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

Analyzing the Impact of Momentum on Tennis Matches Using Decision Tree Classification: A Case Study of Wimbledon

Xueyu Zhu[1, *, #], Kaixing Liu[2, #]

[1]Geography Department, Tianjin Normal University, Tianjin, China, 300000

[2]School of Economics & Management, South China Normal University, Guangzhou, China, 510006

*Corresponding author: ShirleyZhu1227@126.com

#These authors contributed equally.

Abstract. In this paper, a Momentum Scoring Model is developed to quantify the abstract concept of Momentum through a set of important factors such as the base point, the hold point, and the break point. The model has the flexibility to capture player's in-game scoring changes with high accuracy and has a visualization process that is easy for tennis coaches and tennis players to understand. In order to statistically test whether momentum score changes are beyond random fluctuations, this paper validates the data by using the permutation test and shows that momentum plays a role in tennis matches. The paper also extracted several key metrics affecting momentum shifts, such as server, p2_score, etc., through an inductive learning algorithm like decision tree. Based on the above model, this article provides some preparation methods for tennis players. Such as analyzing the opponent's playing patterns, enhancing individual psychological preparation, optimizing serving and receiving strategies, etc.

Keywords: Tennis, Momentum Model, Permutation Test, Decision Tree Classification.

1. Introduction

The concept of momentum in sports refers to the phenomenon where athletes or teams seem to experience streaks of superior performance. Understanding and quantifying momentum can provide valuable insights for strategic decision-making in sports. In tennis, this involves analyzing periods where players dominate play, possibly leading to significant impacts on match outcomes.

The study of momentum in sports has a rich history. Gilovich, Vallone, and Tversky (1985) explored the "hot hand" fallacy in basketball, challenging the idea that players' success rates increase after a series of successful plays. Despite their findings suggesting that perceived streaks are often statistically random, the concept of momentum persists in sports psychology and coaching strategies[1].Further studies have attempted to establish a more concrete understanding of momentum. For example, Silva et al. (2013) examined the psychological aspects of momentum, finding that athletes' perceptions of momentum can influence their confidence and performance[2]. Morgulev et al. (2019) study psychological momentum in NBA overtime games, finding that teams tying the game do not have a higher chance of winning in overtime, suggesting limited impact of psychological momentum on game outcomes[3]. Gifford and Bayrak (2023) use decision tree and logistic regression models to predict NFL game outcomes, identifying key variables like turnovers and rushing yards. Their models, based on data from 2002 to 2018, show high accuracy, underscoring the value of advanced analytics in sports decision-making and performance evaluation[4]. Zhao and Zhang (2023) examine the "hot hand" phenomenon in elite recurve archery, finding that hitting the bullseye generates momentum, enhancing subsequent performance. This effect is driven by psychological and physiological factors[5].

Tennis provides a unique opportunity to study momentum due to its discrete scoring system. Several studies have explored how momentum influences match outcomes and player performance. Depken, Gandar, and Shapiro (2022) conducted an empirical analysis of 66,262 professional tennis matches, focusing on strategic and psychological momentum in best-of-three contests. They found that both forms of momentum significantly impact match outcomes, with psychological momentum being particularly influential in later stages of a match[6]. Gupta, Krishnamurthy, and Deb (2024) analyze Wimbledon tennis matches, identifying key factors such as winning service points, average distance traveled, and rating points in the first set as significant predictors of match outcomes. Using a LASSO regression model, they demonstrate that first set data can effectively forecast match outcomes early. These findings offer practical insights for improving performance and strategy in sports[7].

Psychological momentum refers to the boost in confidence and performance following a series of successful actions. Meier et al. (2020) investigate the distinction between psychological and strategic momentum in men's professional tennis, utilizing data from 4,930 game-by-game observations of Grand Slam matches. The findings indicate that the probability of winning a game increases after converting a break point, with this momentum effect being negatively impacted by interruptions. This suggests that psychological momentum is the primary driver of performance increases following a break point[8]. This paper introduces innovations through the development of a quantifiable momentum scoring model, the application of statistical methods such as permutation tests to validate the model, and the use of machine learning techniques like decision tree classification to predict momentum shifts with high accuracy. These contributions provide practical strategies for players and coaches to optimize performance, supported by empirical data and visualization tools. This approach advances the understanding and application of momentum in tennis, distinguishing itself from previous studies by offering a comprehensive and actionable framework.(Source:https://www.contest.comap.com/undergraduate/contests/mcm/contests/2024/problems/)

2. Momentum Scoring Model

2.1. Background to the Momentum Scoring Model

In sports competition, Momentum refers to a team's or an individual's drive in a game. Momentum can be a positive boost in a game, helping them gain an advantage and win the game. Momentum is not just at one point in the game, but is a concept that is constantly changing as the game progresses. It will change its direction and strength. Turning points in a match can often change the direction of momentum. For example, in tennis, a significant break of serve can change the momentum of a match, causing a player who was trailing to come back from a deficit to win.

Momentum is an abstract concept. However, it is possible to quantify the concept through some key metrics in the game. Such as consecutive points, breaks of serve, successful holds, and wins of important points (e.g., tiebreakers). By modeling momentum in sports, it can help viewers better understand sports matches and help players achieve better results.

2.2. Components of the momentum model

In our momentum model, as shown in fig.1 a player's momentum score (S_{momentum}) at each scoring point can be calculated by using the following formula:

$$S_{\text{momentum}} = S_{\text{base}} + S_{\text{hold}}(\text{if hold point win}) + S_{\text{break}}(\text{if break point win}) + S_{\text{con}}(\text{if consecutive points win}) \tag{1}$$

In this formula, Base Point (S_{base}) =1 mark, Hold Point (S_{hold}) =2.7 mark, Break Point (S_{break}) =4.9 mark. When a player wins three or more points in a row, they receive an additional Momentum Bonus, Consecutive Points Bonus (S_{con}) =2 mark

Because momentum scores are cumulative, a player's total momentum score in a match is the sum of the momentum scores on each scoring point.

$$S_{total}=\sum_{i=1}^{n} S_{momentum_i} \tag{2}$$

Where n is the total number of scoring points in the match, $S_{momentum_i}$ is momentum score on the i^{th} scoring point.

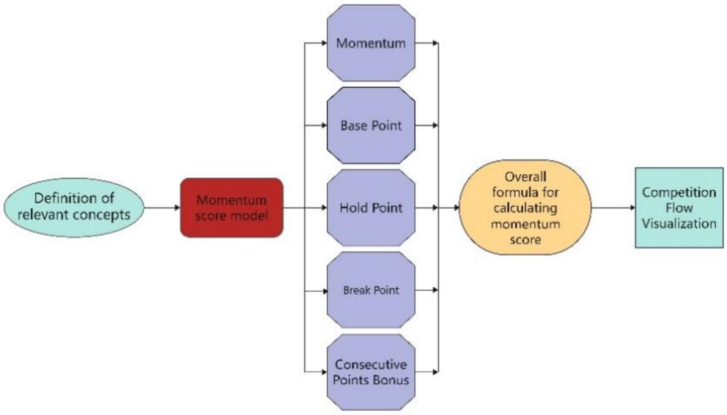

Fig.1 Flowchart

2.3. Practical Applications of the Momentum Model

This paper tested the significance of the relationship between the constructed indicator system and the labels through the statistical LOGISTIC method. On the basis of the tennis match dataset, labels for these metrics and whether a player received a score can be calculated, thus providing sufficient data to support subsequent modeling. On the basis of the derived theoretical formulas, the actual solution can then be carried out using Python.

Taking the example of Carlos Alcaraz's victory over Novak Djokovic in the men's singles final of the 2023 Wimbledon Championships in 2023-1701 as an example. With the help of the momentum model, the following momentum integral table can be derived:

Table 1 Momentum Score Table

set_no	game_no	point_no	server	complex_momentum_p1	complex_momentum_p2
1	1	1	2	0	0
1	1	2	2	5.9	0
1	1	3	2	11.8	0
1	1	4	2	11.8	1
...	...	...	...	...	...
5	10	332	1	634.9	551.2
5	10	333	1	635.9	551.2
5	10	334	1	636.9	551.2

As shown in the Table 1, the last two columns show the momentum scores of player 1 and player 2, and based on the scores, you can measure which player performed better.

2.4. Momentum Score Visualization

Still using the example of match 2023-wimbledon-1701, in the study, adding visualization code to the momentum model produced the following fig. 2.

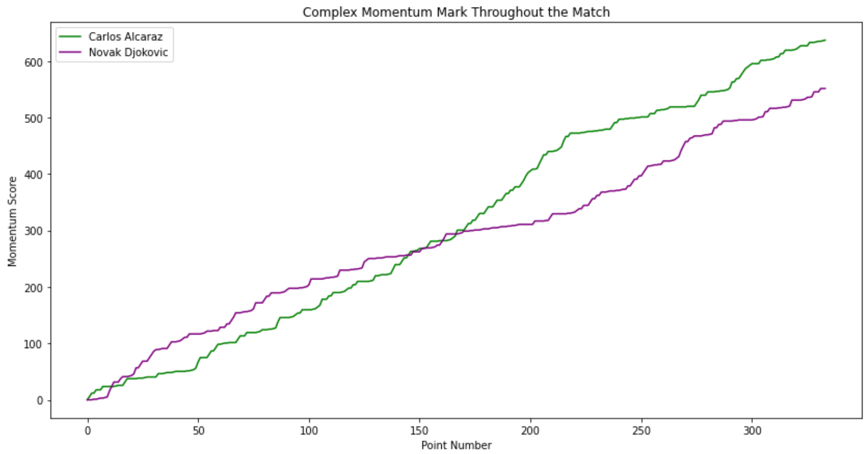

Fig.2.Momentum Score Chart for 2023-wimbledon-1701

From the fig.2, it can be concluded that the momentum change situation is very consistent with the actual match between the two, Djokovic from the beginning of the momentum advantage to the final momentum of the fire completely behind, losing the game.

This momentum model also applies to other matches. Taking 2023-wimbledon-1404 match as an example, to get the following fig.3. The momentum change also matches well with the actual match between the two, with Christopher Eubanks dominating the first two games, then falling behind and losing the match quickly.

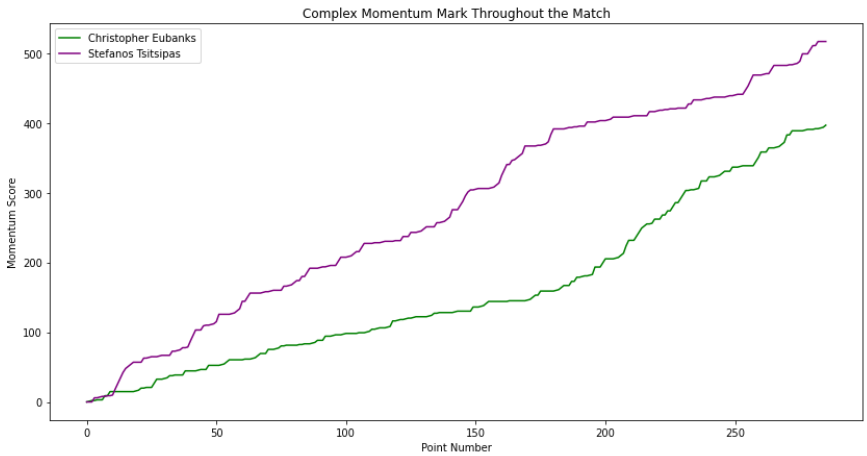

Fig.3. Momentum Score Chart for 2023-wimbledon-1404

2.5. Permutation Test

To determine whether momentum is a random phenomenon in terms of winning or losing a game, a permutation test was used to assess whether changes in momentum scores exceed the level of change expected under randomness[9]. Assuming there are two sample groups, A and B, the observed test statistic Tobs is defined as the difference between the means of the two groups. Tper represents the test statistic for each randomly permuted data set,as shown in fig. 4.

$$P=\frac{\text{quantity}(\text{Tperm}>Tobs)}{\text{The number } of\ all\ permutations} \tag{3}$$

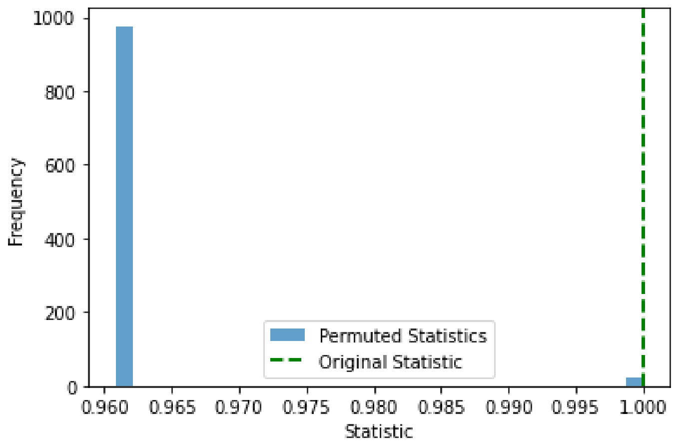

Fig.4 Permutation Test Results

The permutation test results are as above. The original Statistic is 1.0, indicating that in the original data, the percentage of maximum momentum scores aligning with the match winner is 100%. The P-value is 0.024, signifying that in all the permutation tests, the proportion of maximum momentum scores consistent with the match winner seldom matches the proportion observed in the original data. Since the P-value of 0.024 is less than the predetermined significance level of $\alpha = 0.05$, it suggests that observing a proportion of maximal momentum scores aligned with the winners as high as that in the original data is unlikely under random permutations.

This suggests that the raw observations are unlikely to have occurred by chance and that the consistency between momentum scores and match winners is statistically significant. Therefore, momentum plays a role in tennis matches and is not completely random. This indicates that momentum shifts in matches and players' winning streaks are unlikely to be merely random events, but are instead related to the actual dynamics and outcomes of the match.

3. Look for indicators that will turn the game around

3.1. Feature extraction

In this paper, match_id, set_no, game_no, point_no, server, p1_score, and p2_score were defined as base features. Additionally, consecutive_points, break_point_win, and momentum_diff were defined as high-level features.

Identifying turning points in a match is crucial for predicting changes in match dynamics. Turning points can indicate shifts in game momentum, strategy, or mental state. Turning points are labeled as instances where there is a significant change in momentum. This is determined by setting a threshold, with a momentum difference exceeding 4.9 points considered a turning point.

After the extraction process, the resulting feature table is obtained, with a portion of it displayed below Table 2 and Table 3.

Table.2. Feature table (Partial)

match_id	set_no	game_no	point_no	server	p1_score	p2_score
2023-wimbledon-1303	2	2	57	1	40	0
2023-wimbledon-1306	1	4	26	1	40	40

Table.3. Feature table (Partial)

match _id	break_ point	break_point _win_p1	break_point _win_p2	consecutive_ points_p1	consecutive_ points_p2	momentu m_diff	turning _point
2023-wimbledon-1303	0	0	0	1	2	-1	0
2023-wimbledon-1306	1	0	0	3	1	2	0

3.2. Decision Tree Classification Model

In section 4.1, feature data for all matches have been extracted. The match "2023-wimbledon-1405" will be used as an example for actual machine learning classification prediction[10].

From the Table 4 below, it can be seen that the model performs exceptionally well, achieving 99% accuracy, recall, and precision.

Table.4. Results of the model evaluation

	accuracy rate	recall rate	precision rate	F1
training set	1	1	1	1
test set	0.99	0.99	0.99	0.988

3.3. Analysis of results

The bar chart below shows fig.5 the percentage importance of each factor, highlighting their ranking in terms of influence on the direction of the match. Among these factors, the server ranks the highest, followed by p2_score, consecutive_points, break point, momentum_diff, and p1_score. These six metrics can therefore be used to determine when the flow of the match will shift.

In the model built in this paper, SERVER is found to be the most relevant factor, playing a stronger positive role on momentum when the player is serving. Conversely, when the opponent is serving, the probability of the opponent winning increases, negatively impacting the player's momentum. However, scoring a break point significantly boosts the player's momentum. Additionally, scoring points and consecutive points have a relatively strong positive effect on momentum, suggesting that players should be encouraged to score multiple times in a winning situation. Furthermore, the transition points tend to occur when a player has a decisive advantage, typically following a break of serve and occurring on their own serve, with a correlation index of $R^2 = 0.7$.

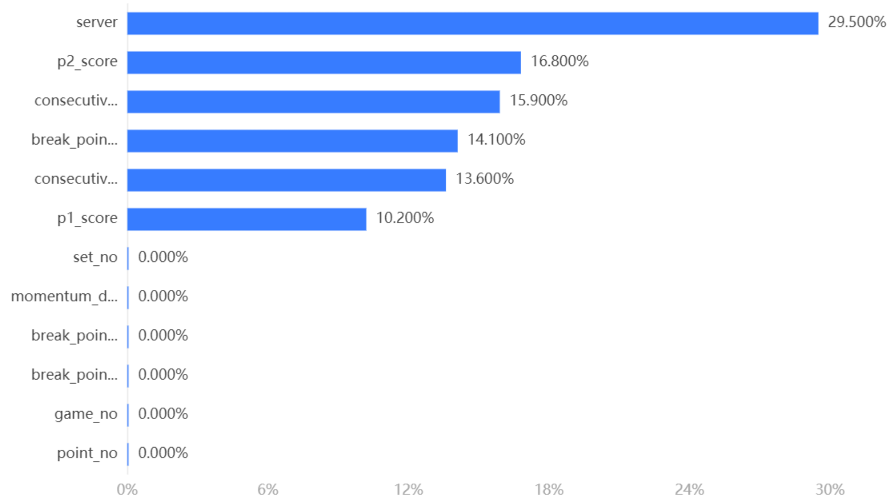

Fig.5 Importance of Characteristics

4. Match Recommendations

Predicting turning points using a decision tree classification model and achieving 99% accuracy demonstrates that the model is highly effective in capturing patterns in match data. This understanding can be utilized to identify key factors in momentum shifts and provide players with the following preparation strategies.

4.1. Analyzing opponents' playing patterns

Utilize the decision tree model to analyze key moments and triggers for momentum shifts in the opponent's past games. If the model reveals that the opponent tends to lose momentum after reaching a particular score, the coach can advise players to strategically aim for that score and capitalize on the resulting opportunity to attack.

4.2. Enhanced Psychological Preparedness

Players should be prepared to handle changes in momentum, both their own and their opponents'. Mental toughness training can aid players in remaining calm during a match, particularly when the momentum is against them.

4.3. Optimizing serving and receiving strategies

The results of the model analysis in this paper indicate that momentum shifts often occur when a player is broken. Serve and serve receive are key factors leading to momentum changes. Therefore, players need to focus on these aspects to ensure optimal performance during critical moments.

4.4. Adjusting the pace of the game

Players may need to adjust the pace of play during a match. For instance, speeding up the pace when momentum is shifting in their favor or slowing it down when momentum is against them can help disrupt their opponent's rhythm.

4.5. Analyzing Match Data

Coaches can conduct in-depth analysis of game data as the match progresses, utilizing real-time assessments from machine learning models and "momentum" reports to identify players' strengths and areas for improvement. By thoroughly understanding game data, coaches can more accurately guide their players' training and game strategies.

5. Conclusion

"Momentum" is often viewed as a difficult quantity to measure and control in sports. Momentum can play a constructive role in a game, helping a player gain an advantage and win the game. Therefore, it is extremely important to study the role of momentum in sports competitions. This study build a momentum score model for analysing the role of momentum for players in a tennis match. The concept of momentum is quantified by defining several metrics, including the base point, hold point, and break point, and developing a formula to calculate the momentum score. By applying the momentum model to different races, it is found that the change in momentum is very consistent with the actual races, and it is concluded that the momentum model has strong practical applicability. The permutation test proved that momentum does play a role in tennis and is not a completely random event. Finally, by using a decision tree classification model to predict the direction of FLOW in a tennis match, the article derives several factors that are most likely to influence the fluctuation of the match, such as the SERVER factor, which has the highest correlation. Since the Decision Tree Classification Model has up to 99% accuracy in predicting turning points, the model is used in the article as a basis for some preparation strategies for tennis players. For example, analyzing the opponent's playing patterns, enhancing individual psychological preparation, optimizing serving and receiving strategies, etc.

References

[1] Elia Morgulev, Ofer H. Azar, Michael Bar-Eli. Does a “comeback” create momentum in overtime? Analysis of NBA tied games[J]. Journal of Economic Psychology, 2019,75, 102126.

[2] Matt Gifford, Tuncay Bayrak. A predictive analytics model for forecasting outcomes in the National Football League games using decision tree and logistic regression[J]. Decision Analytics Journal,2023, 8, 100296.

[3] Yangqing Zhao, Hui Zhang. Does success breed success? An investigation of momentum in elite recurve archery[J]. Psychology of Sport & Exercise,2023, 66, 102397.

[4] Depken, C. A., Gandar, J. M., & Shapiro, D. A. Set-level strategic and psychological momentum in best-of-three-set professional tennis matches[J]. Journal of Sports Economics,2022, 23(5), 598-623.

[5] Kapil Gupta, Vijayshankar Krishnamurthy, Soudeep Deb. What elements of the opening set influence the outcome of a tennis match? An in-depth analysis of Wimbledon data[J]. IIMB Management Review.2024

[6] Philippe Meier, Raphael Flepp, Maximilian Ruedisser, Egon Franck. Separating psychological momentum from strategic momentum: Evidence from men’s professional tennis[J]. Journal of Economic Psychology, 2020,78, 102269.

[7] Xiaojin Zhang. Improved algorithm for permutation testing[J]. Theoretical Computer Science, 2024,986, 114316.

[8] Songming Zhang, Xiaofeng Chen, Xuming Ran, Zhongshan Li, Wenming Cao. Prioritizing causation in decision trees: A framework for interpretable modeling[J]. Engineering Applications of Artificial Intelligence, 2024,133(Part C), 108224.

[9] Thomas Gilovich, Robert Vallone, Amos Tversky. The hot hand in basketball: On the misperception of random sequences[J]. Cognitive Psychology, 1985,17(3), 295-314.

[10] Silva, P., Hardy, J., & Crace, R. The effectiveness of pre-performance routines in sport: A meta-analysis[J]. Journal of Applied Sport Psychology, 2013,26(1), 94-108.