WEP Warwick Evans Publishing

Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 6
ISET 2024

Research on Momentum Prediction of Tennis Match Based on Scoring Model and Multilayer Perceptron

Jiaze Dong [*,#], Qilang Feng [#]

School of Electrical Engineering and Automation, Hefei University of Technology, Hefei, China, 230009

* Corresponding author: 2022210590@mail.hfut.edu.cn

[#] These authors contributed equally.

Abstract. During a tennis match, players may encounter various circumstances that can significantly influence their performance. This force, known as "momentum," can drastically alter the game's flow and ultimately determine the winner. Addressing the challenge of analyzing and predicting momentum shifts, this study applied the Entropy Weight TOPSIS method combined with fuzzy comprehensive evaluation to assess player performance across technical, physical, and psychological factors. These methods provided a comprehensive analysis of the players' performance metrics, revealing patterns linked to momentum changes. Following this, Pearson correlation analysis was conducted to identify key performance indicators strongly associated with match outcomes. These indicators were then integrated into a Multilayer Perceptron (MLP) model, which demonstrated high predictive accuracy, with an R-value of 0.868, effectively capturing momentum shifts in real-time. The study's findings offer valuable insights for improving strategic decision-making during matches and hold potential applications in other fields that require real-time predictive analysis.

Keywords: Entropy Weight TOPSIS Method, Fuzzy Comprehensive Evaluation, Multilayer Perceptron Prediction.

1. Introduction

Tennis, a globally popular sport, involves dynamic shifts in gameplay known as "momentum," which significantly influence match outcomes by reflecting the psychological and physical states of players [1]. While factors like fatigue, player psychology, and strategies have been studied regarding momentum shifts, a gap remains in quantitatively analyzing and predicting these shifts in real time—a crucial need for developing strategies to maintain or regain match control.

Previous studies have examined various aspects of tennis performance analysis, highlighting gaps that still need addressing. Cui Y. explored the multifactorial effects on game performance in Grand Slam tournaments, emphasizing the importance of match preparation and the impact of fatigue and strategy adaptation on momentum shifts [2]. Shrom et al. (2023) discussed the impact of lifestyle challenges and mental health on professional players, linking these factors to performance fluctuations [3]. Additionally, Lv X et al. developed momentum prediction models based on CatBoost regression and random forest algorithms, highlighting the role of potential energy in predicting player performance [4]. Zang Q et al. examined the prediction of match outcomes using entropy weight-TOPSIS combined with machine learning models, revealing the importance of a comprehensive analytical framework [5]. Ge Y. quantitatively modeled momentum in tennis using factor analysis, underscoring the challenges of predicting momentum shifts in competitive matches [6].

Building on these findings, the research aims to integrate and advance the analysis of momentum using the Entropy Weight TOPSIS method combined with fuzzy comprehensive evaluation [7]. This approach will consider technical, physical, psychological, and situational factors [8]. Additionally, this article uses an MLP model, enhanced by Pearson correlation analysis, to predict momentum shifts in real time, offering a more precise tool for coaches and players [9]. By combining these advanced techniques with traditional methods, this article seeks to bridge existing gaps in tennis analytics, providing a comprehensive model to enhance strategic planning and in-game decision-making for elite players and their coaching teams. Data source: 2024 MCM/ICM Problem C attachment.

2. The basic fundamentals of Multi-Criteria Evaluation Models

2.1. Topsis model based on entropy weight method:

First, we utilize the Entropy Weighting Method (EWM) to calculate the weights of the final-level indicators. This method provides higher credibility and accuracy compared to subjective weighting.

We begin by normalizing the indicators at the final level. Let x_{ij} denote the jth indicator of the ith evaluation object, and its normalized value be denoted as x_{ij}'. The normalization process is carried out using the following formulas:

If the indicator is a negative direction indicator:

$$x_{ij}'=\frac{\max(x_j)-x_{ij}}{\max(x_j)-\min(x_j)} \tag{1}$$

If the indicator is a positive direction indicator:

$$x_{ij}'=\frac{x_{ij}-\min(x_j)}{\max(x_j)-\min(x_j)} \tag{2}$$

Next, we calculate the entropy value e_j for each indicator:

$$e_j=-k\sum_{i=1}^{n} p_{ij}\ln(p_{ij}) \tag{3}$$

where $p_{ij}=\frac{x_{ij}'}{\sum_{i=1}^{n} x_{ij}'}$ and $k=\frac{1}{\ln(n)}$.

Finally, we calculate the weight w_j for each indicator:

$$w_j=\frac{1-e_j}{\sum_{i=1}^{m}(1-e_j)} \tag{4}$$

where e_j is the entropy value of the jth indicator, which is derived from the normalized values of the indicators.

After standardizing the indicators in a positive direction, define the maximum value:

$$X^{+}=(\max\{x_{11}',x_{21}',\cdots x_{n1}'\},\max\{x_{12}',x_{22}',\cdots x_{n2}'\},\cdots\max\{x_{1m}',x_{2m}',\cdots x_{nm}'\}) \tag{5}$$

Define minimum value:

$$X^{-}=(\min\{x_{11}',x_{21}',\cdots x_{n1}'\},\min\{x_{12}',x_{22}',\cdots x_{n2}'\},\cdots\min\{x_{1m}',x_{2m}',\cdots x_{nm}'\}) \tag{6}$$

Define the distance between the ith (i=1, 2, …, n) evaluation object and the maximum value:

$$D_i^+=\sqrt{\sum_{j=1}^{m} w_j*(X_i^+-x_{ij}')^2} \tag{7}$$

Define the distance between the ith (i=1, 2, … , n) evaluation object and the minimum value:

$$D_i^+=\sqrt{\sum_{j=1}^{m} w_j*(X_i^--x_{ij}')^2} \tag{8}$$

So, we can calculate the score of the ith (i=1, 2, … , n) evaluation object:

$$Scoure1_i=100*\frac{D_i^-}{D_i^++D_i^-} \tag{9}$$

It is obvious that $0 \leq Scoure1_i \leq 100$, and the larger the $Scoure1_i$, the smaller the Di+, that is, the closer it is to the maximum value.

2.2. Fuzzy comprehensive evaluation based on entropy weight method

Step 1: Data standardization.

Standardize the data processing to ensure that all indicators are on a comparable scale. This step is essential to ensure that the indicator does not dominate the evaluation due to its large scope. Typically, we use z-score standardization.

$$z=\frac{x-\mu}{\sigma} \tag{10}$$

Step 2: Calculation of the fuzzy membership function.

The fuzzy membership function is calculated, as shown in Formula 11.

$$\mu(x)=\begin{cases} 0 & z_{ij}\leq a_j \\ \frac{z_{ij}-a_j}{b_j-a_j} & a_j<z_{ij}\leq b_j \\ \frac{c_j-z_{ij}}{c_j-b_j} & b_j<z_{ij}\leq c_j \\ 0 & c_j<z_{ij} \end{cases} \tag{11}$$

where a_j, b_j, and c_j are three key points for the jth indicator, and x_{ij} is the standardized value of the ith evaluation object on the jth indicator. The function μ(x) calculates the degree to which the indicator value x_{ij} belongs to a fuzzy set. This value ranges from 0 to 1.

Step 3: Fuzzy evaluation function.

$$E_j=\sum_{i=1}^{n} w_j*\mu_{ij} \tag{12}$$

where E_i represents the comprehensive evaluation result, and μ_{ij} represents the output value of the membership function. w_j is the weight determined by the entropy weight method.

Step 4: Converting E_i to a 0-100 Scale.

We need to transform the comprehensive evaluation score Ei, which is originally in the range of [0, 1], to a scale of [0, 100].

$$Scoure2_i=E_i*100 \tag{13}$$

2.3. MLP

Multi-layer perceptron is a network structure based on a single-layer perceptron. The single hidden layer (n neurons) neural network is built in Figure 1.

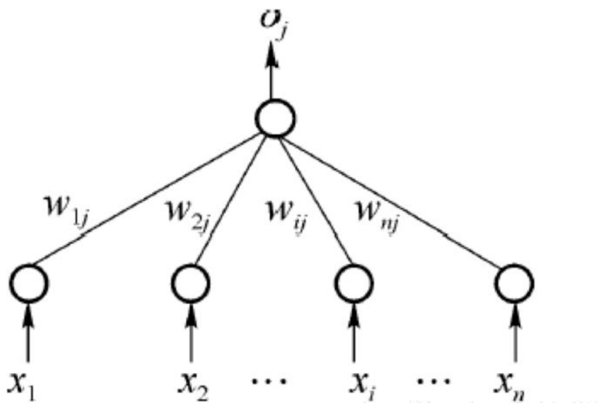


Figure 1. Single-Layer Perceptron Mode

The MLP is a forward-structured artificial neural network, including an input layer, output layer, and multiple hidden layers. The neural network diagram of t he 3-layer perceptron is shown in Figure 2.

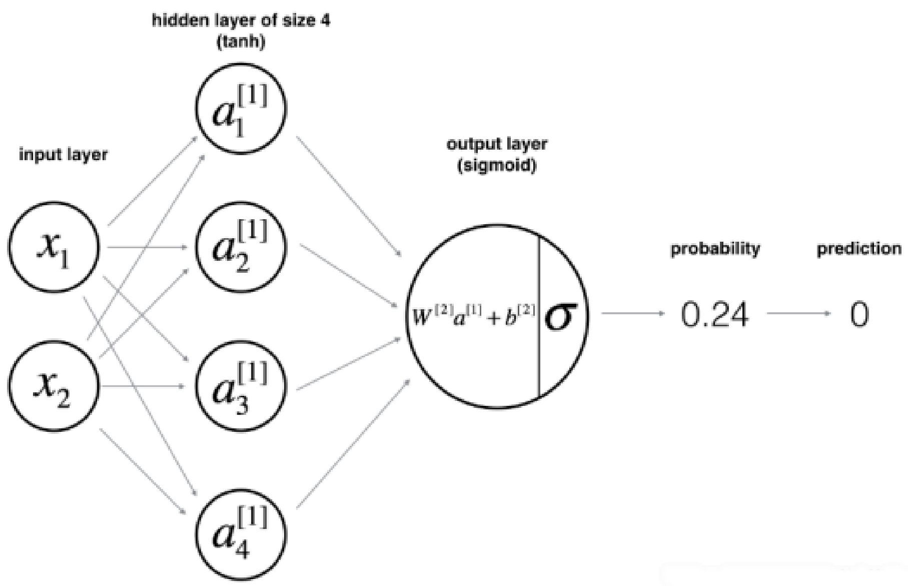


Figure 2. The Neural Network Diagram of the 3-Layer Perceptron

The activation function:

$$g(h)=\sigma(h)=\frac{1}{1+e^{-h}} \tag{14}$$

The multilayer perceptron formula:

$$Y_j = g\left(\sum_i w_{ij}^k * x_i^{k-1}\right) \quad (15)$$

where Y_j represents the $_{jth}$ component of the output layer; x_i^{k-1} is the input to the $_{ith}$ neuron in the k-1 hidden layer; w_{ij}^k are the connection weight between the $_{ith}$ component in the k-1 hidden layer and the $_{jth}$ neuron in the k_{th} hidden layer.

3. Results

3.1. TOPSIS scoring results and analysis based on entropy weight method

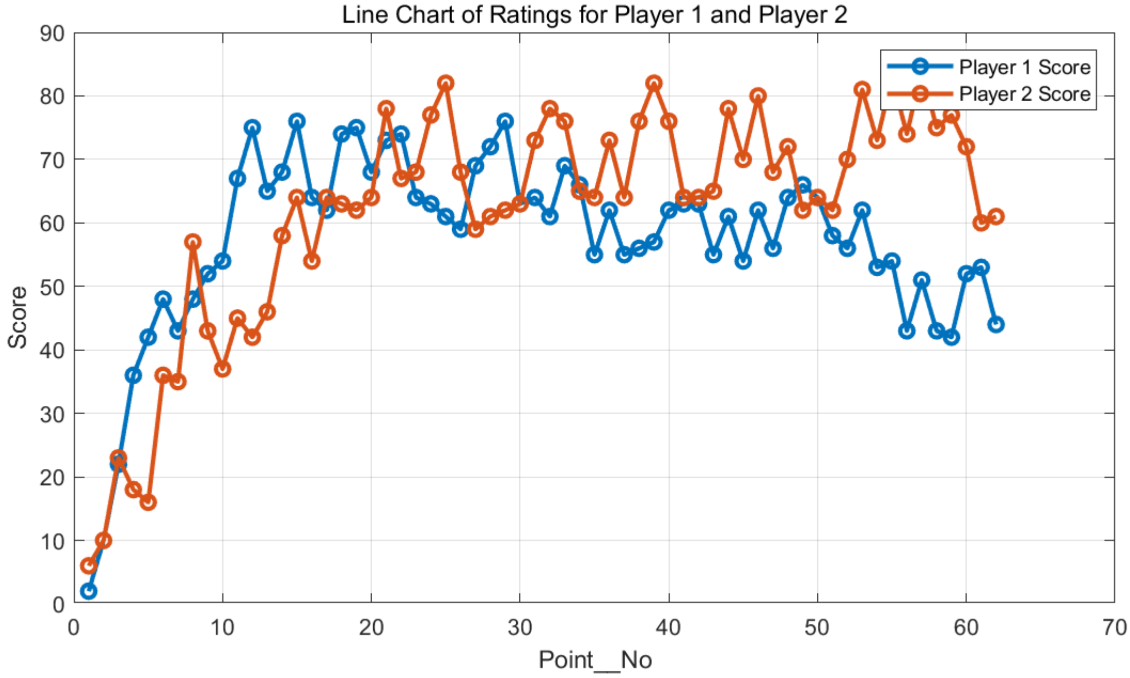


Figure 3. TOPSIS scoring based on entropy weight method

At the beginning of this section, we utilized the Entropy Weight TOPSIS method to assess the players' performances. This approach allowed us to derive the weighted scores of the players' performances across various indicators [10]. As shown in Figure 3, the results indicate that in the first 23 rounds, player 1 generally received higher scores, suggesting that he adapted to the game conditions more rapidly. However, between rounds 24 and 63, player 2's scores were relatively higher, indicating a stronger overall performance during this period. Notably, player 2 achieved the highest score in the 56th match. This was a significant point in the game, where despite player 1 making a serve mistake, player 2, acting as the receiver, managed to win the match through four successful returns.

The scoring trends of player 1 and player 2 during this period showed largely opposite patterns, suggesting that their game states influenced each other. Player 2's ability to capitalize on Player 1's mistakes and maintain high-performance levels during critical rounds is indicative of strong strategic adaptation. Conversely, player 1's fluctuating scores suggest periods of adjustment and possibly reduced performance due to player 2's increased pressure.

As the game progressed, both players' performances fluctuated, reflecting the momentum shifts typical in competitive matches. These fluctuations are consistent with common match dynamics and are visually represented in Figure 3. Additionally, the trends highlight how each player's ability to adapt and respond to the other's strategies significantly impacts the overall game outcome. This interplay underscores the competitive nature of the match and emphasizes the importance of tactical adjustments and resilience throughout the game.

3.2. Fuzzy comprehensive evaluation results and analysis based on entropy weight method

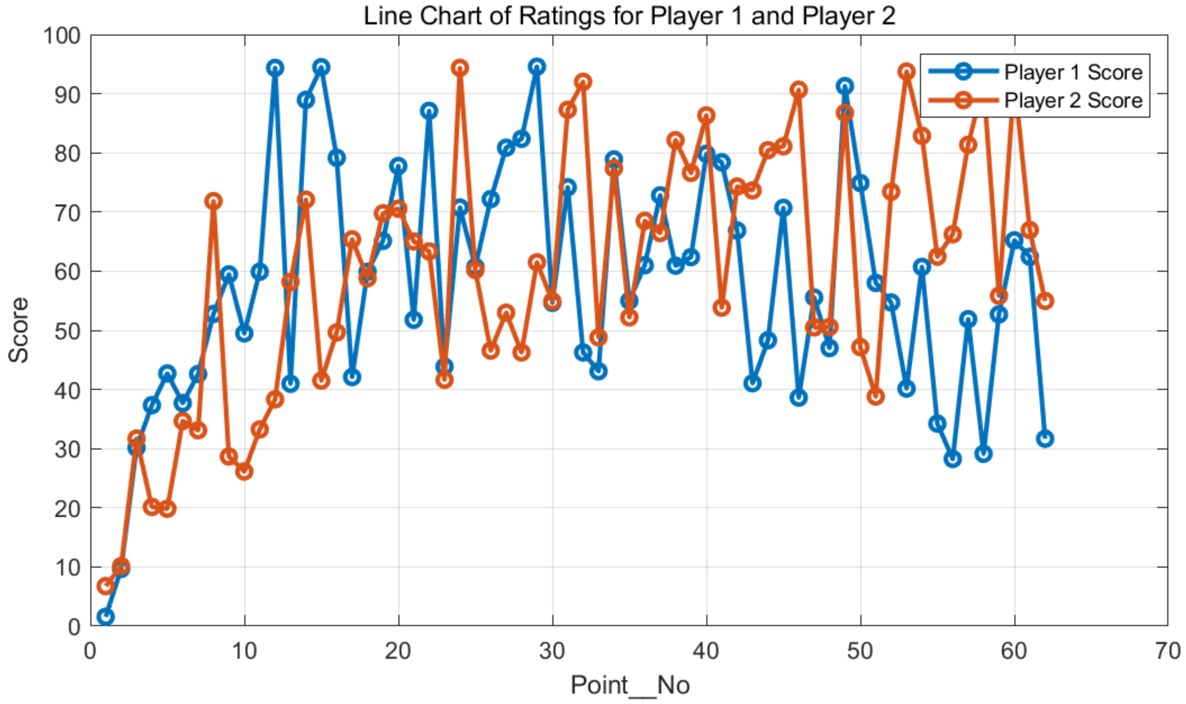


Figure 4. Fuzzy comprehensive scoring based on the entropy weight method

The fuzzy comprehensive evaluation method, combined with the entropy weight approach, provided a detailed assessment of the players' performances over time. By calculating the entropy weights, we objectively determined the significance of each performance indicator, which in turn influenced the overall fuzzy evaluation scores.

As illustrated in Figure 4, the evaluation results highlight that Player 1 had a strong start but began to lose momentum after round 23, where Player 2's performance became more dominant. This shift is particularly evident from round 24 to round 63, where player 2 consistently outperformed player 1.

The similarity in results between the fuzzy comprehensive evaluation and the TOPSIS method based on entropy weights demonstrates the robustness and reliability of our model. Both methods independently identified the same key trends and shifts in player performance, which adds credibility to the analysis. This consistency across different evaluative approaches suggests that our model effectively captures the dynamics of player performance throughout the game, making it a trustworthy tool for performance analysis in similar contexts.

3.3. Establishment and Analysis of Multi-Layer Perceptron

Before constructing the MLP model, we conducted Pearson correlation analysis to identify the key performance indicators most strongly correlated with match outcomes [11].

$$r=\frac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\overline{X})^2} * \sqrt{\sum_{i=1}^{n}(Y_i-\overline{Y})^2}} \quad (16)$$

In the formula, r represents the correlation coefficient, n is the number of samples, X_i and Y_i represent the two sets of attribute values of the ith sample respectively. When r=1, X and Y are completely related, indicating that there is a linear function relationship between them; when r>0.8, it is highly related; when r<0.3, it is lowly related; otherwise, it is moderately related. The correlation coefficients are calculated by SASS. For this study, we focused on identifying indicators with high positive correlations (r>0.8) with the match outcomes, as these are most likely to contribute to the accuracy of our predictive model.

Pearson correlation coefficients were calculated to determine the linear relationships between different indicators and the final performance results. Only indicators with a high correlation coefficient ($r > 0.8$) were selected as inputs for the MLP model. This preliminary step was crucial in refining the model's predictive capabilities by focusing on the most impactful factors.

After identifying the significant indicators, we applied the MLP model to predict momentum shifts within the match. The MLP model, trained on the identified key indicators, was designed to capture the non-linear relationships and complex dynamics that influence momentum in tennis.

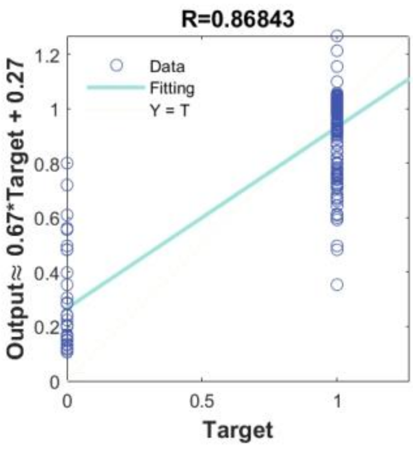
Table. 1. Comparison of Actual and Predicted Values

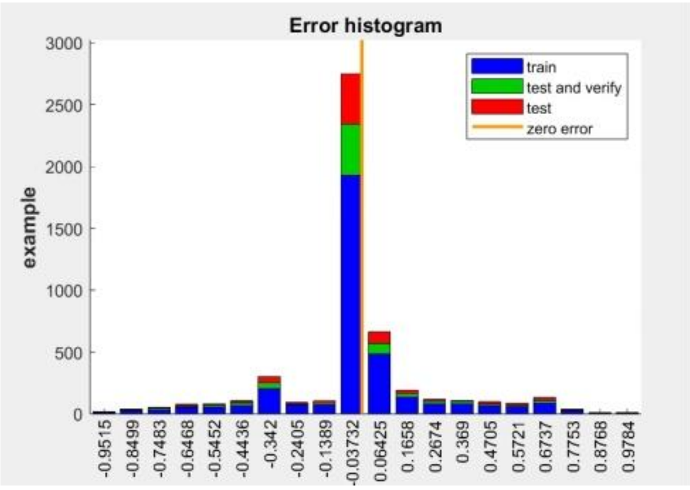
Serial Number	Actual Value	Predicted Value
503	1	0.9333
504	1	0.8123
505	0	0.0285
506	0	0.0056
507	0	0.0276

The results of the analysis, as shown in Table. 1, indicates that the MLP model accurately predicts the momentum shifts in the match. The table includes three columns: the Serial Number, which represents the match identification number; the Actual Value, which shows whether Player 1 won the match (1 for a win, 0 for a loss); and the Predicted Value, which indicates the probability of Player 1 winning as predicted by the model.

For instance, Serial Number 503 shows an Actual Value of 1, indicating that Player 1 won, with a Predicted Value of 0.9333, suggesting a high probability of winning according to the model. Similarly, Serial Number 505 has an Actual Value of 0, indicating a loss for Player 1, with a Predicted Value of 0.0285, closely aligning with the actual outcome.

This analysis highlights that as one player gains momentum, the other tends to lose it, creating a dynamic and fluctuating match environment. The MLP model effectively captures these interactions, offering valuable insights for real-time decision-making in tennis matches.





(A) (B)

Figure 5. Coefficient of Determination & Error Histogram

The comparison of the predicted outcomes with the actual match results is illustrated in Figure 5. The model achieved an R-value of 0.868, indicating a strong correlation between predicted and actual performance. The error distribution, primarily centered around 0.06, confirms the model's accuracy. These findings are presented in Figure A and supported by the Error Histogram (Figure B). The analysis confirms that the MLP model can reliably predict momentum shifts and offers valuable insights for real-time decision-making during matches.

In this paper, the Coefficient of Determination for the MLP model reaches 0.868, which is greater than 0.8. This suggests that the MLP model meets the predictive conditions, indicating a strong fit between the model predictions and the actual data. Additionally, the Error Histogram displays a normal distribution, with the majority of errors concentrated within 0.06. This indicates that the MLP training has yielded excellent results.

The predictive performance of the MLP model is further demonstrated by comparing the predicted binary classification results with the original data, as illustrated in the figure below. The blue color represents data where predictions match the actual outcomes, while the red color represents data where predictions do not match the actual outcomes. Among the matched data, validation confirms an accuracy of 87.63% of the total, indicating a high level of predictive efficacy.

4. Conclusions and outlooks

This study tackled the challenge of predicting momentum shifts in tennis matches by developing a model that combines the Entropy Weight TOPSIS method, fuzzy comprehensive evaluation, and an MLP model. The model's high R-value of 0.868 demonstrates its effectiveness in accurately identifying key momentum shifts, offering valuable insights for real-time decision-making in sports. These findings could be applied beyond tennis, potentially influencing strategies in other sports and fields requiring real-time predictions.

However, the study has limitations. The reliance on historical data limits the model's responsiveness to unexpected factors like injuries or weather changes. Incorporating real-time data from player wearables could improve accuracy. Additionally, expanding the model to include more performance indicators, such as psychological factors, could provide a more comprehensive analysis. Further optimization of the MLP model and validation across a broader range of matches would enhance its generalizability.

In conclusion, while this research marks significant progress in sports analytics, addressing these limitations and exploring further improvements will broaden its applicability and impact.

References

[1] Iván P, Adrián P, Daniel T, et al. Match analysis and probability of winning a point in elite men's singles tennis. [J].PloS one,2023,18(9):e0286076-e0286076.

[2] Cui Y. Exploring matches performance of elite tennis players: the multifactorial game-related effects in Grand Slams[D]. Ciencias, 2018.

[3] Shrom S J, Cumming J, Fenton S J. Lifestyle challenges and mental health of professional tennis players: an exploratory case study[J]. International Journal of Sport and Exercise Psychology, 2023, 21(6): 1070-1090.

[4] Lv X, Gu D, Liu X, et al. Momentum prediction models of tennis matches based on CatBoost regression and random forest algorithms[J]. Scientific Reports, 2024, 14(1): 1-17.

[5] Zang Q, Li Z, Hu J, et al. Prediction of winning and losing in tennis match based on entropy weight-TOPSIS and machine learning model[J]. Transactions on Computer Science and Intelligent Systems Research, 2024, 5: 1529-1535.

[6] Ge Y. Quantitative modelling of momentum in tennis based on factor analysis[J]. Transactions on Computer Science and Intelligent Systems Research, 2024, 5: 1204-1213.

[7] Luo Qian, Li Yongmei, Wang Tenghua, et al. Constructing an evaluation system of rational medication use indicators in clinical departments based on the improved entropy weight method combined with TOPSIS method [J]. Clinical rational use of medication, 2024, 17(15): 170-172+ 177.

[8] Christopher Y, Lyndell B, Dan D , et al. Understanding passing network characteristics and their link to match outcome in elite Netball.[J].Journal of sports sciences,2023,41(16):1-9.

[9] Malalur S S, Manry T M, Jesudhas P. Multiple optimal learning factors for the multi-layer perceptron[J].Neurocomputing,2015,1491490-1501.

[10] Xia Lei. Research on specialised physical characteristics and training strategies of competitive tennis [J]. Sports Science and Technology Literature Bulletin, 2024, 32(04): 91-94+201.

[11] Liu Y, Mu Y, Chen K, et al. Daily activity feature selection in smart homes based on pearson correlation coefficient[J]. Neural processing letters, 2020, 51: 1771-1787.