Journal of Computing and Electronic Information Management
ISSN: 2413-1660 | Vol. 12, No. 2, 2024

Momentum Study Based on Logistic Regression Models and BP Neural Networks

Cheng Li [1, *], Yuquan Ji [2], Jingqi Liu[3], Weisong Zhang [2]

[1] School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo, China
[2] School of Physics & Electronic Information Engineering, Henan Polytechnic University, Jiaozuo, China
[3] School of Mechanical and Power Engineering, Henan Polytechnic University, Jiaozuo, China

[*] **Corresponding author**: Cheng Li (Email: 1478895462@qq.com)

Abstract: In the context of a game, "momentum" usually refers to a set of events (e.g., consecutive points scored) that create momentum or a trend in a game, which may have a significant impact on the outcome of the game. The purpose of this paper is to investigate how momentum affects the outcome of a game through mathematical modeling. First, an AHP-TOPSIS model is built to calculate the score of each athlete at each moment to quantify and visualize momentum. Second, based on real game data, a Kruskal-Wallis H model test is established to assess the correlation between momentum and score. To predict the fluctuation of players' momentum, logistic regression model was used to predict the momentum fluctuation. Finally, GABP neural network and genetic algorithm are established to visualize the performance data of the players and predict the inflection points during the game. The model proposed in this paper effectively solves the problem of momentum during the game and improves the ability to predict and analyze the game results.

Keywords: Logistic Regression; AHP-TOPSIS; Kruskal-Wallis H; BP Neural Networks.

---

1. Introduction

Momentum in a game is the gradual dominance of a player or a team in a game, usually through a series of scoring runs, winning key plays or using control tactics. Momentum can change rapidly during a match, not only in relation to scoring, but also in relation to changes in players' mentality, confidence, energy and tactics [1]. Addressing how momentum affects the direction and outcome of a match, this paper proposes a series of models to quantify momentum and examine the factors that influence it.

First, to capture the flow of the game when scoring occurs, an AHP-TOPSIS model is developed in this paper. According to the characteristics given by the data, the performance score of each player at each time point in each game is calculated, the momentum of good and bad players is evaluated, and the result of the change of the momentum of the game is deduced to realize the quantification and visualization of the momentum. Secondly, in order to verify the role of "momentum" in the match, this paper establishes the Kruskal-Wallis H-model test to evaluate the correlation between momentum and score, and in order to predict the fluctuation of players' momentum, the trend of momentum change is taken as the target value, and the momentum fluctuation is predicted by logistic regression model. Finally, in order to predict fluctuations in matching and calculate which features have the strongest relationship with fluctuations, this paper uses neural networks and genetic algorithms to visualize athletes' performance data and predict inflection points during matches. Experimental results show that the model proposed in this paper predicts with high accuracy and speed. According to the analysis of influencing factors provides a reference meaning for coaches and players.

2. AHP-TOPSIS prediction modeling

2.1. Factor analysis

In order to make a more representative classification, it is necessary to reduce the dimensionality of the features. In this paper, factor analysis is used to digitize interdependencies within continuous attributes, grouping a large number of complex relationships into a small number of composite factors [2]. The original system is represented as a new combination of factors by analyzing the effect of variance contribution among variables to reproduce the internal connections. The construction model is as follows:

$$\begin{aligned} x_1 &= b_{11}F_1 + b_{12}F_2 + \cdots + b_{1n}F_m + e_1 \\ &\ldots \\ x_1 &= b_{m1}F_1 + b_{m2}F_2 + \cdots + b_{mn}F_m + e_m \end{aligned} \tag{1}$$

x is the observable random vector. F is a common factor, an unmeasurable random vector.

Matrix form models:

$$\begin{aligned} x_1 &= b_{11}F_1 + b_{12}F_2 + \cdots + b_{1n}F_m + e_1 \\ &\cdots \\ x_1 &= b_{m1}F_1 + b_{m2}F_2 + \cdots + b_{mn}F_m + e_m \end{aligned} \tag{2}$$

Thus, the factor scoring model can be written as:

$$F_1 = w_{11}x_1 + w_{12}x_2 + \cdots w_{1m}x_m \tag{3}$$

Calculating the variance of the common factor shows that there are four factors with eigenvalues greater than 3, with a cumulative interpretation of 74.990361%. As can be seen in Figure 1, the subsequent eigenvalues drop very low, i.e., the 4 components are the "inflection points" of the graph.

The results of the ordinal matrix of factor scores are shown in Table. 1, provided by rotating the matrix of components: the first factor variable error and has a larger load, rotational speed, and whether or not to serve so defined as a drive factor; in the second factor variable, running distance and the number of shots have a larger load, so they are defined as process factors. The third factor variables of volleyball scoring, and irresistible ball have large loadings and thus they are defined as hitting factors.

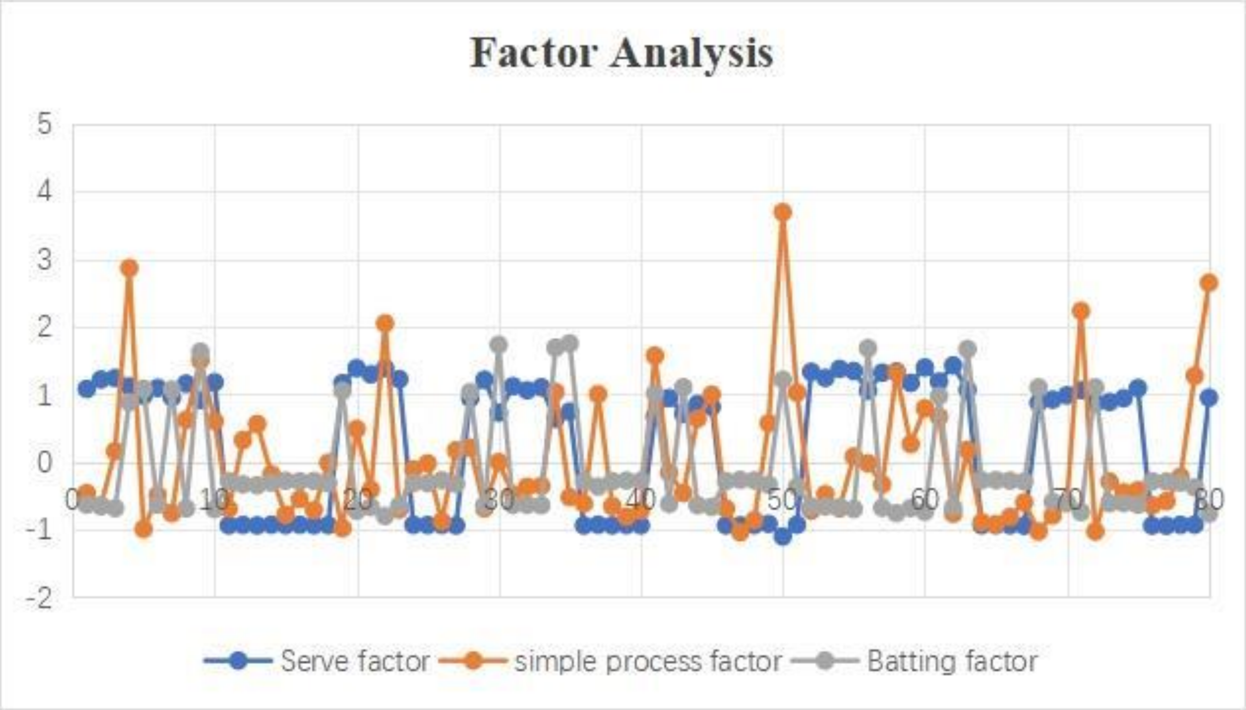

Figure 1. Factor analysis

Table 1. Factor score ordinal matrix

	F1	F2	F3
Lead game	-0.001	-0.035	-0.002
Single shot error rate	0.326	0.008	-0.056
Volley	-0.103	0.018	0.658
Distance the ball travels	-0.015	0.542	-0.036
Batting	0.023	0.546	-0.018
Service speed	0.358	0.006	-0.053
Serve or not	0.365	0.018	-0.074

2.2. AHP-TOPSIS modeling

After the construction of evaluation indexes, the score of each athlete at each moment can be calculated based on the AHP-TOPSIS model. TOPSIS method can rank a limited number of evaluation objects according to the degree of proximity of the evaluation object to the ideal target [3], but it is difficult to determine the weight of the indexes when analyzing the "momentum" of the athlete. The AHP method can overcome this shortcoming to a certain extent, while the TOPSIS method can consider many factors of supply generation, the AHP method can give the weight of each influential factor more comprehensively.

Based on the judgment matrix R constructed by the expert scoring method for the five features of the supply, the largest feature root is noted as $\lambda_{\max}$, then the consistency index can be noted as:

$$A = \frac{\lambda_{\max}-n}{n-1} \tag{4}$$

Suppose we determine the weights W of each evaluation attribute by expert judgment, normalize the decision matrix X min-max, and map the values of each attribute to intervals 0,10,1. The normalization matrix is denoted by X.

$$x'_{ij} = \frac{x_{ij}-\min(X_j)}{\max(X_j)-\min(X_j)} \tag{5}$$

Calculate the ideal and negative ideal solutions for each column in the normalized decision matrix:

$$\begin{cases} Z^{+} = \max(X') \\ Z^{-} = \min(X') \end{cases} \tag{6}$$

Calculate the Euclidean distance between the ideal and negative ideal solutions for each scenario.

$$\begin{cases} D_i^{+} = \sqrt{\sum_{j-1}^{m} \left(x'_{ij} - z_j^{+}\right)^2} \\ D_i^{-} = \sqrt{\sum_{j-1}^{m} \left(x'_{ij} - z_j^{-}\right)^2} \end{cases} \tag{7}$$

Calculate the ratio of the distance of the negative ideal solution to the distance of the ideal solution for each team's composite evaluation.

$$C_i = \frac{D_i^{-}}{D_i^{-}+D_i^{+}} \tag{8}$$

2.3. Analysis of results

Athlete Performance Judgment Matrix The results of the calculations are shown in Table. 2.

Table 2. Athlete performance judgment matrix

Criterion	Service factors	Process factor	Batting average	Weights
Service factors	1	0.5	3	0.3269
process factor	2	1	5	0.5586
batting average	0.3256	0.23	2	0.1223

The consistency ratio of 0.0173 calculated for the athletes' performance is less than 0.10, so the judgment matrix is reasonable.

After processing the data and training the model, we can analyze the results predicted by the model and get the results of the game momentum change to quantify and visualize the momentum, the results are shown in Figure 2. You can view the scoring trend of each player in the game, as well as their scores and historical trends. Then, considering each player's score, scoring trend and average score, we can analyze the result of the game in detail and predict the winner of the game.

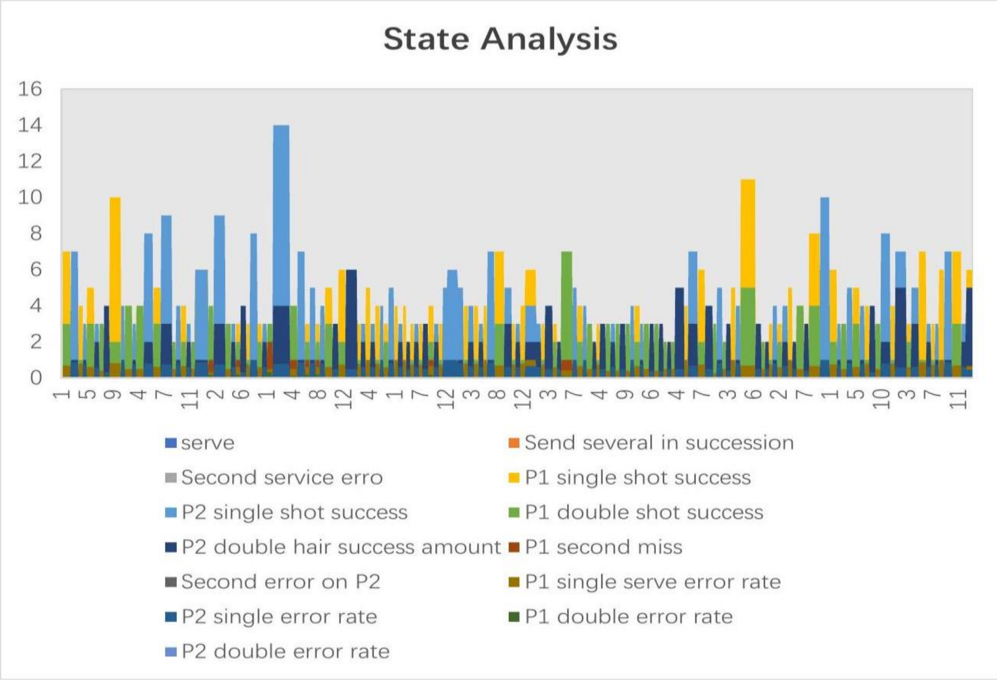

Figure 2. State analysis

3. Kruskal-Wallis H test model

3.1. The Kruskal-Wallis H model

First, the data were grouped based on momentum events, and then the total or mean score for each group was calculated. Next, we will use the Kruskal-Wallis H-test to compare whether there are significant differences in scores between these groups [4].

When the test level is equal to the number of repetitions, namely:

$$m_1 = m_2 \cdots = m_r = m \qquad (9)$$

m is the number of repetitions at each level, which is the ratio of the maximum and minimum variances of r samples. Under the condition of equal variance, quantile of M distribution can be obtained by random simulation method, which depends on the level number r and the degree of freedom $f = m - 1$ of the sample variance, so the distribution can be recorded as $M(r, f)$, then:

$$\delta_1^2 = \delta_2^2 \cdots = \delta_r^2 \qquad (10)$$

The value of M should be close to 1, and when the value of M is large, the difference between the differences between the various approaches is large. The larger the value of M, the larger the difference between the approaches is, then the null hypothesis is rejected. It follows that for a given level of significance, the rejection domain of the test is:

$$L_1 = \{M > M_{1-\delta}(r, f)\} \qquad (11)$$

$M_{1-\delta}(r, f)$ is the M quantile of the $1 - \delta$ distribution.

If the data in an unordered variable does not follow a normal distribution or satisfy the homogeneity of variance, the Kruskal-Wallis H test can be considered to determine the correlation. The Kruskal-Wallis H test is used to check whether there are significant differences in the distribution of multiple populations. The general idea is to mix multiple sets of sample data and sort them in ascending order to find the rank of each variable value. Then, test statistics were constructed by examining whether there were significant differences in the mean values of each rank.

$$M = \frac{12}{N(N+1)} \sum_{i=1}^{k} n_i \left(\grave{R}_i - R\right)^2 \sim \chi^2(k-1) \qquad (12)$$

3.2. Analysis of results

Kruskal-Wallis H test is used to compare whether the medians of two or more independent samples are equal. Here, we can use momentum scores as the group variable and game outcomes as the observed variable, and then perform a statistical test to determine if momentum scores are correlated with game outcomes. The results are shown in Table. 3, Table. 4 and Figure 3.

Table 3. Test of normality

Aggregate	Score or not	Athletes		
Score	3000.0	Statistics	Degree of freedom	Significance
	1	0.236	7256	0
	2	0.149	7256	0

If the value of the test is less than 0.05, the scores do not follow a normal distribution.

Table 4. Kruskal-Wallis H test

Aggregate	15698
Testing statistic	3000.0
Degree of freedom	1
Two-side test	0

The results show that the significance of disorder index is less than 0.001. it can be assumed that the level of different athletes is not identical. That is, the performance of a player is related to his score. By comparing the scoring performance of the players, the visualization shows that the momentum change is not random but related to the scoring status of the players.

Figure 3. Scores

4. Neural network predictive modeling

4.1. GABP model

The data is used as a target to build a predictive model that predicts which player will have an advantage at a given point in time, and thus to understand what is going on in the game. In the coding part of the genetic algorithm, IGABP successively represents the weights and thresholds of the neural network as vectors that constitute the expression of individual genes [5]. Since the structure of the network and the number of weights and thresholds to be determined are already determined during the algorithm's operation, the length of the chromosomes remains constant during the iteration process. Using the forward propagation process of the neural network, the individual fitness is calculated directly, eliminating the computational effort required for training and improving the optimization efficiency of the algorithm [6].

Similar to the neural network, to build to serve factor, process factor, shot factor, px_games (won) the number of small office, whether to serve serve (1), whether hit cow force (irresistible ball), number of strokes as input features, whether as a turning point for the output characteristics of the relational model. In order to explore the relationship between serve factor, process factor, stroke factor, px_games (number of games already won), whether to serve (serve 1), whether to hit an irresistible ball, the number of strokes and whether it is a turning point, the importance of characteristics is reversed.

$$E_k = C_k * \Omega \tag{13}$$

According to the weight value, the influence of the input layer on the hidden layer is judged, so whether the input feature k pair is the inflection point.

$$C_k = [c_{k1} \quad c_{k2} \quad \cdots \quad c_{km}] \tag{14}$$

E_n is the influence of input feature k on whether it is an inflection point. Note whether the input feature k is the contribution rate of the inflection point.

$$PC_k = \frac{E_k}{\sum_{i=1}^{n} E_i} \times 100\% \tag{15}$$

4.2. Analysis of results

The genetic algorithm is used to optimize the BP neural network, and the forward propagation process of the BP neural network is used to calculate the fitness of each individual in the iterative process, so as to improve the optimization efficiency of the algorithm. In the coding part of the genetic algorithm, IGABP continuously represents the weights and thresholds of the neural network as a vector that constitutes the expression of individual genes. Since the structure of the network and the number of weights and thresholds needed to be determined have been determined during the algorithm operation, the length of chromosomes remains unchanged during the iteration process. The experimental results are shown in Figure 4，Figure 5 and Figure 6. According to the analysis of the simulation diagram, when the game turns around, the score and the consecutive score have a strong positive effect on the momentum, and the predicted momentum shows upward curve. When winning the game, players should be encouraged to increase the momentum again and score many times to win the game. The server has an impact on the momentum, and when the other side serves, the probability of the other side winning will become higher, which has a negative effect on the momentum of the own side.

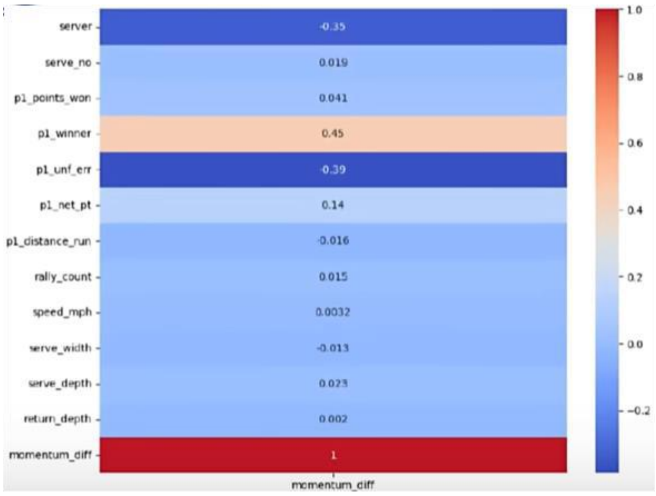

Figure 4. Heat map of momentum change

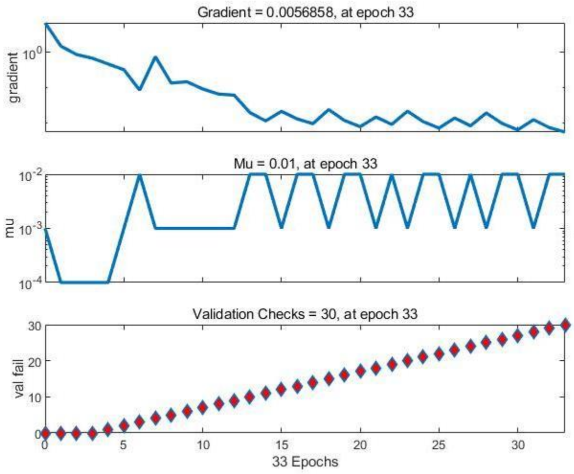

Figure 5. 33 iterations

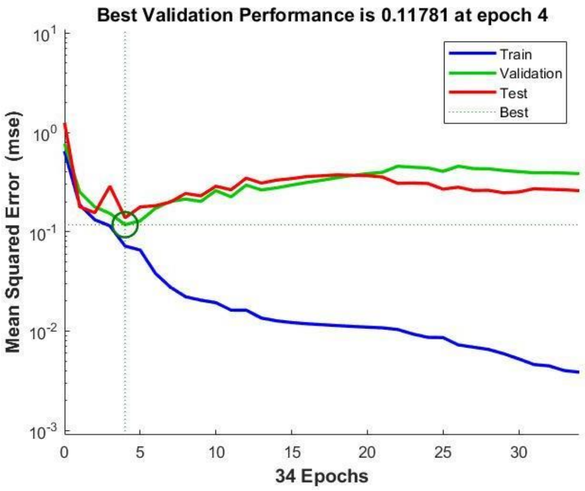

Figure 6. Inflection point forecast

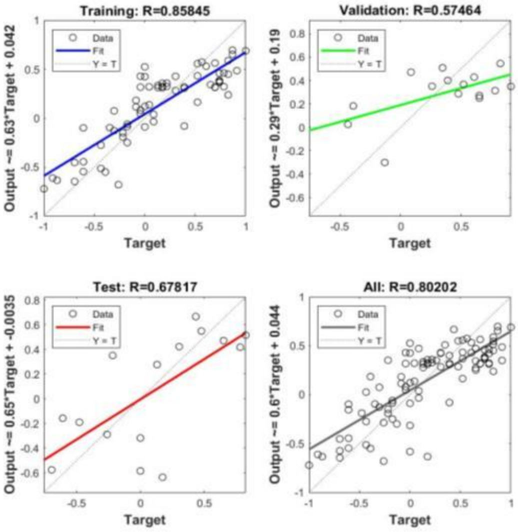

Figure 7. Training

5. Conclusion

In this paper, several models are proposed to solve the problem of momentum change prediction in tennis matches, and these models are analyzed and evaluated in detail. Firstly, factor analysis is utilized to downsize the data and identify the key factors affecting the momentum change, which provides an important reference for the subsequent model building. The AHPTOPSIS model was utilized to evaluate the performance of the athletes, and the scores of the athletes were comprehensively considered from multiple perspectives to improve the comprehensiveness and accuracy of the evaluation. Then, the Kruskal-Wallis H-test model was applied to evaluate the correlation between momentum and score, which further revealed the regularity during the game. Finally, the GABP neural network model was utilized to predict inflection points to help coaches and players better respond to changes in the game. The model proposed in this paper is of great significance for improving the analysis and prediction of tennis matches, and it can also help coaches and players better understand the match process so as to develop more effective tactical strategies.

References

[1] Den Hartigh, Ruud JR, and Christophe Gernigon. "Time-out! How psychological momentum builds up and breaks down in table tennis." Journal of sports sciences 36.23 (2018): 2732-2737.

[2] Gorsuch, Richard L. Factor analysis: Classic edition. Routledge, 2014.

[3] Tiwari, Varun, Prashant Kumar Jain, and Puneet Tandon. "An integrated Shannon entropy and TOPSIS for product design concept evaluation based on bijective soft set." Journal of Intelligent Manufacturing 30 (2019): 1645-1658.

[4] MacFarland, Thomas W., et al. "Kruskal–Wallis H-test for oneway analysis of variance (ANOVA) by ranks." Introduction to nonparametric statistics for the biological sciences using R (2016): 177-211.

[5] Tongle, Xu, Wang Yingbo, and Chen Kang. "Tailings saturation line prediction based on genetic algorithm and BP neural network." Journal of Intelligent & Fuzzy Systems 30.4 (2016): 1947-1955.

[6] Wang, Lihua, et al. "Optimal parameters selection of back propagation algorithm in the feedforward neural network." Engineering Analysis with Boundary Elements 151 (2023): 575-596.