| Problem Chosen | 2024 MCM Summary Sheet | Team Control Number |
|---|---|---|
| C | | 2428322 |

# An Exhaustive Assessment Framework for Momentum in Tennis Competitions


**Summary**

In sports, "momentum" usually refers to a positive, increasing energy or momentum. Momentum plays a big part in the game, but it is difficult to predict. Therefore, we have developed a Comprehensive Evaluation Model of Momentum based on data from the 2023 Wimbledon Gentlemen's Championships to predict how momentum will be measured during tennis matches.

Initially, we selected eight indicators to analyze the momentum change including serve advantage, double faults, breaks of serve and so on. We used the data imported from *math_id*_1701 to observe the trend of momentum by **Time Series Analysis**. In order to accurately assess the contribution of each indicator to the momentum change trend, we adopt the **Entropy_Weight Method**, which is an objective assignment method, to standardize the data to find out the entropy value of each indicator. Meanwhile, we analyzed the entropy value of each indicator, it can be observed that winning broken serves and number of sets have a greater impact on the momentum of athletes. Then the differential momentum value is applied to the **CUMSUM** detection algorithm to draw a momentum change graph, and the turning points are marked with different colors. By virtue of observing where the turning points are located, it is possible to identify which players are performing better at particular times of the match and how well they are doing. It was also found that the serving side was more likely to win.

Subsequently, in order to judge whether the momentum and winning streak of the athletes in the game are random or not, we use the method of Run Test to solve the Z and P values of the momentum and winning streak on the basis of the above operation. The solved values were compared with the significant level to conclude that swings in plays is non-random while runs of success is random. In doing so, We found that **Machine Learning Algorithms** can be used to find some metrics that would help determine when the flow of a game was shifting from one side to the other. Therefore, we constructed the **Adaboost** regression model, the **GBDT** regression model, and the **Xgboost** regression model, by contrast, we found that the **Xgboost** model worked the best. So we trained the **Xgboost** model to make valid predictions out of the dataset, and then used the **Shap** model to more accurately assess the contribution of each feature to the model's predictions, followed by upward merging into an overall feature importance. It is clear from the Figure 7 that *p*2_*distance_run*, *p*1_*distance_run* have the largest feature importance. In addition, *p*1_*score*, *speed_mph* also affect the momentum to a relatively large extent. Based on these factors, we made some suggestions for athletes who are about to compete.

In an effort to contribute to the strategic development of the team, we have compiled our research findings into a memorandum for the coach. Our objective is that these insights will inform and inspire the formulation of training methodologies and game strategies for the athletes under his guidance.



**Keywords**: Momentum in Tennis Competitions, ,Entropy_Weight Method, Run Tests, Machine Learning Algorithms,Xgboost

# Contents

# 1 Introduction

## 1.1 Problem Background

Sports hold a significant place in human society, serving not only as a form of physical exertion but also as a lifestyle and a spiritual endeavor. Sports events, as a crucial component of sports, garner widespread attention and affection from the public. During these competitions, athletes often encounter an intangible phenomenon that leads to remarkable performance fluctuations, subsequently influencing the course of the game. This phenomenon is referred to as ‘momentum’.

The 2023 Wimbledon men’s singles final is eagerly awaited. In a contest marked by unexpected developments, Djokovic’s undefeated run at Wimbledon since 2013 was ended by the emerging Spanish talent, Alcaraz. In the realm of sports, "momentum" typically denotes a positive, escalating force or energy. It characterizes an athlete’s or team’s performance and demeanor during a match, as well as the assertiveness and control they exhibit. When an athlete or team is experiencing an upward trend, they often demonstrate heightened morale, self-assurance, and resolve. This enables them to adopt a more assertive stance in the game, seize opportunities, and maintain their success. Consequently, examining the factors that influence momentum can aid in identifying shifts in athletes’ momentum during a match, thereby assisting athletes in making adjustments and enhancing their performance.

## 1.2 Restatement of The Problem

In the 2023 Wimbledon Gentlemen’s Final, Novak Djokovic’s undefeated streak at the Wimbledon Open since 2013 was brought to an end by the emerging Spanish star, Alcaraz. Throughout the match, momentum served as a crucial factor, even dictating the match’s trajectory. To understand how different events within a match generate or alter momentum, we need to address the following questions, guided by the provided data.

- Construct a model that can evaluate which player performs better and how well at a given time of the game, and visualize the flow of the game.
- This model is designed to discern the randomness of an athlete’s momentum and winning streak during a match.
- The model’s objective is to identify indicators that can assist observers in determining when the game’s balance is shifting from one player to another. Utilize data to construct models that predict momentum swings, identify the most pertinent factors, and make recommendations accordingly.
- The developed model will be tested in one or more extended matches to assess its ability to predict match fluctuations. If the model’s performance is subpar, we will identify other potential factors to incorporate into future models. Additionally, we will determine whether the model can be generalized to other matches.

# 2 Assumptions and Notations

## 2.1 Assumptions

- Assumption 1: We will disregard factors external to our study, such as the influence of off-field spectators, the athletes' physical condition, and environmental variables.
- Assumption 2: We operate under the presumption that the data we utilize are independent. However, it's worth noting that in reality, there may exist some degree of correlation among the factors, which could potentially impact the final outcomes of the analysis.
- Assumption 3: We assume that the predictive models we have chosen are optimally suited to the problem under investigation and analysis. It's important to remember that the effectiveness of a model can only be ascertained through rigorous testing and validation.

## 2.2 Notations

Table 1: Notations

| Symbol | Definition |
|---|---|
| $M(t)$ | magnitude of momentum at moment t |
| $m(t)$ | difference in momentum between M(t)and M(t-1) |
| $r0$ | number of observed trips |
| $N$ | Derivative of Change In Serve Advantage Weights |
| ΔM | Momentum Change |
| ΔS | Change in weight of serve advantage |

# 3 Preliminary Preparation

## 3.1 Data Preprocessing

Upon comprehending the rules of tennis, we acquainted ourselves with the sport's technical terminology and processed the intricate data from the Wimbledon featured matches.csv file, guided by two provided documents.

Initially, we constructed a sub-table to individually record each match's data and computed the scoring advantage of the two competitors for subsequent analysis. Our literature review revealed that the server often holds the initiative in tennis matches, influencing the players' on-court momentum. Furthermore, the scoring and error situations on the court can impact the players' emotional state and mindset, potentially affecting their momentum and even the game's control[1][2].

### 3.2 Description of Definitions

Momentum: "Momentum," typically a positive, increasing energy or force, is gained by an athlete on the field.

Turning point: We hypothesize that when the cumulative sum of momentum crosses the x-axis over time, this moment can be considered a turning point.

## 4 Model Building

### 4.1 Momentum - Time Series Modeling

An athlete's momentum plays a crucial role in their performance during a match. To delve into the specific impact of momentum, we analyzed it by selecting a particular match. For this topic, we used the dataset from the 2023 Wimbledon Men's Singles Tennis Final as an example. We selected eight key metrics: Sets and Games won, Point advantage, Double faults, Break points, Unforced errors, Winning points, and ACE. Each indicator reflects the athlete's performance in one aspect, and analyzing them collectively provides a comprehensive and three-dimensional view of the athlete's performance.

Take the loss function L(w) with a single parameter w as an example, provided that this loss function is differentiable:

$$w^{*}=\text{argmin} L\,(w) \tag{1}$$

The main purpose is to solve for w that minimizes the above Loss, which in effect is to find the tangent line L with a slope of 0. (Here there may be a local minimum with a slope of 0 as well, which leads to the momentum below.) Using gradient descent can be a good solution to this problem (graphical gradient descent). Based on this idea, we first presuppose a simple additive model :

$$\mathrm{M} = \sum \mathrm{Ti} \tag{2}$$

(Ti is a different impact indicator), to determine the impact of each indicator on the momentum.

Indeed, not all indicators contribute equally to the momentum trend. To accurately assess each indicator's contribution, we assigned weights to these indicators. We assumed an additional weight of 0.175 for the serving side on each score based on the title. For the other metrics, we employed the entropy weighting method, an objective method for analyzing weights.The entropy weight method determines the weights based on the data's variability degree, avoiding the arbitrariness and bias of subjective assignment. We use the entropy value to judge each indicator's dispersion degree and understand the indicator's relative importance in the overall evaluation system. The smaller its entropy value, the greater the indicator's dispersion degree, and the greater the indicator's influence (i.e., weight) on the comprehensive evaluation. If all the values of a certain indicator are equal, the indicator does not play a role in the comprehensive evaluation. Therefore, we use information entropy as a tool to calculate each indicator's weight, providing a basis for the comprehensive evaluation of multiple indicators.

Firstly, we construct the indicator system and determine the indicator matrix:

$$X = [X1, X2, X3, X4, X5, X6] \tag{3}$$

Table 2: Notations

| Variable Name | Definition |
|---|---|
| P1_advantage | X1 |
| P1_double_fault | X2 |
| P1_break_pt_won | X3 |
| P1_unf_err | X4 |
| P1_winner | X5 |
| P1_ace | X6 |

taking the performance of 1701 field athletes as an example, where each column is an indicator.

At the same time, we select its first inning performance:X=(10 rows of 2.0).Then we normalize the initial indicator matrix X. X1,X3,X5,X6 are positive indicators, X2,X4 are negative indicators. (The specific process of calculation was realized through Python language. ) This approach ensures a fair and objective evaluation of each player's performance.

Positive Indicators:

$$Z_{ij} = \frac{x_{ij} - minX_i}{maxX_i - minX_i} \tag{4}$$

Negative Indicators:

$$z_{ij} = \frac{maxx_i - x_{ij}}{maxx_i - minx_i} \tag{5}$$

Calculate The Entropy Hj for Each Index:

$$H_j = -\frac{\sum_{i=1}^{m} p_{ij} \ln(p_{ij})}{lnm} \tag{6}$$

Calculate entropy weight wj according to entropy value

$$w_j = \frac{1 - H_j}{\sum_{j=1}^{n}(1 - H_j)} = \frac{1 - H_j}{n - \sum_{j=1}^{n} H_j} \tag{7}$$

Through the analysis and calculation using the entropy weight method, we discovered that the three indices - scoring, ace, and break serve - significantly influence an athlete's momentum. Scoring, being the most direct outcome of a match, serves as a crucial indicator for momentum shifts. ACEs and break serves represent an athlete's serve-attack and

$$\begin{bmatrix} Z_{11} & Z_{12} & \cdots & Z_{1n} \\ Z_{21} & Z_{22} & \cdots & Z_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{m1} & Z_{m2} & \cdots & Z_{mn} \end{bmatrix}$$

Table 3: Individual Data Weights for Player 2

| Index | Weight |
|---|---|
| p1_points_advantage | 0.007351 |
| p1_serve_advantage | 0.043201 |
| p1_break_points_won | 0.271763 |
| p1_unforced_errors | 0.016542 |
| p1_winners | 0.046163 |
| p1_sets_won | 0.400722 |
| p1_games_won | 0.004087 |
| p1_ace | 0.206252 |
| p1_double_fault | 0.003917 |

Table 4: Individual Data Weights for Player 2

| Index | Weight |
|---|---|
| p2_points_advantage | 0.013236 |
| p2_serve_advantage | 0.023533 |
| p2_break_points_won | 0.226787 |
| p2_unforced_errors | 0.004895 |
| p2_winners | 0.067570 |
| p2_sets_won | 0.334404 |
| p2_games_won | 0.004140 |
| p2_ace | 0.318906 |
| p2_double_fault | 0.006529 |

serve-receive-defense abilities, respectively. These factors largely dictate the match's pace and direction, subsequently affecting the athletes' momentum.

We employed time series analysis and the customized *calculate_momentum* function (an improved time window function) to compute the momentum changes of the two players in the same match. This process was visualized for better understanding (as shown in Figure 1, Figure 2). This comprehensive approach ensures a robust and accurate evaluation of each player's performance and the game's dynamics.

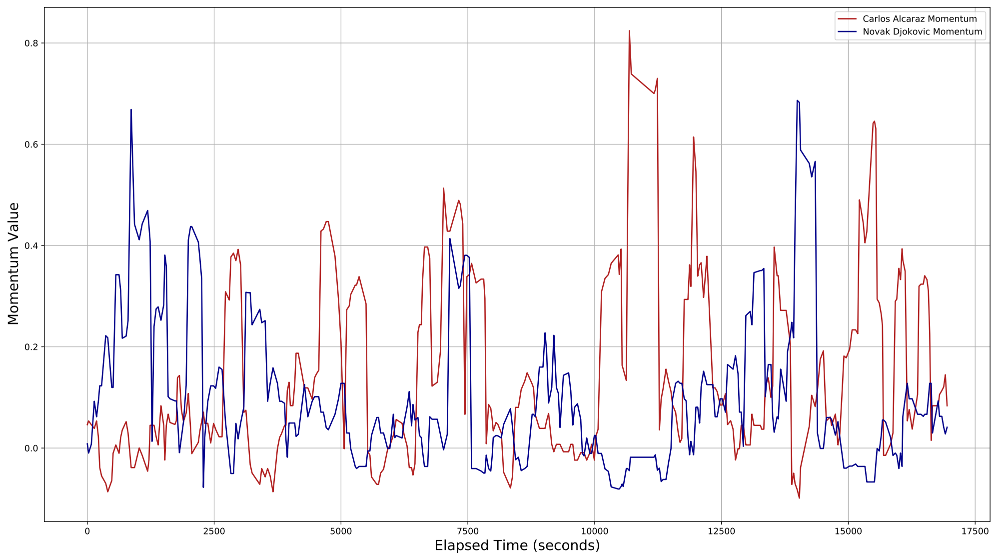



Figure 1: Momentum Change Throughout the Match

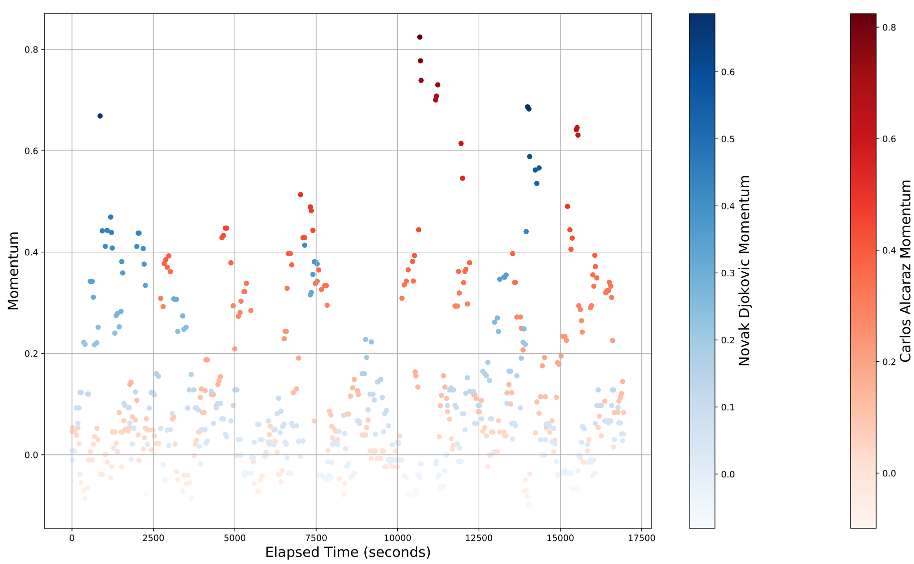



Figure 2: Momentum Change with Color Gradient Throughout the Match

## 4.2 Turning Point - CUMSUM Detection Algorithm

To examine the validity of the coach's skepticism, we opted to enhance the existing model by incorporating the CUMSUM detection algorithm. This algorithm enables us to identify pivotal moments during a game where distinct "momentum shifts" or consecutive winning patterns occur. By utilizing this approach, we aim to analyze whether the coach's doubts hold merit.

The CUMSUM detection algorithm, known as a statistical process control method, evaluates a dataset by calculating the cumulative sum of the data points, denoted as

$$C(t) = \sum \mathrm{m}(t) \tag{8}$$

This algorithm assesses the data's distribution and trends, enabling the observation of changes in the overall process. The fundamental principle involves computing the cumulative sum of deviations between each data point and the initial data point, and then plotting it on a control chart. By implementing this algorithm, we can establish a more comprehensive and precise assessment system for analyzing momentum. This system allows us to accurately identify crucial turning points that signify the occurrence of a momentum shift or a winning streak.

First, the data is processed as a differenced sequence

$$m(t)=M(t)-M(t-1) \tag{9}$$

by comparing the difference data derived from the leveled data and filling in NaN values with zeros. This operation captures the variance between two consecutive records, representing a first-order difference. Subsequently, the momentum of both players in the match is accumulated and processed using the CUMSUM algorithm. If the cumulative sum crosses the zero point again and changes sign compared to the previous value (In other words,transitioning from a winning streak to a losing streak or vice versa), it is identified as a turning point. Lastly, the momentum changes and turning points are visualized, marking these points with different colors (Fig. f3). By observing the data trend, we can distinctly discern the fluctuations in momentum during the game, thus accurately identifying periods when players have momentum and when opponents are gradually gaining the upper hand.

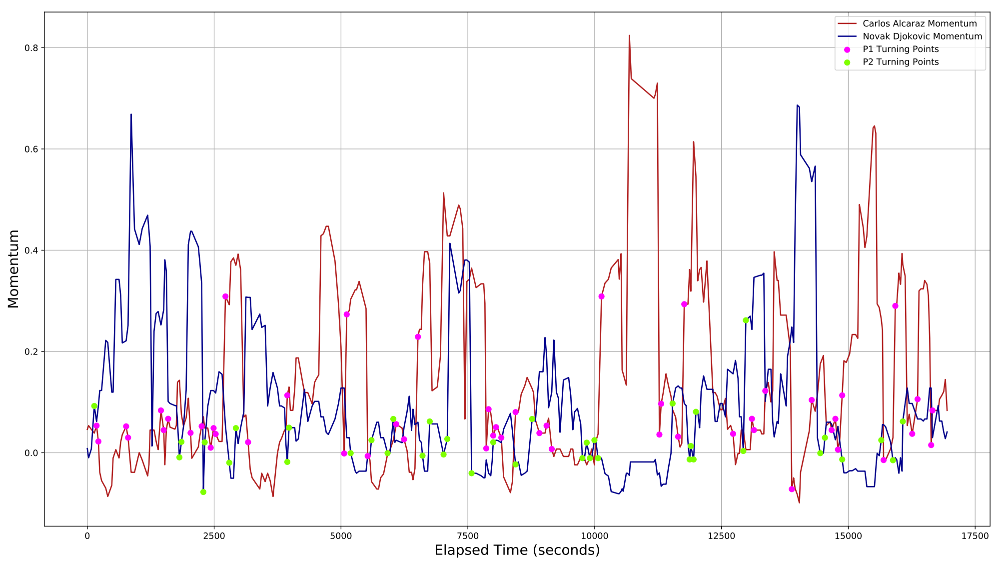



Figure 3: Momentum Changes and Turning Points

## 4.3 Randomness Judgments - Run Tests

To validate the accuracy of the results, a non-parametric statistical test, such as the Mann-Whitney U test, is employed to determine the randomness of momentum and streaks[3].

Consecutive records with identical values are considered as runs, and the total number of runs is compared to evaluate data randomness and the similarity of the two distributions[4]. A higher number of runs indicates less randomness in the data and significantly different distributions between the two aggregates. Similarly, longer run lengths signify reduced randomness and substantially distinct distributions between the two aggregates. We further assess data randomness and the equality of the overall distributions by calculating the Z-statistic and corresponding p-value.

Run r approximates normal distribution, testing statistics[5].

$$Z = \frac{r - E(r)}{\sigma_r} \tag{10}$$

where

$$E(r) = \frac{2n_1 n_2}{n_1 + n_2} + 1 \tag{11}$$

$$\sigma_r = \sqrt{\frac{2n_1 n_2 (2n_1 n_2 - n_1 - n_2)}{(n_1 + n_2)^2 (n_1 + n_2 - 1)}} \tag{12}$$

When n1 and n2 are large, Z approximately follows the standard normal distribution.

The Z statistic quantifies the disparity between observed and anticipated run lengths. It is computed as

$$z_c = \frac{\overline{x} - u_0}{\sigma / \sqrt{n}} \tag{13}$$

where O represents the observed run length, E denotes the expected run length, and SQRT signifies the square root function. If the Z statistic exceeds the critical value at a confidence level of 95%, it indicates that observed run lengths significantly deviate from expected run lengths, suggesting reduced data randomness and markedly different distributions between the two aggregates.

Likewise, the p-value gauges the significance of the variance between observed and anticipated data. It is calculated as

$$p = 2[1 - \Phi_0] \tag{14}$$

wherein X denotes the random variable and O denotes the observed data. If the p-value falls below the significance threshold of 0.05, it indicates significant disparity between observed and expected data, reflecting diminished data randomness and distinctly disparate distributions between the two aggregates. The assessment of data randomness and the parity of the two distributions can be further evaluated by computing the Z-statistic and p-value.

Table 5: Run Test Results

| Name | Sample | Z | P |
|---|---|---|---|
| p1_momentum | 334 | -11.609 | 0.000007 |
| p2_momentum | 334 | -13.713 | 0.000005 |
| p1_turning_points | 334 | 0.282 | 0.677 |
| p2_turning_points | 334 | -0.315 | 0.578 |

## 4.4 Machine Learning

### 4.4.1 Data Pre-Processing

Initially, we removed all null values, outliers, and text-based data from the dataset. The Backward Differencing Method was employed for this purpose: rows exhibiting null differences were discarded and the dataset was reindexed subsequently. We then segregated the variables into two sets: X, comprising all independent variables, and Y, representing the momentum of p1[6].

### 4.4.2 Machine Learning Multiple Model Comparison Processing

The dataset was partitioned into a training set and a test set. We evaluated the model's performance by plotting the actual versus predicted values and analyzing the mean absolute error (MAE)[7].Resulting in the findings illustrated in Figure 4, Figure 5.

Table 6: MLP Neural Network Model

| Type | Value |
|---|---|
| Mean Absolute Percentage Error of The Training Set | 1.262 |
| Mean Absolute Percentage Error of The Testing Set | 1.069 |
| Average Absolute Error | 0.377 |
| r2_score | -36.429 |

Table 7: Linear Regression Model

| Type | Value |
|---|---|
| Mean Absolute Percentage Error of The Training Set | 2.642 |
| Mean Absolute Percentage Error of The Testing Set | 3.495 |
| Average Absolute Error | 0.081 |
| r2_score | -0.345 |

Table 8: Decision Tree Regression Model

| Type | Value |
|---|---|
| Mean Absolute Percentage Error of The Training Set | nan |
| Mean Absolute Percentage Error of The Testing Set | nan |
| Average Absolute Error | 0.114 |
| r2_score | -2.45 |

Table 9: Random Forest Regression Model

| Type | Value |
|---|---|
| Mean Absolute Percentage Error of The Training Set | 4.519 |
| Mean Absolute Percentage Error of The Testing Set | 7.560 |
| Average Absolute Error | 0.072 |
| r2_score | -0.251 |

Table 10: Adaboost Regression Model

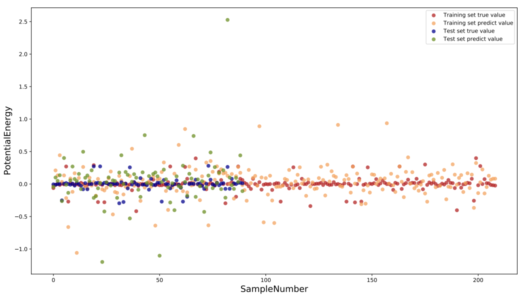
| Type | Value |
|---|---|
| Mean Absolute Percentage Error of The Training Set | inf |
| Mean Absolute Percentage Error of The Testing Set | 3.937 |
| Average Absolute Error | 0.096 |
| r2_score | -0.904 |

Table 11: GBDT Regression Model

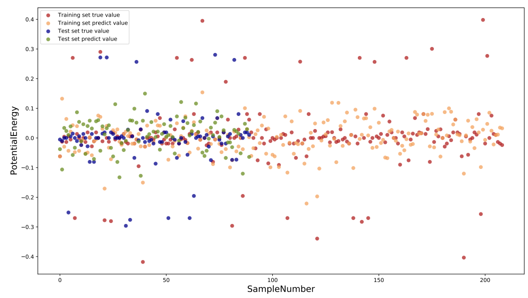
| Type | Value |
|---|---|
| Mean Absolute Percentage Error of The Training Set | 5.813 |
| Mean Absolute Percentage Error of The Testing Set | 2.320 |
| Average Absolute Error | 0.037 |
| r2_score | 0.667 |

Table 12: Xgboost Regression Model

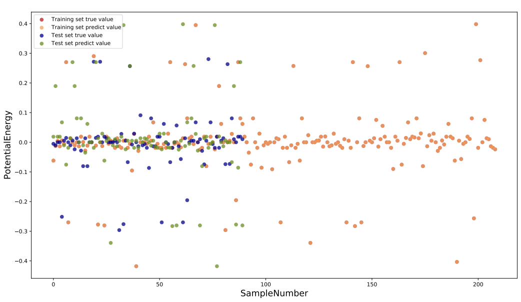
| Type of Error | error value |
|---|---|
| Mean Absolute Percentage Error of The Training Set | 0.207 |
| Mean Absolute Percentage Error of The Testing Set | 0.221 |
| Average Absolute Error | 0.001 |
| r2_score | 0.999 |



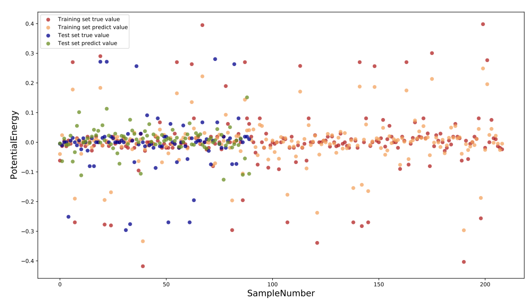

(a) Neural Network




(b) Linear Regression




(c) Decision Tree Regression



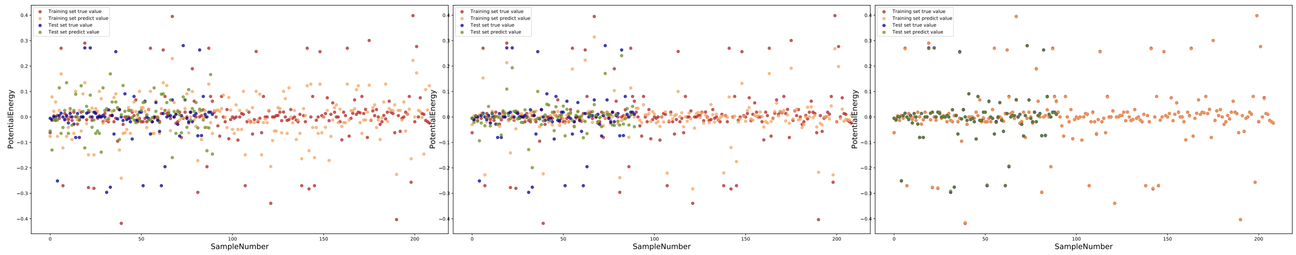

(d) Random Forest Regression

Figure 4: True Vs. Predict

### 4.4.3 Model Comparison - Xgboost Model

Upon comparing the MAE across various models, it was determined that the XGBoost regression model outperformed others by yielding the lowest MAE. It should be noted that due to the inherent randomness in the selection of data, the feature importance derived from tree-based machine learning models can exhibit variability, indicating an instability in feature importance within the context of machine learning.

The essence of the XGBoost algorithm lies in its application of gradient boosting. This approach incrementally builds new learners on existing ones to minimize the loss function, thereby enhancing model performance. Furthermore, the algorithm employs a tree structure for feature transformation and selection, with Classification and Regression Trees (CART) serving as the foundational learners. Consequently, we subjected our data to analysis using the XGBoost regression model, resulting in the findings illustrated in Figure 6.




(a) Adaboost Regression (b) GBDT Regression (c) Xgboost Regression

Figure 5: True Vs. Predict

Xgboost Regression:

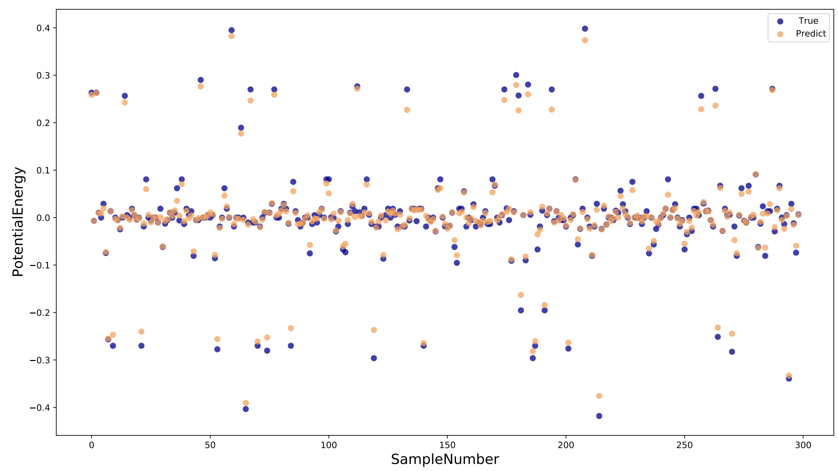



Figure 6: True Vs. Predict

## 4.5 Shap Modeling

### 4.5.1 Building Shap Modeling

It is essential to clarify that our methodology does not incorporate machine-learned feature importance due to the potential instability arising from correlations between two features. In practical scenarios, the interaction effects among features can substantially influence the outcomes of predictions. As a result, we adopt an alternative analytical approach by leveraging SHAP (SHapley Additive exPlanations) model interaction value analysis for the interpretation of machine learning model predictions.

This technique is underpinned by the Shapley value, a concept derived from cooperative game theory, designed to quantify the individual contributions of features to the overall model prediction. The primary objective is to disentangle the interaction effects present between features from their main effects, thereby enabling a more precise evaluation of each feature's contribution to the model prediction. This process commences with an initial analysis of individual samples to ascertain the feature importance for each. Subsequently, these insights are amalgamated, leading to an overarching view of feature importance, which is illustrated in the subsequent figure (Figure 7).

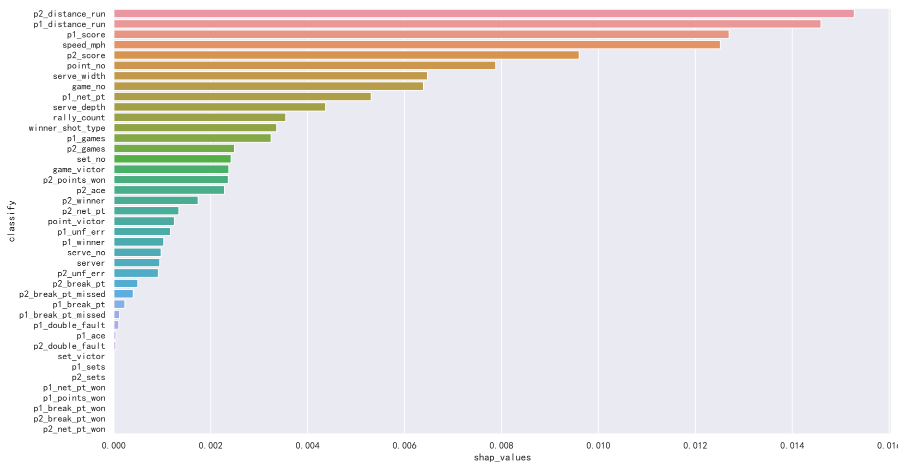



Figure 7: Degree To Which Each Factor Affects Momentum

Employing SHAP model interaction value analysis provides a nuanced perspective on how interactions between features affect model predictions, offering a more sophisticated understanding than what is typically afforded by conventional feature importance metrics.

The bar lengths depicted in the graph serve as indicators of the extent to which various factors influence momentum; a longer bar signifies a stronger correlation. Analysis of the graph reveals that *p2_distance_run*, *p1_distance_run*, *p1_score*, and *speed_mph* are identified as having greater feature importance. This suggests that these metrics play a pivotal role in determining the race's dynamics, essentially influencing the shift in momentum from one competitor to another.

### 4.5.2 Strategies and Recommendations

Utilizing the SHAP model, it was discerned that variables such as *p2_distance_run*, *p1_distance_run*, *p1_score*, and *speed_mph* exert significant influence on momentum. Based on these findings, additional factors were identified for inclusion in our future models. Consequently, we propose the following recommendations for athletes to enhance their performance:

- Comprehensive Physical Training: Athletes should regularly engage in training that encompasses strength, speed, endurance, flexibility, and coordination. Such holistic physical conditioning is essential for improving an athlete's capability in covering distances effectively.

- Technical Mastery: Proficiency in technical skills, particularly in serving, receiving serves, and executing aces, is paramount for a tennis player's ability to score while minimizing errors during play. Emphasizing technical skill development can significantly impact performance.
- Tactical Adaptability: Adopting offensive or defensive strategies based on the opponent's strengths and weaknesses is crucial. Athletes should tailor their serving tactics and adapt their gameplay—altering rhythm, direction, and strategy—to exploit opponents' vulnerabilities, thereby leveraging their own strengths and depleting the opponent's stamina.
- Mental Resilience: The psychological state of an athlete plays a pivotal role in the game's outcome, especially under high-pressure conditions. Athletes with emotional stability, confidence, determination, and a robust ability to resist distractions are more likely to excel. Thus, enhancing mental fortitude through targeted training is vital to ensure athletes remain unfazed by external factors and maintain peak performance levels during competition.

# 5 Sensitivity Analysis

At the beginning, we assumed that the weight of the player's serve advantage was 0.175. Since this was set subjectively, we conducted sensitivity analysis for this. We set the weight of serve advantage as (0.1-0.2) and adjust the simulation parameters in increments of 0.75. Make sure the momentum value is around 0.4, while introducing weak positive trends and fluctuations to recalculate sensitivity to draw a new chart. The influence of momentum on the weight of serve advantage is analyzed. The sensitivity is obtained by the formula:

$$\boxed{N = \frac{\Delta M}{\Delta S}} \tag{15}$$

The results show that the momentum value has a slight change around 0.4. With the increase of service advantage weight, the momentum value has a slight positive change. So increasing the serve advantage can increase the momentum.

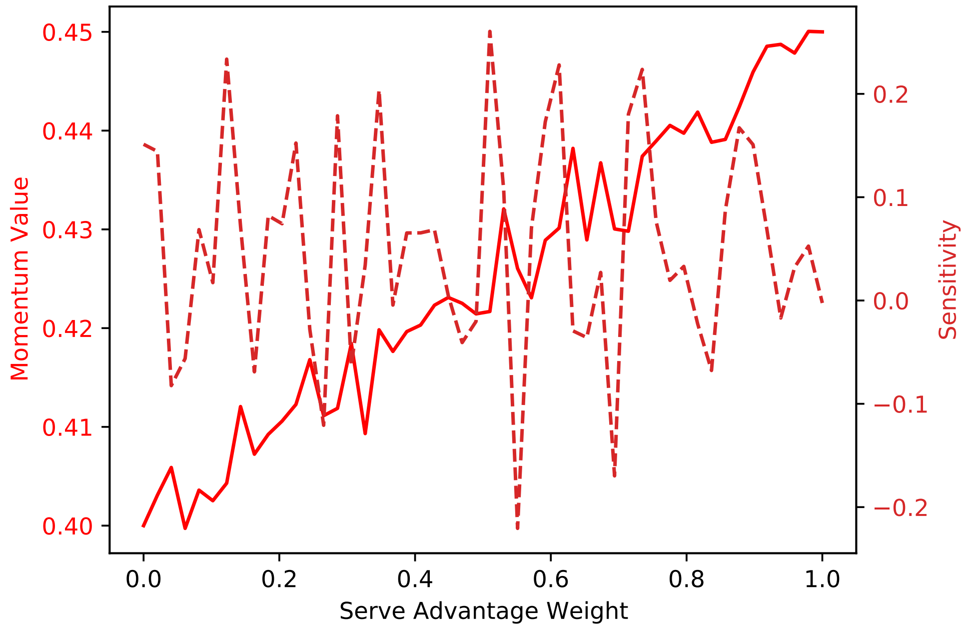



Figure 8: Degree To Which Each Factor Affects Momentum

# 6 Pervasive testing

## 6.1 Forecast accuracy

In the above operation, the Xgboost model has been trained, so we directly construct an array of data for prediction.

Nine variables are put together to construct a model for violent solutions. For more than six hundred, we use a genetic algorithm. It searches for optimal solutions by modeling natural selection and genetic mechanisms. It searches using a population approach, which contains multiple individuals (solutions), each with a fitness value. Operations such as selection, crossover and mutation are used to iteratively generate new individuals, while eliminating those with low fitness, and finally obtaining the individual with the highest fitness as the optimal solution. The score is counted as 1 and the lost score is counted as 0. Then the original data is replaced and put into the Xgboost model for prediction, which gives a prediction value of 2.024695,which is similar to the actual value. Therefore, the model is more accurate.Then the original data is replaced and put into the Xgboost model for prediction。We can get that the best combination was (0,1,0,1,0,0,0,0,0) and the best sore was 0.29930353.

## 6.2 Exploration of universality

In exploring the generalizability of the resulting model to other games (e.g., women's games), tournaments, court surfaces, or other sports (e.g., table tennis), we begin by distinguishing between the theoretical framework of the model and its practical application. In data science, the theoretical framework of a model is usually somewhat universal, and the

Table 13: True Vs. Predict

| True | Predict |
|---|---|
| 0.088843278 | 0.09277342 |
| -0.00837873 | -0.002227305 |
| 0.088843278 | 0.0948741 |
| 0.012807814 | 0.017289933 |
| 0 | 0.05569276 |
| 0.033280148 | 0.033280148 |
| -0.089800096 | -0.07803993 |
| -0.080464547 | -0.06600786 |
| 0.016757461 | 0.023373215 |

principles and methods can be applied to a wide range of contexts. In real life, however, there are often many factors that affect the use of mathematical models. The universality of modeling frameworks stems mainly from their algorithms and theoretical foundations, such as machine learning algorithms, statistical methods, and data processing techniques. These frameworks are usually not designed to target specific types of data or domain-specific problems, but rather aim to solve generalized prediction or classification tasks. Thus, other campaigns that can be defined as regression tasks, where the model theoretically has universally applicable methods for explaining the model's decisions, are not subject to a specific application domain in order to employ the same framework. However, when the model is trained on a specific dataset, its generalizability is limited.

Since each sport, game type, and field surface has its own unique characteristics and patterns, models need to capture these uniqueness to make accurate predictions. Therefore, models trained on specific types of game data may not be directly applicable to other contexts and their effectiveness may be compromised. The consideration of model generalization is to evaluate the generalizability of the model in different contexts, and the ability of the model to generalize. If a model performs well on several different datasets, then we can assume that the model has strong generalization ability and generalizability. On the contrary, if the model only works well on a specific dataset, then its generalization is questionable.

Taking table tennis game as an example, the players' on-court serving, receiving, attacking, defending and moving distances, and serving speeds all affect the size of their on-court momentum, and thus their scores. Since their scores are also arranged according to a certain time, it meets the regression characteristics, so it can be predicted according to the time series, and our model can be applied.

# 7 Model evaluation

## 7.1 Strengths

- Employing the polar deviation method entails applying distinct standardization algorithms to each fundamental indicator within the player's serve error and winning indicator system. This approach effectively minimizes discrepancies among indicators in terms of units and magnitudes.

- Utilizing the CUMSUM detection algorithm enables the transformation of the abstract notion of momentum into quantifiable data, thereby mitigating the impact of subjective factors. Consequently, this process renders the assessment results more objective and credible.
- The SHAP model provides a visual representation elucidating the relative contribution of each feature towards the predictive outcomes, facilitating a deeper understanding of the model's decision-making foundation.

## 7.2 Weaknesses

- The current approach overlooks the variance in distance covered by athletes during serve and receiving actions and fails to consider the physical exertion experienced by athletes during matches.
- When applying the swim test, the sample is still not large enough, which may affect accuracy.
- The serve advantage has not been objectively analyzed for its weight, and the assumption of additional weights is taken directly, with some error.

## 7.3 Further Improvements

While our model exhibits a degree of generalizability, it is essential to verify its applicability through real-world data. Emphasis must be placed on the specificity of data in practical implementations, necessitating proper adjustment and validation of the model to ensure its continued efficacy and accuracy within novel application environments.

# Memorandum

**To:** Coaches
**From:** Team2428322
**Date:** February 5,2024
**Subject:** Recommendations based on Comprehensive Evaluation Model of Momentum in Tennis Matches

Dear coach,

We are Team2428322.

In the realm of sports, athletes or teams often perceive a sense of momentum or dominance during a game, a phenomenon challenging to quantify. We refer to this as "power momentum." To gain deeper insights into momentum and enhance athletes' on-court performance, we have developed a mathematical model aimed at predicting momentum fluctuations in tennis matches. This model is based on data derived from the 2023 Wimbledon Tennis Championships.

Among numerous indicators, we carefully selected eight specific metrics for analysis in order to study momentum shifts. These indicators include serve advantage, double faults, broken serves, unforced errors, winners, aces, games won, and sets won. Initially, we employed a simplistic cumulative model to evaluate their impact on momentum. Recognizing the unequal significance of these indicators, we assigned individual weights to accurately measure their contributions to momentum changes. Based on a review of pertinent literature and win-loss patterns, we observed that in tennis matches, the serving side typically seizes greater initiative and momentum. Consequently, we augmented the weight of each point scored by the serving side by 0.175. Normalizing the data, we computed the weights for each indicator and presented the process of momentum change visually. Comparative analysis of the weights revealed that breaks, winners, and winning sets exerted a more substantial influence on player momentum.

Secondly, following the aforementioned procedure, we calculate the variance in momentum values between consecutive moments. This calculation enables us to graph the momentum fluctuations and delineate pivotal moments using distinct colors. By examining these inflection points, it is feasible to discern the superior performance of individual players at specific junctures during the match. Furthermore, this analysis allows for an assessment of the likelihood of victory for the serving side. Upon conducting a thorough comparison of significance levels, we have ascertained that momentum exhibits non-random behavior, while streaks demonstrate randomness within the context of our study.

During our research, we observed that certain metrics appeared to be indicative of a shift in game momentum from favoring one side to favoring the other. Consequently, we conducted an analysis using multiple machine learning models and selected the most effective one for training. This enabled the dataset to generate valid predictions. By employing the selected model, we were able to assess the contribution of each feature to the model's predictions more accurately. The results were then consolidated into a single graph to comprehensively present the findings. The graph clearly demonstrates that $p2\_distance_run$ and $p1\_distance_run$ have the most significant impact on momentum. Moreover, $p1\_score$ and $speed\_mph$ also exert a relatively substantial influence on momentum.

The level of momentum an athlete possesses during a game directly impacts their performance and likelihood of winning. Our research has identified the critical factors that influence the magnitude of momentum. Based on these factors, we have formulated recommendations aimed at aiding athletes in effectively managing the challenges and pressures of competition to enhance their overall performance.During regular training sessions, it is advisable to prioritize the development of physical attributes, including strength, speed, endurance, flexibility, and coordination. This training regimen contributes to improving the athlete's ability to cover greater distances, subsequently increasing their momentum.In tennis, technical proficiency plays a pivotal role in scoring. By enhancing athletes' technical abilities in serving, receiving, and executing ACE shots, while simultaneously minimizing errors on the court, athletes can cultivate higher levels of confidence and determination, thus amplifying their momentum. Furthermore, in both training and competition, it is essential to develop serving tactics that are tailored to the opponent's characteristics. This involves strategically capitalizing on the opponent's weaknesses, modifying the tempo and direction of play, and employing other tactical maneuvers to maximize the utilization of the athlete's own strengths while depleting the opponent's physical resources.Additionally, the athlete's psychological state holds equal importance in determining the outcome of a tense and intense game. Athletes who exhibit emotional stability and possess strong resilience against external distractions are more likely to perform at their optimal level. To ensure they are not affected by external factors on the court and maintain stability, it is crucial to prioritize mental training interventions.

We anticipate that our findings and suggestions will provide assistance for your future training sessions and forthcoming competitions.

Yours sincerely,

Team2428322

# References

[1] Huang Zhiying. Research on tennis match performance prediction model based on BP neural network. 2022. Fujian Normal University,MA thesis.doi:10.27019/d.cnki.gfjsu.2022.000427.

[2] Chen Yan, et al. Kinesiological study on the first and second serving skills of female athletes of the national tennis team. Bulletin of Sports Science and Technology Literature 31.09(2023):9-12+211.doi:10.19379/j.cnki.issn.1005-0256.2023.09.003.

[3] Aslam, Muhammad. "The run test for two samples in the presence of uncertainty." Journal of Big Data 10.1 (2023): 166.

[4] Sun Jianwei, et al. Run test and its application and discussion in epidemiology. Preventive Medicine Forum 23.02(2017):100-103+106.doi:10.16406/j.pmt.issn.1672-9153.2017.02.006.

[5] Pandis, Nikolaos. "Calculating the P value and carrying out a statistical test."American journal of orthodontics and dentofacial orthopedics 148.1 (2015): 187-188.

[6] Xia Runhai, and Wang Kaiyan. "Machine Learning and Intelligent Decision Support System." Journal of Weifang University,003.002(2003):57-59.

[7] Meng， Yuan， et al. "What Makes Online Reviews More Helpful: An Interpretation Framework Using XGBoost and SHAP Values." Research in E-Commerce Theory and Applications 16.3 （2020）： 466-490.

# Report on Use of AI

1. Github CoPilot (Feb 3, 2024 version)
   Query1: Could you help me polish my academic paper?
   Output: Of course, I'd be happy to help you polish your academic paper. Please provide the section or the specific text that you'd like me to help with. Remember to remove any personal or sensitive information from the text before sharing.

2. Github CoPilot (Feb 3, 2024 version)
   Query1: Since each sport, game type, and field surface has its own unique characteristics and patterns, models need to capture these uniqueness to make accurate predictions. Therefore, models trained on specific types of game data may not be directly applicable to other contexts and their effectiveness may be compromised. The consideration of model generalization is to evaluate the generalizability of the model in different contexts, and the ability of the model to generalize. If a model performs well on several different datasets, then we can assume that the model has strong generalization ability and generalizability. On the contrary, if a model only works well on a specific dataset, then its generalizability is questionable.

   Translated with www.DeepL.com/Translator (free version)
   Output: Since each sport, game type, and field surface has its own unique characteristics and patterns, models need to capture these uniqueness to make accurate predictions. Therefore, models trained on specific types of game data may not be directly applicable to other contexts and their effectiveness may be compromised. The consideration of model generalization is to evaluate the generalizability of the model in different contexts, and the ability of the model to generalize. If a model performs well on several different datasets, then we can assume that the model has strong generalization ability and generalizability. On the contrary, if the model only works well on a specific dataset, then its generalization is questionable.