

Momentum Ongoing Model in Tennis with Quantification Optimization and Regression

Summary

Since 1877, Wimbledon symbolizes British tennis. Momentum is vital in tennis, affecting players' confidence and match outcomes. Data analytics and machine learning now provide concrete insights, aiding predictive models. Understanding momentum is crucial for optimal performance, effective strategy, and grasping players' psychology.

For Task 1, to capture the flow of play and analyze tennis players' performance, we first conducted data preprocessing. We then focused on key performance indicators such as First Serve Success Rate and Unforced Errors. By integrating a **Success Rate model** with a **Weight Model** optimized by **LAHC(The late acceptance Hill-Climbing) Algorithm**, we established a **SEMOLA Model(Success Rate Model combined with Weight Model optimized by LAHC)** that quantifies athletes' performance, taking into account both visible and less obvious factors. We used a **Forest-RC Model** to predict players' win probabilities. The Forest-RC Model addresses data sparsity issues and the LAHC Model ensures jumping out of the local optimum solution and accelerates the model's computational speed respectively. Our approach transcends mere victory-defeat assessments, delving into a holistic appraisal of athletic prowess. It sheds light on the strategic and psychological underpinnings that sway match outcomes, thereby furnishing a richer, more nuanced understanding of competitive play.

For Task 2, we explored the relationship between scoring and momentum in tennis matches, adopting a systematic approach to quantify momentum from various angles, including **Consecutive Scoring, Break Points, and First Serve Success Rates**. We used **chi-square tests** to establish a significant correlation between consecutive scoring and match outcomes. Subsequently, we also employed **ANOVA** to validate the relationship between outcomes of the scoring evaluation model and momentum. Our findings underscore the critical role of scoring continuity and effective break strategies in competitive tennis, offering significant implications for training and player development.

In Task 3, we utilized the **FP-Growth tree algorithm** for analyzing and modeling match fluctuation factors, focusing on identifying the most relevant indicators through frequent pattern mining. By employing the Kulczynski measure (Kluc) and the Imbalance Ratio (IR), **Break Points Won, Winner and Net Play Success Rate** were found to be the most frequent of all indicators. We then constructed a **Dual-objective Optimization and Prediction Model** to classify match fluctuations as positive, negative, or neutral based on momentum variables, applying the **NSGA-II algorithm** for parameters optimization and Logit Regression to evaluate the explanatory power of selected variables. This method provides strategic insights into managing match momentum.

For Task 4, we tested the predictive model's validity and generalizability across tennis matches. The model showed predictive prowess with AUC values indicating its capability to forecast momentum shifts, despite needing further refinement. Validation against additional matches and data from the U.S. Open women's event confirmed the model's accuracy, affirming its effectiveness.

Keywords: SEMOLA(Success Rate Model, Weight Model, LAHC) Forest-RC ANOVA
Frequent pattern mining NSGA-II Dual-objective Optimization

Contents

1	Introduction	2
1.1	Background	2
1.2	Problem Restatement and Problem Analysis	2
1.3	Our work	3
2	Assumptions and Notations	4
2.1	Assumptions and justification	4
2.2	Model Notations	5
3	Task 1	6
3.1	Data Preprocessing and Analysis	6
3.1.1	Addressing the missing values and the outliers	6
3.1.2	Data analysis	6
3.2	SEMOLA Model	7
3.2.1	Success Rate Model	8
3.2.2	Weight Model	10
3.3	Visualization and Outcome	12
4	Task 2 & Task3	15
4.1	Task 2	15
4.1.1	Differential Analysis	15
4.2	Task 3	17
4.2.1	Match Fluctuation Factors Analysis: FP-Growth tree	17
4.2.2	Dual-objective Optimization and Prediction Model	20
4.2.3	Strategic Advice Based on Historical Momentum Differential in Matchups	22
5	Task 4: Model Test	22
5.1	Model Validation	22
5.2	Model Generalizability	22
6	Sensitive Analysis	23
7	Conclusion	23
7.1	Strengths and Weaknesses	23
8	Memorandum for The Chief Administrator	24

1 Introduction

1.1 Background

The Wimbledon Championships is an annual British tennis competition that was established in 1877. The tournament takes place on the outdoor grass courts of the All England Lawn Tennis and Croquet Club (AELTC) in the Wimbledon suburb of London. The men's singles was the first event contested in 1877. Since the inaugural tournament, the rules for men's singles have undergone several changes. From 1878 to 1921, the competition began with a knockout stage, with the winner challenging the defending champion. If the champion was absent, the competitor would win by default. In 1922, the challenge round system was abolished, and matches were played as best of five sets. Initially, a single game win after a five-all tie was sufficient for a set victory; from 1884 to 1970, a two-game advantage was required. The tiebreak system was introduced in 1971, first used after an eight-all tie, and from 1979, it was applied after a six-all tie.

In tennis, momentum plays a pivotal yet often unnoticed role, determining the course of a match. It encompasses both psychological and strategic advantages gained through successful actions. Psychologically, momentum boosts confidence and determination while pressuring opponents. This influences decision-making and performance. Advancements in big data and machine learning have made momentum research more scientific. Researchers now collect and analyze extensive match data to quantify momentum's impact and create predictive models for match outcomes. These studies enhance result prediction accuracy and offer valuable insights for athlete training. Athletes can better comprehend momentum patterns, allowing them to harness or reverse it during actual competitions.

In conclusion, as a key factor in tennis matches, the study of momentum will continue to provide important perspectives and tools for enhancing athlete performance, optimizing match strategies, and deepening our understanding of the psychological and strategic aspects of the sport.

1.2 Problem Restatement and Problem Analysis

- Task 1:

(1) We need to develop a model that captures the flow of play at the moment scoring occurs and apply it to one or multiple matches. The model we establish should identify which player is performing better at specific times in the match and how well they are performing. Notably, in tennis matches, the server has a significantly higher probability of winning points. We need to incorporate this aspect into our model somehow. This issue falls within the realm of predictive modeling and statistical analysis, particularly in the field of sports data analytics. To address such problems, we consider creating a scoring evaluation model to depict players' performances and capture the flow of play as scores are made. This not only aids in understanding the dynamic shifts during a match but also provides a basis for strategic decision-making for coaches and players.

(2) Utilize our constructed model to visually describe the match flow.

- Task 2:

Tennis coach is skeptical about the role of "momentum" in matches, positing instead that a player's fluctuations in performance and successful plays are random. Using our model/metrics, evaluate this hypothesis. The concept of momentum encompasses the advantage a player builds through

a series of successful actions or consecutive scores, both psychologically and in game strategy. It's one of those intangible factors in a match. In this section, we need to investigate whether gaining and losing momentum during the intense confrontations of singles tennis matches affects the outcomes. By developing a model to observe if players' fluctuations and successful plays are random, we can assess the tennis coach's hypothesis.

- Task 3:

(1) This section primarily aims to explore which factors are most correlated with fluctuations during a match. (2) By developing a model to predict these fluctuations, we must include crucial indicators that assist coaches in determining when the flow of play is likely to shift from favoring one player to another.

(3) Considering the variance in "momentum" from past matches, how would you advise a player for a new match against another? In this section, based on momentum fluctuations, we need to provide relevant advice or strategies to increase the likelihood of scoring for the player.

- Task 4:

(1) Test the model you developed on one or more additional matches. How accurate are your predictions of match fluctuations? We should evaluate the model's effectiveness using metrics like AUC. If the model sometimes performs poorly, we should identify any factors that may need to be included in future models.

(2) How generalizable is our model to other matches (e.g., women's matches), tournaments, court surfaces, and other sports like table tennis? This requires testing the universality of our model using external data.

- Task 5:

Write a one to two page memo to the coach summarizing our findings on the role of "momentum" in matches and recommendations, as well as how to prepare players to cope with the flow of play in tennis matches.

1.3 Our work

Our work is shown in Figure 1. The specific steps are as follows:

Step 1: Data preprocessing to detect any missing or anomalous values and address them, yielding the final dataset. To more accurately analyze and understand players' performances, we condensed the data into 9 indicators.

Step 2: We capture the flow of play occurring during scoring by establishing a scoring evaluation model $G = W \times S$, which combines a weight model (W) and a success rate model (S). This model helps us determine which player performs better at specific times in the game and assess the quality of their performance.

Step 3: Provide a model-based visualization to describe the match flow.

Step 4: Quantify "momentum" to explore its role in matches.

Step 5: Use the model to observe if players' fluctuations and successful runs are random, thus evaluating the tennis coach's hypothesis.

Step 6: Investigate the factors most relevant to in-match fluctuations .

Step 7: Predict in-match fluctuations through modeling.

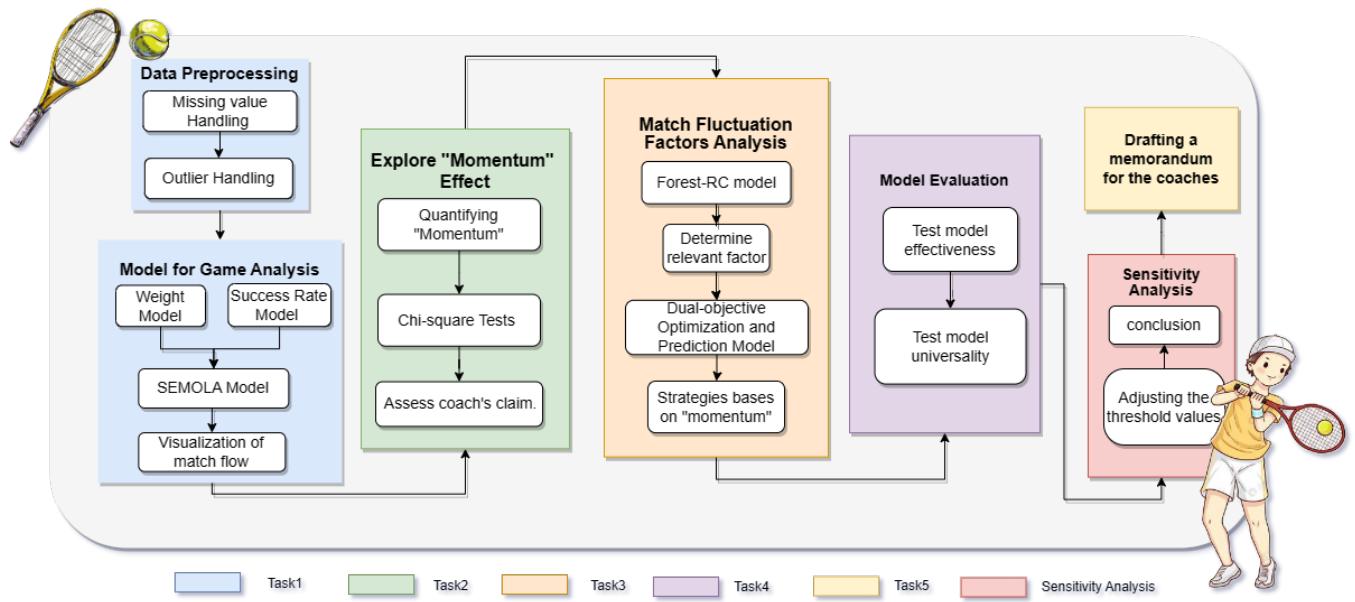


Figure 1: Our Work

Step 8: Offer strategies based on momentum fluctuations to increase the likelihood of scoring for the player receiving the strategy.

Step 9: Evaluate the model's effectiveness and identify any factors that may have been overlooked if the model underperforms.

Step 10: Test the model's universality.

Step 11: Sensitivity analysis.

Step 12: Write a one to two page memo to the coach summarizing your findings on the role of "momentum" in matches and recommendations, as well as how to prepare reserve players to cope with the flow of play in tennis matches.

2 Assumptions and Notations

2.1 Assumptions and justification

- **Assumption1:** An athlete's true level of skill cannot be solely determined by a single victory.
Justification: The level of a player's performance in a game may not always be at its peak.
- **Assumption2:** The actual level of performance an athlete displays on the field is influenced by latent factors.
Justification: Well-trained athletes can be affected psychologically by matters outside the competition.
- **Assumption3:** Assumed that halftime breaks provide similar benefits to both parties, without considering the impact of halftime breaks on either side's momentum.

Justification: The number of halftime breaks is limited, and since rest periods are equal and brief for all athletes, the differences are considered negligible.

- **Assumption4:** Reliability of Historical Data Assumption: Assumes that the historical price data used is accurate, complete, and reliable.

Justification: If the historical data is inaccurate, incomplete, or unreliable, any conclusions drawn from it may be flawed, leading to incorrect insights.

- **Assumption5:** Assuming that all the indicators included in the given dataset encompass most of the information about the flow of the match.

Justification: The rationale for this assumption is to ensure that any analysis, forecast, or decision based on this historical data is built upon a foundation deemed to be accurate and trustworthy.

2.2 Model Notations

Table 1 and Table 2 presents the primary metrics considered in our study. Other symbols appearing in the text are defined below their respective formulas.

Table 1: Notations 1

Symbol	Definition	Specific Significance
V	victory	Player won a game this point
C_A	Aces	The number of times an untouchable winning serve was hit.
C_W	Winners	The number of shots that result in a point won by a player where the opponent is unable to touch the ball with their racket.
C_{DF}	Double Faults	Occurs when a player fails to get their first and second serve into the opponent's service box, resulting in the loss of the point.
C_{UE}	Unforced Errors Count	Errors in play that are not due to the opponent's good play, typically resulting from a player's own mistake during a rally.
C_{FS}	First Serve Count	The total number of first serve attempts a player makes during a match.
C_{FSS}	First Serve Success Count	The total number of first serve attempts a player successfully makes during a match.
C_{SS}	Second Serve Count	The total number of second serve attempts a player makes during a match following a failed first serve.
C_{SSS}	Second Serve Success Count	The Number of points won on the second serve after a first serve fault.
C_{Ne}	Net Approaches	The number of times a player approaches the net in an attempt to play a volley or a smash.

Symbol	Definition	Specific Significance
C_{NeS}	Net Points Success Count	The number of points a player wins during net approaches.
C_{TPS}	Total Points Success Count	The total number of points a player has won in a match.
C_{BPS}	Break Points Success Count	The number of times a player wins a point that results in breaking the opponent's service game.

Table 2: Notations 2

Symbol	Definition	Specific Significance
R_A	The percentage of Aces	The percentage of direct point-winning serves.
R_{FSS}	First Serve Success Rate	The percentage of a player's first serves that successfully land in the service box.
R_{SSS}	Second Serve Success Rate	The percentage of second serve attempts that are successfully placed into play.
R_{NPS}	Net Play Success Rate	The percentage of points won at the net out of total net approaches.
R_{BPW}	Break Points Won Rate	The percentage of points a player successfully broke the opponent's service game.

3 Task 1

3.1 Data Preprocessing and Analysis

3.1.1 Addressing the missing values and the outliers

Addressing missing values and outliers is crucial in data analysis. Missing values can lead to biased results, while outliers can distort the accuracy of statistical models. We used techniques like imputation and trimming help handle missing values and outliers, ensuring more reliable and robust data analysis and decision-making. For example, according to the match rules, the scores can theoretically be among 0, 15, 30 and 40, but scores in the given data also range from 0 to 9, which may owe to incorrect records. Considering its rarity, removing the whole line will not have a huge influence.

3.1.2 Data analysis

Key performance metrics, such as the serving efficiency, effectiveness, and error rate of the serving player, are crucial tools for evaluating a player's performance. Like in other sports, tennis involves a wealth of statistical data, where metrics like first-serve success rate, first-serve points won, and aces serve as critical indicators for assessing a player's competitive level. These metrics not only reflect

a player's fundamental skills but also reveal their strategic choices and mental resilience during the game.

Moreover, data on serve speed, receiving success rate, net success rate, and break point conversion highlight a player's playing style, offering in-depth insights into their tactical preferences and on-court behavior. An increase in unforced errors and double faults often indicates fluctuations in a player's mental state or a decline in physical stamina, both of which are significant risk factors in a match.

To analyze and understand a player's performance more precisely, all these data can be summarized into nine key indicators, as follows: Aces(C_A), Winners(C_W), Double Faults(C_{DF}), First-Serve Success Rate(R_{FSS}), Second-Serve Success Rate(R_{SSS}), Unforced Errors(C_{UE}), Net Points Success Count(C_{Nes}), and Break Points Won(C_{BPW}). These indicators, detailed in Table 1 and 2 , each carry specific significance, covering various aspects from basic skills to advanced tactics. Through these indicators, coaches and players can better assess their strengths and weaknesses, devise tailored training plans, and formulate match strategies. This not only aids players in making quick and effective adjustments during matches but also enables them to continually improve throughout their training journey, ultimately elevating their overall competitive level and match performance.

3.2 SEMOLA Model

In discussing how to quantify an athlete's true level of ability and determine who performs better during a specific period in a match, we are faced with a highly complex and challenging issue. Commonly, the outcome of a match, whether a win or a loss, is seen as the sole indicator of an athlete's true talent. However, this simplified method overlooks many key factors, such as the intricate details of an athlete's performance during a match, their consistency, and certain "invisible" factors like psychological state and competitive experience, all of which significantly impact an athlete's actual performance.

Therefore, to more accurately assess an athlete's level, we propose a comprehensive model that considers not only the athlete's in-match performance but also attempts to capture those less observable elements. Our model is based on two main components: a Weight Model W and a Success Rate Model S, which mutually define a Score Evaluation (SEMOLA Model) ($G = W \times S$), quantifying the athlete's performance at each scoring moment.

The **Success Rate Model (S)** focuses on assessing the probability of the athlete's victory. It centers on evaluating the effectiveness of a player's actions in a game. It quantifies the successful outcomes of specific actions, such as serves, returns, or strategic plays, within a match. By assessing the ratio of successful to attempted plays, the model provides a nuanced view of performance that transcends mere win-loss records. It considers the intricacy of skill execution and decision-making that contributes to each point won, offering a detailed measure of a player's competency in various aspects of the game.

The **Weight Model (W)**, in contrast to the Success Rate Model, which predominantly measures outcomes, the weight model dives deeper into the nuances of a player's performance. It takes into account a broader range of factors that influence a player's game, such as consistency, pressure handling, and adaptability. This comprehensive approach allows for a more granular analysis of a player's abilities, highlighting strengths and pinpointing areas that may require further development.

Generally, the defined Score Evaluation Model (SEMOLA) provides a quantified assessment at each scoring instance, reflecting not just the athlete's instantaneous performance but also the context of the match and the athlete's stability. Notably, our model uniquely acknowledges that the athlete who loses a match does not necessarily have a lower real ability than the one who wins. This is because

sports outcomes are influenced by a multitude of factors, including luck, referee decisions, weather conditions, etc., all of which can affect the final result.

By using this comprehensive evaluation method, we can provide a more complete and objective assessment of athletes' performances, thereby assisting coaches, players, and even sports analysts to better understand the process of the game and the athletes' true levels. The application of this method is not limited to post-match analysis but can also offer real-time feedback during the match, helping coaches and athletes adjust their tactics and strategies to enhance the overall level and spectatorship of the game.

In summary, our proposed model quantifies athletes' performances through an innovative approach that considers not only the scores in the match but also a comprehensive assessment of both visible and invisible factors. This method offers a deeper and more comprehensive understanding of athletes' true levels during a match, breaking away from the traditional limitation of evaluating based solely on win or loss, and bringing new insights and standards of evaluation to the field of sports.

3.2.1 Success Rate Model

In the process of in-depth data analysis, we encountered a common but challenging problem: the matrix composed of variables is extremely sparse. The direct consequence of a sparse matrix is the significant increase in the complexity of training and modeling, especially when our target model needs to capture the flow of play. To address this issue, we adopted a strategy of modeling each point as a unit. The initial intention of this approach was to alleviate the sparsity by reducing the granularity of analysis; however, even so, the processed data matrix still exhibited a high degree of sparsity. Under such circumstances, the direct application of a linear regression model fell far short of the standard, and traditional decision tree models were also ineffective in dealing with sparse data, unable to extract the key weights for prediction.

Faced with these challenges, we turned to an improved **Forest-RC** model. The Forest-RC model, also known as the Oblique Decision Tree model, is specifically designed to overcome the problems posed by sparse datasets, providing a more flexible classification method than traditional decision trees. Unlike standard decision trees, which make judgments on a single attribute at each node, the Forest-RC **fits a linear classifier at each decision node**. This means that its classification boundaries are no longer just vertical or horizontal lines but can consist of multiple oblique segments, greatly enhancing the model's ability to handle complex data structures and making the predictions more accurate and reliable.

To elaborate on our methodology, the model employs a crucial formula that underpins the decision-making process at each node within an oblique decision tree. Here is the pivotal equation:

$$\sum_{k=1}^p a_k x_k + a_{p+1} \leq 0, \text{ where } a_1, a_2, \dots, a_{p+1} \in \mathbb{R}. \quad (1)$$

This formulate 1 delineates a hyperplane in a p -dimensional space, where each a_k represents the weight assigned to the corresponding feature x_k , and a_{p+1} is the bias term. The set of weights a_1, a_2, \dots, a_{p+1} , belonging to the real numbers (\mathbb{R}), are optimized during the tree's construction. When the weighted sum of the input features and the bias term is less than or equal to zero, the model predicts one class; otherwise, it predicts another. This approach allows for more complex and nuanced decision boundaries compared to traditional decision trees, which tend to create boundaries parallel to feature

axes. The capability to draw oblique boundaries enables the model to capture more intricate patterns in the data, often leading to improved classification accuracy.

Based on commonly referenced factors for measuring an athlete's individual ability in literature, we filtered the independent variables related to the winning probability in our dataset, which include: $C_A, C_W, C_{DF}, R_{FSS}, R_{SSS}, C_{UE}, R_{NPS}, C_{BPW}$

The first serve success rate is calculated as:

$$R_{FSS} = \frac{C_{FSS}}{C_{FS}} \quad (2)$$

The second serve success rate is calculated as:

$$R_{SSS} = \frac{C_{SSS}}{C_{SS}} \quad (3)$$

The net play success rate is calculated as:

$$R_{NPS} = \frac{C_{NeS}}{C_{Ne}} \quad (4)$$

Given the binary nature of our model's output, which delineates victory and defeat as 1 and 0 respectively, we designate the game outcome between two players as our dependent variable. This methodology constructs a dataset of positive and negative instances, thereby preventing data skewness and facilitating the foundation for advanced regression analysis. And the equation 5 represents the results of the regression.

$$W = 0.2099C_A + 0C_W - 0.0191C_{DF} + 1.6554R_{FSS} + 1.0576R_{SSS} \\ - 0.2680C_{UE} + 0.2190R_{NPS} + 0.2362C_{BPW} - 0.9317 \quad (5)$$

The model adopts a linear equation at every bifurcation point as a partitioning criterion, effectively segmenting the dataset into x distinct branching nodes from an initial set of approximately 2000 training instances. This segmentation strategy is pivotal for calculating the likelihood of each contestant's victory at various scores, thereby enabling a comprehensive understanding of competitive dynamics. The ROC curve, delineated in Figure 2, serves as a testament to the model's predictive proficiency and its capacity to discriminate between potential outcomes with a high degree of accuracy.

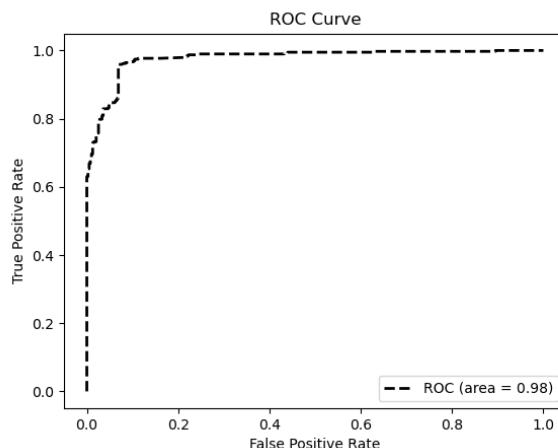


Figure 2: ROC Curve demonstrating the model's discriminative capability

Figure 2 displays a Receiver Operating Characteristic (ROC) curve, which is a graphical plot that illustrates the diagnostic ability of a binary classifier system. The curve plots the True Positive Rate (TPR) against the False Positive Rate (FPR) at various threshold settings. The dotted line represents the ROC curve, and the area under the curve (AUC) is a measure of the model's accuracy, which in this case is 0.98. An AUC of 1 represents a perfect model, while an AUC of 0.5 suggests no discriminative ability. The high AUC in this graph indicates that the model has an excellent capacity to distinguish between the two classes, which are likely to be 'win' and 'loss' in this context.

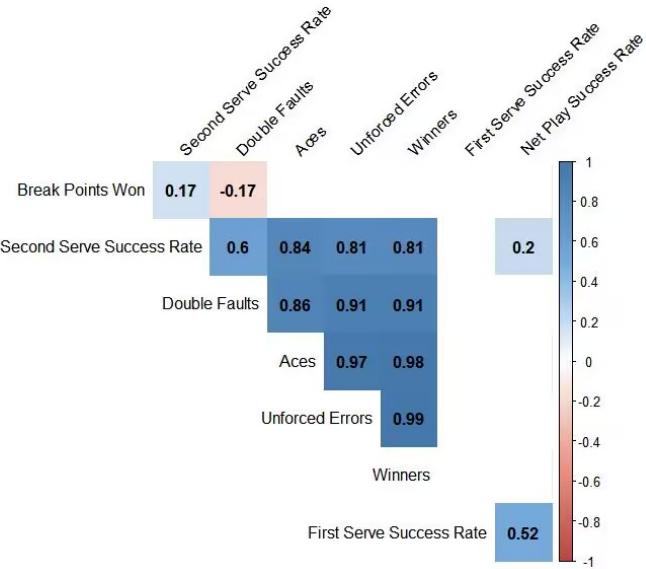


Figure 3: Correlation Matrix of Tennis Match Statistics

Figure 3 provides a correlation analysis of these variables, where cooler colors indicate positive correlations, warmer colors signify negative correlations, and crosses represent insignificance in correlation. It is observable that the number of Winners (C_W) and Unforced Errors Count (C_{UE}) show significant correlation, which could be the reason for the coefficient of (C_W) being zero in the model.

3.2.2 Weight Model

Compared to the Success Rate Model, the Weight Model more significantly focuses on assessing an individual player's abilities. Specifically, the weight model works in tandem with the Success Rate Model, together defining a scoring evaluation model. This model employs a **dual-standard** approach to prevent certain data from significantly influencing the outcomes, thereby enhancing robustness. Additionally, as the weight model revolves around parameter optimization, we utilized a Hill-Climbing algorithm with a strategy combined with the late acceptance and greed to optimize and fit the parameters of the weight model. The initial model we selected is represented by Equation 6:

$$S(X, Y, Z, S1, S2) = \left(aX + bY + cZ + \frac{f}{1 + e^{d \times S1}} \right) \frac{S2}{S1} \quad (6)$$

In the model 6, a , b , c , d , and f represent the parameters, while X , Y , Z , $S1$, and $S2$ are the independent variables. Within a game, X is the Ace Rate(R_A), Y is the Break Points Success Rate(R_{BPW}), Z is the Net Play Success Rate(R_{NPS}), $M1$ is the distance run by the player, and $N1$ is the number of points, resulting in $S1 = M1/N1$ which denotes the average running distance per point. $M2$, $N2$, and $S2$ similarly define these variables for the opponent. Through Figure 3, we can observe that the four indicators—Double Faults, Aces, Unforced Errors, and Winners—exhibit strong correlations with one another. Furthermore, upon reviewing related literature, we understand that Aces can be indicative of a player's individual skill level. Consequently, we have chosen Aces to represent these four indicators, designating it as variable X . Additionally, from the Remaining indicators, we have selected two that reflect individual player capabilities:Break Points Success Rate and Net Play Success Rate , which we designate as variables Y and Z , respectively. These choices are grounded in a desire to encapsulate both the skill and strategic prowess exhibited during play. Additionally, the dependent variable is the score obtained at the end of each point.

Beyond common metrics such as Aces, Net Success Counts, and Winning Shot Numbers, we've factored in the skill level between players, particularly reflected in the distance covered while running. Since running depletes stamina, and skilled players are adept at conserving energy during play to be prepared for subsequent points in the match, an increase in running distance for a player suggests greater pressure applied by the opponent. Hence, the ratio $S2/S1$ accounts for the relative movement distance, providing insight into the tactical exchanges between players. This model dimension introduces a physical component into the skill assessment, acknowledging that endurance and strategic movement are indicative of a player's competitive edge. Additionally, the expression $\frac{f}{1+e^{d \times ST}}$ accounts for player 1's running distance. We recognize that when a player is evenly matched with an opponent and both exhibit high skill levels, such a deadlock more accurately reflects certain qualities and abilities of the players. Also, considering that athletes, especially outstanding ones, is good at making full use of their energy, our model use expression $S2/S1$ to penalize those who use extra energy. Therefore, we provide a reward in our model for this situation. **The expression $S2/S1$ and $\frac{f}{1+e^{d \times ST}}$ mutually constrain each other**, linking the levels of both players within our model.

At the moment, it is time to optimize the parameters. Although the model seems to show the positive relationship between the independent variables and the dependent ones, there may be local optimal solutions in the search space, which hinders us from reaching global optimal solution. In case of this issue, we employed the **Late Acceptance Hill-Climbing (LAHC) heuristic strategy**. LAHC is a stochastic local search method that explores the solution space by delaying the acceptance of poorer solutions during the search process, thereby **possibly reaching an optimized solution** and effectively shortening the running time of the Forest-RC model.

The Late Acceptance Hill Climbing (LAHC) method integrates aspects of both hill climbing and greedy searching. It conducts comparisons between a potential solution and the solutions from previous iterations, merging the principles of late acceptance with those of greediness. This alternation from greedy to non-greedy approaches or just the opposite facilitates the identification of a superior solution. The fundamental equation is presented below:

$$C_i < C_{i-L_h} \text{ or } C_i \leq C_{i-1} \quad (7)$$

In formula 7, C_i^* represents the candidate cost, C_{i-1} is the current cost, and C_{i-L_h} is the cost from L_h iterations ago, with $f(i \% L_h)$ where $\%$ denotes the modulo operator [5].

The algorithm 1 employs non-greedy rules for searching during the early stages and switches to

greedy rules in the final stages.

Algorithm 1 Late Acceptance Hill-Climbing Algorithm

```

1: Produce an initial solution  $s$ 
2: Calculate initial cost function  $C(s)$ 
3: Specify  $L_h$ 
4: for all  $k = 0$  to  $L_h - 1$  do  $f_k := C(s)$ 
5: end for
6:  $I = 0; I_{idle} = 0$ 
7: while  $I \leq 1000000$  and  $I_{idle} \leq 0.02$  do
8:   Construct a candidate solution  $s^*$ 
9:   Calculate a candidate cost function  $C(s^*)$ 
10:  if  $C(s^*) \geq C(s)$  then
11:    Increment the idle iterations number  $I_{idle} = I_{idle} + 1$ 
12:  else
13:    Reset the idle iterations number  $I_{idle} = 0$ 
14:  end if
15:  Calculate the virtual beginning  $v = I \bmod L_h$ 
16:  if  $C(s^*) < f_v$  or  $C(s^*) \leq C(s)$  then
17:    Accept the candidate:  $s = s^*$ 
18:    Update the fitness array  $s = s^*$ 
19:  else
20:    Reject the candidate:  $s = s$ 
21:  end if
22:  if  $C(s) < f_v$  then
23:    Update the fitness array  $f_v = C(s)$ 
24:  end if
25:   $I = I + 1$ 
26: end while
  
```

After optimization and fitting process, we have developed a Weight Model:

$$S(R_A, R_{BPW}, R_{NPS}, S1, S2) = \left(34.375R_A + 61.875R_{BPW} + 9.375R_{NPS} + \frac{78.75}{1 + e^{0.10625 \times S1}} \right) \frac{S2}{S1} \quad (8)$$

The MSE score of formula 8 is 59.6439. By integrating this with the Success Rate Model from the initial question, we arrived at the scoring evaluation model:

$$G = W \times S$$

3.3 Visualization and Outcome

In our detailed examination of the "2023-wimbledon-1301" match, we scrutinized the showdown between Carlos Alcaraz and Nicolas Jarry. The game concluded with a triumphant win for Carlos Alcaraz. A closer look at the subsequent analysis figure 4 , which charts the duo's performance across

eight key metrics, indicates a clear dominance by Alcaraz over Jarry. Not only did Alcaraz secure a greater number of Break Points Won, but he also surpassed in Net Play Success, First Serve Success Rate and Second Serve Success Rate, and executed a higher number of Winners. A critical aspect of his win was his ability to maintain control, committing fewer Unforced Errors compared to his counterpart. Such statistical superiority, as displayed in the performance chart, leads to a well-supported conclusion that Alcaraz's play was substantially more effective than Jarry's during this match. This comprehensive statistical breakdown not only illustrates the specific areas where Alcaraz excelled but also underscores the importance of consistency and precision in winning matches at such high-level competitions.

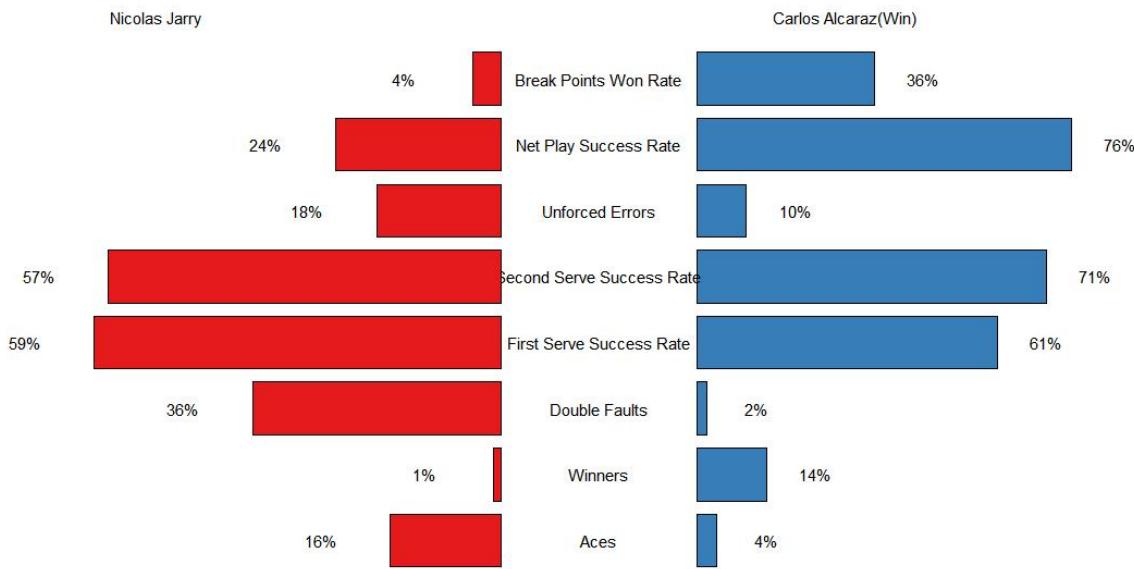


Figure 4: performance in the "2023-wimbledon-1301"

Furthermore, using the scoring evaluation model, we assessed the performances of both players and obtained the following Scoring Chart and Win Probability Graph. Referencing Figure 5, a thorough analysis reveals that Carlos Alcaraz consistently outshone Nicolas Jarry during their encounter. Alcaraz's superiority is quantitatively evidenced by his higher average score of 18.85, a slight yet decisive edge over Jarry's 18.33. This numerical advantage reflects not just isolated moments of brilliance but a sustained level of high-caliber play throughout the duration of the match. Leveraging the insights provided by our scoring evaluation model, we were able to predict Alcaraz's triumph based on his performance metrics. The model's predictions were not only reflective of in-match dynamics but also corroborated by the match's final result, thereby validating the efficacy of our analytical approach. The model's predictive accuracy not only illustrates its utility in forecasting outcomes based on performance data but also highlights the close competition between the two athletes, underscoring the competitive nature of the sport where even marginal improvements can be the key to victory.

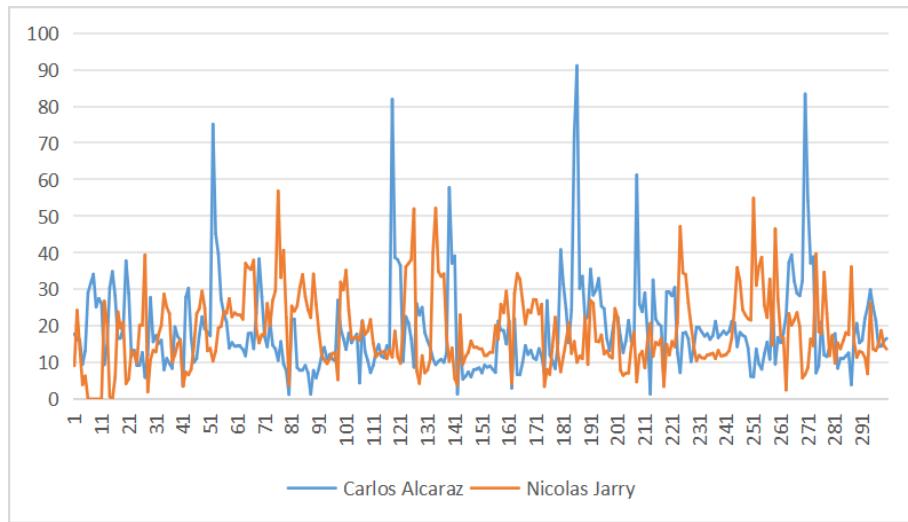


Figure 5: Scoring Chart

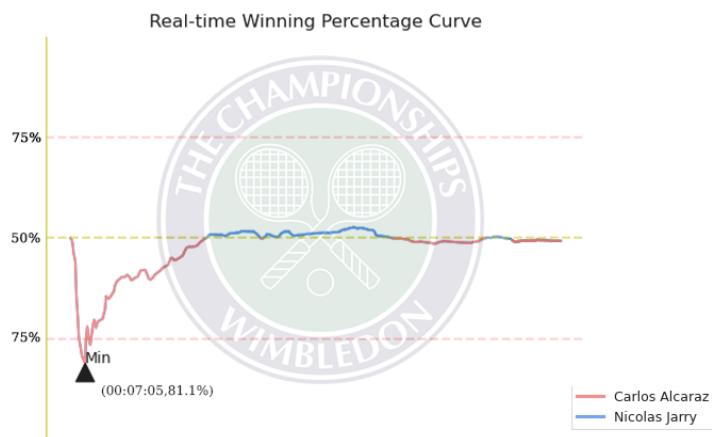


Figure 6: Win Percentage Curve

Upon calculating the win probability, we generated a real-time win probability chart encompassing the entirety of the match. Our model resets the cumulative variables so that the performance took place in the last game would not influence the current game. The Figure 6 clearly demonstrated that Carlos Alcaraz held a distinct advantage in the likelihood of winning during the initial phases of the game, showcasing a performance that was markedly superior to that of his competitor. Such early dominance is often indicative of a player's strong start and ability to assert control from the outset. As the match progressed into its latter stages, however, the win probability between Alcaraz and Jarry leveled off, suggesting a tightening of the contest. This convergence of win probabilities may reflect a shift in momentum or a strategic adjustment from Jarry, who might have increased his level of play to match Alcaraz's earlier performance. Despite this late-game equilibrium, Alcaraz's strong early lead may have been a decisive factor in his overall victory. The win probability chart serves not only as a narrative of the match's flow but also as a testament to the dynamic nature of tennis, where early leads can set the tone for the match, yet resilience can bring a player back into contention.

4 Task 2 & Task3

4.1 Task 2

Task 2 aims to explore the correlation between scoring and momentum in tennis, necessitating a methodical approach to quantify momentum. This endeavor involves analyzing momentum from various aspects of the game, such as consecutive scoring[1], breaking serve, and first serve success rate, to provide a comprehensive view of its impact on game dynamics.

Quantifying Momentum through Consecutive Scoring: The momentum generated from consecutive scoring serves as a critical indicator of a player's dominance on the court at any given moment. Assigning a value based on the number of consecutive scores:+1 for three consecutive scores, +2 for four, and +3 for five—offers a quantifiable measure of positive momentum. Conversely, consecutive losses of points are similarly quantified but with negative values(-1 for three consecutive losses, -2 for four, and -3 for five) to represent a loss in momentum. This approach facilitates a dynamic assessment of momentum shifts throughout a match, highlighting how streaks of scoring or losing points can significantly influence the psychological and strategic state of play.

Quantifying Momentum through Consecutive Breaking Serve: Breaking an opponent's serve marks a pivotal moment in tennis, capable of significantly shifting momentum. A successful break of serve is marked as +1, reflecting an immediate gain in momentum, while two consecutive breaks of serve are recorded as +2. Conversely, being broken by an opponent is a considerable setback, denoted as -1 for a single break and-2 for consecutive breaks. This metric underscores the strategic battle of serves and returns, illustrating how breaking serve affects not only the score but also the mental and strategic dynamics of the match. This analysis covers the entire match, capturing the ebb and flow of momentum with each serve and return game.

Quantifying Momentum through First Serve Success Rate: The first serve success rate is a fundamental aspect of tennis that directly influences momentum. A high success rate on the first serve is a strong indicator of a player's serving efficiency and their ability to maintain control over their service games. A success rate of 90% or above is considered excellent, earning a score of +3, while rates equal to or greater than 75% but less than 90% are scored as +2. A success rate above 60% receives a +1, reflecting a competent serving performance that contributes to maintaining momentum. This metric emphasizes the importance of the first serve in building and sustaining momentum, highlighting its role in setting the tone for each point and game throughout the match.

By systematically quantifying momentum through these three key aspects, Task 2 seeks to provide an in-depth analysis of how momentum influences scoring and the overall outcome of tennis matches. This approach not only captures the immediate impact of scoring patterns, serve breaks, and serving efficiency but also offers insights into the strategic and psychological foundations of momentum in tennis. Understanding these dynamics can significantly inform coaching strategies, player development, and match preparation, offering a nuanced view of the complex interplay between scoring and momentum in the competitive environment of tennis.

4.1.1 Differential Analysis

Considering that the score and the three variables related to momentum mentioned above can all be considered as categorical variables, a chi-square test is utilized, with results shown in Table 3.

Table 3: Chi-square test results

momentum	χ^2	DF	P-value
Consecutive Scoring	82.316	24	2.59×10^{-8}
Consecutive Breaking Serve	20.226	12	0.0629
First Serve Success Rate	11.172	8	0.1922

For the relationship between consecutive scoring and scoring, the chi-square value is 82.316 with 24 degrees of freedom and a P-value of 2.59×10^{-8} , indicating a highly significant correlation between consecutive scoring and match scoring.

Regarding the relationship between consecutive breaks and scoring, the chi-square value is 20.226 with 12 degrees of freedom and a P-value of 0.0629. This suggests that the association between consecutive breaks and scoring is not as significant as that of consecutive scoring, but it still has statistical significance.

Lastly, for the relationship between the first serve success rate and scoring, the chi-square value is 11.172 with 8 degrees of freedom and a P-value of 0.1922, indicating that the correlation between the first serve success rate and scoring is not significant. This may be because the first serve success rate is influenced by a variety of factors and has a lower direct correlation with match scoring.

From the above analysis, we can see that in tennis matches, **consecutive scoring** is a very crucial momentum indicator that affects match scoring and the final outcome. While consecutive breaks are also related to scoring, their impact is less than that of consecutive scoring. As for the first serve success rate, although intuitively we consider it an important factor affecting the game, in this analysis, its direct correlation with scoring is not significant. This may suggest that in training and match strategy, there should be more focus on how to improve the ability to score consecutively and break effectively, while the focus on improving the first serve success rate may need to be considered in conjunction with other factors. Thus, we may put more emphasis on it in the following task.

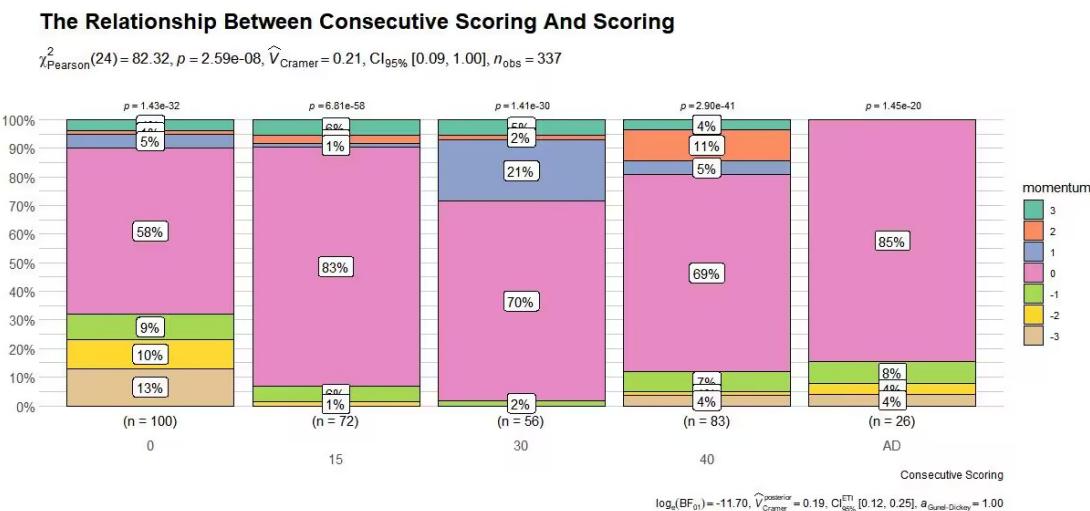


Figure 7: Analysis of Consecutive Scoring Impact on Game Outcomes

Figure 7 presents a bar chart illustrating the relationship between consecutive scoring and overall scoring in a game. The x-axis indicates the consecutive scoring streaks, while the y-axis shows the percentage of total scores. Each bar is color-coded to represent different levels of momentum, with the legend on the right explaining the corresponding values. Notable statistical values such as the Pearson chi-square, p-value, Cramér's V, and the number of observations (n_{obs}) provide evidence of the statistical significance and strength of the relationship. The p-values at the top of each bar indicate a very low probability that the observed association is due to chance, suggesting a strong correlation between consecutive scoring and winning.

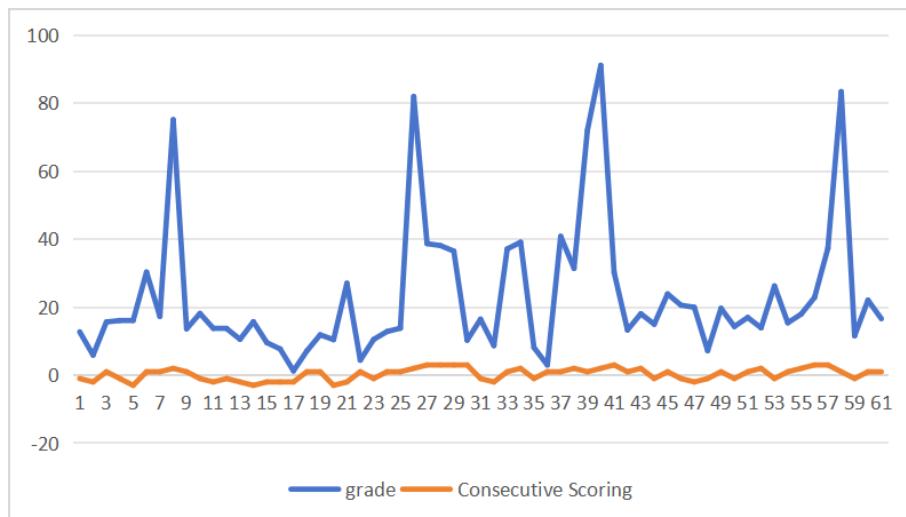


Figure 8: Comparative Trend Analysis of Grade and Consecutive Scoring Fluctuations

Meanwhile, we used our scoring evaluation model to obtain Player1's score in this game and plotted a line graph depicting his score against the quantified momentum. From the fluctuations in the Figure 8, we infer that the score's variability closely mirrors the momentum's fluctuations. To validate our hypothesis, we considered conducting an ANOVA variance analysis to determine the relationship between these two variables. First, we performed a Shapiro-Wilk normality test on the scores, yielding a $p\text{-value} < 2.2\text{e-}16$, indicating the scores' conformity to normal distribution. Finally, conducting the ANOVA variance analysis resulted in a $p\text{-value} = 1.8\text{e-}09$, signifying a significant correlation between the defined momentum and the scores obtained from the evaluation system. This reinforces the pivotal role of momentum in the game.

4.2 Task 3

4.2.1 Match Fluctuation Factors Analysis: FP-Growth tree

To identify the most relevant indicators, we will establish a frequent pattern mining model. Viewing each point as a unit, we will treat the recorded indicators of both players during each point as an item, analogous to mining shopping data. In the study of frequent pattern mining, well-designed association rule algorithms are crucial. Traditional measures often include support and confidence, with the formula for support illustrated as follows:

$$textsup(A) = \frac{\{\text{Events with } A\}}{\{\text{All Events}\}} \quad (9)$$

However, these measures may not reflect the true relationship in certain specific circumstances. We will introduce an improved new algorithm later on and elucidate the shortcomings of using support as a measure.

Given the sparsity of the indicator data, we have employed an enhanced **FP-Growth tree** algorithm, which addresses the instability of the original algorithm and significantly speeds up the computation process.

The steps of the Frequent Pattern Mining Model are as follows:

Step 1: Establish the frequent pattern tree (FP-tree): Scan the entire dataset to calculate the support count, sup(item), for each item. Items that meet or exceed a predefined minimum support threshold are retained. Sort these items in descending order of support count and construct the FP-tree by inserting transaction sequences according to this order, while incrementing the count of existing nodes when the same item sequence recurs.

Step 2: Determine the conditional pattern base for each item: For items in the FP-tree, trace the node links to find all the prefix paths leading up to each item. Each path is a sequence of nodes from the root of the FP-tree to the node just before the target item node. These prefix paths, collectively known as the conditional pattern base, represent all the transactions in which the target item appears.

Step 3: Obtain frequent itemsets from the conditional pattern base: Use the conditional pattern bases to construct conditional FP-trees for each item. Then, recursively repeat the process for each conditional FP-tree to find subsets of frequent itemsets. During each recursion, combine the item for which the conditional FP-tree was made with the frequent itemsets found in the tree to generate new itemsets.

Step 4: Organize the data into a table: Compile the frequent itemsets into a structured format, such as a table, listing the itemsets along with their corresponding support counts. This table serves as a comprehensive summary of all frequent itemsets discovered in the dataset, which can be used for further analysis or to generate association rules.

Based on the data provided in the table, we initially utilized the measure of support to quantify the prevalence of data indicators.

Table 4: Support for Various Tennis Match Indicators

Items	Support
Break Points Won	0.168518
First Serve Success Rate	0.186120
S1	0.502813
Second Serve Success Rate	0.104747
Unforced Errors	0.134757
Winners	0.167075
No Fluctuation	0.765041
Fluctuation	0.234959
Aces	0.045015
Double Faults	0.017963
Net Play Success Rate	0.002597

The table 4 indicates that the data itself exhibits **skewness** between the no-fluctuation and fluctuation conditions, exacerbated by data sparsity, resulting in significantly higher support for almost all indicators under the no-fluctuation condition compared to the fluctuation condition.

Therefore, we introduce the Kluczynski measure (Kluc) and the Imbalance Ratio (IR) to better characterize the association rules between cases and states. The formulas for the Kluc measure and IR measure are as follows:

$$Kluc(A, B) = \frac{P(A, B)}{2} \left(\frac{1}{P(A)} + \frac{1}{P(B)} \right) \quad (10)$$

$$IR(A, B) = \frac{|textsup(A) - textsup(B)|}{textsup(A) + textsup(B) - textsup(A \cup B)} \quad (11)$$

By employing both the Kluc and IR measures in tandem, we can adequately depict the degree of association between the two. The Kluc measure 10 is particularly useful in situations where the distribution of cases is uneven, as it considers the relative frequency of cases across different states. The IR measure 11, on the other hand, quantifies the degree of imbalance within the data, reflecting the proportion of instances in one state relative to others[4].

Together, these measures provide a more nuanced understanding of the data, allowing for a balanced analysis that accounts for its disparities in the dataset. This approach enables us to uncover more accurate and representative patterns within the data, thus facilitating more informed decision-making and strategic planning based on a comprehensive view of the association rules.

Table 5: Support, Kluczynski Measure (Kluc), and Imbalance Ratio (IR) for Tennis Indicators

pair1	pair2	support	Kluc	IR
No Fluctuation	Aces	0.028351	0.333433	0.921096
Fluctuation	Aces	0.016664	0.220558	0.72137
No Fluctuation	Break Points Won	0.079065	0.286263	0.6981
Fluctuation	Break Points Won	0.089453	0.45577	0.211578
Fluctuation	Double Faults	0.017602	0.501464	0.97606
No Fluctuation	Double Faults	0.000361	0.010808	0.859183
Fluctuation	First Serve Success Rate	0.161304	0.538755	0.732944
No Fluctuation	First Serve Success Rate	0.024816	0.119476	0.123248
Fluctuation	Net Play Success Rate	0.001082	0.209041	0.994636
No Fluctuation	Net Play Success Rate	0.001515	0.29489	0.984413
Fluctuation	S1	0.387246	0.638167	0.29778
No Fluctuation	S1	0.115568	0.360853	0.430493
Fluctuation	Second Serve Success Rate	0.096667	0.52461	0.854064
No Fluctuation	Second Serve Success Rate	0.00808	0.055761	0.392647
Fluctuation	Unforced Errors	0.134613	0.587442	0.823701
Fluctuation	Unforced Errors	0.000144	0.000842	0.27113
No Fluctuation	Winners	0.091617	0.334057	0.711441
Fluctuation	Winners	0.075458	0.386398	0.207864

Due to the sparsity of the data, we set a low support threshold to detect subtle association rules.

The improved algorithm compensates for the increased computation time that arises from selecting a small support threshold.

Conclusion:

As the Kluczynski measure (Kluc) approaches 1, there is a stronger positive association between two items; conversely, as Kluc approaches 0, the association is stronger in the negative direction. Additionally, the Imbalance Ratio (IR), as defined by the formula, indicates that the closer IR is to 1, the stronger the data imbalance; conversely, the closer IR is to 0, the weaker the data imbalance. Table 5 showcases the maximum support, Kluc, and IR measures for various items under fluctuating conditions. The ideal case for two items with strong association would have a support of 1, Kluc of 1, and IR of 0.

When employing the combination of Kluc and IR measures, indicators associated with fluctuation, such as Break Points Won, Net Play Success Rate, and Winners, become apparent, revealing that they are difficult to capture solely through support. However, Kluc is unable to significantly affect indicators with a large support disparity, such as Double Faults and Unforced Errors. Ultimately, based on the values of Kluc and IR, we have ranked and selected the following three indicators shown in Table 6.

Table 6: Ranked Indicators Based on Kluc and IR Measures

Indicator	Support	Kluc	IR
Break Points Won	0.089453	0.45577	0.211578
Winners	0.075458	0.386398	0.207864
Net Play Success Rate	0.001515	0.29489	0.984413

Table 6 indicates that **Break Points Won** and **Winners** perform well, while **Net Play Success Rate**, due to excessive data skew, requires further verification to ascertain its impact on fluctuation.

4.2.2 Dual-objective Optimization and Prediction Model

Next, we construct a predictive model based on mostly the three variables as well as other factors to quantify fluctuation, we categorize it into positive fluctuation, negative fluctuation, and no fluctuation based on our analysis of momentum variables from question two. For momentum we defined in Task 2 is a categorical variable, we may lose a lot of information. Therefore, we try to modify SEMOLA Model by adjusting weights of the three variables and ultimately get two new variables so that the variable momentum can be quantified continuously. Then, logistic regression is conducted accordingly.

When our side has a surge of momentum, we desire to maximum the difference $\Delta\Delta G$ between two players' last and current score difference($\Delta G'$ and ΔG), for a surge of momentum is consistent with the surge of the score obtained by SEMONA Model.

Conversely, when the opponent has the momentum, we wish to minimum the difference $\Delta\Delta G$ between two players' last and current score difference($\Delta G'$ and ΔG)

Based on this, we construct a dual-objective optimization model:

- $\text{Maximize } \Delta\Delta G == (G1' - G2') - (G1 - G2)$ when our side has a strong momentum.

- Minimize $\Delta\Delta G = (G1' - G2') - (G1 - G2)$ when the opponent has a strong momentum.

This model allows us to strategically account for the dynamic shifts in momentum during the match.

We define $m_i^j = 0.2190489R_{NPS} + 0.23620803C_{BPW}$, where $i = 1$ represents the opponent's data, $i = 0$ represents our data and j indicates the j -th point. $m_i^{j'}$ denotes the data from the previous point. $n_i^j = 9.375R_{NPS} + 61.875C_{BPW}$ is defined in the same way. So do $p_i^j = \left(34.375R_A + \frac{78.75}{1+e^{d \times 0.10625 \times S^j}}\right) \frac{S_2}{S_1}$, $q_i^j = 0.20991536C_A + 0C_W - 0.0191275C_{DF} + 1.65536857R_{FSS} + 1.05755607R_{SSS} - 0.26796471C_{UE} - 0.93167905$ and $q_i^{j'}$. Specially, $S^j = \frac{S_2^j}{S_1^j}$ and $S'^j = \frac{S_1^j}{S_2^j}$.

Upon substituting the model from the first question, the equation simplifies to:

$$\begin{aligned} \text{Minimize } & [(m_0^{j'} + q_0^{j'})(n_0^{j'} + p_0^{j'})S'^j - (m_0^{j'} + q_0^j + a(x_{10} - x_{10'})) \\ & + b(x_{20} - x_{20'}))(n_0^{j'} + p_0^j + c(x_{10} - x_{10'}) + d(x_{20} - x_{20'}))S^j] - \\ & [(m_1^{j'} + q_1^{j'})(n_1^{j'} + p_1^{j'})S^j - (m_1^{j'} + q_1^j + a(x_{21} - x_{21'})) \\ & + b(x_{31} - x_{31'}))(n_1^{j'} + p_1^j + c(x_{21} - x_{21'}) + d(x_{31} - x_{31'}))S^j] \end{aligned}$$

Maximize by analogy, replacing x_1, x_2 with the opponent's net play success rate and break points won.

Here a, b, c, d are parameters to be optimized, with x_1, x_2 representing the values of net play success rate and break points won rate (noting the subscripts for consistency).

We employed the **NSGA-II** algorithm for optimizing these four parameters. NSGA-II uses an elitist strategy, generating a population with diversity and sharing strategies, making it an excellent genetic algorithm. The result is shown in Table 7 with Spacing-score=0 and HV-score=7.6316.

Table 7: Adjusted parameters

Before	Now	Before	Now
$0.2190R_{nps}$	$0.2012R_{nps}$	$9.3750R_{nps}$	$11.0373R_{nps}$
$0.2362C_{bpw}$	$-0.2096C_{bpw}$	$61.8750R_{bpw}$	$33.3364R_{nps}$

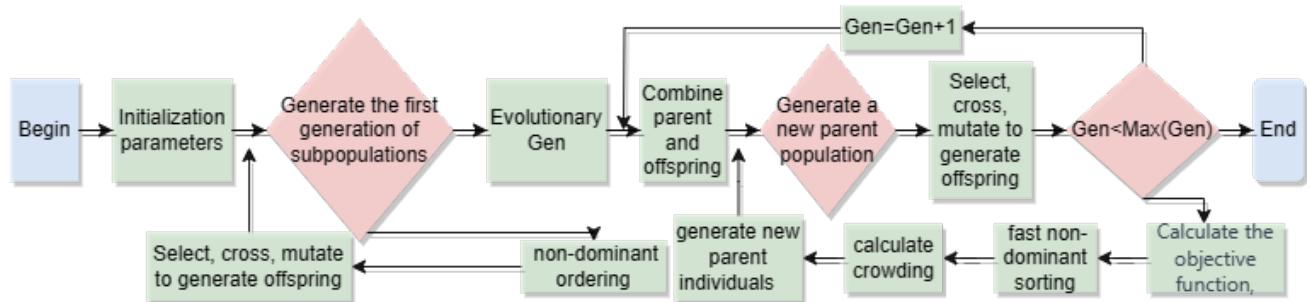


Figure 9: Flow chart

Figure 9 illustrates the flowchart of the genetic algorithm. We consider two players' G0 and G1 depicting the momentum continuously calculated by the modified SEMOLA Model, as independent

variables. The fluctuation defined above works as an implicit variable. The logistic regression is conducted and fits well on the test set. The formula is given in the following.

$$f(G0, G1) = 1 / (1 + \exp(-0.1832 \cdot G0 + 0.1104 \cdot G1 - 0.5305))$$

4.2.3 Strategic Advice Based on Historical Momentum Differential in Matchups

Considering the fluctuations in momentum on the court, we primarily categorize them into three scenarios: experiencing positive momentum, experiencing negative momentum, and maintaining a stable state. Therefore, we will take the three states of the other player into account.

- When facing an opponent experiencing positive momentum, as indicated by our previous analysis, players are advised to revise their strategies and contemplate their tactics attentively. Maintaining focus is crucial to prevent distractions that could impact subsequent gameplay and to minimize errors. By seeking opportunities to execute successful net plays, break points, and winning shots, players can stay composed under pressure. Enhancing the success rates of these critical plays could potentially alter the course of the match. Such strategic shifts, coupled with a calm demeanor, could effectively counter the opponent's momentum and tilt the balance in their favor.
- When the opponent is stable, it's advisable to avoid prolonged stalemates, conserve energy, and seize opportunities for scoring breakthroughs.
- When the opponent is in a negative fluctuation, we recommend the player to keep a balanced mindset and adopt an aggressive style to expand their advantage and maintain their form.

5 Task 4: Model Test

5.1 Model Validation

To investigate the Validation of our model, we applied it to an additional match and analyzed its performance in predicting positive and negative momentum swings. The Area Under the Curve (AUC) values were found to be 0.628 for positive momentum and 0.574 for negative momentum. These figures, while indicative of some predictive ability, **suggest room for improvement** compared to the results obtained from the test set. Factors such as strategic changes, physical endurance, and the opponent's performance might be crucial in enhancing the model's accuracy. Incorporating these elements could potentially yield a more robust predictive model, offering a better understanding of the dynamics within a tennis match.

5.2 Model Generalizability

We accessed data from the "2023-usopen-5101" in "2023-usopen-points"[6]. The U.S. Open, held annually as the fourth and final Grand Slam tournament, usually takes place in New York from late

August to early September. The competition encompasses five main categories: men's singles, women's singles, men's doubles, women's doubles, and mixed doubles, along with junior events. Since 1978, the tournament has been hosted at the USTA Billie Jean King National Tennis Center in New York.

In an attempt to verify the versatility of our model, we applied it to a women's tennis match from the U.S. Open. The model's predictive capability yielded an AUC of 0.688 for positive momentum, indicating a good level of accuracy. For negative momentum, the AUC was a comparable 0.666, further affirming the model's reliability. These results demonstrate that our model holds interpretative power and can be extrapolated to other sporting events. The consistent AUC values across different datasets suggest that the model captures key momentum-related dynamics that are not merely sport-specific but could potentially be universal in competitive sports.

6 Sensitive Analysis

To evaluate the robustness of our model, particularly focusing on predicting positive momentum, we conducted a sensitivity analysis. This involved adjusting the threshold values used in the Receiver Operating Characteristic (ROC) analysis and observing the subsequent changes in sensitivity. The goal was to assess how sensitive our model is to variations in threshold settings, a crucial factor in understanding its reliability under different operational conditions.

Table 8: Sensitivity Analysis at Different Thresholds

Threshold	Sensitivity
-0.9970	0.8667
-1.0070	0.8667
-0.9870	0.8000
-0.9920	0.8667
-1.0020	0.8667

The results from the sensitivity analysis, as shown in the table, indicate that alterations in the threshold values lead to relatively minor changes in sensitivity. This consistency in sensitivity across different thresholds suggests that our predictive model maintains its stability and accuracy even when faced with variations in the criteria used to classify positive momentum. Therefore, we can conclude that our model demonstrates a high level of robustness, making it a reliable tool for predicting momentum swings in sports analytics. This stability ensures that our model can be effectively applied in various scenarios without the need for frequent recalibration.

7 Conclusion

7.1 Strengths and Weaknesses

Task 1's methodology, integrating a Weight Model with a Success Rate Model into a Scoring Evaluation Model, offers a comprehensive approach to player performance analysis in tennis, addressing both overt and subtle game aspects. The LAHC-optimized Weight model enhances predictive accuracy

and computational efficiency, effectively tackling data sparsity. This approach provides a holistic view of players' abilities, extending beyond simple win-loss metrics to include strategic and psychological factors influencing matches. However, the complexity of model optimization could limit its broader applicability without further validation and adaptation and repeated variables may damage the robustness of the model.

Task 2 Model focuses on quantifying momentum in tennis using consecutive scoring, break points, and first serve success rates. Its strengths lie in its comprehensive approach to capturing the dynamic nature of momentum through measurable indicators, offering a nuanced understanding of its impact on match outcomes. However, its reliance on chi-square tests for validation, while statistically rigorous, may not fully account for the complexity of in-match conditions and player psychology.

Task 3 Model employs the FP-Growth algorithm to analyze and model match fluctuation factors, focusing on identifying the most relevant indicators for fluctuations. Its innovative use of Kulczynski measure and Imbalance Ratio enhances data association analysis, providing a more detailed exploration of fluctuations than traditional models. However, the complexity of the algorithm and advanced metrics used may limit its accessibility to practitioners not familiar with these methods.

8 Memorandum for The Chief Administrator

References

- [1] Moss, B., & O'Donoghue, Peter. *Momentum in US Open Men's Singles Tennis*. International Journal of Performance Analysis in Sport, 15(3), 884-896, 2015.
- [2] J. Han, M. Kamber, *Data Mining: Concepts and Technique*, Morgan Kaufmann Publishers Inc., 2006. DOI: 10.1002/9781118029145.ch1.
- [3] Wu, J. M., Zhan, J., Chobe, S. (2018). Mining Association Rules for Low-Frequency Itemsets. *PLoS One*, 13(7), 1-24.
- [4] Han, J., Kamber, M. (2006). Data Mining: Concepts and Technique. *Morgan Kaufmann Publishers Inc.* DOI:10.1002/9781118029145.ch1.
- [5] Edmund K. Burke, Yuri Bykov, *The late acceptance Hill-Climbing heuristic*, European Journal of Operational Research, Volume 258, Issue 1, 2017, Pages 70-78, ISSN 0377-2217, <https://doi.org/10.1016/j.ejor.2016.07.012>.
- [6] Jeff Sackmann's Tennis Slam Point-by-Point Data, https://github.com/JeffSackmann/tennis_slam_pointbypoint/blob/master/2023-usopen-points-mixed.csv, last accessed October 10, 2023.



Momentum

To: coaches
From: Group 2406434
Date: February 4th, 2024
Subject: Strategies for Momentum: Data-Driven Insights in Tennis Performance

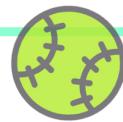
Background

1

In tennis, momentum significantly influences match outcomes, impacting players' confidence and game strategy. It involves psychological and strategic advantages gained through successful actions or scores. Momentum shifts can affect players' mental states, leading to changes in performance and decision-making. Advances in big data and machine learning have allowed for a more systematic study of momentum, enabling researchers to quantify its impact and predict match outcomes more accurately. These insights help athletes and coaches make real-time adjustments and improve training methods. Overall, momentum is crucial for enhancing performance, refining strategies, and understanding the psychological dynamics of tennis.

Advice

2



Findings on the Role of Momentum:

Our analysis reveals that momentum, characterized by a series of successful plays, significantly impacts a player's psychological state and overall match performance. Players with positive momentum are observed to exhibit enhanced focus, increased confidence, and a higher likelihood of maintaining a winning streak.

Key insights include:

- Momentum's Psychological Impact: Players gaining momentum show improved mental resilience and decision-making capabilities.
- Influence on Match Flow: Momentum shifts can alter the course of the game, turning potential losses into victories.
- Strategic Advantage: Effective use of momentum can intimidate opponents and create opportunities for aggressive play.

Recommendations:

- Momentum Building Strategies: Encourage players to capitalize on early successes to build momentum. Techniques such as aggressive serving and capitalizing on opponents' errors can help establish an early lead.
- Mental Resilience Training: Incorporate psychological resilience training to help players maintain momentum and recover from setbacks more effectively.
- Analyzing Opponents' Tendencies: Train players to recognize and exploit moments when an opponent's momentum is waning, using targeted strategies to regain the upper hand.

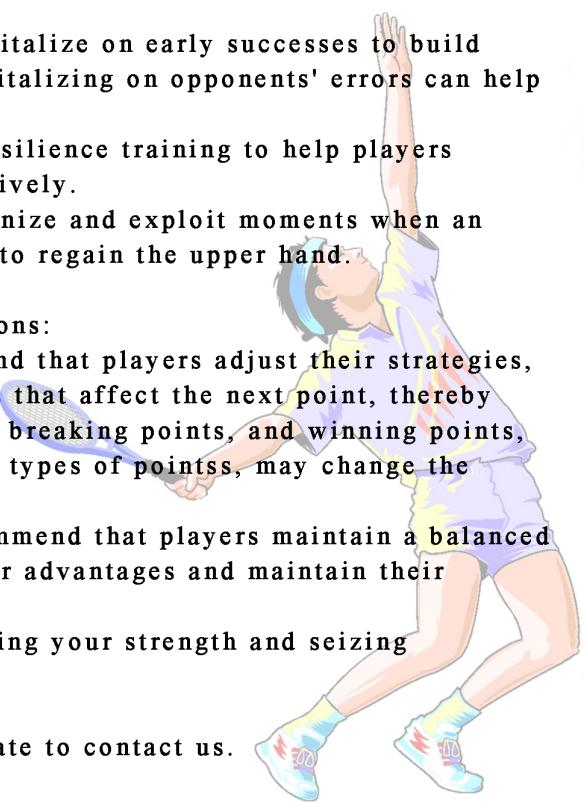
Countermeasures for the Flow of Play:

We divide the flow of play into the following three situations:

- When the opponent is in a positive flow, We recommend that players adjust their strategies, think about tactics, and stay focused to reduce distractions that affect the next point, thereby reducing mistakes. Seeking opportunities to hit net points, breaking points, and winning points, and striving to improve the success rate of the above three types of pointss, may change the situation.
- When the opponent is in a negative fluctuation, We recommend that players maintain a balanced mentality and adopt a positive playing style to expand their advantages and maintain their condition.
- When the play is at a stalemate, We recommend maintaining your strength and seizing opportunities to seek for scoring breakthroughs.

For further discussion or clarification, please do not hesitate to contact us.

[2406434 Group]



Report on Use of AI

1. Microsoft ChatsNow(GPT3.5,2024 Version)

Query1: Can you understand the following sentence in the following? This alternation from greedy to non-greedy approaches or the opposite facilitates the identification of a superior solution.

Answer1: Yes, I can understand this statement. The meaning of this quote is that a shift from greedy to non-greedy methods, or vice versa, helps to find a superior solution. This quote emphasizes that using different methods and strategies when solving a problem can help us find a better solution.

2. OpenAI ChatGPT (Nov 5, 2023 version, ChatGPT-4)

Query1: How to Set Your Own Thresholds for ROC Functions in R?

Answer1: In R, when using functions related to the Receiver Operating Characteristic (ROC) curve to assess the performance of classification models, the setting of thresholds is typically done implicitly, as the ROC curve itself is plotted by calculating the model's True Positive Rate (TPR) and False Positive Rate (FPR) at various thresholds. In other words, the ROC curve offers a comprehensive overview of model performance across all possible thresholds.

However, if you're using the pROC package or a similar package to generate an ROC curve, you might want to calculate certain performance metrics, such as sensitivity or specificity, based on specific thresholds. In this case, even though setting thresholds directly while plotting the ROC curve is not common practice, you can use the ROC object you created to obtain performance metrics based on specific thresholds.

Here is an example of performing ROC analysis and obtaining performance metrics based on a specific threshold using the pROC package:

First, ensure the pROC package is installed and loaded: `install.packages("pROC") library(pROC)` Then, create an ROC object using your model's prediction outcomes and actual results:

Assuming you have the model's prediction probabilities and actual labels `predictions=c(0.1, 0.4, 0.35, 0.8) labels=c(0, 0, 1, 1)`

Create an ROC object `roc_obj = roc(labels, predictions)` At this point, if you wish to obtain performance metrics based on a specific threshold, you can use the `coords` function:

Obtain performance metrics based on a specific threshold, e.g., `threshold = 0.5 coords(roc_obj, "best", ret=c("threshold", "sensitivity", "specificity"))` Note that the "best" parameter attempts to find the optimal threshold, but you can also directly specify a particular threshold, such as:

Specify a threshold of 0.5 `coords(roc_obj, x=0.5, input="threshold", ret=c("threshold", "sensitivity", "specificity"))` Here, `x=0.5` is the specific threshold you've set. The `ret` parameter specifies the performance metrics you wish to retrieve from the function.

In this manner, you can obtain performance metrics related to a specific threshold based on the results of ROC analysis, even though the setting of the threshold is not directly accomplished when plotting the ROC curve.