Problem Chosen	2024 MCM/ICM Summary Sheet	Team Control Number
C		2422432

Swing the Game: A Deep dive into Tennis Match Momentum

Summary

In sports competitions, suspenseful situations often arise, creating fluctuations in the overall dynamics. The intense showdown in the 23rd Wimbledon Men's Singles final epitomized this tension to the utmost degree, sparking fervent discussions among spectators about the factors influencing these fluctuations. Experts introduced the term "**Momentum**" to analyze these dynamics. Therefore, grounded in our research on momentum, we conducted a data analysis of the outcomes of tennis matches.

Firstly, we conducted **data preprocessing** for 4 sequential steps: EDA, Indicator Selection, Data Transformation and Data Augmentation.

For Question I, specifically addressing the dataset requirements for Long Short-Term Memory **(LSTM)**, we performed targeted deep data processing. Considering the significant impact of serves and receptions, an attention module was additionally incorporated to enhance the weightage of the "server". Subsequently, we established the LSTM model to quantify player performance and obtained results for relative performance. Finally, we visualized the quantified results as shown in Figure 6.

For Question II: We established **Random Forest Models based on SHAP Explanations** to quantify momentum, assigning weights to all influencing factors according to their importance and performing weighted calculations. Figure 9 illustrates the quantified momentum results. Subsequently, we evaluated the impact of momentum. Through forward-solving correlation using the **Chi-square Test** and reverse-testing randomness using the **Run Test**, we concluded that swings in play and runs of success are **not random**. Moreover, runs of success are significantly influenced by positive momentum.

For Question III: We introduced the **LightGBM** model to predict the volatility of relative momentum. We evaluated the predictive performance using a **Confusion matrix** and **ROC** curve, achieving an accuracy rate of **96.22%**. This indicates a strong performance in accurately predicting the model. Based on the fluctuation of relative momentum within a match, we provided recommendations for players in two aspects: general patterns and opponent analysis, offering comprehensive guidance for players.

For Question IV: Based on the official dataset in Table 9, we extended the **quantified momentum** and **LightGBM** prediction model to WS and MD of tennis, MS of badminton. According to our analysis, potential factors to be included in future models are player abilities, match schedules. Our model shows strong generalization in tennis matches but has limited capability when applied to datasets from other sports, especially those with insufficient features.

Finally, We have summarized our work into a **memorandum** intending to provide coaches with a preparatory assessment based on the impact of momentum.

Key Words: Momentum; LSTM; Random Tree; SHAP; LightGBM

Contents

1 Introduction . . . **1**
- 1.1 Problem Background . . . 1
- 1.2 Restatement of the Problem . . . 1
- 1.3 Related Work . . . 2
- 1.4 Our Approach . . . 3

2 Assumptions and Justifications . . . **4**

3 Notations . . . **4**

4 Data Preparation . . . **4**
- 4.1 Exploratory Data Analysis . . . 4
- 4.2 Indicator Selection . . . 5
- 4.3 Data Transformation . . . 6
- 4.4 Data Augmentation . . . 6

5 Performance Evaluation . . . **7**
- 5.1 Model Preparation . . . 8
- 5.2 Quantification of Performance based on LSTM . . . 8

6 Momentum Measurement . . . **10**
- 6.1 Quantify Feature Importance . . . 10
- 6.2 SHAP Measures . . . 11

7 Assessment of Momentum's Impact on Performance . . . **12**
- 7.1 Correlation Analysis . . . 13
- 7.2 Stochastic Analysis . . . 14

8 Prediction for Fluctuations in Relative Momentum . . . **16**
- 8.1 Prediction for Swings . . . 16
- 8.2 Factor Analysis . . . 17
- 8.3 Advice on a New Match Base on Swings of **Momentum** . . . 17

9 Model Testing and Generalization . . . **18**
- 9.1 Data preparation . . . 18
- 9.2 Model Generalization . . . 19

10 Sensitivity Analysis . . . **19**

11 Model Evaluation and Further Discussion . . . **20**
- 11.1 Strengths . . . 20
- 11.2 Weaknesses And Further Discussion . . . 20

12 Memorandum . . . **21**

1 Introduction

"I'll often be able to jump out to a quick start so that by the time my opponent pulls himself together I'm on a roll that's tough to stop. "[5]

—— Connors

1.1 Problem Background

The men's singles final at Wimbledon that concluded in the summer of 2023 is considered an extraordinary battle, witnessing the regret of the defending champion and the rise of the post-2000 generation young talent. It also seemed to foreshadow a changing of the guard in the world of men's tennis.

It is noteworthy that this match was replete with dramatic and tense momentum changes. The seasoned veteran Djokovic, who initially displayed a strong and dominant performance, ultimately suffered defeat. However, the young talent Alcaraz, who surged later in the match, did not secure victory easily, with closely contested scores persisting until the fifth set to determine the outcome.

This final has become one of the classics in the history of tennis, sparking fervent discussions about the role and generation of **momentum** in sports competition. By studying the impact of **momentum**, and tracking and analyzing changes in scores in past matches, we can enhance our understanding of match development, aiding players in achieving better results in tennis competitions.

(a) Grand Slam Tennis

(b) Wimbledon Title Game

Figure 1: Tennis Match[1]

1.2 Restatement of the Problem

Tennis is a classic sports event with many complexities. Based on the Wimbledon score dataset provided in the question, coupled with an in-depth analysis and research on the problem, the restatement of the questions can be articulated as follows:

Question 1:

1. Develop a model to capture momentum changes during score fluctuations in a

match, dynamically quantifying player performance in alignment with tennis event rules.

2. Visualize the match process.

Question 2: Establish a model to assess the impact of momentum on swings and consecutive victories.

Question 3:

1. Develop a model to predict match fluctuations.
2. Identify factors significantly influencing fluctuations.
3. Provide recommendations for the next new match based on momentum performance.

Question 4: Test our model with different matches to evaluate the predictive effectiveness and generalization of the model.

Question 5: Write a one to two-page memorandum. This memorandum is intended for communication with tennis coaches regarding the role of "momentum" and providing recommendations for better preparation.

1.3 Related Work

This question primarily pertains to the momentum in the study of tennis. In recent years, there has been fervent discussion regarding the role and definition of momentum in sports.

Firstly, integrating findings from previous research, this paper defines momentum as a dual-level alteration in intellectual, physiological, emotional, and behavioral aspects induced by a single event or a series of sudden occurrences, consequently leading to a transformation in performance.[6]

Building upon this definition, momentum can be characterized as a **short-term** evaluative criterion.

Moreover, despite the abundance of research on the impact of momentum and the continuous evolution of momentum models, there is a scarcity of methods quantifying momentum in tennis matches using real-world data.

In the study of momentum acquisition, one research [6]utilized momentum chain1.3 to provide a vague yet comprehensive overview.

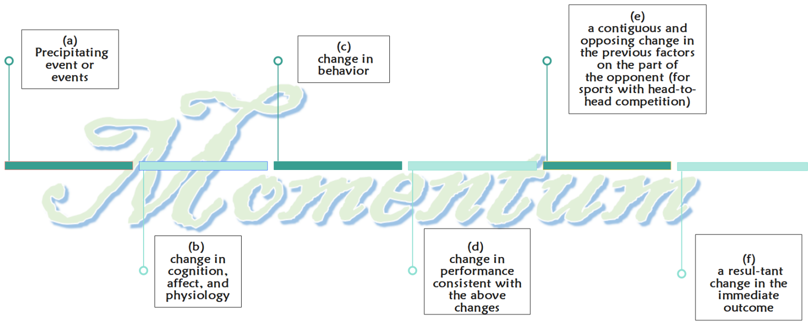


Figure 2: Momentum Chain

However, some studies, such as O'Donoghue (2009)[4], categorize the influencing factors of momentum into **3** classes and provide detailed explanations:

- Positive factors, Negative factors, and Other factors.
- Factors at the individual level, including confidence, luck, and scoring, as well as environmental factors like inspiration, coordination, and expectations, serve as triggers for both positive and negative momentum. Positive factors can lead to an uplifting favorable momentum, while negative factors result in a downturn of unfavorable momentum.

1.4 Our Approach

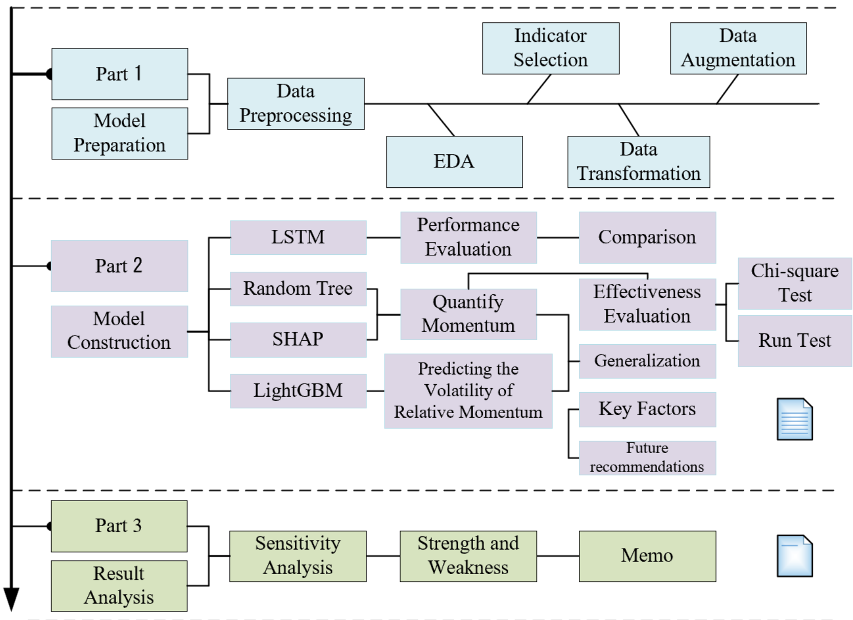


Figure 3: Flow Chart of Our Work

2 Assumptions and Justifications

In order to simplify the problem, we have made the following fundamental assumptions, each of which is assumed to be accurate and reasonable:

Assumption 1: The impact of one factor on momentum is singular.
Justification: The influences of various factors are complex and cannot be simplistically dichotomized. From a psychological perspective, there are assertions of positive inhibition and negative promotion, but their effects are relatively negligible compared to positive promotion and negative inhibition.[5] Therefore, in subsequent bidirectional categorizations of the impact of factors on momentum, only the predominant effect is considered.

Assumption 2: Gender variables are neglected in model establishment.
Justification : Although research has confirmed the impact of gender on momentum fluctuations, with male players having a greater positive momentum compared to female players, the gender of opponents in traditional sports matches is consistent. Therefore, the influence of this factor is overlooked in the model.

3 Notations

Table 1: Notions and Symbol Description

Symbol	Description
Score	the numerical representation of points
Game	a unit of play in tennis
Set	a collection of games in tennis
AD	short for "advantage"
Relative performance	The difference in performance
Relative momentum	The difference in momentum

4 Data Preparation

Through a detailed analysis of the Wimbledon match dataset provided in the question, we can extract rich information about match situations. By examining fundamental information, match states, serving, scoring, player activity data, etc., we can analyze the dynamics of the matches and the performances of the players:

4.1 Exploratory Data Analysis

- **Missing Values:**

Missing values were identified using the 'missingno' module from the 'Seaborn' library, and the distribution and patterns of missing values were visualized using bar charts (see Figure 4). Upon observing the visualizations, it was noted that the last four measurements of player activity data (speed_mph, serve_width, serve_depth, return_dept) exhibited varying degrees of missingness, all below 25% and did not significantly impact the dataset.
Considering the specific context of the problem, as these four variables revolve around the specific evaluation of serve and return actions, their impact on match performance can be summarized by other factors (such as unreturnable winning serves). Therefore, in subsequent analyses, they are categorized as factors not considered, and we can disregard and not handle the missing values.

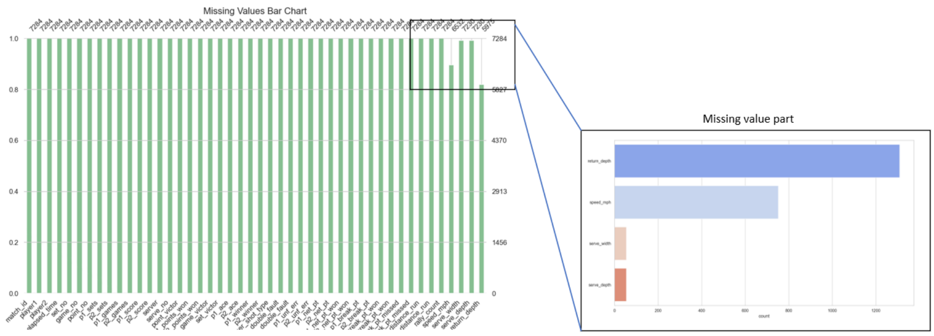


Figure 4: Visualization of Missing Values Statistics

- **Outlier Detection:**

  Based on the initial screening of the data in Excel, we did not identify any values that deviate from the inherent definitions of the variables. Given the multifaceted nature of sports events and the variability associated with objective factors in each match, we conclude that data that appears seemingly anomalous falls within a normal range. Variable values consistent with the fundamental logic of numerical variables will be retained to ensure the authenticity and validity of the data.

4.2 Indicator Selection

Regarding the nature of variables (nominal/ordinal/scale) and the importance of variables in measuring performance, we categorized the given variables in the dataset into **5 overall classes**: Score Situation, Positive Triggering Events, Negative Triggering Events, Not Included in Performance, and Factors Not Considered. The specific classification criteria are shown in Table 2.

Among these, the first three categories were considered, while the last two were excluded. The considered categories will be further subdivided into two based on a binary assessment of performance. In this **subdivision**, Score Situation and Positive

Triggering Events were considered **positive performance**, whereas Negative Triggering Events represent **negative performance**.

Table 2: Data set classification criteria

Data Classification	Classification Criteria	Consider or not
Score Situation	Measurement of match scores, such as P1_Score	Y
Positive Triggering Events	Behaviors with a positive impact on momentum, scale-type, such as Ace	Y
Negative Triggering Events	Behaviors with a negative impact on momentum, scale-type, such as Double fault	Y
Not Included in Performance	Game information category, such as Name	N
Factors Not Considered	Variables with minor impact or substitutability, such as Net_pt	N

4.3 Data Transformation

- Due to the Score data of both players containing **non-numeric information** such as **'AD'**, and considering the unique scoring rules in tennis, it is inconvenient to directly measure the relationship between scores and winning games. Therefore, a change of scale was applied by defining the scoring rules within each set as the number of points won being equal to the number of games won. Specifically, mapping '0', '15', '30', '40', 'AD' to numerical values **'0', '1', '2', '3', '4'**.
- **Distance Normalization**: Due to the wide range of distance data, which is on a different scale compared to other variable indicators, a Min-Max normalization method was applied for better performance. Utilizing the formula1, the data is linearly scaled to the [0,1] interval, eliminating scale differences, enhancing robustness, balancing feature scales, and improving data stability.

$$Y = \frac{X - \min}{\max - \min} \tag{1}$$

In this formula, *max* and *min* represent the maximum and minimum values of the original data, respectively.

4.4 Data Augmentation

Based on an understanding of sports background and a thorough analysis of match performance issues, we introduced 4 new variables to augment the dataset, as shown in Table 3:

Table 3: Supplementary Variables

Category	Varibles	Definition
Score	Successive Victories	Consecutive Score Instances (2 represents consecutive two scores, and so on)
	Break Successive	Whether to Break Opponent's Consecutive Scores (1 = Yes, 0 = No)
Game	Successive Victories	Defined similarly as before
	Break Successive	Defined similarly as before

It is noteworthy that the data mentioned above is not artificially introduced as new data. Instead, new variables are introduced based on the existing data, presenting a **fresh perspective** on the relationships between variables. This approach increases the number of effective indicators and allows for a more accurate measurement of athlete stability, thereby enhancing the model's effectiveness.

5 Performance Evaluation

By capturing match scores, we can quantify player performance to some extent and conduct visual analysis. Tennis is a classic sport with temporal relationships during matches, so we introduced the Long Short-Term Memory (**LSTM**) network, capable of identifying temporal relationships.

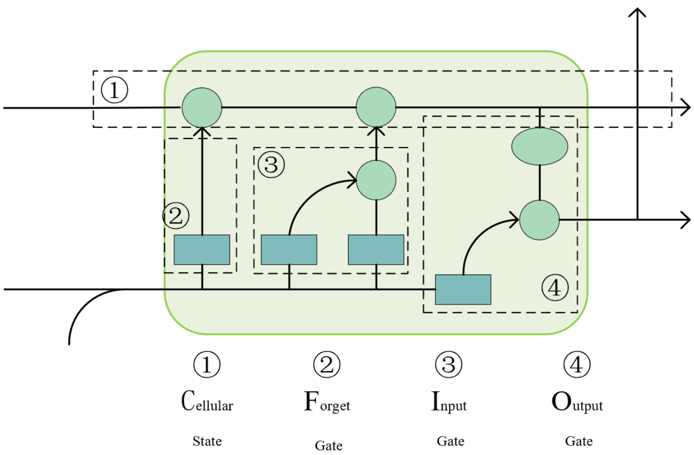


Figure 5: LSTM Neural Network tructure

LSTM is a specialized recurrent neural network (RNN) architecture that overcomes the limitations of traditional RNNs in capturing long-term dependencies in sequential data, converging faster and more easily to optimal solutions. To address the issue of poor performance in handling long sequences, LSTM introduces special storage units and controls access to memory units through three logical gates: the input gate, output gate, and forget gate. The input gate manages the flow of information into the memory unit, the output gate regulates the flow of information out, and the forget gate determines whether information should be erased.[3]

In this approach, historical data of the match time series are input into an LSTM model to learn patterns in the sequence, predicting the probability of each player winning the next game. This enables the quantification of performance.

5.1 Model Preparation

- **Data Preprocessing:** Due to the model's requirement for the time recognition data type to be 'float', the time format in the 'elapsed_time' column was converted to seconds.

  Two new features, 'p1_next' and 'p2_next', were added to record whether each player wins the next point. Standardization is applied to the features.

  The dataset was split into training and testing sets in a 4:1 ratio.

- **Definition of Transformation:** For the convenience of quantifying player performance and evaluating the model's standards, we will predict the probability of each player winning the next point. This probability serves as a concrete representation of how well a player is performing. The difference in probabilities indicates the relative level of performance.

- **Addition of Attention Module:** Given the significant impact of serving and receiving in tennis matches on player scores, we have incorporated an attention module to enhance the weight of the 'serving' aspect in the model. This aims to give more emphasis to the importance of serving during predictions.

5.2 Quantification of Performance based on LSTM

Initially, we initialized an empty sequential model using the **Sequential** model in **Keras**;

Subsequently, an **LSTM layer** and an **output layer** were incorporated by the 'add' method: the LSTM layer is responsible for handling the sequential data of the match, while the output layer is tasked with transforming the learned representations of the model into the final output. Our objective is to obtain the probability of each player winning the next point. However, the output layer typically has only one unit, hence utilizing the **sigmoid** activation function to generate probability values within the range [0, 1] , as illustrated in the formula 2.

$$Y = \sigma \left(W_{\text{out}} \cdot H_t + b_{\text{out}} \right) \tag{2}$$

Whereas, σrepresents the sigmoid activation function, and 'Wout' and 'bout' respectively denote the weights and bias of the output layer.

The output values of 'Y' determine the probability of a player winning the next point. This probability value will also serve as a concrete representation of the player's performance.

Given that the assessment of performance (whether winning the next point or not) can be considered as a binary classification problem, 'adam' optimizer and binary cross-entropy loss function were selected during compilation. Accuracy was chosen as the metric for evaluation. Binary cross-entropy loss measures the disparity between the model's output and the actual labels, and its calculation is defined by the formula 3.

$$\text{Binary Crossentropy Loss } = -(y \cdot \log(\hat{Y}) + (1-y) \cdot \log(1-\hat{Y})) \tag{3}$$

Table 4: Notations used in Equation 3

Variable	Definition
$\hat{Y}$	The model output, representing probabilities after sigmoid activation
y	True label in binary classification (0 or 1)

Note: For y, it will be represented by the p_next value in preprocessing, where a score is denoted as 1 and a loss as 0.

The model was trained using the training dataset, specifying 'epochs' as **10**, indicating that the entire training dataset is iteratively trained 10 times. 'Batch_size' determines the number of samples used in each iteration, set here as **32**.

Finally, the well-trained LSTM model was employed to predict the test dataset, obtaining the predicted probability of winning the next point. Model performance was evaluated using mean squared error, successfully quantifying player performance.

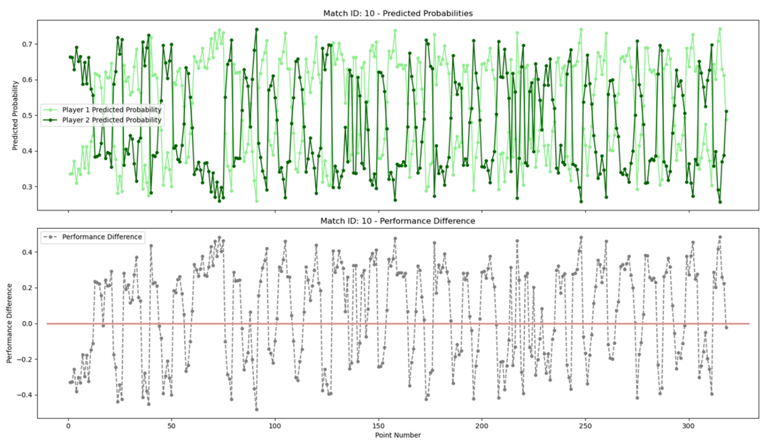


Figure 6: Quantification and Comparison of Player Performance

In Figure 6, the performance of players in the match with ID=10 is illustrated. Simultaneously, their comparative performances are displayed below, with the standard '0' split axis indicating superior performance for Player 1 above and superior performance for Player 2 below.

6 Momentum Measurement

For each match in the Wimbledon_featured_matches.csv dataset, 46 features are provided. After preprocessing, we have identified 26 features that effectively quantify momentum. These variables, along with their corresponding data, were aggregated to form a new dataset.

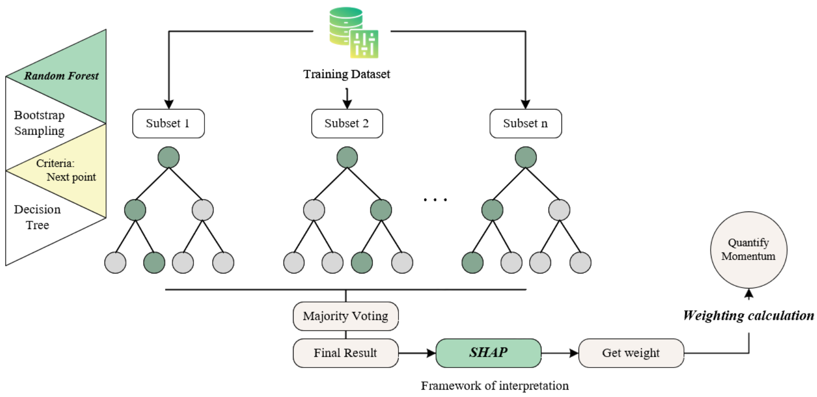


Figure 7: The Comprehensive Model Workflow

Considering the varying importance of each feature in capturing momentum, we introduced a comprehensive model integrating the **Random Forest algorithm with SHAP** . This model is designed to assess the significance of different features and assign corresponding weights to them. The workflow of the comprehensive model is illustrated in the accompanying Figure 7.

6.1 Quantify Feature Importance

Random Forest is an ensemble learning algorithm that uses decision trees as base learners. For a given dataset, Random Forest employs the Bagging technique, randomly sampling training data with replacement to construct sub-datasets, and independently trains decision trees.

Feature importance is quantified by considering the contribution of each feature to the model's performance during the construction of each decision tree in the Random Forest. In this context, model performance is measured by accuracy metrics, defining **scoring events** as a **relative reference point** and **winning the next point** as a **criterion** for evaluating improved **momentum**.

Through the utilization of Random Forest, we obtained the importance levels of 30 pre-selected effective variables within the new dataset. The visualization, depicted in the following Figure 8, categorizes the indicators into 3 levels of importance and presents them in descending **order**.

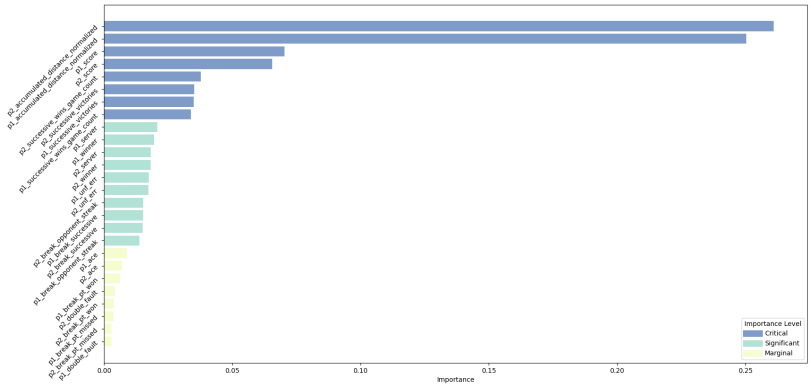


Figure 8: The Importance Level

Due to the algorithmic randomness, our model allows for an objective evaluation of feature importance, minimizing the impact of subjective judgments. This property endows the model with excellent noise resistance capabilities.

6.2 SHAP Measures

SHapley Additive exPlanations (SHAP) is a method based on Shapley values theory from game theory, used to explain the predictions of machine learning models. In this context, we have obtained weights quantifying the impact of different feature values on momentum based on the Random Forest algorithm. These weights are incorporated into the calculation of SHAP values.

We can conceptualize the goal of using SHAP to quantify momentum as a **cooperative game** with multiple participants, where each feature acts as a participant. However, due to the varying contributions of each participant, we employ Shapley values to fairly allocate each participant's **contribution** to the overall gain.

To obtain Shapley values, we performed a weighted average over all permutations of the subsets derived from the combination of influential indicator data (participants). The involved formula is as follows:

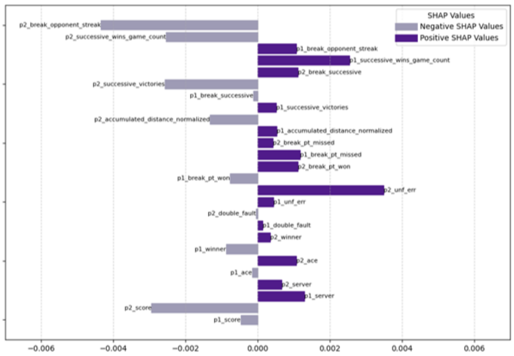
$$\phi_i(x) = \sum_{S \subseteq N \setminus \{i\}} w_i \cdot \frac{|S|! \cdot (|N| - |S| - 1)!}{|N|!} [f(S \cup \{i\}) - f(S)] \tag{4}$$

- where:
  - $\phi_i(x)$ is the Shapley value for player i.
  - S is any subset of the player set N that does not include player i.
  - w_i is the weight associated with player i.

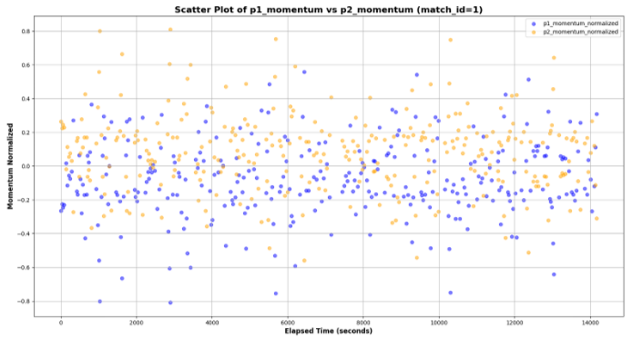
- $|S|$ represents the size of set S, i.e., the number of players in it.
- $|N|$ is the total number of players.
- $f(S)$ is the value or utility of coalition S.

Combining the Random Forest algorithm with the SHAP method, we obtained the final weights for each indicator shown in Figure 9. We successfully quantified and normalized the momentum by applying these weights in a weighted calculation, providing a comprehensive assessment.

Figure 9 Illustrates Quantitative Momentum Analysis of the First Match.



（a）The SHAP Importance level



（b）The Distribution of Momentum in Match 1

Figure 9: Quantization and Distribution of Momentum

7 Assessment of Momentum's Impact on Performance

Firstly, in this context, based on our analysis of match fluctuations, we consider them as pivotal turning points where significant changes in match outcomes may occur. These fluctuations can be reflected in a player's relative performance. Therefore, we interpret *swings in play* as the zero points for the player performance comparison function. Through the solution to the first question, we have quantified this concept.

Additionally, *runs of success* can be measured using two newly introduced variables: the count of successive victories within each game and the count of successive game wins in each set.

$$\text{(Swings in play)} \rightarrow \text{(performance\_difference\_change)} \tag{5}$$

$$\text{(Runs of success)} \rightarrow \text{(successive victories, successive wins game count)} \tag{6}$$

Insights from previous research have provided valuable guidance. Our assessment of the impact of momentum on player performance will unfold in two aspects: positive and negative.

7.1 Correlation Analysis

Beginning with a positive approach to address the issue, given that the fluctuations in match situations are discrete variables, we introduced the chi-square test to measure the correlation between momentum and player performance.

Additionally, since turning points emphasize variability, to enhance the rigor of our argument, we also incorporated **'momentum reversal'** (a discrete variable) into the process, with the implementation outlined as follows:

Step 1 Define the null hypotheses $H0$ respectively:

Independence between **'momentum'** and **'swings in play'**, between **momentum reversal** and **'swings in play'**, and between **'momentum'** and **'runs of success'**.

Step 2 Data preprocessing:

Since the chi-square test requires both variables in the correlation analysis to be discrete, we discretized the continuous variable 'momentum' into five levels (refer to the table), meeting the data requirements.

Table 5: Discrete division of momentum

Label	*Very Low*	*Low*	*Medium*	*Medium*	*Very High*
Range	(-1, -0.6)	(-0.6, -0.2)	(-0.2, 0.2)	(-0.2, 0.2)	(0.6, 1)

Step 3 Create a contingency table for the data. (Rows represent different levels of one categorical variable, and columns represent different levels of another categorical variable.)

Step 4 Calculate the chi-square statistic and p-value:

Firstly, the expected frequencies are the anticipated observed frequencies in each cell under the assumption of independence between the two variables. They were calculated using the formula 7. Using the obtained expected frequencies, apply them to the formula (8 to calculate the chi-square statistic.

$$E_{ij} = \frac{R_i \times C_j}{N} \tag{7}$$

Here, E_{ij} represents the expected frequency in the cell at the i-th row and j-th column, R_i is the total for the i-th row, C_j is the total for the j-th column, and N is the total sample size.

$$\chi^2 = \sum_{i=1}^{k} \sum_{j=1}^{m} \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}} \tag{8}$$

where k is the number of rows, m is the number of columns, O_{ij} represents the observed frequency, and E_{ij} represents the expected frequency.

Next, compare the calculated chi-square statistic with the chi-square distribution to obtain the p-value.

Step 5 Result analysis:

Table 6: Chi-Square Test results 1

Characteristic	Chi2 Statistic	P-value	Degrees of Freedom
successive_victories	5018.785179	0.00E+00	20
successive_wins_game_count	690.7988257	9.386167E-125	32

Table 7: Chi-Square Test results 2

Classification	Chi2 Statistic	P-value	Degrees of Freedom
Group 1	5.44062568	0.244998038	4
Group 2	10.78734278	0.244998038	8

Note: Group 1 refers to Shifts in Momentum & Turning Points;
Group 2 refers to Momentum & Turning Points;

Analyzing Table 6, we observed that the p-values for the 'successive wins' feature are all significantly less than 0.05. This provides sufficient evidence to reject $H0$, indicating a highly significant correlation between a player's successive wins and 'momentum'.

Examining Table 7, whether testing the correlation between 'momentum' and 'swings in play' or 'momentum reversal' and 'swings in play', the p-values exceed 0.2, greater than 0.05, suggesting a lack of correlation between these variables.

7.2 Stochastic Analysis

Taking a reverse error approach, we introduced a run test, examining the swings in play and runs of success for one player to determine whether they exhibit randomness. (That is, the null hypothesis $H0$ assumes randomness). Take the test of randomness for consecutive victories as an example:

Step 1 Define a run by grouping consecutive values in the sequence that share the same direction. In the context of a tennis match, when consecutive victories occur, they can be classified as a run. Next, for a given sequence of match results, we tally the actual observed number of consecutive victory runs.

Step 2 Calculate the expected number of consecutive victory runs in a theoretically random sequence of the same length, based on the assumption of randomness.

$$E(R) = \frac{2n_1 n_0}{n} + 1 \tag{9}$$

- where:
  - $E(R)$ represents the expected number of consecutive victory runs.

- n_1 is the count of consecutive victories in the match result sequence.
- n_0 is the count of consecutive defeats in the match result sequence.
- n is the total length of the match result sequence.

Step 3 Use standardized difference values to calculate the test statistic, measuring the difference between the sample data and the null hypothesis:

$$Z = \frac{R - E(R)}{\sqrt{\text{Var}(R)}} \tag{10}$$

Here, R is the actual number of consecutive victory runs, $E(R)$ is the expected number of consecutive victory runs, and $\text{Var}(R)$ is the variance of the consecutive victory runs.

To determine the significance of the test statistic, we calculated the p-value using the test statistic and the cumulative distribution function. If the p-value is less than the predetermined significance level (e.g., 0.05), we have sufficient evidence to reject the null hypothesis, indicating that the sequence is non-random. The results of the test statistic and p-value are presented in Table 8.

Table 8: Run test results

Characteristic	Test Statistics	p-value
p1_successive_victories	-4.177960901	2.941342E-05
p2_successive_victories	-5.307547785	1.111099E-07
p1_successive_wins_game_count	-64.97921161	0.000000E+00
p2_successive_wins_game_count	-64.96707587	0.000000E+00
performance_difference_change	1.993313483	4.622713E-02

Note: In the analysis of consecutive wins,$p = 0$ is a result of numerical precision limitations in computing, indicating that the p-values are extremely small.

By analyzing the above table, we can observe that the p-values for both "swings in play" and "runs of success" are less than 0.05. Therefore, we have sufficient evidence to reject the null hypothesis, indicating a rejection of the notion that the sequence is random.

In conclusion, through a comprehensive examination involving positive correlation analysis and reverse randomness testing, we conclude that in the match, both the swings in play and runs of success exhibit non-random patterns for a given player. Furthermore, the significant influence of consecutive victories on positive momentum suggests that the tennis coach's statement lacks validity.

8 Prediction for Fluctuations in Relative Momentum

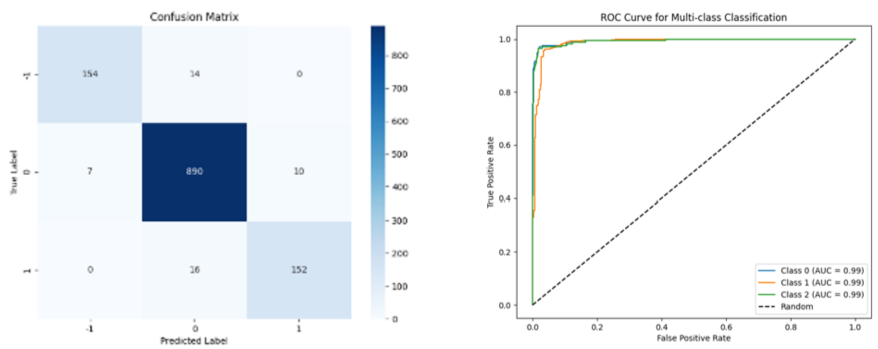
For question 3, firstly, we need to define 'change of the play flow'. Since both 'change of the play flow' and Momentum are short-term concepts, and based on Section 7, the metric for measuring Momentum is winning the next point, which is indeed a short-term assessment of the battle advantage. Therefore, this transition can be quantified using 'Relative Momentum', indicating the side with greater momentum is more favorable during the game.

8.1 Prediction for Swings

Building upon the quantified results of momentum in Section 7, with robust support from Python's extensive libraries, we introduce the LightGBM model to predict relative momentum during the match.

LightGBM is a machine learning algorithm based on a gradient boosting framework, well-suited for efficiently handling large-scale datasets and high-dimensional features. Its compatibility with our training and prediction requirements, based on high-dimensional datasets, aligns with the need to effectively process and analyze our data.

Due to the commonality in measuring momentum, we inherited the dataset comprising 26 indicators from the second question. Eighty percent of the data from the 31 matches were allocated as the training set for the LightGBM predictive model, with the remaining 20% designated as the testing set. The relative values of the quantified momentum, as addressed in the second question, were introduced into the model's training set as the actual change quantity.



(a) (b)

Figure 10: Model performance evaluation

We utilized a pre-trained model to predict the test set and assessed the model's performance on the test set using the confusion matrix and ROC curve. The model achieved a prediction probability of **96.22%**, indicating excellent results. Figure 11 illustrates the prediction results for the 30th match.

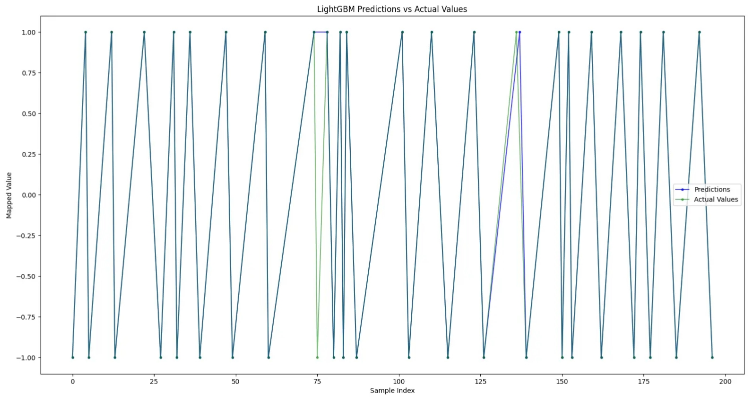


Figure 11: Prediction Result of Match 30

8.2 Factor Analysis

Based on the model output, we selected the top 5 factors with the highest correlation.

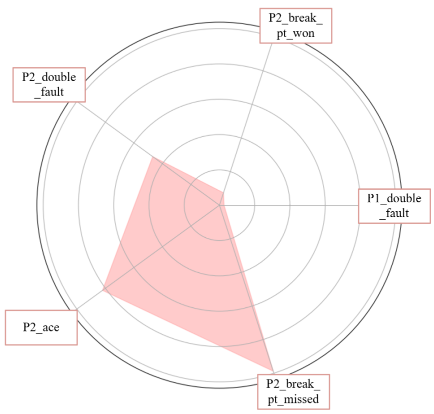


Figure 12: Important factor

8.3 Advice on a New Match Base on Swings of Momentum

In the previous summary, we investigated the role of momentum in player performance and observed significant variations in momentum fluctuations. Confronted with pre-match analysis for a new competition against a different player, we approached it by extracting crucial data from past matches and analyzing both general and individual aspects. Through a comprehensive examination of momentum fluctuations, we have formulated detailed recommendations:

- General Patterns of the Match:

  Studying the positive and negative changes in relative momentum throughout the entire game allows us to ascertain the distribution of situational fluctuations. This information can then be utilized for tactical planning.

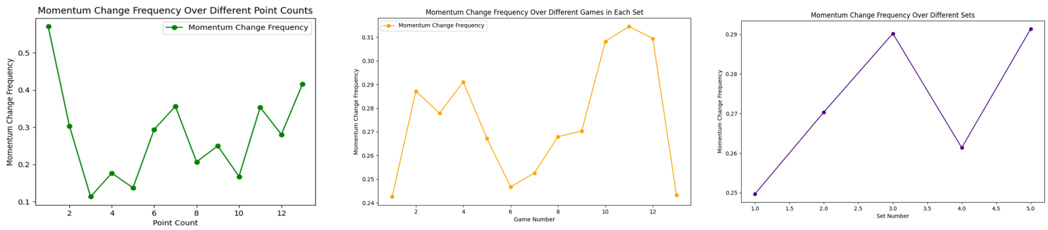


Figure 13: Momentum Swings of Schedule

The graph reveals that momentum fluctuates more significantly during individual games, with more pronounced variations in sets. Therefore, it is advisable to display a proactive approach, capitalize on advantages, and maintain composure, especially during the initial stages of the match.

- Opponent's Analysis:

  Quantifying momentum is a course full of complexity. Ben (2015)[3] explicitly pointed out in their study that the influence of momentum varies among different athletes. Therefore, it becomes crucial to assess the impact of momentum on both the opponent and oneself before the match. This facilitates the formulation of targeted tactical strategies and psychological guidance preparations.

9 Model Testing and Generalization

9.1 Data preparation

First, we collected match data for WTA, US Open, and badminton men's singles from Table 9 shown. The collected data underwent thorough processing to address missing values and outliers. Scores of matches were transformed, following the methodology outlined in the previous section. For tennis matches, the score transformation remained consistent, while for badminton matches, scores were segmented per ball.

Table 9: Data and Database Websites

Data Name	Database Websites
Women's Single in WTA	https://www.kaggle.com/datasets/thedevastator/ wta-tennis-match-results-from-2016-and-2018
2023-usopen-points-doubles	https://github.com/JeffSackmann/ tennis_slam_pointbypoint /blob/master/ 2023-usopen-points-doubles.csv
Badminton Men's Singles	https://www.kaggle.com/datasets/sanderp /badminton-bwf-world-tour

9.2 Model Generalization

The generalization of the quantitative momentum model and LightGBM prediction model extended to tennis women's singles (across different court materials), men's doubles, and badminton men's singles matches. Women's singles and badminton matches lacked detailed ball-by-ball information. Ultimately, 19 and 25 features were considered for women's singles and men's doubles, achieving prediction correctness rates of **66.38%** and **84.75%**. Future models may benefit from incorporating player strength (ranking) and the impact of schedules on momentum.

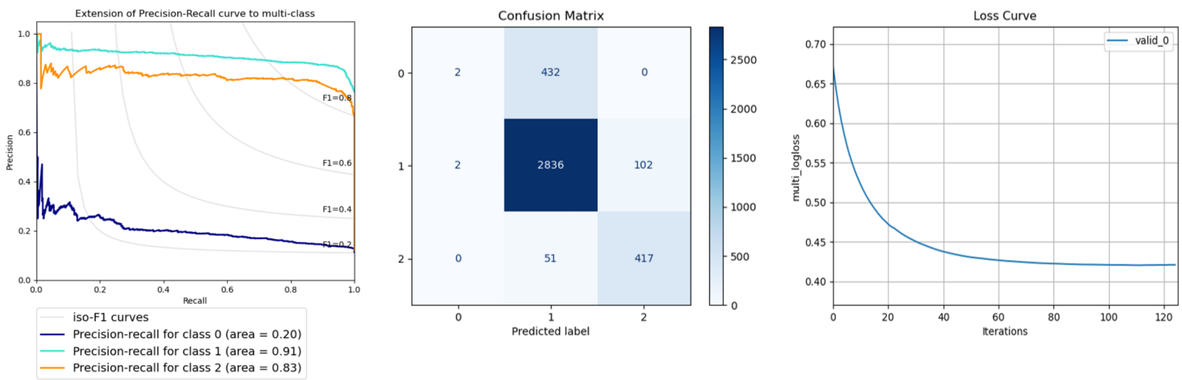


Figure 14: Model performance evaluation

However, Challenges arose when applying the predictive model to badminton men's singles, despite successfully calculating momentum. The dataset recorded only set scores, lacking details like whether the player smashed. This absence of crucial information made predicting momentum changes challenging.

In summary, our model generalizes well across tennis matches, dependent on data feature columns' quantity and quality. Its generalization to datasets from other sports with insufficient features is weaker.

10 Sensitivity Analysis

Through the preceding SHAP analysis, it is determined that the two features exerting the most significant numerical impact on momentum are $p2\_unf\_err$ and $p2\_break\_opponent\_streak$. We adjusted the weights of these two critical features by increasing them to positive and negative directions by **15%**, respectively. Subsequently, we computed the differences between the adjusted p1_momentum values and the original values.

Encouragingly, the differences between the adjusted p1_momentum values and the original values are extremely small, only 1.67×10^{-17}, representing a minuscule percentage of the original values at $2.59 \times 10^{-12}\%$. This implies that even with a 15% positive or negative adjustment in the weights of the two most influential features,

$p2\_unf\_err$ and $p2\_break\_opponent\_streak$ the impact on the variation of $p1$_momentum remains negligible.

This result reveals that in the process of constructing the momentum model, adjusting the weights of the two most impactful features, $p2\_unf\_err$ and

$p2\_break\_opponent\_streak$, by 15% in either direction has minimal influence on the changes in $p1\_momentum$. It indicates that individual feature weight changes have a minor impact on the model's fluctuations in the current scenario, demonstrating the **stability and adaptability** of the model.

11 Model Evaluation and Further Discussion

11.1 Strengths

1. The model exhibits strong stability and adaptability. It considers 26 features when quantifying momentum, ensuring that it remains functional without being overly affected by changes in a few features.

2. It achieves an exceptionally high accuracy rate in predicting swings in the match. Utilizing LightGBM for predictions, the model can reach an accuracy rate of 96.22%.

3. The model demonstrates good generalization capabilities across other tennis matches. It performs well when applied to tournaments, women's singles, doubles, and tennis matches on different surfaces.

11.2 Weaknesses And Further Discussion

1. The model for quantifying momentum has high demands on the dataset content. It requires the dataset to record match data by point; otherwise, the model's performance might be suboptimal.

2. There is room for improvement in its generalization ability. When applying the model to other sports events (such as badminton matches), the lack of data might prevent the model from being properly established.

12 Memorandum

To: Coaches

From: Team # 2422432

Subject: Understanding and Utilizing Momentum in Tennis Matches

Date: February 6, 2024

Tennis, as a classic sports event, often witnesses various suspenseful fluctuations in matches, and the emergence of these fluctuations is crucial, significantly influencing the outcome of the competition. Reflecting on the intense competition in the men's singles final at Wimbledon in 2023, it prominently highlighted the importance of shifts in momentum, sparking passionate discussions among the public about the factors influencing these fluctuations. Experts have introduced the concept of "momentum" to analyze these dynamics. Clarifying the generation and impact of "momentum," that is, identifying the lifeblood of match fluctuations, can assist in planning better preparation strategies and formulating correct game strategies.

Building on this foundation, our team conducted a comprehensive study based on the 2023 Wimbledon men's singles data, focusing on the concept of "momentum" and successfully quantifying momentum to predict match fluctuations.

We initially utilized the LSTM model to quantify player performance, then allocated reasonable weights to influencing factors through a random forest model based on SHAP explanations, successfully quantifying momentum. Combining these two approaches, we obtained the correlation between momentum and player performance, exploring the role of momentum.

In our research, we fully considered the impact of serve and return games on winning, added an attention module, and conducted a comprehensive analysis of the role of momentum from both positive and negative perspectives, achieving a thorough and insightful examination of the issue.

Based on our research, we have drawn the following conclusions:

- Positive momentum has a beneficial effect on player performance, especially in terms of consecutive scoring.
- Momentum is most influenced by the player's physical condition (e.g., fatigue level) and psychological state (e.g., score).
- Relative momentum fluctuations follow certain patterns during the course of a

match, especially in the early and late stages of a set, with significant fluctuations during tiebreaks.

Therefore, considering the strong impact of momentum on player performance, we recommend the following for planning a reasonable preparation strategy for your team:

- Pre-match Analysis:

  1. Identify fluctuation patterns, focus on mental conditioning, and conduct adaptive training.
  2. Given that momentum is influenced by physical fitness, enhance physical training.

- In-Match Attention:
  Pay attention to your momentum state, adjust as needed to ensure focus and maintain emotional control.[2]

  1. When momentum is favorable: Maintain a stable state, observe changes in the opponent, and control the court.
  2. When momentum is unfavorable: Slow down the pace, intensify strategic thinking, and alter tactics.
  3. During momentum shifts: Adjust tactics promptly, maintain your own rhythm.

We hope that your team can benefit from these recommendations and achieve better results in tennis competitions.

Yours Sincerely,

Team # 2422432

References

[1] 2023 wimbledon men's singles final: In a five-set epic battle, alcaraz terminates djokovic's consecutive victories streak to clinch the championship. https://olympics.com/zh/news/wimbledon-2023-carlos-alcaraz-dethrones-novak-djokovic-mens-final.

[2] Momentum - The Hidden Force in Tennis | Semantic Scholar.

[3] Momentum in US Open men's singles tennis: International Journal of Performance Analysis in Sport: Vol 15, No 3.

[4] Peter O'Donoghue and Emily Brown. Sequences of service points and the misperception of momentum in elite tennis. *International Journal of Performance Analysis in Sport*, 9(1):113–127, April 2009. Publisher: Routledge _eprint: https://doi.org/10.1080/24748668.2009.11868468.

[5] John M. Silva, Charles J. Hardy, and R. Kelly Crace. Analysis of Psychological Momentum in Intercollegiate Tennis. *Journal of Sport and Exercise Psychology*, 10(3):346–354, September 1988. Publisher: Human Kinetics, Inc. Section: Journal of Sport and Exercise Psychology.

[6] Jim Taylor and Andrew Demick. A multidimensional model of momentum in sports. *Journal of Applied Sport Psychology*, 6(1):51–70, March 1994. Publisher: Routledge _eprint: https://doi.org/10.1080/10413209408406465.

Report on Use of AI

1. OpenAI ChatGPT (Feb 4, 2024 version, ChatGPT-3.5,)

Query1: Please help me explain Tennis rules

Output1: Certainly! Tennis is a racquet sport played between two individuals (singles) or two pairs (doubles) on a rectangular court divided by a net. The basic rules of tennis include scoring, serving, and gameplay. Here's a brief explanation:

1. Scoring: - The game is scored using a system of points: 15, 30, 40, and game. The first player or team to win four points wins the game. - If both players or teams reach 40 points (known as deuce), a player or team must win two consecutive points to win the game.

2. Sets: - A match is typically divided into sets, and to win a set, a player or team must win six games, with a margin of at least two games. - If the set reaches a 6-6 tie, a tiebreak may be played to determine the winner of the set.

3. Matches: - Matches are usually the best of three sets (for women) or the best of five sets (for men) in major tournaments. - To win a match, a player or team must win the majority of sets. In a best-of-five-set match, the first player or team to win three sets wins the match.

4. Serving: - Players take turns serving, and the server must stand behind the baseline and within the singles or doubles sidelines. - The serve must land within the opponent's service box, diagonally opposite, and must clear the net. If the serve hits the net but lands in the service box, it's called a "let," and the server gets another chance.

5. Faults and Double Faults: - If a server fails to get the ball into the opponent's service box, it's called a "fault." Two consecutive faults result in a "double fault," and the opponent wins a point.

6. In and Out: - If the ball lands outside the boundaries of the court, it's considered out. If the ball lands on the lines, it's considered in.

7. Tiebreaks: - In tiebreak situations, players take turns serving two points at a time, and the first player or team to reach seven points (with at least a two-point advantage) wins the tiebreak.