Problem Chosen	2024 MCM/ICM Summary Sheet	Team Control Number
C		2403615

Riding the Tempo: Unveiling the Momentum of Tennis

Summary

With the aim of conducting an in-depth analysis into the relationship between momentum and the flow of match in tennis, we devise a comprehensive model to quantify the player's momentum and apply various mathematical methods to solve the four tasks in total.

For Task 1 and 2, we establish an evaluating model with copious indicators. Based on the cost-benefit principle, we classify the contributors into two types. It should be noted that **we construct a submodel for Consecutive Scoring** employing vector to meet the properties it should satisfy. After normalizing the data, we calculate the weights of the contributors by **CRITIC**. In particular, we take into account the effect of service by defining it as an additional coefficient for the positive contributors and a deduction for the negative. Thus, the **Momentum Index(MI)** is represented by the linear weighted sum of the processed contributors and visualized to depict the match flow. Then, we study the role of momentum by examing how momentum is related with runs of succeess and swings in play. Through our analysis and multiple graphs, we believe that **swings in play and runs of success are measurably related with momentum**, which overturns the coach's statement.

For the first subquestion of Task 3, we devise a prediction model based on **machine learning** after making clear the definition of swings. We first employ neural network prediction, whose accuracy of predicting swings is **91%**. Considering the aim to identify the factors that are most related with the swings in match and improving the accuracy further, we use **Decision Tree** to forecast the swings in play, whose accuracy reaches **94%**. Derived from the decision tree, **the most related factor is Consecutive Scoring, followed by Winners**. Furthermore, we innovatively develop a model based on a high-dimensional vector space in anology with the working principle of word vector to forecast the general trend of the match. **For the second subquesion**, we first devide the players into two categories according to their characteristic of momentum based on **SVM**, then provide concrete instructions for players when facing different opponents in two categories respectively. A **case study** is conducted about the optimal strategies when facing top players like Djokovic. Comparing the suggested strategies for Alcaraz and the real situation in Wimbledon, we not only **vertify the correctness of our model**, but also **justify the victory of the Spaniard**.

For Task 4, we apply our model to other matches and sports to study its universality. Apart from the **high accuracy(above 98%)** in predicting other matches in Wimbledon, the result showcases a **comparatively high precision(90%) in men's matches with similar competiton formats**, while exhibits a **drop in accuracy considering the difference in court, gender and sport**. Concisely, our model still has reference value for predicting swings in diverse competitions.

Finally, the strengths and weaknesses of the model are discussed, and the sensitivity analysis is conducted to examine the stability of the mode, which shows that our model is robust.

Keywords: Momentum; CRITIC; Machine Learning; Neural Network; Decision Tree; SVM; High-dimensional Vector Space;

Contents

1 Introduction

1.1 Problem Background

Momentum, which refers to the the force that keeps an object moving or keeps an event developing after it has started[1], is generally believed to have an impact on the performance of players in competitive sports. Consider the 2023 Wimbledon Gentlemen's Final, in which the victory of the 20-year-old Spanish rising star Carlos Alcaraz could be partly attributed to "Momentum".

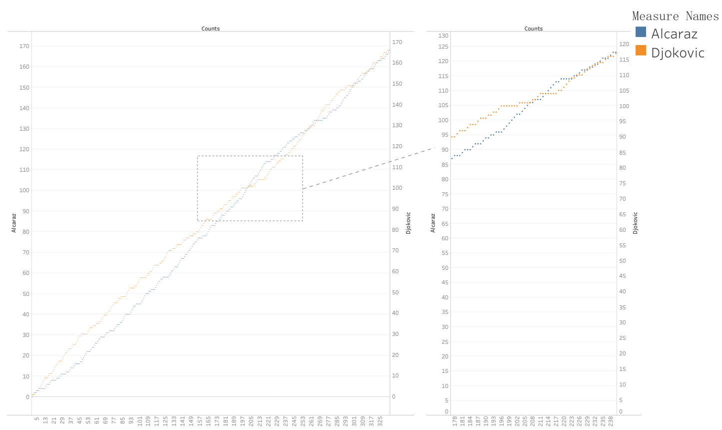


Figure 1: Points Flow of 2023 Wimbledon Gentlemen's Final

As shown in Figure 1, there exhibits both consecutive scoring and persistent downturn in the points flow of either Djokovic or Alcaraz, which can indicate the existence of momentum. Hence, the effect of momentum is worth studying, which justifies our research.

1.2 Restatement of the Task

In response to the question which revolves around momentum, we have broken down the whole problem into four parts.

Part 1 involves constructing a model to record the momentum flow of the match and applying it to several matches to reflect the performances of two players and the edge of the advanced over the other. Indicators should be selected to reflect and quantify momentum in this model. A visualization of the match flow should be provided at the end of this part.

Part 2 involves assessing whether momentum plays a contributing role in the process of a match and if momentum affects when the tables turn and the final outcome by using the model we devised.

Part 3 invloves discovering the indicators which lead the turning of the table. In the meantime, we are expected to establsh a model to predict the swings in the match and elaborate on the primary factors. Besides, we need to devise instructive suggestions for players to prepare for matches against new opponents based on the previous data of momentum.

Part 4 involves applying the model we developed to other matches to test how well we predict the swings in matches. Meanwhile, we may examine include indicators in future models if the result shows poor correctness. Apart from that, the applicability of our models in other matches or other competitive sports ought to be taken into account.

1.3 Literature Review

There are mounting previous research on the effect of momentum on the flow of match. For instance, H Dietl and C Nesseler discovered that players will gain momentum as long as they take control of the match, while they will also have significantly lower chances to win the game when they lose control, which they called "anti-momentum".[2] However, the sign of controlling the match is blurry, and thus the article did not unveil the relationship between the momentum and the performance. A Goyal and JS Simonoff studied the hot hand effect, which refers to the phenomenon that consecutive previous scoring may forecast the current one and also found that this carryover effect is most obvious in the clay court, followed by grass court[3](where the Wimbledon Championships were held). However, the research only focused on the effect of consecutive scoring on the match flow, which is not so comprehensive as the concept of momentum.

In short, albeit the various research, scholars have not drawn a conclusion in the comprehensive effect of momentum on the match flow, which justifies our research.

1.4 Our Work

Our work mainly includes as Figure 2.

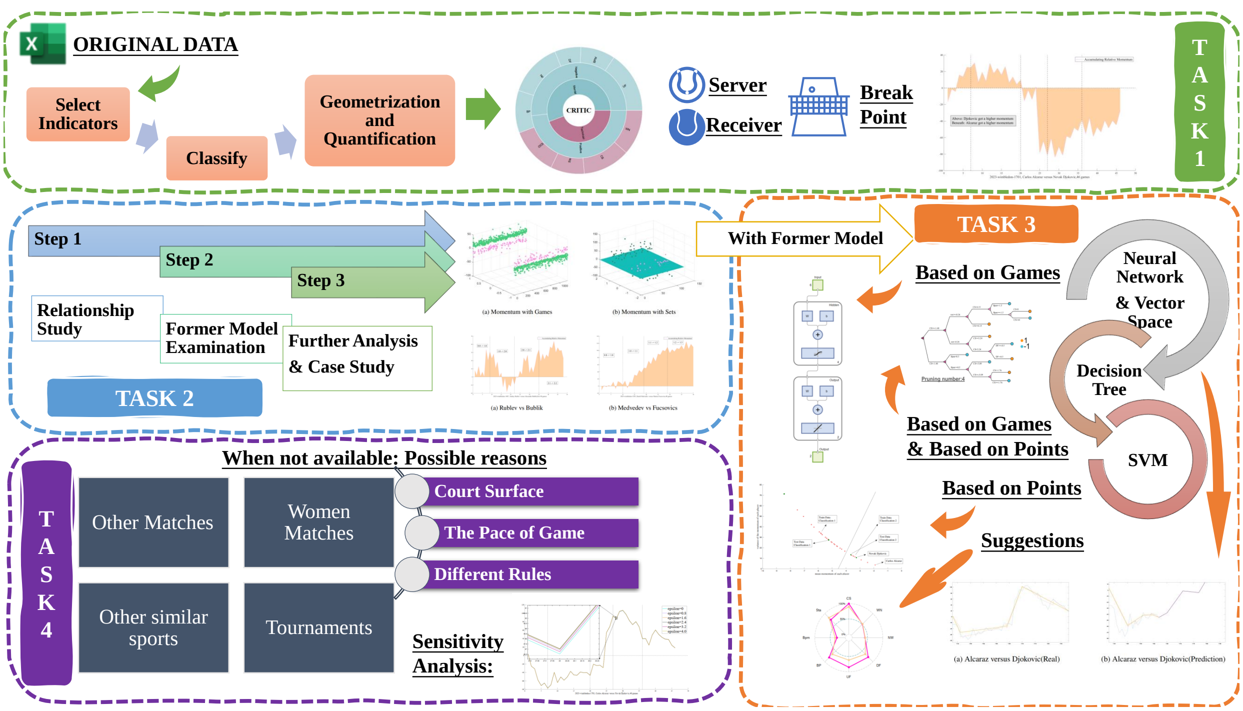


Figure 2: Flow Chart of Our Work

2 Model Preparation

2.1 Assumptions and Explanations

Assumption 1: We will not consider the momentum change caused by the each rally within one point that evolves a multibeat duel.
Explanation: The momentum change is insignificant enough to be dismissed, and detailed information about the performance of each rally is impossible to collect.

Asssumption 2: We assume that the relative momentum can be quantified by the differnce between the momentum index of the two players.
Explanation: Since momentum is introduced to depict the situation of the match and applied to capture the match flow, we use the relative ones instead of the absolute ones in computation and visualization.

Assumption 3: We assume that the off-site factors, enumerating the weather, court, fans and fatigue of players, will not affect the momentum of them.
Explanation: Regarding those objective factors, they are equal to all players and will contribute little to player's performances(Except for some special cases. Consider clay court to Nadal). Meanwhile, fans are expected to be polite and high qualified, justifying our neglecting their affect on player's momentum.

Assumption 4: We assume that the different tactics employed by players(Alcaraz's frequent surfing for example) will not cause the change of momentum of both sides.
Explanation: The tactics of top athletes are carefully planned, and we do not attribute score fluctuations or momentum changes solely to tactical errors.

2.2 Notations

Table 1: Notations Used in this Paper

Symbol	Definition
MI	Momentum Index
CS	Consecutive Scoring
WN	Winners
NP	Net_pt Won
DF	Double Fault
UF	Unforced Fault
BP	Break Point
Bpm	Break Point Missed
Sta	Stamina Consumption
x_i	the i-th Contributor(Unprocessed)
x_i^*	the i-th Contributor(Processed)

*There are some variables that are not listed here and will be discussed in detail in each section.

3 Task 1: Capturing the Flow of Match (Model 1)

3.1 Selecting Contributors for MI

After calculating the possibility of winning a service game(83.92%) by the data provided, we will consider serving games and receiving games as additional factors in the model, rather than just influencing factors.

Given that there are multiple contributors to the momentum of a player, we select several important factors for the **Momentum Index** and classify them according to their type in influencing momentum. It should be noted that based on Cost-Benefit Principle, we devide each primary contributor into positive ones and negative ones, which will be discussed in the third subsection of this section. The concrete indicators selected is shown in Table 2.

Table 2: Contributors

Primary Contributors	Secondary Contributors	Tertiary Contributors
Technical	Positive	Consecutive Scoring(CS) Consecutive Winning of Games Winners(WN) Service Break
	Negative	Opponent's Consecutive Scoring Stamina Consumption(Sta)
Mental	Positive	Net_pt won(NP) Break Point(BP)
	Negative	Double Fault(DF) Unforced Fault(UF) Break Point Missed(Bpm)
Serve	Positive	Service Game
	Negative	Receiving Game

- **Especially, we devised several sub-models to evaluate the complicated contributors. The process are as follows.**

(a) **Consecutive Scoring(CS)**

Step 1: In order to quantify the effect of consecutive scoring as well as opponent's consecutive scoring, we introduce a sub-model based on vector to visualize this indicator. We let one point gained be the vector $(0, 1)$ added from $(\frac{\sqrt{6}}{2}, 0)$, one point lost be the vector $(0, -1)$ added from $(-\frac{\sqrt{6}}{2}, 0)$, and we use the difference vector between a previous vector, which is in the opposite direction to the current vector, and the current vector to represent the change in momentum brought about by this consecutive scoring (or losing points).

Step 2: We then accumulate the difference vectors obtained, and let the Y-axis of the final vector represent the change in the player's momentum.

To make it clear, we present the whole process in a concise flow chart as Figure 3. **It should be noted that we calculate the effect of consecutive winning of games in the same way as consecutive scoring.**

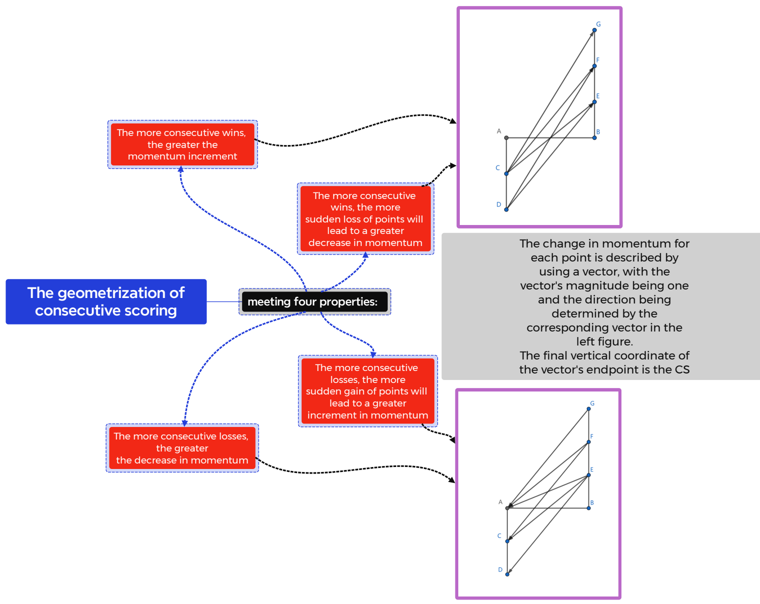


Figure 3: The Geometrization of Consecutive Scoring

According to the aforementioned steps and the data of the Final of 2023 Wimbledon Championships, we visualize the momentum change of Djokovic and Alcaraz in the first set stemming from consecutive scoring in Figure 4. The result is generally align with the fact that Djokovic swept past the Spaniard by 6-1 in the first set, showing the correctness of our analysis.

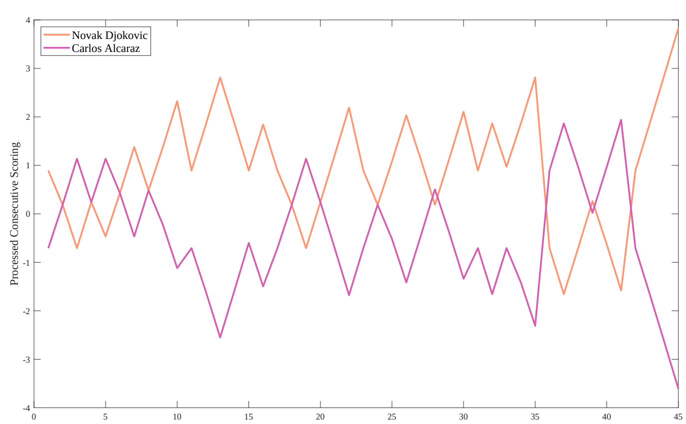


Figure 4: Processed Consecutive Scoring in Set 1, 2023 Wimbledon Gentlemen's Final

(b) **Stamina Consumption(Sta)**

Step 1: In order to quantify the stamina consumption of players, we employ the distace each player runs for the point to express the stamina consumed for this point. Considering that we are calculating the relative values of their momenta, we subtract the running distances of the two players as Equation 1 to reflect the difference in their physical exertion.

$$Distance^{*} = p1\_distance\_run - p2\_distance\_run \tag{1}$$

Step 2: After gaining the relative distance, we normalize the data by using **Arctangent Normalization** as Equation 2. The method is used because it preserves the sign of the data while ensuring that the processed data falls within the range of -1 to 1.

$$Sta = (\tan^{-1} Distance^{*}) * 2/\pi \tag{2}$$

3.2 CRITIC: Calculate the Weight of each Contributors

In order to comprehensively measure the objective weight of different factors based on the comparative strength of evaluation indicators and the conflict between indicators, while taking into account the variability and correlation between indicators, we employ the **Criteria Importance Through Intercriteria Correlation(CRITIC)** to calculate the weights of different factors.

Step 1: Assuming there are n samples to be evaluated, with p evaluation indicators, we form the original indicator data matrix.

$$X = \begin{pmatrix} x_{11} & \dots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \dots & x_{np} \end{pmatrix} \tag{3}$$

Step 2: Then, we use **Min-Max Normalization** to normalize the data.

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{4}$$

Step 3: Next, we calculate the variability(S_j) and conflict(R_j) of the index. S_j here represents the standard deviation, and R_j here means the correlation between conttributor i and j.

$$\begin{cases} \bar{x}_j = \frac{\sum_{i=1}^{n} x_{ij}}{n} \\ S_j = \sqrt{\frac{\sum_{i=1}^{n} (x_{ij} - \bar{x}_j)^2}{n-1}} \end{cases} \tag{5}$$

$$R_j = \sum_{i=1}^{p} (1 - r_{ij}) \tag{6}$$

Step 4: Finally, we can obtain the weights of each indicator according to Equation 7.

$$W_j = \frac{S_j \times R_j}{\sum_{j=1}^{p} (S_j \times R_j)} \tag{7}$$

The result is shown in Figure 5 and Table 3.

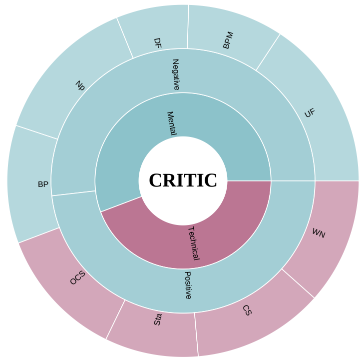


Figure 5: Weights of the Contributors

Table 3: Weights

Contributors	CS	WN	NP	DF	UF	BP	Bpm	Sta
Weights	0.2407	0.1157	0.1375	0.0667	0.1569	0.1083	0.0880	0.0861

3.3 Establishing the Comprehensive Model for Momentum Index

Based on the aforementioned discussion, we may establish the comprehensive model for Momentum Index(MI) which capures the flow of match according to the following steps:

Step 1: In this part, we **take into account the effect of service game and receiving game**. Since the possibility of winning a service game is 83.92% in 2023 Wimbledon Championships, we treat being in service game as a boost for the player, while being in the receiving game the opposite.

Step 2: According to Cost-Effect Principle, we assign a coefficient greater than $1(k_1)$ to represent the bosst on possitive contributors and a coefficient less than $1(k_2)$ to represent the deduction in negative ones in a service game, while in a receiving game, the coefficients are on the contrary.

Step 3: We let p be the ratio of points won by the servers and the receivers, in which p equals 5.2. Then, we get (k_1) and (k_2) under the constraints of these components.

$$\begin{cases} k_1 + k_2 = 2 \\ \frac{k_1}{k_2} = \sqrt{p} \end{cases} \tag{8}$$

Thus, we may optimize each contributors by mutiplying its coefficient under different circumstances to enhance the accuracy of our model.

Step 4: Having obtained the weights of different weights and the final index for contributors, we use Linear Weighted Sum to calculate the MI:

$$MI = \sum_{i=1}^{n} w_i x_i^* \tag{9}$$

3.4 Model Application: Alcaraz Versus Djokovic

To examine our model, we apply our model to 2023 Wimbledon Gentlemen's final. We readily discover that our model for Momentum precisely depict the match flow, that is, the difference in points of the Spaniard and the Czech. As shown in Figure 6, the Momentum Index matches well with the point difference between the two players, effectively reflecting the situation on the court, that is, the performance of the players at any given point.

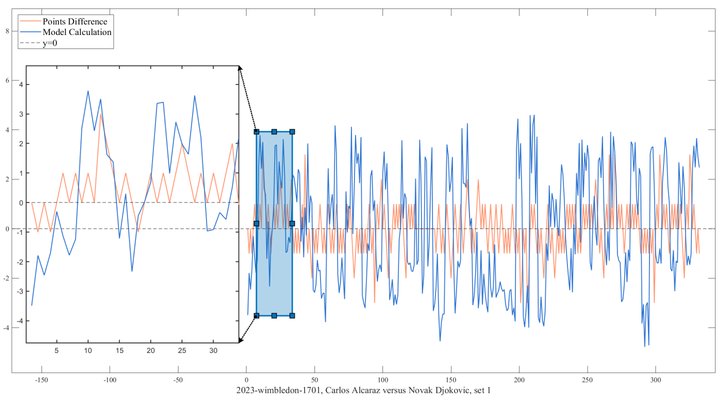


Figure 6: MI of each Point and the Actual Difference in Points

Analysis

As depicted in Figure 6, in the first set, Djokovic had a significant lead in momentum, which corresponds to his actual match performance where he swept his opponent 6-1. In the second set, the battle went to a tie-break, with the situation becoming tense, just as the chart shows frequent exchanges of momentum between the two players on the court. After experiencing a reversal in the second set, the Spaniard gained the lead in momentum and easily defeated Djokovic in the third set, consistent with the negative momentum of Djokovic shown in the middle section of the chart. Regarding the last two sets, both sides exchanged leads, with neither side achieving a decisive lead in any set. However, the momentum Alcaraz built up from being behind to reversing the situation in the first three sets was reflected in the final outcome, leading to his ultimate victory.

To depict the match flow on the court more macroscopic, we accumulate the momentum of Djokovic in each game to embody the overall situation. We visualize the result(Djokovic's momentum compared to Alcaraz) in Figure 7, which matches very closely with the real match flow.

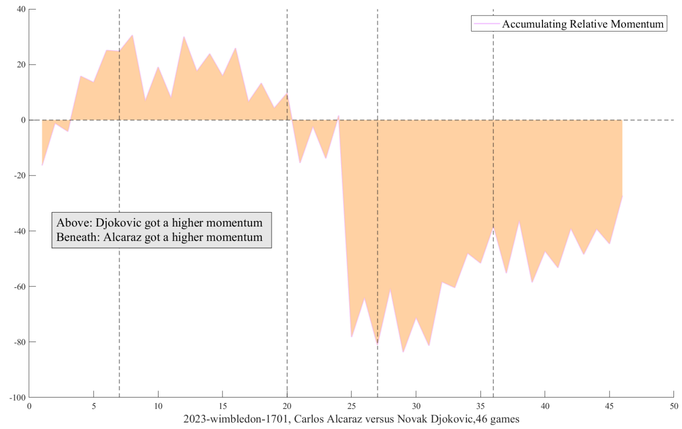
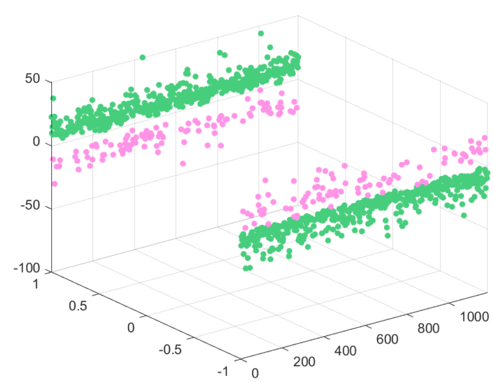


Figure 7: Accumulated Momentum

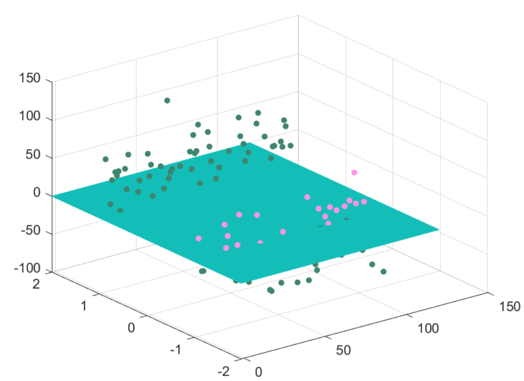
4 Task 2: Momentum's Role in Matches

4.1 Momentum with Runs of Success

To depict the relationship between runs of success and momentum, we look at the relationship between momentum and the victory of games and sets in 2023 Wimbledon Champioonships. We visualize the result in Figure 8.



(a) Momentum with Games



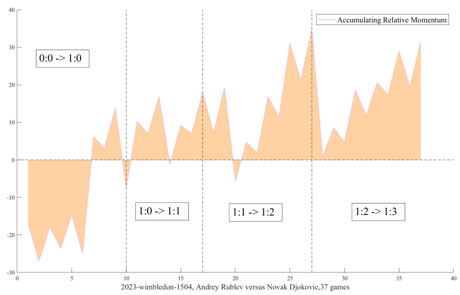
(b) Momentum with Sets

Figure 8: Momentum with the Victory of Games and Sets

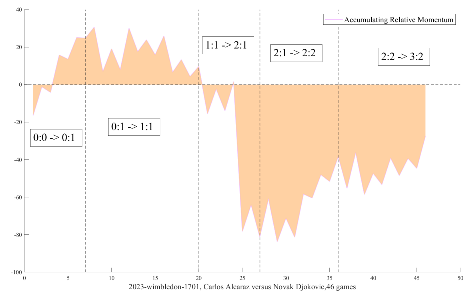
As shown in Figure 8, the green dots represents the games and sets that are won with a higher momentum or lost with occupying a lower momentum, which embody the cases where momentum is cordinated with the runs of success, while the pink ones are not. In fact, we discovered through computation that **983** out of **1188** games**(83%)** and **96** out of **117** sets**(82%)** are won(or lose) with a cordinated momentuum, which shows that the runs of success of a player is not random. Instead, we assert that **runs of success are largely related with momentum**.

4.2 Momentum with Swings in Play

To depict the relationship between momentum with swings in a match, we select several games with obvious swings and compares the moment swings happen with the change of momentum. The result is shown in Figure 9 and Figure 10.



(a) Rublev vs Djokovic

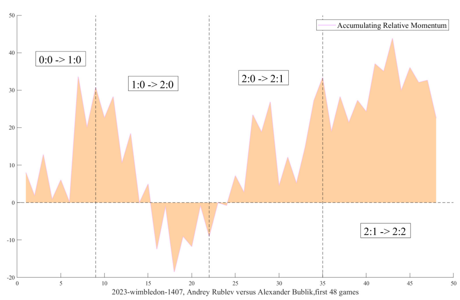


(b) Alcaraz vs Djokovic

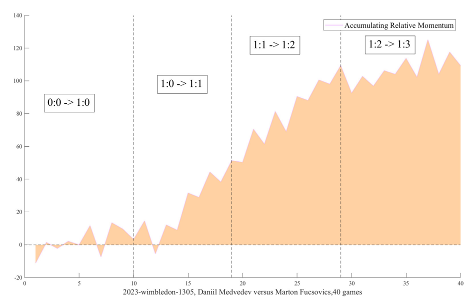
Figure 9: Momentum with Swings in Play(I)

From Figure 9, we may easily discover that **each turning point is accompanied by a surge in momentum.** When trailing, the relative indicator of player momentum tends to be negative, which aligns with the actual situation. After experiencing a surge in momentum, players often see a significant improvement in performance, typically manifested by winning multiple games in a row and eventually catching up or even surpassing their opponents in the overall score.

Further Analysis



(a) Rublev vs Bublik



(b) Medvedev vs Fucsovics

Figure 10: Momentum with Swings in Play(II)

It is worth noting that although minor increases in momentum may not often result in a tangible change in the on-court score, when accumulated to a larger extent, it can lead to an almost irrecoverable transformation. This is usually demonstrated by players winning consecutive sets from a trailing position until they secure victory.

Furthermore, we also observed some special cases where there was no apparent decline in momentum when trailing. Based on our model indicators, we speculate that this might be because players intentionally conserve their energy in the initial few games or sets, or due to the intense nature of the early stages where players often narrowly lose by a few points, thus not displaying a significant drop in morale. These observations align with the actual situations as well.

To sum up, we draw the conclusion through our model that **swings in play and runs of success are measurably related with momentum,** and momentum does play an important role in the match. As a result, the statement of the coach is incorrect.

5 Task 3: Forecasting the Turning Points

5.1 Swings-Prediction Model Based on Machine Learning(Model 2)

To predict the swings in a certain match, we must first determine the indicator for a swing. Based on Figure 6 to 9, we let the consecutive two games won by the trailing side be the sign of a swing. Given the fluctuating momentum and significant score volatility in tennis matches, using models like Grey Forecast Model or ARIMA model for prediction may not yield high accuracy.

Therefore, we employ **Neural Network Prediction** from machine learning to conduct the prediction. The flow chart of the process of the neaural network is shown as Figure 11.

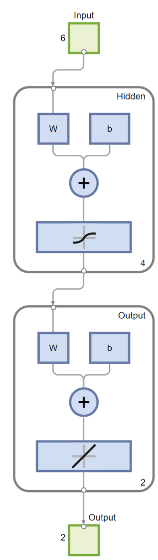
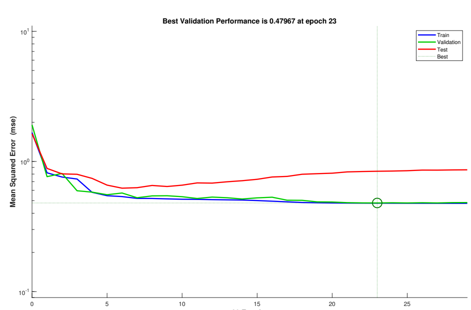


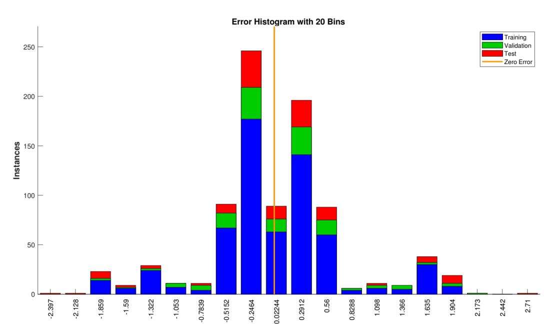
Figure 11: Neural Network Prediction

We choose the win-loss result and the relative momentum of the previous three games as the input of the neural network, and let the result of the following two games as the outcome. It should be mentioned that we calculate player one's victory as positive one, while the victory of player two is calculated as minus one.

The error histogram and the validation performance of the training result are shown in Figure 12. Through computation, we obtain the accuracy of predicting



(a) Validation Performance



(b) Error Histogram

Figure 12: Evaluation of the Neural Network

the right winner of the following two games:

$$p_1 = 90.7\%;\ p_2 = 90.7\% \tag{10}$$

However, there are also some problems with neural network predictions.

- The accuracy of the prediction is not particularly high, which indicates that there is still room for improvement in our model.
- Neural network prediction is not capable of identifying the most related factors in our model.

Therefore, we employ **Decision Tree**, another kind of machine learning, to optimize the prediction while also identify the most related factors for the swings. In this situation, we obtain the structure of decision tree based on all matches given of 2023 Wimbledon Championships as Figure 13.

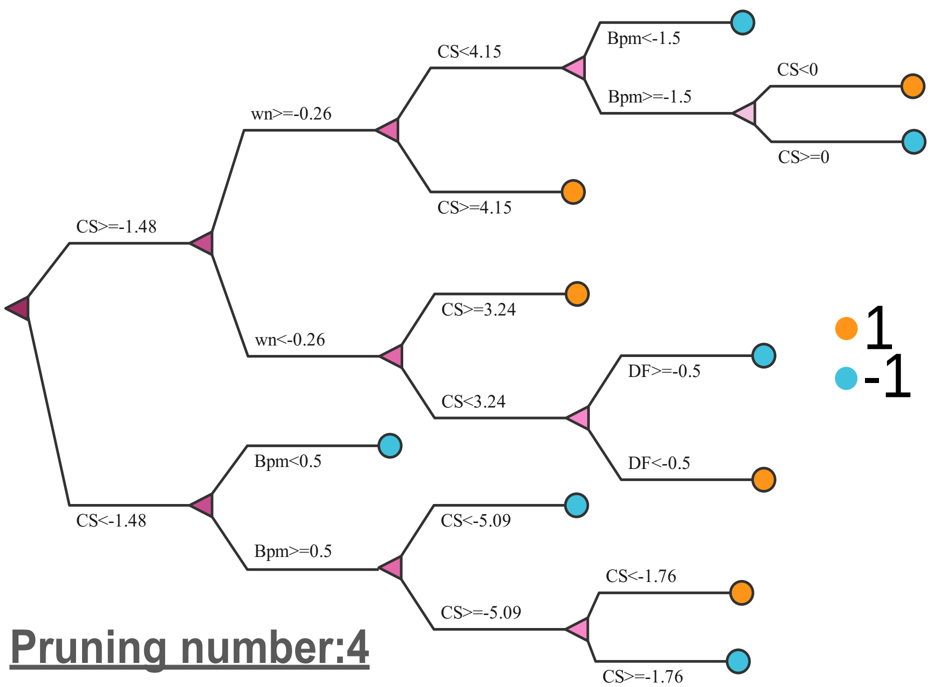
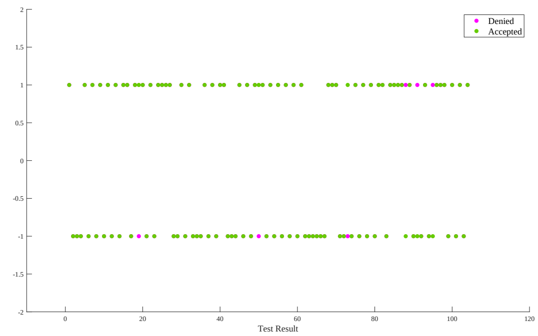
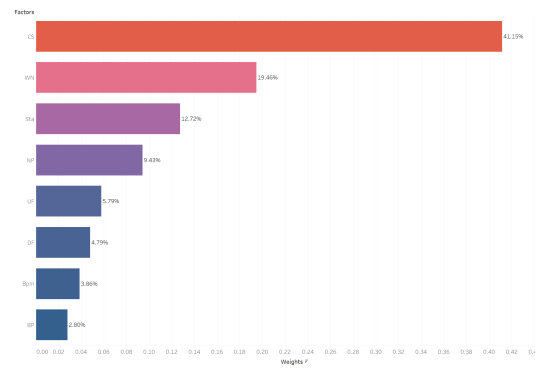


Figure 13: the Structure of Decision Tree based on All Matches

Under this structure of the decision tree, we obtain the accuracy of our prediction and the importance degree of each indicator as Figure 14. The positive one indicates the victotry of player one of one game, while the minus one stands for the victory of player two.



(a) Accuracy



(b) Importance Degree

Figure 14: The Result of a the Decision Tree based on All Testing Data

According to Figure 14, we observe that the accuracy of the decision tree is **94.23%**(98 out of 104), which shows that our model boasts a high accuracy in forecasting the swings in matches, and the top three relevant factors are **Consecutive Scoring**(CS)**, Winners**(WN)**, and Stamina Consumption**(Sta).

5.2 Trend-Prediction Model based on Vector Space(Model 3)

To dig deeper into the forecast of the flow of match, we want to establish a prediction model by learning the flow of game whose data has already been provided. Meanwhile, we notice that the working mode of Artificial Intelligence includes converting the input word to a word vector, searching for the most confident word vector within an existing vector space, and then combining and outputting it through algotithms like neural networks. As a result, we make an appropriate analogy and establish the trend-prediction model according to steps as follows.

Step 1: In order to facilitate the subsequent fitting and prediction, we first processed the calculated momentum data so that the processed momentum data reflects the overall trend of the match rather than the detailed momentum changes for every single game.

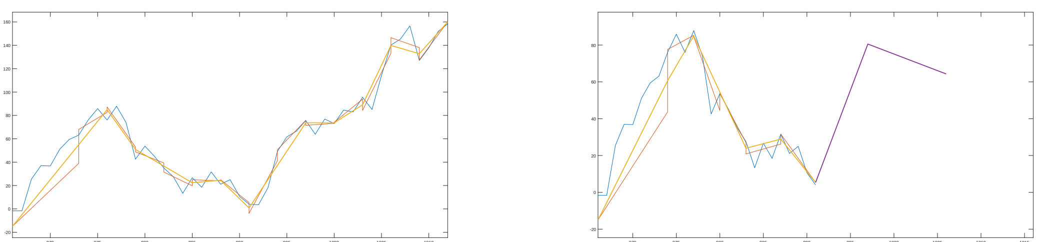
When the match reaches the n^{th} game, we perform a linear regression on the momentum from the $(n-3)^{th}$ to the n^{th} game, then on the momentum from the $(n-4)^{th}$ to the n^{th} game, and so on, up to the momentum from the $(n-9)^{th}$ to the n^{th} game. Among these seven linear regression processes, we select the one with the smallest fitting error (from the $(n-k)^{th}$ to the n^{th} set) and draw a fitted line segment. Then, we update n to n-k and repeat the above process until the entire match is fitted. In this way, the momentum graph becomes a line graph with minimized fluctuations.

Step 2: Next, due to the discontinuity at the connection points in the simplified momentum trend graph, further fitting is required. In the graph previously obtained, each breakpoint records the slope of the fitted line segment for the previous section. Starting from the last discontinuity point, we search backward and ignore the ones with similar slopes and keep the ones with dissimilar slopes. We then take the average of the upper and lower momentum values to obtain a line graph with minimized fluctuations and no breakpoints at the vertices, **which can represent the momentum pattern of the match**.

Step 3: We calculate the slope k_i of the i^{th} segment of the final line graph, and after scaling, obtain k_i^*. Considering that the line graph can be divided into a maximum of 15 segments among the 31 matches, we represent the pattern of a match in a 15-dimensional vector space using the vector $(k_1^*, \cdots, k_{15}^*)$ (with 0 for missing segments).By doing so, we obtain a set of 31 vectors as samples.

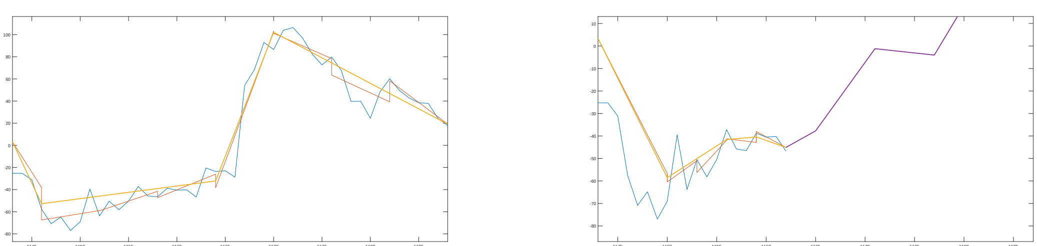
Step 4: Now, let's say a match has reached the n^{th} game, and we have fitted the momentum graph into a line graph with m segments using the previous method. The pattern can be represented by an m-dimensional vector v. In the sample set, we only consider the first m dimensions. We then calculate the cosine value between v and the 31 sample vectors in the m-dimensional vector space, and select the one with the highest value as the closest pattern. We then apply the subsequent match patterns of the closest sample to the current match for prediction.

Below are some of our predictions made according to the aforementioned steps compared with the real situation.



(a) Medevedev versus Eubanks(Real) (b) Medevedev versus Eubanks(Prediction)

Figure 15: Prediction 1



(a) Alcaraz versus Djokovic(Real) (b) Alcaraz versus Djokovic(Prediction)

Figure 16: Prediction 2

As shown in Figure 15 and Figure 16, the blue lines represents the difference between the real momentum index of two players, while the yellow ones stands for the fitted index. In the right column, the purple lines are the predictions we make according to the aforementioned steps. After comparison, we find that although our predictions cannot be accurate to every point of gain or loss, they can accurately reflect the macro trend of the game.

5.3 Pregame Instructions Based on SVM(Model 4)

To carry out a comprehensive suggestion for a player on how he or she may go into a new match against a different player, we should first understand the types of momentum shifts in the opponent's game. We implement this idea through **Support Vector Machine**.

Step 1: We create a two-dimensional space with the average momentum and momentum fluctuation as its two dimensions. Based on the data provided in the 2023 Wimbledon Championships and **Model 1**, we calculate the cumulative values of momentum for each player in each game and use their mean and variance as the two dimensions. Thus, the data is processed.

Step 2: Among the 32 players, we randomly select 28 of them as the training samples, and the four left as testing samples. After classification, we devide the players into two categories based on their performance of momentum, one is those with high mean of momentum index and low fluctuation which indicates player's comparatively high level of momentum, while the other the opposite. The result is visualized in Figure 17.

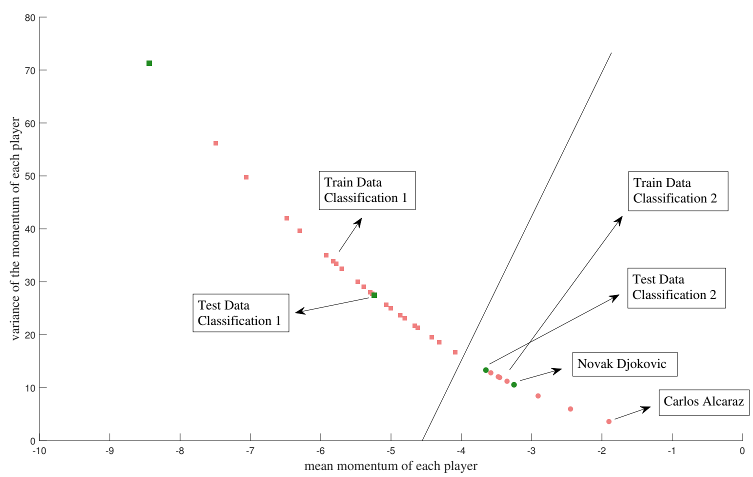


Figure 17: Result of SVM

From Figure 17, we discover that the error rate of our classification is **0%**, which indicates the accuracy of our model. We also observe that these top-ranked seeded players in the world, like Novak Djokovic and Carlos Alcaraz, all belong to the second category, which represents players with excellent momentum performance, which indicates the strong practicability of our model.

Based on the classification and our discussion, we may devise some pregame instructions for a player when going into a new match with different opponents.

When faced with players in the first category, that is, those players who have a comparatively low performance in Momentum Index:

- If the momentum level of the targeted player is classified as the first category, it is advisable to adopt a defensive strategy. This is because although the opponent's own strength may not be strong enough, the targeted player's aggressive style of play may lead to mistakes, resulting in a significant drop in momentum. On the other hand, this drop in momentum will allow the opponent's momentum to rise, downplaying their fluctuations and disadvantages. This makes it difficult to find breakthroughs and reverse the momentum.

  Instead, it is advised to use a defensive strategy, like focusing service games and avoiding the breaking serve by the opponent. These efforts may prevents you from falling into a significant disadvantage. Afterwards, by exploiting the weaknesses in the opponent's fluctuating momentum, the player may seize opportunities when the opponent makes mistakes or experiences a drop in momentum, and launch the attacks. This approach will be beneficial for breaking the opponent's momentum and turning the tide.

- If the momentum level of the targeted player is classified as the second category, it is recommended to take the initiative and adopt a more aggressive approach. By doing so, the player can strive to enhance the momentum in the early stages of the match, thereby suppressing the opponent and putting them under greater pressure. This will increase the likelihood of their momentum fluctuating and making mistakes.

  It should be noted that when having a sufficient advantage in momentum, it is advisable to play conservatively in order to reduce the consumption in stamina and preserve that advantage.

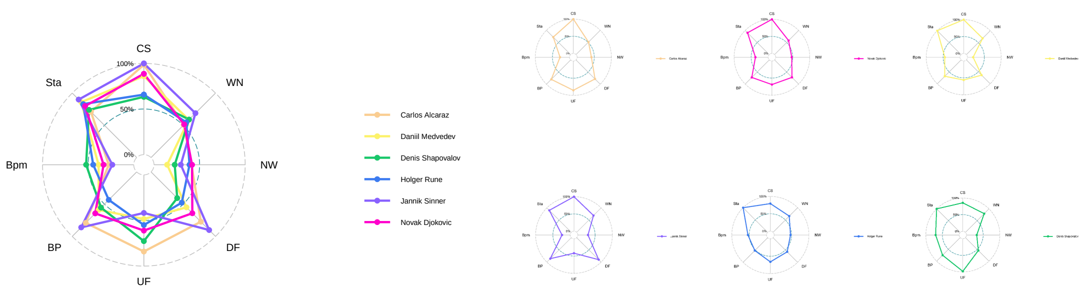
When faced with players in the second category, that is, those players who have a comparatively high performance in Momentum Index:

- If the momentum level of the targeted player is classified as the first type, it is recommended to take the initiative and make active attacks. Although having significant fluctuations in the player's own momentum may be a disadvantage, it also provides more opportunities to boost momentum. By daring to launch attacks, the opponent's mindset may be affected, potentially leading them to make mistakes or creating opportunities for the player to reverse the momentum, which may ultimately result in a victory.

- If the momentum level of the targeted player is classified as the second type, where both players are strong, maintaining the advantage in momentum can be quite challenging. Both players will be striving to break each other's serves, so it is important to minimize faults as much as possible and not to let our mentality be affected. During the match, it is crucial to stay excited and improve the momentum by achieving breakthroughs through breaks of serve and winning key points. However, it is important to observe the situation closely. When trailing by a significant margin, it is advisable to reduce excessive running, conserve energy for counterattacks, and minimize the depletion of momentum.

5.4 Case Study: Facing Top Players

To provide practical advice for players when facing different opponents, we study the characteristics of competition of six different top tennis players, namely Novak Djokovic, Carlos Alcaraz, Danil Medvedev, Jannik Sinner, Denis Shapovalov, Holger Rune, according to the eight dimensions mentioned in Table 3. After conducting **Min-Max Normalization**, we visualize the result in Figure 18.



(a) Radar Map of All Players Selected (b) Radar Map Each

Figure 18: Radar Map of Scores in Eight Indicators of the Top Players

Consider Alcaraz. When the Spaniard faced Djokovic in the final with no way out, there did exist such maneuvers for him to take to increase the chance of victory according to the aforementioned analysis. According to Figure 18, we can draw the conclusion that Djokovic scores considerably low in Unforced Fault and Double Fault, which indicates that Alcaraz may increase the momentum and secure victory by minimizing personal mistakes and utilizing strategies such as breaking serve and scoring consecutive points.

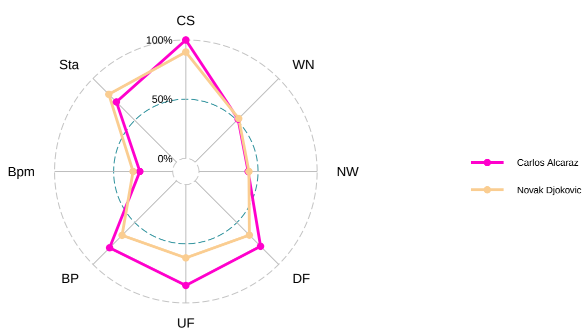


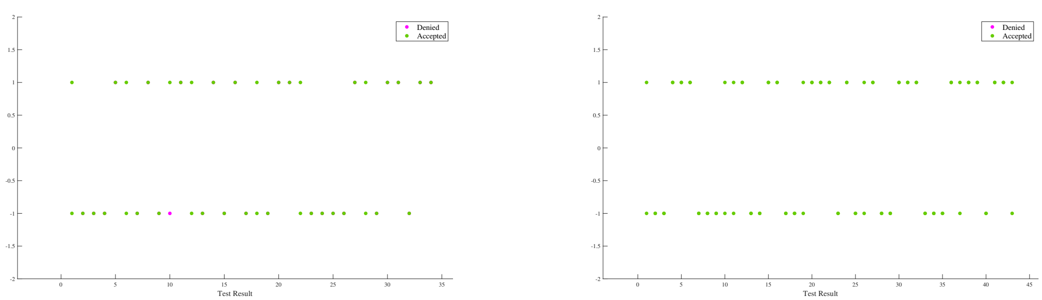
Figure 19: Comparison between Djokovic and Alcaraz

From Figure 19, we discover that Alcaraz did score higher than Djokovic in **Consecutive Scoring, Break Point, Unforced Fault and Double Fault**, which echoes with our suggestions carried out in the previous subsection. The discrepancy seen in Figure 19 also justifies Alcaraz's final victory, proving the accuracy and correctness of our model.

6 Task 4: Model Examination and Generalization

6.1 Testing Model 2 on Other Matches

Based on the **Decision Tree**, we apply the model to 2023-Wimbledon-1701 (Djokovic versus Alcaraz) and 2023-Wimbledon-1502 (Medvedev versus Eubanks) to test the model's accuracy. The result is shown in Figure 20.



(a) Djokovic versus Alcaraz (b) Medvedev versus Eubanks

Figure 20: Accuracy in Predicting other Matches in 2023 Wimbledon

From Figure 20, we discover that the accuracy of our prediction model based on Decision Tree is **98%**(40 out of 41) and **100%**(41 out of 41), and they showcase similar importance degree of contributors, which indicates that our model can predict the swings in play quite precisely.

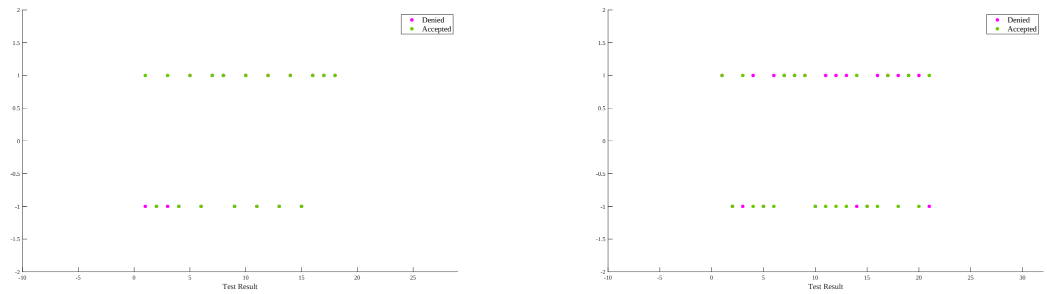
6.2 Generalization of the Model

Competition Formats

To test the generalization of model in different competiton formats, we look at the **Gentlemen's final of 2017 Miami Masters** (Roger Federer versus Rafal Nadal)[4] and **Women's final of 2016 French Open** (Serena Williams versus Garbine Muguruza)[5]. The former one follows a best-of-three format, resulting in a **faster pace**. The latter takes place on the **clay court**, and the **differences between men's and women's matches** also pose challenges to the generalization of the model. Similar with previous situations, we visualize the result in Figure 21.

From Figure 21, we find that the acuuracy of predicting swings is **90%**(18 out of 20) in Gentlemen's final of 2017 Miami Master and **66%**(21 out of 32) in Women's final of 2016 French Open. The reasons might be concluded as follows.

- The Miami Masters is a best-of-three match, which makes the pace of the game faster compared to Wimbledon. This puts greater pressure on players, leading to more errors and a



(a) Federer versus Nadal

(b) S.Williams versus Muguruza

Figure 21: Accuracy in Predicting Swings in Different Matches

decrease in consecutive points won. The distribution of players' physical exertion also differs from that of Wimbledon. Additionally, there are fewer break point opportunities. The combined effect of these factors results in using historical data from Wimbledon for predicting the Miami Masters causing some bias and a decrease in accuracy. However, the decrease in accuracy is minimal, indicating that the model is still suitable for predicting outcomes in the Miami Masters.

- The French Open, like the Miami Masters, follows a best-of-three format, so the changes in relevant factors in the latter are also reflected in the French Open. Additionally, the clay surface at the French Open results in lower ball speed compared to Wimbledon, which requires players to have better performance in receiving shots and leads to fewer winners. Furthermore, the differences in stamina between women and men can also lead to a decrease in the accuracy of prediction models. Despite the decrease in accuracy, it still stands at 66%, providing valuable reference for predicting match fluctuations.

Sports Type

To study the practibility of our prediction model in other sports, we choose table tennis as the optimal subject due to its similarity with tennis. Without loss of generality, we chose the Men's final of **Tokyo Olympics** as the sample, with the two sides being Ma Long and Fan Zhendong[6]. Similarly, we present the accuracy in Figure 22.

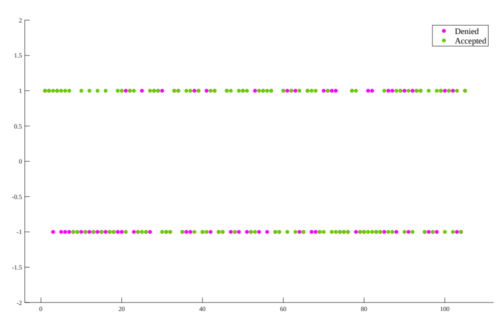


Figure 22: Accuracy in Predicting Swings in Men's final of Table Tennis in Tokyo Olympics

From Figure 22, we discover that accuracy of predicting the swings is **60%**(62 out of 105). which indicates a significant decrease in accuracy.

- Table tennis matches, although the name is similar to tennis matches, have many differences. Firstly, there is no service game in table tennis matches. Instead, the two players take turns serving, and the serving side has a relatively small impact on the probability of winning the point. Secondly, consecutive points and winning points have a significant impact on a player's momentum in table tennis matches. In addition, we have not taken into account the representative factors in table tennis matches, such as the roar after scoring, crucial timeouts to interrupt the opponent's hot streak, and other influences on fluctuations during the match. Albeit the flaw, our prediction model has reference value for predicting fluctuations in the match.

7 Strengths and Weaknesses

Strengths

- In Model 1, we have taken into consideration a wide range of factors with Cost-Benefit principle which comprehensively consider both positive and negative factors, allowing us to reasonably quantify momentum, which closely aligns with the direction of the match.
- Model 3 is capable of handling the trend of momentum, weakening fluctuations, and reflecting the overall trend of the match. By analyzing the overall trend, we can predict thw future flow of plays.

Weaknesses

- We do not consider the impact of off-site factors or the differences between players, which may underly the fluctuations in the obtained Momentum Index and the slight fluctuations and the inability to completely accurately simulate the course of the match.
- The sample size is relatively small, and sometimes there may be errors in handling fluctuations, leading to prediction inaccuracies.

8 Sensitivity Analysis

To verify that the fine tuning of contributors' weights has little effect on the trend of Momentum Index(2023 Wimbledon Gentlemen's final is chosen as a sample), random perturbations of 0 to epsilon are added to the eight selected contributors. As shown in Figure 23, under different perturbations, the contours of the figures are basically unchanged. This shows that our model is mot so insensitive to weights and robust, justifying the rationality of our results.

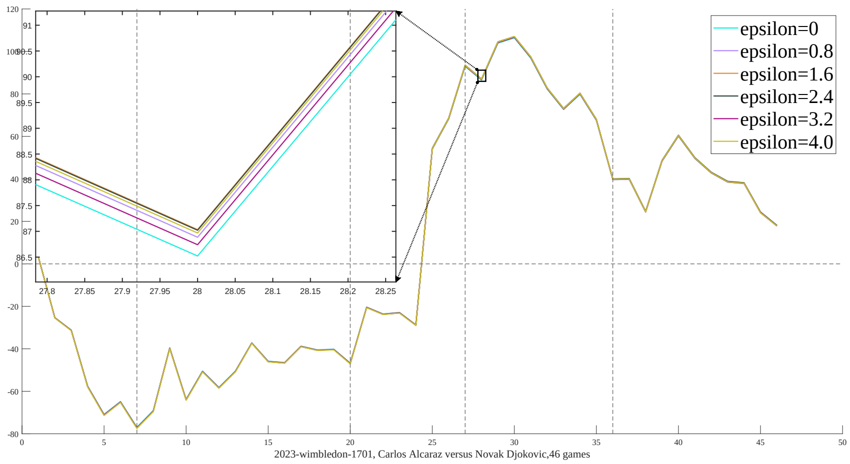


Figure 23: Momentum Index after Adding Different Perturbations to the Weights of Contributors

References

[1] cambridge. https://dictionary.cambridge.org/zhs/. 2024.2.2.

[2] Helmut Dietl and Cornel Nesseler. Momentum in tennis: Controlling the match. *UZH Business Working Paper Series*, (365), 2017.

[3] Arjun Goyal and Jeffrey S Simonoff. Hot racquet or not? an exploration of momentum in grand slam tennis matches. *arXiv preprint arXiv:2009.05830*, 2020.

[4] ATP. https://www.atptour.com/en/. 2024.2.4.

[5] WTA. https://www.wtatennis.com/. 2024.2.4.

[6] ITTF. https://www.wtatennis.com/. 2024.2.4.

Appendices

Some of our key codes are listed below:

```
%Decision Tree
datasum=zeros(1188,8); i=1;j=1;
while i<=7284 && j<=1187
    id=index(j+1);
    while i<id
        datasum(j,:)=datasum(j,:)+data(i,:); i=i+1;
    end j=j+1; end
datasum(1188,:)=sum(data(7279:7284))
j=1;Train_Data=zeros(7284,8);Train_Label=zeros(7284,1);
for i=1:1188
    if isequal(abs(wl(i,1)),1)
        Train_Data(j,1:8)=datasum(i,1:8);Train_Label(j,:)=wl(i,:);j=j+1;
    end
    ends
Train_Data=Train_Data(1:j-1,:);Train_Label=Train_Label(1:j-1,:);
Train_Data=Train_Data(1:1000,:),Train_Label=Train_Label(1:1000,:)
Tree = ClassificationTree.fit(Train_Data,Train_Label);
view(Tree);view(Tree, 'Mode', 'graph')
Tree_pre=predict(Tree,Test_Data)
scatter([1:1:j-1],Tree_pre,"magenta","filled")
hold on
scatter([1:1:j-1],Test_Label,[],[102,204,0]./255,"filled")
xlabel("Test Result"), legend("Denied","Accepted")
```

```
% Prediction of the trend of matches based on a High-dimensional Vector Space
  for i=0:k-2
      v=k-i; D(1,i+1)=fs*(C(v,1)-C(v-1,1))/(C(v,2)-C(v-1,2)); end
  for i=1:31
      tr=(D*transpose(VSPACE(i,1:k-1)))/(norm(D)*norm(VSPACE(i,1:k-1)));
      trust(i,1)=tr; end
  trustmatch=0; maxtr=0;
  for i=1:31
      if trust(i,1)>maxtr
          maxtr=trust(i,1); trustmatch=i;
      end
  end
disp(trustmatch); disp(trust(trustmatch,1));
DOT=CUT(trustmatch,:); zong=DOT(1,16);
delta=gmend-DOT(1,zong-k+1);
for i=1:zong-k+1
    DOT(1,i)=DOT(1,i)+delta;
end
DY(1,zong-k+1)=C(1,1);
for i=0:zong-k-1
    dx=DOT(1,zong-k-i)-DOT(1,zong-k+1-i);
    DY(1,zong-k-i)=DY(1,zong-k+1-i)+dx*VSPACE(trustmatch,k+i)/fs;
end
plot(transpose(DOT(1,1:zong-k+1)),transpose(DY(1,1:zong-k+1)),'LineWidth',2);
```

Momentum In Tennis

Introduction

Momentum in a tennis match can determine the flow of play to some extent, and whether a player can win. Therefore, the ability to change strategies in real time based on momentum can help players maintain their advantage or turn the tables in a tough situation.

Momentum's Role

Through our research, we found that players with high momentum tend to win the game. Momentum is determined by some objective factors, including technical aspects, mental aspects, and serve. Specifically, it is related to factors such as consecutive scoring, winners, break points, fault, net_pt won, and stamina consumption. For the serving games, positive factors that increase momentum will have less impact, while adverse factors will be magnified. The opposite is true for receiving games. Momentum is closely related to the direction of the game. By observing changes in momentum and underlying factors, we can predict the turning point of the game.

Advices

We found that consecutive scoring, winners, and stamina consumption have the greatest impact on momentum. Therefore, when these three factors change in a game, it is worth paying attention to and thinking about how to respond, so as to make momentum advantageous for oneself.

When our players score continuously, if the momentum is at a disadvantage, this is a great opportunity to overtake. The game situation will immediately undergo a reversal. Players should seize the opportunity and complete the comeback. If the momentum is at an advantage, this will be a great opportunity to increase the momentum. A slight momentum advantage is not enough to support victory, but a significant advantage will help players win. When our players lose points continuously, they should immediately adopt a defensive strategy to protect the momentum from falling sharply.

When our players hit a winner, cheering for them can boost their mentality and help them play better and increase the momentum in the following rounds. If it is the opponent's winner, we should train our players not to panic, keep calm, and avoid the momentum from continuing to decline.

When the opponent's playing style makes our players run more and consume more physical strength, we should train our players to adjust their pace and reduce physical exertion.

No matter the outcome,

facing the match positively is the best momentum!