Prediction of potential energy on player scoring trend in tennis match based on Hopfield neural network model

Haodong Bian [1, *, #], Shijia Ni [1, #], Yuchen Liu [2, #]

[1] School of Electrical Engineering, China University of Mining and Technology, Xuzhou, China, 221116

[2] School of Mathematics, China University of Mining and Technology, Xuzhou, China, 221116

* Corresponding Author Email: 13884713503@163.com

[#]These authors contributed equally.

Abstract. In tennis, it is very important to catch the direction and trend of the match when the score happens. This paper aims to establish a mathematical model to predict the dynamic change of match progress and the appearance of match points, and explain the role of "momentum" in the match. By mining the real-time match data of 2023 Wimbledon Tennis Open, based on Hopfield neural network and entropy weight method, the technical level, psychological pressure and fatigue level of players in the match are selected as the influencing factors, and the mathematical model between the "momentum" of players and the change of the match flow is established, and the change of the match schedule can be predicted more accurately. The change of momentum may affect the motivation of players, and the choice of a reasonable strategy may be a magic weapon to win the match. It is hoped that this can help the pre-competition training of tennis players and promote the vigorous development of tennis events.

Keywords: Entropy Weight Method, Hopfield Neural Network, the Flow of Match.

1. Introduction

Tennis is a sport with a long history and is considered one of the most popular sports in modern times. It is widely popular worldwide and has a large number of professional and amateur athletes. Tennis events are a series of professional competitions held worldwide in which players score points by returning a ball served by an opponent on a court. The ball must go over the net and land in the opponent's court, while not going out of bounds. In addition to the field conditions, weather conditions and other external objective factors, the player's skill level, psychological pressure, health level, tactics and strategies, motivation and emotional state and other factors to a large extent determine the outcome of the game.

In previous researches on the prediction of tennis match results, there have been a lot of mature models, such as the review of models in literature [1,2]. Many research papers have mostly used hierarchical Markov model, logistic regression model, random forest model and other machine learning models, such as the research content in literature [3-5]. However, the model used usually requires the input of parameters of each player's serving ability, which are usually calculated from the player's scoring records in historical matches, and the proof can be reflected in the literature [6]. In previous studies, penalty points, cumulative points, average points, winners, venues, gender and other factors were often used as model parameters for prediction, and these factors had a certain impact on the performance of tennis players, as can be seen from the study of the above factors in literature [7-11]. These results provide reference for coaches and athletes to formulate competition strategies, which can be more appropriate in different competition situations.

However, this kind of research cannot well solve the interference factors brought by the intensity of the game or the individual state, such as the physical and energy consumption of the athlete in a certain stage, the average score over a period of time, the psychological pressure of serving failure, etc. These factors often affect the competitive level of players. Moreover, most predictions in such studies are not in real time. In order to capture the flow of the game and determine the performance of the players, this paper tries to establish a comprehensive evaluation model of the players'

performance level from three aspects: technical level, psychological pressure and fatigue level. First of all, according to the characteristics of the players in the game, extract the current winning rate of the players, the probability of the opponent serving to win the match point, the probability of sending the untouchable ball winner, the average score and high score rate, the double fault probability of the serve, the average distance of the run point. Then, the entropy weight method is used to assign weights to each index, and the obtained weights are standardized, and a mathematical model is established to predict the real-time performance of the players' scores.

2. The basic fundamental of entropy weight method and Hopfield neural network.

2.1. Fuzzy comprehensive evaluation model based on entropy weight method

This article establishes a fuzzy comprehensive evaluation model based on entropy weight method. The entropy weight method is used to determine the weights of various evaluation factors. This method avoids the bias caused by subjective judgment, making weight allocation more scientific and reasonable.

(1) Firstly, the fuzzy comprehensive evaluation deals with the de-dimensionalization of each index. There are ρ indicators: $X_1, X_2, ..., X_\rho$ among them $X_i = \{x_1, x_2, ..., x_m\}$ .Suppose the value after the standardization of the indicator data is $Y_1, Y_2, ..., Y_n$

The positive indicator is:

$$Y_{ij} = \frac{X_{ij} - \min(X_i)}{\max(X_i) - \min(X_i)} \tag{1}$$

The negative indicator is:

$$Y_{ij} = \frac{\min(X_i) - X_{ij}}{\max(X_i) - \min(X_i)} \tag{2}$$

(2) Calculate the ratio of each indicator under each scheme

Assuming each level has m secondary indicators and n data obtained. Under the same indicator, calculate the proportion of each value to the total value, using the following formula:

$$p_{ij} = \frac{Y_{ij}}{\sum_{i=1}^{n} Y_{ij}} (i = 1, 2, ..., m) \tag{3}$$

(3) Calculate the information entropy of each indicator

According to the definition of information entropy, the information entropy of a set of data is:

$$E_j = -\ln(n)^{-1} \sum_{i=1}^{n} p_{ij} \ln p_{ij} \tag{4}$$

If $p_{ij} = 1,\ E_{ij} = 0$.

Determine the weight of each indicator

According to the formula of calculating information entropy, calculate the information entropy $E_1, E_2, ..., E_m$ for each indicator.

Step 1: Calculate the weights of each indicator through information entropy:

$$\omega_i = \frac{1 - E_j}{k - \sum E_j} \tag{5}$$

Where k is the number of indicators? $k = m$

Step 2: Calculate weights by calculating information redundancy:

$$D_j = 1 - E_j \tag{6}$$

Then calculate the weight of the indicators:

$$\omega_i = \frac{D_j}{\sum_{j=1}^{m} D_j} \tag{7}$$

Step 3: Calculate the comprehensive score for each proposal finally:

$$s_i = \sum_{j=1}^{m} \omega_j x_{ij} \tag{8}$$

2.2. The structure of Hopfield network

The Hopfield theorem describes the properties of stable states in Hopfield networks. This theorem states that under appropriate conditions, the iterative process of a Hopfield network will converge to a local minimum point, which corresponds to a stable state of the network's energy function. If the connection weight matrix of the network is symmetrical and the energy function satisfies the Lyapunov stability condition, then regardless of the initial state, as long as the system is large enough, the network will eventually reach an attractor state, that is, a memory state. Figure 1 shows the structure of Hopfield neural network.

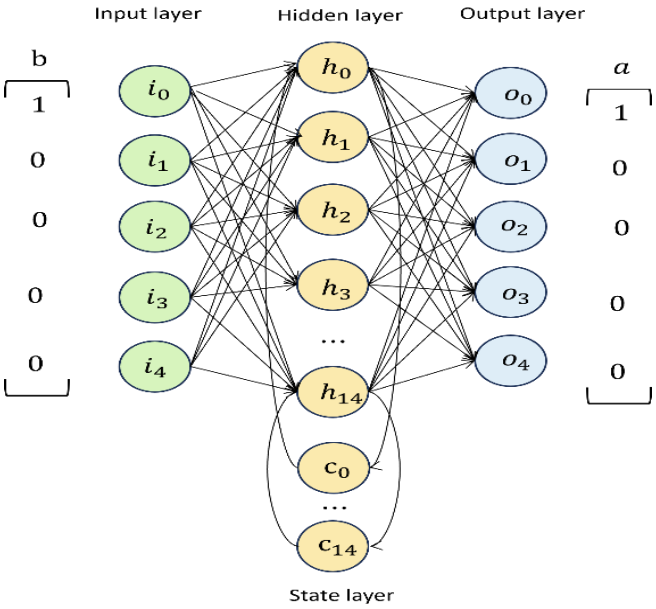


Figure 1. The structure of Hopfield network.

Assuming that Hopfield contains n neurons, the interconnection weight from neuron j input to neuron i input is represented as w_{ij}, and the input of the transfer function of any neuron i at time t is represented as $u_i(t)$, the output at time t is represented as $v_j(t)$, and the deviation value is b_i, then the input of the transfer function at time t can be represented as:

$$u_i(t) = \sum_{j=1}^{n} w_{ij} v_j(t) + b_i \tag{9}$$

The output state of the corresponding neuron is:

$$v_i(t+1) = f(u_i(t)) \tag{10}$$

The transfer function f can take the form of a step function or a sign function. If the sign function is taken, the neuron output $v_i(t+1)$ of the Hopfield network takes a discrete value of 1 or -1, that is:

$$v_i(t+1) = \begin{cases} 1, \sum_{j=1}^{n} w_{ij} v_j(t) + b_i \geq 0 \\ -1, \sum_{j=1}^{n} w_{ij} v_j(t) + b \leq 0 \end{cases} \tag{11}$$

The Hopfield network operates according to dynamic rules, and its working process is the evolution of states. It evolves from the initial state in the direction of decreasing Lyapunov energy function until it reaches a stable state and obtains a stable network output. The operation of Hopfield network includes two working modes: serial and parallel.

(1) Initialize the network.

(2) Randomly select a neuron from the network.

(3) Calculate the input $u_i(t)$ of the neuron according to the formula.

(4) calculate the output of the neuron $v_i(t+1)$ according to the formula, while keeping the outputs of other neurons in the network unchanged.

(5) Determine whether the network has reached a stable state. If it has reached a stable state or meets the given conditions, end the training. Otherwise, proceed to Part 2 to continue execution.

The stability of the feedback network is a very important indicator. If the state of the network no longer changes after a certain moment, it is said to have reached a stable state. At this point, the network output of the next moment is equal to the network output of the current moment, expressed as:

$$v_i(t+1) = v_i(t) \tag{12}$$

Sufficient conditions for the stability of Hopfield networks. To make the Hopfield network exist in a stable state, symmetric connections are drawn from the network, which satisfies both:

$$w_{ij} = w_{ji} \tag{13}$$

And the neurons themselves have no connections, that is:

$$w_{ij} = 0 \tag{14}$$

The Hopfield network with symmetric link weights is shown in the figure 2.

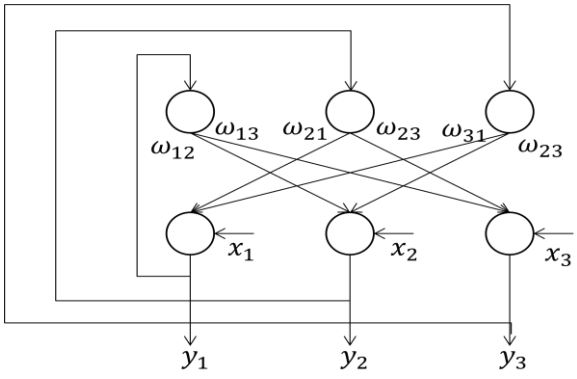


Figure 2. Schematic diagram of Hopfield network with zero diagonal symmetric weight matrix.

Due to its feedback, the Hopfield network is a nonlinear dynamical system that operates in a nonlinear dynamical manner, evolving in the direction of decreasing energy function until reaching a steady state. The energy function of its network is a Lyapunov function, defined as:

$$E = -\frac{1}{2}\sum_{i=1, i\neq j}^{n}\sum_{j=1, j\neq i}^{n} w_{ij} v_i v_j + \sum_{i=1}^{n} b_i v_i \tag{15}$$

Among them, w_{ij} is the interconnection weight of neurons i and j. Under the above parameter conditions, the energy of the Lyapunov function continuously decreases during the operation of the network, and finally reaches a stable equilibrium state.

3. Results

3.1. Establish a model to capture the flow of match

First, we establish an indicator evaluation system. The selection of indicators is made in the following three areas:

1) Technology level

- Number of wins W_{ik}
- Stability of win H_{ik}
- Average of score $\overline{C}_{ik}$
- Percentage of high points E_i
- Average percentage of points won F_{ik}
- Stability of performance G_{ik}
- Probability of win J_{ik}
- Probability of an untouchable wining serve L_{ik}
- Probability of an untouchable winning shot S_{ik}
- Probability of approaching the net P_{ik}

2) Psychological pressure

- Probability of winning a match point when the opponent serves V_{ik}
- Probability of winning the point when the opponent serves U_{ik}
- Probability of losing the point when the opponent serves I_{ik}
- Probability of two missed serves R_{ik}

3) Level of fatigue

- Average time to win T_{ik}
- Average distance ran during the point $\overline{O}_{ik}$
- Stability of distance ran during the point A_{ik}
- Total points won at the net Q_{ik}

Then, calculate the entropy weight through the fuzzy comprehensive evaluation model based on entropy weight method.

Table 1. P1&P2Indicator weights.

Targets	Weight (player 1)	Targets	Weight (player 1)	Weight (player 2)
W_1	6.131%	P_1	0.857%	0.551%
H_1	0.024%	Q_1	15.701%	11.038%
$\overline{C_1}$	6.441%	V_1	19.977%	24.283%
E_1	5.073%	U_1	5.831%	8.072%
F_1	0.125%	I_1	5.741%	7.961%
G_1	0.013%	R_1	1.993%	2.234%
$\overline{O_1}$	1.426%	A_1	1.018%	1.068%
J_1	4.604%	T_1	0.757%	0.807%
L_1	16.996%	S_1	7.292%	3.281%

Through entropy weight method and fuzzy comprehensive evaluation model, the weight distribution table is shown in Table 1. The probability of winning match point when the opponent serves, the probability of winning the untouchable ball, the total score at the net, and the probability of winning the point when the opponent serves are the highest weighting factors. Player 1 performed best at 00:27:00, 2:26:30 and 3:49:13, scoring 60% of the points at 00:27:00, 15% higher than the opponent, accumulating 2 points in front of the net, 1 point higher than the opponent, and running a cumulative distance of 643 meters, 120 meters less than the opponent.

As shown in Figure 3, the trend of the curves represents the change of the player's performance level, the higher the peak, the better the performance level. The intersection of the two curves signals a shift in the momentum of the game and the emergence of a match point.

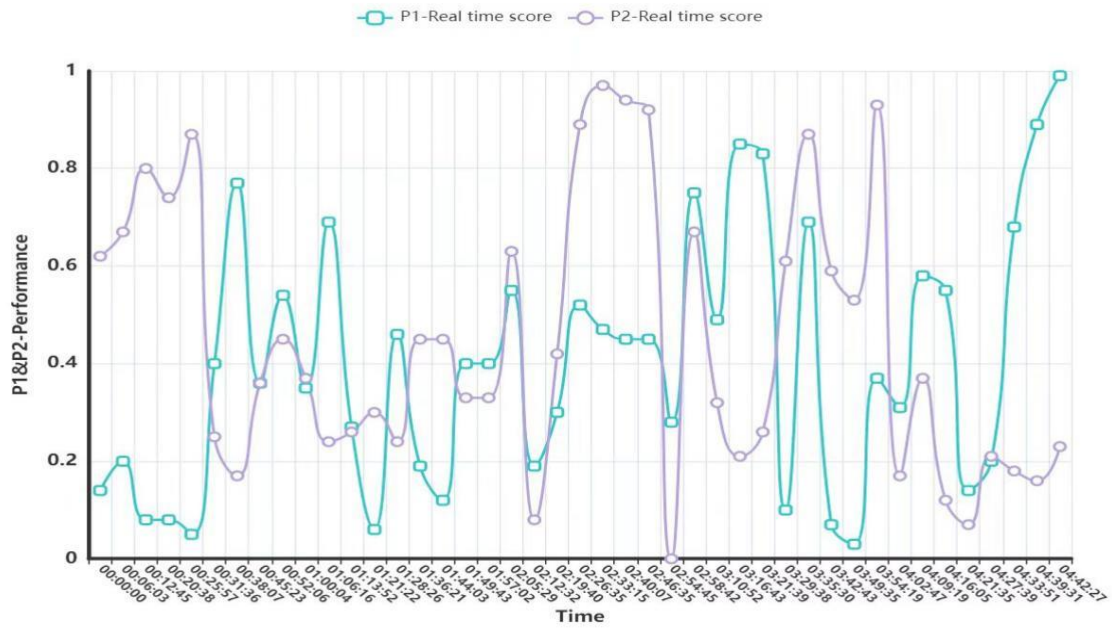


Figure 3. Visualization of 2023-wimbledon-1701.

3.2. Assessing the role of "momentum"

Considering that there are many factors affecting momentum, this paper adopts the following 11 indicators to quantify the momentum of players, and lists six sets of data, as shown in Table 2.

Table 2. 11 indicators to quantify the momentum of players.

Target	D_1	Average scores(X1)	0.00	0.00	5.00	7.50	6.00	6.67
	D_2	Number of consecutive wins(X2)	0	0	0	0	0	0
	D_3	Number of consecutive points(X3)	0	1	2	0	1	0
	... ...	... ...	... ...	... ...	... ...	... ...	... ...	... ...
	D_{11}	Percentage of high points(X11)	0.3	0.3	0.3	0.3	0.3	0.3
Quantification of momentum	Y	Whether or not to win the next game(Y)	1	1	0	1	0	0
Note: Due to space constraints, only part of the data for 2023-wimbledon-1701 player 1 is listed as a representation, see appendix for more details.								

In this paper, Pearson correlation was used to analyze the relationship between each index and momentum change. The result is shown in Figure 4, “the momentum indicator” and “the athlete's next game whether to win” has a strong correlation.

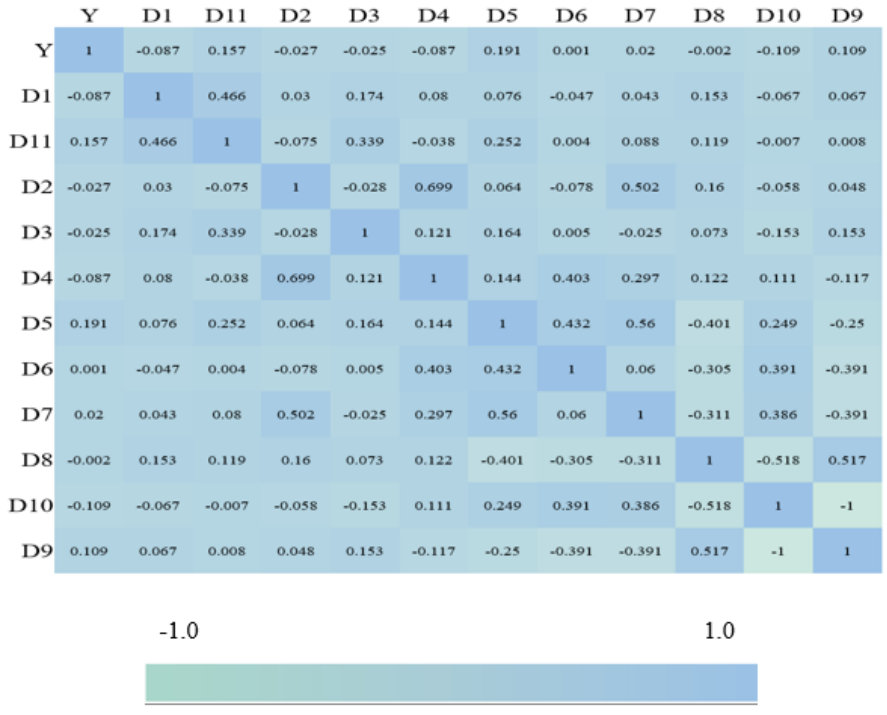


Figure 4. Visualization of 2023-wimbledon-1701.

An evaluation of the prediction effect of model based on Hopfield network was carried out to obtain the prediction results for the non-participating training data. The resultant data were first transformed with the discrete wavelet demy function.

$$\phi_{jk}(x) = 2^{\frac{j}{2}}\phi(2^j x - k) \tag{16}$$

The final prediction results were determined based on the transformed data means and their prediction success rates were analyzed in python, and some of the final results are shown in the table 3.

The charts show that it is possible to accurately predict a player's next state of play based on the selected indicators, which in turn shows that there is a correlation between "momentum" and a player's state of play during a match.

Table 3. The final prediction results analyzed in python.

	Actual value	Predicted value	Proposed prediction rate
1	1	1	75.30%
2	1	0	74.50%
3	0	0	82.00%
4	1	1	89.40%
5	0	1	85.40%
6	0	0	86.90%
7	1	1	83.50%
8	0	1	82.00%

In this paper, the fluctuation of the data of selected representative indicators when the player's game status changes is shown in the table 4.

Table 4. The fluctuation of the data of selected representative indicators.

Change of state	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9	D_{10}	D_{11}
Lose points in last game and score in the next	0.00	0.45	0.00	1.00	-10.00	0.00	0.00	0.05	0.00	0.72	0.28
	7.50	0.65	1.00	2.00	-9.00	0.00	0.18	0.05	0.00	0.52	0.28
Score in the last game and lose points in the next	7.50	0.60	1.00	0.00	13.00	0.09	0.19	0.05	0.00	0.62	0.38
	5.00	0.45	0.00	1.00	12.00	0.09	0.14	0.05	0.00	0.73	0.38

4. Conclusions and outlooks

This paper finds that it is possible to accurately predict a player's next state of play based on the selected indicators, which in turn shows that there is a correlation between "momentum" and a player's state of play during a match.

When the state is from losing to winning, the average score, high scoring rate and the number of consecutive wins in the current game will increase, and the difference with the opponent as well as the probability of losing will decrease, according to the results of Established model which reflects the dynamic changes of the match, the personal performance coefficient of the player rises from 0.37 to 0.41, while the state is from winning to losing, the above indexes change in the opposite way, the average score, high scoring rate and the number of consecutive wins in the current game will decrease, and the difference with the opponent as well as the probability of losing will increase. In summary, this paper believes that the momentum in the process of the player's game has an impact on the state of the game. The player should pay attention to the role of their own momentum when they are in a continuous state of scoring, trying to get more points. While in the continuous process of losing points and a large gap between the opponents, players should change the game strategy and adjust the mentality according to the form.

In the future, this paper hopes that this model can be applied to more field games, such as badminton, softball, etc., in order to promote the development of this sport.

References

[1] Šarčević A, Vranić M, Pintar D, et al. Predictive modeling of tennis matches: a review[C]//2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO). IEEE, 2022: 1099-1104.

[2] Wilkens S. Sport's prediction and betting models in the machine learning age: The case of tennis [J]. Journal of Sports Analytics, 2021, 7(2): 99-117.

[3] Sim M K, Choi D G. The winning probability of a game and the importance of points in tennis matches [J]. Research quarterly for exercise and sport, 2020, 91(3): 361-372.

[4] Gao Z, Kowalczyk A. Random forest model identifies serve strength as a key predictor of tennis match outcome [J]. Journal of Sports Analytics, 2021, 7(4): 255-262.

[5] Ünal T. Predicting tennis match outcome: a machine learning approach using the SRP-CRISP-DM framework [D]. Middle East Technical University, 2023.

[6] Gollub J. Forecasting serve performance in professional tennis matches [J]. Journal of Sports Analytics, 2021, 7(4): 223-233.

[7] Martínez-Gallego R, Crespo M, Ramón Llin J, et al. Men's doubles professional tennis on hard courts: Game structure and point ending characteristics [J]. 2020.

[8] Cui Y, Liu H, Liu H, et al. Data-driven analysis of point-by-point performance for male tennis player in Grand Slams[J]. Motricidade, 2019, 15(1): 49-61.

[9] O'Donoghue P, Ingram B. A notational analysis of elite tennis strategy [J]. Journal of sports sciences, 2001, 19(2): 107-115.

[10] Gu W, Saaty T L. Predicting the outcome of a tennis tournament: Based on both data and judgments [J]. Journal of Systems Science and Systems Engineering, 2019, 28: 317-343.

[11] Bayram F, Garbarino D, Barla A. Predicting tennis match outcomes with network analysis and machine learning[C]//SOFSEM 2021: Theory and Practice of Computer Science: 47th International Conference on Current Trends in Theory and Practice of Computer Science, SOFSEM 2021, Bolzano-Bozen, Italy, January 25–29, 2021, Proceedings 47. Springer International Publishing, 2021: 505-518.