

Prediction of Physical Ability of College Students Based on the BP Neural Network

Xu Zhang¹

¹ College of Physical Education and Arts Humanities, China University of Petroleum (Beijing), China

Correspondence: Xu Zhang, College of Physical Education and Arts Humanities, China University of Petroleum (Beijing), China.

Received: November 19, 2023

Accepted: February 1, 2024

Online Published: May 29, 2024

doi:10.5539/ies.v17n3p92

URL: <https://doi.org/10.5539/ies.v17n3p92>

Abstract

The prediction of physical ability is a key point to understand the physical training effect of college students. This paper uses the error Back Propagation neural network algorithms to investigate the college students' physical test results, and predicts the future trends of the results. The findings indicate that, in future ten years, the physical test grades of both men and women will slip to a certain extent. It implies that the indicators of speed, power, endurance level and physical strength of college students keep going down. This phenomenon should be paid attention to. It is necessary to strength the exercise consciousness, and stimulates the enthusiasm of students' exercise.

Keywords: physical ability, BP neural network, college students

1. Introduction

In recent years, the physical health status of college students has been paid more and more attention to by the society, major colleges and universities actively carry out various forms of college sports. A large volume of historical data on college students' physical health and sports achievements is collected. With further data mining, it is meaningful and important to forecast health status of college students. (Zhang et al., 2013) The forecasts will not only help college lecturers to set reasonable and effective physical education curriculum goals for students, in order to improve their physical fitness, but also tracks the physical health status of college students. Zhang (2018) used the big data analysis to evaluate college students' sports learning engagement factors. Then he concluded that more attention should be paid to the sports construction of college students. Therefore, designing an efficient big data prediction model is an important research problem in the field of college sports. In recent years, the results of various physical tests show that the physical fitness of college students has a downward trend year by year. The average height and weight of grade 2017 are higher than grade 2009. The peak values of average height and weight in grade 2017 are also higher than those in grade 2009. By comparing the physical test scores, it shows that the physical test scores of male students in grade 2017 are lower than those in grade 2014, but the situation of female students is different. There is little improvement in the 50-meter run and the 1000-meter race, but the scores for other tests fall sharply.

The reason lies in the fact that college students in the process of entering the society, they just start to live independently. And the phenomenon of the contemporary college students staying up late and staying overnight becomes popular in college students. College students have insufficient physical exercise, lacking of sleep and mental stress are considered to be the main reasons for the decline of college students' physical health. The decline of physical fitness of contemporary college students has no advantage in the long run. It is necessary for physical educators to pay attention to the future physical trends of college students in China, and to make targeted physical training plans and projects.

2. Literature Review

Since the reform and opening up, China's economic construction has made brilliant achievements. At the same time, the per capita income of urban and rural residents and per capita consumption level has also increased greatly. Zhang Li (2002) argues that, the improvement of living conditions has a positive impact on the physical fitness of cohort people. But in recent years, due to the rapid economic development, e-commerce services and network functions have become more popular and accessible, people prefer the delivery service than self-pickup.

Due to the boom of the e-competition industry, more college students prefer online games rather than physical exercise in leisure time. Survey data show that only 5.1% of the total number of contemporary college students adheres to daily exercise. Many college students choose irregular life, as a result of increased learning stress and the accelerated pace of school life. Combined with a decrease in exercise, there is evidence that the unhealthy situation of contemporary college students continues to increase. It can be seen that, with the rapid economic development and better living conditions, the nutrition of contemporary college students has improved significantly. However, the physical fitness of college students continues to decline, in the form of speed, explosiveness, endurance level and strength (Tan & Zhang, 2019; The State Council, 2014; Wang, 2014).

In the face of the grim medical condition of college students today, the state and all sectors of the society launched a number of measures to improve the physical condition of college students. Since 1990, the state has begun to implement the standard of physical exercise (later changed to the national student physical health standards), with the body shape, physical function and physical fitness of the trinity of evaluation methods to make a systematic evaluation of students' physical health. The state hopes to urge students to carry out physical exercise, develop strong physique and improve their physical fitness through these normative standards. Over time, all sectors of society became more aware of the importance of physical health. In recent years, the 'Healthy China 2030' Outline, issued by the State, clearly states that 'Healthy China' has become a national development strategy. In recent years, schools have carried out a wide range of sports activities and sports program skills training, while vigorously promoting the construction of student sports associations. At the same time, the national fitness boom is also rising in the whole society. In the process, the public service system of national fitness has been improved and the modernization of national fitness governance has been promoted. The fitness infrastructure has been improved and the number of sports events throughout the country has steadily increased.

Through the use of competition, more college students participate in sports. They can find the fun of sports and improve the enthusiasm of physical exercise. The development of campus basketball and campus football is based on entertainment and fitness, with athletics as a demonstration. It provides a platform for students to cultivate interest, self-conscious in sports, and physical fitness. More and more students are benefiting from this project (Zhu & Liu, 2020; Wang, 2019). In 2016, the State Council Issued the National Fitness Plan (2016-2020), which stated that "... improving the physical fitness and health level of the whole nation is the fundamental goal..." Today, the whole society set off a boom of national fitness and healthy, which stimulates their enthusiasm for exercise in various ways. Improving the physical fitness of future college students has become a hot topic. Zhao (2016) argues that, from the psychological aspects, college students should participate more in physical exercise, which have a positive impact not only for physical health, but also for mental health (Wang et al., 2020). Fricke et al. (2018) made a research on the effects of incentives to exercise on student performance in college. The findings indicate that financial incentives will encourage college students to exercise. College students were willing to participate more in off-campus physical activities.

The application and development of college students' sports achievement prediction method is divided into the following three stages. The initial stage is the predictions about the physical fitness and physical achievement of college students based on historical statistics and experience, which are collected by physical teachers in colleges and universities. The second stage is mainly the application of the mathematical modeling method based on statistical theories. Bandyopdhyay and Chattopadhyay (1981) used Harvard step test to access the physical fitness of college students. Zhao (2001) established a multi-return prediction model of 30m running, short running four-level jump, medium-range eight-step technology, deep squat and athletic performance. Zhu (2003) predicts the sports performance through GM (1,1) gray prediction model. Huang (2005) uses the main component regression analysis to establish a model to predict the performance of 50m and 800m sports. Liu (2014) puts forward a model based on nonlinear screening factors and weighted sports performance prediction which improved the prediction accuracy of sports performance. Carter (2012) examines the relationships between the motivational variables and academic and athletic performance. The combination of variance (ANOVA) and multivariate analysis of variance (MANOVA) is used to predict the academic performance of student-athletes. Most of these prediction methods can only describe the linear or upward trend change characteristics of sports performance, while the characteristics of its nonlinear trend change scant recognition ability is low. Therefore, the prediction accuracy is sometimes not ideal. With the development of artificial intelligence and machine learning, predictive models based on deep learning and data mining have been widely used in various fields. It overcomes the shortcomings of traditional models and improves the speed and accuracy of prediction. Zhang et al. (2018) proposed a sports competition performance prediction model based on fruit fly algorithm to optimize the extreme learning machine, the efficiency is greatly improved. Wang (2017) used the support vector machine to establish a prediction model of college students' sports performance. Particle group algorithm is used to select model

parameters, and it is applied to a university's sports performance modeling and prediction used to predict sports performance; Zhang et al. (2018) establishes the prediction model of College Students' sports achievement by using machine learning algorithm-support vector machine. Used particle swarm optimization algorithm is used to select model parameters. As a result, the conclusion shows that machine learning algorithm can overcome the shortcomings of the traditional model and improve the accuracy of prediction of College Students' sports performance. Zhang et al. (2013) and Zhang (2017) used BP neural network and sports performance prediction methods for the prediction, which improves the prediction accuracy. However, the sports achievements of college students are not only influenced by linear factors, but also by many nonlinear factors. Artificial neural networks, as a research method in nonlinear fields, have powerful learning functions and can learn historical data, so as to realize the prediction of future conditions. The algorithm of error reverse propagation neural network has a strong nonlinear modeling ability and can be adjusted adaptively. Using this algorithm, this paper investigates the physical health and sports performance of students at a Chinese University.

3. Analysis Methods and Ideas

3.1 Neural Network with Reverse Propagation of Error

Error-reverse propagation neural network, also known as BP neural network, is an important branch of machine learning. Its structural model is a network of simple neurons connected to each other in different ways. Through neural network learning algorithms, the received information is processed and the memory is formed. The learning process is divided into two stages: the first stage is to enter the known learning sample. The current network receives information, and uses the weight and threshold of the iteration results to process the information from the first layer to the last layer. Finally, it gets the output of the last layer of neurons, that is, the forward transmission of the working signal. The second stage is to calculate the error between the actual output of the current network and the expected output. Then, the error is passed from the last layer to the first layer. The error signal modifies the weights and thresholds between layers. The two stages alternate over and over again until the error reaches convergence.

In this paper, the error reverse propagation neural network prediction model, which is proposed by Zhang (2019), is used. This model is based on machine learning algorithm, and includes input layer, implied layer, output layer and transfer function between layers. The neurons are fully connected. There is no connection within the layer, and the implied layer of the BP network can be infinitely increased, theoretically. The topology model of the three-layer BP neural network is shown in the following illustration:

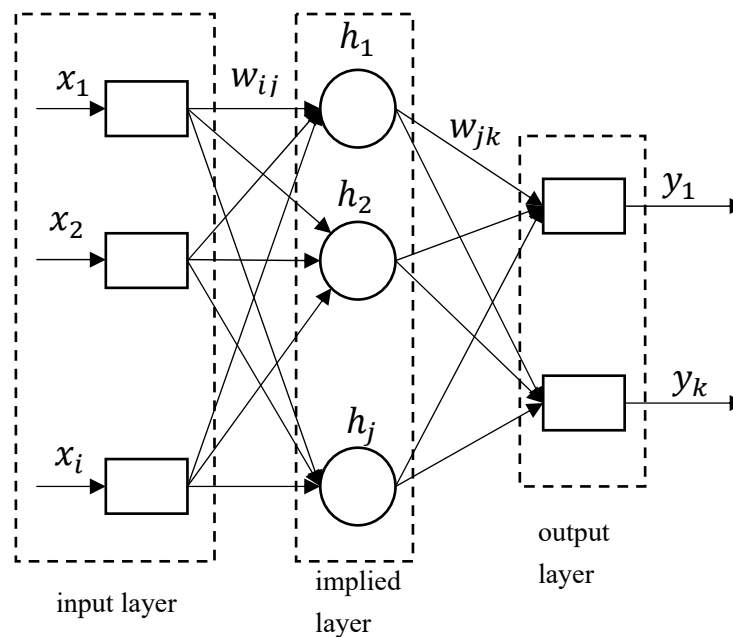


Figure 1. Three-layer neural network topology graph

As shown in Figure 1, the input layer has i neurons, the input vector is $X = [x_1, x_2, \dots, x_i]^T$; the implied layer has j neurons, the vector is $H = [h_1, h_2, \dots, h_j]^T$; the output layer has k neurons, the vector is

$Y = [y_1, y_2, \dots, y_k]^T$. w_{ij} is the weight between the input layer and the implied layer, w_{jk} is the weight between the implied layer and the output layer. The threshold of each neuron of the implied layer is θ_j , and the neuron threshold of the output layer is θ_k . Thus, the output of the j neuron in the implied layer can be expressed as:

$$h_j = f_1(\sum_{i=1}^J w_{ij}x_i + \theta_j) \quad (1)$$

The output of the k neuron in the output layer is:

$$y_k = f_2(\sum_{j=1}^K w_{jk}h_j + \theta_k) \quad (2)$$

$f_1(\cdot)$ represents the activation function of the implied layer. f_1 is the Sigmoid-type activation function, and f_2 is the line-type activation function. The activation function is an important part of the error-reverse propagation neural network and must be continuous and differentiable. The function uses an S (Sigmoid) log or tangent and linear function, where the Sigmoid function can implement any nonlinear mapping between input and output. Among them, the most used is the Sigmoid function, which can map the data in the range of $(-\infty, +\infty)$ to the interval $(0,1)$. it can be expressed as:

$$\text{logsig}(x) = \frac{1}{1+e^{-x}} \quad (3)$$

During the error reverse propagation, the error function of the output layer can be represented as:

$$E = \frac{1}{2} \sum_{k=1}^K (d_k - y_k)^2 \quad (4)$$

The formula d_k indicates the desired output value of the k neuron in the output layer. In this paper, the output layer contains an output unit, which explains the error propagation process with only one output unit. Error function w_{jk} takes the derivatives to the connection weight value between the implied layer and the output layer, we can get:

$$\frac{\partial E}{\partial w_{jk}} = \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial f_2} \frac{\partial f_2}{\partial w_{jk}} \quad (5)$$

Then,

$$\frac{\partial E}{\partial w_{jk}} = - (d_k - y_k) y_k (1 - y_k) h_j \quad (6)$$

Thus, the connection weight between the implied layer and the output layer w_{jk} is adjusted to:

$$\Delta w_{jk} = \eta (d_k - y_k) y_k (1 - y_k) h_j \quad (7)$$

η is the learning speed. w_{ij} , which is the connection weight between the input layer and the implied layer, is then determined by error function of the output layer.

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial f_2} \frac{\partial f_2}{\partial w_{ij}} = - (d_k - y_k) y_k (1 - y_k) \frac{\partial f_2}{\partial w_{ij}} \quad (8)$$

f_2 , which is the activation value between the implied layer and output layer, is determined by h_j and connection weight value w_{jk} . When taking derivative to h_j , we get

$$\frac{\partial f_2}{\partial w_{ij}} = \frac{\partial f_2}{\partial h_j} \frac{\partial h_j}{\partial w_{ij}} = w_{jk} h_j (1 - h_j) x_i \quad (9)$$

then,

$$\frac{\partial E}{\partial w_{ij}} = - (d_k - y_k) y_k (1 - y_k) w_{jk} h_j (1 - h_j) x_i \quad (10)$$

The connection weight value between the input layer and the implied layer w_{ij} is

$$\Delta w_{ij} = \eta (d_k - y_k) y_k (1 - y_k) w_{jk} h_j (1 - h_j) x_i \quad (11)$$

The establishment of the model generally requires the following steps: first, sample data collection and grouping. Second, the preprocessing of data, third, determination of the number of implied layers and their nodes, forth, the construction and training of neural networks, and last, prediction simulation test and result error analysis.

3.2 Data Collection and Grouping

The physical fitness test scores of students in grades 07 to 17 of the Chinese University are used as sample data to predict the physical fitness test scores of students in the next 10 years. The annual data is divided into two groups of men and women, in each. The average is used as the observation of each year. In the construction of neural networks, the number of neurons in the network input layer is set to 3. The number of neurons in the output layer is set to 1. It means three calendar years data to predict the next year's data, and so on. In order to reduce the impact of data range on the forecast results, all the data sub-categories obtained above were standardized, and they were normalized into the range of $[-1, +1]$. Taking the lung capacity of male as an example, the input matrix after normalization processing was as follows:

$$X = \begin{bmatrix} 1 & 0.8964 & 0.9309 & \dots & 0.7878 & 0.7582 & 0.7927 \\ 0.8964 & 0.9309 & 0.8964 & \dots & 0.7582 & 0.7927 & -0.9320 \\ 0.9309 & 0.8964 & 0.8618 & \dots & 0.7927 & -0.9320 & -0.9689 \end{bmatrix}$$

The output matrix is as follows:

$$Y = [0.8964 \quad 0.8618 \quad 0.7878 \quad \dots \quad -0.9320 \quad -0.9689 \quad -1]$$

The design of the number of hidden layer nodes in the network is more complex, and it has always been the difficulty of neural network construction. Although the improvement of network accuracy can be achieved by increasing the number of neurons, too many hidden layer neuron nodes will lead to a long model training time, and to reduce the prediction efficiency. Too few hidden layer neuron nodes will lead to poor network model fitting and the accuracy of prediction will be very low. Therefore, the number of implicit layer neuron nodes needs to be carefully selected. In this paper, the determination of the number of hidden layer neuron nodes is determined by trial method. When the number of implied layer nodes is 7, the model achieves the fastest convergence speed. In the neural network, the activation function of the implied layer is the scheming function of the S (Sigmoid) type, and the activation function of the output layer is the linear function. Referring to the process of the network algorithm, we can get a flow chart of the neural network:

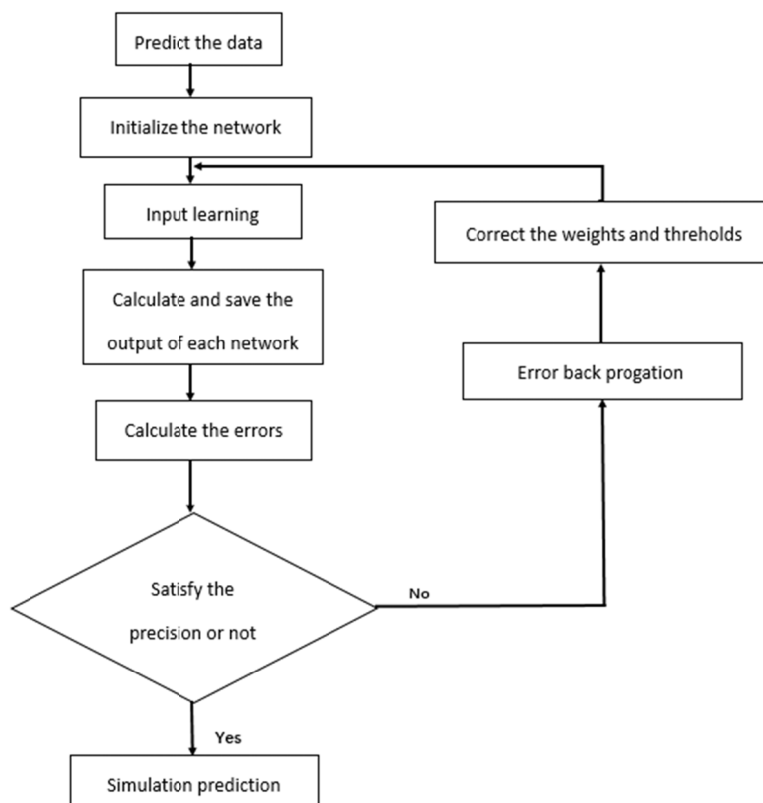


Figure 2. The flow chart of BP neural network

4. Simulation and Results Error Analysis

4.1 The Simulation Results of the Male Group

Figure 3 shows sports test scores of males for the next ten years predicted by the Error-reverse propagation neural network. The overall performance of physical tests for males showed an obvious downward trend, and the standing long jump performance showed a trend of decline for years. Lung capacity and pull-up scores also keep decreasing with fluctuations. This phenomenon indicates that college students' physical fitness and strength level are not ideal to a certain extent. Among them, lung capacity scores may decline to a record low of 3,658 by 2029, and pull-ups may decline to an all-time low of 2.7 in the same year. Although the 50-meter dash results show some volatility in the future, the overall trend shows an increase. It is forecasted that, in 2028, the results reached the worst 7.32 seconds, reflecting the continued decline of the explosive power of male. Long-distance running, measured by 1000 meters, fluctuates considerably over the next decade, but still shows a trend of increasing. Finally, there is an overall decline in the sit-and-reach performance over the next decade. All these results suggest that it is necessary to pay sustained attention to the physical fitness of college students.

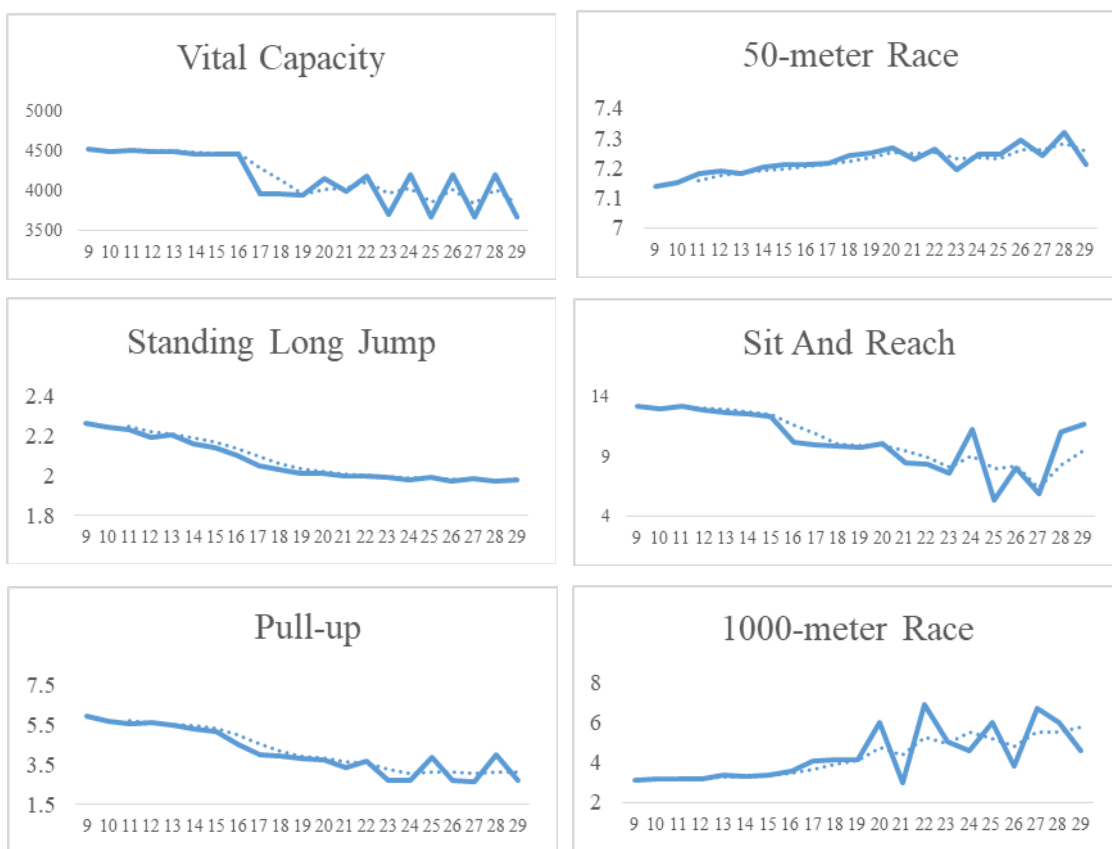


Figure 3. Prediction of physical fitness test results of male students in the next 10 years

4.2 The Simulation Results of the Girl Group

Figure 4 shows the physical examinations of females in the next ten years, which are predicted by the Error-reverse propagation neural network. Similar to the results of males, the physical fitness of females is unoptimistic as well. The forecasting results show that the lung capacity and standing long jump performance of females in the next ten years show a fluctuating trend and an improvement in performance in 2029. Lung capacity achievement will reach 3751.1, and the standing long jump performance will reach 1.6. Overall, the two measures were still lower than the average for grades 09 to 19. After a short period of decline, the sit-and-reach score will show a trend of continuing volatility over the next decade, with no significant decline. But it is worth noting that the sit-and-reach score is expected to hit a record low of 11.9 in 2029. The time of 50-meter race of females shows a trend of continuing increase, similar to the trend of males. The long-distance running, measured by 800 meters race, the results keep

the decreasing trend. Combined with the long-distance results of males, it indicates that the endurance level of college students keeps a declining trend in the long run.

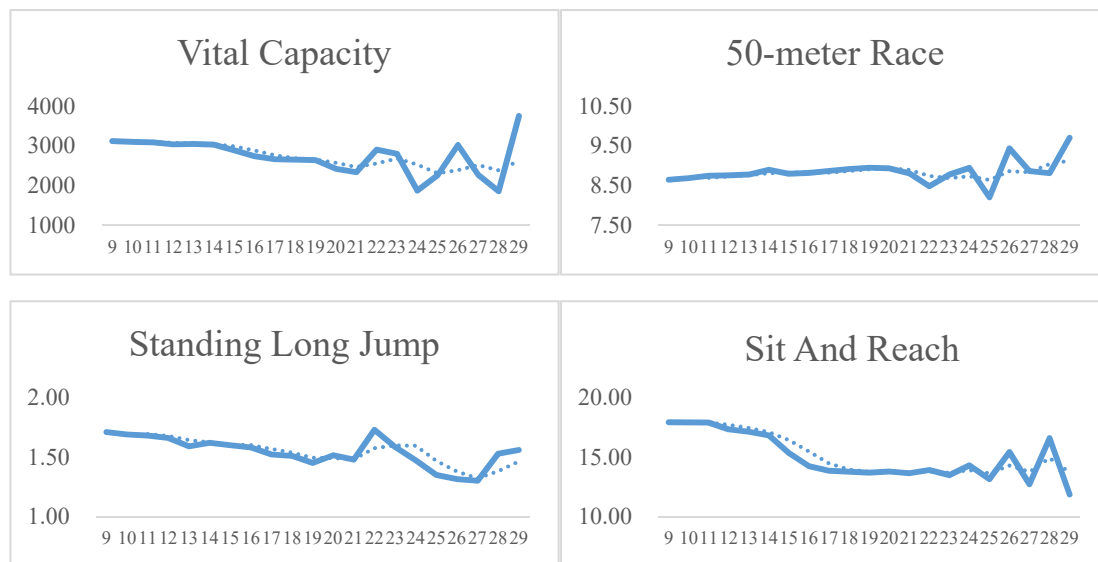


Figure 4. Prediction of physical fitness test results of female students in the next 10 years

Overall, the sports test scores of male and female students in the next ten years showed a decline in different degrees. Speed, explosive power, endurance level and strength continued to decline, which is consistent with Wang (2018). Due to the increased learning pressure and the acceleration of the pace of college life, it is difficult for college students to maintain regular life style. Coupled with the reduction of exercise, which ultimately leads to the decline of physical fitness, it is necessary to strengthen the health consciousness, and improve the physical fitness of college students. In order to enhance students' interest in sports, college students are encouraged to take physical exercise through different kinds of activities and competitions.

5. Conclusions and Recommendations

Based on the data of the physical health and sports scores of the university students, this paper uses the neural network with error reverse propagation, to predict the physical fitness of college students in the next ten years. we determined the number of neuron nodes in the implied layer by "testing method" and defined the activation function of the hidden layer of the neural network as a tangent function of S (Sigmoid), and the activation function of the output layer as a linear function. Based on the convergence speed, three calendar years data to predict the next year's data, the results show that in the next ten years, male and female students' physical test scores are expected to decline to varying degrees. The speed, explosiveness, endurance level and strength and other physical fitness indicators of college students continue to decline. It is necessary to strengthen the health consciousness, stimulate the college students' consciousness and enthusiasm for physical exercise, in order to improve the physical fitness of college students steady.

In order to improve the health level of all-people, while facing the current situation of the decline of college students' physical fitness, the state has made a series of measures to improve the physical fitness. Since 1990, the state has begun to implement the standard of physical exercise (later the national student physical health standards), which requires college students to actively participate in exercise. Inspired by this standard, many college students walked out of the dormitory and classrooms, and take part in some exercises. At present, under the campaign of healthy China, the state pays more attention to the role of physical education, with a view to improving the physical fitness of college students through physical education. China's college physical education mode also changes from test-oriented education to quality education, and becomes more scientific and reasonable. With the high enthusiasm of national fitness, campus football, campus basketball and other activities are in full swing. More students participate in football, basketball and other sports, which aims to enhance the physical fitness and encourage them to put enthusiasm into sports. In the background of The Asian Games, the Winter Olympics and other events, sports and exercises becoming the part of many people's life. Campus sports activities

will effectively promote the healthy exercise of college students, and improve their physical fitness. Although the physical fitness of college students shows a downward trend in recent years, there is every reason to believe that reversing the decline in the health of students is not an elusive goal. It is expected that, by 2022 and 2030, more than 50% students can meet the national physical health standards. The physical fitness of college students will be improved significantly in the future. And college students will gradually improve the enthusiasm to participate in physical exercise, through scientific and reasonable exercise programs. It will definitely help college students to achieve a steady improvement in physical fitness.

References

- Bandyopdhyay, B., & Chattopadhyay, H. (1981). Assessment of physical fitness of sedentary and physically active male college students by a modified Harvard step test. *Ergonomics*, 24(1), 15-20. <https://doi.org/10.1080/00140138108924826>
- Carter, C. M. (2012). *Academic and athletic motivation as predictors of academic performance of Division I college student-athletes*. The University of Oklahoma.
- Fricke, H., Lechner, M., & Steinmayr, A. (2018). The effects of incentives to exercise on student performance in college. *Economics of Education Review*, 66, 14-39. <https://doi.org/10.1016/j.econedurev.2018.06.009>
- Huang, X. M. (2005). The forecast model of college students' sports results. *Journal of Shangluo Teachers College*, 19(4), 52-53.
- Liu, H. (2014). Modelling and Predicting Sports Performances with Nonlinear Screening Factors and Weighting. *Computer Applications and Software*, 31(11), 105-108.
- Ministry of education of the people's Republic of China. (2010). *Ministry of education's announcement on the results of 2010 National Students' physique and health survey*. Retrieved from <http://old.moe.gov.cn/publicfiles/business/htmlfiles/moe/s5948/201109/124202.html>
- Niu, Z. (2020). On the promotion of College Students' physical health under the background of healthy China. *Think Tank Era*, 2020(5), 245-246.
- Tan, N., & Zhang, X. (2019). Research on the evolution and differences of students' physical health standards in China. *Sports World (Scholarly)*, 2019(11), 157-158.
- The State Council. (2014). *Opinions on accelerating the development of sports industry and promoting sports consumption*.
- Wang, J. (2017). Machine learning based prediction and analysis of college sports performance. *Modern Electronics Technique*, 40(17), 116-119.
- Wang, W. (2014). Analysis of the Content and Construction Path of Public Fitness Service System. *Journal of Harbin Institute of Physical Education*, 32(5), 48-51.
- Wang, X. (2018). *A Study on the Physical Quality of Chinese College Students Population*. Jilin University.
- Wang, X. (2019). Analysis on the influencing factors of the construction of campus basketball culture with middle school characteristics. *TIYU SHIYOU*, 42(6), 69-71.
- Wang, X., Sun, S., & Zhang, C. (2020). *Research on the Construction and Path of School Sports Cultural Objectives under the Promotion of Campus Football*.
- Zhang, C. (2017). Sports performance prediction model based on glowworm algorithm optimizing neural network. *Modern Electronics Technique*, 40(15), 94-100.
- Zhang, F., Jiang, Q., & Zhou, B. (2019, June). Prediction and Analysis of College Students' Sports Achievements Based on Support Vector Machine and Particle Swarm Optimization. In *Proceedings of the 5th International Conference on Frontiers of Educational Technologies* (pp. 128-132). <https://doi.org/10.1145/3338188.3338203>
- Zhang, L. (2002). A study on residents' living standard and college students' Physique in the past 20 years of reform and opening-up. *Journal of Beijing University of Physical Education*, 25(5), 600-602.
- Zhang, Q. (2018). Evaluation Model of College Students' Sports Learning Engagement Factors Based on Big Data Analysis. *Kuram ve Uygulamada Egitim Bilimleri*, 18(5), 2097-2107.
- Zhang, W., Mou, Y., Gao, Z., & Liu, Z. (2018). Research on the Performance Prediction of the Sport Competition Based on the Optimization Learning Machine. *Microcomputer Applications*, 34(3), 58-61.

- Zhang, Y. (2019). Prediction of college sports performance based on improved grey neural network. *Electronic Measurement Technology*, 11(42), 86-90.
- Zhang, Y., Yuan, X., & Gong, X. (2013) Research on Sports Achievements Predictive Based on the BP Nerve Network. *Bulletin of Science and Technology*, 29(6), 149-151.
- Zhao, B. (2001). Model System on Main Training Targets of Men's Jumping. *Journal of XI'An Institute of Physical Education*, 18(1), 81-82.
- Zhao, J., Jin, L. (2016). Discussion on the development trend of College Students' physical quality. *Heilongjiang Science and Technology Information*, 2016(35), 115-115.
- Zhu, H., Liu, J., Wang, G., & Li, M. (2003). Gray Forecast of GM(1,1) Model and Its Application in Forecast of Sports Performance. *Journal of Capital College of Physical Education*, 15(1), 118-121.
- Zhu, Y., & Ji, L. The Time Change and Implementation Path of National Fitness. *Journal of Harbin Institute of Physical Education*, 38(1), 33-37.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Canadian Center of Science and Education.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

Open access

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.