Problem Chosen	2024	Team Control Number
C	MCM/ICM Summary Table	2408343

Capture the “Momentum”, Take Control of the Game

Summary

To quantify the performance of each player in the match, we firstly analyze the related papers and data provided, and then construct a comprehensive **evaluation index system**, which consists of 3 dimension and 9 indicators totally, including The Comprehensive Performance dimension, The Technical Skills dimension and The Physical and Psychological Quality dimension. Secondly, we allocate the weights of each indicators with two methods: through professionals and papers**Entropy Weight Method (EWH) method**. Thirdly, we develop the **EWH-Topsis evaluation model** to calculate the scores of players’ performance. Finally, we selected 3 matches and for each we drew the **time sequence figure** of players’ performance scores, which makes the situation of the players in a given time is evidently presented. We can find through the curves that the performance scores of players constantly fluctuate.

We begin with measuring momentum through swings, runs of success, and score difference, and add them up by weight. Then we explore the non-randomness of swings and runs of success with **Augmented Dickey Fuller test**. The results of the test prove that they are stationary rather than random. Finally, we analyze the correlation between momentum and game results using the **Pearson correlation coefficient**. According to the correlation coefficient, there is a positive moderate correlation between them, indicating that the greater the momentum, the more likely it is to win the game.

To capture the momentum swings, we firstly measure the momentum swings of a match as the fluctuation of the momentum scores difference between two players, and the change of the flow of the play appears when the trend of the difference is about to change, that is the peak or the bottom on the time sequence curve of the momentum scores difference. Secondly, we select 9 indicators and use **Factor Analysis** to reduce the dimension of these indicators, which at last generate 3 new factors: ,,, .Thirdly, we predict the momentum change with the **Multiple Linear Regression Model**, We let the 3 factors constructed formerly be 3 independent variables, and let the momentum scores difference between the two players be the dependent variable, and input 20 matches data to calculate all the coefficients. The R^2 of this model is 0.03 .To optimize the fit ,we additionally use **BP Neural Network Prediction Model** to develop the Prediction Model, we input the same 20 matches data to predict other 10 matches, and the R^2 of this model is ,which means the latter model is better.

We use the model established before to predict on the data from the other three matches to test the predictive performance of the model. The results show that the prediction accuracy is around 70%, with a high accuracy. To optimize the model, we suggest adding three factors, including age, professional experience, and recent matches, to indicators for predicting momentum fluctuations. Finally, we apply the predictive model to other types of competitions and found that the model has a certain degree of universality.

Keywords: EWH-TOPSIS Evaluation Augmented Dickey Fuller Test Model Augmented-Dickey Fuller Test Multiple Linear Regression Model BP Neural Network Prediction Model

Contents

1 Introduction

1.1 Problem Background

The momentum of a match refers to the relative advantage and positive trend exhibited by a player or team in the match, usually reflected in scoring, ball control, offense, and other aspects.In some tennis matches, to our astonishment, some great players may lose to rising stars. Analyzing their game process, we found that the seemingly dominant side can sometimes experience incredible fluctuations in multiple points or even multiple games, which are usually attributed to momentum.

The psychology of athletes may be influenced by the momentum of the competition, especially at critical moments such as the match point. What's more， it also affects the emotions of the audience, creating a different atmosphere on the field, and affecting players' performance, thereby affecting the results of the match. At the same time, the performance of different players and the momentum of the match change over time. Understanding and grasping the change of the momentum and studying its relationship with the success of the match are of great importance for formulating tactical strategies, adjusting the state during the match and predicting the results.

1.2 Restatement of the Problem

Considering the background information, the main tasks of this paper are as follows:

• Problem 1 Based on the data of the matches, establish a mathematical model to describe the flow of play as points occur. The model identifies when player is performing in a match， as well as how much better they are performing. A visualization based on the model to depict the match flow needs to be included.

• Problem 2 Analyze the correlation between the "momentum" and the success of the matches to verify whether the swings in play and runs of success by one player are random.

• Problem 3 Develop a model to predict the swings in the matches and analyze the most relevant factors. Provide suggestions to a player according the differential in past match "momentum" swings.

• Problem 4 Test the model on other types of matches to evaluate predictive ability and identify other factors to improve the model.

• Problem 5 Summarize the findings by writing a report, including a memo for coaches with an introduction to the role of momentum in competitions and suggestions for athletes preparing for the match.

1.3 Our Work

We mainly build three models to solve these problems.

For problem 1, we selected nine indicators were from three aspects of Comprehensive performance, technical skills, physical and mental state to construct an athlete performance indicator system and establish an evaluation model. Use the model to describe the match flow, evaluate the performance of athletes, and finally visualize the results.

For problem 2, We begin with measuring momentum through swings, runs of success, and score difference, and add them up by weight. Then we explore the non-randomness of swings and runs of success with Augmented Dickey Fuller test. Finally, we analyze the correlation between momentum and game results using the Pearson correlation coefficient.

For problem 3, We first measure the momentum fluctuation of the game as the difference in momentum scores between two players. When the trend of the difference is about to change, the process of the game will change. Secondly, we selected 9 indicators and used factor analysis to reduce their dimension. Finally, we used multiple linear regression models and BP neural network models to predict momentum changes, and compared the goodness of fit to determine which model had the best predictive performance.

For problem 4, We use the model established before to predict on the data from the other matches to test the predictive performance of the model and consider more factors that can serve as indicators to optimize the model. Then we apply the model to other types of competition data to test its universality.

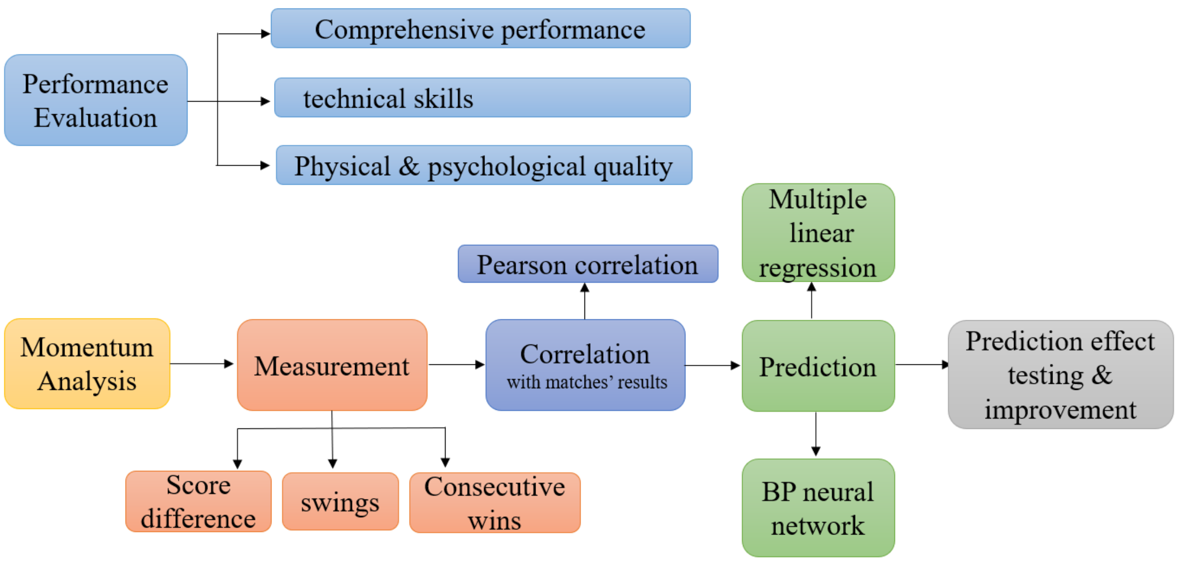


Fig 1. Our Work

2 Assumptions and Justifications

- **Assumption 1.** Servers Have an Advantage in the Match

•**Reason 1.** Servers exhibit a higher scoring rate during their serving games, indicating a higher probability of scoring for servers in the match.

•**Assumption 2.** Momentum Influences Match Outcomes, and Player Performances Are Random During Matches

•**Reason 2.** Player momentum is influenced by various factors such as consecutive points won and the point difference with the opponent.

•**Assumption 3.** Player Performance Is Affected by Their State During the Match

•**Reason 3.** Player performance can be influenced by factors such as physical condition, technical proficiency, error rates, and whether they are in peak condition. For example, in matches where players are in better states, their scoring rates are higher.

•**Assumption 4.** Fluctuations Exist in the Match, Signifying Differences in Player Performance at Different Times, But These Are Not Attributable to Momentum and Are Random Occurrences

•**Reason 4.** Player performance may vary due to multiple factors such as physical fatigue, tactical adjustments, and opponent performance. Analysis of match data indicates that players' scoring rates and performances differ in various games and match stages, but these fluctuations are more likely attributed to random occurrences rather than momentum.

3 Notations

The key mathematical notations used in this paper are listed in Table 1.

Table 1: Notations used in this paper

Symbol	Description	Unit
$Diff_{mom}$	the momentum scores difference between Player 1 and Player 2 in each set	/
u_i	The indicators of momentum swigs	/
T_q	the final score of the evaluation object	/
e_q	Entropy	/
s_q	Difference Coefficient	/

4 Data Preprocessing And Description

In the process of analyzing tennis match data, we first conducted a detailed assessment of missing values in the dataset. We discovered that key performance indicators such as rally_count, serve_width, serve_depth, and return_depth all had missing values. To maintain the integrity and continuity of the data, we decided to use linear interpolation to fill these gaps. This method is based on the linear relationship between adjacent observations, calculating and filling in missing values, and is a standard practice for handling continuous numerical data.

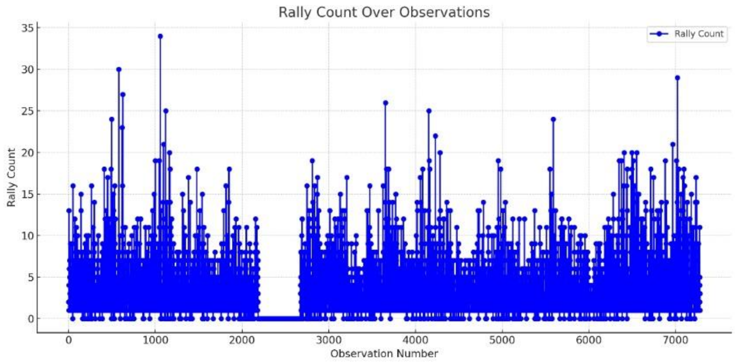


Fig 2. Rally Count Over Observations

5 Tennis Player Performance Evaluation: EWH-TOPSIS Evaluation Model

5.1 The Establishment of TOPSIS Evaluation Model

5.1.1 Construct evaluation index system

By reading and analyzing relevant literature, this paper constructs the performance evaluation index system of tennis players. The evaluation index system consists of 9 indicators, including comprehensive performance, technical skills, physical and psychological quality respectively. Among them, the comprehensive performance includes the cumulative scores of the player on the court and the score of the time point (the cumulative scores up to the point, the score of the point); The technical skills include 4 indicators related to serving skills, hitting skills and error rate (points of serve wining, points of hit winning, error rate of serving and number of unforced errors); Physical and psychological quality includes two sub-indexes to measure physical quality (running meters of one hit, serving speed) and match point chance grasp to measure psychological quality. We then did a series of calculation to get the specific numerical values of each index precisely.

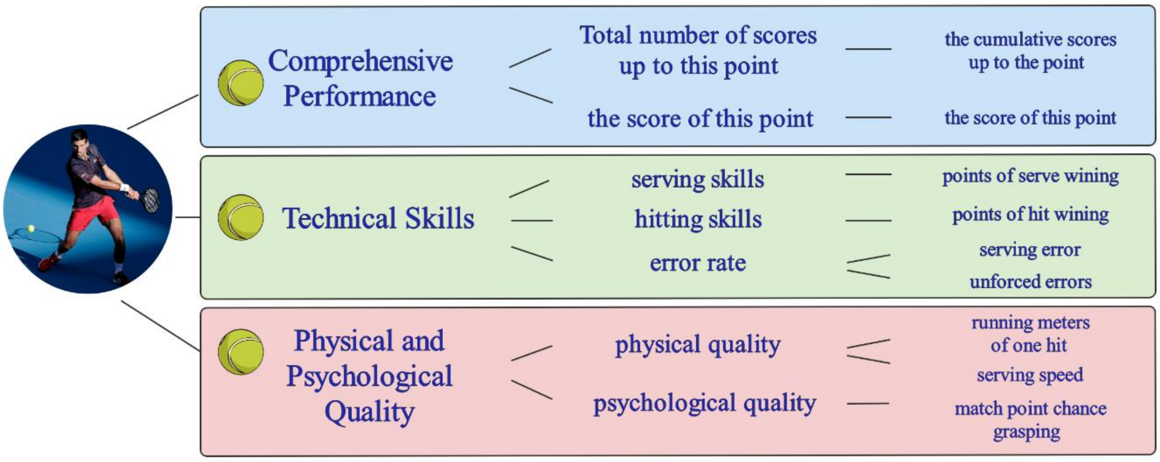


Fig 3. Evaluation Index System

5.1.2 Determine the Weight of the Evaluation Index

- **Through professionals**

When allocating the weight of each indicators, through the discussion with relevant professionals, coaches or team members, we fully understand the characteristics of tennis games and athletes, formulate a more reasonable weight, reflecting the authenticity and scientific evaluation of performance. The weight of each indicators we design is as below:

Table2. The weight of each indicator

The performance of athlete	Comprehensive Performance (0.3)	the cumulative scores up to the point(0.15)
		the score of this point(0.15)
	Technical Skills(0.5)	points of serve wining(0.15)
		points of hit wining(0.15)
		serving error(0.10)
		unforced errors(0.10)
	Physical and Psychological Quality(0.2)	running meters of one hit(0.10)
		serving speed(0.05)
		match point chance grasping(0.05)

- **Entropy Weight Method （EWM）**

The entropy weight method is a method of calculating objective weights to determine the probability of occurrence of a certain state among n independent possible states.

$$P_{pq} = x_{pq} / \sum_{P=1}^{n} x_{pq}$$

Calculate entropy:

$$e_q = k \sum_{P=1}^{n} P_{pq} \ln\left(P_{pq}\right),$$

Where $k = \frac{n}{-\ln n}$,

Determine the difference coefficient of indicators:

$$s_q = 1 - e_q, \qquad \text{q=1,2…m}$$

For each q, the smaller the difference coefficient s_q, the greater the entropy e_q. The index q in evaluation is relatively insignificant. Conversely, the greater the difference coefficient s_q is, the smaller the entropy e_q, and the bigger the role of index q in the evaluation.

Calculate the weight of each indicator:

$$w_q = 1 - s_q / \sum_{q=1}^{m} \left(1 - s_q\right), \qquad \text{q=1,2…m}$$

Use the entropy weight method to calculate the weight of each indicator, as shown below:

Table3.the weights of index

index	Entropy e_q	Difference Coefficient s_q	Weight
the cumulative scores up to the point	0.995	0.005	20.504%
the score of this point	0.765	0.235	12.874%
points of serve wining	0.886	0.114	14.395%
points of hit wining	0.654	0,364	10.556%
serving error	0.521	0.479	8.997%
unforced errors	0.651	0.387	9.895%
running meters of one hit	0.698	0.302	10.589%
serving speed	0.273	0.727	4.763%
match point chance grasping	0.515	0.485	7.427%

Multiply the weight of each indicator w_q with the specific data of all indicators to obtain a weighted evaluation matrix H, that is:

$$H = (r_{pq})_{n\times m} = (w_q x_{pq})_{n\times m}$$

5.1.3 Forward Process the Index and standardize Data

First of all, {max-x} is used to forward process the data of error rate of serving and number of unforced errors. As the other indicators are all forward indicators, there is no need to forward process.

Attentionally, considering that the player serving has a much higher probability of winning the point, We take into account who is the server in every point, and deal with the data by giving less scores to the winner who simultaneously is the server in every point than the winner who is not the server.

Then with the number of evaluation objects is p and the number of evaluation indicators is q, a positive data matrix of p*q is denoted as:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1q} \\ x_{21} & x_{22} & \cdots & x_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pq} \end{bmatrix}$$

For the object p , the indicator q, x_{pq} is its corresponding data , After standardizing the matrix X , we can get the matrix Z, where $z_{ij} = x_{ij} / \sqrt{\sum_{i=1}^{n} x_{ij}^2}$.

5.1.4 Calculate the score of the evaluation object

Determine positive ideal solutions $O+$ and negative ideal solutions $O-$:

$$O^+ = \max\{(r_{1q}, r_{2q}, \dots, r_{nq})\}$$

$$O^- = \min\{(r_{1q}, r_{2q}, \dots, r_{nq})\}$$

Calculate the distance between each evaluation index and positive ideal solution and negative ideal solution.

$$D^{+} = \sqrt{w_q \sum_{p=1}^{n} (r_{pq} - O^{+})^2}$$

$$D^{-} = \sqrt{w_q \sum_{p=1}^{n} (r_{pq} - O^{-})^2}$$

Calculate the relative proximity T_q between the evaluation object and the optimal solution, that is, the final score of the evaluation object:

$$T_q = \frac{D^{+}}{D^{+} + D^{-}}$$

5.2 The Result visualization

To capture the flow of the match, we visualize two players' performance scores in one match by using the above models we constructed.

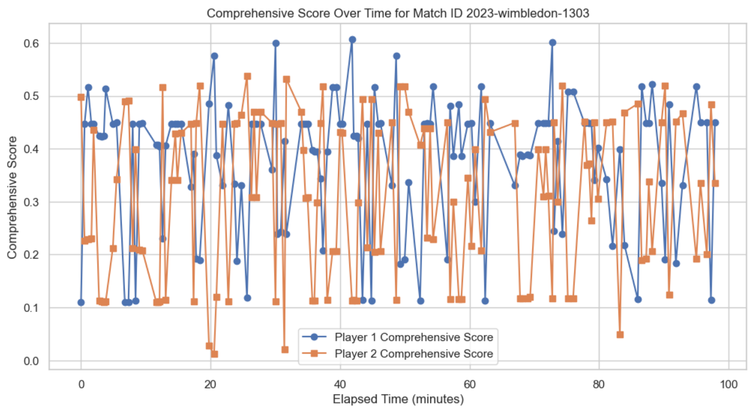
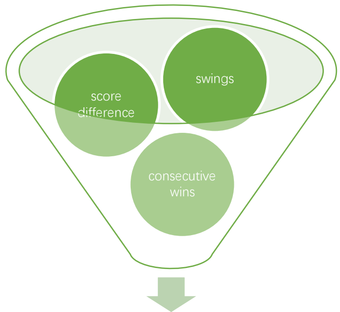


Fig 4. The Performance Score Time Sequence of 2 Player in two Matches

6 The Role of Momentum: Correlation Analysis Model

6.1 Measure momentum

Before conducting correlation analysis, it is essential to measure the momentum. Taking into account the background of the issue and the definition of momentum, we have chosen three key indicators, namely score difference, swings and consecutive success, to jointly quantify the momentum.



momentum score

We define swings as the oscillations in score, specifically the difference between the scores of the current game and those of the preceding game.

swings = score of the current game - score of the preceding game

The measurement of consecutive success is based on the consecutive occurrence, and different scores are assigned according to the consecutive count. We assess consecutive successes from two perspectives: consecutive successes of games and consecutive successes of sets. They are assigned different weights and summed up. Due to the greater significance of set scores in determining the ultimate victory, consecutive successes within a set account for 70%. Winning two sets in a row is worth 5 points, winning three sets in a row is worth 15 points, and the point of winning some games in a row is scored below.

Number of consecutive wins	2	3	4	5	6	7
Momentum score	1	3	6	10	15	21

Scoring rules for different number of consecutive wins

Score difference is utilized to gauge the player's advantage at the end of the game. The calculation of score difference should consider the difference between the two sides within the set and the difference in the match. Given the more pronounced impact of score difference between sets on the final result, it is assigned a weight of 75%, while the score difference within a set accounts for 25%.

score difference = the player’s score – the opponent's score

We believe that these three indicators have a similar impact on momentum. Therefore, the weights are evenly distributed, and the momentum score is obtained by adding the weights after standardization and normalization.

6.2 Explore the non-randomness of swings and runs of success with ADF test

Firstly, we visually observe the trend of the data by plotting the time series graph, making an initial assessment of its randomness.

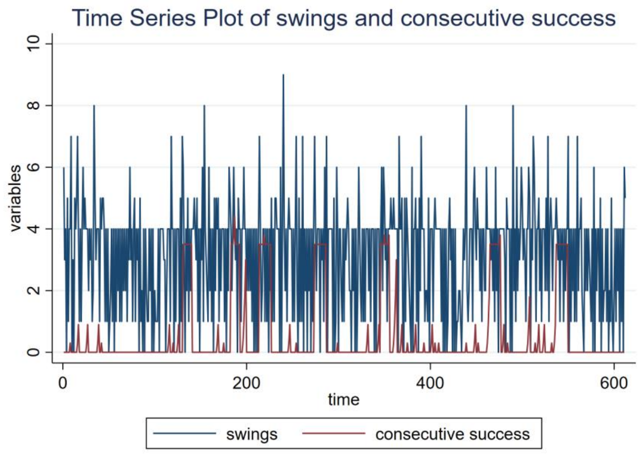


Fig 5. time series plot of swings and consecutive success

Then we use ADF test to test the stationarity of the data, and the results are as follows. ADF test, also known as Augmented Dickey Fuller test, is a statistical test method used to test the existence of unit roots in time series data. The unit root indicates that time series data has non stationarity, meaning that the mean or variance may vary over time. The absence of unit roots means that the changes in data remain at a certain level over time. This causes time series variables to exhibit a trend rather than random fluctuations.

Table4. the result of ADF test for swings data

H0: Random walk without drift, d = 0

	Test statistic	Dickey-Fuller critical value 1%	5%	10%
Z(t)	-7.431	-3.634	-2.952	-2.610

MacKinnon approximate *p*-value for Z(t) = **0.0000**.

Table5. the result of ADF test for data

H0: Random walk without drift, d = 0

	Test statistic	Dickey-Fuller critical value 1%	5%	10%
Z(t)	-8.312	-3.634	-2.952	-2.610

MacKinnon approximate *p*-value for Z(t) = **0.0000**.

The ADF test value is significantly lower than the critical value, and the p-value is lower than the significance level, rejecting the null hypothesis that the data is stationary rather than random.

6.3 Analyze the correlation between momentum and results with regression model

Firstly, the momentum changes of both players in each game were drawn to observe the momentum changes of both teams, and the correlation between momentum and the match result. From the figure, we can see that the momentum of both sides always has a big difference, showing a high and a low state, and the game result is also a win and a loss. Therefore, from the intuitive view, there seems to be some correlation between momentum and the game result.

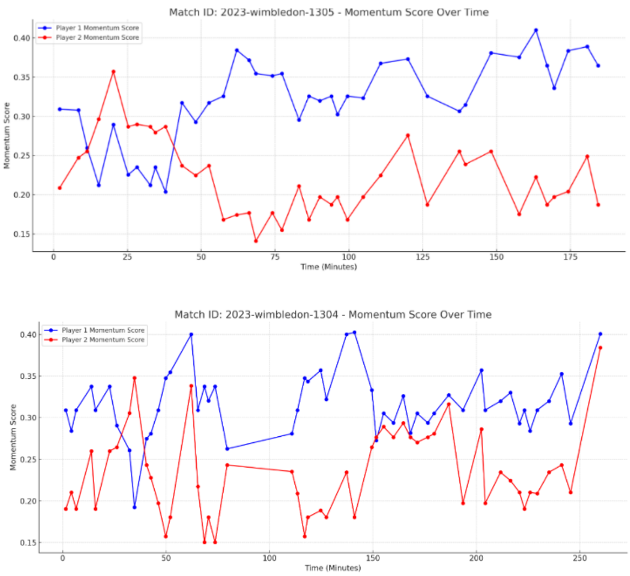


Fig 6. the time sequence of momentum scores in 2 matches

Then we further analyzed the correlation between the game results (win with 1 and lose with 0) and momentum. This type of variable is called a binomial variable and can be measured by calculating the Point-Biserial Correlation Coefficient to measure the strength of its relationship with the continuous variable. Although Pearson correlation coefficient is usually used to measure the linear relationship between two continuous variables, we can still use Pearson correlation coefficient here, because although the victory indicator is binary, it can still be regarded as a continuous variable on the [0,1] interval in statistical analysis, which can be understood as the probability of winning. Pearson correlation coefficient is calculated as follows:

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

In the above equation, r represents Pearson correlation coefficient, n represents the number of samples, and X_i and Y_i represent the samples of the two groups of variables respectively. The value of r is between 0 and 1, and the absolute value of r closer to 1, indicating the greater the degree of correlation. The positive and negative correlation coefficients represent positive correlation and negative correlation respectively.

The results show that the correlation coefficient between Player 1's momentum score and Player 1's game result is 0.317, while the correlation coefficient between Player 2's momentum score and Player 2's game result is 0.320. This indicates that there is a positive moderate correlation between momentum and game result， which means that players with higher momentum scores are more likely to win the game. Use these data to draw a graph to visually show the relationship between momentum score and game victory.

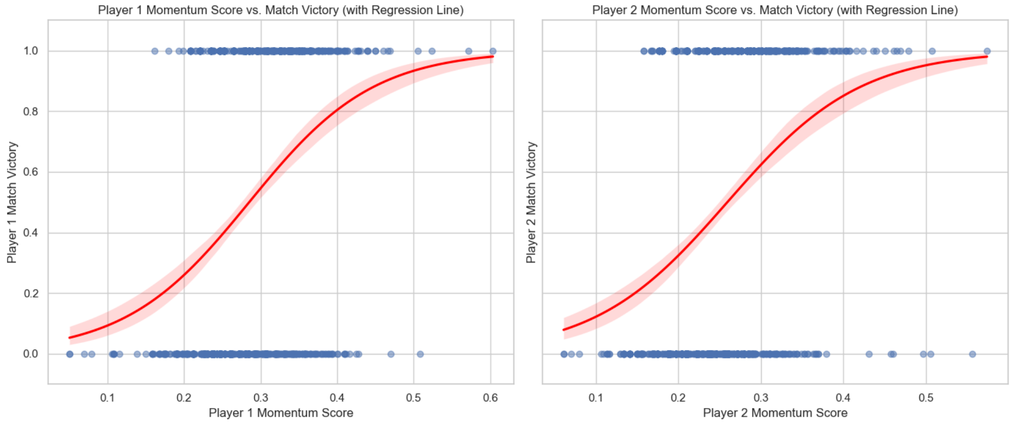


Fig 7. Scatter plot of correlation between momentum and match outcome

The graph includes regression lines and confidence intervals to more clearly show the relationship between players' momentum scores and game victory. Through these graphs, we can visually see: There is a positive relationship between momentum scores and game wins.The confidence interval provides information about the predictive accuracy of the regression line, indicating that most data points cluster around the predictive line, increasing our confidence in the validity of this correlation.

7 Capture the “Momentum” Swings: Factor Analysis and Multiple Linear Regression Model / BP Neural Network

7.1 The measurement of “momentum” swings

For the third problem, we use the momentum scores difference between Player 1 and Player 2 in each set($denoted\ as\ Diff_{mom}$) as an indirect indicator of the momentum flow of the set. In particular, if the momentum difference is greater than zero, it means that the current situation on the field is biased in favor of player 1; Conversely, if the momentum difference is less than zero, the situation on the field is biased in favor of player 2. Furthermore, we observe that the absolute increase in the difference reflects the extent to which the leading momentum of the player on the field is evident. Therefore, a larger absolute value indicates that the player has achieved a more significant lead in the current situation.

In order to more intuitively reveal the change of the situation on the field, we plot the time series of the momentum difference. By analyzing the inflection points of the timing charts, we are able to identify the moments when momentum shifts. When the sequence map shows an evident peak or bottom point(specifically, the multiplization of the the peak point values and the next bottom point values should be less than 0), it indicates that the situation has changed at the corresponding point in time, which provides us with key information. This analysis method is helpful to understand the change of competitive state between players in the match, and provides an objective and visual basis.

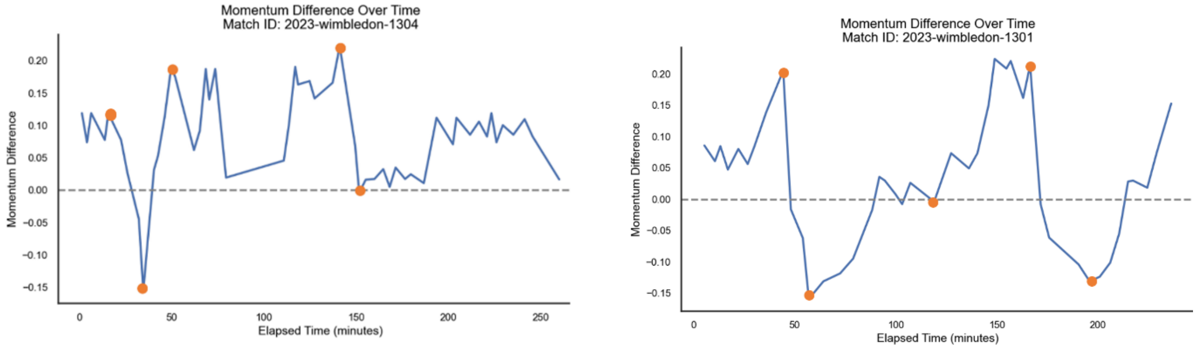


Fig 8. the time sequence of $Diff_{mom}$ in two match

As shown in fig , the curve present the variation of the flow of the play, and the orange point refer to the evident change of the momentum situation. In the 1304 match, at first , the match seems favor player 1, but after about 20 minutes, the momentum preference change to player 2, we can see from the figure that the $Diff_{mom}$ drop to negative and reach the bottom in about another 20 minutes.

7.2 Find the Related Factors: Evaluation Model and Factors Analysis

7.2.1 Evaluation factors

By reading tennis related literature, the following indicators are constructed as factors affecting momentum change:

Table6. The initial factors of momentum difference:

momentum difference factors	The mean service speed
	the total point of the set(u_1)
	the mean running distance of P1 and P2(u_2)
	the point difference between P1 and P2(u_3)
	the difference in untouchable winning serve number between P1 and P2(u_4)
	the difference in untouchable winning shot number between P1 and P2(u_5)
	the difference in performance scores (see problem 1) between P1 and P2(u_6)
	the difference of error rate between P1 and P2(u_7)
	The difference of performance scores between P1 and P2(u_8)

By taking the above indicators into account, we can more comprehensively understand the factors that affect the momentum shift in tennis matches, so as to provide better game strategies and training directions for players and coaches.

7.2.2 Factors analysis

However, because the feature dimension is too high, it is very complicated for both subjective and objective methods to establish the evaluation system, so we use factor analysis to reduce the dimension:

- **Correlation analysis of indicators.** The premise of factor analysis is that there is correlation between variables, so it is necessary to carry out KMO test and Bartlett sphericity test for sample data.

Table7 KMO and Bartlett Test

The number of KMO sampling appropriateness		0.759
Bartlett sphericity test	Approximate chi-square	2612.687
	degree of freedom	36
	Distinctiveness	0.000

The test results are shown in Table 2. As can be seen from Table 2, the KMO detection value is 0.759, greater than 0.5; The P-value of Bartlett sphericity test is 0, which is less than 0.05, indicating that there is a significant correlation between the variables, and the data is suitable for factor analysis.

- **Extract common factors.** The criterion for extracting common factors is that the eigenvalue is greater than 1.

Table8 extraction results of factors

ingredient	Initial eigenvalues			Extract the sum of squares of the load			Sum of squares of rotational loads		
	total	Variance %	cumulation %	total	Variance %	cumulation %	total	Variance %	cumulation %
1	2.190	24.336	24.336	2.190	24.336	24.336	2.137	23.740	23.740
2	1.899	21.097	45.433	1.899	21.097	45.433	1.915	21.276	45.016
3	1.391	15.459	80.893	1.391	15.459	60.893	1.429	15.877	60.893
4	0.951	10.565	82.458						
5	0.791	8.791	86.249						
6	0.728	8.090	88.338						
7	0.475	5.277	93.616						
8	0.429	4.771	98.387						
9	0.145	1.613	100.000						

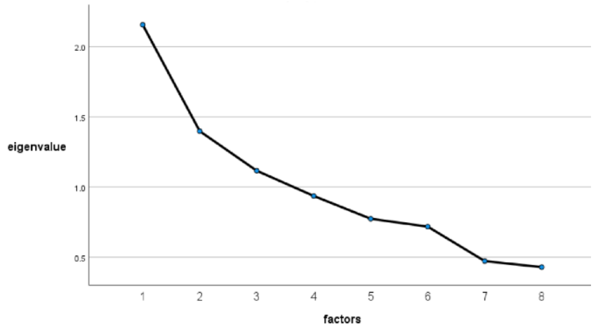


Fig 9. Scree plot

The extraction results of common factors are shown in Table 3. It can be seen from Table 3 that the eigenvalues of the first three factors are all greater than 1, and the cumulative variance contribution rate reaches 80.893%, indicating that these four factors can well represent the main characteristics of the hidden debt risk of local governments.

• **Determination of common factors and their weights.** The correlation coefficient between each factor and the common factor can be determined by the factor load matrix after rotation. The product of the raw data of each index and the component score coefficient is the common factor score. The component score coefficient matrix is shown in Table 5.

Table9 Common Factor Coefficient Matrix

	Common Factor		
	1	2	3
The mean service speed	-0.093	-0.339	0.262
the total point of the set	0.042	0.523	-0.010
the mean running distance	-0.035	0.504	0.211
the point difference	0.091	-0.030	-0.752
the difference in untouchable winning serve number	0.414	0.090	-0.111
the difference in untouchable winning shot number	0.390	0.122	-0.105
the difference in error rate	0.199	0.000	0.414
the difference in performance scores	0.330	-0.194	0.067

As can be seen from Table 9, there is a significant correlation between the difference in untouchable winning serve number, the difference in untouchable winning shot number, the difference in error rate, the difference in performance scores and common factor 1; and there's a significant correlation between momentum difference, the point difference and common factor 2, and there's a significant correlation between the total point of the set, the mean running distance, By analyzing the factors comprehensively and with the common factor coefficient matrix in Table9, we can get 3 new factors :

$$\begin{cases} F1 = -0.093u_1 + 0.042u_2 - 0.035u_3 + 0.091u_4 + 0.414u_5 + 0.390u_6 + 0.199u_7 + 0.330u_8 \\ F2 = -0.339u_1 + 0.523u_2 - 0.504u_3 - 0.030u_4 + 0.090u_5 + 0.122u_6 - 0.194u_8 \\ F3 = 0.262u_1 + -0.010u_2 + 0.211u_3 - 0.752u_4 - 0.111u_5 - 0.105u_6 + 0.414u_7 + 0.067u_8 \end{cases}$$

7.3 Predicting the Momentum Swings

7.3.1 The Establishment of Multiple Linear Regression Model

We let the 3 factors constructed formerly be 3 independent variables, and let the momentum scores difference between the two players be the dependent variable.

$$Diff_{mom} = b_1F1 + b_2F2 + b_3F3 + \varepsilon$$

Then, we input 20 matches data to calculate all the coefficients, and get the specific regression formula.

$$Diff_{mom} = 0.0031F1 - 0.0018F2 - 0.0020F3 + 0.0963$$

Table10. The coefficient of regression variables

$Diff_{mom}$	Coefficient	Std. err.	t	P>t	R^2	Root MSE
F1	0.003144	0.003942	0.8	0.025	0.0302	0.1064
F2	-0.0018	0.000419	-4.31	0.000		
F3	-0.00207	0.001209	-1.72	0.086		
_cons	0.096262	0.013227	7.28	0.000		

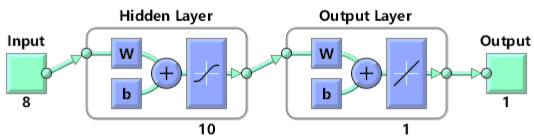
As can be seen from the table, the regression coefficient is significant and the$R^2$is 0.0302.

7.3.2 The Establishment of BP Neural Network Prediction Model

Apart from predicting with factors and Regression model, we use BP Neural Network Prediction Model to predict the momentum difference between two players directly.

・**Model Description**

BP neural network prediction model is a model that learns a large amount of input-output corresponding data through the fastest descent method, and constantly backpropagates to adjust network parameters and optimize its own prediction effect. The topology of BP neural network is shown in Figure 1, including input layer, hidden layer and output layer.



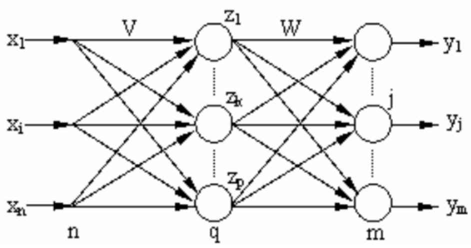


Fig 10. The Structure of Neural Network

The BP neural network prediction step diagram is shown below:

Let the input layer of BP network consist of m nodes, the hidden layer has q nodes, the output layer has n nodes, the weight between the input layer and the hidden layer is v_{kj}, the weight between the hidden layer and the output layer is w_{jk}, and the transfer function of the hidden layer is $f1$, and the transfer function of the output layer is $f2$, then the output of the hidden layer point is:

$$z_k = f_1(\sum_{i=0}^{n} v_{ki}x_i)$$

The output of the output layer node is：

$$y_j = f_2(\sum_{k=0}^{q} w_{jk}z_k)$$

• **The Prediction Outcome of Momentum "Swings" with BP Neural Network**

we input the same 20 matches data to predict other 10 matches, the fit outcome of BP Neural Network is shown as Fig., and the R^2 of this model is 0.0937, which is much better than the multiple linear regression model.

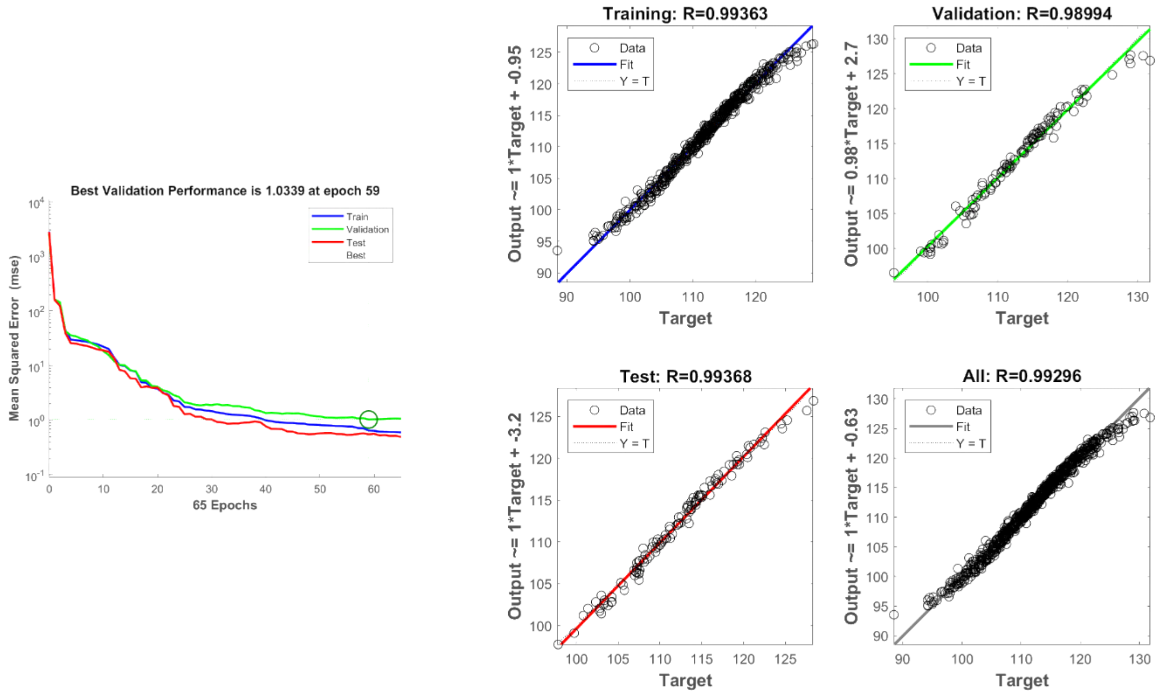


Fig 11.fitting process and Plot Regression

7.4 The suggestions for a player going into a new match

Based on the factors of the above regression analysis and the analysis of tennis player related literature, the following suggestions are put forward.

• Pay attention to the dynamic change of the momentum score difference swings drawn from the above model we constructed and adjust the tactics according to the flow of play. Be confident when you are ahead and look for opportunities to come back when you are behind.

• Strive to maintain a high speed when serving to increase the difficulty of the opponent's response, and choose to change the serving speed at a critical moment to increase the uncertainty of the opponent.

• Control the rhythm of the game and avoid too frequent turns to maintain the stability of physical strength and concentration. At the right time, you can choose to break the rhythm of your opponent by controlling the number of turns.

• Minimize the total distance of your own running through efficient movement and tactical arrangements to maintain physical and mental status. At the same time, observe the fatigue of the opponent and use tactics to guide them into more runs.

• Emphasize the opportunity to win in the service game by changing the direction and speed of the serve to create a situation that is difficult for the opponent to predict. At the same time, strengthen the defense of the opponent's service game and reduce the chances of winning.

• Control your own error rate, especially in key games and important moments. Observe the mistakes of the opponent and be good at using the other side's mistakes to gain points.

8 Generality of Our Prediction Models

8.1 The results of the prediction with the data provided

We use our Regression model and BP neural network model to predict the momentum sings in other matches, one of the results is shown in the following figure, as we can see, the curves of prediction and real value fluctuate with similar trend generally, which means our model has a good performance. But there's also some unprecise prediction. As calculated, the accuracy of the prediction of 10 matches is 70.13%.

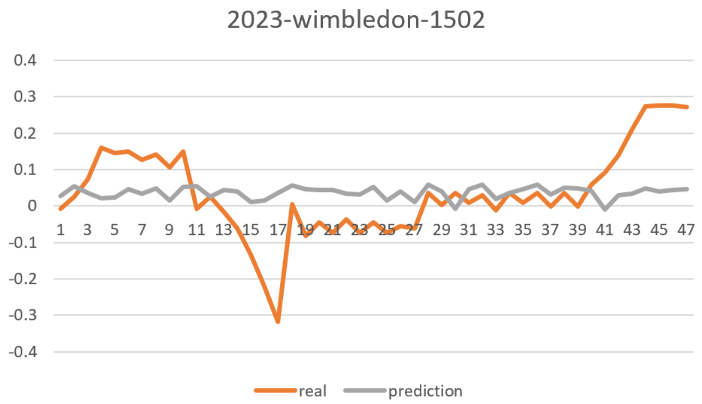


Fig12. the prediction and real value in match 2023-wimbledon-1502

Table11. Part of the Specific Values of Prediction

Real Value	Prediction Value
-0.00807	0.026917
0.02596	0.054732
0.073209	0.035435

0.160598	0.021887
0.144417	0.022797

8.2 The results of the prediction with data from other type of sport

In order to test the generality of the prediction model, we queried the data of women's tennis matches and applied it to the multiple linear regression prediction model to women's tennis matches, as can be seen from the figure, the prediction effect is very good, and the accuracy is 80.64%.Therefore, our model is versatile for predicting momentum swings.

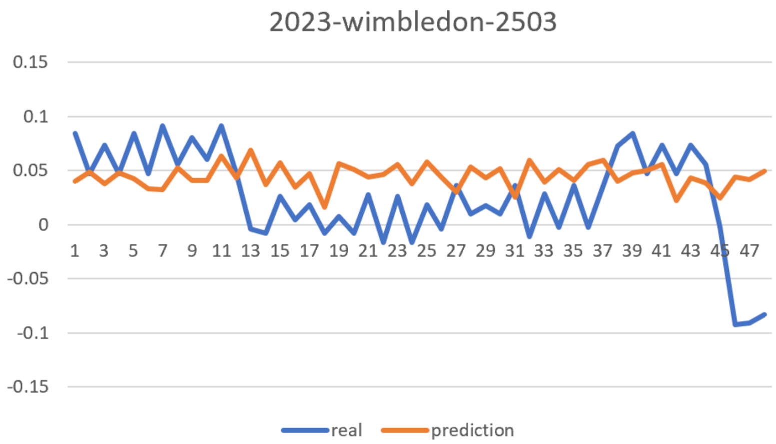


Fig13. the prediction and real value in match 2023-wimbledon-2503

Table12. Part of the Specific Values of Prediction

real	prediction
0.084427	0.040007
0.047132	0.048631
0.073209	0.037861
0.047132	0.047992
0.084427	0.042796
0.047132	0.033343

8.3 The correlation between the match result and the momentum score

We built two models, the first to assess the correlation between momentum in a match and the outcome of a match, while the second model to study the relationship between momentum shifts and key indicators of the match. To verify the generalizability of these two models, we chose to use Wimbledon women's race data for validation. The validation results show that there is a significant positive correlation between momentum and match results, which verifies the good versatility of these models in different scenarios.

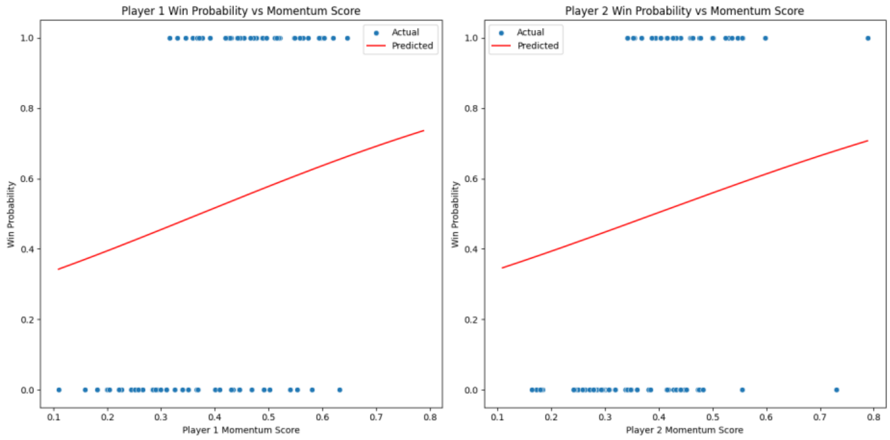


Fig14. the validation results

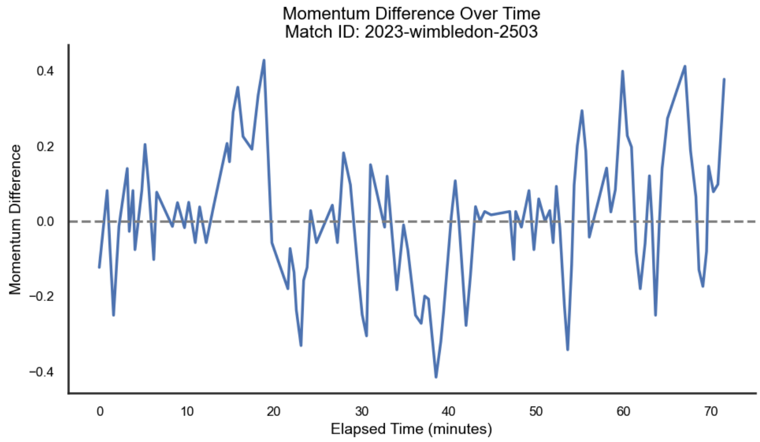


Fig15. the time sequence of momentum scores difference

9 Model Evaluation

9.1 Strengths

• **Comprehensive consideration of weights** We use entropy weighting method, which is objective weighting method based on relevant information and practical considerations to determine weights, constructing a more reasonable performance evaluation system for athletes.

• **Comparison of prediction methods** When predicting momentum fluctuations, we combine factor analysis with regression analysis, and use two methods such as regression analysis and BP neural network to make predictions. By judging the R-squared, we find that BP is better.

• **Visual display** Due to space limitations, for results that are inconvenient to display, we select a portion for visual display, which is more intuitive;

• **Practical significance** Based on practical game data, combined with actual situations and related research, a model is constructed to provide practical suggestions for players and coaches, which has certain practical reference significance

9.2 Weaknesses

• **Lack of consideration for individual differences** Measure momentum based on the competition situation, without considering the player's own factors, such as stress resistance, physical health, recent psychological state, and pre competition fatigue level;

•**Lack of theoretical derivation** Our model provides empirical rules and relationships between momentum and competition results based on competition data, lacking scientific theoretical derivation and requiring further research

10 Conclusion

In this article, after significant preprocessing of competition data, we quantified and organized the indicators. Firstly, a player performance evaluation model was established, and the EWH TOPSIS model was used to evaluate player performance, and the evaluation results were visualized. Secondly, measure momentum and use Pearson correlation coefficient to analyze its correlation with game results. The results showed a positive correlation between momentum and victory in the competition, with players with greater momentum having a greater likelihood of winning. Finally, we use two models, multiple linear regression and BP neural network, to predict momentum fluctuations. Based on the goodness of fit, we found that the latter has better predictive performance. After testing, the prediction model has good predictive performance, with an accuracy rate of over 70%. Further testing indicates that the model proposed in this article still has universality for other types of competitions.

In order to make our model more rigorous and accurate, we applied it to other data and adjusted the indicators based on the results, and explored its advantages and disadvantages for reference.

In addition, based on our research findings, we introduced tennis coaches to the definition of momentum in competition, its measurement methods, and its key role in the game. We also provided them with some suggestions, such as changing the serving speed at critical moments, maintaining physical strength and focus, etc., hoping to be helpful to tennis players, which also makes our research more meaningful!

Memo

To: Tennis coaches
From: Team 2408343
Date: Feb 05,2024
Re: Clarify the importance of "momentum" in ball game, and give helpful advice for players to prepare for a match.

The "momentum effect" in sports refers to a scenario where a team is more likely to achieve success if they have been performing well in recent games. Winning streaks are commonly characterized as a team gaining momentum. This concept shares similarities with the "hot hand," frequently discussed in basketball, where a player is believed to have an increased likelihood of successfully making shots.

Main Conclusions

The notion of the "hot hand" belief has garnered considerable attention in academic literature. Our team has also conducted meticulous research on this phenomenon, and we are now poised to share our findings and insights：

- *Momentum do play a great role in the Sport, which may determines the outcome of the match.*

- *There are many factors that will influence the momentum in ball game: including the mean service speed, the total point of the set, the mean running distance of two players, the point difference between two players, he difference in the two players' number of untouchable winning serve or shot , the difference of error rate.*

Suggestions to players

To utilize the momentum effect, we advise the player in following aspects:

- *Adjust your strategy and mindset based on the feedback. When the score is ahead, stay focused and avoid slacking off; When you are behind, adjust your tactics and try to come back. Focus on key game play to ensure a score advantage at key moments.*

- *Strive to maintain a high speed when serving to increase the difficulty of the opponent's response, and choose to change the serving speed at a critical moment to increase the uncertainty of the opponent.*

- *Control the rhythm of the game and avoid too frequent turns to maintain the stability of physical strength and concentration. At the right time, you can choose to break the rhythm of your opponent by controlling the number of turns.*

- *Minimize the total distance of your own running through efficient movement and tactical arrangements to maintain physical and mental status. At the same time, observe the fatigue of the opponent and use tactics to guide them into more runs.*

- *Improve the efficiency of winning shots, focusing on winning at key moments, while defending opponents' winning shots. Observe your opponent's winning habits and adjust your tactics to limit their advantage.*

- *Control your own error rate, especially in key games and important moments. Observe the mistakes of the opponent and be good at using the other side's mistakes to gain points.*

Feb 05,2024

References

[1]Zhao Mingnan, Jiang Haoqi, Qi Bing, et al.Theoretical model construction and scale compilation of emotional intelligence of Chinese tennis players[J].Psychological Research,2024,17(01):42-53.DOI:10.19988/j.cnki.issn.2095-1159.2024.01.005.

[2]Yang Linfeng,Yan Jun.Tennis serve comprehensive evaluation system based on attitude recognition[J].Industrial Control Computer,2023,36(12):24-26.)

Contemporary Sports Science and Technolo-gy,2023,13(30):28-30.DOI:10.16655/j.cnki.2095-2813.2023.30.008.

[3]ZHANG Rong. Tennis match prediction and player analysis[D].Yunnan Univer-sity,2023.DOI:10.27456/d.cnki.gyndu.2022.000807.

[4]Ram Sumit Kumar,Nandan Shyam,Sornette Didier.Significant hot hand effect in the game of cricket[J].scientific reports,2022

[1]Jonas V ,Marc C D ,Jan A .Does time series momentum also exist outside traditional financial markets? Near-laboratory evidence from sports betting[J].Journal of Behavioral and Experimental Economics,2023,104

[5]P. M B ,C. T S ,K. L L , et al.The state of diversity in the Association for Applied Sport Psychology: Gaining momentum or still swimming upstream?[J].Journal of Applied Sport Psychology,2022,34(6):1372-1386.

[6]Scott ,Dascombe ,Delaney , et al.Running momentum: a new method to quantify prolonged high-intensity intermittent running performance in collision sports[J].Science and Medicine in Football,2017,1(3):244-250.

[7]Briki W .Rethinking the relationship between momentum and sport performance: Toward an integrative perspective[J].Psychology of Sport Exercise,2017,3038-44.