Design Studies and Intelligence Engineering
L.C. Jain et al. (Eds.)

doi:10.3233/FAIA250351

Machine Learning-Based Personalized Physical Education Curriculum Design

Chen WU[a,1]
[a]*Xi'an FanYi University,Xi'an 710105,China*

Abstract. In order to solve the problem of low accuracy of personalized prediction of current sports course model, personalized sports course design based on machine learning is proposed. In this paper, the machine learning algorithm-support vector machine is used to establish the college students' sports performance prediction model, and the particle swarm algorithm is used to select the model parameters, and finally the model is applied to the sports performance modeling and prediction in a university. The experimental results show that the prediction accuracies of all college students' sports scores are more than 90%, which are much larger than 85% in the practical application range, which indicates that the model is versatile and can be applied to the prediction of college sports scores in practice. Conclusion: The machine learning algorithm can overcome the shortcomings of the traditional model, improve the prediction effect of college sports performance, and the prediction results can guide the reform of college sports.

Keywords.sports training, machine learning algorithms, predictive modeling, particle swarm algorithms

1. Introduction

With the rapid development of science and technology, digital technology has penetrated into various fields, including physical education. The introduction of digital technology has revolutionized physical education, greatly enriched teaching methods and improved teaching effect [1]. Digital teaching platform is an important infrastructure for physical education. Through the construction of online teaching platform, teachers can realize remote teaching and online guidance, and students can learn anytime and anywhere [2]. The platform usually contains course materials, video tutorials, online tests and other functions, which can help improve students' independent learning ability and participation [3]. Sports data analysis is an important part of sports teaching. Through the collection of athletes' sports data, such as speed, distance, heart rate, etc., athletes' physical condition and technical performance can be comprehensively analyzed. This helps coaches to develop personalized training plans and improve the training effect of athletes [4]. Virtual reality technology provides an immersive learning experience for sports teaching. Through virtual reality equipment, students can simulate a variety of sports scenes in the virtual environment, realistic training.

One of the tasks of machine learning (ML) is to achieve classification, which

[1] Corresponding Author.Chen WU,Xi'an FanYi University,E-mail:309770854@qq.com.

involves predicting target variables in previously invisible data [5]. The purpose of classification is to predict the target variable by establishing a classification model based on the training data set, and then use the model to predict the value of the test data variable. This type of data processing is called supervised learning, because the data processing phase is guided to class variables when building models [6].

Personalized physical education teaching strategies for students based on physical education test data can not only improve the teaching efficiency and intelligence level, but also optimize the allocation of teaching resources, enhance students' learning interest, and promote the overall development of students [7]. However, in the process of practical application, there are still problems such as bias in the selection and design of test items, errors and omissions in the process of data collection, teachers' lack of ability to utilize the data for teaching improvement, lack of in-depth understanding of students' interests, strengths and potentials, and neglect of the influence of students' mental health factors [8]. Therefore, future research should focus on these problems, such as optimizing the design of test items, improving the accuracy of data collection and processing, improving teachers' ability to apply data, exploring students' interests, strengths and potentials, and paying attention to students' mental health [9]. At the same time, the application of modern information technology and artificial intelligence technology should be strengthened in order to improve the intelligent level of physical education teaching and provide students with better quality and more personalized physical education teaching services.

2. Literature review

With the continuous improvement of people's living standard, the material life of college students is more abundant, they lack of exercise, and some college students' physical fitness has declined to different degrees [10]. College students' physical education is an important way to improve the physical quality of college students, and college students' sports performance prediction can help the university sports management department to reasonably open the relevant courses and develop the most scientific training mechanism, so how to establish a high precision sports performance prediction model has aroused great attention of the majority of colleges and universities [11].

The most original prediction methods of college physical education achievements are mainly realized by physical education teachers. They use statistical methods to calculate and estimate college students' physical education achievements. However, because college students' physical education achievements are affected by many factors, physical education teachers cannot fully consider them. Even if various factors are fully considered, the calculation process is quite complex, and the efficiency of sports achievement prediction is low, Can not meet the requirements of modern college students' physical training. With the deepening of information technology, automation technology and artificial intelligence theory research, Lis, A. and others have proposed a large number of automatic prediction models for college sports scores [12]. The prediction of college sports performance is essentially a regression problem, so the current model can be divided into two types: linear regression model and nonlinear regression model. The most widely used linear model is the multiple linear regression model, which estimates the relevant parameters of the model according to the historical sports performance of college students, establishes a college students' sports prediction

model, and then predicts the sports performance of a college student. When there are few influencing factors, the prediction accuracy is high [13]. The modeling of college physical education performance is a very complex process. The influencing factors include college students' own factors, as well as external factors. Therefore, there is no change rule to follow in college physical education performance, and the randomness is large, while the linear model cannot describe the randomness of college physical education performance. The nonlinear model uses the randomness theory to model college students' sports performance, and the established model has higher prediction accuracy. At present, it is mainly the neural network in the machine learning algorithm [14]. Neural network is a machine learning algorithm based on the principle of experience risk maximization. Its nonlinear fitting ability is quite strong, and it can infinitely fit the changing characteristics of college students' sports scores, and has achieved good application results. However, it is based on the principle of experience risk maximization, so a large number of samples of college physical education achievements are needed to obtain the ideal prediction results of college students' physical education, while the historical samples of a student's physical education achievements are quite few, which is difficult to meet this condition [15].

Support vector machine is a new type of machine learning algorithm, which adopts the principle of minimizing structural risk, without the large sample requirement of neural network, and can still obtain good fitting ability under the condition of small sample. In order to improve the prediction effect of college sports scores, a college sports score prediction model based on support vector machine is proposed, and the particle swarm algorithm is used to select the model parameters, and finally the model is applied to modeling and prediction of sports scores of a university, and the results show that the support vector machine solves the deficiencies of the other traditional models and improves the prediction effect of college sports scores.

3. Methodology

3.1 Machine Learning and Sports Learning Diagnostics

Artificial intelligence has promoted the development of ubiquitous learning and adaptive learning, and physical education teaching has been further personalized and fragmented, which makes it difficult to monitor the teaching effect. The contradiction between personalized learning and collective teaching system is becoming more and more prominent, and it is necessary to realize the complementary advantages of the two [16]. Objectively, it is necessary to upgrade the technical monitoring means to provide accurate and systematic big data monitoring support services; subjectively, physical education teachers need to combine students' personality characteristics, reconstruct the teaching content, reengineer the teaching process, and seek a combination of personalized and collective learning of multiple teaching evaluation methods [17].

Machine learning is mainly used for intelligent classification and prediction. At present, machine learning is mostly used in the field of competitive sports, and its impact on physical education is mainly reflected in two aspects [18]. First, deep learning algorithms can be used to recognize the type of sports activities. For example, artificial neural networks can assess individual exercise metabolic equivalents and determine individual activity types (low-intensity activity, exercise, strenuous exercise, and household/other activities); computational modeling can monitor and feedback

muscle status in real time, predict fatigue, and avoid sports injuries [19]. Secondly, deep learning algorithms can be used for sports learning diagnosis and performance prediction; by mining the historical data of training and competition to predict the outcome of the game, providing data basis for hierarchical sports teaching. At present, sports learning diagnosis of deep learning algorithms require high precision and accuracy, and its accuracy needs to be optimized.

3.2 Machine Learning Algorithms for Predictive Modeling of College Sports Performance

Compared with neural networks and other traditional machine learning algorithms, support vector machines have fewer constraints and do not have the defect of "overfitting", which are very suitable for modeling and predicting the sports performance of small samples and nonlinear college students [20].

Let the sample set of college athletic performance is $X = \{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\}, i = 1,2, \cdots, n$

Support vector machine regression is specified as:

$$f(x) = w \cdot \phi(x) + b \tag{1}$$

Where w and b are parameters of support vector machine.

To establish a prediction model for college athletic performance, it is necessary to find the most reasonable w and b, therefore, according to the principle of structural risk minimization, it is converted into the following form:

$$\begin{aligned} &min \quad \frac{1}{2}\| w\|^2 + C\frac{1}{k}\sum_{i=1}^{k}\varepsilon\left(f\left(x_i\right) - y_i\right) \\ &s.t. \quad \varepsilon\left(f\left(x_i\right) - y_i\right) = \begin{cases} |f\left(x_i\right) - y_i| - \varepsilon, & |w \cdot \phi(x) + b - y_i| \geqslant \varepsilon \\ 0, & |w \cdot \phi(x) + b - y_i| < \varepsilon \end{cases} \end{aligned} \tag{2}$$

Where.$\varepsilon()$ is the regression error; C is the error penalty parameter.

In order to simplify the solution process and reduce the computational complexity of modeling, the relaxation factor is introduced, and the quadratic programming form is obtained as follows.

$$\begin{aligned} &\min_{w,b,\xi_i,\xi_i^*} \quad \frac{1}{2}\| w\|^2 + C\sum_{i=1}^{l}\left(\xi_i + \xi_i^*\right) \\ &\text{s.t.} \quad \begin{cases} y_i - w \cdot \phi(x) - b \leqslant \varepsilon + \xi_i, & \varepsilon_i \geqslant 0; i = 1,2, \cdots, n \\ w \cdot \phi(x) + b - y_i \leqslant \varepsilon + \xi_i^*, & \varepsilon_i^* \geqslant 0; i = 1,2, \cdots, n \end{cases} \end{aligned} \tag{3}$$

Further transformation of Eq. (3) using Lagrange multipliers gives,

$$\min_{a^{(*)} \in R^2} \frac{1}{2}\sum_{i,j=1}^{n}\left(a_i^* - a_i\right)\left(a_j^* - a_j\right)k\left(x_i, x_j\right) + \varepsilon\sum_{i=1}^{n}\left(a_i^* + a_i\right) -$$

$$\sum_{i=1}^{n} y_i \left(a_i^* - a_i\right) \tag{4}$$

$k\left(x_i, x_j\right)$ denotes the kernel function.

The regression function of a support vector machine can be described as follows.

$$f(x) = \sum_{i=1}^{n} \left(a_i - a_i^*\right) \left(\phi\left(X_i\right), \phi(X)\right) + b \tag{5}$$

RBF function is selected as the kernel function, which is defined as:

$$k\left(x_i, x_j\right) = exp\ \left(-\frac{\left\|x_i - x_j\right\|^2}{2\sigma^2}\right) \tag{6}$$

Where σ represents the parameter of RBF.

The current optimal solution of the particle swarm optimization algorithm is pbest, and the current optimal solution of the population is gbest. The fitness function describing the advantages and disadvantages of individual particles is constructed as follows:

$$fitness = \frac{1}{2N} \sum_{i=1}^{N} \sum_{j=1}^{D} \left(y_{ij} - t_{ij}\right)^2 \tag{7}$$

$x_{i,d}^k$ and $v_{i,d}^k$ are the speed and position of the kth iteration of particle i, and their updating method is:

$$v_{id}^{k+1} = \varpi v_{id}^k + c_1 Rand\ ()\left(p_{id} - x_{id}^k\right) + c_2 Rand\ ()\left(p_{gbest}^k - x_{id}^k\right) \tag{8}$$

$$x_{id}^{k+1} = x_{id}^k + v_{id}^k \tag{9}$$

Where: c1 and c2 are learning factors; $Rand\ ()$ is a random number; M is the inertia weight.

In order to solve the problem of determining the parameters of support vector machine, particle swarm optimization algorithm was used to optimize the training process of support vector machine online, and the optimal parameters C and σ were determined. According to the optimal parameters C and σ, the prediction accuracy of college students' sports performance was improved.

3.3 Machine Learning Algorithms for Predictive Modeling of College Sports Performance

(1) For a sports program, historical data are collected and processed as follows.

$$x_i^{'} = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{10}$$

In the formula, x_{min} and x_{max} represent the minimum and maximum values of

college students' sports scores.

(2) Determine the parameter C and σ range of support vector machine according to experience.

(3) Initialize the particle swarm, and the position vector of each particle includes a set of parameters C and σ.

(4) According to each set of parameters C and σ, the training samples of college students' sports performance are learned by support vector machine, and the fitness function value of each particle is calculated.

(5) All particles are sorted according to the fitness function value, and the current optimal solution pbest and the current population optimal solution gbest are replaced.

(6) Update the velocity and position of the particle swarm.

(7) The number of iterations of the particle swarm algorithm increases.

(8) If the number of iterations exceeds the predetermined maximum value, then the algorithm is terminated, otherwise return to step (4) to continue the optimization operation.

(9) According to the group optimal solution gbest, the optimal parameters C and σ of the support vector machine are obtained.

(10) According to C and σ, we train college students' sports performance, and establish a prediction model of college students' sports performance based on support vector machine.

(11) The performance of the college sports performance prediction model was tested and analyzed by using college sports performance test samples. In summary, it can be seen that

The workflow of the college sports performance prediction model based on machine learning algorithm is shown in Figure 1.

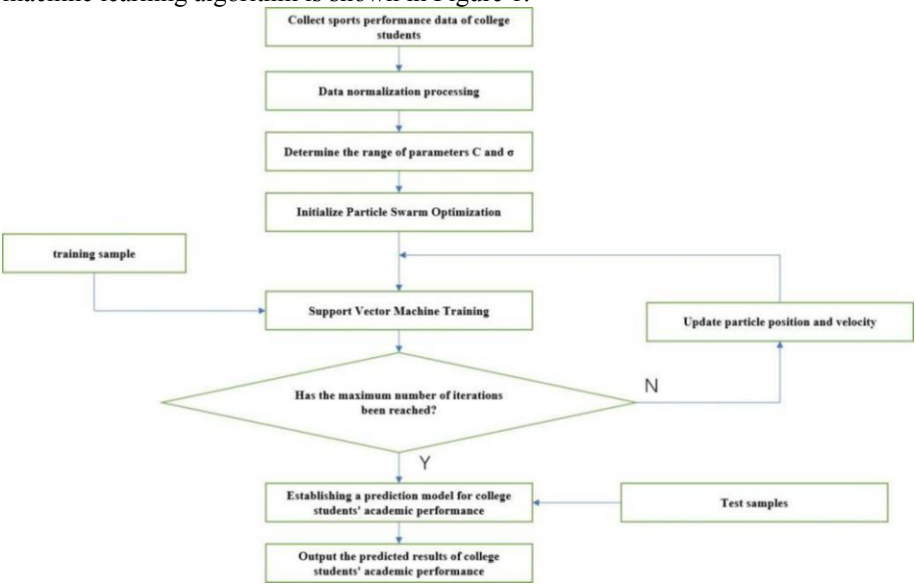

Figure 1. Workflow of the predictive model for college students' sports performance

Digital information technology personalized physical education teaching strategy can provide customized teaching content and training methods according to the characteristics of each student, to meet the individual needs of students, so as to better

promote the overall development of students. Digital information technology makes the interaction between teachers and students more convenient and frequent. Teachers can answer students' questions in time through the online platform, and students can give teachers feedback on their learning at any time, which is a real-time interaction that helps to enhance students' motivation and participation in learning. Through digital information technology, teachers can more comprehensively and accurately understand the students' learning progress and feedback, so as to adjust the teaching content and strategy according to this information in a timely manner, and make the teaching more in line with the actual needs of students. Digital information technology personalized physical education teaching strategy focuses on cultivating students' independent learning ability. By providing rich learning resources and diversified learning methods, digital informatization technology can stimulate students' interest and motivation in learning and encourage students to participate in the learning process more actively.

3.4 Experimental validation

In order to comprehensively analyze the effectiveness and superiority of the prediction model of college students' sports performance based on machine learning algorithm, the college students' sports performance of a university was selected as the experimental object. First, the 100 meter run performance was selected for testing, with a total of 500 scores, 300 of which were 100 meters for establishing the prediction model of college students' sports performance, and others for testing the generalization performance of the model.

4. Results and discussion

In order to test the superiority of the prediction model of college students' sports performance based on machine learning algorithm, linear regression model and neural network were selected for comparative experiments, and their prediction accuracy of boys' 100m running performance was calculated. The accuracy is shown in Table 1. By comparing and analyzing the results in Table 1, the following conclusions can be obtained:

(1) Among the prediction models of all college students' sports performance, the linear regression model has the lowest prediction accuracy. This is because the linear regression model can only describe part of the change characteristics of boys' 100m running performance, and cannot describe the overall change characteristics. The prediction effect of college students' sports performance is poor.

(2) Compared with the linear regression model, the prediction accuracy of the neural network for college students' performance has been improved to a certain extent. This is because the neural network is a machine learning algorithm, which can fit the linear and nonlinear change characteristics of boys' 100m running performance, and achieve better prediction results. However, due to the requirement of a large number of samples, the prediction error of some points is large, and the prediction error increases accordingly, The prediction effect of the model is not good as a whole.

(3) Compared with linear regression model and neural network, support vector machine has the highest accuracy in predicting college students' sports performance. This is because the support vector machine requires fewer training samples, overcomes the defects of linear regression model, solves the limitations of neural network, and

improves the accuracy of boys' 100m running performance prediction, which has obvious advantages.

Table1. Superiority tests of different models

Model	Prediction accuracy/%
linear regression models	80.196
neural network modeling	88.63
support vector machine models	95.23

In order to test the universality of the prediction model of the machine learning algorithm for college students' sports performance, we used the results of a university's male triple jump, male 3000m run, male pull up, female 1000m run, and female standing long jump as the research objects, and counted their prediction accuracy. The results are shown in Table 2. From the prediction accuracy in Table 2, it can be found that the prediction accuracy of all college students' sports performance is more than 90%, far more than 85% of the actual application range, which indicates that the model is highly versatile and can be applied to the actual prediction of college sports performance.

Table 2. Tests of generalizability of predictive models for college students' sports performance

Type	Prediction accuracy/%
Boys' triple jump	93.84
Boys' 3000 m run	94.90
Boys' pull-ups	93.71
Girls' 1000 m run	96.07
Women's standing long jump	95.15

5. Conclusion

This paper proposes a personalized sports course design based on machine learning, sports training is an important factor to improve the physical quality of college students, and sports performance prediction can help universities to develop effective sports training programs. In order to improve the prediction effect of college students' sports performance and solve the defects of the current college students' sports performance prediction model, we constructed a college sports performance prediction model based on machine learning algorithm, and used the support vector machine, which has the best performance in machine learning algorithm, to model and fit the college students' sports performance data, and the results of the specific application examples show that the model has strong general performance and can be applied to various college students' sports performance prediction, and the prediction results are reliable. The results show that the model has strong general performance and can be applied to the prediction of various college sports, and the prediction results are reliable, which can provide valuable information for college sports training.

Reference

[1] Liu, L. . (2022). E-commerce personalized recommendation based on machine learning technology. Mobile information systems, 2022(Pt.5), 1-11.

[2] Doherty, T. , Yao, Z. , & Ahmad A.l. KhleifatHanz TantiangcoStefano TamburinChris AlbertynLokendra ThakurDavid J. LlewellynNeil P. OxtobyIlianna LouridaJanice M. RansonJames A. Duce. (2023). Artificial intelligence for dementia drug discovery and trials optimization. Alzheimer's & dementia: the journal of the Alzheimer's Association, 19(12), 5922-5933.

[3] Li, Z. , & Zhu, B. . (2022). Research on the design and implementation of middle school physical education curriculum based on the process model., 6(2), 49-52.

[4] Wang, D. . (2023). A study on the instructional design of interdisciplinary thematic learning in middle school physical education based on the cases-t model. Journal of Contemporary Educational Research, 7(12), 320-325.

[5] Baena-Morales, S. , Olalla García-Taibo, Merma-Molina, G. , & Ferriz-Valero, A. . (2023). Analysing the sustainability competencies of preservice teachers in spain. Journal of Applied Research in Higher Education, 15(3), 731-744.

[6] Smeltz, L. , Carpenter, S. , & Benedetto, LaurenNewcomb, NoraRubenstein, DanaKing, TonyaLunsford, ChristopherShaw, TanyaDewaters, Ami L. (2024). Introduction to disability and antiableist health care. American Journal of Physical Medicine and Rehabilitation, 103(5), 54-57.

[7] Hao, X. , & Yang, Y. . (2022). Healthy physical education curriculum model and students' extracurricular sports participation ——test based on the trans-contextual model of motivation. BMC Public Health, 22(1), 1-12.

[8]David Gutiérrez, Luis Miguel García-López, Hastie, P. A. , & Segovia, Y. . (2022). Adoption and fidelity of sport education in spanish schools:. European Physical Education Review, 28(1), 244-262.

[9] Muchalski, M. B. H. , Muchalski, H. , Hubert Muchalski Hubert MuchalskiDepartment of Chemistry and Biochemistry, California State University, Fresno, E. San Ramon Ave. MS SB, Fresno, California , United StatesMore by Hubert Muchalski, ,and, Hawker, M. B. M. J. , & Hawker, M. J. , et al. (2024). Quantized grading: an ab initio approach to using specifications grading in physical chemistry. Journal of Chemical Education, 101(2), 474-482.

[10](2023). Integrating curriculum ideology and politics into physical education on students' bipolar depression. CNS Spectrums, 28(S2), 83-83.

[11] Qiao, L. . (2022). Analyzing the teaching strategies for physical education in middle schools under the background of the new curriculum, 6(6), 5.

[12] Lis, A. , Snowdon, L. , & Downey, K. . (2023). Lessons learned after integrating the five steps of evidence-based practice into a doctor of physical therapy curriculum. Journal of Physical Therapy Education, 37(2), 127-131.

[13] Santos, L. . (2024). Evaluation of efnep peer educator facilitators and barriers to teaching physical activity to adult participants. Journal of Nutrition Education and Behavior, 56(8), 87-88.

[14] Williams, J. , Davies, M. J. , Suesee, B. , & Hunt, D. . (2022). Teachers' experiences of teaching the australian health and physical education health benefits of physical activity curriculum and the need for greater reality congruence. Curriculum Perspectives, 42(1), 27-37.

[15] Gao, K. , & Yu, F. . (2023). An improved gra algorithm for teaching quality evaluation of college physical education based on the probabilistic simplified neutrosophic sets. Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology, 45(4), 5199-5210.

[16] Juncheng, L. I. . (2023). A study on improving the teaching ability of college physical education teachers from the perspective of obe. Journalism and Mass Communication, 13(3), 128-133.

[17] Chao, Y. , Guo, X. , & Wang, R. . (2022). Exploration of teaching and education: integration of sports and medicine in physical education in colleges and universities, 6(1), 5.

[18] Liu, G. , & Zhuang, H. . (2022). Evaluation model of multimedia-aided teaching effect of physical education course based on random forest algorithm. Journal of Intelligent Systems, 31(1), 555-567.

[19] Ying, D. , & Liang, H. . (2022). A study of the rain classroom-based teaching mode of college english for art or physical education majors, 7(1), 146-155.

[20] Zeldin, E. R. , Charles, S. C. , Tumin, D. , Lawson, L. , Faulk, C. , & Norbury, J. W. . (2022). Impact of a hybrid-virtual teaching model on the physical examination skills of fourth-year medical students. American journal of physical medicine & rehabilitation, 101(10), 960-964.