ISSN 2616-5775 Vol. 7, Issue 5: 152-158, DOI: 10.25236/AJCIS.2024.070520

Momentum Study Based on CRNN Neural Network and Logistic Regression

Zehao Qu[1,*]

[1]School of Computer and Communication, Lanzhou University of Technology, Lanzhou, China
[*]Corresponding author: 2717782356@qq.com

Abstract: *This paper investigates the influence of momentum on the player's victory or defeat in tennis matches. Through the preprocessing and comprehensive analysis of the match data, the Topsis model based on the entropy weighting method was used to quantify the momentum index, and the visual analysis of the change of players' momentum was realized. Further, the strong correlation between the momentum and the player's score was verified by the Pearson correlation coefficient method. The correlation analysis and CRNN neural network model were utilized to successfully predict the momentum fluctuation, which was validated by logistic regression model with an accuracy of 78.4%. This study provides a reference for coaches to customize better tactics and important suggestions for improving players' competitive level and game performance. By deeply exploring the influencing factors and prediction model of momentum, this study provides new perspectives and methods for the analysis of sports game data and player training.*

Keywords: *Topsis, CRNN Neural Networks, Logistic Regression*

1. Introduction

In tennis, the fierce competition between players and the uncertainty of the match results have always been the focus of attention, especially when there are jaw-dropping reversals in the match. Spectators have found that the players' momentum has an undeniable impact on the trend of the game, which in turn affects the change of the outcome of the match. The advancement of modern technology has made it possible to accurately record such imperceptible game data as serve speed, which provides a new way for in-depth research and exploration of the influence of momentum on the outcome of the game [1]. The purpose of this paper is to explore the mechanism of momentum in tennis matches, and to gain a deeper understanding of the influencing factors of momentum and its relationship with match results. By comprehensively analyzing and evaluating the match data, we adopt the Topsis model based on the entropy weighting method to quantify the momentum indicators and realize the visual analysis of the change of players' momentum. Further, we conducted a correlation analysis by the Pearson correlation coefficient method, which proved that there was a significant similarity between the change of momentum and the player's score. To help coaches assess and analyze athletes and game status, and to improve athletes' competitive level, we successfully predicted the fluctuation of qi by using correlation analysis and CRNN neural network model and verified the prediction results by logistic regression model with an accuracy of 78.4%. This study will provide coaches and players with suggestions for customizing better tactics, providing useful guidance for optimizing training and improving game performance.

2. Player evaluation and result visualization

2.1 Establishment of Topsis model based on entropy weight method

First, we construct a momentum scoring system to quantify momentum by considering the influence of different factors on momentum. Serving power, winning streak, and high level of scoring balls will cause positive momentum improvement for players, while losing, forced errors, and two service errors will cause players' momentum decline.

Based on the existing data, aiming at the problem, we use entropy weight method to objectively empower each indicator, combine Topsis model to quantify the performance of players in the match, and use this model to visualize the game flow [2]. The specific algorithm steps are shown as follows.

Step 1: First, we selected corresponding indicators according to key data features. The main influencing factors of player p1's momentum are the winner (point_victor), serve (server) and winning streak (calculated according to the data), while the secondary factors are the winning serve (pl_ace), winner (pl_winner), two service errors (pl_double_fault) and un-forced errors (pl_ unf_err).

We selected the corresponding indicators for the key data features (indicators are the data when a player's momentum is about to change at a certain moment) and constructed the indicator framework as shown in Table. 1.

Table 1: Athlete indicator frame table

Player	St	Pv	Pa	Pw	Pdf	Pue
Carlos Alcaraz	1	2	0	0	0	0
Alexander Zverev	2	2	0	0	0	0
Nicolas Jarry	1	1	0	1	0	0
Holger Rune	1	2	0	0	0	1
Tommy Paul	1	2	0	0	0	1
Mikael Ymer	1	2	0	0	0	1
Guido Pella	1	1	0	0	0	0

Step 2: Use the indicators in step 1 to build a matrix and carry out forward processing, that is, all indicator types are converted into extremely large indicators.

Deterministic index matrix:$X = [X_1, X_2, \cdots, X_n]$.

Each column has an indicator, and each behavior has an individual value in a different indicator.

$$X = \begin{bmatrix} X_{11}, X_{12}, X_{13}, \cdots, X_{1n} \\ X_{21}, X_{22}, X_{23}, \cdots, X_{2n} \\ X_{31}, X_{32}, X_{33}, \cdots, X_{3n} \\ \vdots \\ X_{n1}, X_{n2}, X_{n3}, \cdots, X_{nn} \end{bmatrix} \tag{1}$$

Step 3: To eliminate the influence between different dimensions, we normalized the forward matrix. Make sure that the normalized matrix includes each value between [0,1]. The standardized formula is:

$$Z_{ij} = \frac{x_{ij} - min(x_{1j}, x_{2j}, \cdots, x_{nj})}{max(x_{1j}, x_{2j}, \cdots, x_{nj}) - min(x_{1j}, x_{2j}, \cdots, x_{nj})} \tag{2}$$

Step 4: Calculate the proportion of the i sample of the JTH indicator and obtain the information entropy and information utility value of each indicator according to the calculation and obtain the entropy weight of each indicator through normalization processing.

For the JTH index, the formula for calculating information entropy is as follows:

$$e_{ij} = -\frac{1}{logn} \sum_{i=1}^{n} p_{ij} \ln(p_{ij}), (i, j\ all < n) \tag{3}$$

The formula of information utility value is:

$$d_j = 1 - e_j \tag{4}$$

After the normalization of the information utility value, the weight of each indicator can be obtained:

$$w_j = \frac{d_j}{\sum_{j=1}^{m} d_j} \tag{5}$$

Step 5: Calculate the score and normalize the score.

Distance between the $i - th$ $(i = 1,2,3, \dots, n)$ evaluation object and the maximum value:

$$D_i^+ = \sqrt{\sum_{j=1}^{m} w_j \left(Z_j^+ - z_{ij}\right)^2} \tag{6}$$

Distance between the $i - th$ $(i = 1,2,3, \dots, n)$ evaluation object and the minimum value:

$$D_i^- = \sqrt{\sum_{j=1}^{m} w_j \left(Z_j^- - z_{ij}\right)^2} \tag{7}$$

We can then get the non-normalized score of the $i - th$ evaluated object:

ISSN 2616-5775 Vol. 7, Issue 5: 152-158, DOI: 10.25236/AJCIS.2024.070520

$$S_i = \frac{D_i^-}{D_i^+ + D_i^-} (D_i^+ smaller，S_i Closer\ to\ the\ maximum\ 1) \tag{8}$$

Then the score is normalized, and finally the final score of each player is obtained, which is a measure of momentum.

2.2 Evaluation results of Topsis model based on entropy weight method

Momentum change is a continuous process that needs to consider the momentum of the previous moment. Based on the concept of entropy, the weight value is obtained by calculating the entropy of each row of data, as shown in Table 2.

Table 2: Weight table of related indicators

Pv	0.1	Pw	0.2
St	0.3	Pdf	0.2
Pa	0.1	Pue	0.1

The momentum at time point t is called Mt, The momentum change is the product of Pv (point_victor=1, then Pv=1, otherwise it is equal to Pv=-1.1), St (server=1 then St=1.2, otherwise St=1), Ws (1 for the first win, 0.2 for each consecutive win), plus the winning serve Pa (0.02), winning ball Pw(0.01), two service errors Pdf(-0.2) and Unforced errors Pue (-0.1) Momentum change score.

The calculation formula is:

$$m = m(t-1) + Pv * (St + Ws) + Pa + Pw + Pdf + Pue \tag{9}$$

The momentum of players in the game can be calculated through the formula, and the game data is brought into the calculation. This paper takes the 2023-wimbledon-1301 game as an example to calculate the momentum of players in the game, as shown in the Figure 1 and Figure 2.

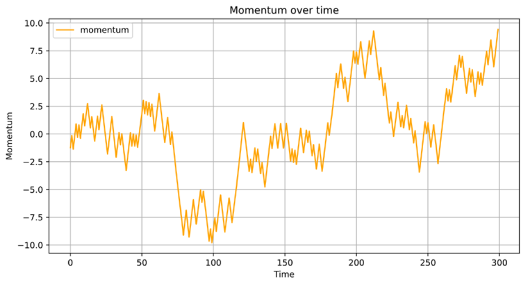

Figure 1: Chart of overall momentum change over the course of a winner's race

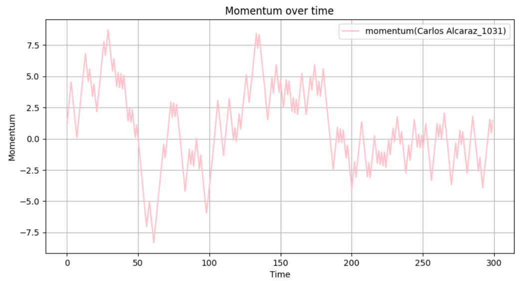

Figure 2: Player Carlos Alcaraz 1031 momentum change chart

As you can see from the figures above, in the first game of the first set, Nicolas Jarry had less momentum than Carlos Alcaraz. Carlos Alcaraz quickly won, and then the momentum reversed in the second game of the first set. We can see that Carlos Alcaraz made several service errors and made several unforced errors, to a certain extent, the subjective factors of the player adversely affected, and at the same time, the lack of priority serve to create a chance to win, Carlos Alcaraz momentum decreased. Thus, our model can capture the flow of points during a match. In this way, players can be quantitatively analyzed and evaluated, and the momentum change of players during the game can be clearly reflected and the visual expression of the game flow can be realized.

ISSN 2616-5775 Vol. 7, Issue 5: 152-158, DOI: 10.25236/AJCIS.2024.070520

3. Analysis of the relationship between momentum and winning

To determine whether there is a random relationship between player "momentum" and player wins, We selected the winner (point_victor), serve (server), and winning streak (calculated according to the data), and the secondary factors were pl_ace (winner), pl_winner (winner), two service errors (pl_double_fault), and unforced errors (pl_unf_err) After quantifying player momentum during the match, correlation analysis was performed with Number_of_winning_streaks. Through the analysis, we found that the variance of each index follows a normal distribution, and the variance among variables is homogeneous, and there is a linear and continuous relationship between the two variables. Therefore, we decided to adopt Pearson correlation coefficient method for correlation analysis [3].

$$r_{XY} = \frac{Cov(X,Y)}{S_X S_Y} \tag{10}$$

Among them: $S_x = \sqrt{\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{n-1}}$, $S_y = \sqrt{\frac{\sum_{i=1}^{n}(Y_i-\overline{Y})^2}{n-1}}$, sample mean $\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n}$,$\overline{Y} = \frac{\sum_{i=1}^{n} y_i}{n}$, sample covariance $Cov(X,Y) = \frac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{n-1}$.

According to the above analysis, this paper selects the number of consecutive scores, the score difference with opponents, and the number of consecutive wins as momentum indicators to conduct correlation analysis with Number_of_winning_streaks, and the results are shown in the following Table. 3.

Table 3: Correlation analysis results table

	point_victor	p1_ace	p1_unf_err	p1_double_fault	server	p1_winner	momentum
Point _victor	1	-.194**	.352**	.163**	.300**	-.377**	-.159**
p1_ace	-.194**	1	-0.068	-0.032	-.203**	.513**	0.022
p1_unf_err	.352**	-0.068	1	.464**	-.153**	-.133*	-0.042
p1_double_fault	.163**	-0.032	.464**	1	-.154**	-0.061	0.040
server	.300**	-.203**	-.153**	-.154**	1	-.240**	-0.043
p1_winner	-.377**	.513**	-.133*	-0.061	-.240**	1	0.088
momentum	-.159**	0.022	-0.042	0.040	-0.043	0.088	1

**. Correlation is significant at the 0.01 level (2-tailed). *. Correlation is significant at the 0.05 level (2-tailed).

As can be seen from the above table, the number of consecutive scores and the score difference with opponents is significantly correlated with Number_of_winning_streaks. In which swings in play and runs of success by one player are not random. In which swings in play and runs of success by one player are not random. So, the coach underestimated the role of momentum in a game.

The quantified momentum data was visualized with the players' scores, and the data of 1301 and 1302 games were selected as the research objects. The results in Figure 3 and Figure 4 showed that the trend of change had a strong similarity, indicating that the change of momentum was not random, but had a relatively high correlation with the players' scores.

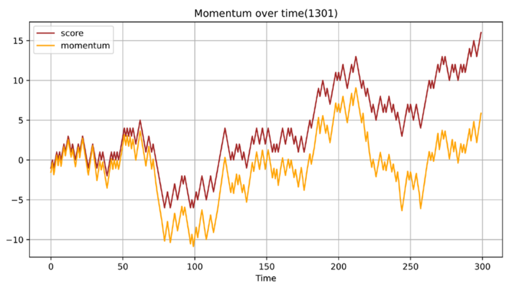

Figure 3: 2023-wimbledon-1301 Momentum and player scores

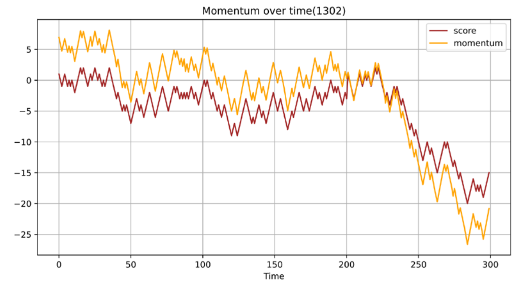

Figure 4: 2023-wimbledon-1302 Momentum and player scores

4. Analysis of momentum influencing factors

4.1 Establishment of the neural network model

To explore the influencing factors of momentum, we first need to predict the status of players corresponding to different momentums. We chose the number of consecutive wins, the number of consecutive winning games and the difference between the two players as input indicators. Whether the winning streak can continue next time is One or two categorical variables are used as output indicators. By comparing different models, we found that the CRNN Neural Network Model can accurately and effectively solve this problem [4]. Through Neural Network training, we can get the player's scoring status. After obtaining the relevant information, we can perform error analysis to obtain the accuracy of the model. Finally, by counting the characteristics of relevant indicators when the player's status changes, the factors that affect the change of the player's momentum are obtained, and suggestions are given to the players.

Step 1: Initialize the parameters, that is, initialize the weights and biases of each node, to obtain more accurate and ideal results during network training.

Step 2: Build the network for the CNN part and RNN in CRNN.

Step 3: Select appropriate neural network layers and activation functions and set appropriate hyperparameters. Hyperparameters include learning rate, batch size, number of iterations, etc.

Step 4: Use the training data to train the model. Continuously adjust hyperparameters during the training process to achieve the best training effect. After training, the model is tested using test data to evaluate the accuracy and generalization ability of the model.

Step 5: Apply the trained model to actual predictions and optimize and adjust it.

4.2 Establishment of logistic regression model

Logistic regression is a binary classification algorithm suitable for predicting whether the situation will change or remain unchanged [5]. Logistic regression models can predict these changes in advance based on various factors within the game.

A logistic regression model was constructed by selecting server, pl _ points _ won, pl _ winner, pl _double_ault, pl _ unf _ err.

4.3 Model results

This paper seeks to determine if there are some indicators that can help determine when momentum is about to shift in favor of one player to another. As analyzed above, we learned that the scoring situation is significantly related to the momentum, so we believe that the transition of momentum is the moment when the player will win consecutively in the next game. The number of consecutive wins, the number of consecutive points, and the difference in player scores can all have an impact on momentum as dependent variables. The choice is then made to use the number of consecutive wins, the number of consecutive scores, and the difference in player scores as momentum indicators, and whether a score is scored in the next game as the state quantity of the game. Continuous scoring is recorded as 1, and non-consecutive scoring is recorded as -1. We select data from 1301 games as the indicator.

ISSN 2616-5775 Vol. 7, Issue 5: 152-158, DOI: 10.25236/AJCIS.2024.070520

Then perform correlation analysis using Pearson correlation coefficient and obtain the following Table. 4.

Table 4: Pearson correlation coefficient

Correlations				
	Number of winning streaks	Scored continually	Winning streak	Player score difference
Number of winning streaks	1	-.114*	-0.017	.238**
Scored continually	-.114*	1	0.029	.138*
Winning streak	-0.017	0.029	1	-.183**
Player score difference	.238**	.138*	-.183**	1

*. Correlation is significant at the 0.05 level (two-tailed). **. Correlation is significant at the 0.01 level (two-tailed).

According to the above table, we can see that the number of consecutive wins, the number of consecutive points, the score difference value and other indicators have a significant correlation with whether the next game has a winning streak. The transition of momentum is the moment when a player wins the next game in a row. This assumption is correct. Based on this, we believe that the player's game results can be predicted through momentum. We choose to predict through the CRNN neural network model to obtain the player's subsequent scoring status. The actual results are in good agreement with the predicted results. The prediction result graph and the actual result graph, the prediction error graph is as shown in Figure 5 below:

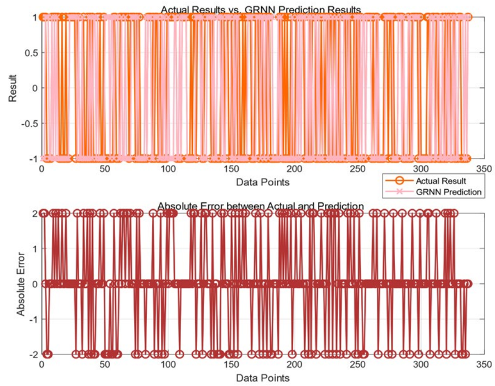

Figure 5: Neural network prediction result chart

To better verify the accuracy of our predictions, we use Logistic Regression for secondary predictions. First, judge the relevant factors. Through the correlation coefficient matrix and heat map (as shown in Figure 6), you can see the correlation between each feature and 'momentum_diff'. The heat map uses color depth to express the strength of the correlation. Dark colors indicate positive correlation, and light colors indicate negative correlation. This can help coaches analyze relationships between data features.

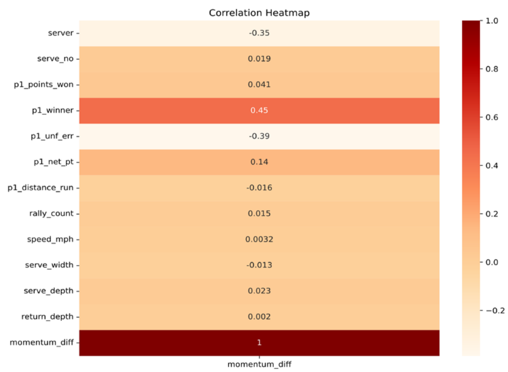

Figure 6: Correlation analysis between multiple features and player music

ISSN 2616-5775 Vol. 7, Issue 5: 152-158, DOI: 10.25236/AJCIS.2024.070520

A logistic regression model is then used to predict changes in the 'momentum' data. First, through the difference operation, the difference of the 'momentum' column is calculated and stored in the 'momentum_diff' column. Then, the target variable $'y'$ is determined based on the sign of the 'momentum_diff' column, with values greater than 0 being assigned a value of 1 and values less than or equal to 0 being assigned a value of -1. Split the feature variable data set $'X'$ and the target variable $'y'$ into the training set and the test set. The test set accounts for 20% and the random seed is 42. Initialize the logistic regression model and fit the model on the training set. Finally, we get Model accuracy: 0.7844886753603294.

4.4 Competition suggestions

As a coach, it is crucial to focus on tactics. Players are reminded to stay focused, pay attention to the details of every hitting action, and be fully prepared to face situations where they may lose points, to reduce service errors and the psychological impact and unforced errors they bring. When the point difference is large or the opponent has a winning streak, it is especially important to adjust the mentality, because this is the moment when the opponent has a strong momentum, and it is easier for us to lose points. In this case, effective measures need to be taken to deal with it.

5. Conclusions

Through the in-depth discussion in this study, we successfully adopted the Topsis model and the Pearson correlation coefficient method to quantify and validate the mechanism of the influence of qi, which provides coaches with more bases for assessing and analyzing the state of players. In addition, using correlation analysis and CRNN neural network model, we successfully predicted the fluctuation of qi and verified it through logistic regression model, achieving a considerable accuracy rate. This study provides effective tactical customization and training suggestions for coaches and players, which is expected to help optimize game performance and improve athleticism. In conclusion, this study not only theoretically deepens the understanding of the mechanism of the role of momentum in tennis matches, but also provides new perspectives and methods for the analysis of sports match data and player training.

References

[1] Hu Y, Wang Z, Li X. Impact of policies on electric vehicle diffusion: An evolutionary game of small world network analysis [J]. Journal of Cleaner Production, 2020, 265: 121703.
[2] Tian C, Song M, Tian J, et al. Evaluation of Air Combat Control Ability Based on Eye Movement Indicators and Combination Weighting GRA-TOPSIS[J]. Aerospace, 2023, 10(5): 437.
[3] Armstrong R A. Should Pearson's correlation coefficient be avoided? [J]. Ophthalmic and Physiological Optics, 2019, 39(5): 316-327.
[4] Fu X, Ch'ng E, Aickelin U, et al. CRNN: a joint neural network for redundancy detection[C]//2017 IEEE international conference on smart computing (SMARTCOMP). IEEE, 2017: 1-8.
[5] Landwehr N, Hall M, Frank E. Logistic model trees[J]. Machine learning, 2005, 59: 161-205.