Retraction

Retracted: Modeling and Analysis Method of National Fitness Big Data for Basketball Projects Based on a Multivariate Statistical Model

Security and Communication Networks

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1].

This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Q. Zhao, "Modeling and Analysis Method of National Fitness Big Data for Basketball Projects Based on a Multivariate Statistical Model," *Security and Communication Networks*, vol. 2022, Article ID 2591633, 11 pages, 2022.

WILEY | Hindawi

Research Article

Modeling and Analysis Method of National Fitness Big Data for Basketball Projects Based on a Multivariate Statistical Model

Qi Zhao [ID]

Department of Physical Education, Shanghai Publishing and Printing College, Shanghai 200093, China

Correspondence should be addressed to Qi Zhao; zhaoqiww@sppc.edu.cn

How to start from the fitness needs of people and effectively improve the precision of the supply of public fitness services everyone is an important issue that needs to be solved first at the current stage. This requires us to proceed from the reality, conduct accurate research, and find a method that can match the current problem. In this paper, taking basketball projects in national fitness as an example, by introducing a proposition about the development of small basketball events, the corresponding big data modeling and analysis methods are studied. The research methods and research objectives involved in this paper are based on the relevant parameters of the multivariate statistical model. First, the article introduces the calculation principle of the multiple linear regression model. We introduce the concept of variance inflation factor involved in this principle and carry out the modeling and analysis of big data based on this variable. In order to illustrate the application effect of big data in this kind of research, this paper introduces three different big data technologies, including immune selection optimization algorithm, particle swarm optimization algorithm, and Elman neural network, to predict and analyze the variance inflation factor (VIF) corresponding to the small basketball project. The analysis results show that the Elman network exhibits certain advantages in terms of computing convergence time. And, as the number of calculation steps increases, the superiority of the Elman network is more obvious. As far as the prediction performance is concerned, the square of the correlation coefficient corresponding to the immune selection optimization algorithm is the largest and the sum of the squares of the residuals is the smallest, showing superior prediction performance.

1. Introduction

National fitness [1, 2] and healthy China is an important strategy for national development in the new era. From the internal logic point of view, the goal of national fitness is to achieve a healthy China. In addition, healthy China has become an institutional guarantee to meet the needs of the Chinese people for a better life by means of the extensive development of national fitness.

National fitness and national health involve two major elements of sports and health [3, 4], so the deep integration of the two cannot be achieved by the sports department alone. This requires multiple forces such as health care, education, culture, finance, gardening, and construction to work together to achieve the goal of deep integration [5].

With the advent of the new century, internet technology [6, 7] has achieved unprecedented development and has become an important force leading the third industrial revolution. It has produced a remarkable impact in the world, bringing a new experience to the production and life of the whole society. Among them, the big data technology involved in internet technology, or digital technology, is also widely used in all aspects of social life. With the in-depth integration of digital technology and national fitness public services, digital technology has gradually become an important form of social governance by empowering national fitness public services with precise supply, intelligent service, and intelligent governance. It also provides a strong technical support for promoting the high-quality development of China national fitness public services. In this context, all parts of the country have strengthened the informatization

construction of the national fitness public service system. This not only enhances the specific application of digital technology in the field of national fitness public services, but also improves the governance capacity and level of national fitness public services. It can be seen that the empowerment of national fitness public service governance through digital technology is of great significance for creating a social governance pattern of coconstruction, cogovernance, and sharing, and realizing the modernization of national fitness public service governance. However, the modernization of public service governance for national fitness empowered by digital technology in China is still in its infancy. Therefore, the problems of supply and demand deviation and imbalance of public services for national fitness need to be solved urgently.

Looking back at the relevant research in China in recent years, the digital governance of public services for national fitness can be roughly divided into the embryonic stage from the beginning of the 21st century to 2013 and the initial development stage from 2014 to the present.

Based on the abovementioned analysis, we can see that the modernization of public service governance for national fitness in the digital era is centered on digital technology, guided by the needs of people's livelihood and based on the relationship between supply and demand. This method can provide digital and intelligent governance services for the public service field of national fitness. This implementation can achieve the long-term goal of modernizing national governance.

However, it is undeniable that the digital governance of national fitness public services, as a complex systematic project, involves many participants. Therefore, its corresponding technical governance is more difficult. Although digital technology has been introduced into the public service governance of national fitness, due to the lag in the governance concepts of the government, society, and other multiple subjects and the lack of technological innovation capabilities, the problem of weak digital infrastructure construction in the field of national fitness public services has emerged. In addition, the bottleneck phenomenon encountered in the application of technological innovation is also more prominent. Therefore, it is necessary to solve or improve the abovementioned problems by improving the big data calculation method of national fitness.

Technologies such as big data [8, 9] and cloud computing [10, 11] are widely used in different fields such as engineering and social sciences. For example, traditional intelligence techniques are used in forecasting and research processes in various fields. In addition, statistical knowledge similar to multivariate statistical models can also be applied to the big data modeling and analysis process of national fitness. This paper takes the basketball project in national fitness as an example and first determines the main research content and research objectives of the article through a multivariate statistical model. Then, by introducing three new big data technologies, the research function obtained by the multivariate statistical model is predicted and analyzed, in order to provide suggestions for the reform and progress of national fitness.
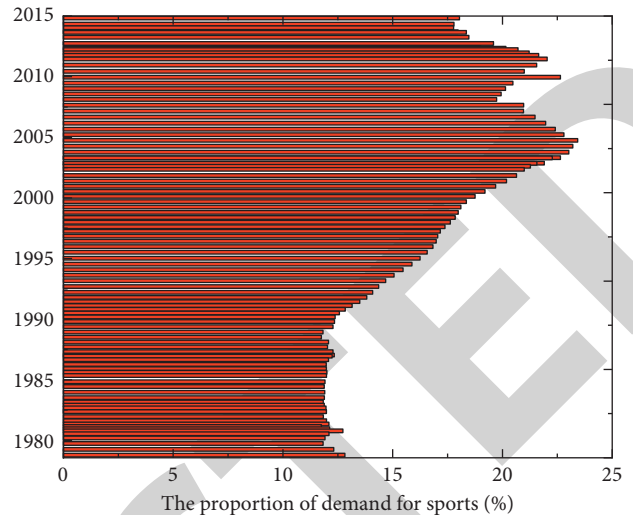

FIGURE 1: Trends in the demand for sports among the masses.

2. National Fitness Development in the New Era

In the new era, we should promote the transformation and upgrading of the national fitness public service system under the guidance of precise supply.

As far as the reality is concerned, problems such as unbalanced distribution of mass sports resources and insufficient supply still exist. It is mainly manifested in the lack of equalization of public services for national fitness between urban and rural areas, regions and different groups, and the inability to meet the diverse needs of the masses caused by the insufficient targeting ability of the traditional supply model. This hinders the further implementation of the national fitness strategy to a certain extent. Figure 1 shows the changing trend of the Chinese people's demand for sports in the past 30 years. As shown in Figure 1, after entering the 21st century, the Chinese people's demand for sports has gradually increased, and it has remained within a relatively high proportion.

Among them, after 2005, the desire of Chinese people to participate in fitness seems to have reached a peak. The reason for this phenomenon may be due to China's successful application for the 2008 Olympic Games.

The unbalanced economic and social development leads to the staged structural dislocation in the development of sports. Therefore, sports policies should be targeted by categories and angles. At present, a large number of innovative practice methods have been adopted in the field of sports, and comprehensive and precise interventions have been carried out in school sports, rural sports, and competitive sports. These types of interventions are based on their respective specific issues. Through accurate positioning of the problem, solutions and measures can be selected, and a series of precise solutions can be proposed.

The rapid development of high-tech means such as big data, cloud computing, and block chain has promoted the transformation and upgrading of the economy and society, and brought revolutionary innovation to traditional industries. High-tech that pays attention to individual needs

and individualized development can meet the individual sports needs of the masses in various ways. Therefore, the use of big data and other high technologies in the field of national fitness to respond to the increasingly diverse sports demands of the masses provides a technical possibility for the transformation and upgrading of the public service system for national fitness.

Finally, the development of national fitness is inseparable from the precise control and guidance of the government. The requirement of precise implementation of government policies runs through all aspects of the development of socialism with Chinese characteristics in the new era. The government's precise policy preference has achieved excellent results in promoting economic and social development, which also provides lessons for the modernization of the sports governance system. The national fitness public service, a field involving the public, has begun to implement precise policies, laying a solid foundation for the precise transformation of the national fitness public service system.

Figure 2 shows the various influencing factors of national fitness in the new era.

As shown in Figure 2, the realization of the national fitness goal is the result of the joint action of four main factors. Figure 2 shows that the main influencing factors mainly include the following four. They are the change of the main contradiction, precise intervention, government intervention, and the application of big data. Among them, the application of big data technology is the most important factor.

Data thinking [12] is the foundation and data application is the superstructure. Data analysis is a stepping stone, and its power is not limited to research itself. It is more important to apply technology at a higher level. We are committed to the multidimensional, systematic, and in-depth application of these basic knowledge. With the wide application of big data technology, the implementation of national fitness is inseparable from the blessing of this new technology. Big data technology is a platform based on artificial intelligence or intelligent algorithms. Compared with the data thinking mentioned in this section, big data technology is equivalent to an evolutionary processing of this data thinking, so that this way of thinking about problems is supported by new technologies. This approach is more reasonable and accurate.

Taking the basketball project in national fitness as an example, the specific purpose is to provide constructive opinions and reform ideas for the development strategy of China's small basketball events. To accomplish this goal, first of all, based on statistical knowledge, this paper first determines the goal and content of the research by introducing a multivariate statistical model. Then, by introducing several typical big data technologies, we compare their effects in the application of multivariate statistical models.

In order to clearly distinguish the layout of the article, here is a certain elaboration of the overall layout of the manuscript. In the third chapter of the article, we introduce the knowledge of multivariate statistical models. In the fourth chapter of the article, three typical big data techniques are introduced throughout the manuscript. The fifth part of
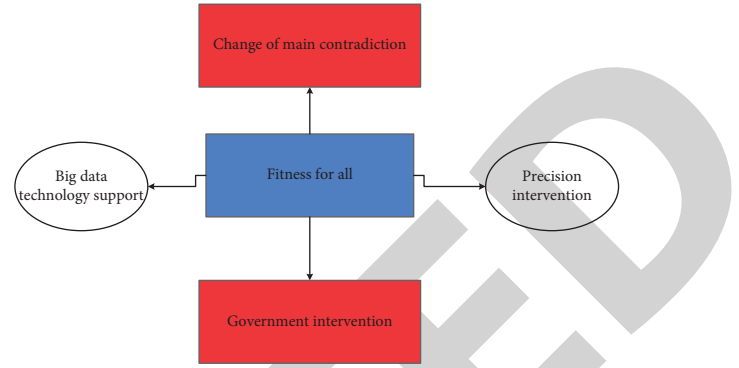

Figure 2: Knowledge modules in the training of data thinking.

the article shows the practical application of the theoretical research part. According to the reviewer's request, the author has added some necessary content. The specific additions are as follows.

3. Multiple Linear Regression Model [7, 13]

The multiple linear regression model can be expressed as follows:

$$W = \lambda_0 + \lambda_1 y_1 + \lambda_2 y_2 + \ldots + \lambda_{m-1} y_{m-1} + \mu, \quad (1)$$

where y represents the dependent variable, $y_1, y_2, \ldots y_{m-1}$ represent m-1 independent variables, $\lambda_0, \lambda_1, \lambda_2, \ldots, \lambda_{m-1}$ are m unknown parameters, and μ is an unobservable random variable with zero mean and variance $\sigma 2$ greater than 0. It can also be called the error term. It is well known and usually assumed that $\mu \sim N(0, \sigma 2)$.

After the multiple regression model is initially established, whether it really explains the relationship between the predictor variable and the dependent variable needs to be tested for significance.

The decision linearity exponent R^2 describes the proportion of the total change in Y that is reflected by the value of the linear function of the independent variable. The result of the coefficient of determination is between 0-1. As we all know, the larger the coefficient of determination, the better the fitting effect. The coefficient of the determination formula is as follows:

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} = 1 - \frac{\left(\sum_{i=1}^m \left(\widehat{y}^{(i)} - y^{(i)} \right)^2 \right)/m}{\left(\sum_{i=1}^m \left(\widehat{y}^{(i)} - y^{(i)} \right)^2 \right)/m} = 1 - \frac{MSE(\widehat{y}, y)}{\mathrm{Var}(y)}, \quad (2)$$

where MSE stands for mean squared error and SSE is called residual sum of squares. The SSR is called the regression sum, which reflects the sum of squared deviations of the linear function of the independent variable in each group of observations. The SST is called the sum of squares of the total deviation and is used to measure the degree of difference of the dependent variable itself. It can also be called the total change in data.

Before the overall mathematical experiment begins, we should first calculate the statistic t_i; then, we can look up the t

distribution table according to the given significant level α, degrees of freedom $n - k - 1$, and get the critical value t_α or $t_{\alpha/2}$. If $t > t - \alpha$ or $t_{\alpha/2}$, the regression coefficient b_i is significantly different from 0. On the contrary, the test result is not significantly different from 0. The calculation formula of the statistic t can be expressed as follows:

$$t_i = \frac{b_i}{s_y \sqrt{C_{ij}}} = \frac{b_i}{s_{bi}}. \tag{3}$$

There are two common types of variables: quantitative variables and categorical variables. Among them, categorical variables are also called attribute variables, that is, the score of the variable is an attribute or can be classified. However, categorical variables should not be used directly in regression analysis. Because the equal spacing between discrete values assigned to categorical variables masks differences between categories. It is well known that dummy variables are one of the classic ways to solve this problem. Anyone with k attributes can be defined as a set of k dummy variables whose value is 1 or 0. This requires us to construct a dummy variable. It is worth noting that the conversion calculation discards a dummy column. Only then can we get a full rank matrix.

One of the main assumptions of multiple linear regression models is that the independent variables are not strongly correlated with each other, otherwise, multicollinearity problems will arise. A major problem with multicollinearity is that it causes the significance of the multiple linear regression coefficients to deviate from the true direction. To determine whether there is multicollinearity between two variables, the most common way is to use the VIF to correct.

The variance inflation factor refers to the ratio of the variance in the presence of multicollinearity among the explanatory variables to the variance in the absence of multicollinearity. VIF is obvious that the larger the variance inflation factor, the more severe the collinearity is. The calculation method can be expressed as follows:

$$VIF_i = \frac{1}{1 - R_i^2}. \tag{4}$$

In the formula, VIF represents the variance inflation factor of the independent variable x. It mainly represents the degree of data dispersion of this data series. The empirical judgment method shows that when $0 < \text{VIF} < 10$, there is no multicollinearity. When $10 \leq \text{VIF} < 100$, there is strong multicollinearity. When $\text{VIF} \geq 100$, there is severe multicollinearity. R_i^2 is the coefficient of repeated measures for the regression of other independent variables when x_i represents the ith independent variable.

4. Big Data Modeling and Analysis Methods

4.1. Immune Selection Optimization Algorithm. The biological immune system is a complex adaptive system [14, 15]. The human immune system can recognize pathogens and respond to them, so it has certain abilities of learning, memory, and pattern recognition. This way is similar to an external stimulating antigen to stimulate the human immune system to produce antibodies adapted to it. That is to say, an input variable corresponds to a unique output function. In this way, the principle and mechanism of its information processing can be described by computer algorithms to solve scientific and engineering problems. Algorithmic immunity retains some characteristics of the biological immune system and introduces them to solve optimization problems.

A population suppression process is added to the immune algorithm to control the average concentration of the population and avoid premature convergence of the algorithm to a locally optimal solution. This increases the global optimization capability.

We take the VIF involved in the previous section as a research variable and optimize it through the immune selection optimization algorithm.

A population suppression process is added to the immune algorithm to control the average concentration of the population and avoid premature convergence of the algorithm to a locally optimal solution. This increases the global optimization capability. Figure 3 is the calculation flow chart of the immune selection optimization algorithm.

A typical multipeak function is used to enhance the application of the immune algorithm. The multipeak function can be expressed as follows:

$$S(z) = \sum_{k=1}^{m} \left(600 \left(z_{k+1} - z_k^2 \right)^2 + \left(1 - z_k \right)^2 \right). \tag{5}$$

In the formula, z represents the independent variable that meets the conditions, and S represents the value of the dependent variable after the convergence of multiple nonlinear calculations.

The global minimum point of the multimodal function is obtained when all the independent variables are 1, and the minimum value of the function is 0. The search interval for the independent variable is $(-5, 5)$. The specific parameter settings of the algorithm are shown in Figure 4. As shown in the figure, NP denotes the number of antibody population sizes, G denotes the maximum number of cycles, and NC represents the number of clones. As shown in Figure 4, NP is equal to 220, G is equal to 198, and NC is equal to 238.

The optimization results of 10 trials of the immune algorithm are shown in Figure 5. It can be seen from Figure 5 that the immune algorithm has a good ability to search for multidimensional and multipeak functions. Therefore, it can be applied to solve optimization problems about VIFs in multivariate statistical models.

4.2. Particle Swarm Optimization Algorithm. Particle swarm optimization algorithm [16, 17] is used to simulate various biological social behaviors such as biological reproduction and upgrading. This artificial intelligence algorithm consists of a finite number of particles that are unrelated to each other. These particles automatically search for a single best position and a global best position according to the optimal problem solution, according to the optimization criteria
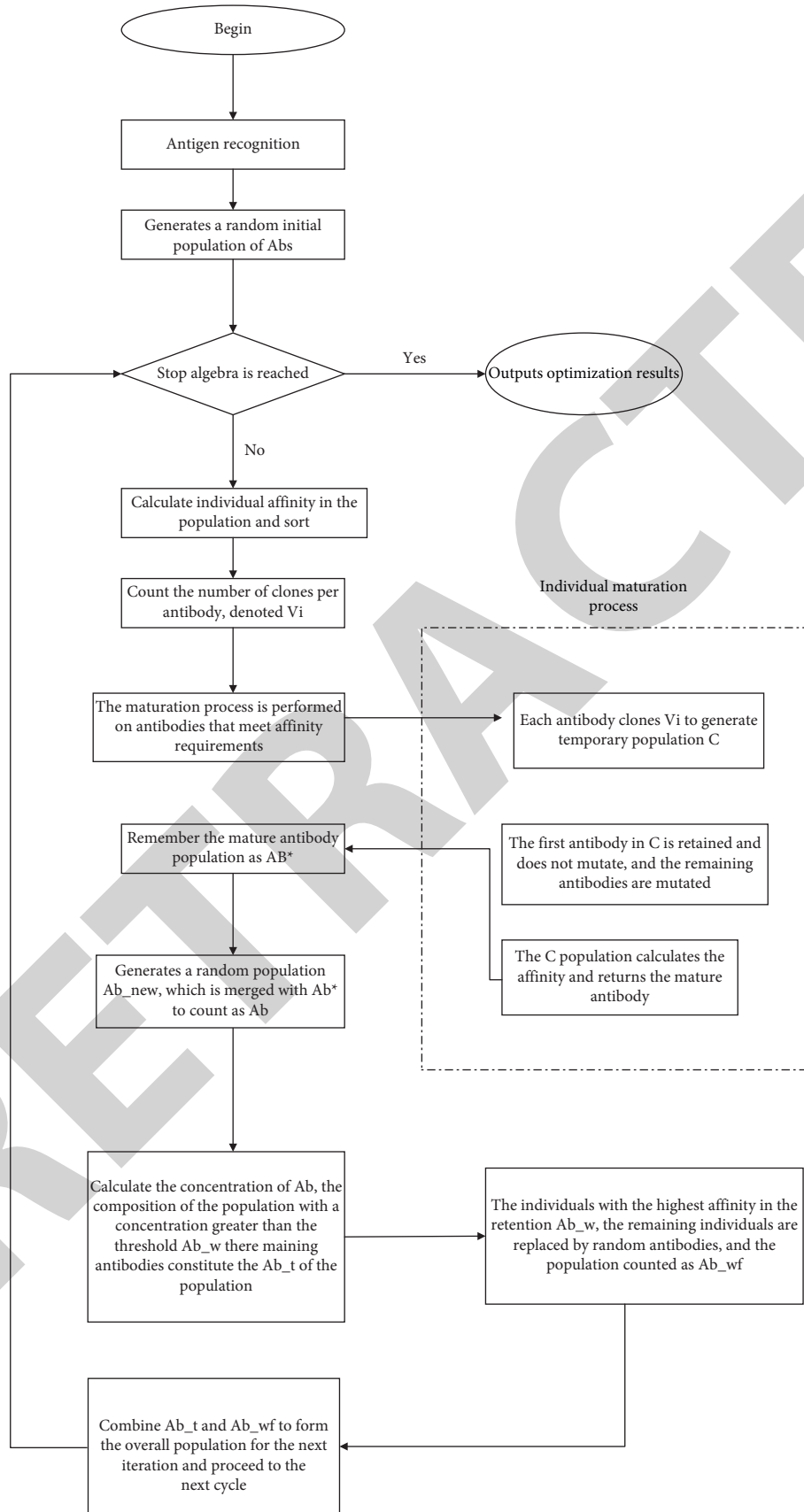
Figure 3: The calculation flow chart of the immune optimization algorithm.

found in nature. Each iteration the researchers got during the computation was reupdated based on the particle's position and velocity. The calculation steps for updating the relative motion trajectory of each particle can be expressed as follows.

$$L_i = \lambda \times L_i + D_1 R_1 \left(Q_{\text{best}} - Y_i \right) + D_2 R_2 \left(K_{\text{best}} - Y_i \right),$$
$$K_i = K_i + L_i. \tag{6}$$

4.3. Elman Networks. Neural networks [18, 19] are widely used for their large-scale parallel distributed structure, learning ability, and generalization ability. The main advantages are as follows: nonlinear analysis capability, convenient input/output mapping, adaptive capability, evidence response, background information, strong fault tolerance, VLSI (very-large-scale integrated) implementation, analysis and design consistency, and neural biological analogy. This paper takes the Elman neural network as an example to describe the implementation process of traditional neural network prediction in detail.

The calculation process of the Elman network can be expressed as follows.

For the input layer, the Elman network can be represented as follows:

$$x_i^0 = x_i(k),$$
$$\begin{cases} s_i^1 = \sum_{j=1}^{n^0} w_{ij}^0 x_j^0(k) + \sum_{j=1}^{n^1} w_{ij}^2 c_j^0(k) \\ x_i^1 = f1\left(s_i^1(k) \right). \end{cases} \tag{7}$$

Similarly, the implicit value in the network can be expressed as follows:

$$-\frac{\partial F(j)}{\partial w_{ij}^0} = -\sum_{l=1}^{r} \frac{\partial F(j)}{\partial y_l(j)} \cdot \frac{\partial y_l(j)}{\partial w_{ij}^0} = \sum_{l=1}^{r} e_l(j) \cdot f2'\left(s_l^3(j) \right) \cdot w_{ij}^1(j) \cdot \frac{\partial x_i^1(j)}{\partial w_{ij}^0}. \tag{8}$$

Through comprehensive calculation, we can get the following:

$$\begin{cases} -\frac{\partial F(j)}{\partial w_{ij}^0} = \sum_{l=1}^{r} e_l(j) \cdot f2'\left(s_l^3(j) \right) \cdot w_{ij}^1(j) \cdot \chi_{ij}^i(j) \chi_{ij}^i(j) = f1'\left(s_i^1(j) \right) \cdot \left(x_j^0(j) \right) + \sum_{m=1}^{n_1} w_{im}^2 \cdot \chi_{ij}^m(j-1). \end{cases} \tag{9}$$

The key to the nonlinear ability and learning ability of the network lies in the continuous correction of the weights. There are two methods for recurrent network training, one is the batch mode, the other is the online mode, the Elman network adopts the latter.

5. Example Verification and Analysis

This paper takes a small basketball project as an example to study the application of multivariate statistical models in the process of big data analysis and modeling. The small basketball is a basketball game tailored by the Chinese Basketball Association for young people. By changing the rules and equipment of adult basketball, basketball can adapt to the physical and mental development characteristics of young people of all ages. It is also a social sports project for children and teenagers that is mainly promoted by the Chinese Basketball Association. The new era of small Chinese basketball began to gradually take shape after Yao Ming became the new chairman of the Chinese Basketball Association in 2017.

Although the development of small basketball events in China is in full swing, the development of small basketball events has not kept up with the development of small basketball events, and there are still many problems in the development of small basketball events. These problems mainly include the long-term cycle of large-scale competitions held by the country; the small number of competitions; the small number of referees

and coaches; the imperfect rules and other problems. This paper analyzes the pros and cons, and opportunities threats faced in the development of small basketball events in China from a strategic perspective. Through this method, we hope to provide strategic reference and reference for promoting the scientific, stable, and sustainable development of small basketball events in China.

This paper systematically studies the internal strengths (strengths), internal weaknesses (weaknesses), external opportunities (opportunities), and external threats (threats) faced in the development of small basketball events in China. In addition, we analyze the possible strategic combinations in the development of small basketball events in China.

Among them, internal advantages and external threats are two objective facts, which will not change with the development of the event. However, internal disadvantages and external opportunities can be improved according to the nature of the event.

The direct impact matrix of the VIF of the four factors faced in the development process of Chinese small basketball events can be drawn in Figure 6.

Among them, the number of data sets corresponding to each factor is 6,000, so the total number of data sets studied in this paper is 24,000.

The direct influence matrix of VIF is composed of four main elements: influence degree, influenced degree, centrality degree, and cause degree. The measurement criteria of the influence degree of the four main elements, the degree of
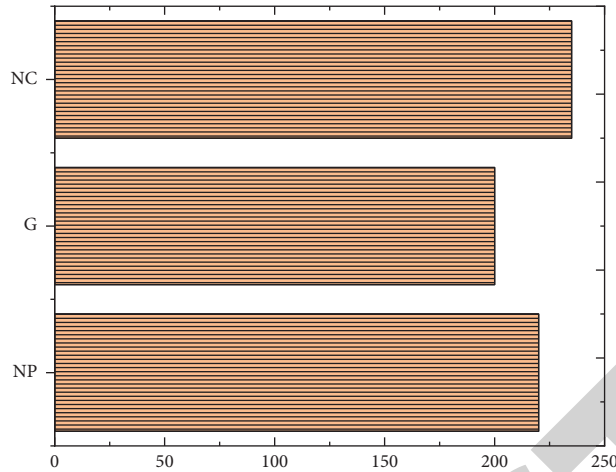
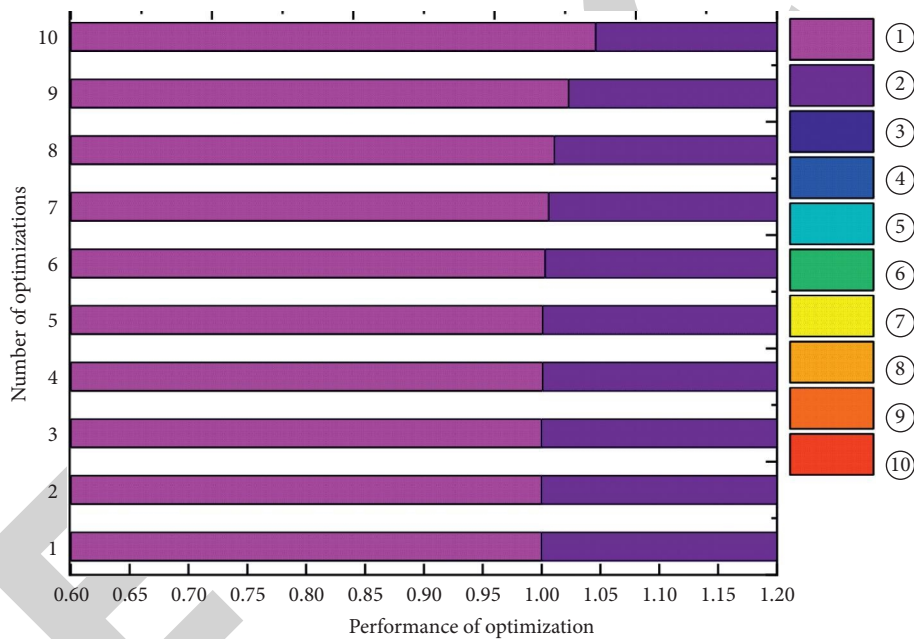FIGURE 4: Specific application parameters of the immune algorithm.

FIGURE 5: Function optimization results.

influence, the degree of being influenced, the degree of centrality, and the degree of cause, in the system can be calculated from the comprehensive influence matrix. We can add the elements of each row in the comprehensive influence matrix to get the influence degree of the corresponding variable. The sum of the influence degree and the influenced degree is the centrality degree. A large number of studies have proved that the greater the centrality, the stronger the effect on the research target. That is to say, the larger the corresponding value of VIF, the stronger the linearity of the corresponding multivariate statistical mode. The difference between the influence degree and the influenced degree is the cause degree. This effect is called the causal factor of the variable. The smaller the cause degree, the smaller the influence factor is susceptible to the influence of other influencing factors, and the corresponding VIF is larger. The

variables represented by this phenomenon are called outcome factors. Figure 7 shows the VIF metrics corresponding to four factors.

As shown in Figure 7, the cause degree can reflect the categories of the influencing factors. The larger the causality degree of the influencing factor, the more the factor is the Causal Index in the influencing factor system. In addition, the smaller the causality degree of the influencing factor, it means that the factor is the Result Index in the influencing factor system. We can obtain the multivariate statistical model of the influencing factors corresponding to the four factors through the theory of the multivariate statistical model mentioned previously. We take weaknesses as an example, and the corresponding multivariate statistical model is obtained by mathematical fitting as follows:
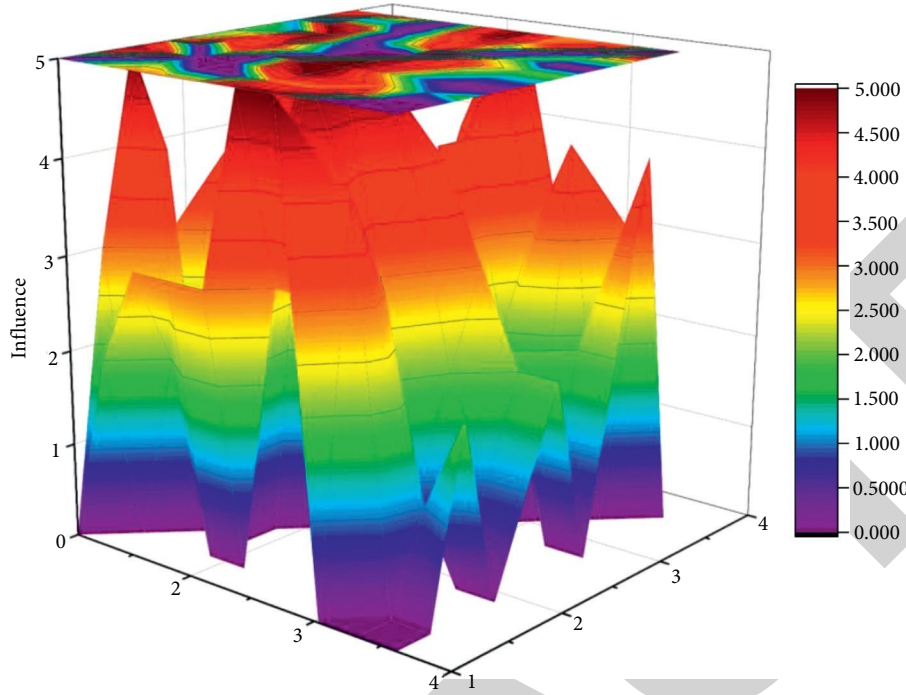
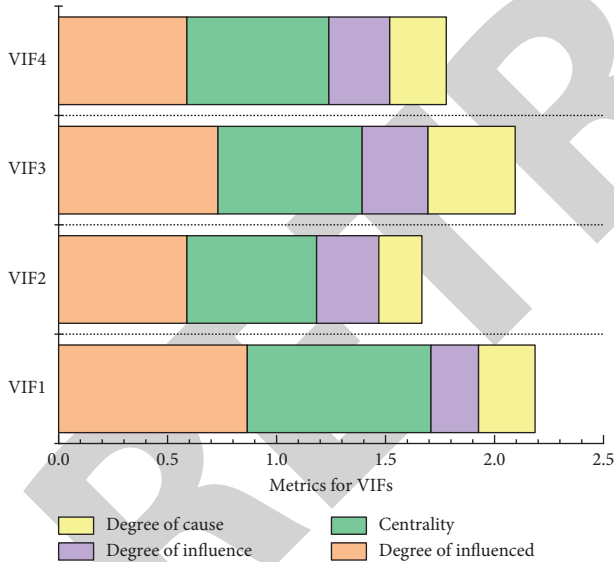FIGURE 6: Convergence comparison of algorithms.

FIGURE 7: Influencing factors of new media copyright development in sports events.

$$y = 0.78 + 4.5x_1 + 3.6x_2 + 3.8x_3 + 4.9x_4. \quad (10)$$

In the formula, y represents the weaknesses corresponding to the small basketball event, x_1 represents the influence degree corresponding to the factor, x_2 represents the influence degree corresponding to the factor, x_3 represents the centrality corresponding to the factor, and x_4 represents the cause degree corresponding to the factor. And, the fitting process of formula (10) can be approximated by the fitting curve. The specific linear fitting curve is shown

in Figure 8. It can be seen from Figure 8 that for the four independent variables, for the weaknesses corresponding to the small basketball items, the linear fitting relationship is good, and the scatter points are basically distributed on both sides of the straight line.

Similar to weaknesses in the multivariate statistical model, the other three influencing factors can also find their corresponding multivariate statistical models through similar methods. Due to the limitation of the length of the article, the calculations are not explained one by one.

In order to better analyze the degree of linearity in the multivariate statistical model, three big data technologies such as immune selection optimization algorithm, particle swarm optimization algorithm, and Elman neural network are used to predict and analyze the corresponding VIF value in the multivariate statistical model. Similar to the abovementioned analysis, we still use the corresponding internal disadvantages in the small basketball project for analysis.

It is quite necessary to predict the computational convergence ability of the three techniques before performing intelligent computation. To this end, when the calculation iteration steps of the three big data technologies are 200, 400, 600, 800, and 1000 respectively, we compare the convergence completion time of the three artificial intelligence algorithms. The specific test results are shown in Figure 9.

As can be seen from Figure 9, the Elman network has the least computation time. In addition, with the increase of the number of computational iterations, the computational superiority exhibited by the Elman network continues to increase.

Through the three abovementioned intelligent algorithms: immune selection optimization algorithm, particle swarm optimization algorithm, and Elman network, the
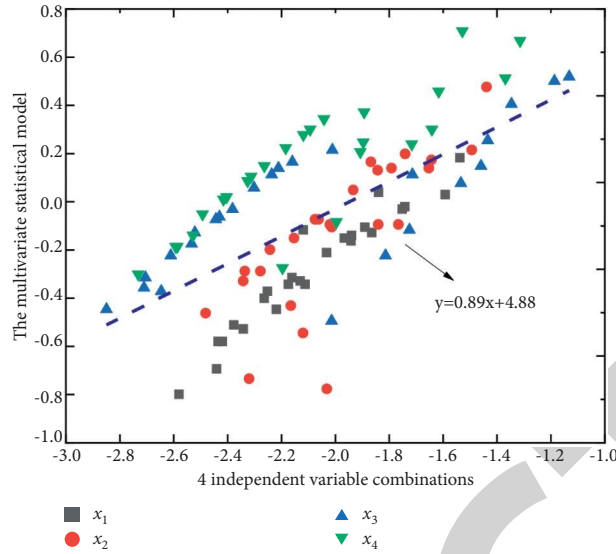
FIGURE 8: The fitting curve of the multivariate statistical model represents the relationship.
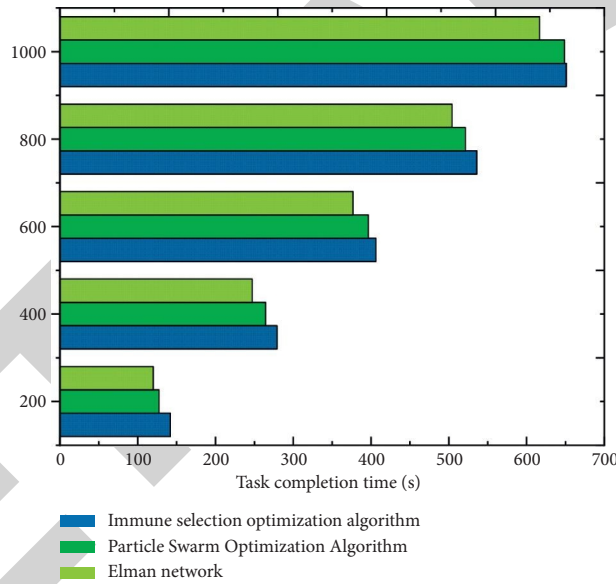

FIGURE 9: Comparison of task completion time.

corresponding VFI value in the multivariate statistical model is predicted and studied. It is well known that the squared correlation coefficient (r2) and root mean squared difference (RMSE) (SSE) are two typical predictors. The next step is to compare r^2 and SSE of the three algorithms. The closer R^2 is to 1, the smaller SSE, and the higher the prediction accuracy. Table 1 shows the prediction performance of the three artificial intelligence algorithms. Among them, the determination coefficient can be obtained by the following calculation method:

$$r^2 = \frac{\left[\sum_{i=1}^{n} (x_i - x_m)^2\right] - \left[\sum_{i=1}^{n} (x_i - y_i)^2\right]}{\left[\sum_{i=1}^{n} (x_i - x_m)^2\right]}. \quad (11)$$

The root mean square difference can be obtained by the following calculation method:

$$RMSE = \sqrt{\frac{1}{n} \times \sum_{i=1}^{n} \left|(y_i - x_i)^2\right|}. \quad (12)$$

It can be seen from Table 1 that compared with the three algorithms, the square of the correlation coefficient corresponding to the immune selection optimization algorithm is the largest, the maximum value is 0.9813, the root mean squared difference is the smallest, and the minimum value is 11. This shows that the prediction effect of the immune selection optimization algorithm is the best and it can be

TABLE 1: Comparison of prediction performance of three big data technologies.

Big data technology	The effect of regression analysis	
	Squared correlation coefficient (r^2)	Root mean squared difference (RMSE)
Immune selection optimization algorithm	0.9813	0.11
Particle swarm optimization algorithm	0.9473	0.21
Elman neural network	0.6851	0.36

used as a key technology for evaluating the multivariate statistical model of small basketball projects. At the same time, the multivariate statistical model introduced in this paper combined with the prediction method of big data technology provides a new processing idea for the modeling and analysis of national big data.

6. Conclusion

(1) It is foreseeable that under the current social development background, if the national fitness wants to maintain a sustainable development trend, it must recognize the mainstream trend of current social development. At the same time, we should combine our own development plan, with the strong support of new technologies, change the path and method of development, strengthen concept innovation, improve our own system, and adopt scientific management and operation methods. Only in this way can the sustainable development of national fitness be guaranteed.

(2) National fitness in the new era requires the blessing and assistance of technology. The application of big data, cloud computing, and other technologies can effectively improve the vitality of national fitness. In this paper, the multivariate statistical model is introduced to study the influence degree of the internal strengths (strengths), internal weaknesses (weaknesses), external opportunities (opportunities), and external threats (threats) corresponding to the small basketball project. The research results show that the four main factors have a good multilinear relationship with their corresponding degree of influence, degree of centrality, and degree of cause.

(3) The article introduces three different big data technologies: immune selection optimization algorithm, particle swarm optimization algorithm, and Elman neural network, to predict and analyze the variance inflation factor VIF corresponding to the small basketball project. The prediction effect shows that for the research cases introduced in this paper, the prediction effect of the three intelligent algorithms is ranked from good to bad: immune selection optimization algorithm, particle swarm optimization algorithm, and Elman neural network. A large number of facts have proved that big data technology can be effectively applied to the eye-catching evaluation research of national fitness. At the same time, this algorithm can also provide a certain reference for the modeling and analysis of national fitness big data.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declared that they have no conflicts of interest.

References

[1] F. B. Ortega, J. R. Ruiz, I. Labayen, C. J. Lavie, and S. N. Blair, "The Fat but Fit paradox: what we know and don't know about it," *British Journal of Sports Medicine*, vol. 52, no. 3, pp. 151–153, 2018.

[2] R. T. Bartee, K. A. Heelan, and B. L. Dority, "Longitudinal evaluation of aerobic fitness and academic achievement among schoolchildren," *Journal of School Health*, vol. 88, no. 9, pp. 644–650, 2018.

[3] K. Kaasalainen, K. Kasila, J. Komulainen, M. Malvela, and M Poskiparta, "Changes in psychosocial factors and physical activity among Finnish working-age men in the adventures of joe Finn campaign," *The International Quarterly of Community Health Education*, vol. 39, no. 1, pp. 39–49, 2018.

[4] E. Fuzeki, T. Engeroff, and W. Banzer, "Author's reply to lopez: comment on "health benefits of light-intensity physical activity: a systematic review of accelerometer data of the national health and nutrition examination survey (nhanes)"," *Sports Medicine*, vol. 48, no. 3, pp. 749-750, 2018.

[5] A. S. Harrison, L. Tang, and P. Doherty, "Are physical fitness outcomes in patients attending cardiac rehabilitation determined by the mode of delivery," *Open Heart*, vol. 5, no. 2, Article ID e000822, 2018.

[6] X. Q. Jiang, T. T. Xu, and X. J. Dong, "Campus data analysis based on positive and negative sequential patterns," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 33, no. 5, 2019.

[7] B. Sánchez-Torres, J. A. Rodríguez-Rodríguez, D. W. Rico-Bautista, and C. D. Guerrero, "Smart campus: trends in cybersecurity and future development," *Revista Facultad de Ingeniería*, vol. 27, no. 47, pp. 104–112, 2018.

[8] A. J. Ferrer, J. M. Marques, and J. Jorba, "Towards the decentralised cloud: survey on approaches and challenges for mobile," *ACM Computing Surveys*, vol. 51, no. 6, pp. 51–36, 2019.

[9] S. Fatima and S. Ahmad, "An exhaustive review on security issues in cloud computing," *Ksii Transactions On Internet and Information Systems*, vol. 13, no. 6, pp. 3219–3237, 2019.

[10] W. Zhang, X. C. Zhang, H. L. Shi, and L. Zhou, "RDDSACCGA: a reliable data distribution solution assisted by cloud computing based on genetic algorithm," *Cluster Computing*, vol. 22, no. S5, pp. 11069–11077, 2019.

[11] P. Fraga-Lamas, M. Celaya-Echarri, P. Lopez-Iturri et al., "Design and experimental validation of a LoRaWAN fog

computing based architecture for IoT enabled smart campus applications," *Sensors*, vol. 19, no. 15, 2019.

[12] F. Wan, "Matched or unmatched analyses with propensity-score-matched data?" *Statistics in Medicine*, vol. 38, no. 2, pp. 289–300, 2019.

[13] K. Sakunthala, S. Iniyan, and S. Mahalingam, "Forecasting energy consumption in Tamil Nadu using hybrid heuristic based regression model," *Thermal Science*, vol. 23, no. 5 Part B, pp. 2885–2894, 2019.

[14] Y. Que, W. Zhong, H. L. Chen, X. Chen, and X. Ji, "Improved adaptive immune genetic algorithm for optimal QoS-aware service composition selection in cloud manufacturing," *International Journal of Advanced Manufacturing Technology*, vol. 96, no. 9-12, pp. 4455–4465, 2018.

[15] Z. Zareizadeh, M. S. Helfroush, A. Rahideh, and K. Kazemi, "A robust gene clustering algorithm based on clonal selection in multiobjective optimization framework," *Expert Systems with Applications*, vol. 113, pp. 301–314, 2018.

[16] L. Wang and Y. Q. Liu, "Application of simulated annealing particle swarm optimization based on correlation in parameter identification of induction motor," *Mathematical Problems in Engineering*, Article ID 1869232, 9 pages, 2018.

[17] Y. Zhao, R. L. Shan, and H. L. Wang, "Research on vibration effect of tunnel blasting based on an improved hilbert–huang transform," *Environmental Earth Sciences*, vol. 80, 2021.

[18] Z. Tang and H. Che, "Big data analysis and modeling of higher education reform based on cloud computing technology," *Security and Communication Networks*, vol. 2022, Article ID 4926636, 11 pages, 2022.

[19] X. G. Li and J. F. Wang, "Traffic detection of transmission of botnet threat using bp neural network," *Neural Network World*, vol. 28, no. 6, pp. 511–521, 2018.