CONSTRUCTION OF AN EFFICIENT EVALUATION MODEL FOR ATHLETIC ATHLETES' COMPETITIVE ABILITY BASED ON DEEP NEURAL NETWORK ALGORITHM

Yuhan Niu*

School of Human Settlements and Civil Engineering, Xi’an Jiaotong University, Xi'an, Shaanxi, 710000, China - Xi'an Jiaotong University City College, Xi'an, Shaanxi, 710000, China

18602977530@163.com

Reception: 05/11/2022 **Acceptance**: 01/01/2023 **Publication**: 01/02/2023

Suggested citation:

N., Yuhan (2023). **Construction of an efficient evaluation model for athletic athletes' competitive ability based on deep neural network algorithm**. *3C Empresa. Investigación y pensamiento crítico, 12(1)*, 111-131. https://doi.org/10.17993/3cemp.2023.120151.111-131

https://doi.org/10.17993/3cemp.2023.120151.111-131

ABSTRACT

This paper analyzes the data of this year's athletes' physical fitness test scores and manages the classification of different physical qualities of the farmer. In order to reduce the manual calculation and increase the prediction efficiency, as well as to unify the scoring criteria of previous years, this paper proposes a comprehensive performance prediction model based on deep neural network algorithm. First, principal component analysis is used to transform multiple attributes with strong correlation into independent attributes that are not related to each other, and to reduce the time and space for model training by eliminating redundancy. Second, a back propagation (BP) neural network algorithm is used to build a physical fitness test prediction model, and the model is applied to the test dataset for model performance evaluation. Finally, the physical fitness test model was applied to other years for comprehensive performance prediction, and the differences between the model prediction results and the actual teachers' manual calculation results were observed. The results showed very good prediction results for 2021, in which 92.95% of the data had an absolute value of error less than 2 and only 0.06% had an absolute value of error greater than 4, which indicated that the prediction performance of the model was extremely significant. At the same time, a new athletic athletic scoring standard was also developed based on the neural network BP model to provide a more scientific theoretical basis and guidance for the evaluation of athletic ability of athletes.

KEYWORDS

BP neural network; athletic ability; track and field; prediction; evaluation

https://doi.org/10.17993/3cemp.2023.120151.111-131

PAPER INDEX

https://doi.org/10.17993/3cemp.2023.120151.111-131

1. INTRODUCTION

Athletics is a sport with a long history, with the largest number of gold medals in the Olympic Games and the most widely practiced sport in the world [1, 2]. Therefore, track and field is known as the mother of sports, and it plays a pivotal role in the development of each sport. Track and field is the cornerstone of the Olympic movement and best embodies the Olympic concept of "faster, higher, stronger". In the 2008 Beijing Olympic Games, China's athletes fought hard and achieved very impressive results, ranking first in the world with 51 gold medals, 21 silver medals and 28 bronze medals, achieving a historic breakthrough and reflecting the strong strength of a major sports nation. But at the same time, we should also clearly see that China's track and field projects, such as men's 110m hurdles, women's middle and long distance running, women's chain ball, etc. have been declared defeated in this Olympic Games, which is a big regret for Chinese sports in this Olympic Games.

Athletics is an important constraint to the progress of Chinese sports from a large sports country to a strong sports country [3]. For this reason, an ambitious goal of Chinese sports in the post-Olympic era is to accelerate the pace of revitalizing Chinese athletics and promoting the perfection of Chinese sports [4]. Athletics can be studied from several perspectives, such as the study of competition performance. Athletic performance is the result of athletes' participation in the competition, which is the substantial reflection of athletes' comprehensive ability, especially the reflection of athletic ability, and also the direct reflection of training results. The intuitive nature of athletic performance fully demonstrates the results of coaches in the whole sports training and competition activities, which is an objective criterion for scientific diagnosis of training and competition results, as well as a basis for the implementation of scientific training methods [4]. Therefore, the analysis of track and field results can provide coaches with an objective basis for training control, which is of great significance for the development of indoor track and field in China [5]. zheng et al. used gray theory and methods to analyze the results of Chinese and foreign outstanding decathletes by multi-level correlations, revealing the correlations between total performance and speed, jumping, throwing, and endurance categories, between each classification, and between total performance and each constituent item. Yu [6] et al. compared and analyzed the factor loading matrices of the top 150 athletes in the world in 2006, and concluded that the intrinsic factors limiting the development of the athletes' performance were the relatively weak basic quality of the athletes, especially the significant differences in the strength factor, and the overall low level. Chen [7] established a multiple regression prediction model for decathlon based on the data of the decathlon competition in track and field championships. The predicted values extrapolated from the dynamic trends of athletes' performance in each individual sport were highly correlated with the actual values with high accuracy and no variability.

Artificial neural networks, also known as neural networks, have their origin in neurobiology [8, 9]. Neural networks, which are composed of a large number of nerve cells, are a simplification, abstraction and simulation of the human brain [10]. Researchers have built artificial neural models by simulating the response processes of nerve cells [11]. A neural network is a parallel interconnected network composed of simple units that can simulate the interactive responses of the biological nervous system to the real world [12]. In recent years several studies have tried to apply neural networks to sports and athletic ability analysis, i.e., to improve the original algorithm combined with BP neural networks for modeling and combining it with practical applications, mainly for prediction or evaluation of some indicators [13]. As yan et al. provided a basis for the use of neural network modeling in biomechanics, opening a wide prospect of research in this area. They explored the problem of generalized inverse transformations of information in sports biomechanics by using neural network techniques to model the transformation of the characteristic quantities and the original information, taking shot put sports as an example.

The purpose of this paper is to trace the characteristics of various nonlinear functions by analyzing each physical test ability of track and field athletes and the all-around analysis standard published by IAAF, and then select the appropriate function for progressive scoring, so that the progressive method is not constrained by the average performance of athletes, i.e., it is universal. Based on the decathlon scoring scale developed by the progressive scoring method, the scoring criteria were revised based on the neural network method, and its good nonlinear mapping ability was used to try to make the scoring of each item more scientific and reasonable. Based on the successful establishment of the evaluation model of athletic ability of track and field athletes, it provides help for athletes to develop targeted training plans.

2. RESEARCH METHODS AND MODELS

2.1. KEY DATA PROCESSING THEORY

2.1.1. DATA STANDARDIZATION

Along with the progress of human society, the fields of study that humans are involved in have become more and more complex. In fact it has become very difficult to describe things in detail using individual attributes, and it is necessary to consider the problem from a holistic point of view, thus giving birth to the method of multi-indicator evaluation. In the system of multi-indicator evaluation, there are different units of attributes, and there are significant differences in magnitude and order of magnitude between them. Therefore, it is necessary to standardize each attribute of the original data, thus ensuring the reliability of the analysis results.

Currently, there are various methods of data standardization [14, 15], among which the most typical method is data normalization, which is to map the data to the [0, 1] interval uniformly. In this paper, the z-score standardization method is used to standardize the athlete physical measurement data to eliminate unit restrictions and dimensional relationships between variables and also to reduce the prediction time and increase the prediction accuracy [15, 16]. z-score standardization is based on the mean and standard deviation of the original data to standardize the data (Equation 1). This standardization method mainly converts the original data into standard normally

https://doi.org/10.17993/3cemp.2023.120151.111-131

distributed data with mean 0 and variance 1, and is suitable for scenarios with large data volumes.

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

where x_i is the original data, $\mathrm{E}(x)$ is the mean of the data, and s_i is the standard deviation of the data. The conversion of raw data into dimensionless evaluation index values by standardization makes the values of each index at the same level and facilitates the ability to sum weight attributes of different units or magnitudes.

2.1.2. DATA CORRELATION ANALYSIS

In today's era of big data, data correlation analysis can quickly and efficiently discover the intrinsic connections that exist between different things [17, 18]. Correlation, which refers to the pattern that exists between two or more variables in some sense, aims to explore the intrinsic information hidden in the data set. The correlation coefficient can be used to describe the relationship between variables, where the sign of the correlation coefficient indicates whether the direction of the relationship is positive or negative, and the magnitude of its value represents the strength of the relationship between the two variables, where the correlation coefficient is 0 when there is no correlation at all and 1 when there is a perfect correlation. There are various methods for calculating correlation coefficients in correlation analysis, including Pearson correlation coefficient [19], Spearman correlation coefficient [20], partial correlation coefficient [21], Kendall correlation coefficient [22], and so on. In this paper, the Pearson correlation coefficient is used to calculate the magnitude of correlation between the attributes. Standardization of the original data does not change the correlation between the attributes of the original data. Therefore, the standardized data can be used to directly calculate the correlation between the attributes. The correlation coefficient is calculated as shown in Equation 2:

where xi and yi are the data of the two attributes involved in the calculation, and $\bar{x}$ and $\bar{y}$ are the mean values of the corresponding attributes.

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

2.1.3. PRINCIPAL COMPONENT ANALYSIS ALGORITHM

Principal component analysis is a commonly used technique for reducing the dimensionality of a data set, allowing exploration and simplification of certain complex

https://doi.org/10.17993/3cemp.2023.120151.111-131

relationships between variables. It can transform several original variables with strong correlation characteristics into several uncorrelated variables through coordinate transformation, and the uncorrelated ones calculated as principal components. The principal component is a linear combination of the original variables, and its model is shown in Figure 1.

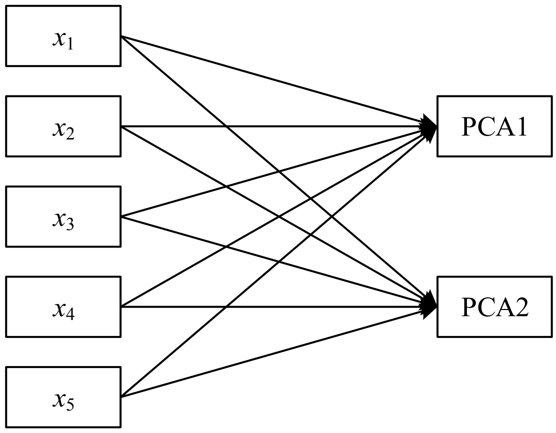

Figure 1 Principal component analysis (PCA) model

The principal components reflect most of the characteristics of the primitive variables and can remove strong correlations between the primitive variables. The principal components are linear combinations of the original variables. The first principal component explains the most variance of the primitive variables, while the second principal component explains the second variance of the original variables and is orthogonal to the first principal component, which means it is completely uncorrelated. By analogy, the remaining principal components are all orthogonal to each other. Suppose, there exist p variables in the original data, namely x_1, x_2,…,, which are linearly combined to form a new p mutually independent principal component variable y_p, whose mathematical model expression is shown in Equation (3).

$$\begin{cases} y_1 = \mu_{11}x_1 + \mu_{12}x_2 + \cdots + \mu_{1i}x_i + \cdots \mu_{1p}x_p \\ y_2 = \mu_{21}x_1 + \mu_{22}x_2 + \cdots + \mu_{2i}x_i + \cdots \mu_{2p}x_p \\ \qquad \vdots \\ y_i = \mu_{i1}x_1 + \mu_{i2}x_2 + \cdots + \mu_{ii}x_i + \cdots \mu_{ip}x_p \\ \qquad \vdots \\ y_p = \mu_{p1}x_1 + \mu_{p2}x_2 + \cdots + \mu_{pi}x_i + \cdots \mu_{pp}x_p \end{cases} \tag{3}$$

The model is changed into matrix form as shown in Equation (4):

$$y = \begin{bmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{1p} \\ \mu_{21} & \mu_{22} & \cdots & \mu_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{p1} & \mu_{p2} & \cdots & \mu_{pp} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} = Ux \tag{4}$$

https://doi.org/10.17993/3cemp.2023.120151.111-131

where the sum of squares of the principal component coefficients μ_{pp} is 1, as shown in Equation (5):

$$\mu_{i1}^{2} + \mu_{i2}^{2} + \boxed{?} + \mu_{ip}^{2} = 1 \tag{5}$$

Next, to obtain the principal component values y_p, the principal component coefficients μ_{pp} are to be calculated, and first the covariance matrix is calculated from the original data as shown in Equation (6):

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

Since the data have been standardized, then the variance s^2 of the original data is 1 see equation (7), the

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

Changing equation (7) to equation (8), the

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

At this point, we can obtain equation (9)

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

From Equation (9), the correlation coefficient matrix of the original data is actually equivalent to the covariance matrix. The eigenvalue λ_i of the covariance matrix represents the variance of the principal components, while the eigenvalue of the covariance matrix with the principal component coefficients μ_{pp} is calculated from the correlation coefficient matrix and the eigenvalue y is calculated by the formula $y = x\mu^T$. Theoretically, a smaller number of principal components is selected to replace the original full data based on the contribution of principal components, and the number of principal components with a contribution of 90% is generally selected, but in this paper, the same number of principal components as the number of original variables will be selected, which does not throw away the information of the original data and removes the influence of strong correlation between the original variables on the model training. Next, the neural network model is built using eight new variables that are not correlated with each other, which increases the persuasiveness of the model accuracy and excludes the influence of the data itself factors on the model building and parameter optimization.

2.2. BP NEURAL NETWORK MODEL OPTIMIZATION

 https://doi.org/10.17993/3cemp.2023.120151.111-131

BP neural network as a part of neural network is a supervised learning algorithm [23]. It is a multilayer nonlinear feedforward network trained by an error back propagation learning algorithm. The network consists of an input layer, an implicit layer and an output layer. The BP learning algorithm consists of two processes, a forward propagation of the data and a backward propagation of the error signal.

2.2.1. FORWARD PROPAGATION OF DATA

The interconnection pattern between neural networks is formed by interconnecting neurons, and the initial weights between each connection are randomly assigned by the computer. The forward propagation phase of the data signal is the process of the original data signal from the input layer through the implicit layer to the output layer, i.e., the output of the upper layer nodes is used as the input of the lower layer nodes. The basic structure of the forward propagation of the BP neural network is shown in Figure 2.

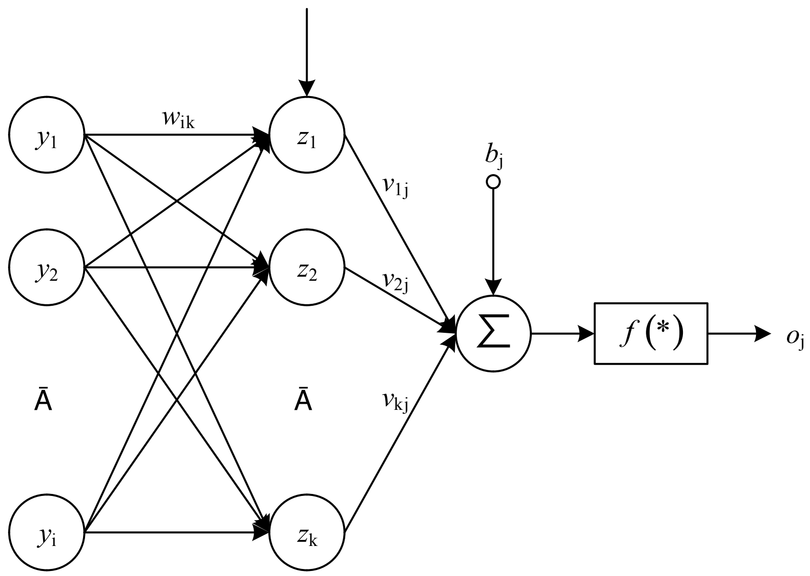

Figure 2 Basic structure of forward propagation stage BP neural network

Figure 2 shows that each neuron cell y_i has a corresponding computational weight w_{ik}. The output value of the input layer in the hidden layer $net1_k$ is obtained by summing and weighting the input values, connection weights and threshold b_k, calculated as Equation (10) shows:

$$net1_k = \sum_{i=1}^{n} w_{ik} y_i + b_k = w_{1k} y_1 + w_{2k} y_2 + \boxed{?} w_{nk} y_n + b_k \qquad (10)$$

In the process of prediction, better prediction accuracy can be obtained by applying activation function processing. There are many kinds of activation functions, such as step function, Sigmoid function, tanh function, and ReLU function. In this project, the Sigmoid function is used to activate the output information. The output value of the input layer is activated by the activation function as ($net1_k$), then the implicit layer zk is obtained by Equation (11).

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

Next, the implicit layer data is passed to the output layer as an input layer. The output value $net2_j$ is obtained from the weighted sum of the implied layer value and the connection weight υ_{kj} between the implied layer and the output layer, plus the threshold b_j as shown in Equation (12):

$$net2_j = \sum_{k=1}^{n} v_{kj} z_k + b_j = v_{1j} z_1 + v_{2j} z_2 + \boxed{?} v_{nj} y_n + b_j \tag{12}$$

Equation (13) activates the output values to obtain the data of the final output layer.

$$o_j = f\left(\text{net}\,2_j\right) = \frac{1}{1+\exp\left(-net2_j\right)} \tag{13}$$

2.2.2. ERROR BACK PROPAGATION

The error function is used to detect whether the training process of the neural network is finished when the signal is passed to the output layer. The condition for the neural network to stop is that the error function limit is satisfied or a set maximum number of iterations is reached. When the output error function is less than the predefined value, the training will stop. If the condition is not satisfied, the error will be back-propagated. The error function (E) is used to measure the magnitude of the error between the actual output d_j and the desired output o_j, which is calculated as shown in Equation (14).

$$\text{E} = \frac{1}{2}(d-o)^2 = \frac{1}{2}\sum_{j=1}^{n}\left(d_j - o_j\right)^2 \tag{14}$$

The error signal obtained at each layer is used to adjust the weights of the connections between neurons. Equation (3.15) simulates the process of error back propagation.

$$\text{E} = \frac{1}{2}(d-o)^2 = \frac{1}{2}\sum_{j=1}^{n}\left(d_j - o_j\right)^2 \tag{15}$$

The error is reduced along the gradient direction by continuously adjusting the connection weights and thresholds. After calculating the change value $\Delta\upsilon_{jk}$ of the weight connection value between the implicit layer and the output layer, each connection weight is updated as shown in Equation (16) and Equation (17).

$$\Delta v_{kj} = -\eta \frac{\partial E}{\partial v_{kj}} = -\eta \frac{\partial E}{\partial net2_j} \frac{\partial net2_j}{\partial v_{kj}} = -\eta \frac{\partial E}{\partial net2_j} z_k \tag{16}$$

https://doi.org/10.17993/3cemp.2023.120151.111-131

$$v_{kj} = v_{kj} + \Delta v_{kj} \tag{17}$$

Further extending the error to the input layer as in equation (18)

$$E = \frac{1}{2}\sum_{j=1}^{n}\left(d_j - f\left(\sum_{k=1}^{n} v_{kj} z_k\right)\right)^2 = \frac{1}{2}\sum_{j=1}^{n}\left(d_j - f\left(\sum_{k=1}^{n} v_{kj} f\left(\sum_{i=1}^{n} w_{ik} y_i\right)\right)\right)^2 \tag{18}$$

The connection weights between the input layer and the implied layer are updated as shown in Eqs. (19) and (20).

$$\Delta w_{ik} = -\eta \frac{\partial E}{\partial w_{ik}} = -\eta \frac{\partial E}{\partial net1_k} \frac{\partial net1_k}{\partial w_{ik}} = -\eta \frac{\partial E}{\partial net1_k} y_i \tag{19}$$

$$w_{ik} = w_{ik} + \Delta w_{ik} \tag{20}$$

After all the weights are readjusted, signal forward propagation will continue to be executed. When the model reaches the convergence criterion, the training is stopped, the model is built, and the model parameters are adjusted to optimize the model. The established model is used to predict the physical fitness test data, and the error magnitude between the predicted and actual values is calculated to verify the feasibility of the model, and then the model is applied.

3. RESULTS AND DISCUSSION

3.1. VALIDATION OF THE PHYSICAL TEST SCORE PREDICTION MODEL

The physical fitness test data were preprocessed and standardized, and principal component analysis was used to eliminate strong correlations between the data. The model was built using BP neural network after transforming the original data using principal component analysis method. In this project, the physical fitness test data of a provincial track and field team in 2019 were selected to build the model, in which 80% of the athlete samples were used as the training set and the remaining 20% were used as the test set to evaluate the model. The scoring criteria and methods for male and female distance athletes are different, so the male and female distance athletes test data were separated and separate models were built for prediction. The model was continuously adjusted and optimized using 80% of the training set, and the final model parameters are shown in Table 1.

https://doi.org/10.17993/3cemp.2023.120151.111-131

Table 1 The parameters of model

Parameter Name	Parameter Value
Number of neurons in the input layer	8
Number of neurons in the output layer	1
Number of neurons in the hidden layer	11
Threshold	0.005
Learning rate	0.1
Maximum number of iterations	1.0e10
Training algorithm	rprop+
Error function	sse
Activation function	logistic

As shown in Table 1, the threshold value is used as a conditional value for training stop, which specifies the predetermined value in the error function. The maximum number of iterations forces the training to stop when the predetermined value is not always reached and the iteration cannot be stopped. The algorithm used for training, "rprop+", is the error back propagation algorithm with weights, also known as BP neural network algorithm [24]. The error function "sse" is used to calculate the magnitude of the error at the end of the forward propagation. The activation function uses the parameter "logistic" for the Sigmoid activation function. Where the number of neurons in the hidden layer is determined by the mean square error (MSE) with equation (21).

$$h = \sqrt{m+n} + \alpha \tag{21}$$

Where m in equation (21) is the number of input neurons, n is the number of output neurons, and α is a constant in the range 1 to 10, so the number of hidden layers h takes values in the range 4 to 13.

In order to increase the accuracy of the model, prediction models for male and female students were built separately, and the accuracy of the model was evaluated using the data from the test set. The test set data that were not involved in the model training process were brought into the model for prediction separately, and the prediction results for both boys and girls were finally combined to observe the overall model prediction performance. A sample of 40 athletes was randomly selected from the test set of 2019 data, and the difference between the actual values of these 40 athlete samples and the predicted values of the model was compared as shown in Figure 3. The line graph shown in Figure 3 shows the comparison between the predicted and actual values of the randomly selected sample of 40 athletes in the test set. The solid line is the actual value and the dashed line is the predicted value. It is obvious that the two lines have a high overlap rate and only a few samples can be

https://doi.org/10.17993/3cemp.2023.120151.111-131

viewed with significant errors. The results show that the prediction model established in this project is very accurate and has good performance, and the value of mean square error MSE of the model is 1.361713.

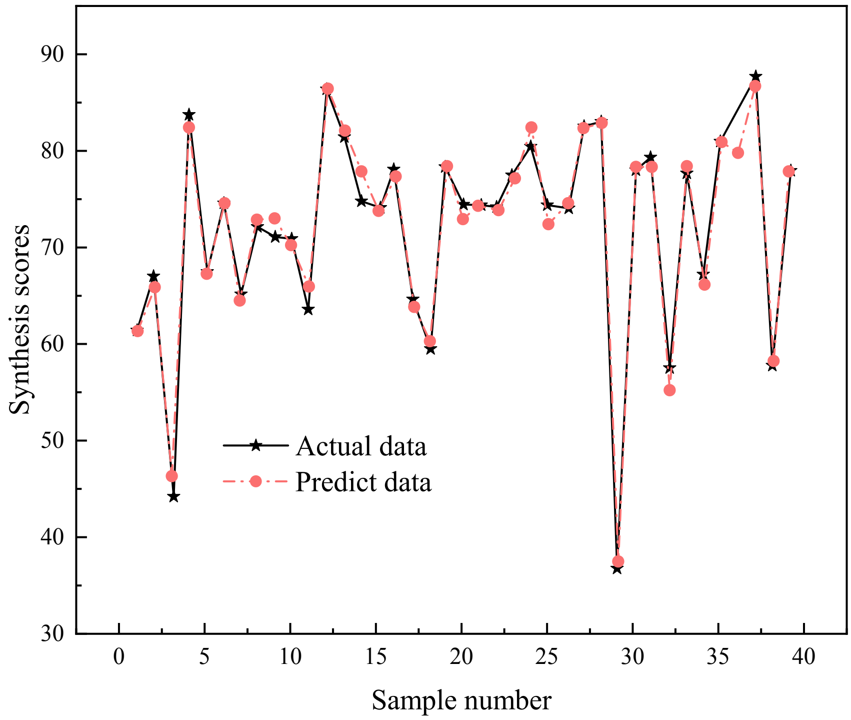

Figure 3 Comparison of predict data with actual data samples

3.2. APPLICATION OF PHYSICAL TEST PERFORMANCE PREDICTION MODEL

After the performance evaluation of the model, the next step is to observe the practical application of the model. In this paper, we chose to apply the model developed from the 2019 athletes to the 2021 athletes to predict the overall performance of the 2019 students, to better observe the changes in the overall performance of the 2019 athletes under a uniform grading scale, and to observe the actual application of the 2019 model in another year. The 2021 athletes' test scores were pre-processed and standardized, and the strong correlation was reduced using principal component analysis. The data of male and female students were separated and brought into the model of male and female students built from the 2019 data separately to obtain the model prediction results. The absolute values of the absolute errors between the 20% test set of 2019 and the predicted results of 2021 data were boxed into six intervals, [0,1), [1,2), [2,3), [3,4), [4,5), [5,∞), and their distributions are shown in Table 2.

https://doi.org/10.17993/3cemp.2023.120151.111-131

Table 2 Percentage of absolute error distribution predicted by the model established 2019 data

	[0,1)	[1,2)	[2,3)	[3,4)	[4,5)	[5,∞)
2016	63.38	27.57	5.27	1.7	0.04	0.02
2019	39.84	28.16	15.71	8.06	4.06	4.17

As shown in Table 2, the prediction results for 2021 are very good, in which 92.95% of the data have an absolute error value less than 2, and only 0.06% have an absolute error value greater than 4. This indicates that the prediction performance of the model is very high. When the model built from the 2021 scoring criteria is applied to 2021, the accuracy of the model prediction decreases, and the percentage of absolute error distribution in the range of 0 to 1 decreases by 23.54%, but the overall prediction result is still considerable. If the scoring criteria of previous years are strictly adhered to, the model criteria of 2019 must be fully applicable to 2021, and the accuracy of the model prediction will be very close to the prediction results of the 2019 test set, while the significant decrease in prediction results indicates that there is a difference between the scoring criteria of 2019 and 2021 due to human calculation.

This model was built based on data from 2019, and the weights between models used the 2019 scoring criteria. When the 2019 scoring scale was used to predict the 2021 composite score, the results showed differences in the scoring scale. Apparently, the scoring criteria for the composite scores of the physical fitness tests were not uniform due to the manual involvement of the coaches in the calculations. The different scoring standards in previous years resulted in the composite score of the physical fitness test not accurately reflecting the changes in the basic physical fitness of the farrier, and it was difficult to infer whether the physical fitness of the farrier had improved or decreased over time, which obviously did not maximize the value of the physical fitness test data from previous years. Using the model to predict the composite score, the constant relationship between the measured item data and the composite score is explored through the learning ability of the model, which can well avoid the whole complicated process of traditional physical fitness test calculation while saving the time of composite score calculation. It is necessary to use the model for predicting the physical fitness test scores because it is possible to see the changes in the physical fitness of the farriers more clearly by classifying them according to the composite score levels after the uniform scoring criteria and combining the radar chart visualization data information.

3.3. ANALYSIS OF TRACK AND FIELD SCORING CORRECTION SCHEME

This section first analyzes the problems of the current decathlon scoring scale published by the IAAF, then proposes our basic correction scheme and develops a new scoring scale using the progressive scoring method and the neural network model. Finally, the existing scoring scale, the scoring scale developed by the

https://doi.org/10.17993/3cemp.2023.120151.111-131

progressive scoring method, and the scoring scale developed based on the neural network model are compared and analyzed, and finally a more scientific and reasonable set of scoring scale is screened and determined.

3.3.1. SCORING REASONABLENESS ANALYSIS

As the level of all-around sports continues to improve, the more difficult it is to improve, the more time and energy athletes pay, in order to motivate athletes and promote the development of all-around sports performance to a higher level, so it should be reflected and rewarded through the incremental value. That is, the all-around performance of each single item from low to high points of the corresponding coordinate points of the line, should be a positive parabolic type with the sports performance and the score increases rapidly. At present, some of the projects with the improvement of the level of sports, but the value of the corresponding linear growth relationship, and even some of the project sports performance incremental rate is gradually reduced.

Scoring from a low score loses its meaning in all-around sports competitions. Because its score is too low and rarely used, then the score of that part of the performance becomes unused and makes the score sheet too large. For example, men's decathlon such as 100m starting point is 17.83s, shot put 1.53m, high jump 0.77m, pole vault 1.03m. Such results can be achieved for children and teenagers with a little training, not to mention for adult professional athletes. If the scoring table is also applicable to children and teenagers, but the IAAF decathlon is quite difficult, and most of its individual events are beyond the physical ability of children and teenagers due to competition conditions and equipment specifications, so a separate scoring table should be developed to meet the needs of children and teenagers. Thus, it seems that the IAAF scale has no meaningful use for children and teenagers. Secondly, the scoring table has certain restrictions, and with the improvement of the level of all-around sports, the scoring table must be constantly widened, but the wide application of the object extends the minimum score of the scoring table, the result is that the original scoring table is too wide to become even wider, making the scoring table more massive, bringing difficulties to the reasonable preparation of the scoring table, bringing trouble to the evaluation of the scoring table.

Some of the projects due to technical innovation, improvement of equipment, the use of more scientific training methods and means, so that the level of sports to improve quickly, while some of the projects training science and technology and means development is slow, so that the difficulty of the individual scoring is not balanced. In the decathlon, a more reasonable score should be equal or approximate scores for different items of equal difficulty, so as to avoid some athletes to get good results, too biased towards an easy to score items, thus not really promote the development of the decathlon in the direction of the whole. For example, some subjects can be achieved with a little training in a certain scoring section, while some items require some effort.

https://doi.org/10.17993/3cemp.2023.120151.111-131

3.3.2. SELECTION OF SCORING INCREMENTAL FUNCTIONS

According to the progressive idea, with the improvement of sports performance, the corresponding score should be increased sharply to encourage the improvement of the sports level of excellent members. That is, sports performance at low levels and high levels to improve the same interval, the increase in score should not be the same, high levels because of the difficulty, the score increased more, the score should be progressive trend with the improvement of performance. Next, we try to describe this phenomenon by using mathematical functions to find out the correspondence between grades and scores. In higher mathematics, monotonicity is actually the portrayal of some functions when the self-varying process is the change of the dependent variable from small to large changes in a given range when the change in the dependent variable presents a special law. There are generally the following four cases: ① from small to large ② from large to small ③ suddenly large and small ④ unchanged. Which has ①, ② these two special laws of the function is an important type of function - monotonic function. Therefore, the monotonicity of a function is a concept used to describe the tendency of a function to change in a certain range. By studying the monotonicity of the function can be reduced to the study of complex functions to some of the more typical, simple types of functions. This phenomenon can be described as monotonically increasing. A monotonically increasing mathematical function can be used to implement the process of increasing.

Definition of increasing monotonicity: Let there be a function y=f(x), x is called the independent variable and y is the dependent variable. If for any x1, x2 ∊ [a, b] contained in M, if when x1 < x2, there is y1 < y2, that is, f(x1) < f(x2), y increases as x increases, the image of the function rises from left to right, then f(x) is said to be increasing on [a, b], said y is the increasing function on the interval, [ab] is called the monotonic increasing interval of y = f(x). In the interval greater than 0, incremental functions in the primary functions are power functions (such as y = xn), exponential functions (such as y = ex), logarithmic functions (such as y = lnx), tangent functions (such as y = tanx). The four types of functions are shown in Figure 4. From the figure, we can see that as x increases, except for the tangent function, which increases periodically, the y values of the other three types of functions increase to varying degrees, with the exponential function growing the fastest, followed by the power function, and the slowest being the logarithmic function. From Figure 4, we can see that the tangent function periodically increasing, and even appear to calculate the y-value appears negative, so it is not suitable for the development of the score table. We chose the power function, exponential function, and logarithmic function to participate in the initial fitting of the scoring table.

https://doi.org/10.17993/3cemp.2023.120151.111-131

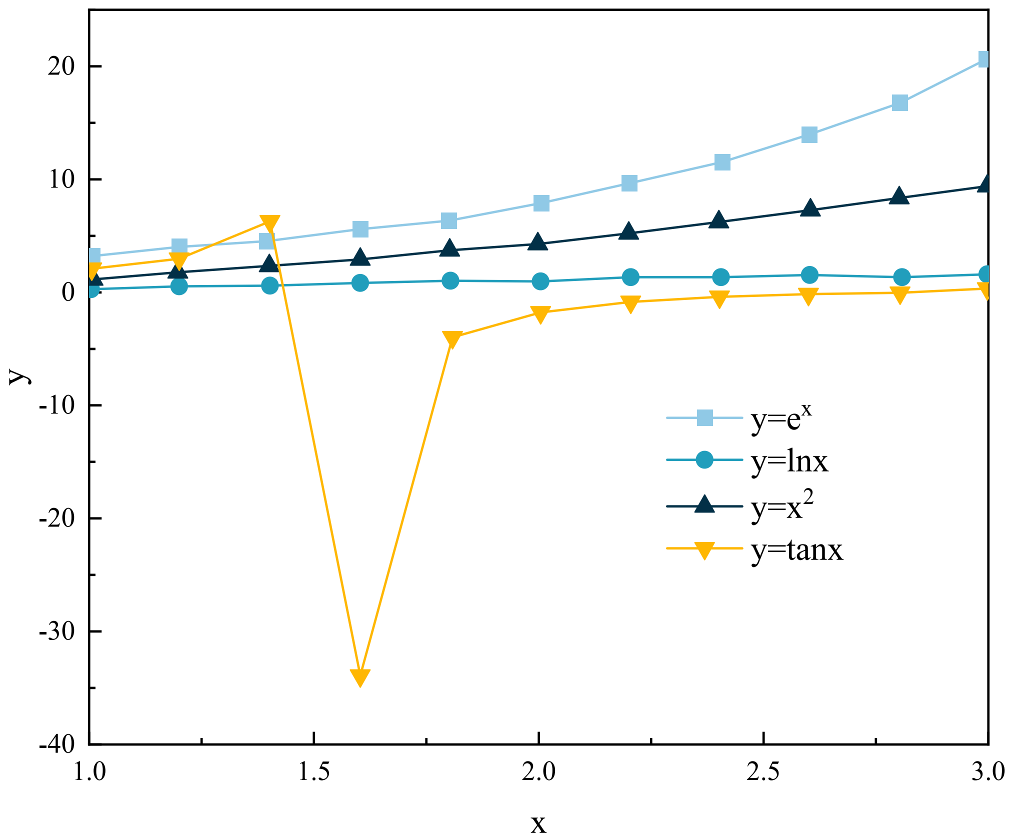

Figure 4 Incremental function diagram

3.4. NEURAL NETWORK DEVELOPMENT OF THE ALLOMETRIC RATING SCALE

Artificial neuronal network is a theorized mathematical model of the neural network of the human brain, an information processing system based on imitating the structure and function of the neural network of the brain [25, 26]. It is an artificially constructed neural network capable of achieving certain functions based on the existing human understanding of the neural network of the brain, which absorbs many advantages of biological neural networks and thus has its special characteristics: (1) highly parallel computing and distributed storage functions: artificial neural networks are composed of many identical basic processing units grouped in parallel, and although the function of each unit is simple, both each Although the function of each unit is simple, both the small unit and the whole neural network have the dual capability of processing and storing information, and these two functions are naturally integrated in the same network, which makes its processing capability and effect on information amazing [27, 28]. (2) Highly nonlinear global action: An artificial neural network is a large-scale nonlinear dynamical system in which each neuron can receive inputs from a large number of other neurons and produce outputs that affect other neurons through parallel networks. It has a strong nonlinear processing capability. Globally, the overall performance of the network is not a simple superposition of local performance, but some kind of collective behavior that exhibits the characteristics of complex nonlinear dynamical systems in general [29]. (3) Good fault tolerance and associative memory function: Artificial neural networks can store information in the weights between

https://doi.org/10.17993/3cemp.2023.120151.111-131

neurons, and through their own network structure can achieve memory of information. This storage is distributed and the extraction of information is collaborative as a whole, each information processing unit contains both a contribution to the collective and cannot determine the overall state of the network, so a failure of the local network does not affect the correctness of the overall network output, which makes the network fault-tolerant [30].

From Figure 4, we can see that the exponential function and the power function are more in line with our design purpose. Therefore, We take the starting and ending points of the new scale initially formulated in the section, remove the ultra-low achievement segments below the starting point, and remove the ultra-high score segments above the stopping point, and use the exponential function and power function to fit the simplified scale in origin9.0, and take the best fitting effect among the many fitting results [31-32]. Then use the function with the best fit, still using the interval of each score segment in the original score table, and bring the scores of the original score table into the fitting function to find the corresponding score, which is the result we get in this section using the progressive idea. The results of the optimal exponential and power function fitting for each item are shown in Tables 3 and 4.

Table 3 Best-fit exponential function

Project	Best-fit function
100m	$y=1+27583.7556*e^{(-0.3196*x)}$
Long Jump	$y=1+71.8068*e^{(0.3345*x)}$
Shot Put	$y=1+149.3857*e^{(0.1071*x)}$
High Jump	$y=1+43.2982*e^{(1.2941*x)}$
400 m	$y=1+17297*e^{(-0.0637*x)}$
110m hurdles	$y=1+10352.2147*e^{(-0.1258*x)}$
Discus	$y=175+27.1458*e^{(-0.0566*x)}$
Pole vault	$y=1+68.0512*e^{(0.4288*x)}$
Javelin	$y=135+38.4222*e^{(0.0528*x)}$
1500m	$y=1+13345.0214*e^{(-0.0208*x)}$

https://doi.org/10.17993/3cemp.2023.120151.111-131

Table 4 Best-fit power functions

Project	Best-fit function
100m	y= 8359871.4050* $x^{-3.7534}$
Long Jump	y= 10.8651* $x^{2.2154}$
Shot Put	y= 26.9875* $x^{1.2345}$
High Jump	y=128.5425* $x^{2.3545}$
400 m	y= 225487521* $x^{-3.3125}$
110m hurdles	y= 14715841* $x^{-2.8412}$
Discus	y= 6.2514* $x^{1.2587}$
Pole vault	y= 40.2512* $x^{1.8998}$
Javelin	y= 3.5413* $x^{1.3087}$
1500m	y= 38965007571* $x^{-3.1875}$

4. CONCLUSION

In this paper, we propose a model for predicting the overall performance of physical fitness test, and successfully apply the machine learning algorithms such as BP neural network and principal component analysis to predict the overall performance of physical fitness test. The following conclusions were obtained:

(1) The use of the prediction model for predicting the composite scores reduces the calculation time of the scores and solves the problem of inconsistent scoring standards due to manual calculation in previous years. The results show that the prediction results for 2021 are very good, with 92.95% of the data having an absolute value of error less than 2 and only 0.06% having an absolute value of error greater than 4, which indicates that the prediction performance of the model is very high.

(2) In addition, this paper also explores the rationality of the athletics scoring method, and by training the existing samples and building a BP neural network model to obtain the expected output, the nonlinear relationship between the score and the athletic performance is better resolved.

(3) By comparing the IAAF scoring scale, the exponential fitting scoring scale, and the scoring scale based on the neural network model, it is shown that the neural network scoring method is better than the former two in terms of scoring progressivity and item balance.

REFERENCES

(1) H Li, **Research on the Status Quo and Countermeasures of Scientific Training of Track and Field Sports**, (2020).

(2) H Cheng, J Lv, T Yuan, **Forecasting Chinese Athletics Performance in Tokyo Olympics from the World Top 20 Athletics Countries in 2018**, *Bulletin of Sport Science & Technology* (2020).

(3) J G Zhan, **Research on Measures Taken for Development of Track & Field Socialization**, *Marketing and Professionalization in China, Journal of Beijing University of Physical Education* (2002).

(4) D Chen, **Cultivation of Athletes' Non-intelligence Factors in Track and Field Training Based on Sports Training Management**, 2019.

(5) Y Ren, J Li, **The Conception of Application of Computer Virtual Reality Technology in Sports Training**, *Journal of Physics: Conference Series* 1861(1) (2021) 012110-.

(6) L Yu, Y He, X Kong, et al., **Projected estimation for large-dimensional matrix factor models, Journal of Econometrics** 229 (2022).

(7) X Chen, R Tu, M Li, et al., **Prediction models of air outlet states of desiccant wheels using multiple regression and artificial neural network methods based on criterion numbers**, *Applied Thermal Engineering* 204 (2022) 117940-.

(8) A A April, J K Basu, D Bhattacharyya, et al., **Use of Artificial Neural Network in Pattern Recognition**, *Springer Berlin Heidelberg* (2020).

(9) S Knight, N Gadda, **Spatiotemporal Patterns in Neurobiology: An Overview for Future Artificial Intelligence** (2022).

(10) H Bagherinezhad, A Farhadi, M Rastegari, **Lookup-based convolutional neural network**, *IEEE*, 2020.

(11) G Jin, M Wang, J Zhang, et al., **STGNN-TTE: Travel time estimation via spatial–temporal graph neural network**, *Future Generation Computer Systems* 126 (2022) 70-81.

(12) I Danihelka, N E Kalchbrenner, G D Wayne, et al., **Associative long short-term memory neural network layers**, 2021.

(13) Z Xing, P Hamilton, A K Story, **Reduction of computation complexity of neural network sensitivity analysis**, 2021.

(14) L Ogiela, M R Ogiela, M Takizawa, **Safety and Standardization of Data Sharing Techniques and Protocols for Management of Strategic Data**, *2017 IEEE 31st International Conference on Advanced Information Networking and Applications (AINA)*, 2017.

(15) X Yang, J Bai, W Sun, **Research on AutoCAD Data Standardization in Coordinate Transformation**, *Urban Geotechnical Investigation & Surveying* (2019).

(16) P Chandio, A B Talpur, **An Assessment of Financial Stability of Textile Sector of Pakistan: An Altman Z Score Approach**, *Sustainable Business and Society in Emerging Economies* 3 (2021).

(17) [P J Schreier, L L Scharf, **Statistical Signal Processing of Complex-Valued Data: Correlation analysis** (2010).

(18) Y Kan, **A Cloud Computing Resource Optimal Allocation Scheme Based on Data Correlation Analysis** (2021).

(19) L Egghe, L Leydesdorff, **The relation between Pearson's correlation coefficient r and Salton's cosine measure**, *Journal of the American Society for Information Science & Technology* 60(5) (2014) 1027-1036.

(20) Y Han, A D Hutson, **A Robust Spearman Correlation Coefficient Permutation Test** (2020).

(21) Y Yan, C Jiang, **Construction of Network Model Based on Partial Correlation Coefficient**, *Journal of Longyan University* (2019).

(22) J Zhou, X Su, H Qian, **Research on Fusion of Dependent Evidence Based on Kendall Correlation Coefficient**, *CSAE 2020: The 4th International Conference on Computer Science and Application Engineering,* 2020.

(23) Y Li, D Zhao, S Yan, **Research on Travel Agency Human Resource Crisis Early Warning Model based on BP Neural Network and Computer Software**, *Journal of Physics Conference Series* 1744(4) (2021) 042061.

(24) S Horikawa, T Furuhashi, Y Uchikawa, **On fuzzy modeling using fuzzy neural networks with the back-propagation algorithm**, *IEEE Transactions on Neural Networks* (1992).

(25) A Lebedeva, A Mironov, A Pisarchik, et al., **Integration technology for replacing damaged brain areas with artificial neuronal networks**, *IEEE 4th Scientific School on Dynamics of Complex Networks and their Application in Intellectual Robotics (DCNAIR)*, 2020.

(26) Y Gao, G A Ascoli, L Zhao, **BEAN: Interpretable and Efficient Learning With Biologically-Enhanced Artificial Neuronal Assembly Regularization**, *Frontiers in Neurorobotics* 15 (2021).

(27) P Meier, K Rohrmann, M Sandner, et al., **Compensation of Measurement Errors for a Magnetoresistive Angular Sensor Array using Artificial Neuronal Networks**, *IEEE Access* (99) (2020) 1-1.

(28) X Yang, T Levi, T Kohno, **Digital Hardware Spiking Neuronal Network with STDP for Real-time Pattern Recognition**, *Journal of Robotics Networking and Artificial Life* 7(2) (2020).

(29) M T Barros, H Siljak, P Mullen, et al., **Objective Multi-variable Classification and Inference of Biological Neuronal Networks** (2020).

(30) J B Bardin, G Spreemann, K Hess, **Topological exploration of artificial neuronal network dynamics**, *Network Neuroscience* (2019) 1-28.

(31) Medina, R., Breña, J. L., y Esenarro, D. (2021). **Efficient and sustainable improvement of a system of production and commercialization of Essential Molle Oil (Schinus Molle)**.3*C Empresa. Investigación y pensamiento crítico*, 10(4), 43-75. https://doi.org/10.17993/3cemp.2021.100448.43-75

(32) Meng Siyu & Zhang Xue.(2021). **Modelling and Simulation of Collaborative Innovation System in Colleges and Universities Based on Interpreted Structural Equation Model**. *Applied Mathematics and Nonlinear Sciences* (1). https://doi.org/10.2478/AMNS.2021.2.00151.