Problem Chosen	2024 MCM/ICM Summary Sheet	Team Control Number
C		2410071

Tracking Momentum: Performance Quantification and Swings Prediction

Summary

Wimbledon Championships is the oldest tennis tournament in the world and has been held at the All England Lawn Tennis and Croquet Club in Wimbledon, London since 1877. Wimbledon is played on outdoor grass courts. For players, grasping the key points of the game is important. This paper gives **definitions of the momentum and swings of momentum**, aiming to build a **quantitative model** based on **psychological and strategic indicators** with scoring results to **predict the swings of momentum** for future use.

Before solving the task, we first construct indicators based on two categories: **Psychological and Strategic features**. Then we subdivide them into five subcategories and summarize 17 features in all. They are computed after examining and filling up raw data.

For Task 1, we build a model based on 17 features and the decision variable of whether the player wins the point. We employ **Random Forest** whose accuracy is the highest among several machine learning methods. Also, we visualize the importance of each feature and find that, the indicator: Server or Not ranks second in feature importance, which conforms to reality. We **visualize performance curve** of Alcaraz and Djokovic, which conforms to the match result to some extent.

For Task 2, we define momentum as a relative concept and calculate momentum through **differential between two players' performance**. Randomness test for momentum is based on the **mapping approach** which maps momentum to 0-1 series. Considering that Chi-Squared Test can only find out whether the proportion of 0s or 1s in the population of momentum is the same, we prefer using **Runs Test** for in-depth analysis and we find that the trend of momentum is **not random**. The swings in play are quantified by the method of **Cumulative Sum (CUSUM)** to detect significant change in the cumulative momentum shifting when each point occurs. Supposing that randomness mentioned in the task stands for the exponential distribution, we use **Kolmogorov-Smirnov test (KS test)** to test for randomness, finding that the swings in plays are **not random**. Our visualization also includes upward and downward turning points.

For Task 3, we adopt a new method that is closer to the real situation and determine flows of changes directly from the original data. We apply **CUSUM** to the difference of points gained by two players. Based on the 30 features upgraded from 17 features, we use **Light Gradient Boosting Machine (LightGBM)** to predict turning points after comparing the accuracy of different machine learning methods. We illustrate the importance of each feature. Among them, the top three is **The unforced error ratio of player1 (9.04%)**, **The number of shots during the point (8.90%)**, **The unforced error ratio of player2 (7.65%)**. Also, we obtain indicators contributing most to the occurrence of upward and downward turning points respectively. When analyzing a single player for providing advise, we put one player's data into the same model. Take Alcaraz as an example, the **biggest related feature** is Whether Alcaraz's previous goal was scored so we advise Alcaraz to calm down when losing points.

For Task 4, we make an in-depth analysis of the universality of the model. With the data extracted from **Australian Open Women's Single**, we verify the universality of the model, but features mentioned before need to be adjusted and new features should be added in the new model for better prediction.

With all the interactions between indicators and momentum swings, coaches and tennis athletes can grasp the key points of matches. Several sensible advice could also be made based on the model.

Keywords: Random Forest Kolmogorov-Smirnov Test Runs Test LightGBM

Contents

1 Introduction 3
- 1.1 Background and Literature Review . . . 3
- 1.2 Problem Restatement . . . 3
- 1.3 Our Work . . . 4

2 Assumptions and Notations 4
- 2.1 Model Assumptions . . . 4
- 2.2 Notations . . . 5

3 Data Processing 6

4 Task 1: Match Performance Analysis 6
- 4.1 Problem Analysis . . . 6
- 4.2 Indicators Selection . . . 7
- 4.3 Model Comparison and Selection . . . 8
  - 4.3.1 Model Comparison . . . 8
  - 4.3.2 Random Forest Model . . . 9
- 4.4 Feature Importance Analysis . . . 10
- 4.5 Performance Visualization . . . 10

5 Task 2: Test for Randomness of Momentum and Swings 12
- 5.1 Testing the Randomness of Runs of Success . . . 12
- 5.2 Testing the Randomness of Swings in Play . . . 14

6 Task 3: Predicting Swings in the Match 15
- 6.1 Problem Analysis . . . 15
- 6.2 Detecting Turning Points . . . 15
  - 6.2.1 Establishment of the Model . . . 15
  - 6.2.2 Solution and Result . . . 16
- 6.3 Predicting Turning Points . . . 16
  - 6.3.1 Establishment of the Model . . . 16
  - 6.3.2 LightGBM Model . . . 17
  - 6.3.3 Solution and Result . . . 19
- 6.4 Analysis for a Single Player . . . 20

7 Task 4: Universality of Model 21
- 7.1 Extra Data Collection . . . 21
- 7.2 Model Test on Extra Data . . . 22
- 7.3 Factors to be Considered in Future Models . . . 22

8 Strengths and Weaknesses 23
- 8.1 Strengths . . . 23
- 8.2 Weaknesses . . . 23

1 Introduction

1.1 Background and Literature Review

The Championships, commonly known simply as Wimbledon, is the oldest tennis tournament in the world and is regarded by many as the most prestigious[1]. Winning the Wimbledon trophies, undoubtedly, is a paramount victory for all the tennis players. Matches in the Gentlemen's Singles are best-of-five sets while the matches in the all other events are best-of-three sets.

The Wimbeledon championship is a competition requiring excellent tennis techniques and also something correlated with perseverance and mindset. As the longest match at Wimbledon lasted 11 hours and 5 minutes and the longest final took 4 hours and 57 minutes, set by Djokovic and Federer. Moreover, we have also seen many dark horses overtaking seeded players to get a ticket for the final. A close game of seven and a thrilling come-from-behind victory all indicate that Wimbledon can tip the winning scales in a matter of moments.

Chuck Kriese once said that tennis more than any other sport, is a game of momentum. Momentum can act as a game changer and even indicates the victory, while the swings in Momentum are referred to as turning points[2]. Regarded as a highly-related reference for tennis competition, it can also be classified as strategic momentum and psychological momentum[3], which can help the coach to analyze the situation and seek the critical points of the whole competition. Therefore, with the data from Wimbledon 2023 men's matches after the first 2 rounds, a quantitative model to predict and analyze the momentum according to the scores as well as important indicators for tennis like ace serving is expected.

1.2 Problem Restatement

Based on the background, in this paper we are going to solve the following problems:

- **Task 1:** Develop a model to illustrate the flow of play and visualize it. The model should quantify which player performs better at a given time. Some indicators like 'server' should be considered.
- **Task 2:** Use the model developed or metrics to show that swings and runs of success are random or not.
- **Task 3:**
  - Develop a model that can predict when the flow of play shifts from favoring one player to the other and show most related factors.
  - Give advise on a player going into new match against a different player.
- **Task 4:** Test the model on other matches and analyze its performance. Analyze factors can be included to improve the performance. Measure the generality of the model.
- **Task 5:** Summarize the results in one- to two-page memos, which include advice for coaches on "momentum" and for players to respond to events during a tennis match.

1.3 Our Work

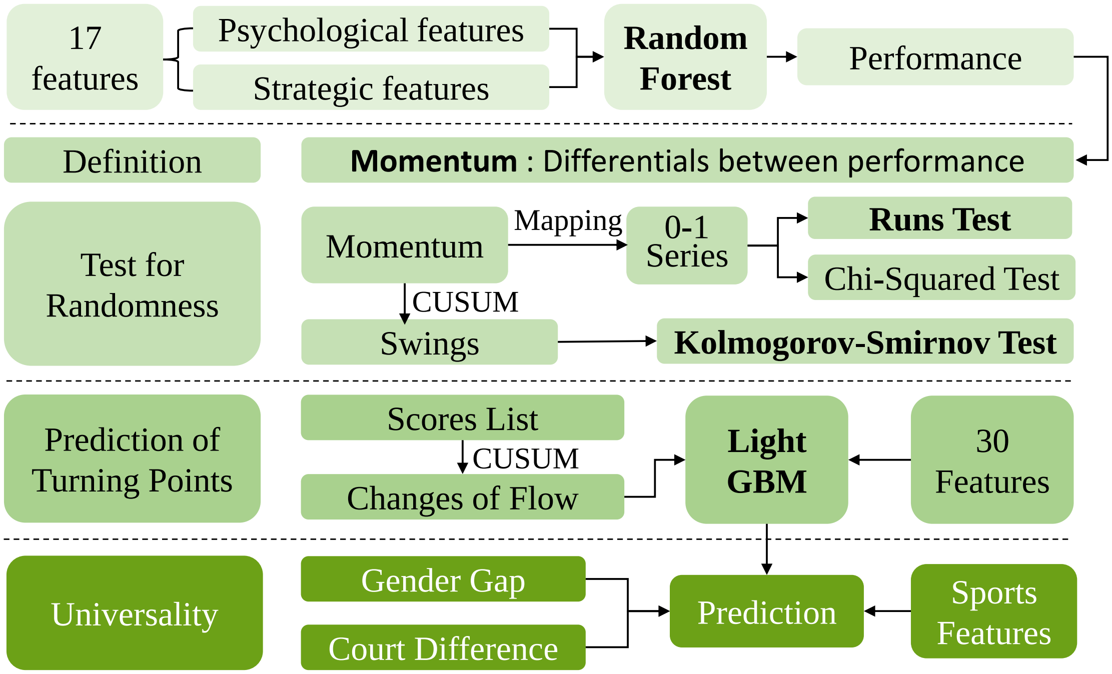

Figure 1: Flowchart of our work

Firstly, we processed the original data to construct indicators that can be classified into two categories: psychological and strategic features. We continued to subdivide these features into five subcategories. Based on 17 features found, with the decision variable of whether the player wins the point, we used Random Forest to reflect a player's performance and visualized the performance curve.

Secondly, we defined momentum as a relative concept and tested randomness of momentum and swings respectively. For the momentum test, we mapped momentum to 0-1 series, considering the defects of Chi-Square Test, we chose Runs Test as the final testing method. For the swings test, we employed CUSUM to find the swings and employed Kolmogorov-Smirnov Test for randomness test.

Thirdly, in order to predict the occurrence of turning points, we adopted a new method that used scores list directly computed from the original data and applied CUSUM for detecting changes of flow. With the 30 features upgraded from Task 1, we employed LightGBM for prediction.

Finally, we made an in-depth analysis of the universality of the model. With the data extracted from Australian Open Women's Singles, we verified the universality of this model to some extent and also put forward more features extracted from gender gap, court difference and sports features for better prediction.

2 Assumptions and Notations

2.1 Model Assumptions

Considering the conditions required for modeling,we make following assumptions:

- **Assumption 1:** No sudden change exists in players' skills and physical quality.

  Justification: We suppose that the existence of ratios(including Ace, Double Fault, Unforced error, etc) belongs to the strategic momentum factors which will not be affected by an interruptions like the intervals and tactical injury suspensions.

- **Assumption 2:** The indicators in one game or from previous points play a more important role in the whole match.

  Justification: The cross-game indicators may be affected by interruptions like the intervals and tactical injury suspensions in which time players can refresh themselves and make adjustments. That's also the reason why we define these indicators as psychological momentum.

- **Assumption 3:** The performance indicators indicate the probability of which player is more inclined to score, relatively, scoring is the most basic means of judgment.

  Justification: Our momentum is based on performance,so the performance indicators act as the bridge between scoring and momentum.

- **Assumption 4:** Our momentum is a relative concept.

  Justification: Our momentum is more inclined to depict the scoring trend for the entire game rather than for individual players. So the momentum in following models is a zigzag line. Our definition for momentum is more like the differentials between two parties.

2.2 Notations

Here are the notations used in this section:

Symbol	Definition
P_1	Performance of Player 1
P_2	Performance of Player 2
M	Momentum
M_m	Mapping of Momentum to 0 or 1
p_0	The Frequency of Occurrence of 0s
p_1	The Frequency of Occurrence of 1s
R	The Total Number of Runs
n_0	The Number of 0s in Runs Test
n_1	The Number of 1s in Runs Test
T_i	The Recording of Turning Points
ΔT_i	Intervals between Consecutive Turning Points
λ	Mean Arrival Rate

Table 1: Notations

3 Data Processing

The dataset `Wimbledon_featured_matches.csv` contains 31 matches from Wimbledon 2023 men's matches after the first 2 rounds with each set and each game. Each row contains the indicators of two players such as the server, ace and so on.

However, in the `speed_mph` column, because of the missing serve and missing data, some squares are recorded as NA, which cannot be considered into the model, so we first want to delete these data for a more precise prediction while the number of NAs reached 754, accounting for more than ten percent of the dataset. Then we choose to replace NA with the previous row's serving speed data based on consideration of accuracy.

The reason why we choose this replacement method is that we find that the average serving speed is 112.40 while the variance is about 12, even if the lowest speed is 72 and the highest speed is up to 141, the serving speeds are mostly centered around 100-130 (see Fig.2). Through the process of data-searching, we find that the serving speeds above the NA speed mostly center in 100-130, which have high frequency of occurrence (except 2023-wimbledon-1310 and 2023-wimbledon-1311 where all the speed is NA, so in these games, we disregard this factor, this may not have a great impact on overall results, as there are 31 matches overall), we can suppose that the NA (because of missing servings) can be replaced by the the speed of the previous row. Therefore, this replacement method is somewhat reliable.

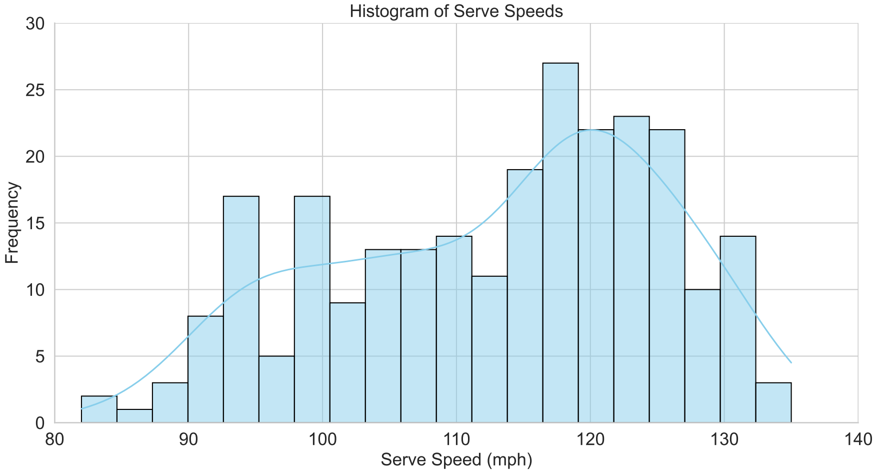

Figure 2: Histogram of serve speeds

4 Task 1: Match Performance Analysis

4.1 Problem Analysis

Since it is difficult to measure 'momentum' during the match and even it is not apparent how various indicators will affect 'momentum', simply using linear combinations in the indicators given in the dataset may be not reasonable. Considering the factors influencing the players' performance and scoring abilities, we extract some data as features from the original dataset and use **machine learning** methods to predict the performance of two players multidimensionally.

4.2 Indicators Selection

According to paper[3], psychological momentum(features) and strategic momentum(features) are distinguished by whether they will be interrupted by exogenous interruptions.

Feat.	Explanation
x_1	Leading score in the whole match
x_2	Leading score in the whole set
x_3	Leading score in the whole game
x_4	Serve or not in one game
x_5	Ace or not in one game
x_6	Double fault or not in one game
x_7	Unforced error or not in one game
x_8	Whether the previous goal was scored or not
x_9	Whether the previous goal was Ace or not
x_{10}	Whether the previous goal was unforced error or not
x_{11}	Ace ratio
x_{12}	Double fault ratio
x_{13}	Unforced error ratio
x_{14}	Ratio of net point won
x_{15}	Break point conversion ratio
x_{16}	Total distance ran during match (meters)
x_{17}	Speed of last serving

Table 2: 17 features explanation

- **Leading scores** ($x_1 \sim x_3$): including the leading scores in the whole match/set/game.

  The leading scores can affect the players' following tactics. We suppose that this influence may become the factors of **psychological momentum**.

- **Decision variables in one game** ($x_4 \sim x_7$): including the variables deciding whether the player acts as the server, whether the serving is an ace, whether the player makes a double-fault, and whether the player makes an unforced error.

  The reason why we choose to define the decision variables only in one game is that the tennis intermission may affect the momentum of players within different games, which is called exogenous interruptions[3]. We also suppose that these influences may become the factors of **psychological momentum**.

- **Decision variables with previous goals** ($x_8 \sim x_{10}$): including the variables deciding whether the previous goal is scored, whether the previous goal is an ace, and whether the previous goal is lost by an unforced error.

These variables, extracted from previous goals, are further indicators from the above features. Similarly, we suppose that these influences may become the factors of **psychological momentum**.

- **Ratios** ($x_{11} \sim x_{15}$): including Ace/Double fault/Unforced error ratios related to the performance of a ball game, also the ratios of scoring methods containing the Ratio of winning the point while at the net and the break point conversion ratio.

  By calculating the ratios of three important factors mentioned before, we can measure a player's skill level. Rather than the psychological momentum, we are prone to suppose that the ratios are consistent with the **strategic momentum**.

- **Other related factors** (x_{16}, x_{17}): something may be related to the fatigue level of the players including the Total distance ran during point (meters) and an important indicator reflecting the serving quality which is called the Speed of last serving.

  These factors reflect both psychological and strategic momentum as they are composite indicators.Also, due to the data recorded in the factors have overstepped the scale,so we standardize the data.

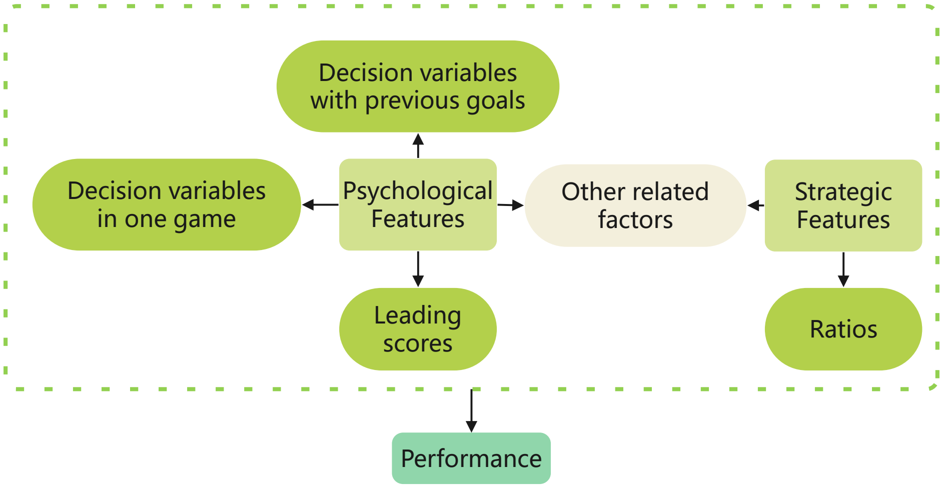

Figure 3: Performance factors

We calculate the correlation coefficients between various features and the label (see Fig.4). From the heatmap, we can find that whether the player serves or not contributes a lot to the label.

4.3 Model Comparison and Selection

4.3.1 Model Comparison

We use the above 17 indicators as features $x_1 \sim x_{17}$. Performance of a player can be partly reflected by whether the player wins the point, so we choose it as labels y. Since the label values are 0 or 1, we

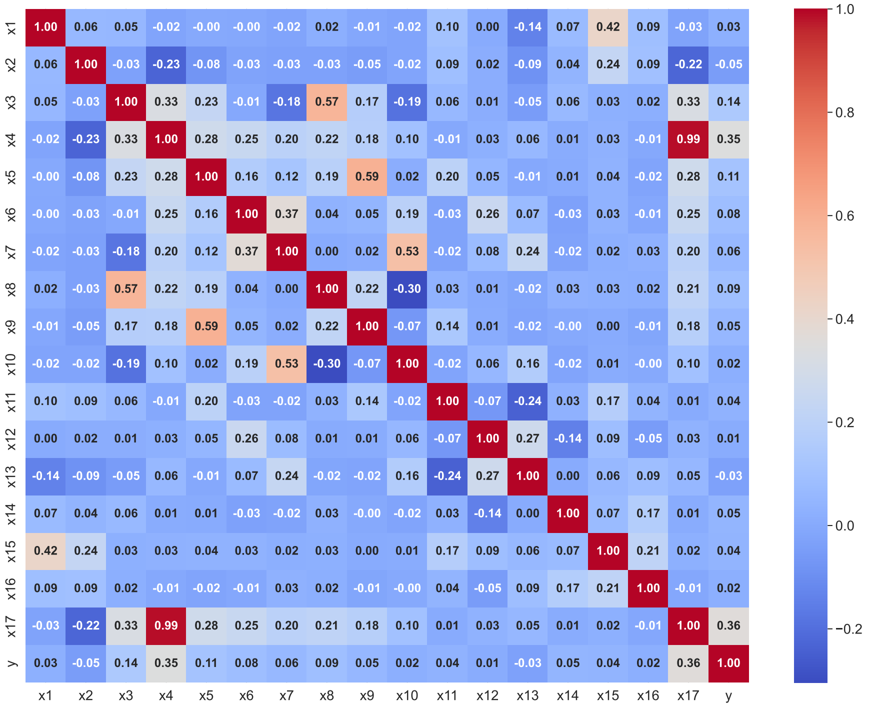

Figure 4: Correlation heatmap

can view this problem as a classification problem. We have tried different machine learning methods and get Tab.3.

Method	Accuracy	Precision	Recall	F1
SVM	0.564	0.564	0.573	0.533
XGBoost	0.641	0.641	0.641	0.641
LightGBM	0.669	0.669	0.670	0.669
KNN	0.640	0.640	0.641	0.640
Logistic Regression	0.666	0.666	0.669	0.665
Random Forest	0.687	0.710	0.655	0.681
Decision Tree	0.660	0.660	0.661	0.660
BP Neural Network	0.671	0.671	0.672	0.671

Table 3: Comparison of different models on predicting performance

4.3.2 Random Forest Model

Through the table, we can conclude that **Random Forest** method performs the best, therefore, we choose it for further analysis. Since the model achieves an accuracy of **68.7%**, we conclude that our model can to some extent reflect a player's performance.

Random Forest (see Fig.5 is an ensemble learning method that combines multiple decision trees to improve predictive accuracy and control over-fitting. Each tree in the forest is built from a random sample of data and at each split, a random subset of features is considered, contributing to the diversity of the model. The final prediction is made by aggregating the predictions from all trees, typically by a majority vote for classification or averaging for regression. This approach leverages the wisdom of the crowd, while individual trees may make errors, collectively they produce robust and generalizable results.

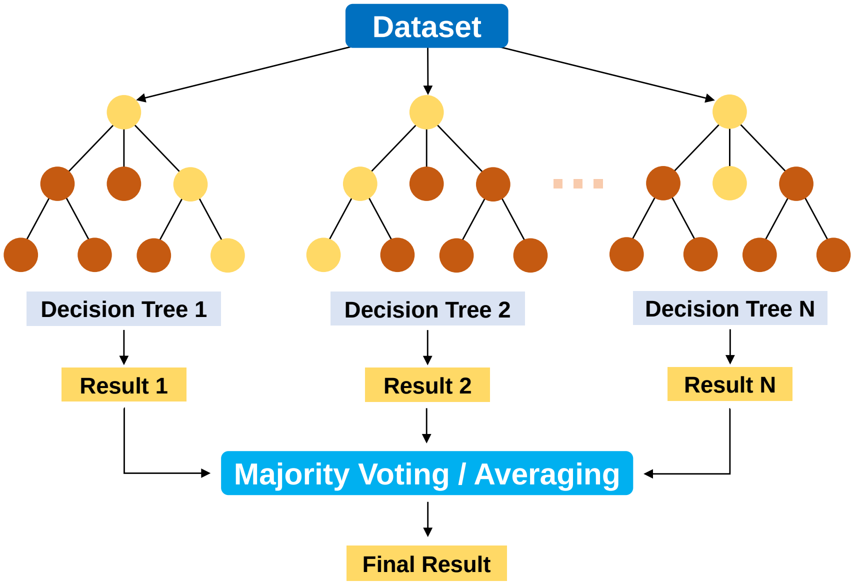

Figure 5: Random Forest

4.4 Feature Importance Analysis

Fig.6 shows the importance of each feature. In our model, x_{17}, x_4, x_{13} are three most important features, which stands for **Speed of last serving**, **Server or not**, **Unforced error ratio**, which corresponds to the truth that in tennis, the player serving has a much higher probability of winning the point/game.

4.5 Performance Visualization

To get the explicit value of 'Performance', we invoke the `predict_proba()` method, using the predicted probability of the label being 1 as the value P_1 and P_2. In Fig.7, we choose match 2023-wimbledon-1701, Carlos Alcaraz vs. Novak Djokovic as an example of visualization.

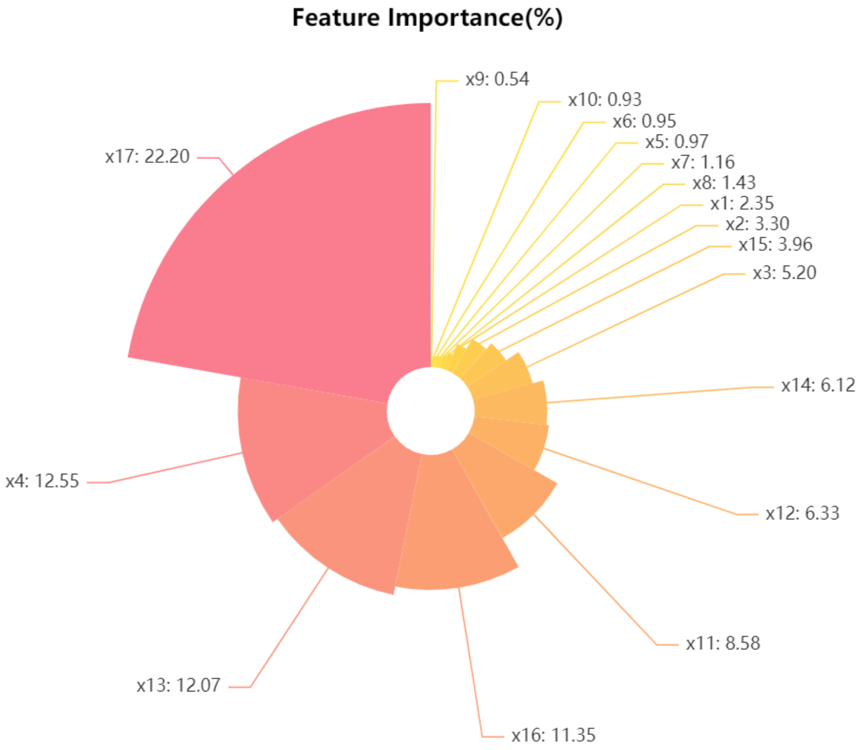

Figure 6: Feature Importance

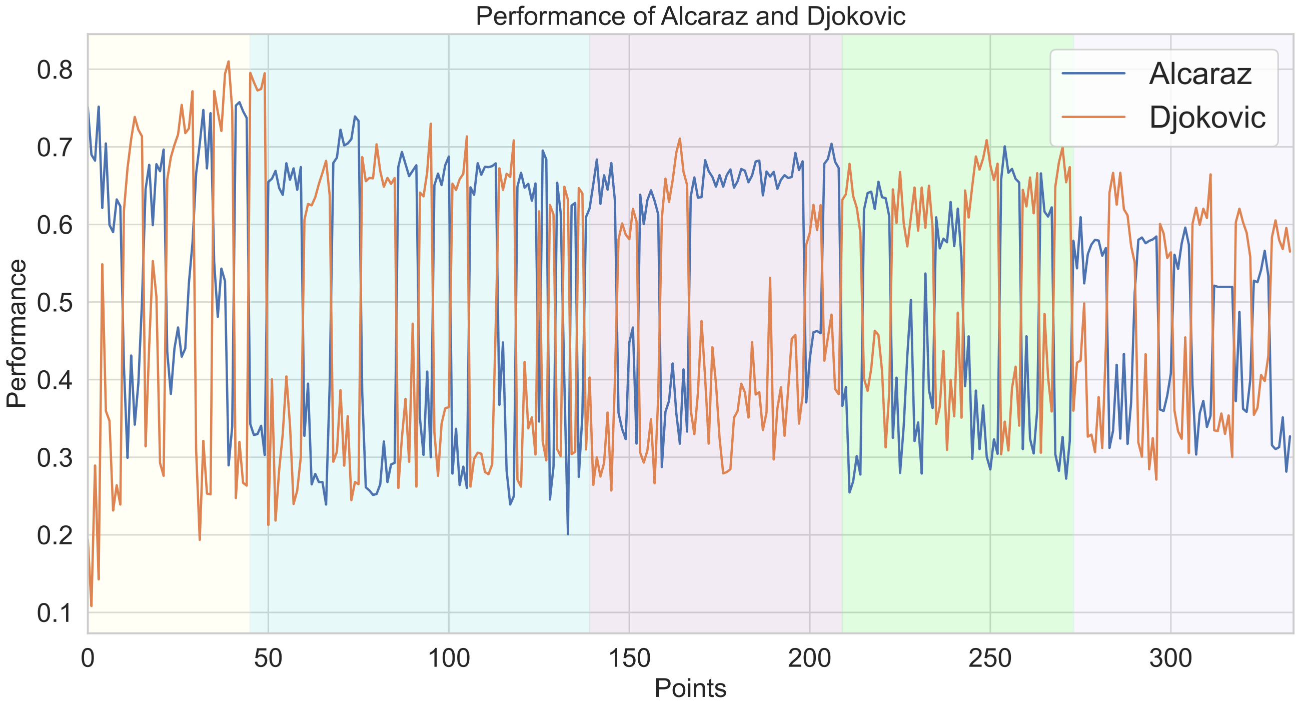

Figure 7: Visualization of Performance

The different background colors signalize the five sets in the battle between Alcarez (Player 1, the blue curve) and Djokovic (Player 2, the orange curve). The predicted performance can be an indicator judging which player wins the set. Taking set-3 (the purple area) as an example, we can see that the trend of the curve for Player 1 is relatively higher than that of Player 2. Actually the result of this set is 6-3, which is corresponding to the performance curve.

5 Task 2: Test for Randomness of Momentum and Swings

5.1 Testing the Randomness of Runs of Success

”**Runs of success**” refers to consecutive occurrences of successful points gained by one player. From our points of view, ”runs of success” depends on the player’s ”momentum”, which can serve as a quantifiable indicator for runs of success. Given the visualization result in Task 1, we have a new definition for momentum. We suppose momentum is a **relative** concept and it is affected by both sides. Therefore, we define momentum as the **differential between the two performances**. Represent in mathematical notation:

$$M = P_1 - P_2$$

Since $P_1, P_2 \in [0, 1]$, $M \in [-1, 1]$. If $M > 0$, the flow of play favors player 1, otherwise, the opposite. Fig.8 illustrates the variation of the momentum between player 1 and player 2 throughout the entire course of the game. The larger the absolute value of the difference, the darker the color. Red color is assigned when the performance of player 1 is higher than player 2, while blue color is assigned when the performance of player 1 is lower than player 2.

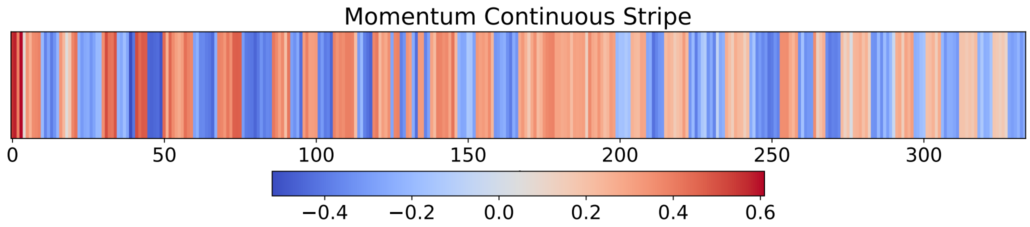

Figure 8: Momentum visualization

However, when examining whether the momentum is random or not, we adopt a **mapping approach**:

we define a new variable

$$M_m = \begin{cases} 0, M < 0 \\ 1, otherwise \end{cases}$$

when player 1’s performance value outweighs player 2, then we set M_m at 1. Similarly, if player 2’s performance value outweighs player 1, we set M_m at 0. In all, the mapped momentum variable M_m is valued at 0 or 1. Then, our goal becomes **testing the randomness of the 0-1 series**.

By the preprocessing of the definition of momentum, now we can verify whether the trend in momentum is random or that there are certain logical patterns to follow when forecasting the momentum. We use two test methods to examine the existence of randomness.

First, we use chi-squared test to test the distribution of 0s and 1s. We counted the frequency of occurrence of 0s and 1s for the momentum throughout the match. Then we performed a chi-square test on the frequency of 0s or 1s. Our assumptions are as follows:

$$H_0 : p_0 = p_1 = 0.5$$

$$H_a : p_0 \neq p_1$$

By performing the chi-squared test, the **p-value is 0.0628**. Under the significant level of $\alpha = 0.05$, the null hypothesis cannot be rejected. This result is consistent with the fact that Carlos Alcaraz and Novak Djokovic are evenly matched.

However, this test can only find out whether the proportion of 0s or 1s in the population of momentum is the same. To tell whether the trends in momentum is random or not, we should test whether the occurrences of 0s or 1s in the sequence of momentum are random. The following Fig.9 depicts the occurrences of 0s and 1s in player 1's momentum. By observing the image, it can be noticed that the probability of consecutive occurrences of 0s and 1s is very high.

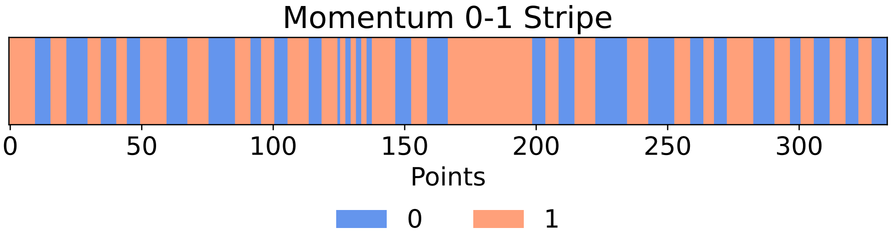

Figure 9: Momentum 0-1 series visualization

We apply **Runs Test** to further test whether the occurrences of 0s and 1s are random. The runs test is a statistical test used to analyze the presence of randomness or patterns in a sequence of data. It focuses on the occurrence of consecutive runs, where a run is defined as a sequence of consecutive data points that share the same characteristic. For example, in the sequence "110011", there are three runs"11", "00", "11". The assumptions of runs test are as follows:

$$H_0 : \text{The sequence is random, } H_a : \text{The sequence is not random}$$

The steps of runs test are as follows:

- Step1 : Counting the total number of runs (R).
- Step2 : Counting the number of 0s (n_0) and 1s (n_1).
- Step3 : Calculating the expected value and variance of R.
- Step4 : Calculate the Z-statistic and p-value.

The calculation methods for the aforementioned statistical measures are as follows:

- $\mathbb{E}(R) = \dfrac{2n_0n_1}{n_0 + n_1} + 1$

- $Var(R) = \dfrac{2n_0 n_1(2n_0 n_1 - n_0 - n_1)}{(n_0+n_1)^2(n_0+n_1-1)}$
- $Z = \dfrac{R - \mathbb{E}(R)}{\sqrt{Var(R)}}$

Then we apply the runs test to the sequence of momentum. The test result turns out that the **p-value is 0.00000** and rejects the null hypothesis. We conclude that the trend of momentum is **not random**.

5.2 Testing the Randomness of Swings in Play

To quantify ”swings in play”, we define ”turning points” to exhibit the difference between two players’ performance. When there is a turning point, the swing occurs.

To detect turning points, we applied the method of **CUSUM** (Cumulative Sum). CUSUM is used for monitoring and detecting shifts or changes in a process over time. The basic idea behind CUSUM is to calculate the cumulative sum of the differences between the observed values and a target value. If the cumulative sum exceeds a specified threshold, it indicates a significant change.

In our model, we define momentum as the differential between performance of two players. Then we use a counter to record the cumulative sum of the changes in momentum when each point occurs. We set the threshold at 0.5. When the counter outweighs 0.5, we record an ”upward turning points”, and reset the counter to zero. When the counter falls below -0.5, we record an ”downward turning points”, and reset the counter to zero. Fig.10 depicts the momentum of the match between Alcaraz and Djokovic and turning points through the game.

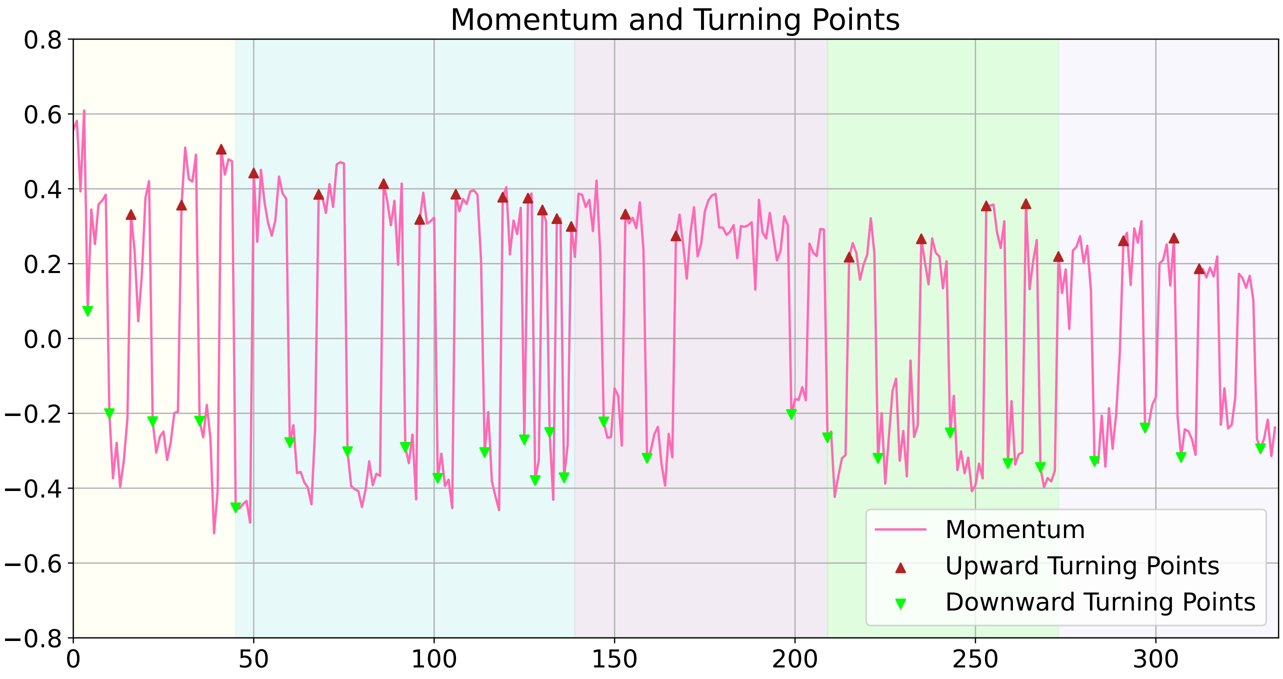

Figure 10: Momentum and turning points

In probability theory and statistics, the exponential distribution (also known as the exponential distribution) is a probability distribution that describes the time between events in a Poisson process, where events occur continuously and independently at a constant average rate. We suppose ”random” mentioned in Task 2 refers to exponential distribution.

To further test whether the occurrences of turning points are random or conform to an exponential distribution, we apply Kolmogorov-Smirnov test. The steps of the KS test are as follows:

- Step1 : Record every turning points (T_i).
- Step2 : Record the intervals ($\Delta T_i = T_{i+1} - T_i$) between consecutive turning points.
- Step3 : Performing an exponential distribution fitting on the data of intervals and estimating the λ parameter. $\left(\tilde{\lambda} = \dfrac{1}{\overline{\Delta T}}\right)$
- Step4 : Perform Kolmogorov-Smirnov test to test whether the distribution of time intervals conforms to an exponential distribution.

The assumptions of Kolmogorov-Smirnov test are as follows:

$$H_0 : \text{The distribution of time intervals conforms to an exponential distribution}$$

$$H_a : \text{The distribution of time intervals doee not conform to an exponential distribution}$$

By performing the K-S test, the **p-value is 0.00003**. Under the significant level of $\alpha = 0.05$, the null hypothesis is rejected, which means that the distribution of time intervals does not conform to an exponential distribution

In conclusion, the occurrences of turning point, (i.e., swings in play) are **not random**.

6 Task 3: Predicting Swings in the Match

6.1 Problem Analysis

A tennis match always has a lot of ups and downs. In order to better monitor the flow of play, predict these swings and decide which factors are most related, we develop a model that weighs the importance of different features. By applying the model, advise can be given to a player when he goes into a new match against a different player.

6.2 Detecting Turning Points

6.2.1 Establishment of the Model

To monitor the flow of play, instead of detecting turning points based on Task 2, we adopt a new method that is closer to the real situation and can determine ”turning points” directly from the original data. If a player has a trend of continuously gaining points, we say it indicates a turning point. This can be attained by the difference of sum of points gained by two players.

To realize the method above, we applied the method of CUSUM, as we mentioned in Task 2. In our model, we use `score_list` to record the **difference of points gained by two players** as the play goes through. When player 1 gains a point, add 1 to the sum of points and addend it to `score_list`. When player 2 gains a point, subtract 1 from the sum of points and addend it to the list. Fig.11 visualizes this process.

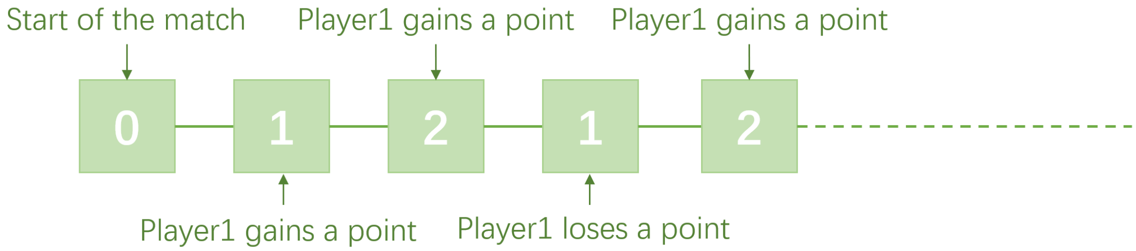

Figure 11: Score list visualization

Then applying CUSUM method, we use a counter to record the cumulative sum of the list when each point occurs. We set threshold at 3. When the counter outweighs 3, we record an ”upward turning points”, and reset the counter to zero. When the counter falls below -3, we record an ”downward turning points”, and reset the counter to zero.

6.2.2 Solution and Result

Fig.12 depicts `score_list` of 2023 Wimbledon Gentlemen’s final and turning points through the game. The figure clearly shows that Alcaraz kept losing points and met four downward turning points in set 1. The second set was tense and consisted of several upward tunring points and downward turning points and was finally won by Alcarez. Alcarez owned five upward turning points in the third set and won this set. Alcaraz seemed in total control as the fourth set started and owned an upward turning point, but a change of direction occurred and the young Spaniard went downward in this set. The final set started with Djokovic maintaining the advantage of the fourth set, but soon Alcaraz won three upward turning points and won the championship.

This figure clearly shows that these ’turning points’ can to a large extent reflect the flow of play.

6.3 Predicting Turning Points

6.3.1 Establishment of the Model

Recalling the features of players we mentioned in previous tasks, by concatenating them and adding one more feature (rally counts), we constructed a total of **30 features** to predict the occurrence of turning points. The list of 30 features are as Tab.4 show.

All these factors have been normalized. Then we put all these 30 features and 1 label (-1, 0, or 1, standing for down turning point, non turning point and up turning point repectively) into difference models and compare their performance (see Tab.5).

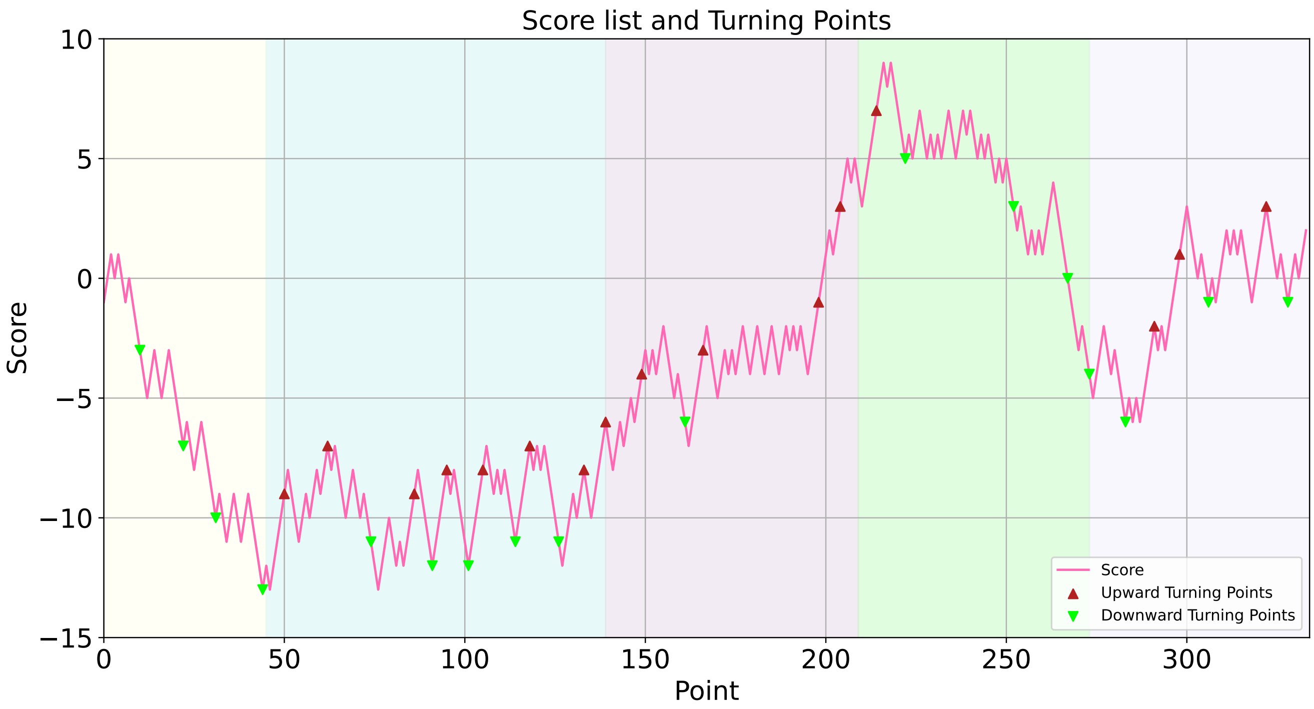

Figure 12: Alcaraz's score list and turning points

Method	Accuracy	Precision	Recall	F1
SVM	0.889	0.889	0.791	0.837
XGBoost	0.907	0.909	0.898	0.882
LightGBM	0.909	0.909	0.902	0.884
KNN	0.887	0.887	0.806	0.838
Logistic Regression	0.889	0.889	0.821	0.838
Random Forest	0.893	0.893	0.899	0.846
Decision Tree	0.893	0.893	0.878	0.880
BP Neural Network	0.888	0.888	0.791	0.837

Table 5: Comparison of different models on predicting turning points

6.3.2 LightGBM Model

We choose LightGBM model which reaches an accuracy of **90.87%**. It somehow proves the efficiency of the model. LightGBM uses a leaf-wise growth strategy (Fig.13), which finds one leaf at a time with the largest splitting gain from all the current leaves, then splits it and so on. While other Gradient boosting decision trees use level-wise growth strategy which traverses the data once can split the leaves of the same layer at the same time but with low efficiency.

Feat.	Explanation	Feat.	Explanation
x_1	p1-Leading score in the whole match	x_{16}	p1-Speed of last serving
x_2	p1-Leading score in the whole set	x_{17}	p2-Ace or not in one game
x_3	p1-Leading score in the whole game	x_{18}	p2-Double fault or not in one game
x_4	p1-Ace or not in one game	x_{19}	p2-Unforced error or not in one game
x_5	p1-Double fault or not in one game	x_{20}	p2-Whether the previous goal was scored or not
x_6	p1-Unforced error or not in one game	x_{21}	p2-Whether the previous goal was Ace or not
x_7	p1-Whether the previous goal was scored or not	x_{22}	p2-Whether the previous goal was unforced error or not
x_8	p1-Whether the previous goal was Ace or not	x_{23}	p2-Ace ratio
x_9	p1-Whether the previous goal was unforced error or not	x_{24}	p2-Double fault ratio
x_{10}	p1-Ace ratio	x_{25}	p2-Unforced error ratio
x_{11}	p1-Double fault ratio	x_{26}	p2-Ratio of winning the point while at the net
x_{12}	p1-Unforced error ratio	x_{27}	p2-Break point conversion ratio
x_{13}	p1-Ratio of winning the point while at the net	x_{28}	p2-Total distance ran during point
x_{14}	p1-Break point conversion ratio	x_{29}	p2-Speed of last serving
x_{15}	p1-Total distance ran during point	x_{30}	Number of shots during the point

Note: p1 stands for player 1's indicators and p2 stands for those of player 2

Table 4: 30 features explanation

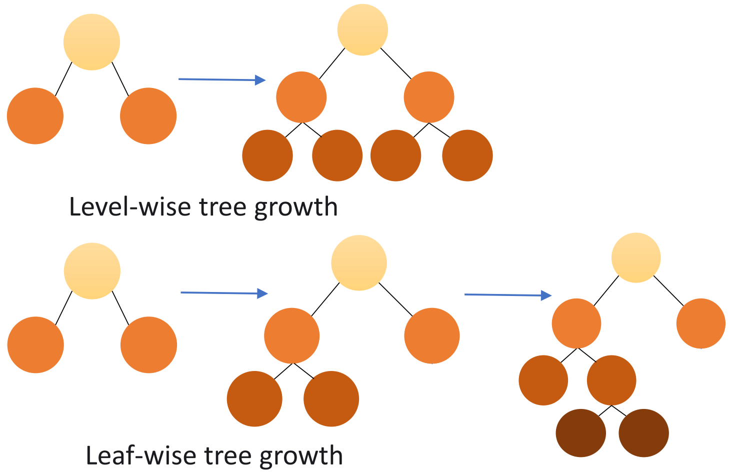

Figure 13: LightGBM

6.3.3 Solution and Result

By applying the LightGBM model, we obtain data of importance of these 30 features. Fig.14 clearly illustrates the importance of each feature.

Among all the 30 features, the top four most important features are:

- **The unforced error ratio of player1 (9.04%):** A possible explanation for its significant importance is that a lower unforced error ratio symbolizes a better overall comprehensive quality, which is more likely to lead to upward turning points in the match.
- **The number of shots during the point (8.90%):** This is probably because when a player gains a points after more number of shots, he is more likely to be encouraged and heads into an upward turning point.
- **The unforced error ratio of player2 (7.65%):** Opponent's unforced errors can also contribute to the occurrence of turning points.
- **Leading score in the whole game of player1 (7.14%):** It is obvious that the more points a player gains in the game, the more likely will he meet an upward turning points.

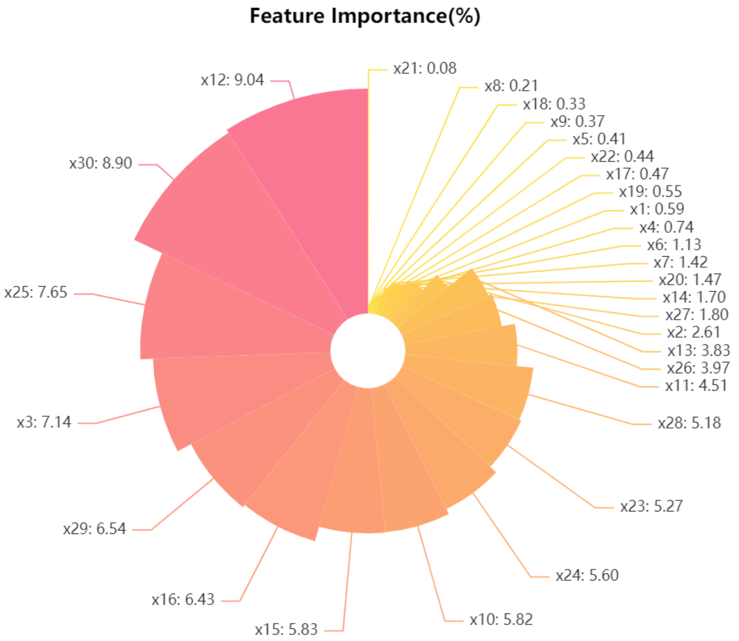

Figure 14: Feature importance

Also, we calculate SHAP value of two classes (up turning points and down turning points, see Fig.15). From these two figures, we can conclude that **scores** indicators contribute most to the occurrence of turning points.

Utilizing the LightGBM model to predict the probability of occurrence of turning points in the 2023 Wimbledon Gentlemen's final, the results are as Fig.16 shows.

From Fig.16 we can see that, the probability curves of both up and down turning points exhibit a high degree of consistency.

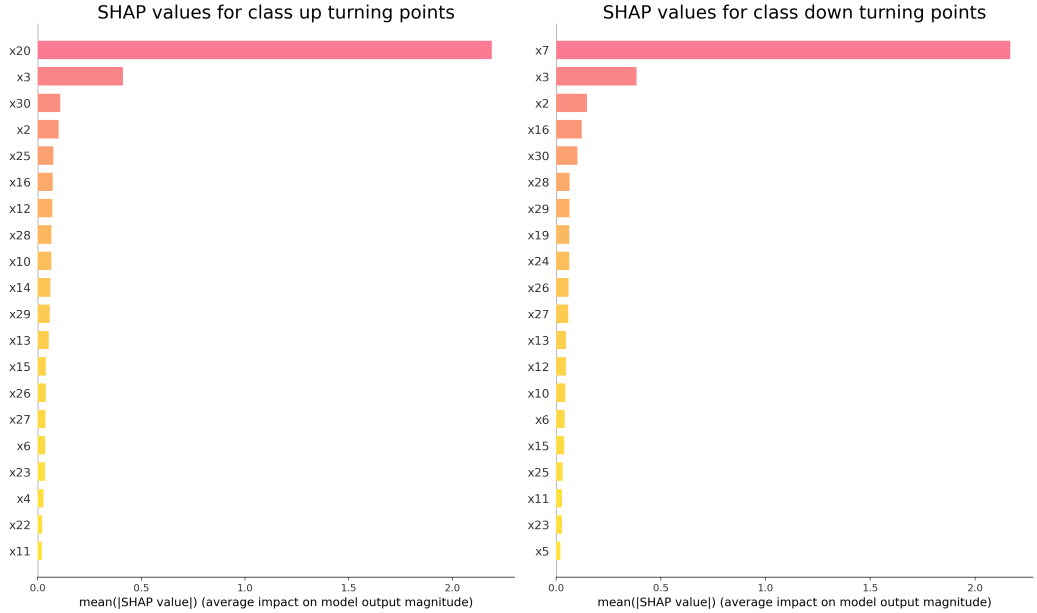

Figure 15: SHAP values of two kinds of turning points

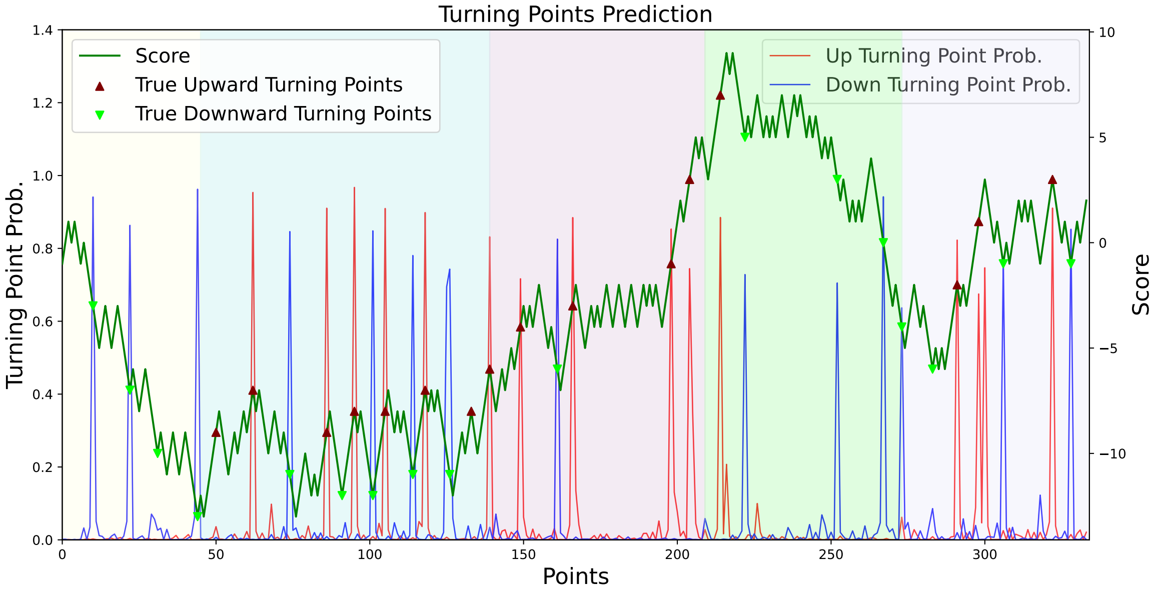

Figure 16: Turning points prediction

6.4 Analysis for a Single Player

Since the feature importance analyzed above may be not applicable to a single player, we calculate the correlation coefficient between a single player's $x_1 \sim x_{30}$ in one particular match and label y. Take Alcarez in match 2023-wimbledon-1701 as an example (See Fig.17). This figure shows that, in this match, x_7 (Whether Alcaraz's previous goal was scored or not) and x_3 (Alcaraz's leading score in the whole game) has the greatest positive correlation with turning points, while x_{20} (Whether Djokovic's previous goal was scored or not) and x_9 (p1-Whether the previous goal was unforced error or not) has

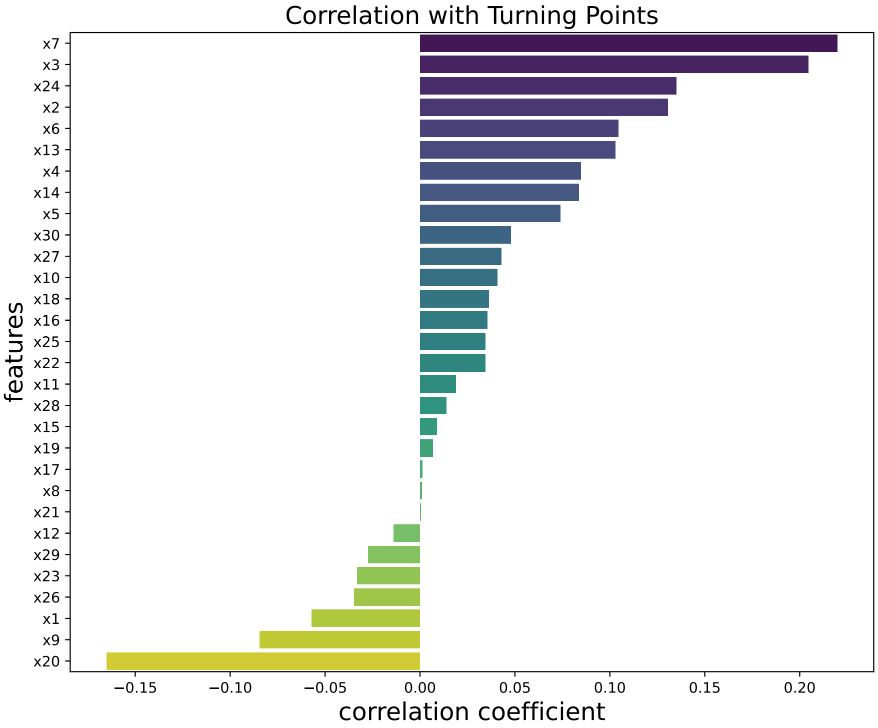

Figure 17: Correlation with Turning Points

the greatest negative correlation with turning points. According to the result, some advise can be given to Carlos Alcaraz to help earn more upward turning points and further control the game in a new match:

- Continuous scoring can boost player's momentum.
- Reducing the impact of opponents' scores may be helpful.
- Reducing player's own unforced error rate.

Also, we can analysis the opponent's data and draw such a correlation figure to find out his weaknesses. Combining these two analysis can be helpful to preparing the next match with a different player.

7 Task 4: Universality of Model

7.1 Extra Data Collection

In order to test the universality of the model or whether the model performs well on other matches, we should find some extra data on women's match.

From the official website of Australian Open[4], we can find detailed data in each match, including how the point is won, serving information and so on. Finally, we select Australian Open Women's Singles Final 2024 as an extra example for in-depth analysis.

We collected the data based on the format of the original dataset except `p1_distance_run` and `p2_distance_run`. We set the two columns to be all zero.

7.2 Model Test on Extra Data

We process the raw data as in Task 1, and use the processed data as input to the model, and we get the result (see Fig.18). The curves can still provide some insights into when turning points may occur, but does not perform as well as in men's matches. This gap is in line with expectations, considering differences among court surfaces and differences between men and women.

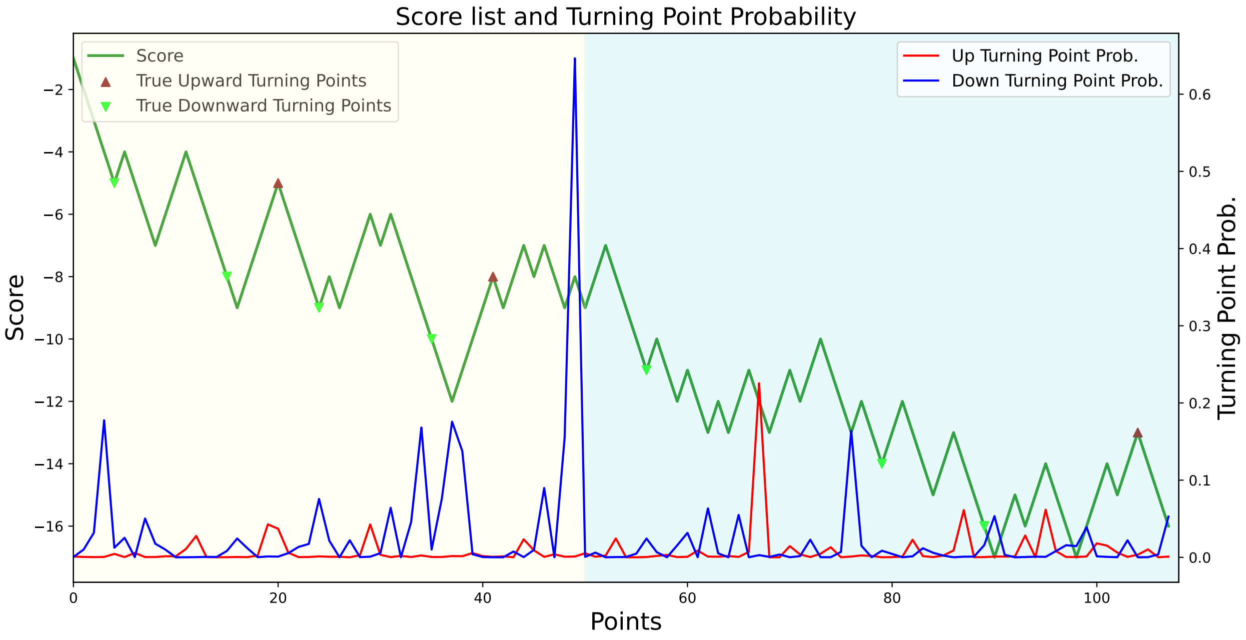

Figure 18: Turning points prediction on Australian Open Women's Singles Final 2024

7.3 Factors to be Considered in Future Models

In terms of **court surfaces**, indicators such as **the height of tennis ball above the ground** and **returning speed** should also be taken into consideration. The court surfaces can be divided into three categories: Grass courts, hard courts and clay courts. As hard courts (Australia Open and US Open) allow for higher ball bounce while grass courts (Wimbledon) allow for low-ball bounces, the height of tennis ball above the ground should be an important feature. Also, the returning speeds of the tennis should be considered as clay courts can cause the tennis ball to move at a lower speed. Taking Nadal as an example, Nadal has a 92% winning rate on red clay while he only has 80% winning rate on hard clay.

When considering **gender gap**, women and men adopt different strategies as men prefer more aggressive attacks while women may prefer more rallies. (This explanation is based on the general performance, we do not deny that certain women tennis player matches male athletes in terms of explosiveness and tennis abilities). Also **formats** of women's and men's matches exist some differences as women single adopt Best of Three format while men adopt Best of Five strategy.

In terms of **other sports**, different sports emphasize different aspects of physical fitness. Comparing table tennis with tennis, the former emphasizes fine control and precision so that indicators like **Spin** are needed when using the prediction model.

Our thought of model is universal and can be applied to other kinds of sports, but the model itself may be not applicable to other sports directly, as certain feature consistent with different scenarios mentioned before needs to be adjusted in the new model for better prediction.

8 Strengths and Weaknesses

8.1 Strengths

- **Feature selections are scientific.** Based the given dataset, we take 17 features into consideration. These indicators are based on the summary data given by official website of The Grand Slam tournaments in tennis, which can reflect players' current mindsets, physical states, and also their overall skill levels.
- **The model contains multidimensional features.** Considering the complex factors influencing the flow of match, we include both psychological and strategic features in our prediction model, which also conforms to the definition of psychological and strategic momentum.
- **Definitions of momentum and turning points are reasonable.** When defining momentum, we prefer momentum as a relative concept when a player has higher momentum then the other player will have lower momentum. Also, we define momentum swings as turning points in Task 2 and use original score lists as the basic data to detect changes of flow in Task 3.
- **The model can predict both up and down turning points.** Instead of only predicting when turning points (or swings in play) may occur, our model can demonstrate the probability of points being up or down turning points.

8.2 Weaknesses

- **Model accuracy can still be improved.**

  For the model developed in Task 1, we calculate the sum of momentum. We predict the outcome of the match simply by assessing the relationship between this value and 0, and then compare the obtained result with the actual outcome. The result is that in total 31 matches, correct predictions account for 54.8%, which has space for improvement.

  For the model developed in Task 3, the high accuracy may be caused by the high proportion of 'non-turning-point's.
- **There still exists gap between models and reality.** In the real world, factors affecting players' performance are far more than number of model features.

References

[1] `https://en.wikipedia.org/wiki/Wimbledon_Championships`

[2] `https://www.youtube.com/watch?v=TDfBzQShh00`

[3] Meier P, Flepp R, Ruedisser M, et al. Separating psychological momentum from strategic momentum: Evidence from men's professional tennis[J]. Journal of economic psychology, 2020, 78: 102269

[4] `https://ausopen.com/`

[5] Fitzpatrick, A., Stone, J.A., Choppin, S. and Kelley, J. (2024), How are elite tennis matches won at Wimbledon? A comparison of close and one-sided contests. Eur J Sport Sci. https://doi.org/10.1002/ejsc.12063

[6] `https://theanalyst.com/eu/2022/03/capturing-momentum-in-tennis/`

[7] Craig Depken; John M. Gandar and Dmitry A. Shapiro, (2023), Set-level strategic and psychological momentum in best-of-five matches in professional tennis, Applied Economics Letters, 30, (6), 735-739

[8] van der Sluis J, Huijgen B, Zaal F, et al. The Effect of Psychological Momentum on Affordances in Tennis[C]//FEPSAC 2022: Sport, exercise and performance psychology: challenges and opportunities in a changing world. 2022.

[9] Vives F, Lázaro J, Guzmán JF, Martínez-Gallego R, Crespo M. Optimizing Sporting Actions Effectiveness: A Machine Learning Approach to Uncover Key Variables in the Men's Professional Doubles Tennis Serve. Applied Sciences. 2023; 13(24):13213.

[10] Elia Morgulev,Success breeds success: Physiological, psychological, and economic perspectives of momentum (hot hand),Asian Journal of Sport and Exercise Psychology,Volume 3, Issue 1,2023,Pages 3-7,ISSN 2667-2391

[11] Jiawang Bai, Yiming Li, Jiawei Li, Xue Yang, Yong Jiang, Shu-Tao Xia,Multinomial random forest,Pattern Recognition,Volume 122,2022,108331,ISSN 0031-3203

[12] Zhang Y, Zhu C, Wang Q. LightGBM-based model for metro passenger volume forecasting[J]. IET Intelligent Transport Systems, 2020, 14(13): 1815-1823.

M E M O

To: Coaches

From: Team # 2410071

Date: 2024/2/6

Re: “Momentum” in tennis match

During a tennis match, “momentum” is a huge factor in determining the flow of play, since player with higher momentum is more likely to feel motivated and encouraged, leading to greater opportunity to carry the edge.

From our perspective, the better the performance of a player compared to the opponent, the higher will his momentum be. After quantifying performance of two players, we can weigh their momentum as the differential of their performance.

You may be skeptical about the role of momentum in the match. In fact, we did a test and the result turns out that the trend of momentum as well as the occurrence of turning points is actually **not random**.

The swing of momentum can to some extent symbolize upward or downward turning points during the game. Then, we developed a model to help monitor and further predict the flow of play. According to our prediction model, some factors play a more vital role in affecting the flow of play.

Taking these important factors into consideration, we put forward some suggestions to help you better prepare players to respond to a potential downward turning points:

✧When a player made consecutive **unforced errors**, the player may request for a timeout to adjust his or her state.

✧When the player missed a point after **multiple shots**, the coach may encourage the player to enhance his or her mood.

✧Train the player's **serving speed** and capability of handling fast serves from the opponent.

✧A successful **Ace** can help promote momentum at a risk of **double default**. Try to serve an Ace when the player has high momentum, but adopt a more conservative strategy under the circumstance of lower momentum.

✧**Running distance** matters in momentum. Coaches should pay attention on physical training so that the player can earn an edge in the second half of the match.

Above are our findings, and hope to help you have a deeper understanding of the role of “momentum” in tennis match. What’s more, “momentum” only serves as a reference to better design strategy. After all, it is the players who are equipped with extraordinary comprehensive quality as well as high momentum that finally win a grand slam.

Yours sincerely,
Team # 2410071

Report on Use of AI

1. OpenAI *ChatGPT* (Jan, 2022 version, ChatGPT-3.5)
   Query1: <*Where can I find tennis match data, such as score data of the Women's Singles Final of the 2023 Australian Open?*>
   Output: < *You can find tennis match data, including score data for the Women's Singles Final of the 2023 Australian Open, from various sources online. Here are some places where you can look for such data:*

   Official Tournament Websites: The official website of the Australian Open (ausopen.com) often provides detailed match data, including scores, statistics, and results for all matches during the tournament.

   Sports Data Providers: There are several sports data providers and companies that offer comprehensive tennis match data. Some well-known providers include Sportradar, Stats Perform, and Tennis Abstract. You may need a subscription to access their data.

   Tennis Websites and Forums: Tennis-specific websites and forums like Tennis.com, TennisForum.com, or Reddit's tennis subreddit may have match data shared by enthusiasts and fans. These sources may not always be official but can still provide useful information.

   Online Tennis Databases: Some websites and databases specialize in tennis match data. Examples include Tennis Explorer (tennisexplorer.com) and Tennis-Data (tennis-data.co.uk).

   API Services: Some API services offer access to sports data, including tennis match data. You can explore APIs provided by sports data companies or organizations.

   Please note that the availability of specific match data may vary, and some sources may require a subscription or payment to access certain data sets. Always ensure that you have the necessary permissions to use and share the data, especially if you plan to use it for commercial or research purposes.>

2. OpenAI *ChatGPT* (Jan, 2022 version, ChatGPT-3.5)
   Query1: < *How to draw an elaborate histogram?*>
   Output: < *Drawing an elaborate histogram involves creating a histogram with various customizations to make it more visually appealing and informative. Here are the steps to draw an elaborate histogram using Python and the Matplotlib library:*

   ```
   import matplotlib.pyplot as plt
   import numpy as np

   # Generate sample data
   data = np.random.randn(1000)

   # Create histogram
   plt.figure(figsize=(10, 6))
   plt.hist(data, bins=20, color='skyblue', edgecolor='black', alpha=0.7)

   # Add labels and title
   plt.xlabel('X-axis Label')
   plt.ylabel('Frequency')
   plt.title('Elaborate Histogram')

   # Add grid lines
   plt.grid(True, linestyle='--', alpha=0.5)

   # Customize the x and y-axis ticks
   ```

```
plt.xticks(np.arange(-3, 4, 1))
plt.yticks(np.arange(0, 120, 20))

# Add a legend
plt.legend(['Data'])

# Customize the plot background
plt.gca().set_facecolor('lightgray')

# Add annotations or text
plt.text(1, 80, 'Mean = {:.2f}'.format(np.mean(data)), fontsize=12, color='red')
plt.text(1, 70, 'Std Dev = {:.2f}'.format(np.std(data)), fontsize=12, color='green')

# Show the plot
plt.show()
```

In this example:

We generate sample data using np.random.randn.
We create the histogram using plt.hist and customize its appearance with various arguments.
Labels, title, grid lines, and ticks are added to improve readability.
A legend is added to indicate what the data represents.
The background color of the plot is changed.
Annotations or text are added to provide additional information.
Finally, plt.show() displays the histogram.
You can further customize the histogram by adjusting colors, fonts, and other parameters according to your preferences.>

3. OpenAI *ChatGPT* (Jan, 2022 version, ChatGPT-3.5)
   Query1: < *How is the format of a memo?>*
   Output: < *A memo, short for memorandum, is a concise and formal written document used for internal communication within an organization. The format of a memo typically follows a standard structure, including the following components:*

   Heading:
   "To:" - The names or titles of the recipients. Include their full names and job titles.
   "From:" - Your name and job title.
   "Date:" - The date when the memo is written.
   "Subject:" - A brief, clear, and specific description of the memo's purpose or topic.
   Example:
   To: All Employees
   From: [Your Name], [Your Title]
   Date: [Date]
   Subject: New Company Policy Updates
   1.Introduction:
   Begin with a concise and clear statement of the purpose of the memo. State why you are writing and what it is about.
   2.Body:
   Provide detailed information, explanations, or instructions related to the subject of the memo.
   Use paragraphs for clarity and readability.
   Use bullet points or numbered lists if applicable.
   Ensure that the content is organized logically.
   3.Conclusion:
   Summarize the key points of the memo.
   Mention any action items or next steps if necessary.

4.Closing:
Express your willingness to answer questions or provide further information.
Include your contact information (email, phone number) if needed.
5.Attachments (if any):
List any attached documents or materials relevant to the memo.
Mention the number of attachments and briefly describe their contents.
6.Signature:
Sign the memo with your name and job title. This is optional in some cases.
7.Distribution List:
Optionally, include a list of individuals or departments who should receive copies of the memo.
Remember that memos should be clear, concise, and professional in tone. They are typically used for conveying important information, announcements, or directives within an organization. >

4. OpenAI *ChatGPT* (Jan, 2022 version, ChatGPT-3.5)
   Query1:< *SyntaxError: Cannot use import statement outside a module*>
   Output: < *To convert an HTML file to a PDF in an environment where tools like Puppeteer can't be directly executed, you can still achieve your goal using Python with libraries such as pdfkit or weasyprint. These libraries are good for handling basic HTML and CSS but might struggle with complex JavaScript or interactive content. Given the limitations of our current environment, I'll guide you through a process using a Python approach, which should work well for static HTML content.*

   Since you've encountered a scenario that suggests you might be dealing with dynamic JavaScript content (given the reference to import statements and modules), and considering the direct execution of JavaScript-rendered content to PDF conversion is not feasible here, I'll proceed with a Python-based solution that is typically used for static content. For fully dynamic content, the conversion process ideally involves rendering the HTML in a browser-like environment first.

   Using pdfkit to Convert HTML to PDF
   pdfkit is a Python wrapper for wkhtmltopdf, which can convert HTML to PDF. It requires wkhtmltopdf to be installed in the system. However, since we cannot install software in this environment, I'll demonstrate how you would use pdfkit if you were to do this on your own system:

   Install wkhtmltopdf:

   On Linux: sudo apt-get install wkhtmltopdf
   On Windows and macOS, download the installer from the wkhtmltopdf website.
   Install pdfkit in Python:

   Run pip install pdfkit.
   Convert HTML to PDF:

   python
   Copy code
   ```
   import pdfkit

   pdfkit.from_file('path/to/rosepro28.html', 'output.pdf')
   ```
   Using weasyprint for Conversion
   Another alternative is weasyprint, which is purely Python and does not require external dependencies like wkhtmltopdf. However, it might not render JavaScript-heavy pages as expected.

   Install weasyprint:

   Run pip install weasyprint.
   Convert HTML to PDF:

python
Copy code

```
from weasyprint import HTML

HTML('path/to/rosepro28.html').write_pdf('output.pdf')
```

Given the current limitations, if your HTML file (rosepro28.html) heavily relies on JavaScript for rendering its content, the most effective conversion would require using a tool like Puppeteer on your local machine, which can accurately capture the dynamic content.

If you want to proceed with a static conversion here, I can demonstrate how to use a Python library to convert your HTML file to PDF, but please be aware that this approach may not perfectly render JavaScript-based content. Let me know how you'd like to proceed!>

5. OpenAI *ChatGPT* (Jan, 2022 version, ChatGPT-3.5)
   Query1:< *how to judge the effect of models of classification>*
   Output: < *Evaluating the effectiveness of classification models is a crucial step in machine learning and data analysis. The choice of evaluation metrics depends on the specific problem you are trying to solve and the nature of your dataset. Here are some common steps and metrics to judge the effect of classification models:*

   1. **Split Your Data**: Start by splitting your dataset into two or three parts: a training set, a validation set, and a test set. The training set is used to train your model, the validation set is used to tune hyperparameters, and the test set is reserved for the final evaluation.

   2. **Select Evaluation Metrics**:
   - **Accuracy**: It measures the ratio of correctly predicted instances to the total number of instances. While it's commonly used, it may not be appropriate for imbalanced datasets.
   - **Precision**: Measures the fraction of true positive predictions out of all positive predictions. It is useful when minimizing false positives is important.
   - **Recall (Sensitivity)**: Measures the fraction of true positive predictions out of all actual positives. It is useful when minimizing false negatives is important.
   - **F1 Score**: Combines precision and recall into a single metric. It's a good choice when you want to balance precision and recall.
   - **Specificity**: Measures the fraction of true negative predictions out of all actual negatives. It is useful for binary classification problems when minimizing false negatives is crucial.
   - **ROC-AUC**: Evaluates the area under the Receiver Operating Characteristic curve, which measures the model's ability to distinguish between positive and negative classes.
   - **Confusion Matrix**: Provides a breakdown of true positives, true negatives, false positives, and false negatives, giving you a more detailed view of your model's performance.

   3. **Train the Model**: Choose an appropriate classification algorithm (e.g., logistic regression, decision trees, random forests, support vector machines, neural networks) and train it using the training data.

   4. **Hyperparameter Tuning**: Use the validation set to fine-tune hyperparameters (e.g., learning rate, regularization strength, tree depth) to optimize your model's performance.

   5. **Evaluate on the Test Set**: Once you've trained and tuned your model, evaluate its performance on the test set using the chosen evaluation metrics. This step simulates how well your model will generalize to new, unseen data.

   6. **Cross-Validation**: In cases where you have limited data, consider using techniques like k-fold cross-validation to assess the model's performance more reliably. Cross-validation provides an average performance metric over multiple splits of the data.

7. **Consider Business or Domain Context**: Sometimes, the choice of metrics and model effectiveness depends on the specific business or domain context. For instance, in medical diagnosis, minimizing false negatives might be more critical than maximizing accuracy.

8. **Visualization**: Visualize your model's performance using ROC curves, precision-recall curves, and confusion matrices to gain a deeper understanding of its behavior.

9. **Compare Different Models**: If you have tried multiple classification algorithms, compare their performance using the same evaluation metrics to select the best-performing model.

10. **Iterate and Improve**: Use the results of your evaluation to iteratively improve your model by adjusting features, hyperparameters, or data preprocessing steps.

Remember that the choice of evaluation metrics and techniques may vary depending on the specific problem and goals of your classification task. It's essential to understand the implications of each metric and select the ones that align with your objectives and constraints.>

6. OpenAI *ChatGPT* (Jan, 2022 version, ChatGPT-3.5)
Query1:< *how to draw a SHAP figure*>
Output: < Creating SHAP (SHapley Additive exPlanations) figures is a valuable way to interpret and visualize the contributions of individual features to model predictions. SHAP values provide insights into how each feature impacts a particular prediction. Here's how to draw a SHAP figure:

1. **Install Necessary Libraries**: Make sure you have the required libraries installed. You will typically need `shap`, `matplotlib`, and the library for your machine learning framework (e.g., `scikit-learn`, `XGBoost`, `LightGBM`) if you're using one.

```bash
pip install shap matplotlib scikit-learn    # Replace with the appropriate library if not using scikit-learn
```

2. **Train a Model**: Train a machine learning model on your dataset. Make sure you have a trained model ready for interpretation.

3. **Generate SHAP Values**: Use the `shap` library to compute SHAP values for a specific instance or a set of instances. The primary function you'll use is `shap.Explainer`.

```python
import shap

# Create a SHAP explainer for your model (replace with your model and data)
explainer = shap.Explainer(model, X)    # 'model' is your trained model, 'X' is your dataset
shap_values = explainer.shap_values(X)    # Compute SHAP values for your data
```

`shap_values` will be a numpy array with SHAP values for each feature and instance in your dataset.

4. **Create a SHAP Plot**: Now, you can create various types of SHAP plots depending on your needs. Two common types are:

- **Summary Plot (Bar Plot)**: To visualize the average impact of each feature across all instances.

```python
shap.summary_plot(shap_values, X, feature_names=X.columns)
```

- **Individual Explanation Plot**: To visualize the SHAP values for a specific prediction for a given instance.

```python
instance_idx = 0    # Replace with the index of the instance you want to explain
shap.initjs()    # Initialize the JavaScript visualization
shap.force_plot(explainer.expected_value[1], shap_values[instance_idx, :], X.iloc[instance_idx, :], feature_names=X.columns)
```

The `expected_value[1]` represents the expected value of the model's output for the positive class, assuming it's a binary classification problem. Adjust this value accordingly for multi-class problems or regression.

5. **Display and Save**: Depending on your environment, the SHAP plots may be displayed directly. You can also save them as image files.

```python
# To save the plot as an image file
plt.savefig("shap_plot.png")
```

6. **Interpret the SHAP Values**: Analyze the SHAP plots to understand the impact of each feature on the model's predictions. Features with positive SHAP values contribute to increasing the prediction, while features with negative SHAP values contribute to decreasing the prediction.

7. **Repeat for Other Predictions**: If needed, you can create SHAP plots for different predictions or instances to gain a comprehensive understanding of your model's behavior.

Remember to replace `model` and `X` with your specific model and dataset. SHAP plots are powerful tools for explaining the behavior of complex machine learning models and understanding the importance of individual features in predictions.>

7. OpenAI *ChatGPT* (Jan, 2022 version, ChatGPT-3.5)
Query1:< *Develop a model that captures the flow of play as points occur and apply it to one or more of the matches. Your model should identify which player is performing better at a given time in the match, as well as how much better they are performing. Provide a visualization based on your model to depict the match flow. Note: in tennis, the player serving has a much higher probability of winning the point/game. You may wish to factor this into your model in some way.>*
Output: < *Developing a model to capture the flow of play in tennis matches involves several steps. We'll need to consider data collection, model development, analysis, and visualization. Here's a structured approach to tackle this problem:*

1. Data Collection

Collect detailed match data including:
- *Point-by-point outcomes.*
- *Server for each point.*
- *Timestamps for each point (if possible, to measure momentum or fatigue).*

2. Preprocessing and Feature Engineering

●*ncode Points* Convert the traditional tennis scoring system (0, 15, 30, 40, game) to a numerical system for easier calculation.
●*erver Advantage* Mark points with who is serving, as the server typically has a higher chance of winning the point.
●*unning Score* Maintain a running total of points for each player.
●*omentum Indicators* Calculate metrics like "points won in the last 10 points" or "break points saved/conceded".

3. Model Development

Momentum-Based Model

1. Point Importance: Assign a weight to each point based on its importance (e.g., break points might be more crucial than others).
2. Serve Bias Adjustment: Adjust the point importance based on the server (e.g., points won on serve might be slightly less valued than points won on return).
3. Moving Averages: Use a moving average (e.g., over 5 points) for each player's performance to smooth out the performance metric and capture momentum.
4. Performance Differential: Calculate the performance differential between players at each point in the match.

Probabilistic Model (Optional)

1. Win Probability: Model the probability of each player winning the match from any given state (scoreline).
2. Elo or Glicko Ratings: Incorporate player ratings like Elo or Glicko to account for player strength.

4. Analysis

●*Critical Junctures* Identify key moments in the match where the momentum or the probability of winning shifts significantly.
●*Serve Impact* Analyze how serving affects the match flow and the performance differential.
●*Player Performance* Summarize how each player performed, highlighting moments when they were underperforming or outperforming their average.

5. Visualization

Create a dynamic visualization to represent the match flow. A line graph is an intuitive choice:

●*X-axis* Time or point number.
●*Y-axis* Performance differential metric.
●*Lines* Two lines representing the moving average of the performance for each player.
●*Shading/Coloring* Use different colors or shading to indicate who is serving. You could also use color intensity to represent the performance differential magnitude.
●*Annotations* Mark critical junctures like break points, set points, and match points with annotations.

This approach provides a structured way to analyze and visualize the flow of a tennis match, taking into account the inherent advantage of serving and the momentum swings that are crucial in understanding match dynamics.>

8. OpenAI *ChatGPT* (Jan, 2022 version, ChatGPT-3.5)
   Query1:< *B: Body, BC: Body/Center, BW: Body/Wide, C: Center, W: Wide*

what do these mean>
Output: < *The abbreviations you've provided appear to represent different categories or labels related to some context. Based on the abbreviations you've given, it seems like they might be related to seating or positions within a seating arrangement. Here's a possible interpretation:*

*- **B: Body**: This could refer to the central or main body of a seating arrangement or object.*

*- **BC: Body/Center**: This might indicate a specific position within the central or middle area of the body.*

*- **BW: Body/Wide**: This could represent a position that is wider or situated towards the sides of the body.*

*- **C: Center**: This likely denotes the exact center position within the seating arrangement or object.*

*- **W: Wide**: This may signify a position that is wider or located towards the outer edges of the seating arrangement or object.*

Please note that the exact meaning of these abbreviations can vary depending on the specific context in which they are used. If you have a specific application or context in mind where these abbreviations are used, providing additional information or context would help in giving a more precise interpretation.>

9. OpenAI *ChatGPT* (Jan, 2022 version, ChatGPT-3.5)
Query1:< *how to assess the effect of features in the model>*
Output: < *Assessing the effect of features in a machine learning model is crucial for understanding the importance of each feature in making predictions. There are several techniques and methods you can use to assess feature importance. Here are some common approaches:*

*1. **Feature Importance Scores**:*
*- **Tree-based Models**: Algorithms like Decision Trees, Random Forests, and Gradient Boosting Trees provide built-in feature importance scores. You can access these scores after training the model.*
*- **Permutation Importance**: This technique involves permuting (randomly shuffling) the values of a single feature and measuring the impact on the model's performance (e.g., accuracy, F1 score, or any relevant metric). Features that have a significant impact when permuted are considered important.*
*- **Feature Importance from Linear Models**: In linear models like Linear Regression or Logistic Regression, the magnitude of the coefficients indicates the feature's importance. Larger absolute coefficients suggest more significant impacts.*

*2. **SHAP Values** (SHapley Additive exPlanations):*
- SHAP values provide a comprehensive way to explain the output of any machine learning model. They show the contribution of each feature to the prediction for a specific instance.
- SHAP values can be visualized using various plots, such as summary plots, force plots, and dependency plots, to understand how features affect predictions globally and locally.

*3. **Correlation Analysis**:*
- Calculate the correlation between each feature and the target variable. Features with higher absolute correlation coefficients are often more important.
- Visualize the correlations using heatmaps or scatterplots to identify linear relationships.

*4. **Feature Selection**:*
- Use feature selection techniques like Recursive Feature Elimination (RFE), SelectKBest, or L1 regularization to automatically select the most important features for your model based on a specific criterion (e.g., mutual information, F-statistic, L1 regularization coefficient).

*5. **Domain Knowledge and Expert Insights**:*
- Consult domain experts or subject matter experts to gain insights into which features are likely to be the

most important based on their knowledge and intuition.

6. **Principal Component Analysis (PCA)**:
  - _Apply PCA to reduce the dimensionality of your dataset. Features with higher loadings on principal components can be considered more influential._

7. **Feature Importance with Neural Networks**:
  - _For neural networks, you can use techniques like Grad-CAM (Gradient-weighted Class Activation Mapping) to visualize which parts of an image are most relevant to the network's predictions._

8. **Recursive Feature Elimination (RFE)**:
  - _RFE is a method that recursively removes less important features and re-evaluates the model's performance at each step. It helps identify the optimal subset of features._

9. **Mutual Information**:
  - _Calculate mutual information between each feature and the target variable to measure their dependence. Higher mutual information indicates higher importance._

10. **Ensemble Methods**:
  - _Use ensemble methods like XGBoost, LightGBM, or AdaBoost that inherently provide feature importance scores based on how often features are used in the ensemble._

11. **Feature Importance Stability**:
  - _Assess the stability of feature importance rankings by bootstrapping your dataset or using cross-validation._

It's essential to consider multiple methods and cross-reference their results to get a robust understanding of feature importance. The choice of technique may vary depending on your dataset, model type, and problem domain.>