Research on momentum conversion in sports

Yakun Zhang¹, Ziren Qi², Yufeng Ye³
¹Hefei University of Technology, Hefei, 230601, China
²Taiyuan University of Technology, TaiYuan Shanxi, 030000, China
³Furen International School, Singapore, 8 Claymore Hill, Singapore 229572, Singapore

Yakun Zhang, Ziren Qi, Yufeng Ye are co-first authors.

Abstract. Momentum, traditionally examined in sports, is an ability to increase or develop continuously, impacting mental and physical performance in a pattern akin to human preferences, which are complex and non-linear. This research explores momentum in sports, particularly tennis, integrating it with the economic concept of diminishing marginal utility to analyze momentum shifts. Using data from the 2023 Wimbledon men's singles, the study dissects momentum into absolute and relative measures, revealing server advantage and precise models with an R-square value of 0.877. It employs macro and micro perspectives to link momentum with success, affirming its significant impact on win rates and performance dimensions. A GARCH model identifies influential factors in momentum transitions, informing strategies for game control. Applying the model to basketball, specifically a 2024 Clippers-Celtics NBA game, maintains strong predictive capabilities, indicating its broader applicability. Practical recommendations for tennis coaches and players, grounded in empirical findings, conclude the study, emphasizing strategic momentum utilization.

Keywords: Diminishing Marginal Utility; XGBoost Model; Tennis; Comprehensive Analysis of Rank Sum Ratio.

1. Introduction

In the 2023 Wimbledon men's singles final, 20-year-old Spanish rising star Carlos Alcaraz takes on 36-year-old veteran Novak Djokovic. After the first four games, the two sides were tied, Carlos Alcaraz won the fifth game, giving Djokovic his first Wimbledon loss since 2013. During the game, some of the best performances tend to happen to players with more "momentum". And that got us thinking.

In our review of the literature, we discovered that many researchers had discussed the effect of momentum on various phenomena that could be defined as concepts of gain and loss, such as sporting events and even stock investments.[1] Psychological momentum as a gain in or gained psychological strength which changes interpersonal perceptions and affects an individual's mental and physical performance. In addition, psychological momentum has a certain moderating effect on success, it will significantly increase future success and facilitate the achievement of goals.[2]

In psychology, there exists a concept known as the theory of internal and external factors. This theory explores the dynamic internal aspects of behavior based on the outcomes of human actions. External factors are identified as the conditions that facilitate the existence and development of entities, while internal factors are the fundamental causes that drive the development and change of these entities.[3] In the context of this discussion, factors such as scoring are external, acting as triggering conditions for motivation. On the other hand, a player's 'momentum' represents an internal factor, serving as the foundational basis of motivation. Both internal and external factors interact with and influence each other.[4]

Inspired by this, we hypothesized that by this concept, the psychological factor of "momentum" would have an impact on the game, possibly disrupting the opponent's rhythm or even turning the game around. However, in the traditional sense, "momentum" is a qualitative indicator. We want to quantify "momentum" by modeling factors such as serving direction, the occurrence of unreturnable shots by the opponent (the ball that cannot be caught by the opponent) and the number of shots in
each round, so as to analyze how it causes fluctuations in the match and affects the result of the match, and explore whether "momentum" is predictable. Whether it's just the product of chance.

![FIG. 1: Structure of Our Work](image)

2. Related Work

2.1 Model Assumption

**Hypothesis 1.** Momentum is an abstract and relative concept.

**Hypothesis 2.** The outcome of a single game and the number of turns in the game represent relative momentum.

**Hypothesis 3.** The marginal utility of momentum is variable.

**Hypothesis 4.** Absolute momentum can moderate relative momentum.

**Hypothesis 5.** Momentum is an abstract and relative concept.

2.2 Data Collection and Description

In this paper, We used three datasets, sifted through the data, and described it.

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Data description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ace</td>
<td>hit an untouchable winning serve</td>
</tr>
<tr>
<td>winner</td>
<td>hit an untouchable winning shot</td>
</tr>
<tr>
<td>double_fault</td>
<td>missed both serves and lost the point</td>
</tr>
<tr>
<td>unf_err</td>
<td>made an unforced error</td>
</tr>
<tr>
<td>net_pt_won</td>
<td>won the point while at the net</td>
</tr>
<tr>
<td>break_pt_won</td>
<td>won the game player 2 is serving</td>
</tr>
<tr>
<td>break_pt_missed</td>
<td>missed an opportunity to win a game player 2 is serving</td>
</tr>
<tr>
<td>distance_run</td>
<td>player 2's distance ran during point (meters)</td>
</tr>
<tr>
<td>rally_count</td>
<td>number of shots during the point</td>
</tr>
<tr>
<td>speed_mph</td>
<td>speed of serve (miles per hour; mph)</td>
</tr>
<tr>
<td>Dp</td>
<td>Number of game rounds</td>
</tr>
<tr>
<td>W</td>
<td>Outcome of the game</td>
</tr>
</tbody>
</table>

Data source: Tencent Sports database

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Data description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Game scoring rate</td>
<td>Count the BOS scoring sessions in a single game/6</td>
</tr>
</tbody>
</table>
The number of times the game passes the ball
The number of game fouls
the number of game offense
Game dunk points
Game error rate

Count the number of balls passed by BOS in a game
Count the number of fouls committed by BOS in a game
Count the number of attacks made by BOS in a game
BOS dunks score points in a single game
BOS Error turn in a single game/6

Data source: https://www.atptour.com/en

Data source: https://www.atptour.com/en

2.3 Player Performance Evaluation

2.3.1 Defining Momentum

2.3.1.1 Defining the relative momentum of player1

We denote the number of rounds in Player 1's ith game as \( D_{pi} \), the outcome as \( W_i \) (assigning 1 for a win and -1 for a loss), and the relative momentum as \( m_{ri} \). The difference in player ability is represented by \( a \) (where \( a > 1 \) reflects a difference in athletic advantage, and at \( a = 1 \), the model reverts to a linear form). In this context, \( a \) is set to 1.2 and is treated as an exogenous variable.

The scoring rate is not considered in this model because winning a game is not strictly correlated with the score, which could lead to misleading judgments about the overall game outcome. Our table illustrates that the final scores can be identical while the winner may differ, often due to the number of rounds in each game. This rationale underpins our decision to base momentum on the duration of individual game rounds.

<table>
<thead>
<tr>
<th>Game number</th>
<th>player1 Score situation</th>
<th>Player2 Score situation</th>
<th>winner</th>
</tr>
</thead>
<tbody>
<tr>
<td>2023-wimbledon-1304</td>
<td>169</td>
<td>168 (+1)</td>
<td>Player2</td>
</tr>
</tbody>
</table>

Based on the above discussion, we construct the following model:

\[ m_{ri} = W_i \cdot d^{(W_i \cdot D_{pi} + 44)} \]

\[ D_{pi} = 4, 5, ..., n \]

2.3.1.2 Moderating effect of absolute momentum

Let's define the ranking of player1 as \( R_1 \) and the ranking of player2 as \( R_2 \). Since we assume that momentum is inherently a relative concept, we want to limit the moderating effects of absolute momentum (K):

\[ K_i = \left( \frac{\ln R_i}{\ln R_1} \right)^{W_i} \]

2.3.1.3 The relative momentum under absolute momentum adjustment (referred to as "momentum" hereafter)

\[ M_{ri} = K_i \cdot m_{ri} \]

To make our function more intuitive, we plotted the momentum of player1 for 2023-wimbledon-1301 match as a function of the degree of backpressure:
We choose to analyze momentum at the game level rather than at the set, match, or point levels for the following reasons:

(1) A more precise time frame allows for a more flexible reflection of momentum changes. The round level lacks explanatory power over extended periods, such as entire sets or matches. We believe that 'momentum' is at least a state that can be sustained over the short to medium term, not merely a fleeting occurrence.

(2) Sets and matches provide weak interpretations and are inflexible in capturing short-term momentum changes. They also tend to average out a lot of data within each group. In this article, 'rally_count' is defined as a concept similar to the degree of suppression, but it is applied at the turn level. Furthermore, metrics such as winning streaks are inherently based on turn-by-turn performance.

2.3.2 Evaluate a Player's Performance by Momentum

To create a meaningful classification, we followed the prompts to distinguish players based on whether player1 is serving. As the server often has an easier time establishing an advantage, we hypothesize that the server also tends to have a better suppression degree.

FIG. 3: Box Chart

We make player1 serve Mr1 and player1 receive Mr2. When M1 serves, there is little difference in momentum stability, but the overall performance is significantly better than its performance at the moment of receiving the ball. Moreover, we can see that both sets of suppression are completely one-way in 25% to 75% of the data, which also shows that it is easy to suppress the opponent through service rallies.

2.3.3 Evaluate a player's performance by other variables

Then we use RSR comprehensive evaluation method to display some evaluation results:

RSR comprehensive evaluation is based on the original index data matrix composed of evaluation objects and evaluation indicators:

\[
X = \begin{pmatrix}
X_{11} & \cdots & X_{1p} \\
\vdots & \ddots & \vdots \\
X_{n1} & \cdots & X_{np}
\end{pmatrix}
\]

According to each specific evaluation index according to the size of its index value, obtaining rank:
It is used to replace the original evaluation index value. Through \( WRSR_i = \frac{1}{n} \sum_{j=1}^{p} W_j R_{ij} \), WRSR, a dimensionless statistic with weight, is obtained, and the merits and demerits of evaluation objects are sorted according to the value of WRSR.

When evaluating a player's performance, relying solely on momentum is not comprehensive. After analyzing the boxplot at the game level, we proceed to examine the player's performance from additional dimensions. Initially, we employ the entropy weight method to assign weights to the selected variables. We observe that the information entropy for variables such as p1_ace, p1_net_pt_won, p1_break_pt_won, and p1_winner is relatively high. These variables are all indicative of positive performance. The diminutive weight assigned to negative performance indicators may be attributed to the prevalence of '0' values in the data, suggesting that the athletes generally maintain good control and commit fewer errors.

Table 3: Entropy Weight Method is Used to Calculate the Weight

<table>
<thead>
<tr>
<th>Entropy weight method</th>
<th>item</th>
<th>Information entropy e</th>
<th>Information utility value d</th>
<th>weight(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p1_ace</td>
<td>0.66</td>
<td>0.34</td>
<td>25.765</td>
</tr>
<tr>
<td></td>
<td>p1_net_pt_won</td>
<td>0.71</td>
<td>0.29</td>
<td>21.964</td>
</tr>
<tr>
<td></td>
<td>p1_break_pt_won</td>
<td>0.527</td>
<td>0.473</td>
<td>35.808</td>
</tr>
<tr>
<td></td>
<td>p1_winner</td>
<td>0.802</td>
<td>0.198</td>
<td>14.963</td>
</tr>
<tr>
<td></td>
<td>p1_double_fault</td>
<td>0.998</td>
<td>0.002</td>
<td>0.144</td>
</tr>
<tr>
<td></td>
<td>p1_unf_err</td>
<td>0.985</td>
<td>0.015</td>
<td>1.148</td>
</tr>
<tr>
<td></td>
<td>p1_break_pt_missed</td>
<td>0.997</td>
<td>0.003</td>
<td>0.208</td>
</tr>
</tbody>
</table>

Below, we take Probit, the probability unit value corresponding to the cumulative frequency, as the independent variable and RSR as the dependent variable to calculate the regression equation. The specific evaluation results are shown in the table. According to the analysis of the results of the F test, the significance P value is 0.000***, showing horizontal significance, rejecting the original hypothesis that the regression coefficient is 0. Meanwhile, the goodness of fit R² of the model is 0.887, and the model performs well, so the model basically meets the requirements. For the collinearity of variables, VIF is all less than 10, so the model has no multicollinearity problem, and the model is well constructed. The formula of the model is as follows:

\[
RSR = -0.388 + 0.115 \cdot \text{Probit}
\]

Table 4: Linear Regression Analysis Results (n=13)

<table>
<thead>
<tr>
<th>t</th>
<th>P</th>
<th>VIF</th>
<th>R²</th>
<th>Adjust R²</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>-5.196</td>
<td>0.000***</td>
<td>-</td>
<td>0.887</td>
<td>0.877</td>
<td>F=86.712 P=0.000***</td>
</tr>
<tr>
<td>9.312</td>
<td>0.000***</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dependent variable: RSR
Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

The ranking table is used to order the evaluation subjects based on the RSR (Relative Standard Ratio) estimates derived from the regression equation. In this instance, we have designated the number of ranking positions as four. The aim of this exercise is to map the data onto a normal distribution curve according to the various rank cases and to employ the correlation division method.
associated with normal distribution for classification purposes. By comparing the RSR fit values with the critical RSR fit values from the previous table, we determine the grade level. A higher number indicates a higher level, signifying better performance.

Table 5: Player Performance Ranking

<table>
<thead>
<tr>
<th>Player</th>
<th>Level</th>
<th>RSR ranking</th>
<th>Probit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stan Wawrinka</td>
<td>4</td>
<td>4</td>
<td>6.927206731</td>
</tr>
<tr>
<td>Guido Pella</td>
<td>4</td>
<td>1</td>
<td>8.98093887</td>
</tr>
<tr>
<td>Carlos Alcaraz</td>
<td>3</td>
<td>6</td>
<td>6.310562867</td>
</tr>
<tr>
<td>Andrey Rublev</td>
<td>2</td>
<td>11</td>
<td>3.55433142</td>
</tr>
<tr>
<td>Daniel Elahi Galan</td>
<td>1</td>
<td>13</td>
<td>2.206674316</td>
</tr>
</tbody>
</table>

3. Randomness Test of Momentum

Many coaches subscribe to the belief that a player's win rate is the result of regular study and practice, rather than the abstract concept of momentum. They contend that a player's performance on the field is an objective factor and dismiss the notions of 'condition' and 'momentum' as non-existent. This view implies that momentum is a normally distributed variable. To test this hypothesis, we first plotted a QQ plot for an initial, intuitive assessment. The result of the experiments show that a stronger suppression of the opponent correlates with a greater overall control of the game, allowing for more impressive performance in each round and every game. However, when considering the overall influence on the game, the coaches' assessments of 'momentum' do not provide a reliable indicator for the macro dynamics of the game.

4. Prediction Model Building and Eigenvalue Ranking

4.1 Observation of momentum volatility based on GARCH model

We introduced GARCH model on the basis of ARCH model, introduced higher order terms of variance residuals, simulated changes in the volatility of time series variables, and described the variance structure in time series data more accurately, in the following form:

\[ y_t = \mu_t + \epsilon_t \]

\[ \epsilon_t = \sigma_t Z_t \]

\[ \sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_j \epsilon_{t-j}^2 \]

Where is the observed value of \( y_t \) time series, \( \mu_t \) is the mean value, \( \epsilon_t \) is the normalized residual of time point \( t \), \( \sigma_t^2 \) is the variance of time point \( t \), \( Z_t \) is the standard normal random variable, \( \omega \), \( \alpha_1, \cdots, \alpha_p \), \( \beta_1, \cdots, \beta_q \) is the GARCH model parameter, \( \omega \) is the constant term, \( \alpha \) controls the influence of the past residual term, and \( \beta \) controls the influence of the past variance term.

We used the GARCH model to observe the time nodes with large momentum fluctuations of player1 in 2023-wimbledon-1301 match.
We can observe that during a momentum upswing, the ball's speed is significantly higher than the overall average, the rally count is lower, there are no errors, and the running distance is shorter compared to a momentum downswing. Additionally, the break rate is much higher than average, with no service errors occurring. Conversely, during a period of momentum decline, the ball's speed noticeably decreases compared to the upswing period, the number of rallies exceeds the overall average, errors are frequent, the running distance is longer, the break rate drops to zero, and service errors are above average.

4.2 Establishing The XGBoost prediction model

After training the model, our initial step is to assess its fit. To circumvent overfitting due to unreasonable data set partitioning, we employ cross-validation. This process involves dividing the data set into k mutually exclusive subsets of approximately equal size. Then, for each of k iterations, we use the union of k-1 subsets as the training set and the remaining subset as the test set. After completing the k training and testing cycles, we calculate the mean of the k evaluation results.

The closer the prediction is to 1, the more accurate the model is when compared with the mean alone. We can see that the $R^2$ of the test set reaches 0.776, indicating that the model fits well and can be used to make predictions.

<table>
<thead>
<tr>
<th></th>
<th>MSE</th>
<th>MAE</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training set</td>
<td>0.033</td>
<td>0.182</td>
<td>0.998</td>
</tr>
<tr>
<td>Test set</td>
<td>0.177</td>
<td>0.216</td>
<td>0.776</td>
</tr>
</tbody>
</table>

4.3 Feature ranking and prediction based on XGBoost model

4.3.1 Feature importance ranking

![Feature Importance Ranking](FIG. 5: Feature Importance Ranking)
We selected the matches from 2023 Wimbledon, specifically those numbered 1301 to 1316, as our sample for further research. To determine which factors were most relevant, we employed the XGBoost model to rank feature importance. Our findings suggest that for a player to sustain favorable momentum, the first priority should be to minimize service errors and capitalize on opportunities to break the opponent's serve. Ball speed and the number of shots exchanged during a point also emerged as significant factors influencing momentum shifts. In contrast, the distance run appears to be less crucial than whether Player 1 commits an unforced error. However, unforced errors are not considered highly significant as they are typically well-controlled by the player (Ax means averaging within the group).

4.3.2 Model prediction

The match designated as '2023-wimbledon-1701match' in the title is the focus of our analysis. To better address the problem, we selected Novak Djokovic's performance in the aforementioned match for our test case. This selection was made to confirm that the model can accurately predict the immediate momentum shifts between the two players. The preliminary analysis indicates that the momentum is continuously shifting between the players, making it challenging to discern who has control of the match, as there is no clear visual representation of this dynamic. To better understand the situation, we will begin by visualizing the cumulative momentum to determine who holds the advantage throughout the match.

![FIG. 6: Novak Djokovic Change in the Predicted and Actual Value of the Instant Momentum](image)

We can observe that both the model's predictions and the actual momentum provide compelling explanations for the on-court events. The match outcome shows that Novak Djokovic began strongly; however, despite a late rally, he was defeated in three of the subsequent four sets. Similarly, our model's momentum predictions indicated an initially strong performance by Djokovic. Yet, even with a later rally, his momentum was consistently countered by Carlos Alcaraz, ultimately resulting in Djokovic's loss in the match.

![FIG. 7: Novak Djokovic Cumulative Momentum Predicted and Actual Trend](image)

Based on our previous statements, momentum shifts should also include momentum reversals. So we set all the values that turn negative to -1, all the values that turn positive to 1, and just watch the direction of the momentum change. We found that the success rate of the model to predict the
momentum change was 91.9%, and the success rate of the model to predict the momentum change was 94.4%, and the accuracy of the model to predict the immediate momentum change was high.

![Heat Plot Predicted by Novak Djokovic Momentum Reversal](image)

**FIG. 8: Heat Plot Predicted by Novak Djokovic Momentum Reversal**

### 4.4 Advice to the players

By using GARCH model, we respectively compared some behavioral factors in the momentum upswing period and the momentum downswing period, and we used XGBoost model to rank the importance of different factors affecting the game fluctuations. Therefore, we give the following suggestions:

1. **Optimize service return strategy:**
   - The error rate of service return and service break rate are important factors in the momentum transformation of the match.
   - The server often holds the initiative in a match. For players experiencing rising momentum, optimizing the stability of their serve and improving their break percentage are effective strategies for maintaining that upward trajectory. Conversely, for players facing a decline in momentum, stabilizing their mental state is crucial. They also need to reduce their error rate on the return of serve and enhance their own break percentage to create opportunities to overcome their opponents.

2. **Adjust the pace of the game:**
   - The number of catch turns in the same game is also an important factor in the change of momentum.
   - Therefore, if a player is in the period of rising momentum, he should speed up the pace of the game, while in the period of declining momentum, he should slow down the pace of the game as much as possible to disrupt the rhythm of the opponent.

3. **Optimize service speed:**
   - We understand that a higher stroke speed can boost a player's momentum by reducing the opponent's reaction time. By increasing the speed of the shot, the player makes it more difficult for the opponent to return the ball accurately. Therefore, coaches may advise players to concentrate on the ball's speed when they have rising momentum, as this can be an effective tactic to quickly outplay the opponent.

4. **Depleting the opponent's stamina:**
   - In our research, we have found that players in a period of rising momentum tend to cover less distance than those experiencing a decline in momentum. Therefore, players with rising momentum could strategically adjust their serving angle and direction to increase their opponents' running distance and deplete their physical strength, thereby gaining better control of the game.

5. **Conducting simulated match training:**
   - During player training, athletes can simulate in-game momentum shifts and practice coping strategies for various momentum scenarios. This approach can help stabilize players' psychological states in actual matches, enabling them to handle such situations more effectively.

6. **Use the characteristics of different players to carry out targeted strategies:**
   - By consulting data on players like Rafael Nadal from Spain and Novak Djokovic from Serbia, who are adept at playing long matches, we can draw certain conclusions and make recommendations based
on data analysis. However, it is often necessary to adopt a tailored coaching approach, 'teaching according to the student's aptitude,' to fully leverage each player's individual potential and secure a win.

4.5 The Universality Test of The Model

To test the universality of our model, we analyzed the game between the Los Angeles Clippers (LAC) and the Boston Celtics (BOS) on January 27, 2024. The Clippers triumphed over the Celtics 115-96 in an away game, handing the Celtics their second home loss of the season and their most significant home defeat to date.

Similar to the 2023 Wimbledon match 1701, we hypothesize that the Clippers' significant road victory can be attributed to their momentum. In this article, we extend the concept of momentum to encompass the entire NBA team, rather than just individual players. Given the shorter duration of NBA games, we propose a new metric: every six exchanges of the ball are equated to a 'game', with the score representing the degree of dominance, and the team ranking serving as a proxy for player ranking.

Due to the substantial differences between basketball and tennis in terms of rules, skill dynamics, and the nature of the games, we have adapted tennis metrics to align with basketball. The basketball equivalents are presented in the following table:

<table>
<thead>
<tr>
<th>tennis</th>
<th>basketball</th>
<th>Calculation method</th>
</tr>
</thead>
<tbody>
<tr>
<td>speed_mph</td>
<td>Game scoring rate</td>
<td>Count the BOS scoring sessions in a single game/6</td>
</tr>
<tr>
<td>rally_count</td>
<td>The number of times the game passes the ball</td>
<td>Count the number of balls passed by BOS in a game</td>
</tr>
<tr>
<td>unf_err</td>
<td>The number of game fouls</td>
<td>Count the number of fouls committed by BOS in a game</td>
</tr>
<tr>
<td>distance_run</td>
<td>The number of game offense</td>
<td>Count the number of attacks made by BOS in a game</td>
</tr>
<tr>
<td>break_pt_won</td>
<td>Game dunk points</td>
<td>BOS dunks score points in a single game</td>
</tr>
<tr>
<td>double_fault</td>
<td>Game error rate</td>
<td>BOS Error turn in a single game/6</td>
</tr>
</tbody>
</table>

Since the fourth quarter of this game entered garbage time (where the score gap was too large, and neither team played their main players), the data from this period lacked statistical significance. Therefore, we selected data from only the first three quarters for our analysis to verify the accuracy of our prediction model. The results showed that the predicted momentum fluctuations closely matched the actual values, consistent with the game's progression. In the early stages, Boston Celtics (BOS) and Los Angeles Clippers (LAC) alternated leads, and for a time, BOS had the higher momentum. However, BOS's momentum significantly declined in the second and third quarters, and the Clippers gained complete control of the game.
At the same time, we analyzed MSE, MAE and R-square to describe the prediction effect. It can be seen that the prediction effect of this model is better than that of the model test set, which proves that the prediction accuracy of the model in the momentum fluctuation in basketball matches is improved.

Table 8: Model evaluation

<table>
<thead>
<tr>
<th></th>
<th>MSE</th>
<th>MAE</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test set</td>
<td>0.107</td>
<td>0.166</td>
<td>0.821</td>
</tr>
</tbody>
</table>

For momentum conversion, we can see that for basketball games, the prediction accuracy for momentum gains decreases (0.944→0.882), while the prediction accuracy for momentum losses increases (0.919→0.938). This may be related to the sample, because BOS rarely has a trend of upward momentum, so a mistake in model prediction will have a great impact on the whole prediction accuracy. Of course, it may also include some future factors, such as physical consumption, the change of the opponent's tactics, etc., but in general, the prediction result is relatively accurate.

5. Model advantages and disadvantages

The model presents several advantages and innovations: it ensures accuracy in predictions, introduces the concept of diminishing marginal utility of momentum, incorporates Wimbledon rankings' moderating effect, and analyzes game-winning factors comprehensively. However, it has limitations, including incomplete coverage of all players, decreased prediction effectiveness when leads alternate, and concerns about the meaningfulness of entropy weighting due to numerous zero values in selected variables.

References


