

# Correlative Analysis and Prediction of Physical Education Data Via Machine Learning: A Case Study on Grade Evaluation Method in University

Jiaxin Hu<sup>1</sup>, Xincheng Zhang<sup>1</sup>, Shuangyue Xiao<sup>1</sup>, Linli Fan<sup>1</sup>, Li Liu<sup>1,\*</sup>, Tao Li<sup>2</sup>, Miaoqi Huang<sup>3</sup>

<sup>1</sup> School of Information Science and Engineering, Dalian Polytechnic University, Dalian, Liaoning, 116000, China

<sup>2</sup> Campus Administration, Liaoning Technical University, Fuxin, Liaoning, 123000, China

<sup>3</sup> Department of Human Resources, Dalian Polytechnic University, Dalian, Liaoning, 116000, China

\* Corresponding author: Li Liu (Email: liu.li@dpu.edu.cn)

---

**Abstract:** Physical education is an important way to strengthen the physical quality of students in the university. Sport performance is an important criterion for judging the physical fitness of college students. But during the process of calculating and collecting the comprehensive score for each student, teachers will use different equipment or various grading formulas to evaluate the grade for many years. Thus, referenced score will lose their values if the measurement data for each student in different years are not unified and sometimes with the subjective factor because of the manual calculations. This article analyzes the details of physical fitness test results and the relationship with the comprehensive score. It is discussed that the comprehensive score can provide serious help for teachers to know students' physical quality level and make reasonable teaching program to each student according to their personal radar chart. Therefore, the artificial influence is excluded, and the prediction model is designed by using the principal component analysis method and the back propagation (BP) neural network technology. The predictive model allows students to evaluate their own physical test score in advance, get a preliminary understanding of their physical fitness. Comparisons are made between the applications of this model in different years and errors are discussed to verify the accuracy. The results indicate that the comprehensive score prediction model supplies one effective approach to unify the scoring standards and improve the computation efficiency in physical education.

**Keywords:** Comprehensive Scores; Radar Chart; Back-Propagation Neural Network; Principal Component Analysis.

---

## 1. Introduction

In recent years, the rapid development of computer industry has led to social progress, and has promoted an increase in the amount of data in different fields. Further information can be obtained by analyzing and processing data. For example, the data analysis of clinical literature on the treatment of functional dyspepsia with acupuncture and moxibustion in the past two decades, summarizing the rule of point selection and treatment methods for the treatment of functional dyspepsia with acupuncture, provides a reference for the clinical application of acupuncture and moxibustion for the treatment of functional dyspepsia[1]; Overtaking incidents between electric vehicles and bicycles, and exploring the relationship between the number of overtaking incidents in existing non-motor vehicle facilities and traffic parameters. Then, it analyzes the actual data including flow, speed and overtaking event characteristics in detail, and the conclusions drawn are applied to the calculation of the number of overtaking events in the mixed non-motor vehicle flow and the setting of the width of the non-isolated non-motor vehicle lane[2]; the collection of 340 MLA (mobile library application) user data, structural equation modeling (SEM) with Bending Moment Structure Analysis (AMOS) software was performed to check quantitative data. These findings are provided guidance for effective decision-making in MLA design and development. The results can be used in the resource allocation process to ensure the successful realization of the library's vision and mission. With the continuous in-depth study of data, various methods of processing and analyzing data have been developed[3]; by

inviting student participation data and past performance data, classification and regression tasks are performed, and then artificial neural networks are used to obtain the highest overall Accuracy, providing comprehensive analysis and comparison, is the latest supervised machine learning technology used to solve the task of predicting student test scores, that is, discovering students in a "high-risk" dropout state, and predicting their future results, such as final exam results[4]. With the continuous development of machine learning and data analysis, various algorithms have also been properly applied. In the last few years, machine learning algorithms[5] [6], such as neural network[7], principal component analysis[8], support vector machine[9], etc., have been successfully applied in prediction models working for numerous research subjects[10]. Among them, neural network algorithm widely uses in fields of sport. In 2009, neural network algorithms were successfully used in popular NBA game to predict whether the basketball team can win based on the historical game data[11]. And the neural network algorithm in machine learning was applied to the selection of players in the annual national draft of the Australian Football League in 2010[12], which increased the probability of winning. It is suggested that neural network algorithms are capable to analyze the athletic ability of the participants, thus assist recruiting managers to make correct decision in the talent identification process, so in 2012, the Back-Propagation neural network algorithm was used to establish a model to predict the performance of historical Olympic athletes[13]. Besides, BP neural network model works very well in predicting the future development of traditional sport[14]. Apart from the neural network algorithm, principal

component analysis (PCA) has also been applied to study rehabilitation, triathlon and horse-riding coordination in sports science[15]. The result shows that the PCA is a useful method to characterize the movement coordination in its entirety[16][17]. For example, complex skiing movements were separated into the major motions by principal component analysis[18][19], which assist the coach to determine the standard postures and motions of the skier in teaching process. The neural network method was employed in combination with other methods as well. For instance, in order to predict more accurate sport performance, a hybrid prediction system is proposed based on genetic algorithm and artificial neural network. This hybrid prediction system was designed to study the effects of physical education on the athletic and language ability, social skills and work ability of children with intellectual disabilities[20][21]. Cluster analysis algorithm and neural network model were implemented to evaluate the athletic ability and performance of aerobics athletes. The evaluation results provide coaches with more helpful information[22][23]. Therefore, this paper combines BP neural network algorithm and principal component analysis to analyze the physical fitness test data of students in university. The predicted results and analyzed conclusions will support teachers to make a precise decision of the teaching program in physical education. Firstly, principal component analysis is applied to transform multiple attributes of strong correlations into independent attributes, then reduce training time and space by removing redundancy. Secondly, in addition to establishing the prediction model by BP neural network algorithms to forecast the students' comprehensive results in other years, the calculation standard will be unified simultaneously by reducing the manual effect.

As we know, millions of people died in the world due to the emergence and outbreak of COVID-2019. According to what has been recorded, human body immunity itself helps the most to fight against the corona virus. In university, doing physical exercises is an essential aspect of improving students' fitness and immunity[24]. Teachers should assign different physical education tasks and tests to better understand the current sensible quality of students and help them make proper training programs. But in the process of calculating the comprehensive results of many tests for each student, it is extremely complex because of various sport categories. Besides, the traditional grade evaluation method depends on teachers' subjective opinion very much and gives the results ineffectively. It is difficult to study the physical fitness level of the students directly by the measurement data precisely. Therefore, to establish a prediction model by using the total objective data like the weight and height, will on the one hand assist teachers to have knowledge of students' fitness condition and make more suitable physical exercises accordingly. Moreover, students themselves can also know their physical situation in advance, and then strengthen relevant training in a targeted manner in physical education classes. On the other hand can normalize the evaluation method of the comprehensive results.

## 2. Overview

It is commonly reported in the physical surveys the phenomenon of declining physical fitness of college students is prevalent all over the world[25][26]. More attention should be drawn on the Physical Education(PE) in university, which can directly reflect the physical health of college students. Thus, to study the students' physical test data and make clear visualization accordingly are very necessary. This will help

the students to know their health condition in detail and provide teaches more information to make teaching programs in PE class.

In general, there are some basic test items of physical fitness including vital capacity, 50-m sprint, sit and reach, standing long jump, 800/1000-m run, pull-up/bent-leg sit-up etcetera. Based on the mean and standard deviation, the sum of z-scores for the specific fitness tests can be calculated and used as a physical fitness index (PFI)[27]. Although this method can help to know the basic physical healthy status of college students, due to the different characteristics and the calculation methods of projects, the grade evaluation methods will be various. Besides, the formation of physical test score sometimes have the subjective components. Therefore, to have an accurate predictive model for college students' physical test score and standardizing analysis is of great research significance for understanding the changes in college students' physical fitness.

## 3. Model and Methods

### 3.1. Dataset

The data set of this study consists of students' physical fitness tests during 2016 to 2019 from Liaoning Technical University in northeast of China. The characteristics of the data include eight events:

**Height(H) and Weight(W):** The circumference, width, thickness and density of the human body are reflected through a certain proportional relationship between height and weight. It is an important indicator for evaluating the developmental level and nutritional status of human body.

**Vital Capacity (VC):** It is an important target for assessing the function of the human respiratory system. The size of vital capacity is closely related to factors of body such as weight, height, and chest circumference and others.

**Sprint (S):** 50m for man and woman: Measure the speed qualities of students, including speed of reaction, movement and muscles. It comprehensively reflects the physical qualities of students such as explosiveness, sensitivity, responsiveness, flexibility and others.

**Standing Long Jump (SLJ):** Test the strength of the legs and abdomen with the explosive power. It measures the strength of the lower limb muscles when jumping forward.

**Sit and Reach (SR):** Sit and reach test reflects flexibility of students. The quality of flexibility depends on the stretch of ligaments, tendons, muscles and skin. The decline of flexibility quality leads to a reduce in the physical fitness of students.

**Stamina Project (SP):** 1km for man and 800 m for woman: Middle and long distance running develops endurance quality of students. It is an aerobic metabolic program with large body load for students. It can reflect the tenacious volitional quality of the students.

**Power Project (PP):** Pull-up for man and sit-up for woman in one minute: Both pull-ups and sit-ups are methods for measuring muscle endurance and strength.

Because prediction efficiency can be accelerated by controlling values in the same range, so firstly the z-score standardization method which is based on the mean  $E(x)$  and standard deviation  $s_i$  of the original data  $x_i$  is used to unify data dimensions,

$$x = \frac{x_i - E(x)}{s_i} \quad (1)$$

In this way, all the original data are converted into dimensionless evaluation index values at the same quantitative level. Then, the correlation between data features is calculated. If there has strong correlation between attributes, the information reflected by the statistical data will overlap to some extent. The correlation coefficient  $r$  is,

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2)$$

$x_i$  and  $y_i$  are two data characteristics that participated in the calculation,  $\bar{x}$ - and  $\bar{y}$ - are mean values separately. Taking the data from the latest year 2019 as an example, the correlation between the attributes is observed in Table 1.

**Table 1.** The correlation between attributes (2019)

	H	W	VC	S	SLJ	SR	SP	PP
H	1.000	0.611	0.633	-0.540	0.598	-0.237	0.131	-0.675
W	0.611	1.000	0.544	-0.269	0.290	-0.188	0.333	-0.544
VC	0.663	0.544	1.000	-0.537	0.574	-0.144	0.131	-0.639
S	-0.540	-0.269	-0.537	1.000	-0.748	0.177	0.079	0.632
SLJ	0.598	0.290	0.574	-0.748	1.000	-0.150	-0.049	-0.662
SR	-0.237	-0.187	-0.144	0.177	-0.150	1.000	-0.123	0.329
SP	0.131	0.333	0.131	0.079	-0.049	-0.123	1.000	-0.249
PP	-0.675	-0.544	-0.639	0.632	-0.662	0.329	-0.249	1.000

Multiple original variables with strong correlation can be transformed into several unrelated comprehensive indicators by coordinate transformation. The new variables are called the principal components. They reflect the features of the original variables. Multiple principal components are extracted according to cumulative contribution rate. The information of the original variables is reflected as much as possible in accordance with actual needs. There are eight indicators in the original data, respectively  $x_1, x_2, \dots, x_8$ . New eight indices are obtained by linear combination transformation. The new indicators fully reflect the information of the original indicators according to the principle of retaining the main amount of information. The new variables  $y_1, y_2, \dots, y_8$  are independent of each other. The mathematical model of principal component analysis is as follows The model is organized into a matrix as follow,

$$\begin{cases} y_1 = \mu_{11}x_1 + \mu_{12}x_2 + \dots + \mu_{1i}x_i + \dots + \mu_{18}x_8 \\ y_2 = \mu_{21}x_1 + \mu_{22}x_2 + \dots + \mu_{2i}x_i + \dots + \mu_{28}x_8 \\ \vdots \\ y_i = \mu_{i1}x_1 + \mu_{i2}x_2 + \dots + \mu_{ii}x_i + \dots + \mu_{i8}x_8 \\ \vdots \\ y_8 = \mu_{81}x_1 + \mu_{82}x_2 + \dots + \mu_{8i}x_i + \dots + \mu_{88}x_8 \end{cases} \quad (3)$$

$$y = \begin{bmatrix} \mu_{11} & \mu_{12} & \dots & \mu_{18} \\ \mu_{21} & \mu_{22} & \dots & \mu_{28} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{81} & \mu_{82} & \dots & \mu_{88} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_8 \end{bmatrix} = Ux \quad (4)$$

and the sum of principal component coefficients is 1.

$$\mu_{i1}^2 + \mu_{i2}^2 + \dots + \mu_{i8}^2 = 1 \quad (5)$$

Therefore, in order to obtain the principal component value  $y$ , it is necessary to calculate the coefficient  $\mu$ . Firstly, the covariance matrix is calculated by the original data,

$$\text{cov} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \quad (6)$$

Results indicate that there has strong correlation between some data features from 2019. For example, the correlation coefficient between Sprint (S) and Standing Long Jump (SLJ) reached 0.7478. Besides, it should be noticed that there will be a negative correlation for items using time unit, because less time means higher score. When using these data sets to train a model, the strong correlation will lead to information redundancy and the weights connected to the input neurons in the network would bring similar effects to the model. Thus, the principal component analysis method will be used to eliminate the strong correlations between data.

### 3.2. Principal Component Analysis (PCA)

Principal component analysis (PCA) is a technique to reduce the dimension of datasets.

Because data has been z-score standardized, so the variance  $s^2$  of the data is 1.

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = 1 \quad (7)$$

Then

$$\sum_{i=1}^n (x_i - \bar{x})^2 = ns^2 = n \quad (8)$$

so, the correlation coefficient  $r$  is,

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{n}\sqrt{n}} = \text{cov} \quad (9)$$

Therefore, the covariance matrix is equivalent to the correlation coefficient matrix. In order to retain most of the data information, all principal components are used for calculation. New eight variables which uncorrelated each other will be retained. Then the new variables are trained by neural network model.

### 3.3. Back-propagation Neural Networks

Artificial Neural Network (ANN), also known as neural network is originated from neurobiology. The network composed of a large number of neural cells, which is an abstraction, simplification and simulation of the human brain. As a part of the neural network, back-propagation (BP) neural network is a supervised learning algorithm. It is a multilayered nonlinear feed-forward network trained by the back-propagation learning algorithm. Network is composed of input layer, hidden layer and output layer. The BP learning algorithm consists of two processes, one is the forward computation of data and the other is reverse propagation of error signal.

Forward Propagation Stage The forward propagation stage means that the source data from the input layer to output layer through hidden layer. The output of the upper layer node is the

input of the lower layer. Then signal is propagated to the output layer to measure whether the training of neural network is completed or no by the error function. Figure 1 shows that the number of nerve cell in the input layer, hidden layer and output layer. The number of nerve cells is  $n$ . The exact value of  $n$  is confirmed according to established model, in this article the numbers are 8, 11 and 1 respectively. Each nerve cell has a weight of linear calculation at each output level. The output value of hidden layer  $net1_k$  is calculated by the input value  $y_i$ , connection weight value  $w_{ik}$  and the threshold value  $b_k$  are set as follows,

$$net1_k = \sum_{i=1}^8 w_{ik}y_i + b_k = w_{1k}y_1 + w_{2k}y_2 + \dots + w_{8k}y_8 + b_k \quad (10)$$

The activation function can be applied to obtain more accuracy in the predicting process. The sigmoid function which is a common nonlinear function is used in this paper. The calculated value  $net1_k$  is activated by activation function  $f(net1_k)$ . The hidden layer  $z_k$  is obtained by formula,

$$z_k = f(net1_k) = \frac{1}{1 + \exp(-net1_k)} \quad (11)$$

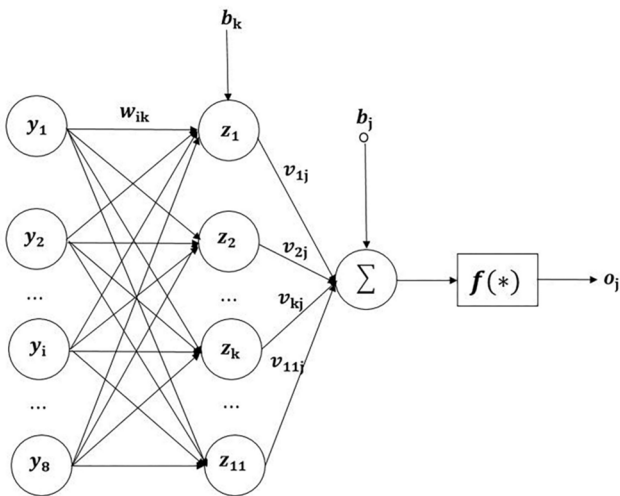
Next, the hidden layer is used as input. The export value of output layer  $net2_j$  is calculated by hidden layer  $z_k$ , connection weight value  $v_{kj}$  between hidden and output layer and the threshold value  $b_j$ ,

$$net2_j = \sum_{k=1}^{11} v_{kj}z_k + b_j = v_{1j}z_1 + v_{2j}z_2 + \dots + v_{11j}z_{11} + b_j \quad (12)$$

The calculated value  $net2_j$  is activated by activation function  $f(net2_j)$  and the output layer  $o_j$  is obtained by formula,

$$o_j = f(net2_j) = \frac{1}{1 + \exp(-net2_j)} \quad (13)$$

Error Back Propagation Stage Training process stops when the output error function is less than the predetermined value. Square error function is used to measure the error size between the actual output  $d_j$  and the expected output  $o_j$ ,



**Fig 1.** Basic structure of forward propagation stage BP neural network

$$E = \frac{1}{2} \sum_{j=1}^n (d_j - o_j)^2 = \frac{1}{2} \sum_{j=1}^n \left( d_j - f\left(\sum_{k=1}^{11} v_{kj}z_k\right) \right)^2 \quad (14)$$

The error is reduced along the gradient direction by adjusting the weights and thresholds. Then the weight correction value  $L \setminus v_{kj}$  is calculated. The weights of each network are updated.

$$\Delta v_{kj} = -\eta \frac{\partial E}{\partial v_{kj}} = -\eta \frac{\partial E}{\partial net2_j} \frac{\partial net2_j}{\partial v_{kj}} = -\eta \frac{\partial E}{\partial net2_j} z_k \quad (15)$$

Then,

$$v_{kj} = v_{kj} + \Delta v_{kj} \quad (16)$$

The error is extended to the input layer,

$$E = \frac{1}{2} \sum_{j=1}^n \left( d_j - f\left(\sum_{k=1}^n v_{kj} f\left(\sum_{i=1}^8 w_{ik}y_i\right)\right) \right)^2 \quad (17)$$

The weight adjustment amount  $\Delta w_{ik}$  between the input and the hidden layer are calculated.

$$\Delta w_{ik} = -\eta \frac{\partial E}{\partial w_{ik}} = -\eta \frac{\partial E}{\partial net1_k} \frac{\partial net1_k}{\partial w_{ik}} = -\eta \frac{\partial E}{\partial net1_k} y_i \quad (18)$$

Then,

$$w_{ik} = w_{ik} + \Delta w_{ik} \quad (19)$$

The forward propagation of the signal is carried out when all weights are readjusted. Training is stopped when the convergence criteria are reached. Next, grade for each student can be predicted based on the model.

### 3.4. Results and Discussion

During the teaching process, a reasonable advice for each student should be proposed according to the data from previous years. In this paper, the physical fitness test data for students enrolled in 2016 and graduated in 2019 is analyzed to see the change of students' fitness condition in the four years. Based on this, a grade prediction model is established to help teachers and students make a better training schedule. Besides, the grade evaluation method is discussed.

## 4. Physical Fitness Test Data Analysis

In order to put forward a reasonable training plan for students with different physical qualities, students are divided into five groups according to comprehensive results category and the average of eight attributes. According to the 'National Physical Health Standard', students' comprehensive results have four categories including excellent (Score above 90), good (Score Between 80 and 89.9), pass (Score Between 60 and 79.9) and fail (Score below 60). Besides, the mean value of eight attributes in four categories are calculated for analysis. Here, female and male students are discussed separately due to the different measurement projects and scoring standards. Based on the four comprehensive results categories and taking the mean value of eight attributes in each category into account, students are rearranged into five groups including 'perfect', 'with weakness', 'middle', 'passing' and 'low score'. The radar chart is used to visually analyze the data of multiple attributes, which can clearly exhibit the change of student's physical fitness in four years. The range of each direction in the radar chart is the minimum or maximum value

of each attribute. Female and male students from each group are randomly selected in 2019 shown in Figure 2 and 3 separately because of their different scoring standards and measuring items. In this way, teachers and students themselves are able to aim around organize the teaching programs precisely. For the students from the ‘Excellent’ category, if their measurement data for each test project is higher than the average, they have excellent physical fitness indicators and belong to the ‘perfect’ group. Figure 2(a) and 3(a) show the selected male and female students from this group in 2019. From the radar chart, it is obvious that the Power Project (PP) of this male student has increased greatly in 2019. One reason would be because this student has strengthened his muscle training. The female’s measuring results in recent years are relatively stable. Intensity training

can be implemented to this kind of students on the basis of safety to inspire the potential of specialty and then be selected as athlete’s candidates to participate in various competitions.

Although in both ‘Excellent’ and ‘Good’ categories, students usually perform very well and have good scores, there are also some students with poor performance in single measurement. If having one score in a test project lower than 60% of the average value, then this student is considered as having personal individual weakness and belong to the ‘individual weaknesses’ group. In Figure 2(b), the Sit and Reach (SR) measurement result of this student is much lower than the average data of 2019. He has a significant weakness in this project. Similar data trend is shown in Figure 3(b), which describes a disadvantage in Vital Capacity (VC) of a female student.

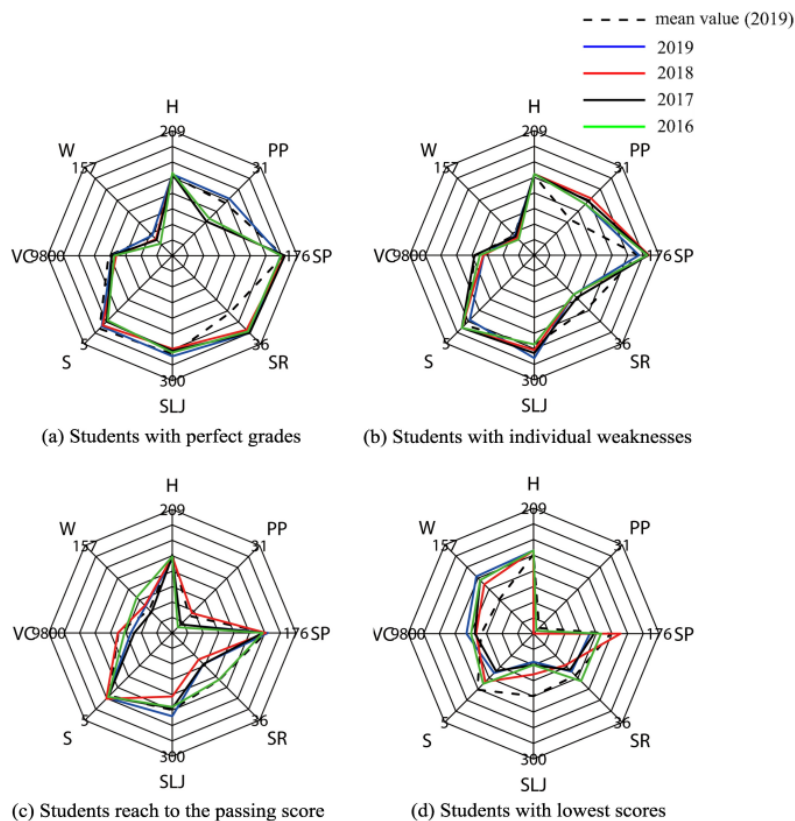


Fig 2. Four years data of selected male student

For the rest students in ‘Good’ categories, they belong to the ‘middle’ group. In the meanwhile, the students in the ‘Pass’ category with a score greater than 65 are grouped into ‘middle’. For students in ‘middle’ group, because their fitness level is at a normal stage, so to take general quantitative exercises under teachers’ supervision is well suggested.

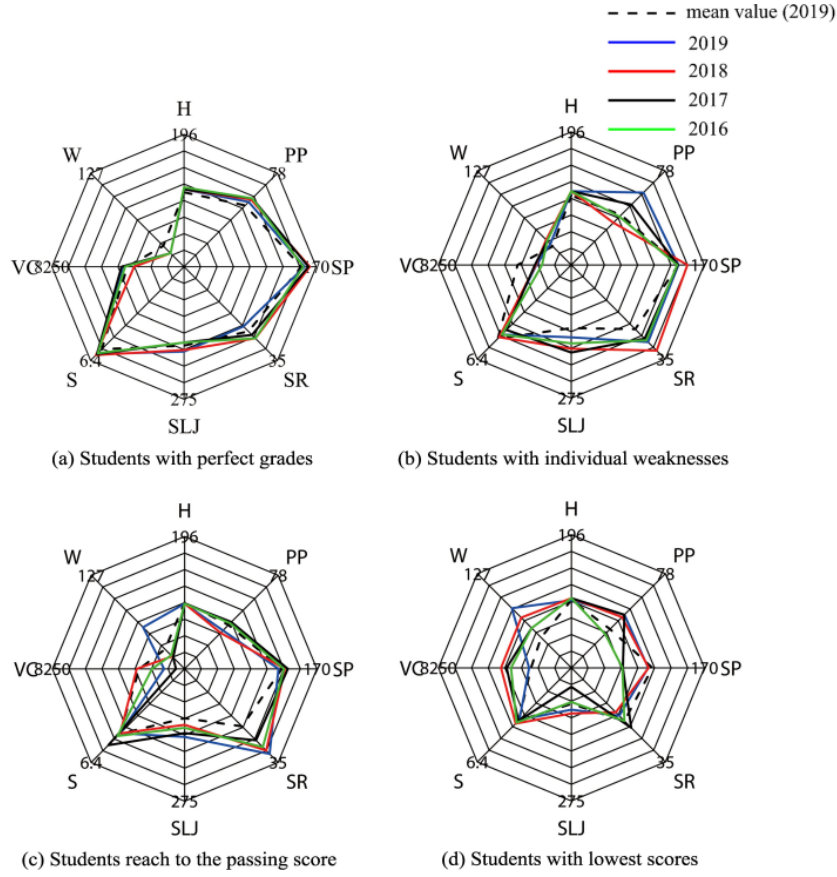
For the students with score between 55 and 65, they are considered as close to the passing score and belong to the ‘passing’ group. Figure 2(c) gives the information of a male student from this group. All physical indicators of this male student are unsatisfactory. The Sit and Reach (SR) test result becomes very low since 2017.

Fig. 3(c) shows the results and change trend of a female student. Shown in radar chart, it is clear that the weight of this student has increased rapidly in 2019. For these two selected students, physical fitness needs to be strengthened through multiple projects because of their low physical indicators.

Students with scores below 55 are regarded as in ‘low score’ group. Figure 2(d) and 3(d) present that the test results are not

stable and mostly lower than the average of ‘Fail’ category. Particularly, weight (W) of these two students are significantly higher than the average. And it should be aware that obesity is a common feature of students from this group, so their physical indicators are declined notably. For the time being, poor physical fitness can cause many physical illnesses and especially these students will face the risk of failing to graduate due to the unqualified grade.

In general, if compare the data of selected students directly by the area of the closed line displayed in the radar chart, the better the physical fitness of the student, the larger area of his personal radar curve he or she will have. The statistical results of physical fitness test over the years can provide a reference for the basic physical condition of students. Accordingly, teachers and students themselves can better understand their actual physical situation and then adjust training programs in time.



**Fig 3.** Four years data of selected female student

Corresponding exercise intensity will be well provided to students in the condition of different physical fitness levels.

#### 4.1. Prediction Results

In the previous section, it was discussed that the comprehensive score can help to better understand the physical condition of students and develop a reasonable training plan. However, sometimes due to the differences in the characteristics of the course and subjective manual calculations. It is difficult to obtain accurate comprehensive score for all students over the years under one standard. In order to unify the scoring standards and eliminate human influence, a comprehensive scoring prediction model is set up based on a variety of machine learning algorithms.

The principal component analysis method is used to eliminate the strong correlations between the standard data. And then 80% of the student samples in 2016 are randomly selected as the training set in BP neural network. The rest 20% of the data are taking as the test set to evaluate the accuracy of the model. The established model is applied to predict the comprehensive score of the physical fitness of the students enrolled in 2016 and are about to graduate in 2019. The application effect of the model is observed by comparing the predicted and actual values shown in Figure.4. The two almost overlapped lines show that the predicted values of selected data from test set are close to the actual ones. To have a better understanding of the prediction reliability, the Absolute Error is calculated by actual data  $o_i$  and forecast data  $o_j$ ,

$$AbsoluteError = o_i - o_j \quad (20)$$

Figure 5 shows the frequency distribution of the Absolute Error (AE). The dashed line gives the trend of normal distribution. It can be seen that the statistics of AE appear a maximum around zero and approximately take up two-third of total area between -1 and 1. The frequency is very small when the error value is greater than |3|. Besides, mean square error(MSE) can also be used to further evaluate the accuracy of prediction. MSE is calculated according to the absolute error and sample size,

**Table 2.** AE regions predicted by the model established with 2016 data

	[0,1)	[1,2)	[2,3)	[3,4)	[4,5)	[5,6)
2016	64.38%	28.57%	6.27%	0.7%	0.04%	0.02%
2019	40.84%	27.16%	15.71%	8.06%	4.06%	4.17%

$$MSE = \frac{1}{n} \sum_{i=1}^n (o_i - o_j)^2 \quad (21)$$

It can be found that the prediction performance of this model is excellent because the MSE value is very small about to 1.362. By considering all the error results, this model is able to be applied to predict comprehensive test scores in other years. First of all, the measured data in 2019 is also standardized and then analyzed by principal component analysis. The data of different gender in 2019 is brought into two models for prediction. The predicted values and absolute errors are calculated. Here, the absolute value of AE is divided into six intervals. The percentage of AE distribution is compared between 2016 and 2019, shown in Table 2. It shows that the predicted results for 2016 are excellent, 92.95 percent of the absolute error is lower than 2, and only 0.06 percent is greater than 4. The accuracy of model prediction

has decreased when the model is applied to 2019. The distribution of absolute error between 0 and 1 decreased by 23.54%. An important reason for the difference in prediction accuracy between 2016 and 2019 is the different scoring standards used by teachers over the years. In addition, it shows that manual calculation has potential influence on the comprehensive score of physical fitness test.

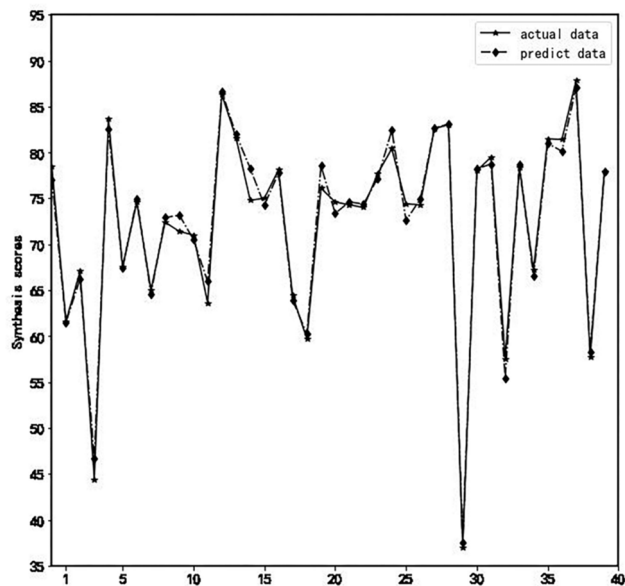


Fig 4. Comparison of predicted data with actual data samples

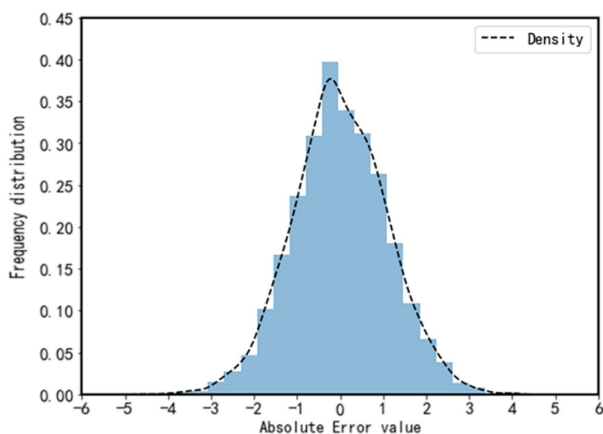


Fig 5. Error value frequency distribution.

The learning ability of the model can be used to explore the constant relationship between the project measurement data and the comprehensive score. This method can avoid the complicated process of traditional physical grade evaluation method and save computation time. Next, combined with visualized data information in radar chart, the changes of students' physical condition can be analyzed clearly, and it is of great significance for the physical education teacher to adjust the next teaching plan to the students. According to the predictive model, students can predict their own physical test results, and the next physical education classes can be targeted for training by themselves, aiming to get better physical test results in the subsequent physical tests. Therefore, it is very necessary to use the model to predict the comprehensive score of physical fitness tests in future years under the condition of keeping the same test projects.

## 5. Conclusion

In this paper, by analyzing the comprehensive score of the physical fitness test and the mean values of each attribute, a scientific grouping method is proposed and discussions based on case study from each group are performed through radar chart illustration. Teachers can make reasonable teaching plans according to the graphic information effectively and give targeted advices to students in different groups. Besides, principal component analysis and BP neural network have been successfully applied to set up a prediction model. Compared with the traditional process, this grade evaluation method improves the calculation efficiency and unifies the scoring standard in different years.

## Acknowledgments

Authors thank for the financial support by Education and Teaching Reform project of Dalian Polytechnic University (Grant: JGLX2021108) and would like to express many thanks to the support of the school of International Education, Dalian Polytechnic University.

## References

- [1] Li YB. Analysis of clinical literature on acupuncture-moxibustion for dyspepsia based on data mining. *J Acupunct Tuina Sci.*, 17:264–269, 2019.
- [2] Yan Xingchen, He Peng, Chen Jun, Ye Xiaofei, and Liu Qingchao. Relationship between number of passing events and operating parameters in mixed bicycle traffic. *Journal of Southeast University (English Edition)*., pages 418–423, 2015.
- [3] Hamaad Rafique, Alaa Omran Almagrabi, Azra Shamim, Fozia Anwar, and Ali Kashif Bashir. Investigating the acceptance of mobile library applications with an extended technology acceptance model (tam). *Computers & Education.*, 145:103732, 2020.
- [4] Nikola Tomasevic, Nikola Gvozdenovic, and Sanja Vranes. An overview and comparison of supervised data mining techniques for student exam performance prediction. *Computers & Education.*, 143:103676, 2020.
- [5] Claude Sammut and Geoffrey I. Webb. *Encyclopedia of machine learning and data mining sentiment analysis*. 10.1007/978-1-4899-7687-1:Chapter 100512, 2017.
- [6] Ioannis Kavakiotis, Olga Tsave, Athanasios Salifoglou, Nicos Maglaveras, Ioannis Vlahavas, and Ioanna Chouvarda. Machine learning and data mining methods in diabetes research. *Computational Structural Biotechnology Journal.*, 15:104–116, 2017.
- [7] Christoph Helma, Tobias Cramer, Stefan Kramer, and Luc De Raedt. Data mining and machine learning techniques for the identification of mutagenicity inducing substructures and structure-activity relationships of noncongeneric compounds. 35:0–0, 2004.
- [8] Singh Poornima, Singh Sanjay, and Gayatri S. Pandi-Jain. Effective heart disease prediction system using data mining techniques. *International Journal of Nanomedicine.*, Volume 13:121–124, 2018.
- [9] Gang Song, Guoqiang Xiao, Yang Jiang and Jianmin Jiang. Support-vector-machine tree-based domain knowledge learning toward automated sports video classification. *Optical Engineering.*, 49:12, 2010.
- [10] Jian. Wang. *Information security in big data: Privacy and data mining*. 2014.

- [11] Earl Bednar Bernard Loeffelholz and Kenneth W. Bauer. Predicting nba games using neural networks. *Journal of Quantitative Analysis in Sports.*, 5:7–7, 2009.
- [12] Angelo Maria Sabatini. Andrea Mannini. Machine learning methods for classifying human physical activity from on-body accelerometers. *Sensors (Basel, Switzerland).*, pages 1154–1175, 2010.
- [13] Bahadorreza Ofoghi, John Zeleznikow, Clare Macmahon, and Dan. Dwyer. Supporting athlete selection and strategic planning in track cycling omnium: A statistical and machine learning approach. *Information ences.*, 233:200–213, 2013.
- [14] Jian Wang. Bp neural network-based sports performance prediction model applied research. *Journal of Chemical and Pharmaceutical Research.*, 2014.
- [15] John. McCullagh. Data mining in sport: A neural network approach. 2010.
- [16] Kerstin Witte, Nico Ganter, Christian Baumgart, and Christian. Peham. Applying a principal component analysis to movement coordination in sport. *Mathematical and Computer Modelling of Dynamical Systems.*, 16:477 – 488, 2010.
- [17] P Federolf, R Reid, M Gilgien, P Haugen, and G. Smith. The application of principal component analysis to quantify technique in sports. *Scandinavian Journal of Medicine and Science in Sports.*, 24:491–499, 2014.
- [18] Balagué Natàlia, González Jacob, and Casimiro. Cardiorespiratory coordination after training and detraining. a principal component analysis approach. *Frontiers in Physiology.*, 7, 2016.
- [19] Myvind, Gloersen, Havard, Myklebust, Jostein, Hallén, Peter, and Federolf. Technique analysis in elite athletes using principal component analysis. *Journal of sports sciences.*, 2017.
- [20] Biao. Xu. Prediction of sports performance based on genetic algorithm and artificial neural network. *International Journal of Digital Content Technology and its Applications.*, 6:141–149, 2012.
- [21] Peng. Sun. Genetic neural network-based study on the impact of sports on the social adaptability of mentally challenged children. *Research Journal of Applied ences Engineering and Technology.*, 5:2771–2778, 2013.
- [22] Yan. Yu. Neural network model aerobics athlete athletic ability and athletic performance assessment. *Information Technology Journal.*, 12:3303–3308, 2013.
- [23] ZH (Hong Zhi-hua)1; Zhou ZY (Zhou Zhi-yong)1. Wang, J (Wang Jian)1; Hong. Clustering analysis of sports performance based on ant colony algorithm. 2014 Fifth International Conference on Intelligent Systems Design and Engineering Applications (ISDEA)., pages 288–291, 2014.
- [24] Wang Jiahong, Ping Xiang, Zhang Dazhi, Weidong Liu, and Xiaofeng. Gao. Evolution of physical education undergraduate majors in higher education in china. *Journal of Teaching in Physical Education.*, 36:51–67, 2008.
- [25] Mónika Kaj, éva Tékus, Ivett Juhász, Katinka Stomp, and Márta. Wilhelm. Changes in physical fitness of hungarian college students in the last fifteen years. *Acta biologica Hungarica.*, 66:270–81, 2015.
- [26] Peter Pribis, Carol A. Burtneck, Sonya O. Mckenzie, and Jerome. Thayer. Trends in body fat, body mass index and physical fitness among male and female college students. *Nutrients.*, 2:1075–1085, 2010.
- [27] Xiaobin Chen, Jie Cui, Yuyuan Zhang, and Wenjia. Peng. The association between bmi and health-related physical fitness among chinese college students: a cross-sectional study. *BMC Public Health.*, 20:1–7, 2020.