

## *Retraction*

# **Retracted: Construction Principles of Physical Fitness Training Objective System Based on Machine Learning and Data Mining**

## **Security and Communication Networks**

Received 31 January 2023; Accepted 31 January 2023; Published 9 February 2023

Copyright © 2023 Security and Communication Networks. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Security and Communication Networks* has retracted the article titled “Construction Principles of Physical Fitness Training Objective System Based on Machine Learning and Data Mining” [1] due to concerns that the peer review process has been compromised.

Following an investigation conducted by the Hindawi Research Integrity team [2], significant concerns were identified with the peer reviewers assigned to this article; the investigation has concluded that the peer review process was compromised. We therefore can no longer trust the peer review process, and the article is being retracted with the agreement of the Editorial Board.

The authors do not agree to the retraction.

## **References**

- [1] J. X. Wang and F. Sun, “Construction Principles of Physical Fitness Training Objective System Based on Machine Learning and Data Mining,” *Security and Communication Networks*, vol. 2022, Article ID 4774372, 10 pages, 2022.
- [2] L. Ferguson, “Advancing Research Integrity Collaboratively and with Vigour,” 2022, <https://www.hindawi.com/post/advancing-research-integrity-collaboratively-and-vigour/>.

## Research Article

# Construction Principles of Physical Fitness Training Objective System Based on Machine Learning and Data Mining

Jian xin Wang and Feng Sun 

*School of Sports Science, Nantong University, Nantong 226019, Jiangsu, China*

Correspondence should be addressed to Feng Sun; sunfeng@ntu.edu.cn

Received 17 February 2022; Revised 12 March 2022; Accepted 28 March 2022; Published 13 April 2022

Academic Editor: Muhammad Arif

Copyright © 2022 Jian xin Wang and Feng Sun. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The objective is to address the issue of simplification of physical education classes offered by large colleges and universities. The evaluation standard of physical education curriculum is not unified. The physical education management system focuses on the functions of collecting information, sorting, and statistics and has low timeliness and guiding significance. This paper puts forward an analysis of the construction principle of physical fitness training target system based on machine learning and data mining. This paper uses informational analysis to statistically analyze the healthy behavior of college students, to guide the physical education of college students, and to propose a model for the analysis of healthy sports behavior of college students based on data mining technology. Create a decision tree template for students whose cardiovascular function does not meet the standard using the decision tree algorithm. The association rule algorithm is used to mine the association of five indexes of physical health, so as to judge the hidden law between students' physical fitness and behavior habits and get the correlation information of various physical health indicators. The simulation results show that, through the prediction of college students' healthy sports behavior data, when the sample point is 5, the original value data is 16, which is higher than the estimated value, the convergence of the overall data characteristic distribution is good, and the disturbance error is low. Therefore, using this method to analyze the application of college students' healthy sports behavior has a high accuracy of sports-related data mining and can effectively guide college students' sports management and training.

## 1. Introduction

In recent years, as the social and economic development of our country has been increased rapidly, the living standards of citizens have been steadily rising and their lifestyles have become more modern, which is a major benefit brought by national development to the people. However, it is the external environment such as lifestyle changes that bring people into a subhealthy state of life; that is, physical activity and physical labor participation gradually decrease, resulting in a large degree of decline in the original healthy body functions, and the incidence rate of hypertension and heart disease increased [1]. Especially for college students, this modern lifestyle makes them stay in the learning and living environment of "sitting quietly and moving less" for a long time, coupled with unreasonable nutritional diet and other

factors, resulting in a serious decline in the physical health of this group [2]. Therefore, the decline of physical health level not only affects the healthy growth of young college students, but also endangers the life safety of the next generation, which makes the country and society have to worry about it. How to change the current situation of physical health of young college students and effectively improve their physical health and physical fitness level is the focus of the current society. College students' sports training plays an important role in improving students' physical quality, cultivating students' sports ability, and training and transporting sports reserve talents for the country. It is becoming more and more important to scientifically analyze and make decisions on the data of individual physical condition and sports quality of college students, so as to formulate scientific physical quality training plans and programs [3, 4].

Physical education is about the health of the whole nation. The development of physical education should start with college students. By improving the physical fitness of university students, we can improve the physical fitness of college students and help healthy college students grow up healthy. The management of college student sports training is the state's economy and people's livelihood is related. Today, with the development of big data information technology, a large statistical analysis of sports training management, combined with the statistical results of a large database of healthy sports training for college students, aims to predict and evaluate healthy physical behavior of college students, accurately make relevant policy decisions, and improve the level of sports management. Data mining originated in the late 1980s. It is a complex cross science, and data mining technology has become an emerging technology after the Internet. It involves many fields of knowledge, such as machine learning, high-performance computing, statistics, database, data visualization, and so on. Therefore, according to the research and development of data mining technology, many large enterprises make it possible to scientifically guide the application of college students' physical training. With the in-depth study of data mining technology, the use of data mining technology is gradually expanding into various fields. Some scholars have applied data mining technology to the education industry. There are a large number of information data that can be mined in the university education system, such as teaching evaluation, student achievement, student information, and so on. However, at present, the research of data mining in higher education is still based on theory, and there are not many formed products really applied to data mining technology [5, 6]. Therefore, the college students' fitness analysis system uses in-depth analysis of student behavior and performance using data collected from student fitness tests and the "Physical Self-Assessment Sheet" based on data mining technology. Displaying the results in the system can not only improve the quality of school education, but also support the development of quality education in our country. Figure 1 is a block diagram of a physical education system.

## 2. Literature Review

In terms of processing student information and student scores, the manual processing mode can not meet the current needs [7]. In such an environment, more scholars and researchers apply data mining technology to college education and physical fitness analysis. Zhao and Zhao used the classification method of decision tree, applied data mining technology to students' achievement information, and constructed a professional ability decision tree model to help teachers have a more accurate and efficient insight into the existing problems in the teaching process, so as to make use of the achievement information to optimize the teaching quality [8]. Xu proposed to adopt decision tree ID3 algorithm and association rule Apriori algorithm for data mining and analysis based on student achievement data [9]. Zhang and Liu studied students' physical health test data from a deeper level by using FP growth algorithm through physical

health test data. The results show that nearly half of the students' weight does not meet the standard, and through the operation results of the algorithm, it is observed that the students lack the training of lower limb strength in physical training, and the vital capacity grade and endurance grade are obviously weak. It is suggested that the students strengthen the training of aerobic exercise [10]. Zhang et al. screened five strong association rules for boys and girls by using the association rule Apriori algorithm based on the body test data of college students. The results show that, under the condition of "total score = pass", more girls fail in the standing long jump and more boys fail in the pull-up [11]. Ba and Qi and others, based on previous research on the meaning and extension of the definition of physical fitness, conducted a more in-depth study of the principles of building a physical fitness target system and index system, focusing on what to do and what not to do and principles to be followed [12]. Jin et al. analyzed through ID3 algorithm and found out the factors related to students' excellent performance. Through the analysis of association rule Apriori algorithm, the influence degree of excellence of a course on other courses is mined [13]. Shen et al. competitiveness refers to an athlete's ability to train effectively and compete effectively and is an organic combination of athletes' physical, skill, mental, and psychological abilities. The process of sports training is a process of comprehensive improvement of athletes' physical abilities [14].

Based on this study, a fundamental analysis of the establishment of a target fitness system based on machine learning and data mining is proposed, as well as the use of data mining technology and data analysis methods to analyze and study college fitness data. The statistics conducted on college students offer a method of modeling and analyzing the healthy sports behavior of college students based on sports training and information mining technology, "analyzing healthy sports behavior and managing the management of college students." The results of this simulation show that the use of this method in the analysis of healthy sports behavior of college students can make the production of sports-related data more accurate and reliable in statistical analysis and make scientific proposals for the establishment of physical education.

## 3. Research Methods

### 3.1. Entity Model Construction

#### 3.1.1. Description of Data Mining Information Fitting Model.

In a big data environment, healthy sports behavior information for college students is extracted, and a database network of healthy sports behavior information for college students is created. The database of sports behavior resources is as follows  $\{e_1, e_2, \dots, e_r\}$ ; seasonal influencing factors for college students' physical activity are provided as shown in

$$\text{dist}(r_i) = \text{dist}(c_i, p_{i+1}) = \text{dist}(\{p_1, p_2, \dots, p_i\}, p_{i+1}). \quad (1)$$

At this time,  $\{\text{dist}(e_1), \text{dist}(e_2), \dots, \text{dist}(e_r)\}$  indicates the health Beneficial Health Benefits Index for College Students.

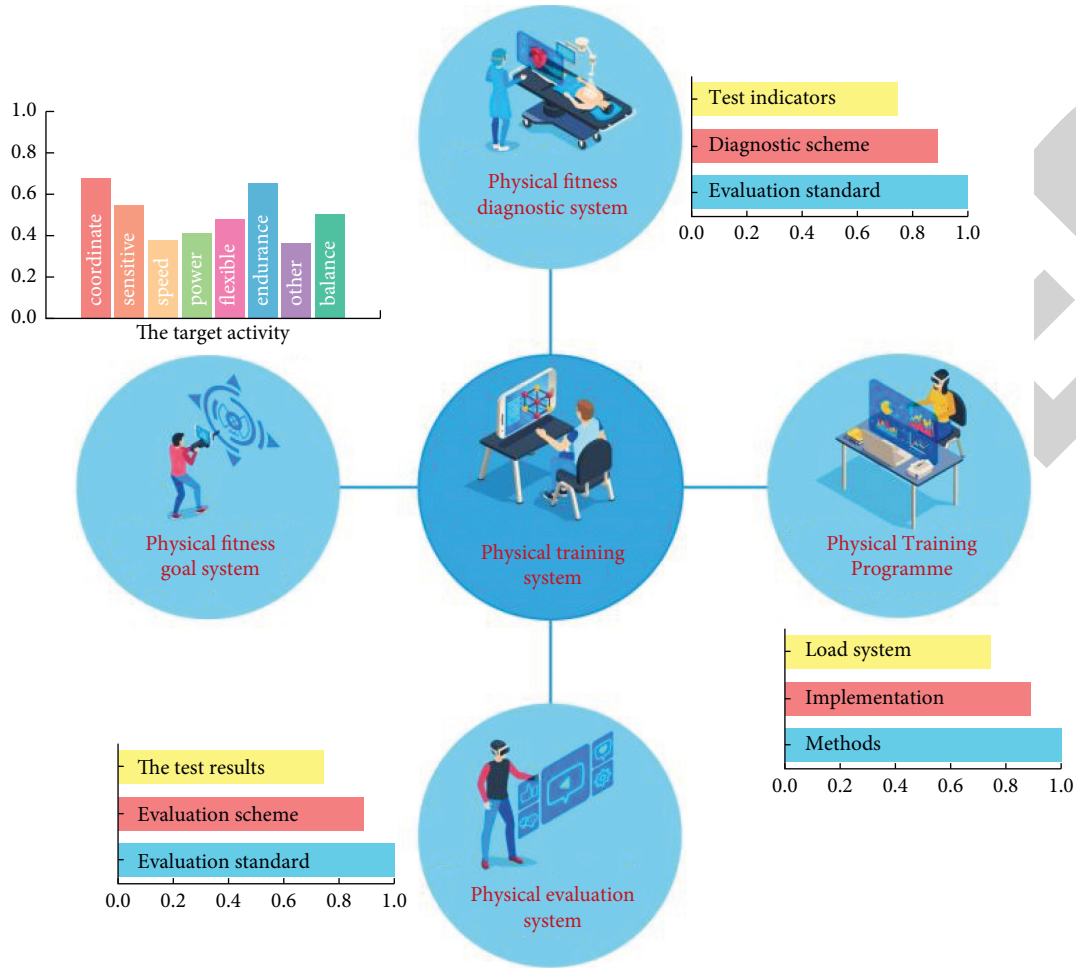


FIGURE 1: Structure of physical training system.

The big data information fitting model is described as equations (2), (3), and (4):

$$R_{\beta}S = U \left\{ E \in \frac{U}{R} | c(E, S) \leq \beta \right\}, \quad (2)$$

$$R_{\beta}S_1 = U \left\{ E \in \frac{U}{R} | c(E, S_1) \leq 1 - \beta \right\}, \quad (3)$$

$$\text{bnr}_{\beta}(S) = R_{\beta}S - R_{\beta}S_1. \quad (4)$$

The characteristic data of college students' healthy sports behavior is affected by fitness equipment, season, and school physical education curriculum. To extract and analyze this data, a set of static and dynamic search models is required to study the physical characteristics of college students' physical activity.

**3.1.2. Fuzzy Decision Feature Extraction.** The fuzzy decision-making method is used to construct the entity model of sports behavior characteristics, and the value range of the feature

distribution of the big data of college students' healthy sports behavior is set as  $N$  discrete feature information points  $A = \{a_1, a_2, \dots, a_N\}$ . The mean time distribution of the index of health and sports behavioral efficiency index of college students is calculated by formula (5).

$$t_m = \frac{1}{E} \int_{-\infty}^{+\infty} t |x(t)|^2 dt. \quad (5)$$

The average frequency range distribution is given by

$$v_m = \frac{1}{E} \int_{-\infty}^{+\infty} t |x(v)|^2 dv, \quad (6)$$

where  $P$  is the order of strength and flexibility in physical exercise.

Historical records of healthy athletic behavior by college students may be consistent with the characteristics of the big data distribution of many influencing factors. An uncertain decision-making constraint model for data mining, taking into account the sparse density of the subnet of healthy sports behavioral information parameters, is designed as shown in (7) and (8):

$$\text{minimize} \left( \sum_{i \in V_{w,+}} c_0^w |b_{i,0}^w - b_i^w| + \sum_{i \in V_{w,-}} c_0^w |b_{i,0}^w - b_i^w| \right), \quad (7)$$

$$\sum_{e \in \delta_{w,-}^-(i)} x_e^w = \sum_{e \in \delta_{w,+}^+(i)} x_e^w + b_i^w, \quad \forall i \in (V_{w,+} \cup V_{w,-}), \quad (8)$$

where  $c_0^w$  is the expected sequence of modes.

**3.2. Theoretical Basis of Machine Learning and Data Mining Technology.** Machine learning (ML) is a study of how computers simulate and implement human learning behaviors in order to acquire new knowledge and skills and how they restructure existing knowledge to continuously improve their performance. Machine learning is a multi-domain interdisciplinary discipline, involving probability theory, statistics, approximation theory, convex analysis, algorithmic complexity theory, and other disciplines, specialized in studying how computers can simulate or realize human learning behavior to acquire new knowledge or skills and reorganize the existing knowledge structure to constantly improve their own performance [15]. Data mining is a simple process of obtaining a valid, innovative, useful, and ultimately understandable pattern from a huge amount of data. Data mining uses a large number of data analysis technologies provided by the machine learning community and data management technologies provided by the database community. Machine learning is an important tool for data mining [16, 17].

**3.2.1. Data Extraction Task.** When data mining technology analyzes and processes data to find valuable information and knowledge hidden under huge data, the main tasks are as follows.

(1). *Concept Description.* By mining the internal information of data, we can get the common and different features between data and then describe the data with these features in the data set. Finally, the different characteristics produced in this class are described differently.

(2). *Classification and Prediction.* The classification method is generally a method of classifying through the established mathematical model or designing the mathematical model according to the research content.

(3). *Association Rules.* The rules of the association reflect the laws and connections that exist between many things. For example, in the supermarket shopping basket problem, association rules can find the correlation between customers buying multiple goods. Association rules are used to find the association relationship between multiple itemsets and infer the remaining attribute information from one or more attribute information [18].

(4). *Cluster Analysis.* Cluster analysis to classify different people according to their unique characteristics. Individuals clustered in the same category have high similarity, but there are differences between different categories.

(5). *Outlier Analysis.* Outlier refers to a small part of uncoordinated data in the data set, which may be caused by measurement or execution errors. Therefore, there are two

basic tasks in dealing with outlier data: one is to define what kind of samples belong to outliers on a given data set; the second is to find an effective way to find such isolated points.

(6). *Trend Classification.* Trend classification is a method to explore the relationship and causality between data in time series data mining. Through trend classification, we can find out the causal laws and trends of some time series in the long-term development process of data.

(7). *Deviation Analysis.* Usually, in the analysis of data, there are some abnormal phenomena that deviate from the traditional data. Through these abnormal phenomena, we can find out the special situations outside the standard class, which can be used to show the differences and changes between the data. Outliers usually appear after cluster analysis.

**3.2.2. Data Mining Taxonomy.** Data mining technology is applied in different fields, involving different disciplines, using different technologies and containing different data. To classify and describe them, see Table 1.

**3.2.3. Data Mining Process.** Data mining is usually divided into four steps:

- (1) Determine the analysis object

Determining the analysis object is the most critical step in the whole process of data mining. Before data mining, it is necessary to determine the purpose of data mining analysis, including experimental objectives and problems to be solved. Only by understanding and learning relevant industry knowledge in advance can we find valuable information and knowledge in the data more effectively according to the demand. If we do not understand the purpose of the research and do it rashly, it will be difficult for the results to meet people's expectations and effects [19, 20].

- (2) Data preprocessing

After defining the analysis object, we preprocess the data; it consists of four steps: data selection, data cleaning, data conversion, and function engineering.

- (3) Data mining

Data mining is the process of analyzing and processing the resulting data set. Feature engineering determines the upper limit of the whole model, and the selection of model and parameters is to wirelessly approach this upper limit. The process of establishing the model needs to select the most appropriate algorithm. Alternative algorithms include machine learning algorithm and deep learning algorithm. Alternative algorithms include classification, regression, cluster analysis, and association analysis. At present, when completing any item, selecting the most appropriate analysis model can greatly improve the efficiency of mining objects.

TABLE 1: Application of data mining.

Classification criteria	Category
Application area	Retail industry, Internet industry, financial industry, communication industry, etc.
Technology used	Statistics, machine learning, visualization technology, etc.
Knowledge type	Decision tree classification, rule induction, cluster analysis, feature rules, etc.
Database type	Object oriented, relationship, transaction, media, etc.

#### (4) Overall analysis

After obtaining the mining results, we should systematically evaluate each stage of the whole mining process. For mining results, delete irrelevant information or redundancy. If the mining result deviates greatly from the expectation, it is necessary to reexamine all the previous steps and sometimes even replace the new analysis model and mining algorithm. Therefore, the overall analysis is a process of model optimization that exists at all times and is carried out many times.

### 3.3. Use Data Mining Methods to Analyze the Physical Fitness of College Students

#### 3.3.1. Description of Physical Problems of College Students.

The main problem is that China's physical education is still a teaching system centered on physical education teachers, and due to the uneven quality of physical education teachers, it will affect the quality of physical education to a certain extent [21, 22]. The backwardness of sports system and statistical tools has also reduced the teaching quality from the following three aspects:

First, the traditional performance evaluation and statistical methods make teachers' tasks complicated and inefficient.

Second, the sports performance and physical examination evaluation system is single, which makes it difficult to implement the results to the problems existing in each student's physical health.

Third, based on the above two points, the heavy task of teachers makes it difficult to give effective guidance and suggestions to students in time. It is difficult to get professional guidance and feedback due to the single evaluation of sports. The flowchart is in Figure 2.

**3.3.2. Data Preprocessing.** The quality of data preprocessing can directly affect the quality of the analysis results. Preprocessing is an important step in all data mining processes such as data quality, preprocessing data selection, data purification, data aggregation, data definition, etc., which can directly affect the quality of analysis [23].

(1). *Data Selection.* There are a large number of redundant and worthless fields in the original data, and the synchronous data analysis of this field information will consume a lot of operation cycle and operation resources and will produce errors to the experimental results. Therefore, data selection is the first step of data preprocessing. The efficiency of data mining can be improved,

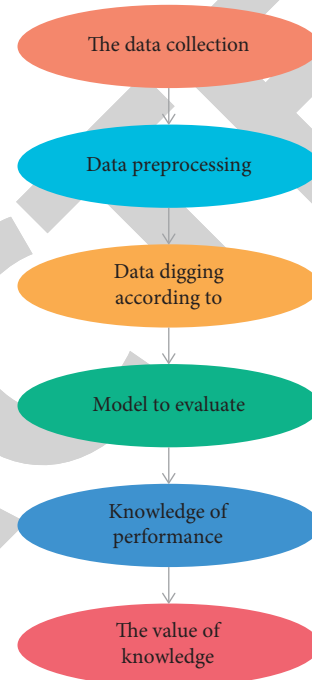


FIGURE 2: Flow diagram for extracting physical fitness test data for college students.

and the hidden internal relations and laws between data information and attributes can be found.

(2). *Data Cleaning.* In practical application, after getting the data, understand its basic situation, determine which data is unreasonable, and then clean it with common data cleaning methods. This process needs to follow the principles of uniqueness, integrity, and legitimacy of the original data set for cleaning. The actual operations mainly include deleting duplicate data, supplementing incomplete data, correcting wrong data, etc., so that the cleaned data is standard, clean, and continuous and meets the requirements of subsequent calculation. Data simplification, deduplication, and standard formatting are completed through data cleaning. The data cleaning model is shown in Figure 3.

(3). *Data Normalization.* Data normalization is a basic operation of data mining. Usually in the process of data mining, different features in the data set have inconsistent dimensions and large differences between values, which will affect the results of data processing. Table 2 shows the standard form for "independent assessment of fitness and health of college students."

(4). *Data Integration.* Data integration is an important step to form a data warehouse. In the process of practical application, due to the different types of databases adopted by the application system, operations such as extraction,

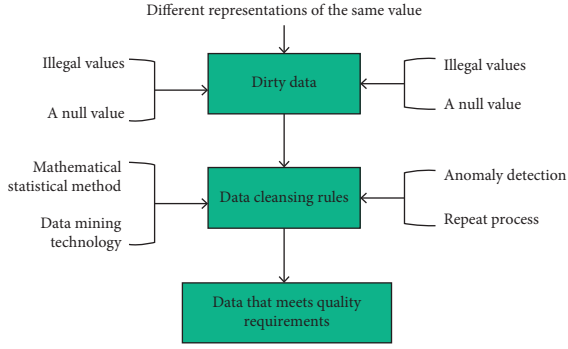


FIGURE 3: Data cleaning principle.

TABLE 2: Specification of college students' physical health self-assessment form.

Attribute	Describe	Attribute value	Code
XB	Gender	Male	XB1
		Female	XB2
XFGN	Cardiopulmonary function	Qualified	XF1
		Unqualified	XF2
JRL	Muscle strength	Qualified	LL1
		Unqualified	LL2
JRNL	Muscle endurance	Qualified	NL1
		Unqualified	NL2
RRD	Softness	Qualified	RRI
		Unqualified	RR2

integration, unified transformation format, deduplication, and merging need to be carried out among various application systems and finally imported into a database [24]. The two tables are combined through the field XH (student number) to generate a new physical health evaluation table for college students, as shown in Table 3.

### 3.4. Data Mining Algorithm and Selection

**3.4.1. Decision Tree Algorithm.** In data mining algorithms, tree model is a particularly common algorithm. Its characteristic is that it can deal with both classification problems and regression problems. In most cases, it is used to deal with classification problems. By learning from the data, the decision tree algorithm classifies and predicts when the values of input variables and output variables are different in different situations. Decision tree is a set of decisions represented by tree structure. The tree model includes root node, leaf node, and nonleaf node. In the actual operation process, leaf nodes are generally represented as a category, leaf nodes connected to the root node are represented as a category, and the path connected between the root node and leaf node is represented as a classification rule of an attribute. Each nonleaf node is represented as a test on an attribute, and each branch represents the output of this characteristic attribute in a value domain [25]. Taking the decision tree algorithm as an example: first, find the appropriate features according to the actual data samples as the root node of the classification, and then judge and classify all other attributes in the data in turn according to the classification standard of

the root node. If the data attribute to be classified is the same as its node attribute, it is divided into the same category; if it is different, it is divided into another category as a new node. Repeat the above operations to form a complete decision tree. See Figure 4.

**3.4.2. Analysis and Application of Decision Tree Algorithm.** As described in the Introduction to the Decision Tree Algorithm, the decision tree algorithm extracts the physical fitness information of college students and calculates the growth rate of information for each characteristic. For the results of the calculations, see Table 4.

The above step J is executed recursively until it cannot be divided, so as to obtain a decision tree on the health status of "cardiopulmonary function" of college students' physical health. Because the data mining algorithm does not consider the problems of data loss and noise in the actual process when generating the decision tree, it needs to prune it after generating the decision tree. Pruning mainly simplifies the decision tree model by controlling the size of the tree, so as to avoid the phenomenon of overfitting to a certain extent. After removing the duplicate branches, the stability and readability of the decision tree model are greatly improved.

**3.4.3. Analysis and Application of Association Rule Algorithm.** The association algorithm is to mine the association between students' physical test data. Then, taking the five indicators in the physical test data (cardiopulmonary function, muscle strength, muscle endurance, softness, and obesity) as an example, eight students are randomly selected from the data list as research samples; see Table 5. Store the unqualified data information in the table into the thing database in turn, and keep the qualified data information, so as to obtain the thing database D. The event database  $d$  contains 6 events, which are sorted according to the dictionary order; see Table 6.

**3.5. Analysis Model of Sports Behavior Characteristics Based on Machine Learning and Data Mining.** The physical fitness data of college students mainly comes from three aspects: physical test data, physical health evaluation, and physical performance. In terms of sports performance, due to the many types of courses offered in different grades and schools, it is difficult to unify the evaluation methods. Through data mining, students' physical fitness data are classified and analyzed, so as to timely grasp the latest trends of students and give targeted guidance, so as to help the school effectively improve students' physical health. The technical structure of the system is divided into a total of three layers; see Figure 5.

A vector machine to support particle herd optimization is used to extract large data of sports behavior, and the particle flight process is given by

$$\begin{cases} v_t = wv_{t-1} + c_1 \text{rand}_1(\cdot)(p_{\text{best}} - x_{t-1}) + c_2 \text{rand}_2(\cdot)(g_{\text{best}} - x_{t-1}), \\ x_t = x_{t-1} + v_t, \end{cases} \quad (9)$$

TABLE 3: Physical health assessment of college students.

XB	XFGN	JRLL	JRNL	RRD	FPCD	YS	ZX	DLXG	TNJK
XB1	XF2	LL1	NL2	RR1	FP1	YS2	ZX2	DL1	TN1
XB2	XF2	LL2	NL1	RR1	FP2	YS2	ZX1	DL1	TN1
XB1	XF1	LL1	NL2	RR2	FP1	YS2	ZX2	DL1	TN2
XB1	XF2	LL2	NL1	RR2	FP2	YS1	ZX1	DL2	TN1
XB2	XF1	LL2	NL2	RR2	FP1	YS1	ZX2	DL2	TN2
XB2	XF2	LL2	NL1	RR1	FP2	YS1	ZX1	DL2	TN2
XB1	XF1	LL2	NL1	RR2	FP2	YS1	ZX1	DL2	TN1
XB2	XF2	LL1	NL1	RR2	FP1	YS2	ZX1	DL1	TN2
XB2	XF1	LL2	NL1	RR2	FP2	YS1	ZX1	DL2	TN2
XB2	XF1	LL2	NL1	RR1	FP2	YS1	ZX2	DL2	TN1

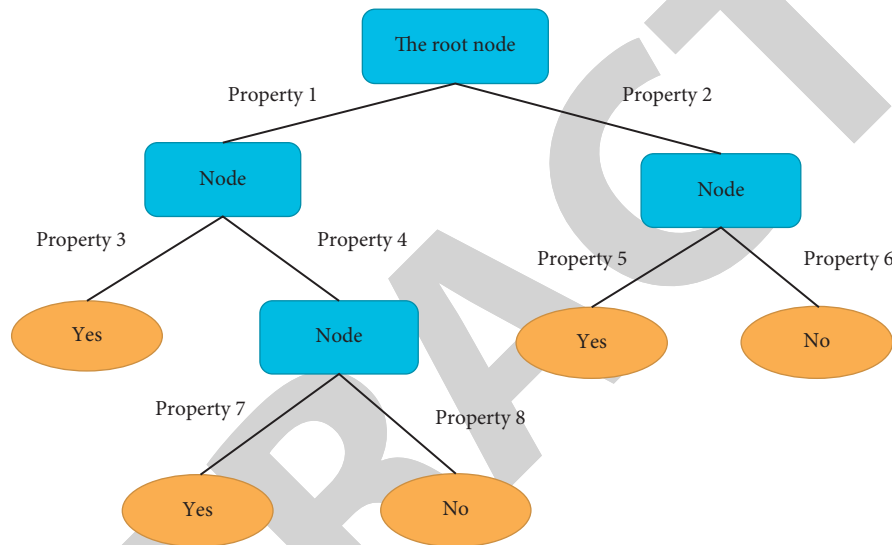


FIGURE 4: Decision tree.

TABLE 4: Calculation results of decision tree algorithm combined with physical fitness data.

Calculation items	YS	ZX	DL	TN
Subset expected information value (E)	0.9417	0.9108	0.8648	0.9402
Information gain value (gain)	0.0037	0.0257	0.0707	0.0063
Split information value (split)	0.8571	0.7500	0.9155	0.9874
Information gain rate (gain ratio)	0.0044	0.0373	0.0772	0.0063

Note. YS: diet; ZX: work and rest; DL: exercise habits; TN: importance of physical education.

TABLE 5: Data samples of association rule Apriori algorithm.

XFGN	JRNL	JRLL	RRD	FPCD
XF2	LN2	LL1	RR2	FP2
XF2	LN2	LL1	RR1	FP1
XF1	LN2	LL1	RR2	FP2
XF1	LN2	LL2	RR2	FP2
XF2	LN1	LL1	RR2	FP1
XF2	LN2	LL1	RR1	FP2

Note. XFG: cardiopulmonary function; JRNL: muscle endurance; JRLL: muscle strength; RRD: softness; FPCD: degree of obesity.



TABLE 6: Transaction database d.

Tid	Items
T <sup>1</sup>	I <sup>1</sup> , I <sup>2</sup> , I <sup>3</sup> , I <sup>5</sup>
T <sup>2</sup>	I <sup>1</sup> , I <sup>2</sup>
T <sup>3</sup>	I <sup>2</sup> , I <sup>1</sup> , I <sup>5</sup>
T <sup>4</sup>	I <sup>2</sup> , I <sup>3</sup> , I <sup>4</sup> , I <sup>5</sup>
T <sup>5</sup>	I <sup>1</sup> , I <sup>4</sup>
T <sup>6</sup>	I <sup>1</sup> , I <sup>2</sup> , I <sup>5</sup>

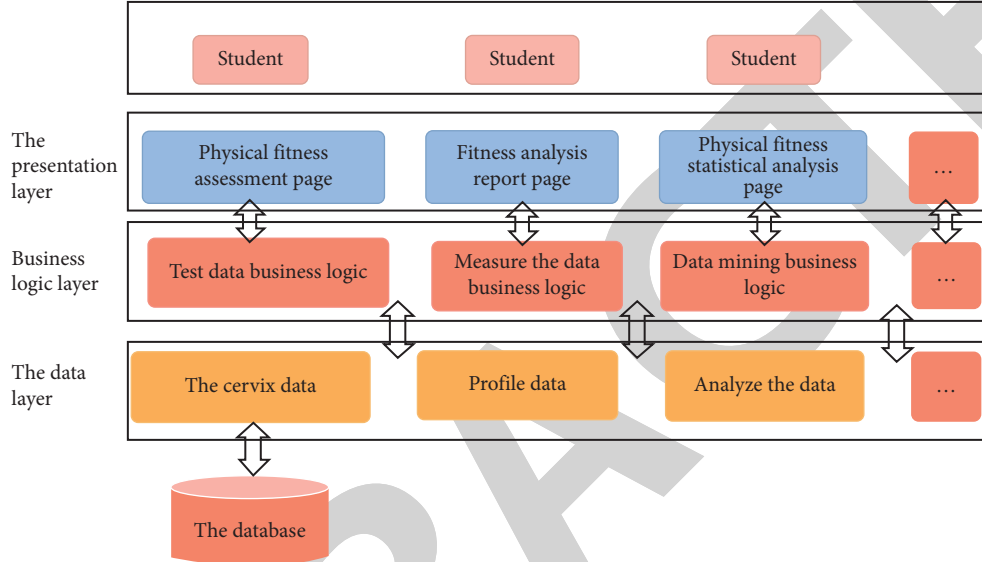


FIGURE 5: General technical architecture of the system.

where  $v_t$  is the running speed of particles;  $x_t$  is the fitness value of particles;  $c_1$  and  $c_2$  are optimal learning operators.

The objective function of the statistical decision to generate data mining is

$$\begin{cases} x_{i,d}^{t+1} = \omega x_{i,d}^{t+1} + c_1 r_1 (p_{a,d} - x_{i,d}^t) + c_2 r_2 (p_{g,d} - x_{i,d}^t), \\ p_{a,d} = \frac{1}{m} \sum_{i=1}^m p_{i,d}, \end{cases} \quad (10)$$

where  $p_{a,d}$  is the average value of the individual optimal positions of the particles.

The individual extreme value  $p_{\text{best}}$  is obtained by using the dynamic inertia weight weighting method. The iterative formula of crossover and mutation of particle swarm optimization is

$$l_i(k) = (1 - \rho)l_i(k-1) + \gamma f_i P_i(k), \quad (11)$$

where  $f_i$  is the error back transfer function of  $X_i$ ;  $P_i(k)$  represents the probability of the  $i$ th particle moving at time  $K$ .

The process of particle cross-optimization is updated according to formula (12) in accordance with the limitations of the convergence condition.

$$\begin{cases} v_{i,d}^{t+1} = \omega v_{i,d}^{t+1} + c_1 r_1 (p_{i,d} - x_{i,d}^t) + c_2 r_2 (p_{g,d} - x_{i,d}^t) + \\ x_{i,d}^{t+1} = x_{i,d}^t + v_{i,d}^{t+1}. \end{cases} \quad (12)$$

Formula (13) is obtained through stability functional:

$$\min_Q \frac{1}{2} \|Q_\Omega - p_\Omega\|_r^2 + \mu \|Q\|. \quad (13)$$

Appropriate adjustment of  $\mu$  value is done to optimize the fitness of college students' healthy sports behavior,  $\mu$ . The selection of value shall meet

$$\|Q_\Omega - p_\Omega\|_r^2 \approx \mu (\Omega) \bar{\sigma}^2. \quad (14)$$

The size of the search step in the downward direction of the gradient is obtained by the sequence of iteration of the particle update using formula (15).

$$\text{dist}(i, j) = \sqrt{\sum_{k=1}^d (x_{ik} - x_{jk})^2 + (f(X_i) - f(X_j))^2}. \quad (15)$$

From this, we can get the characteristics of college students' healthy sports behavior. The data mining process is

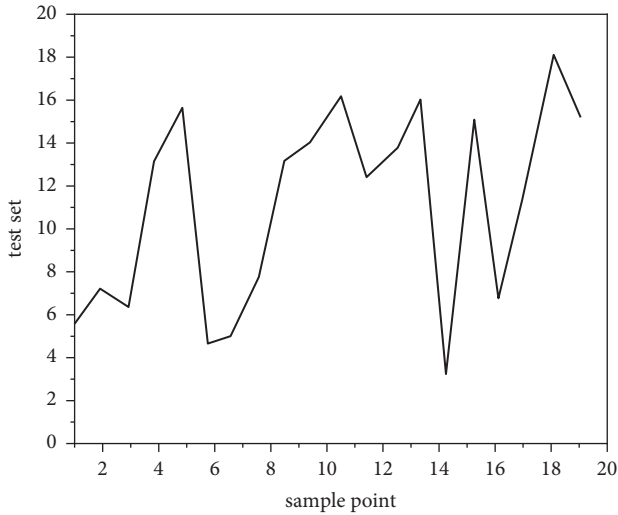


FIGURE 6: Information sampling of original sports behavior data.

$$\min(w, \xi, \xi^*) = \dot{X} \frac{1}{2} \|w\|^2 + \dot{Y} C \sum_{i=1}^n (\xi_i + \xi_i^*), \quad (16)$$

where  $\dot{X}$  is update of the equilibrium coefficient of the particle algorithm;  $\dot{Y}$  is load for data mining;  $H_i(z)$  is the function of an uncertain decision system to disseminate information about healthy athletic behavioral characteristics of college students.

#### 4. Result Discussion

Through the big data analysis and sampling of the university students' healthy sports behavior characteristic database, the statistical characteristics of university students' healthy sports behavior are analyzed. First, define the decision variables: select the seasonal impact factors and determine the relevant variables as follows:  $Z_i$  ( $i = 1, 2, \dots, 8$ ); the importance of healthy sports behavioral education for college students is defined as  $W_{ki}$  ( $i = 1, 2, \dots, 6; k = 1, 2, \dots, 6$ ); the social participation of college students in healthy sports behavior is defined as  $X_{ij}$  ( $i = 1, 2, \dots, 6; j = 1, 2, \dots, 8$ ). According to the above setting, the weight of inertia is 0.78, the correlation coefficient is  $R = 0.5446$ , and the standard error is  $MSE = 0.0321$ . Samples of healthy sports behavior data from college students are taken as a set of experimental data to obtain the sequence of distribution of the first sample of sports behavior information; see Figure 6.

The above data were taken as a set of experiments, data extraction and behavioral hypotheses were made, and the results of the initial data character distribution and hypothesis distribution were shown in Figure 7.

It can be seen from Figure 7 that, through the prediction of college students' healthy sports behavior data, when the sample point is 5, the original value data is 16, which is higher than the estimated value, the convergence of the overall data characteristic distribution is good, and the disturbance error is low.

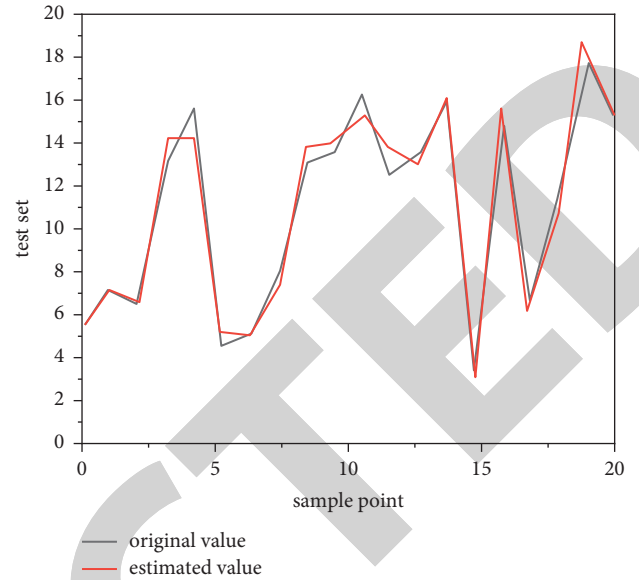


FIGURE 7: Prediction results of sports behavior data mining.

#### 5. Conclusion

Physical education is an important part of college education, and physical education is taught in large colleges and universities, and regular fitness tests are conducted in accordance with national standards. This paper uses machine learning and data mining techniques to statistically analyze the healthy behavioral behavior of college students, guide the management of sports training for college students, and provide a principled analysis of machine learning and data-driven fitness targeting systems mining. First, the ambiguous decision-making method is used to create an enterprise model of sports behavioral characteristics, which is performed using a vector machine to support the extraction of large data of sports behavior and serves as a statistical decision-making function for data mining and then the particle herd method is used to optimize the actual operating parameters of mining to achieve the analysis of large data characteristics of fine mining and healthy sports behavior of college students.

#### Data Availability

No data were used to support this study.

#### Conflicts of Interest

The authors declare that they have no conflicts of interest.

#### References

- [1] V. Kalmanovich, A. Kurshev, and S. Sevodin, "Formation of physical training of future engineers of the construction industry by means of programmed learning," *E3S Web of Conferences*, vol. 274, no. 4, pp. 09015–09019, 2021.
- [2] J. Sun and K. Hu, "The construction of sports public service system for the elderly from the perspective of healthy aging," *Revista Brasileira de Medicina do Esporte*, vol. 27, pp. 66–68, 2021.

- [3] J. Lu and F. Yuan, "The construction of public sports service guarantee system in a healthy community in the background of public health," *Revista Brasileira de Medicina do Esporte*, vol. 27, pp. 62–65, 2021.
- [4] Q. Nian and J. Liu, "Construction of the evaluation system of snowboarders' motor function and the risk evaluation of health sports," *Work*, vol. 7, pp. 1–12, 2021.
- [5] Y. Li, "Construction of intelligent campus tennis players' body data monitoring and injury warning system based on data fusion," *Revista Brasileira de Medicina do Esporte*, vol. 27, pp. 46–49, 2021.
- [6] J. Y. Hong, H. Ko, L. Mesicek, and M. B. Song, "Cultural intelligence as education contents: exploring the pedagogical aspects of effective functioning in higher education," *Concurrency and Computation: Practice and Experience*, vol. 33, no. 2, 2019.
- [7] Z. Zhang and H. Min, "Analysis on the construction of personalized physical education teaching system based on a cloud computing platform," *Wireless Communications and Mobile Computing*, vol. 2020, no. 3, 8 pages, Article ID 8854811, 2020.
- [8] L. Zhao and Y. Zhao, "The construction of the fusion and symbiosis path of infant sports development based on intelligent environment," *Mathematical Problems in Engineering*, vol. 2021, no. 3, 9 pages, Article ID 4942560, 2021.
- [9] H. Xu, "Empirical study on theories and techniques of adolescent physical health promotion under the background of big data," *Mobile Information Systems*, vol. 2021, no. 1, 13 pages, Article ID 3113157, 2021.
- [10] L. Zhang and W. Liu, "Swimming training evaluation method based on convolutional neural network," *Complexity*, vol. 2021, no. 6, 12 pages, Article ID 4868399, 2021.
- [11] Z. Zhang, W. Li, and Y. Zhang, "Automatic construction and extraction of sports moment feature variables using artificial intelligence," *Complexity*, vol. 2021, no. 2, 13 pages, Article ID 5515357, 2021.
- [12] Y. Ba and L. Qi, "Construction of wechat mobile teaching platform in the reform of physical education teaching strategy based on deep neural network," *Mobile Information Systems*, vol. 2021, no. 1, 12 pages, Article ID 3532963, 2021.
- [13] W. Jin, Y. Liu, Y. Jin, M. Jia, and L. Xue, "The construction of builder safety supervision system based on cps," *Wireless Communications and Mobile Computing*, vol. 2020, no. 1, 11 pages, Article ID 8856831, 2020.
- [14] G. Shen, W. Wang, Q. Mu, Y. Pu, Y. Qin, and M. Yu, "Data-driven cybersecurity knowledge graph construction for industrial control system security," *Wireless Communications and Mobile Computing*, vol. 2020, no. 6, 13 pages, Article ID 8883696, 2020.
- [15] S. Sengan, O. I. Khalaf, G. R. Rao, D. K. Sharma, K. Amarendra, and A. A. Hamad, "Security-aware routing on wireless communication for E-health records monitoring using machine learning," *International Journal of Reliable and Quality E-Healthcare*, vol. 11, no. 3, pp. 1–10, 2021.
- [16] D. Viadakovi, M. Hadzima-Nyarko, and S. Marenjak, "The contribution of workers' attributes on sustainability of construction project realization goals—survey on the impact on productivity in Croatia," *Sustainability*, vol. 12, no. 23, pp. 28–33, 2020.
- [17] M. Javed Awan, M. Shafry Mohd Rahim, H. Nobanee, A. Yasin, O. Ibrahim Khalaf, and U. Ishfaq, "A big data approach to black Friday sales," *Intelligent Automation & Soft Computing*, vol. 27, no. 3, pp. 785–797, 2021.
- [18] T. Ma, X. Li, T. Zhu, and C. Cao, "Construction of outdoor sports tourism resource evaluation system for coastal regions," *Journal of Coastal Research*, vol. 106, pp. 182–185, 2020.
- [19] B. Wu, D. Qin, J. Hu, and Y. Liu, "Experimental data mining research on factors influencing friction coefficient of wet clutch," *Journal of Tribology*, vol. 143, no. 12, pp. 1–14, 2021.
- [20] I. Parvez, J. Shen, I. Hassan, and N. Zhang, "Generation of hydro energy by using data mining algorithm for cascaded hydropower plant," *Energies*, vol. 14, no. 2, pp. 298–302, 2021.
- [21] J. Li, "Analysis of the mental health of urban migrant children based on cloud computing and data mining algorithm models," *Scientific Programming*, vol. 2021, no. 5, 7 pages, Article ID 7615227, 2021.
- [22] N. Suo and Z. Zhou, "Computer assistance analysis of power grid relay protection based on data mining," *Computer-Aided Design and Applications*, vol. 18, no. S4, pp. 61–71, 2021.
- [23] L. Tao, "Application of data mining in the analysis of martial arts athlete competition skills and tactics," *Journal of Healthcare Engineering*, vol. 2021, no. 4, 6 pages, Article ID 5574152, 2021.
- [24] B. Wang, "Multimedia filtering analysis of massive information combined with data mining algorithms," *Advances in Multimedia*, vol. 2021, no. 3, 7 pages, Article ID 7461874, 2021.
- [25] L. Liu, "Construction of youth public sports service system based on embedded system and wireless iot," *Microprocessors and Microsystems*, vol. 83, no. 4, pp. 103984–103988, 2021.