Evaluating the role of momentum in tennis based on neural network models

Ruochen Zhang, Tianrui Dong*, Yujie Tang

School of electrical engineering, Northeast electric power university, Jilin, China, 132012

*Corresponding author: d834950869@outlook.com

Abstract. The research on the evaluation mechanism and prediction model of tennis player performance is of great significance and value for improving the quality of matches and cultivating excellent athletes. Firstly, by establishing a scientific and reasonable evaluation mechanism, the technical, physical, and psychological performance of athletes in competitions can be objectively analyzed, providing a basis for developing personalized training plans. Firstly, this article selects five parameters representing the stage performance status of athletes and extracts parameter data. Use clustering algorithm (K-means) to classify data. Based on the calculation of weights for different parameters, a score evaluation formula was established. It also provides indicators for the range of athlete evaluation scores, so that status data can be updated to obtain evaluation values at different times. It also provides a momentum evaluation score calculation function. Based on the calculation function, the total integrated momentum of 100 matches was simulated, combined with the logistic regression algorithm, and the results were compared. It has been found that the use of momentum accumulation evaluation mechanism can predict the success probability of players. Meanwhile, Monte Carlo simulation methods were used to simulate the number of player games under random probabilities. It was found that in 100 games, the performance of players fluctuated and contradicted the actual situation. Therefore, the momentum evaluation mechanism is of great significance for predicting the progress of athletes in competitions.

Keywords: Evaluation Mechanism Model, Monte Carlo Stochastic Simulation, Neural Network.

1. Introduction

In this game and other matches, this article needs to develop a specific tactical system for the players. Therefore, analyzing the situation of players in the game is very important, and establishing a comprehensive model to evaluate the performance of players is of great value[1-2]. Therefore, this article establishes a player performance analysis model based on representative matches and tests the model to predict player performance within a certain range, and provides guidance and suggestions for coaches to formulate tactics[3]. Tennis players will experience various states during matches, including focus, fatigue, and excitement. By analyzing the hitting speed, accuracy, and movement trajectory of athletes, their physical and psychological conditions can be inferred[4]. Establishing a reliable competition prediction model requires comprehensive consideration of factors such as athlete history, opponent competition, and competition environment. Using machine learning algorithms combined with large amounts of data to train models can improve prediction accuracy[5]. This not only helps fans better understand the trend of the game, but also provides coaches with more tactics.

2. Establishment of indicator system model

2.1. Data visualization

This article will draw a winning or losing score curve for two contestants as shown in Fig1. The statistics of the serving times of two players are shown in Fig2.

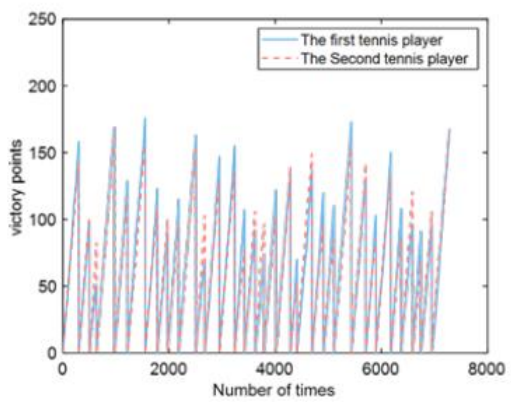


Fig 1: The victory points of two contestants over time

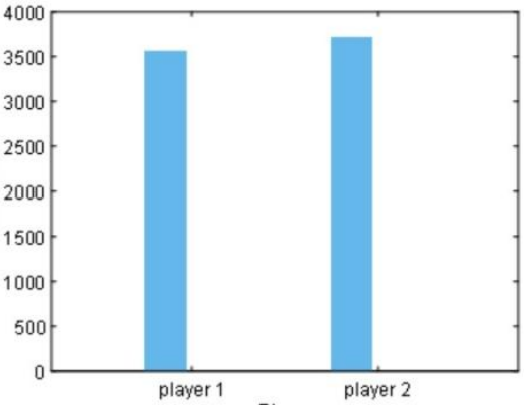


Fig 2: The number of times two players serve

From Figures 1 and 2, it can be seen that the first player has a greater overall chance of winning. Although the second player serves more, they have a higher overall number of losses.

2.2. Data normalization

Since the measurement units of various indicators are not uniform,the data should be unified and standardized before it can be used to calculate comprehensive indicators positive indicators and negative indicators have different meanings[6].

P is a positive index,Let x be the raw data and y be the/th indicator of the I-th state after normalization. Then y; As shown below:

$$p_{ij} = \frac{x_{ij} - \min_i x_{ij}}{\max_i x_{ij} - \min_i x_{ij}} \tag{1}$$

Where $\min_i x_{ij} = \min\{x_{1j}, \cdots, x_{nj}\}$ and $\max_i x_{ij} = \max\{x_{1j}, \cdots, x_{nj}\}$,which mini x_{ij} is the smallest value of P_j, $\max_i x_{ij}$ is the largest value of P_j .

N_j is a negative index

$$p_{ij} = \frac{\max_i x_{ij} - x_{ij}}{\max_i x_{ij} - \min_i x_{ij}} \tag{2}$$

Where $\max_i x_{ij}$ is the smallest value of N_j ,and $\max_i x_{ij}$ is the largest value of N_j .

2.3. Weighting Models of Evaluation Indicators

In order to establish an evaluation mechanism for the impact of the first serve on winning or losing, it is necessary to perform weighted calculations on these indicators. Therefore, there are two weighted models used to calculate weight vectors. One method is entropy weighting. To avoid repetition and compensate for the shortcomings of the EW method, the second method is the power averaging operator method.

2.3.1 Weighted model based on power averaging operator method

The Power Average (PA) operator method uses power operations to provide an aggregation operator, which allows parameter values to support each other during the aggregation process.

According to Table 1, this article obtained this matrix

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{pmatrix} \tag{3}$$

Each row of t e matrix represents all indicators of a country.

h

Then perform the following three steps.

The normalization formula is determined as:

$$x_{ij} = \frac{x_{ij}}{\max_j x_{ij}} \tag{4}$$

Quantify the support relationship xi between two indicators; The relationship function with Xu is defined as follows.

$$R(x_{ij}, x_{ik}) = Ke^{-\alpha(x_{ij}-x_{ik})^2} \tag{5}$$

Where K ∈ [0,1] and α≥ 0.

The weight vector is given by the following formula.

$$W = Pa(X_1, X_2, \cdots X_m) = \frac{\sum_{i=1}^{m}((1+S(X_i))X_i)}{\sum_{i=1}^{m}(1+S(X_i))} \tag{6}$$

Where

$$S(X_i) = \sum_{h\neq i} R(X_i, X_h) \tag{7}$$

Through this model, this article can obtain a more accurate weight vector W. Because the PA operator can reflect the mutual support within the data.

2.4. Comprehensive evaluation model

The two evaluation functions are as follows, For evaluation of the performance of tennis players,the paper have

$$F = w_1 ASTTM + w_2 TTWP + w_3 TNTWE + w_4 TCSVP + w_5 TNTTPL \tag{8}$$

Overall evaluation of tennis players based on a closer assessment of their performance status,we have

$$F' = \sum_{i=1}^{5} w_i' y_i \tag{9}$$

Where

$$W' = (w_1', w_2', w_3', w_4', w_5') = (0.325, 0.175, 0.051, 0{,}159, 0.276) \tag{10}$$

2.5. Metric of Performance Degree for Tennis Players

2.5.1 K-means clustering algorithm

K-means clustering algorithm is an iterative solution clustering analysis algorithm.Steps are as follows.Randomly select k elements from the index system.Let them be the centers of the k classes respectively,Calculate the dissimilarity of the remaining elements to the k cluster centers,and classify these elements into the clusters with the lowest dissimilarity,According to the clustering results,recalculate the centers of the k clusters respectively.The calculation method is to take the arithmetic average of the respective dimensions of all ele- ments in the cluster.All elements in the indexing system have been re clustered based on the new center, repeating the previous steps until the center of each cluster remains largely unchanged[7].

2.5.2 The standard of performance degree (PD)

The total score of each tennis player can be obtained from the above operations.Using the K-means clustering algorithm to cluster the scores of all tennis segments,the standard for the per- formance of tennis players is obtained.Meanwhile,a comprehensive standard was obtained using this method.The performance standards are shown in Fig 3.

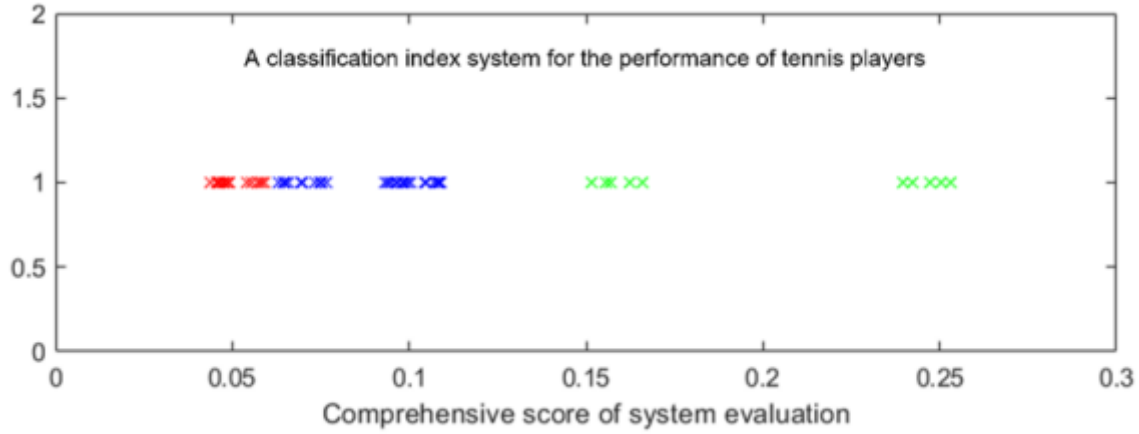


Fig 3: The results of K-means of two systems

The evaluation level of the performance of each tennis player is shown in Fig 4.We use the model in 4.3 to obtain any nation's score.If the score is between 0.135 and 0.3,the performance is good.If the score is between 0.07 and 0.135,the performance is not bad.If the score is between 0 and 0.07,the performance needs great improvement.



Fig 4: The threshold for evaluating the performance of tennis players

3. Assessing the Impact of Momentum on Games

This article describes the exercise situation of tennis players and quantitatively evaluates their performance at different time periods according to requirements. In the second question, we need to evaluate the impact of the word momentum on tennis matches and verify it.

3.1. Establish a momentum rating syste

This article establishes a scoring system for momentum. Since momentum refers to the energy accumulated by an athlete through winning or losing a game, it can be linked to the probability of an athlete winning over time in a normal game. Firstly, quantifying the momentum in the game can establish a model to evaluate the contribution of each scoring point to the player's momentum[8].

$$M_t = \alpha \sum_{i=1}^{t} P_i - \beta \sum_{i=1}^{t} L_i \tag{11}$$

Were M,represents the momentum quantification value at time point t,L;represents the negative contribution of the score lost at time point t to the momentum,a and β are weight parameters used to adjust the impact of scores and losses on momentum.

The overall results of this article are assigned two initial weights, as shown in Table 1 below.

Table 1: Weight assignment table

Weight indicators	Explanation			Initial value
α	The impact	of scoring on	momentum	0.63
β	The impact of	losing points	on momentum	0.37

At the same time,in order to characterize the actual competition trend,we calculate the change in momentum to make the results clearer,as shown below formula.

3.1.1 Probability calculation under the influence of momentum

The paper obtained the following scoring probability calculation formula based on the two players in the entire field

$$\delta_1 = \frac{M_{\text{t-1}}}{M_{\text{t-1}} + M_{\text{t-2}}} \tag{12}$$

Were δ_i represents the scoring probability of the first tennis player,M,-i represents the first tennis player's momentum quantification value at time point t,M,-2 represents the second tennis player's momentum quantification value at time point t,The scoring probability of another tennis player is calculated in the same way[9].

3.2. Momentum evaluation score and regression results

To eliminate the influence of randomness,the paper use a model to simulate the situation of 100 rounds in a single field and provide the results.In the calculation,the paper confirm that the following assumptions are valid.

the paper will not simulate the data for the first 10 games,but use the data provided in the attachment to obtain the scoring momentum of the two athletes in the first 10 games,The data for 90 matches is continuous.

3.2.1 logistic regression model with parameter correction

Under the momentum evaluation model,the momentum changes of two players in 100 matches are as follows Fig 5 and Fig 6 .

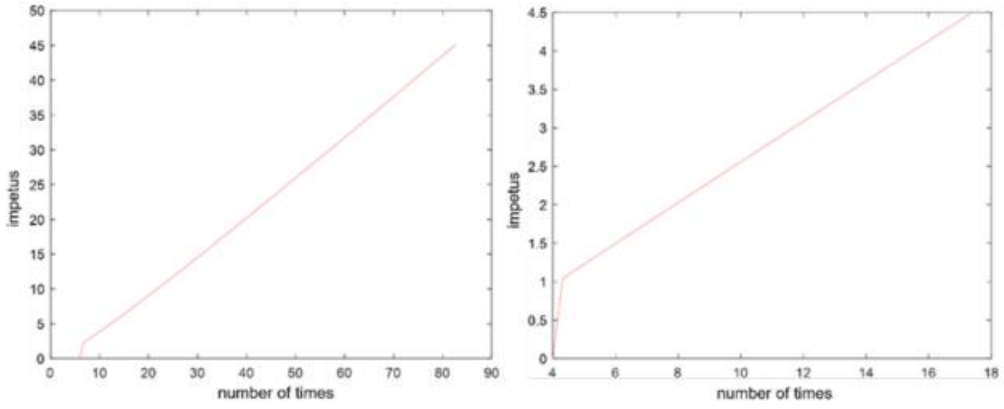


Fig 5: The momentum accumulation of two tennis players under momentum assessment

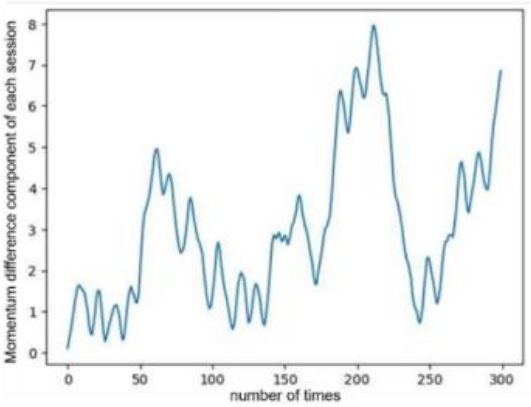


Fig 6: Momentum difference

From the graph,it can be seen that during the competition,when one of the tennis players wins more games,their own momentum accumulates more.the paper only draw a data graph of the number ofwins and momentum as shown in Fig 7. According to the attached data,the first player has more accumulated momentum in 100 games,so the winning chance of the first player in the final game is greater,which is consistent with the actual results as shown in Fig 8.

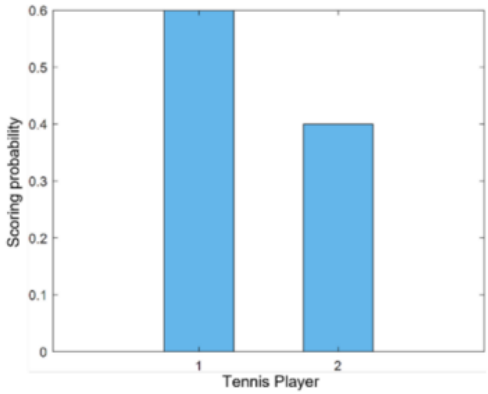


Fig 7: Revised overall scoring probability of two players

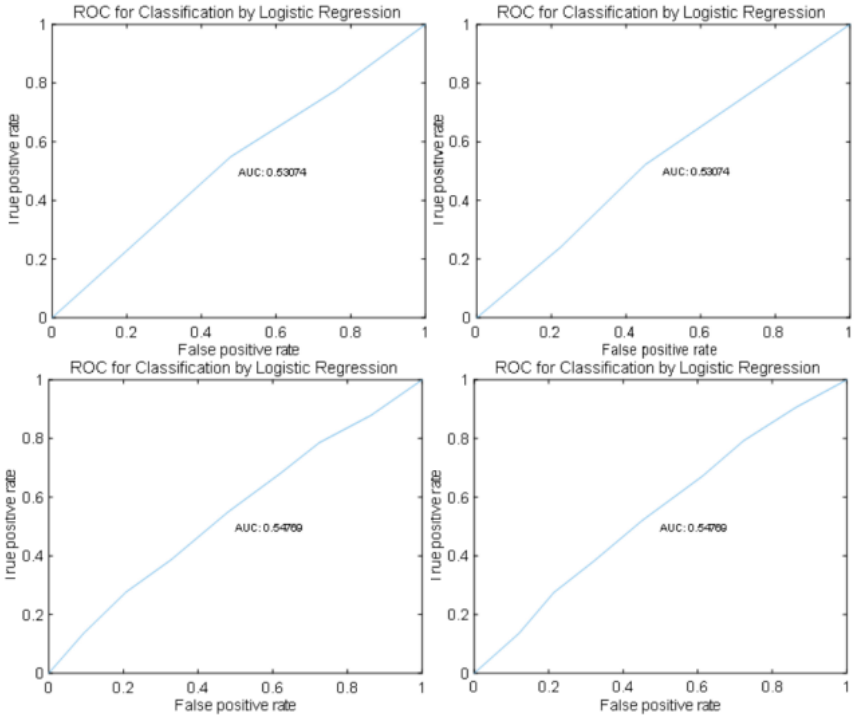


Fig 8: Logistic regression ROC results

At the same time,we selected the winning matches from the actual data for regression analysis,The overall error is small,and the results obtained are all the first tennis player to win,the paper also calculated the overall error of the regression,as shown in Fig 9.The overall error is small and the fluctuation range is small[10].

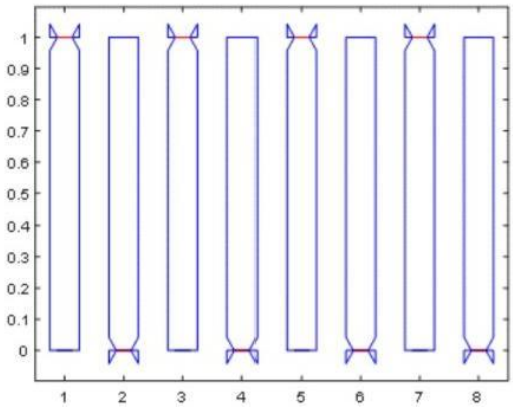


Fig 9: Logistic regression box diagram

3.2.2 Random simulation results

Firstly,for the model,the paper generate a random probability of winning or losing 100 matches according to the [0-1]distribution,and obtain the results shown in the following Fig 10.Since the sum of the winning probabilities of the two athletes is 1,the paper establish a linear relationship between them for random partitioning.In the simulation,the paper generated 1000 datasets and randomly sampled 100 of them as shown in Fig 11.

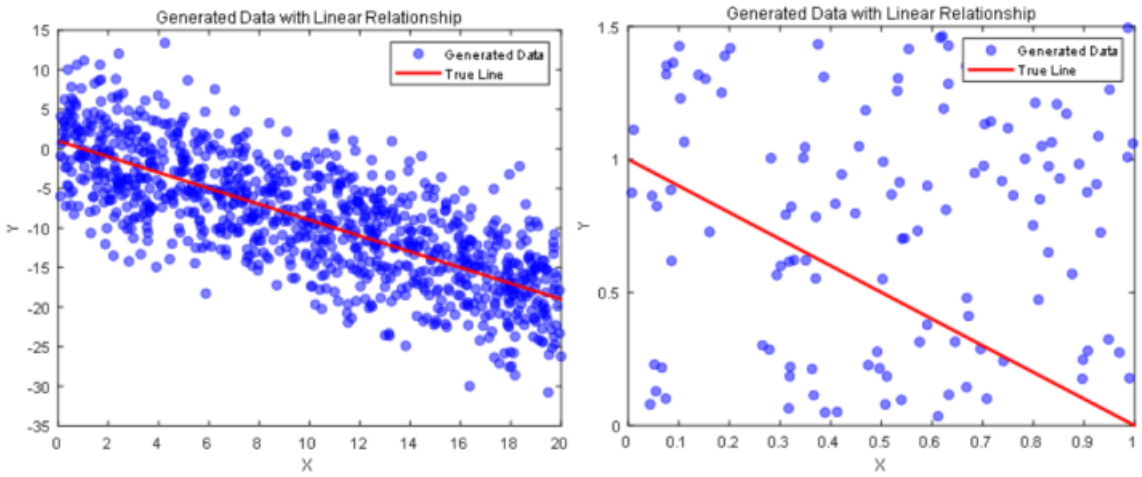


Fig 10: Monte Carlo distribution linear graph of the probability of two athletes winning or losing

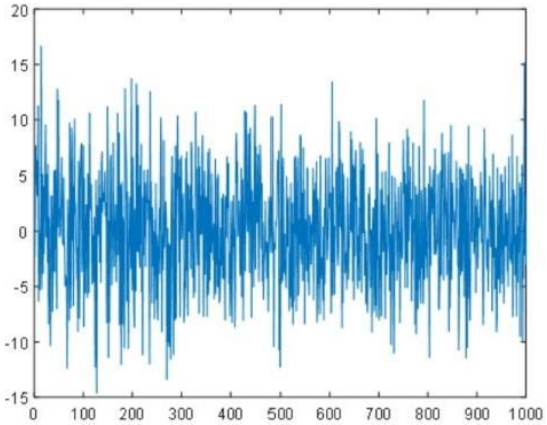


Fig11: Eliminating random normal distribution noise caused by human factors

4. Conclusions

The The entire study summarized the performance mechanism of tennis players under existing data, evaluated the competition trend of tennis players based on the performance mechanism,established a big data network prediction model,which can achieve good data prediction,and proposed creative suggestions for competition sports and coaches,At present,some progress has been made in the research on the performance and evaluation prediction mechanism of tennis players,but there are still some aspects worth paying attention to.Firstly,it is recommended to deepen research on multidimensional factors such as technology, physical fitness,and psychology in order to more comprehensively and accurately evaluate athlete performance.Secondly,it emphasizes the use of advanced technological means,such as motion sensors and artificial intelligence,for real-time monitoring and data collection of athletes,in order to improve the objectivity and accuracy of evaluation.

References

[1] Ye,Rouli,et al.“Nash equilibrium and tennis serve performance:a game theory analysis.” International Journal of Performance Analysis in Sport 23.6 (2023):515-526.

[2] Ulbricht,Alexander,etal.“Impact of fitness characteristics on tennis performance in elite junior tennis players."The journal of strength &conditioning research 30.4(2016):989-998.

[3] Knottenbert, William J., Demetris Spanias, and Agnieszka M. Madurska. A common opponent random model for predicting the results of professional tennis matches. Computer and Mathematics and Applications 64.12 (2012): 3820-3827

[4] Sarcević,Ana,Mihaela Vranić,and Damir Pintar.“A combinatorial approach in predicting the outcome of tennis matches.”International Journal of Applied Mathematics and Computer Science 31.3 (2021):525-538.

[5] Sarcevic,Ana,etal.“Predictive modeling of tennis matches:a review.”202245th Jubilee In- ternational Convention on Information,Communication and Electronic Technology (MIPRO). IEEE,2022.

[6] Gu,Wei,and Thomas L.Saaty."Predicting the outcome of a tennis tournament:Based on both data and judgments.”Journal of Systems Science and Systems Engineering 28 (2019): 317-343.

[7] Kovalchik,Stephanie Ann."Searching for the GOAT of tennis win prediction."Journal of Quantitative Analysis in Sports 12.3 (2016):127-138.

[8] Sipko, Michal, and William Knottenbbelt. "Machine learning for predicting professional tennis matches." MEng Compute Final Project, Imperial College London Phase 2 (2015).

[9] Lautenbach, Franziska et al. "Non automated routines before tennis matches: an intervention study", Journal of Applied Sports Psychology 27.2 (2015): 123-131.

[10] B ü hren, Christoph, and Philip J. Steinberg. The influence of psychological characteristics on continuous game performance: evidence from tennis court experiments. Journal of Economic Psychology 72 (2019): 12-29.