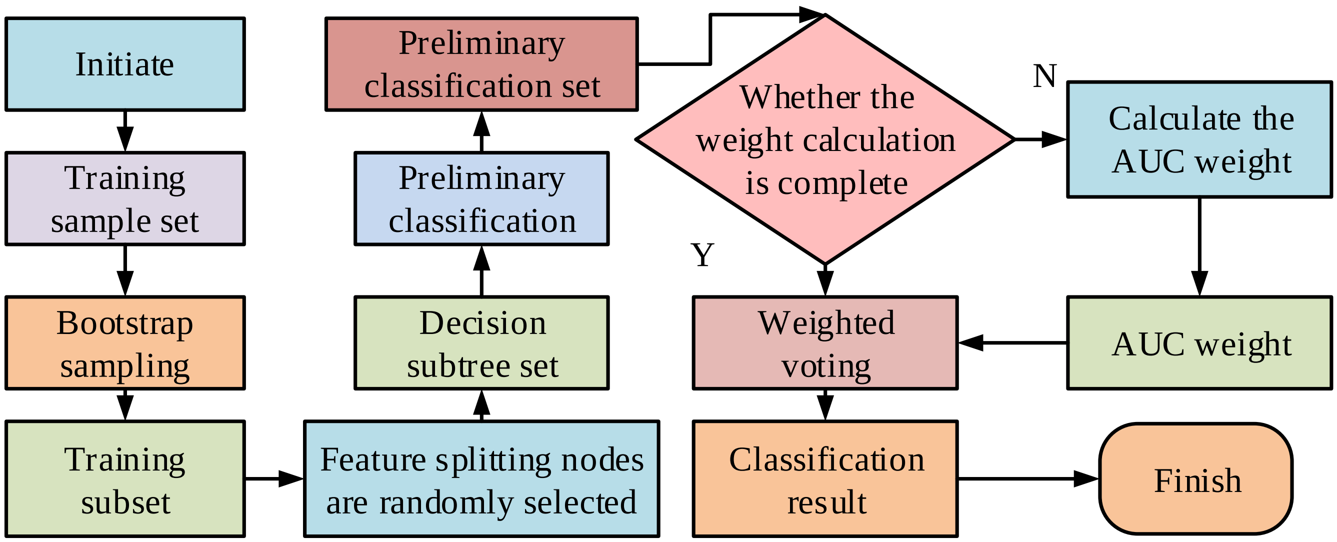

Fig. 3. Steps of the improved RF algorithm.

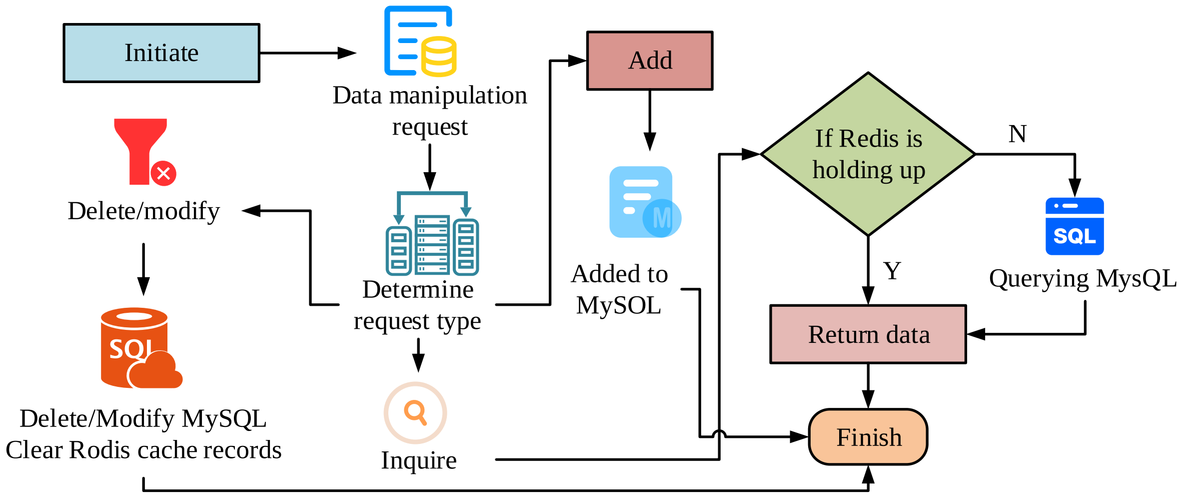

Fig. 4. Workflow of the database module.

In Fig. 4, in the case of high concurrent connections, since MySQL is a traditional relational database, all operations are completed on the hard disk. Therefore, the efficiency of IO is very low, and the database must be used as the buffer. The workflow of the database module is as follows. First, when the user requests a database query, ask whether there is corresponding data in the Redis database. If it exists, it will return directly. If it does not exist, it will query the MySQL database and cache the Redis database after returning data. When the user requests to add an operation, it directly executes the add task in the MySQL database. When the client performs delete and update operations, the MySQL database records are updated directly. If there is a corresponding cache in the Redis database, it is deleted directly. The user's current prescription display module is used to display the user's current exercise prescription. Fig. 5 displays the workflow of the current scheme.

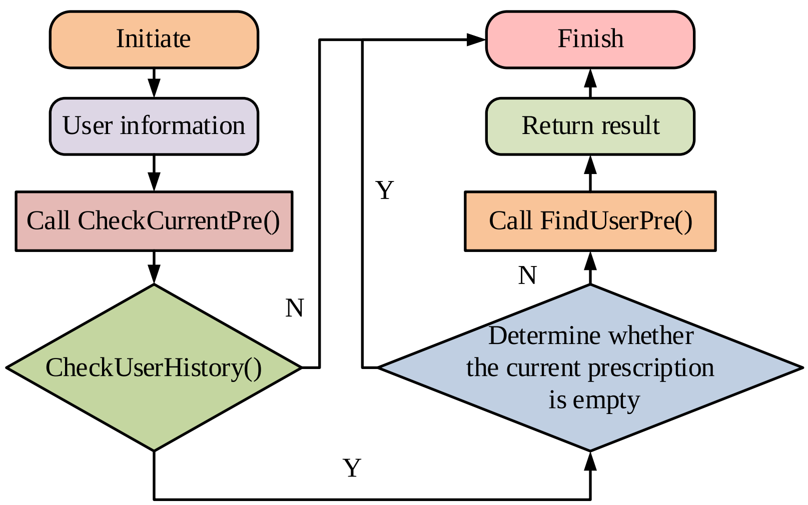

Fig. 5. Workflow in the current scenario.

In Fig. 5, the user only needs to press the current scheme key to submit the request and user information to the platform. After receiving the user's request and information, the platform calls the overall packaging function of $CheckCurrentPre(\ )$. First, the function $CheckUserHistory()$ is called to check whether the user's history file exists. If it exists, whether the ID number of the ID item exe_ Pre_ is empty is queried. If it does not exist or exe_ pre_ id item is empty, the query failure information is returned. If the user record exists and exe_ pre_ id item is not empty, the FindUserPre function is called. According to the motion scheme ID number, the corresponding content is retrieved from the exercise plan library and sent back to the homepage for display.

IV. Performance Experiment Analysis of Personalized Motion Scheme Generation System of Sports Software

The personalized motion scheme generation system uses an enhanced RF algorithm, combines cloud computing resources and machine learning technology. The system analyzes the user's motion and physical data to provide personalized motion recommendations through improved RF and reinforcement learning. The key innovation is the enhancement of the RF algorithm to improve the performance of unbalanced sample sets. The system is trained and verified using a dataset of physical fitness tests in colleges and universities, including multiple test data from male and female students. The cloud platform is implemented by Springboot, React and other technologies. The database module is composed of Redis and MySQL, which supports high concurrency and data caching to ensure low latency and high scalability of the system. The constructed platform is to accurately generate personalized sports programs for users. In order to complete the exercise program database, the experiment collected the physical test data of a university from 2017 to 2022, with a total of 18162 items. There are 9064 for boys and 9098 for girls. Because the data size of this study cannot meet the requirements of deep neural network, the algorithms used to achieve the core functions in the experiment are mainly compared and selected in the traditional machine learning classification algorithms. The performance comparison between RF and other classification algorithms is shown in Fig. 6.

In Fig. 6, in terms of accuracy, RF algorithm value is higher than other four algorithms, which is 87.2%. In terms of recall rate, the ROC value of RF algorithm is still higher than traditional classification methods, which is 0.957. In terms of F1 score, the RF algorithm is 0.861. This algorithm still has a higher F1 score than traditional classification methods such as KNN, SVM, ANN, and decision trees. In conclusion, the RF algorithm has better classification effect than traditional classification algorithms. The experimental test environment is Python 3.6+Sklearm, Intel (R) i5-6300HQCPU@2 Platform. The experimental data are divided into balanced sample set and unbalanced sample set. The specific sample set is shown below.

The total number of samples is 20, with 3 imbalanced negative samples and 10 balanced samples in the sample set. As shown in Fig. 7, the AUC value represents the classification performance under different balance levels of sample sets.

From Fig. 7 (a), when the number of negative categories mistakenly classified as positive categories increases, the AUC value decreases. Therefore, the AUC value has a great impact on the classification effect under the unbalanced sample size. In Fig. 7 (b), AUC decreases as the number of negative classes wrongly classified into positive classes increases. To sum up, AUC value can be used to evaluate the classification effect of algorithms under both balanced and unbalanced sample sets. In addition, it can also be used as a weight factor in the voting process of the improved RF. Figure 8 displays the influence result of RF subtrees.

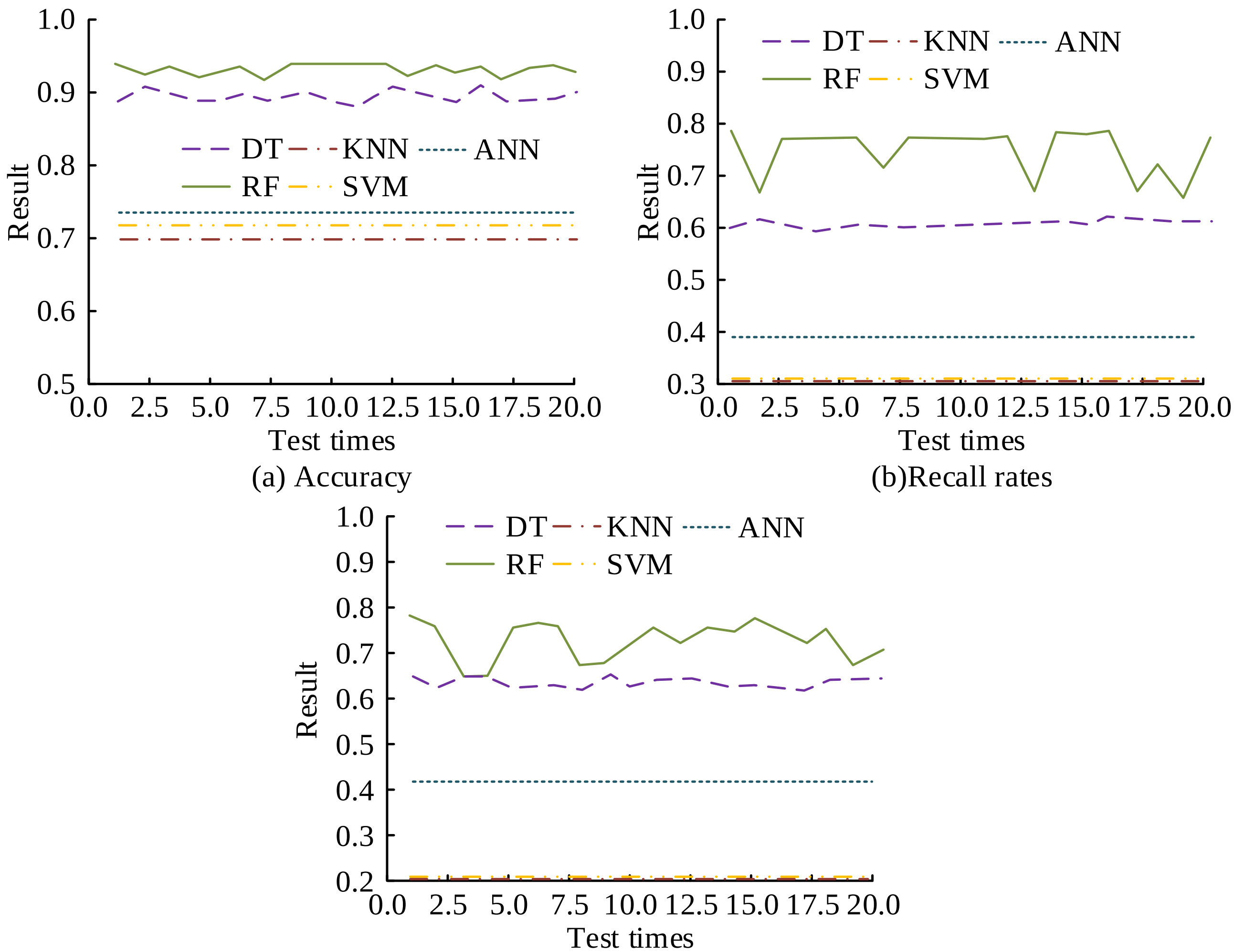

Fig. 6. Performance comparison results of five classification algorithms.

TABLE I. EXPERIMENTAL SAMPLE SET

Sample set		Sample size	Characteristic dimension	Positive sample proposal
Balanced sample set	Banana	5300	2	55.2
	data_banknote_authentication	1371	5	55.5
	Spambase	4600	57	60.6
	Vehicle	846	18	50
Unbal-anced sample set	Ecoli3	336	35	89.6
	Yeast3	1484	8	71.4
	Abalone	580	8	97.7
	Poker_8_9	1459	10	98.3

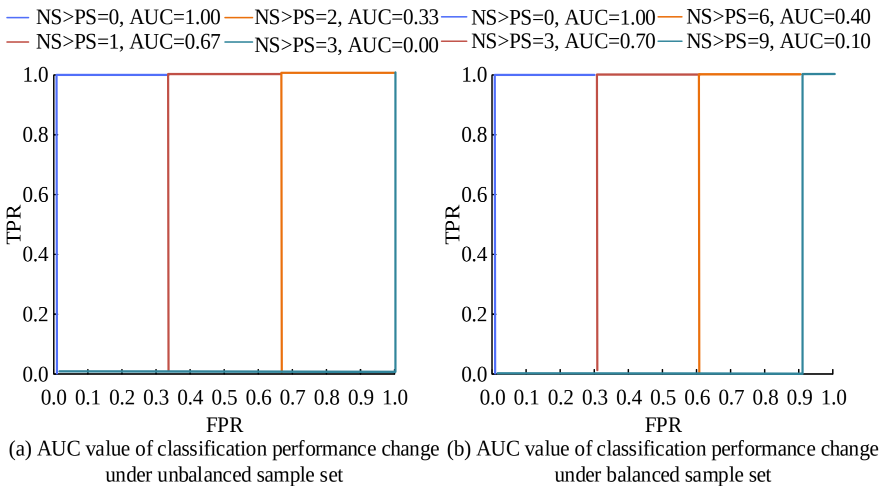

Fig. 7. AUC values representing classification performance in sample sets with different levels of balance.

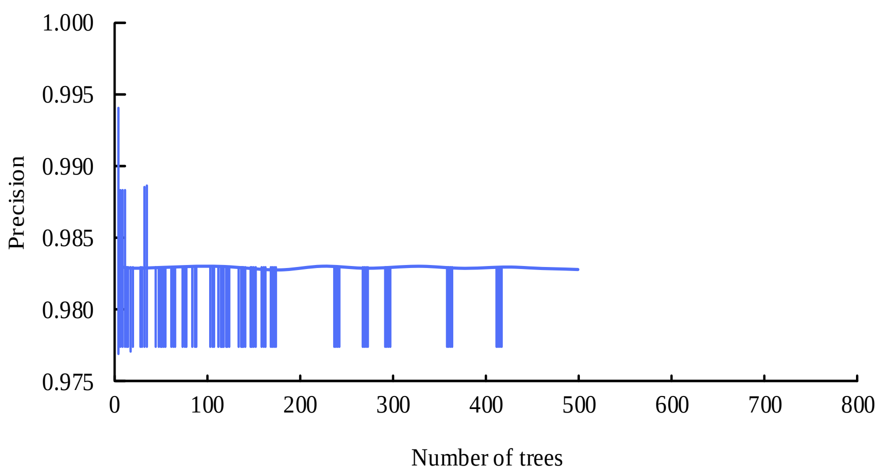

Fig. 8. Influence results of random forest subtrees.

In Fig. 8, after processing with the original RF, Precision gradually converges with the increase of trees. When the subtree is greater than 200, they entered a stable state. Therefore, the number of subtrees in each forest is considered to be 200. Due to the randomness in constructing RF, each algorithm is tested 20 times and then stable values are taken. The experiment uses Recall, Specificity, and G-mean evaluation algorithms to process the results of unbalanced sample sets. Fig. 9 shows experimental result of balanced sample set.

As shown in Fig. 9, in data_ Banknote_authentication dataset, the accuracy of improved RF is 0.985 higher. The precision, recall and F1 score of this algorithm are higher than those of the other two algorithms, with values of 0.977, 0.988, and 0.985. On Spambase, the improved RF algorithm had higher values than the other two algorithms, which are 0.947, 0.977, 0.988, and 0.985, respectively. On Bnana, the improved RF algorithm outperforms the other two algorithms, which are 0.866, 0.898, 0.802, and 0.832, respectively. On vehicle, the improved RF algorithm outperforms the other two algorithms, which is 0.771, 0.763, 0.784, and 0.769, respectively. In conclusion, when the sample set is balanced, the performance of improved RF has certain advantages. Fig. 10 shows the experimental results of an imbalanced sample set.

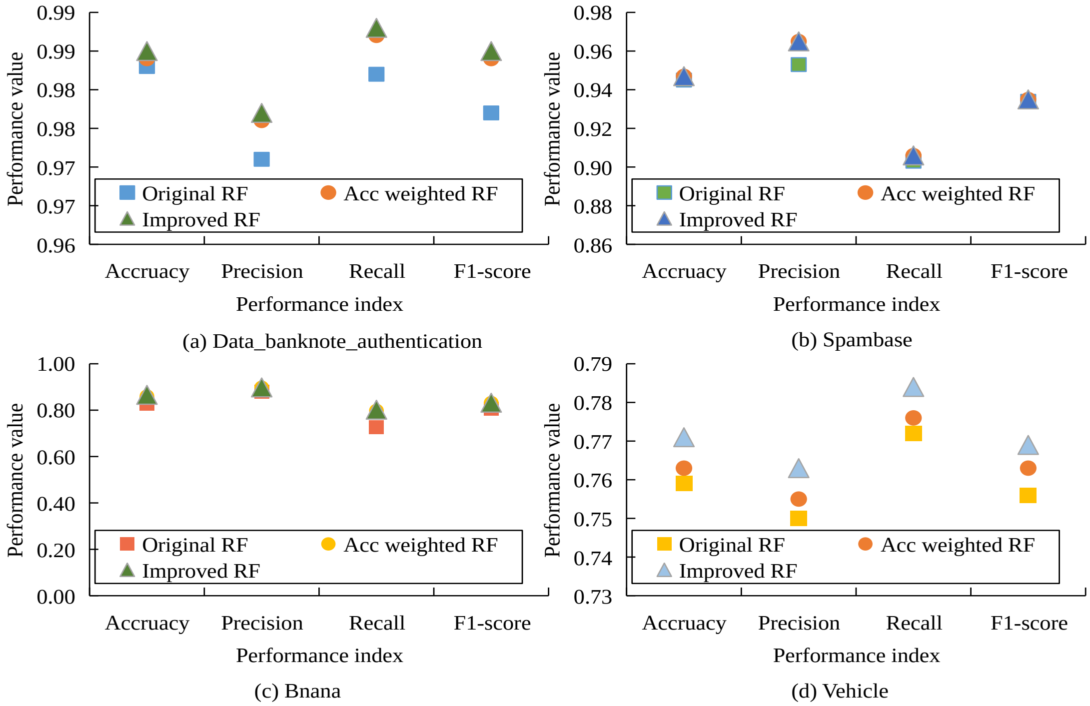

Fig. 9. Experimental results of balanced sample set and unbalanced sample set.

From Fig. 10, the recall, specificity, and G-mean of improved RF are higher than those of the other two algorithms in the Yeast 3 dataset, with values of 0.473, 0.945, and 0.668 respectively. In Ecoli 3, the recall, specificity and G-mean of improved RF are higher than those of the other two algorithms, with values of 0.485, 0.969, and 0.685. In Abalone, the recall, specificity, and G-mean of improved RF are higher than those of the other two algorithms, and their values are 0.994, 0.666, and 0.814, respectively. In Poker_8_9 data set, the recall, specificity, and G-mean of improved RF are higher than those of the other two algorithms, with values of 0.952, 1.0, and 0.975, respectively. To sum up, when the sample set is unbalanced, its specificity representing the accuracy of minority classification is increased by 33.4%, and the comprehensive performance index G-mean is increased by 24.01%. The experimental process is to import the data from the training set into the improved RF algorithm classifier, and train and construct the core function algorithm model. Fig. 11 shows performance test results of sports personalization generation scheme obtained from 20 training sessions.

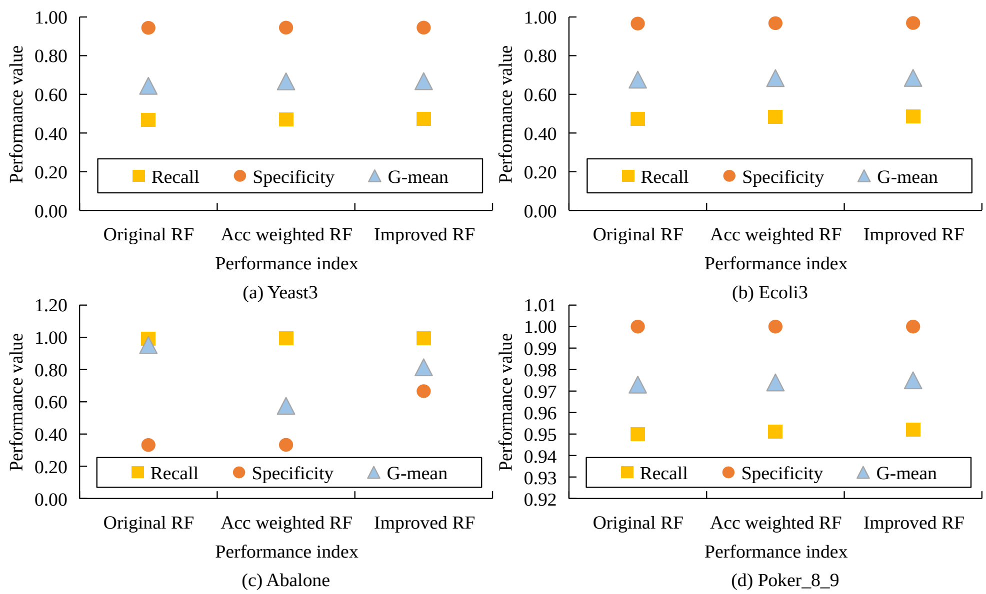

Fig. 10. Experimental results of non-balanced sample set.

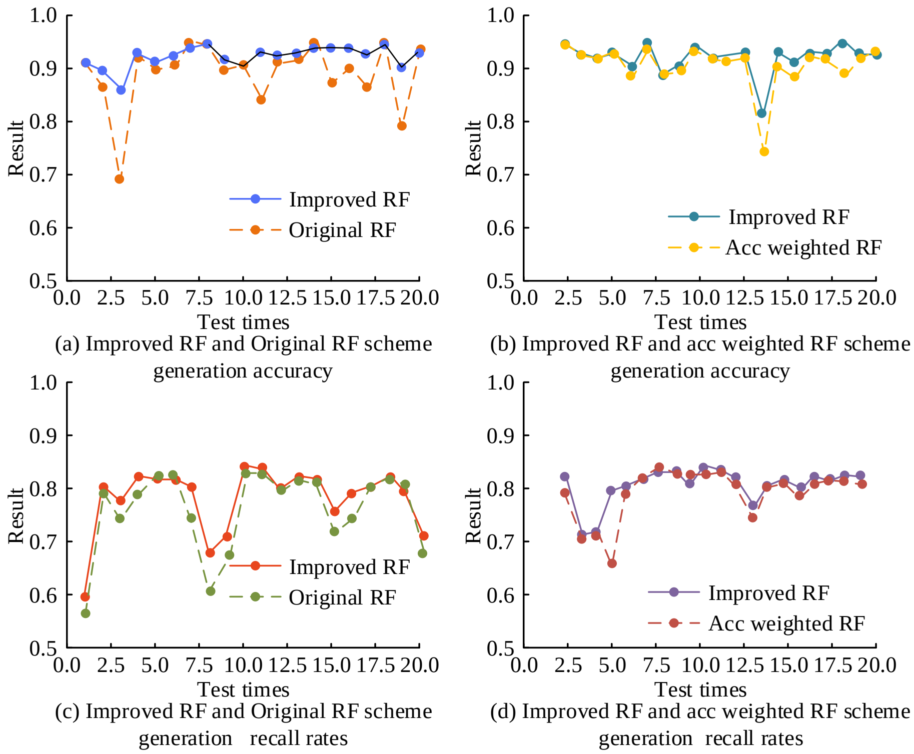

Fig. 11. Performance test results of the motion personalization generation scheme.

In Fig. 11 (a), the accuracy curve generated by the research algorithm is 13 times higher. The highest generation accuracy of the improved algorithm increases by 18.27%. In Fig. 11 (b), the accuracy of the research algorithm is higher than that of the weighted RF for 12 times, with the maximum increase of 4%. The research algorithm is 9.04% higher than the original RF on average, and 2.71% higher than the weighted RF algorithm on average. In Fig. 11 (c), the recall rate of improved algorithm is 13.43% higher than that of the original RF. In Fig. 11 (d), in terms of recall rate, the research algorithm is 7.32% higher than that of the original RF, and 2.68% higher than that of the weighted RF. To sum up, in terms of the accuracy of personalized motion scheme generation of motion software, the improved algorithm reaches 95.05% at most, and its recall rate reaches 83.46% at most.

The ACTIVITYNET Dataset is a large-scale video dataset used to study motor behavior recognition and personalized training plan generation. Video data contains a variety of sports activities, which can be used to analyze and generate personalized sports programs. To more comprehensively verify the performance of the improved RF algorithm proposed in this study in a personalized motion scheme generation system, the research method is compared with several other similar methods on the ACTIVITYNET dataset. The specific results are presented in Table II.

TABLE II. PERFORMANCE OF EACH METHOD'S PERSONALIZED MOTION SCHEME GENERATION ON THE ACTIVITYNET DATASET

Indicators/methods	Traditional RF	Weighted RF	Literature [6]	Deep learning-based method	Literature [11]	Literature [13]	Research method
Accuracy (%)	0.854	0.879	80.5	0.892	85.2	88.1	90.5
Recall	0.841	0.863		0.876			
F1 score	0.847	0.870	0.78	0.884	0.83	0.86	0.91
Computational efficiency (frames per second)	28.500	27.800	30	30.000	25	20	35
Resource consumption (CPU%)	45	42	40	38	35	50	30
Resource consumption (Memory MB)	1300	1250	1200	1400	1000	1500	800

Table II compares the performance of the different methods on the ACTIVITYNET dataset. The research method performs best with an accuracy of 90.5%, an F1 score of 0.91, and a computational efficiency of 35 frames per second. In terms of resource consumption, CPU usage is 30% and memory consumption is 800MB, both of which are the lowest. Compared with traditional RF, weighted RF, and the methods in literature [6], [11], and [13], the research method shows significant advantages in all performance indicators, especially in terms of accuracy, F1 score, and resource consumption. This indicates that the research method has excellent performance and high resource utilization efficiency in the task of personalized motion plan generation.

V. Discussion

This study designed a personalized training plan generation system based on cloud computing and adopted the improved random forest algorithm to achieve in-depth analysis of users' training and physiological data. The results showed that the accuracy and recall rate of the improved algorithm reached 95.05% and 83.46%, respectively. Compared with the exercise energy consumption monitoring methods based on Internet of Things and cloud computing in the literature [6], this study not only focused on energy consumption monitoring, but also focused on providing personalized exercise recommendations through user data. The SVM-based model proposed in literature [11] and [13] performed well in feature recognition, but this study further improved the performance of random forest algorithm on unbalanced sample sets by introducing weight factors and optimizing the voting mechanism, which may be due to the improved algorithm's better handling of sample imbalance, thus improving the accuracy of classification. In addition, literature [8] and [9] discussed the application of deep learning in health monitoring and virtual reality technology. Although the deep learning method was not employed directly in this study, reinforcement learning strategies were utilized to optimize the generation of motion plans and dynamically adjust the plans to adapt to user feedback. This approach simulated the adaptive characteristics of the deep learning model to a certain extent. The enhanced algorithm's improved performance may be attributed to its more efficient capture and utilization of data characteristics, as well as its ability to respond promptly to user feedback, which were pivotal in achieving personalized motion scheme generation.

VI. Conclusion

In view of the high demand of current users for personalized sports advice services in sports health applications, a sports health cloud platform system using cloud computing and improved RF algorithm was designed by combining cloud computing and machine learning algorithms. Experiments showed that in data_banknote_authentication dataset, the accuracy of the improved RF was 0.985 higher. The precision, recall and F1 score of this algorithm were higher than those of the other two algorithms. Similarly, in the Spambase dataset, the Vehicle dataset, and the Bnana dataset, the scores of the research algorithms were higher than the other two algorithms. In conclusion, when the sample set was balanced, improved RF performance had certain advantages. In the Yeast 3 dataset, Ecoli 3 dataset, Abalone dataset, Poker_ 8_ 9 dataset, the recall, specificity and G-mean of improved RF were higher than those of the other two algorithms. When the sample set was unbalanced, the specificity representing the accuracy of minority classification increased by 33.4%, and the comprehensive classification performance index G-mean increased by 24.01%. The research algorithm was 9.04% higher than the original RF on average, and 2.71% higher than the accuracy weighted RF algorithm on average. In terms of the accuracy of personalized motion scheme generation of motion software, the improved algorithm reached 95.05% at most, and its recall rate reached 83.46% at most. This study reduces the dependence on cloud computing resources by improving the RF algorithm, and improves the performance in resource-constrained environments. The system design focuses on modularity and scalability, enabling the rapid integration of new features through plug-ins or services, while optimizing the interface and interaction design to ensure ease of use and accessibility. The initial investment of cloud-based platform is large, but it can reduce maintenance costs and improve resource utilization in the long run. Small users can reduce upfront investment and hardware expenditure through the pay-as-you-go model, lowering the barrier to entry. However, data security and privacy protection need to be taken seriously and may bring additional costs. Despite the flexibility and economics of cloud platforms, as data volumes and users increase, subscription costs may rise as data volume and users increase, requiring small organizations to address the challenges of complex architecture and security management. Therefore, the cloud platform is feasible for small users in the short term, but the long-term cost and technical requirements need to be evaluated. The research has achieved good results. As the number of users and the amount of data increase, how to ensure the stability and accuracy of the algorithm is a challenge. Future research must solve the adaptability of the algorithm to large data sets and improve the generalization ability of the algorithm. In addition, the integrated application of advanced algorithms such as deep learning will further enhance the intelligent level of personalized motion pattern generation. At the same time, research should focus on user privacy and data security to ensure the confidentiality of user information. The exploration of cross-platform data sharing mechanisms to realize data interoperability between different devices and applications will also be an important direction of future research.

Acknowledgment

The research is supported by: Fund Project: 2022 Inner Mongolia Autonomous Region Higher Education Science and Technology Research Project “Research on Innovation and Entrepreneurship Education for Vocational College Students Based on Professional Education” (Project Number: NJSY22422).

References

[1] Vera-Baquero A, Phelan O. Open Source Software as the Main Driver for Evolving Software Systems Toward a Distributed and Performant E-Commerce Platform: A Zalando Fashion Store Case Study. IT Professional, 2021, 23(1):34-41.

[2] Dergaa I, Saad H B, El Omri A, Glenn J, Clark C, Washif, J, et al. Using artificial intelligence for exercise prescription in personalised health promotion: A critical evaluation of OpenAI’ s GPT-4 model. Biology of Sport, 2024, 41(2): 221-241.

[3] Zheng W, Du Q, Fan Y, Tan L, Xia C, Yang F A personalized programming exercise recommendation algorithm based on knowledge structure tree. Journal of Intelligent & Fuzzy Systems, 2022, 42(3): 2169-2180.

[4] Netz Y, Yekutieli Z, Arnon M, Argov E, Tchelet K, Benmoha E, Jacobs J M. Personalized exercise programs based upon remote assessment of motor fitness: a pilot study among healthy people aged 65 years and older. Gerontology, 2022, 68(4): 465-479.

[5] Prabhu N, Jain H, Tripathi A. MTL-FoUn: A Multi-Task Learning Approach to Form Understanding. 2021, 12917:377-388.

[6] Yang C, Ming H. Detection of sports energy consumption based on IoTs and cloud computing.Sustainable energy technologies and assessments, 2021, 46(Aug.):1-11.

[7] Castelli F A, Rosati G, Moguet C, et al. Metabolomics for personalized medicine: the input of analytical chemistry from biomarker discovery to point-of-care tests. Analytical and bioanalytical chemistry, 2022, 414(2): 759-789.

[8] Xie D, Zhang M, Kumar P M, Muthu B A.Wearable energy-efficient fitness tracking system for sports person health monitoring application.Journal of Intelligent and Fuzzy Systems, 2021(3):1-13.

[9] Ling Z. Personalized healthcare museum exhibition system design based on VR and deep learning driven multimedia and multimodal sensing. Personal and ubiquitous computing, 2023, 27(3):973-988.

[10] Saputro A A, Prasetyo G B. PHYSICAL EDUCATION TEACHER KNOWLEDGE LEVEL ABOUT RULES OF BASKETBALL GAME IN JOMBANG DISTRICT SENIOR HIGH SCHOOL. PHEDHERAL, 2021, 18(1): 1-9.

[11] Men Y. Intelligent sports prediction analysis system based on improved Gaussian fuzzy algorithm. Alexandria Engineering Journal, 2022, 61(7):5351-5359.

[12] Wang L, Sun J, Li T. Intelligent sports feature recognition system based on texture feature extraction and SVM parameter selection.Journal of intelligent & fuzzy systems: Applications in Engineering and Technology, 2020, 39(4 Pt.1):4847-4858.

[13] Sun C M D. SVM-based global vision system of sports competition and action recognition.Journal of intelligent & fuzzy systems: Applications in Engineering and Technology, 2021, 40(2):2265-2276.

[14] Wu B. Sports Intelligent Assistance System Based on Deep Learning. Hindawi Limited, 2021, 2021(Pt.11):1-9.

[15] Huang Z, Stakhiyevich P. A Time-Aware Hybrid Approach for Intelligent Recommendation Systems for Individual and Group Users.Complexity, 2021, 2021(2):1-19.

[16] Debnath S. Fuzzy quadripartitioned neutrosophic soft matrix theory and its decision-making approach. Journal of Computational and Cognitive Engineering, 2022, 1(2): 88-93.

[17] Meng X, Ren G, Huang W. A Quantitative Enhancement Mechanism of University Students' Employability and Entrepreneurship Based on Deep Learning in the Context of the Digital Era. Hindawi Limited, 2021, 10:3-12.

[18] Karthikeyan S, Kathirvalavakumar T. Statistical Inference Through Variable Adaptive Threshold Algorithm in Over-Sampling the Imbalanced Data Distribution Problem. 2022, 1404:267-278.

[19] Zhang S, Liu L, Chen Z, & Zhong H. Probabilistic matrix factorization with personalized differential privacy. Knowledge-Based Systems, 2019, 183:4-13.

[20] Ghale-Joogh H S, Hosseini-Nasab S. On mean derivative estimation of longitudinal and functional data: from sparse to dense. Statistical Papers, 2020, 62(4):2047-2066.

[21] Nguyen A D, Nguyen D T, Dao H N, & Tran N Q. Impact Analysis ofDifferent Effective Loss Functions byUsing Deep Convolutional Neural Network forFace Recognition. 2022, 13636:101-111.