Problem Chosen	2024 MCM/ICM Summary Sheet	Team Control Number
C		2418449

Entering the Arena: Unraveling the 'Momentum' Puzzle

summary

In the men's final of the 2023 Wimbledon Open, 20-year-old Spanish star Carlos Alcaraz defeated 36-year-old Novak Djokovic. In this match, the tide turned as first Novak Djokovic won easily and then Carlos Alcaraz took control of the match. The incredible volatility of these matches is often attributed to "momentum."

First, in order to analyze the performance of athletes in a specific time period. We chose the **evaluation model based on entropy weighting method** to find the weights of each index (see Table 3 for details). By analyzing the change of the athlete's comprehensive score in the course of the game, it was found to be consistent with the athlete's actual victory and defeat. This indicates that the evaluation results of the model are more accurate.

Secondly, we established a **correlation analysis model based on gray correlation** degree. It was found that the gray correlation between the athlete's "momentum" and the actual winning and losing situations reached **0.789**. This indicates that the athlete's winning and losing situations are highly correlated with the athlete's "momentum" level in the current time period. Momentum" is highly correlated with the athlete's current time period.

Thirdly, we use the data from the "2023-Wimbledon-1301" match to build a **GSRF model** for training. The optimal hyperparameters obtained from the grid search are max feature="log2" and min sample leaf=3 (see Table 5). The importance of each feature obtained from the random forest training was utilized to make targeted race recommendations for the athletes. It was found that the fluctuations in the matches occurred at **25, 125 and 225 minutes** after the start of the match.

Fourth, in order to evaluate the predictive ability of the GSRF model. We chose the match "2023-Wimbledon-1701" to predict the winners and losers of the athletes. The **accuracy was calculated to be 0.895** with an **AUC of 0.95**. To assess the generalization ability of the model, we found data from women's tennis matches to introduce into the model prediction. The a**ccuracy was calculated to be 0.47** with an **AUC of 0.5**. This shows that the model trained on men's match data is not suitable for generalization to women's matches. Subsequently, the GSRF model was retrained using data from the women's game and the differences in feature importance between the men's and women's games were analyzed. The results of the analysis showed that the women's matches had weaker serve receive and lower serve error rates than the men's matches.

Finally, a memo was written to the coach. The results of the data analysis were summarized and recommendations were made to the coach based on how the athlete could respond appropriately to events that affect the outcome of the game based on the role of momentum during the game.

Keywords: Entropy Weight；Grey Relational Analysis；Grid Search；Random Forest

Contents

1 Introduction

1.1 Problem Background

There doesn't seem to be anyone in the tennis world who doesn't know about Novak Djokovic, the all-time great Grand Slam player who has gone undefeated at Wimbledon since 2013， In the 2023 Wimbledon Gentlemen’s final, 20-year-old Spanish rising star Carlos Alcaraz defeated 36-year-old Novak Djokovic。

The match itself was a remarkable battle. Djokovic seemed destined to win easily as he dominated the first set 6 – 1 (winning 6 of 7 games). The second set, however, was tense and finally won by Alcarez in a tie-breaker 7 – 6. The third set was the reverse of the first, Alcaraz winning handily 6 – 1. The young Spaniard seemed in total control as the fourth set started, but somehow the match again changed course with Djokovic taking complete control to win the set 6– 3. The fifth and final set started with Djokovic carrying the edge from the fourth set, but again a change of direction occurred and Alcaraz gained control and the victory 6 – 4.

In fact, there is a deep mathematical mechanism behind this extraordinary battle.The incredible swings, sometimes for many points or even games, that occurred in the player who seemed to have the advantage are often attributed to “momentum.”

To help coaches identify when a game is about to shift from one player to anothe through metrics such as "momentum," we want to build a predictive model using data from all men's matches after the first two matches of the 2023 Wimbledon Open, and extend it to other sports!

Figure 1:2023 Wimbledon men's final overview

1.2 Restatement of the Problem

Considering the context of the problem and the relevant evaluation indicators, we need to address the following questions：

- **Task 1:**Build a model that captures the flow of the situation as the game is being played and apply it to the game. In addition, it is necessary to determine the performance of a player at a given point in the game and to plot the game's movements.。
- **Task 2:**Develop a model that evaluates the claim that “swings in play and runs of success by one player are random”
- **Task 3:**Develop a model that predicts these swings in the match and to explore the most relevant factors that make a difference in the game. In addition，Given the differential in

past match “momentum” swings， we need to advise a player going into a new match against a different player。

- **Task 4:**The developed model is tested with data from multiple matches and the accuracy of the model in predicting ” momentum” changes during the match is explored. This leads to a discussion of the need to factor in future models. In addition, the generalizability of the model to other games and sports needs to be determined。

1.3 Our Work

Figure 2 shows our modeling framework：

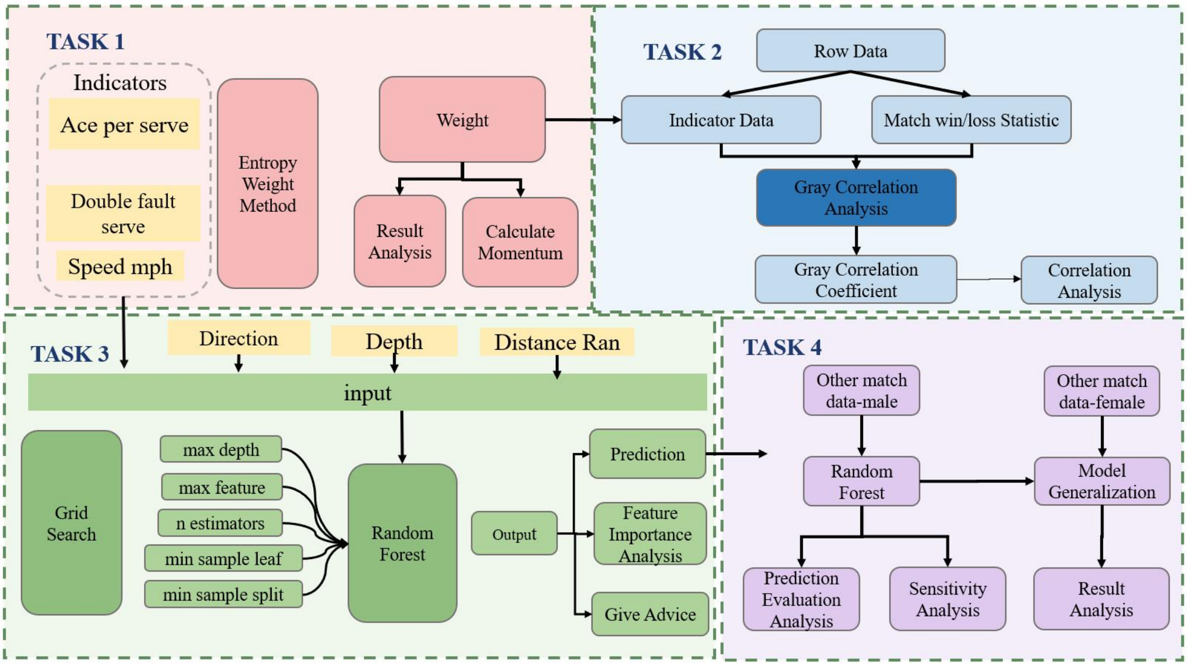


Figure 2:our modeling framework

2 Assumptions and Notations

2.1 Assumptions and Justification

We made several general assumptions to simplify the model. These assumptions, along with their corresponding justifications, are as follows:

Assumption 1: It is assumed that a player's scoring performance is only related to his/her own conditions and the relevant data factors provided by Wimbledon_featured_matches.csv, but not to other external disturbances such as weather conditions, field conditions, external stimuli, etc.

Reason 1: This assumption mitigates the impact of external uncertainty on the most relevant factors for forecasting and improves forecast accuracy.

Assumption 2: Assuming that the competition process is fair and equitable and that the data provided for the competition is true and reliable.

Reason 2: To ensure that the model we built based on the dataset reliably evaluates the degree of player performance and performance advantage.

Assumption 3: It is assumed that the match data profile provided for the 2023 Wimbledon

Men's Singles Final will not have changed by much from previous years.

Reason 3: There is the use of the relative stability of the embodied data to enhance the generalization and practicality of the established model, which is conducive to the generalization of the model.

2.2 Notations

The key mathematical notations used in this paper are listed in Table 1.

Symbol	Description
x_i	Factors for player evaluation
z_i	Indicators for assessing player performance
s_i	Composite score ratio of players
scores	Scores increment within that time period
games	The number of points during that time period
AUC	Performance metrics for classifiers
TPR	Proportion of correct model identifications
FPR	Proportion of model mispredictions

Table 1: Notations used in this paper

3 Data Preprocessing

We are using the official COMAP "Wimbledon_featured_matches" dataset, so we need to preprocess the dataset before solving the problem. An initial examination of the dataset shows that there are some outliers. For example, by looking at the data initially, we see that speed_mph has a numeric value as well as an alphanumeric value NA, which can be interpreted to mean that speed_mph takes the value NA when the server misses the ball, so NA can be replaced with 0.

3.1 Recognition of outliers

Outliers are detected and identified using the interquartile distance method. Points that exceed the upper quartile + 1.5 times the IQR distance or are below the lower quartile -1.5 times the IQR distance are considered outliers. The outliers are identified by observing the given data information about the values in p1_score, p2_score, p1_distance_run, p2_distance_run and speed_mph. We have implemented it with the help of Matlab and the result is shown in Figure 3:

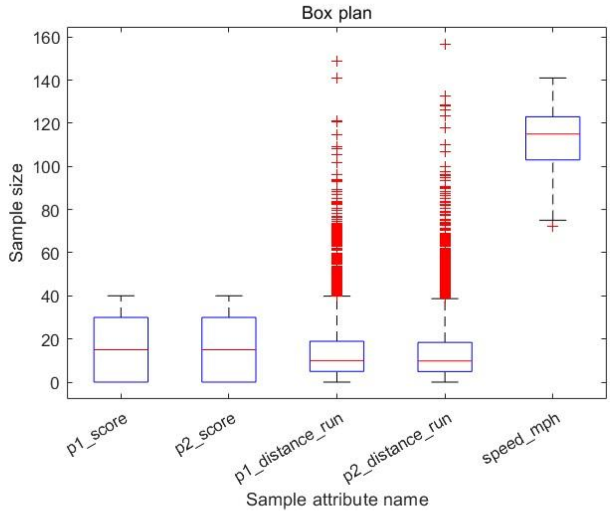


Figure 3 Indicator box plot

According to the game scoring rules, there are several kinds of scores in a game: 0,15,30,40,AD. In addition, the "Wimbledon_featured_matches" p1_score and p2_score have "alternative data" 1,2,3,4,5,6,7,8,9 and other abnormal data.

Analyzing the boxplot, we can find that p1_distance_run and p2_distance_run have a lot of outliers. But p1_distance_run, p2_distance_run will change as rally_count changes. The larger the rally_count of the game, the larger the distance the player moves, so we don't consider the outliers in it. And there are outliers in the data values of speed_mph that are not allowed by the box plot, so we filter them.

3.2 Handling of abnormal data

There are many ways to deal with outliers, and here we consider outliers as missing values and utilize the method of missing value processing for processing correction. Since this approach can utilize the data information in a more reasonable way, we fill the missing values using the median. The results are shown in Table 2 below:

Processed data	Original data	Data after cleaning
speed_mp	72	112

Table 2:Handling of outliers

4 Model 1:Evaluation Model Based on Entropy Weight Method

4.1 Problem analysis

4.1.1 Define the evaluation factors of players

Combining the information provided in reference [1] and the title. We can find that the probability of scoring on serve is much higher in tennis matches. Therefore, we consider to include the information related to serve into the evaluation indexes, such as the proportion of serving points, the proportion of serving errors, the proportion of our points when the opponent serves, and so on.

Instead of using the given data directly, we used the data obtained after the correlation calculation as the input data of the model. Players' performance is also related to the psychological pressure caused by the score gap in the match, the type of opponent's serve (e.g., opponent's serve errors, untouchable winners, etc.), and other factors.

Considering all these factors, we choose the factor xi as an indicator to evaluate the performance of a player in a certain period of time (5-minute intervals are chosen as the time unit), and the definition of the factor is shown in Table 3:

The values of i	**Description**
1	Scoring efficiency per unit time
2	Server-to-winner ratio per unit time
3	Ace percentage per unit time
4	Untouchable winning shot ratio per unit time.
5	Missed serve and lost point ratio per unit time
6	Unforced error rate per unit time
7	The ratio of players makes it to the net per unit of time
8	Ratio of net reach to player's score per unit time in the current game
9	Winning ratio when the opponent serves
10	Average serving speed of the player

Table 3:Definition of x_i

In addition, for the convenience of modeling as well as analysis, the ten evaluation factors can be classified into three categories based on their degree of similarity. The categorization table of the factors drawn is shown in Figure 4:

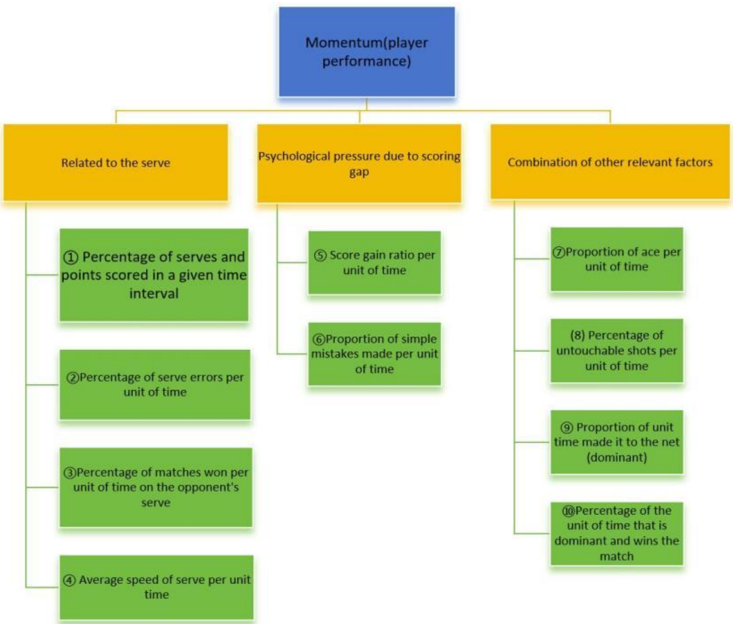


Figure 4: Categorization of evaluation factors

The factors affecting "momentum" are selected as xi, and their weights are calculated using the entropy weighting method. Momentum is quantified by a composite score; the higher the momentum value, the better the player's performance.

4.1.2 Data analysis

As an example, the data 2023-wimbledon-1301 was selected, player1 was Carlos Alcaraz, and player2 was Nicolas Jarry. evaluation metrics for that time period were selected for processing at five-minute intervals from the start of the match.

The score gain ratio per unit time can be expressed as:

$$x_1 = \frac{scores}{games} \tag{1}$$

4.2 Construction of evaluation model

Entropy weight method is an objective weighting method, which can obtain relevant weights according to the data itself. According to the principle: the greater the degree of variation of the index, the lower the probability of event occurrence, the higher the information entropy (uncertainty of event occurrence), and the higher the weight should be given.

1.Data standardization

Indicator normalization can be written as the following formula:

$$x_{ij}^{'} = \frac{|X_{ij} - \min(X_{1j}, X_{2j}, ..., X_{nj})|}{\max(X_{1j}, X_{2j}, ..., X_{nj}) - \min(X_{1j}, X_{2j}, ..., X_{nj})} \tag{2}$$

2.Find the ratio of each index under each scheme

There are n objects to be evaluated i.e. momentum y in various time intervals and m evaluation indicators i.e. xj (j=1,2,3 .... ,10) Calculate the weight of the ith object under the jth indicator and consider it as the probability used in the relative entropy calculation.

$$\mathrm{y}_{ij} = \frac{x_{ij}^{'}}{\sum_{i=1}^{m} x_{ij}^{'}} \left(i = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, m \right) \tag{3}$$

3.Calculate composite score

$$\begin{cases} e_j = -\dfrac{1}{\ln m}\sum_{i=1}^{m} y_{ij}\ln y_{ij} \\ w_j = \dfrac{1-e_j}{m-\sum_{j=1}^{m} e_j} \\ s_i = \sum_{j=1}^{n} y_{ij} w_{ij} \end{cases} \tag{4}$$

Where e_j is the information entropy, j refers to the jth indicator; w_j is the weight of the jth indicator.And the weight is an objective representation of the importance of the indicator so that we can evaluate the degree of player performance.

Finally, we calculate the comprehensive score s_i of the ith player, and the comprehensive score is the indicator for evaluating the degree of player performance. The results of the weights of the relevant indicators are organized in Table 4.

index of correlation	**Income weight**
x_1	0.011644
x_2	0.031414
x_3	0.162098
x_4	0.056018
x_5	0.225218
x_6	0.061388
x_7	0.079291
x_8	0.093800
x_9	0.259166
x_{10}	0.019962

Table 4:Results of entropy weight method

4.3 Analysis of model results

We apply python to the solution, which allows us to calculate the weights of the indicators xi needed for the entropy weighting method and to calculate the composite score of each player; the higher the player's composite score, the better the performance.

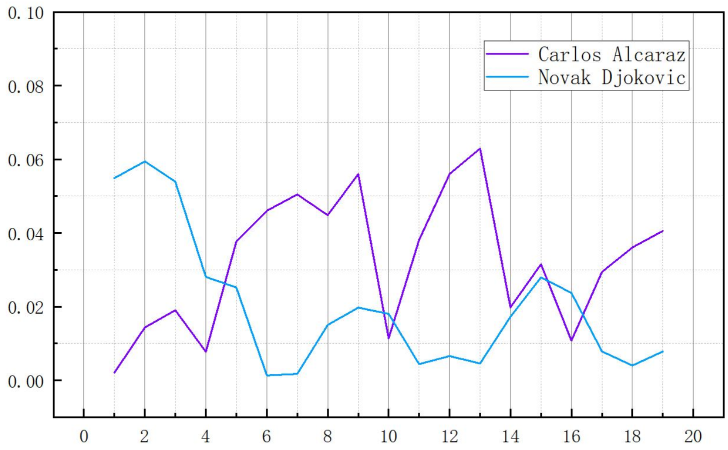


Figure 5: Combined scores of the two players

Taking the 2023 Wimbledon men's final between Carlos Alcaraz and Nicolas Jarry as an example, the comprehensive scoring curve drawn is shown in Figure 5.

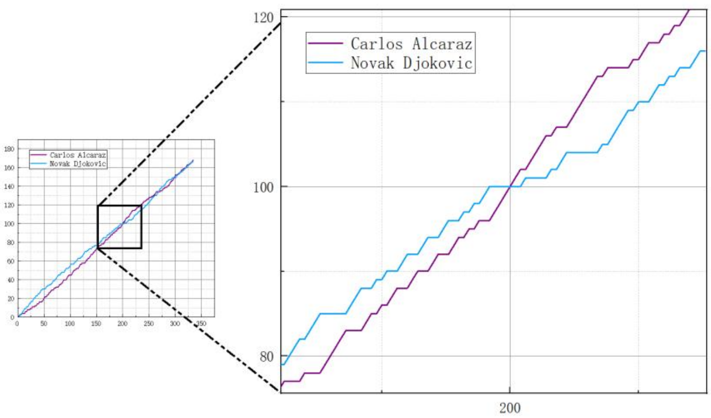


Figure 6: Scoring of Carlos Alcaraz and Nicolas Jarry

Analyzing the combined score graph, it can be seen that Carlos Alcaraz's combined score is higher than Novak Djokovic's in the time period between 5-15 (1:00:04-3:45:09), which shows that the player is gradually catching up with the score.

From the figure of the scoring situation of the match, it can be seen that before the 200th round (point), the increase rate of Carlos Alcaraz's score is higher than that of Novak Djokovic, which coincides with the fact that the player's combined score is higher than that of his opponent's at this time, and in the 200th round (point), which is located in the turning point of the score catching up, Carlos Alcaraz succeeded in catching up with the score. score, at which point the time interval was 2:54:24.

From the comparison analysis of the two graphs, it can be seen that the performance of the players coincides with the overall score, and the evaluation model is more reliable.

Figure 6 gives a clear picture of the performance of the two players in a given time period. The model is more reliable as it agrees with the performance of Carlos Alcaraz and Novak Djokovic given in the title.

In addition, the model can further reflect the magnitude of the player performance advantage and game trends.

We do this by considering z_i, the percentage of combined player scores. This is used as a metric to assess performance advantage and match trend, defined as:

$$Z_i = \frac{s_i}{\sum_{j=1}^{j} s_i} (i = 1, 2; m = 2) \tag{5}$$

The plotted percentage of combined scores is shown in Figure 7. Analyzing Figure 7, it can be concluded that when a player is more dominant, the percentage of the corresponding area is greater than 50%. The larger the percentage accounted for, the greater the advantage of that player.

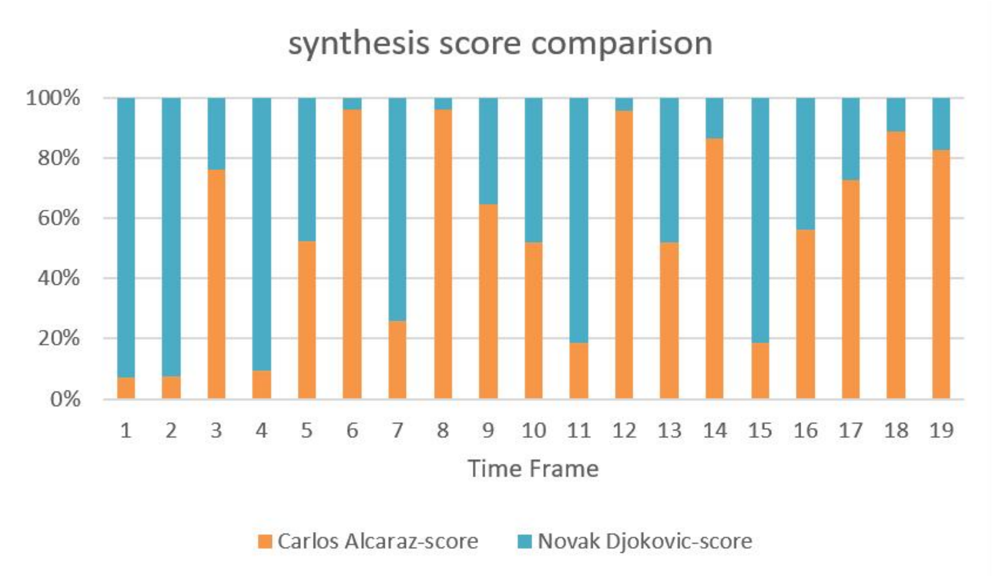


Figure 7:Combined score ratio of the two players

4.4 Uncertainty analysis of impact models

The dataset we have and the relevant metrics selected may not fully reflect the skills and characteristics of the players. This is because there is a great deal of variation among players, which may limit the model's ability to generalize to the entire group of participants.

In addition, the size of the selected dataset, although acceptable, is still relatively limited. It may lead to overfitting or underfitting of the model when dealing with a large amount of unknown data.

5 Model 2：Correlation Analysis Model Based on Grey Relational Analysis Method

5.1 Correlation analysis modeling

5.1.1 Spearman's correlation coefficient

Based on the model established in Task 1, in order to further evaluate the player's performance, we chose another game and calculated the player's score for each time period of the game according to the evaluation model in Task 1. The greater the score, the greater the “momentum” and the greater the probability of winning.

By calculating the percentage of games won by each athlete in each time period. The degree of correlation between momentum scores and game results is assessed through statistical analysis, and determining the spearman correlation coefficient between momentum scores and game results can be transformed into calculating the spearman correlation coefficient between the ratio scores between p1-score/p2-score and the probability of winning the game, which in turn determines whether there is a correlation between momentum and winning and the degree of that correlation。

5.1.2 Gray correlation analysis description

Gray correlation analysis is a statistical method used to assess the relationship between two or more sets of variables, which takes into account the similarity or dissimilarity of trends between the variables and accordingly quantifies the degree of association between them. This method is suitable for analyzing relationships between systems or factors that change over time or under different conditions. Typically it consists of three steps：

Step1: Identification of reference and comparison series： This operation identifies a portion of the array that represents the behavioral characteristics of the system and the different factors that influence those behaviors.

Step2: Dimensionless processing： The different physical significance of the factors in the system may lead to inconsistent data quantiles, which is not conducive to direct comparison. Therefore, dimensionless processing is required for effective comparison and analysis.

Step3: Calculating the gray correlation coefficient： This operation calculates the correlation coefficient, which is measured as the difference between two series at a specific moment in time (point on the curve), and it can be written as follows：

$$\xi(X_i) = (\Delta_{min} + \rho\Delta_{max}) / (\Delta O_{i(k)} + \rho\Delta_{max}) \tag{6}$$

Where $\xi(X_i)(k)$denotes the correlation coefficient between the kth parameter of the ith subsequence and the kth parameter of the parent sequence (i.e., the 0 sequence), ρ is the resolution coefficient, which takes values in the range of [0,1].

5.1.3 Modeling method selection

Since the stability of Spearman's correlation coefficient is related to the sample size, its estimation may not be accurate enough when the sample is small, especially in very small samples that may be affected by extreme values. Gray correlation analysis, on the other hand, remains more robust when the data sample is small, is not easily limited by the sample size, and is applicable to data with different distributions.

In terms of the analysis method, our data treatment is to select time as the independent variable. Although both Spearman's coefficient method and gray correlation analysis can be used to deal with nonlinear relationships. However, gray correlation analysis better reflects the degree of association between the variables, and it is more capable of dealing with nonlinear relationships. Therefore, we choose gray correlation analysis to develop the correlation analysis model.

5.2 Analysis of Model Results

First, the p1-score/p2-score and the probability of winning the game are used as the comparison and reference sequences, respectively, and are dimensionless, and then the gray correlation coefficient is calculated with the help of python according to equation (6)， where ρ is taken as 0.5,the results of the gray correlation analysis are plotted as shown in Figure 8：

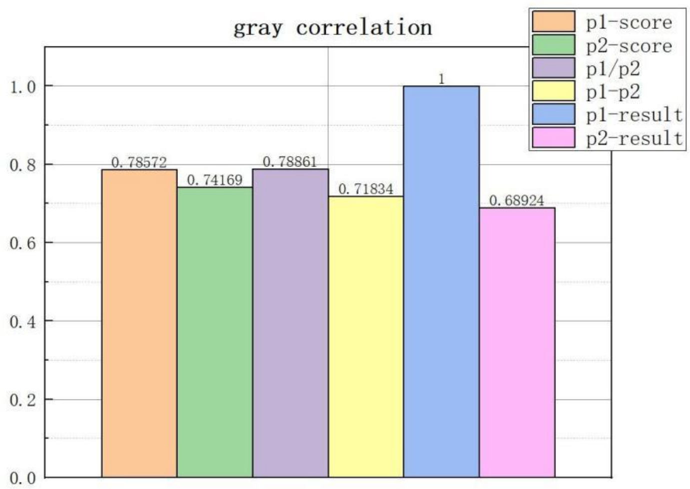


Figure 8:Gray correlation coefficient

The closer the gray correlation coefficient is to 1, the stronger the correlation is, as seen in the above figure, the correlation coefficients of the gray correlation coefficients between p1-score/p2-score and p1-p2 and p1-result are 0.78861 and 0.71834, respectively, which proves that there is a strong correlation between the two and p1-result. Therefore, the statement that “swings in play and runs of success by one player are random” is ill-considered.

6 Model 3：GSRF Prediction Model

6.1 Predictive Modeling

6.1.1 Factor Selection

Based on the selection of factors in Model 1, we have some trade-offs in the selection of factors.

We included the distance moved by the player per unit of time, the percentage of the total number of serves in the direction of the serve per unit of time, and the percentage of the number of points scored from the position of the different serves per unit of time as selection factors。

In addition, during model training, since the number of points scored per unit of time is directly correlated with the winners and losers of the matches, resulting in an insignificant degree of correlation of other features, which is not conducive to our analysis, we chose to remove this selection factor.

6.1.2 Description of the GSRF algorithm

The GSRF algorithm is composed of Random Forest and GridSearchCV.

Random forest is a supervised learning algorithm integrated by multiple decision trees [2] for solving classification and regression problems。It combines multiple decision trees, each of which is trained on a random subset of the training data. This means that each tree is learned based on different samples and features, thus increasing the diversity of the model. In addition, the Random Forest algorithm can handle input samples with high-dimensional features and is highly resistant to interference and overfitting, and can effectively run on large data sets, showing good accuracy [3].

GridSearchCV is a parameter optimization algorithm commonly used to systematically search for combinations of hyperparameters of a model to find the best combination to optimize the model's performance. This tool evaluates the model by performing an exhaustive search on a specified grid of parameters, cross-validating each set of parameters, and calculating performance metrics.

We use GSRF model for classification analysis prediction. Based on the selected relevant factors, a random forest classification model is constructed. The model is trained with the match data, and the dependent variable is the probability of the player's win, so as to predict the win/loss situation of the athlete in a unit of time, and from this, we can derive the degree of correlation between each selected factor and the change of the situation.The algorithmic flow of the GSRF is shown in Figure 9：

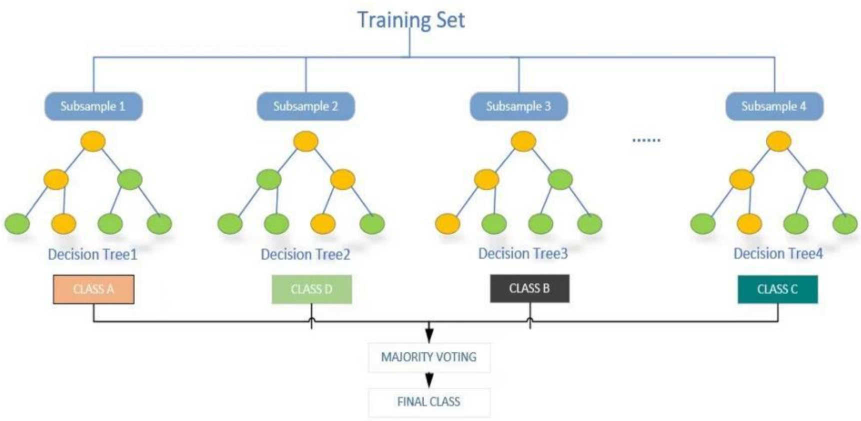


Figure 9:The algorithmic flow of GSRF

6.1.3 GSRF hyperparameter selection

To find the optimal hyperparameters for GSRF, we used the RandomForestClassifier algorithm in the collection module of the scikit-learn (sklearn) machine learning library, and the GridSearchCV algorithm in the sklearn.model_selection module. The optimal hyperparameters are shown in Table 5：

max depth	max feature	min sample leaf	min sample split	n estimators
15	log2	3	10	50

Table 5:The best hyperparameter combination by GridSearchCV

6.2 Analysis of Result

6.2.1 Analysis of Model Result

We trained the GSRF model using player data from the "2023-wimbledon-1301" match, with the aim of selecting the factors that are most relevant to situational change. Then, we use the trained GSRF model to predict the correlation between the selected factors and the situation changes. The plotted correlation level is shown in Figure 10：

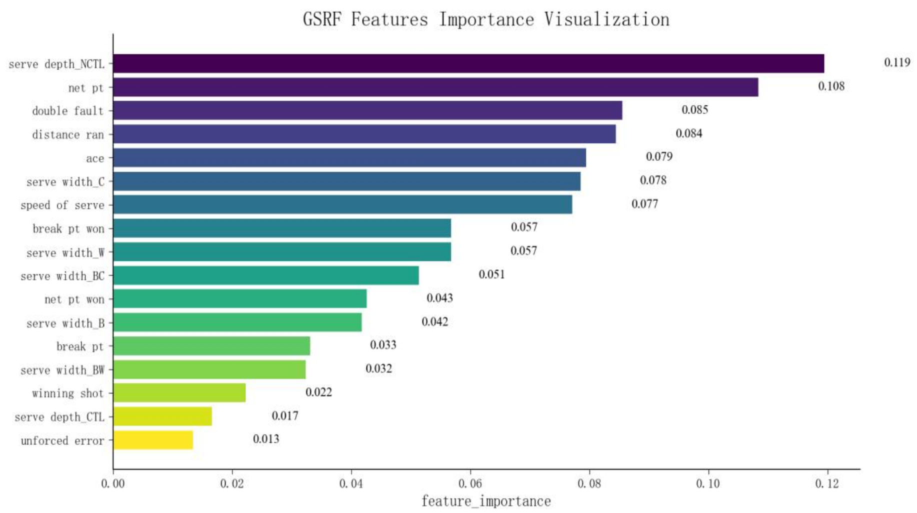


Figure 10:Feature importance

Analyze Figure 10 shows that serve depth_NCTL and net_pt have a relatively high correlation with situational change, at 0.119 and 0.108, respectively, which leads to the conclusion that these two factors are the most relevant factors for situational change,In tennis, a player runs to the net usually to execute a net tackle, allowing the player to hit the ball before it jumps to the ground, which reduces the opponent's reaction time and improves scoring chances.

6.2.2 Prediction evaluation analysis

We used 80% of the player data from the "2023-wimbledon-1301" game as a training set and the remaining 20% as a test set. Subsequently, we plotted the ROC curve as shown in Figure 11, using python as a tool：

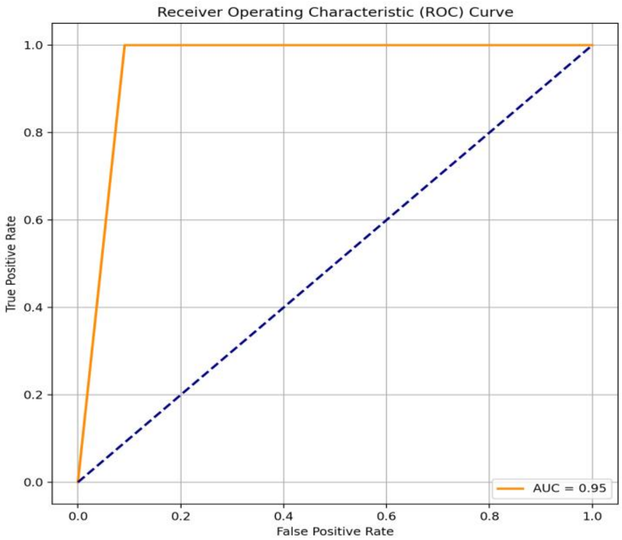


Figure 11:Model evaluation ROC curve

AUC is used to measure classifier performance. When AUC is closer to 1, the classifier performance is better. At this point, the AUC is approximated to 1, indicating that the predictive model is highly accurate.

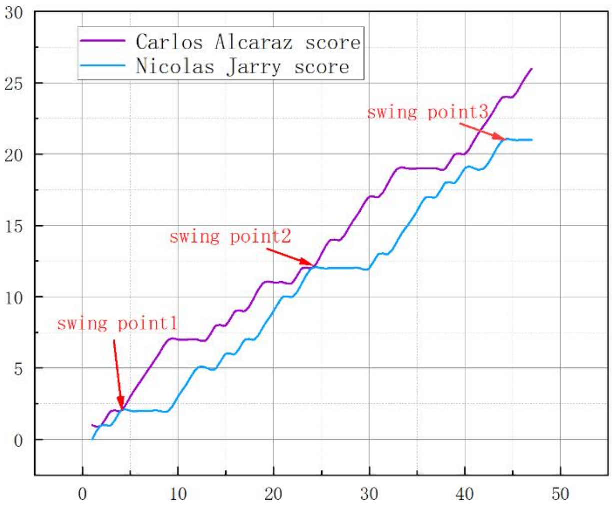


Figure 12:Swing point analysis

We successfully predicted the winner of the match "2023-Wimbledon-1301" using the GSRF model. In order to obtain the location of the swing points during the match, we plotted the cumulative predicted scores of the two athletes, as shown in Figure 12 above.

In the figure above, the locations of the swing points all appear to be located at points where the trend in scoring for a particular athlete is significantly increasing (i.e., the slope of the curve is increasing). For example, at swing point 1, Carlos Alcaraz's scoring trend is significantly higher than Nicolas Gary's, which suggests that this starts at time unit 5 (around 25 minutes into the match). At that point Carlos Alcaraz was ready to start making a push to try and extend his lead.

It can also be concluded from the graph that Carlos Alcaraz was at a higher level of scoring than Nicolas Jarry and ended up winning the final game as well.

6.3 Suggestions for players

First, we ranked the selected factors according to their relevance to the evolution of the game situation and considered the top five factors in terms of relevance to give recommendations. The top five factors are serve depth_NCTL, net pt, double fault, distance ran and ace and the following recommendations are made：

Suggestion 1： When serving, you can try to stay away from the line as far as possible, the farther away from the net, the greater the range of control, when the opponent player counterattacks, you can leave yourself room to hit back. When receiving the ball, you can get closer to the net, so you can move in and out freely and hit the ball relatively well.。

Suggestion 2： Try to hit topspin during your game, this is because topspin increases the arc of the ball and therefore the range. This increases the distance you can run around the court and the topspin is more challenging and difficult for your opponent to deal with.

Suggestion 3： Focus on serving, in the first serve you can choose according to the opponent's state, while the second serve you should choose your own most familiar way to serve to improve their own fault tolerance rate.

7 Generalization of the model

7.1 Accuracy analysis of predictions

When assessing binary classification models for prediction, ROC curves and confusion matrices are commonly employed to evaluate model performance. Accordingly, we utilized them to assess player data from the "2023-wimbledon-1701" match.

The ROC curve is plotted with TPR on the vertical axis and FPR on the horizontal axis. Each point on the curve corresponds to the model's performance at different thresholds. The closer the curve is to the upper-left corner, the better the model's performance.。

TPR is the ability of the classifier to correctly identify positive examples, the closer the TPR is to 1, the better the model performance is. the formula for TPR can be expressed as:

$$TPR = \frac{TP}{TP + FN} \tag{7}$$

where TP denotes the number of correctly recognized positive examples and FN denotes the number of incorrectly recognized positive examples as negative examples.

FPR is the proportion of samples that the model incorrectly predicts to be positive out of all the samples that are actually negative cases.The closer the FPR is to 0, the better the model performs.The formula for FPR can be expressed as:

$$FPR = \frac{FP}{FP + TN} \tag{8}$$

where FP denotes the number of negative instances incorrectly recognized as positive instances and TN denotes the number of negative instances correctly recognized. The ROC curve we plotted using python as a tool is shown in Figure 13:

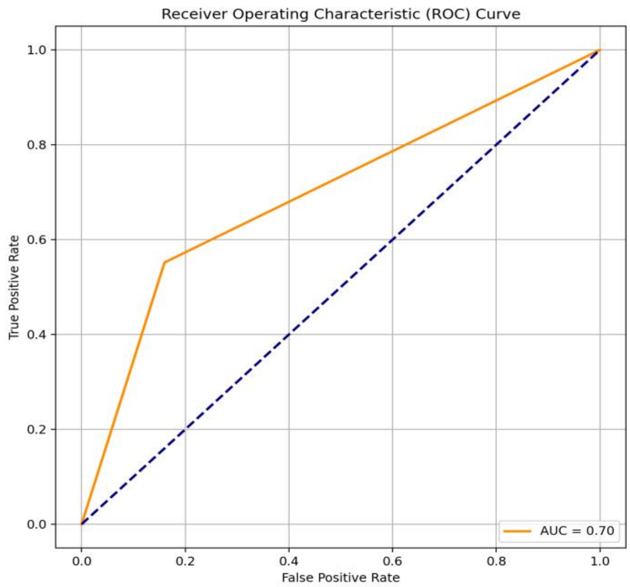


Figure 13:ROC curves for accuracy analysis

By calculating the area under the ROC curve (AUC) and obtaining a value of 0.70.We can infer that the classifier's performance is relatively good, with a high level of accuracy. AUC values approaching 1 indicate strong classifier performance.

In addition, the confusion matrix displays the correspondence between the model's predictions and the true results for different classes. Each row of the confusion matrix represents the actual class, while each column represents the predicted class. Using Python as a tool, the confusion matrix we plotted is shown in Figure 14:

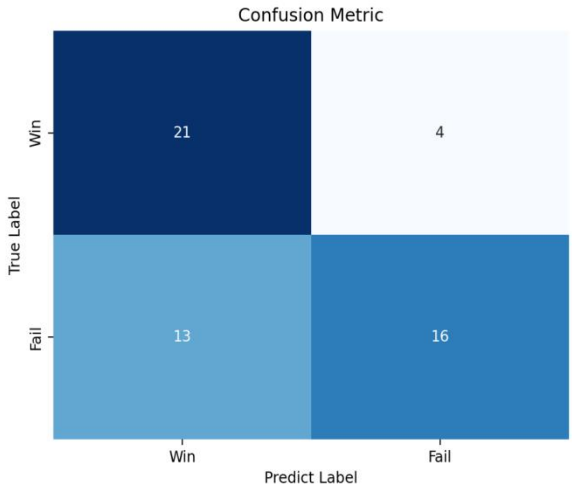


Figure 14:Confusion matrix

The accuracy of the confusion matrix is the proportion of correctly categorized samples to the total number of samples. The formula for the accuracy rate can be expressed as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

The result of the calculations is $Accuracy \approx 0.69$ .Based on the above analysis, we can assume that our forecasting model is comprehensive and accurate.

7.2 The inclusion of future model factors

Research has indicated[6]that due to the complexity and dynamic nature of "human movement," in addition to quantitative metrics during the competition, the relationship between athletes' performance levels and their psychological states is tightly interwoven. It is evident that different psychological states in sports competitions exhibit varying levels of efficacy. Elevated pre-competition anxiety levels, both on an individual and societal expectation level, can diminish athletes' pre-competition state attention levels, accompanied by a heightened sense of self-awareness. Consequently, this impedes the complete immersion of athletes in the competition, hindering the manifestation of optimal psychological state characteristics during the game. The following outlines different conditions that may contribute to the anxiety states experienced by athletes.

1. Participation experience: The ability to self-regulate improves with more competition so that athletes do not become overly nervous during competition. Therefore, athletes with more experience in competitions are more likely to be able to perform their skills more consistently and with greater ease than athletes with less experience in competitions. On the contrary, athletes with less competition experience will be more prone to psychological tension due to their first time or fewer competitions, resulting in muscle stiffness and abnormal movements. As a result, they will not be able to perform their techniques properly.

2. Off-site spectators: In sports competitions, the number of spectators is not the most important aspect. What matters is the message conveyed to athletes through their presence. A larger audience can convey the message to athletes that "many people care about your performance," and this kind of information has a positive effect on athletes.

3. Self-regulation skills: Athletes with strong self-regulation abilities can effectively control the pace of the game during both competitions and training. Rationally adjusting their emotions allows them to establish a positive psychological state for a stable execution of techniques.

7.3 Generalizability Testing of the Model

In the training of the GSRF model, we are based on the data of 2023-wimbledon male athletes. This time, we selected the data of the game "2023-wimbledon-1301" to train the model. Through the evaluation of the model, we can see that the prediction model has a better prediction effect, and can accurately predict the winners and losers of the matches between different male athletes. The model is able to reflect the change of each athlete's "momentum" and the time when the score begins to appear a reversal of the trend.

To evaluate the model's generalization ability, we collected data from the 2023 Wimbledon women's matches and selected the match "2023-wimbledon-2601" featuring Elina Svitolina against Marketa Vondrousova. Firstly, we performed data preprocessing to obtain metrics similar to those used for male athletes. Subsequently, utilizing the aforementioned model, we predicted the outcome of the women's match. The prediction results are depicted in the following Figure:

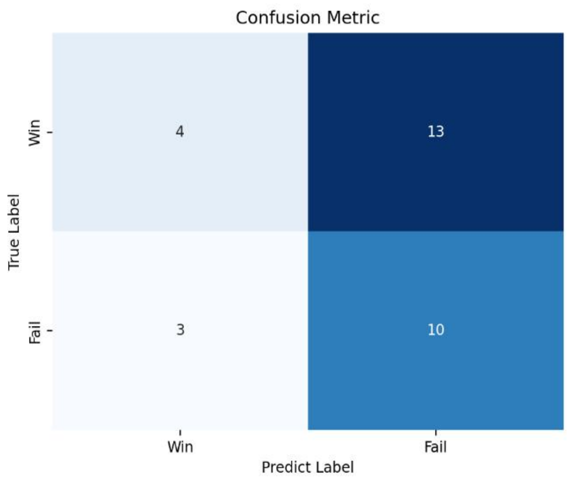


Figure 15: Confusion matrix

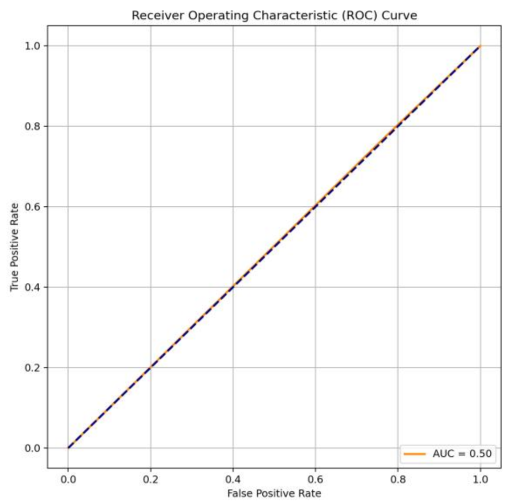


Figure 16:ROC curve for predict

The result obtained by the confusion matrix is Accuracy=0.47.From this, it can be observed that the accuracy is not very high。The result of calculating the area under the ROC curve is AUC=0.5.From this, it can be concluded that the predictive capability of the model did not meet the expected standards.

In summary, the prediction model for match outcomes derived from the training data of male athletes does not translate well to forecasting match results for females. Therefore, we hypothesize that different features have varying impacts on male and female athletes. Consequently, we retrained a Random Forest classification model based on match data from female athletes. Subsequently, we utilized the trained GSRF model to predict the correlation between selected factors and changes in match situations. The correlation plot is depicted as Figure 17.

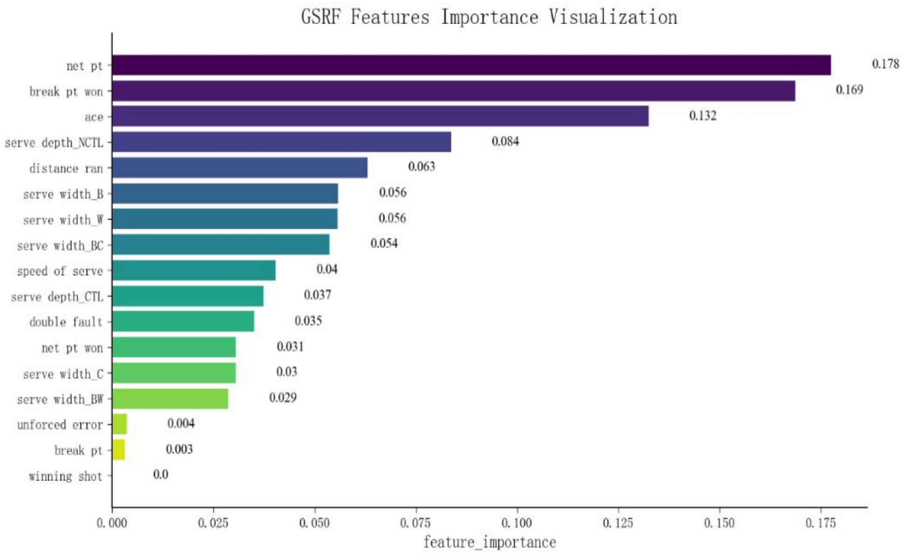


Figure 17: Importance of selected factors

Analyze the degree of correlation between the factors selected for the women's and men's tennis matches and the change in situation. We selected the top three factors in terms of correlation and plotted Table 6:

Female Athlete	Male Athlete	Rank
Net pt	serve depth_NCTL	1
Break pt won	Net pt	2
ace	Double fault	3

Table 6: Top three factors associated with women and men

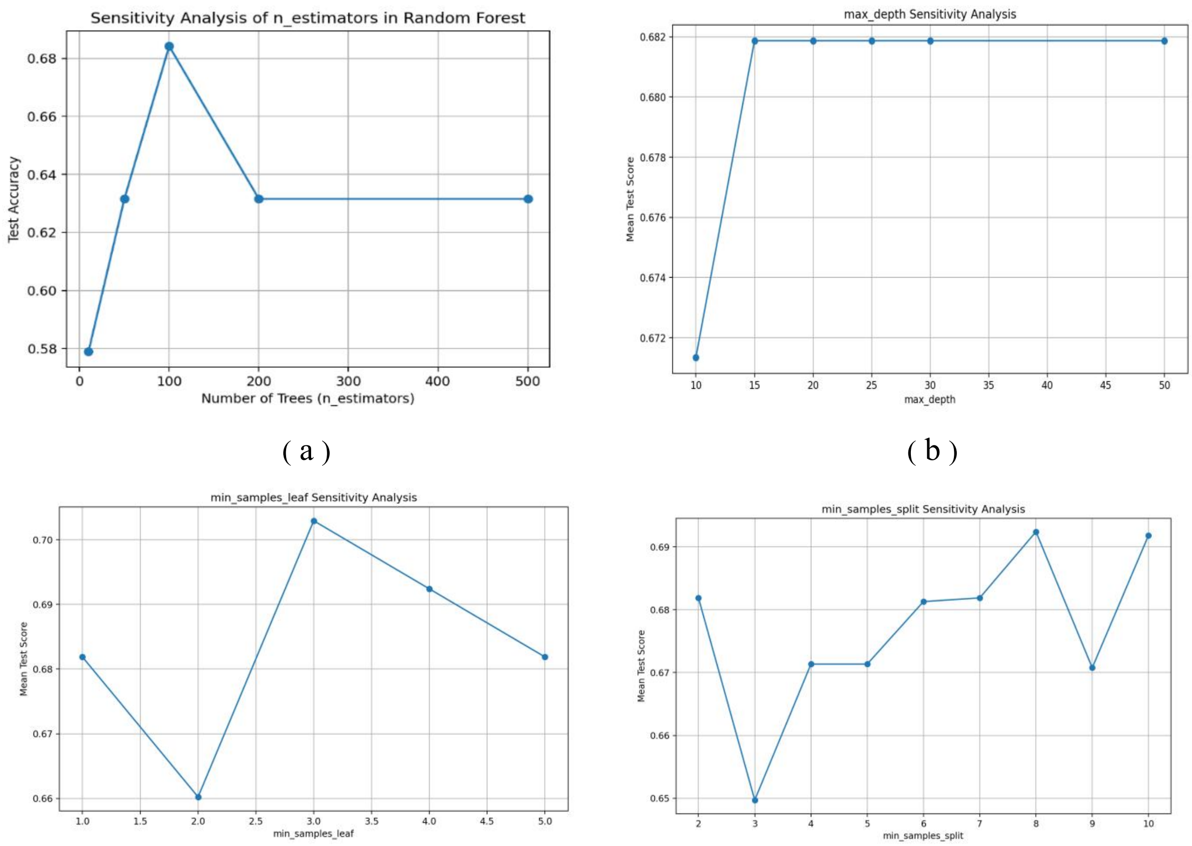
Through the analysis of Table 6,it becomes evident that the 'net pt' feature is of utmost importance for both male and female athletes. This highlights the significance of successfully positioning oneself at the net to execute volleys or smashes is a crucial attacking strategy aimed at capitalizing on the opponent's weaker positioning or slower reaction time to secure points. For female athletes, the 'ace' feature plays a crucial role in winning matches. Conversely, the impact of the 'ace' feature for male athletes is relatively weaker in determining the outcome of the match.

From this observation, it is evident that the ability of female athletes to handle untouchable shots tends to be weaker than that of male athletes, thereby potentially influencing the scoring situations in this scenario. On the other hand, for male athletes, double faults have a significant impact on the outcome of the match, while conversely, this feature is not as pronounced for females. As far as our knowledge extends, the intensity of men's matches tends to be greater than that of women's matches, and according to the data, the serving speed of male players is notably higher than that of female players. This, to some extent, can contribute to an increased error rate in serving, consequently affecting the scoring situations for athletes.

In summary, due to the differences in the intensity of competition between male and female athletes, there are variations in the importance of features influencing the outcomes of competitions for these two groups. Therefore, applying a model trained on male athletes to predict outcomes for female athletes may result in suboptimal predictive performance.

8 Sensitivity Analysis

For the GSRF binary classification model, we conducted sensitivity analysis on several indicators, namely n_estimators, max_depth, min_samples_leaf,and min_sample_split. When analyzing each parameter, we ensured that the other parameters were kept at their default values. We performed a 5-fold cross-validation on the dataset, taking the average accuracy on both the training and testing sets as the result for each parameter. The resulting line chart is depicted in Figure 18.



(a) (b) (c) (d)

Figure 18: Sensitivity analysis chart

Analysis of the above line chart reveals that, with the increase in parameter values, both n_estimators and max_depth exhibit a trend of stabilizing influence on the model's predictive outcomes. At this point, the model's accuracy converges to a constant value, indicating insensitivity to further variations in these two parameters. Additionally, as the parameter values increase, the impact of min_sample_leaf and min_sample_split on the model's predictive performance undergoes continuous fluctuations, suggesting that the model is more sensitive to changes in these two parameters.

9 Model Evaluation and Further Discussion

9.1 Strengths

1. Entropy Weighting Method： The entropy weighting method in Model 1 has several advantages when used to evaluate the model. Since the entropy weighting method is an objective assignment method, it does not require the prior determination of subjective weights for individual criteria. After we have determined the evaluation factors of the players, it can automatically calculate the weights by using the concept of information entropy, which reduces the burden of subjective judgment. In addition, the entropy weighting method is able to consider more comprehensively the amount of information that each evaluation factor contributes to the degree of a player's performance, which helps to improve the comprehensiveness and accuracy of the model.

2. Grey Relational Analysis Method： When the sample is small, the accuracy of the Spearman coefficient method decreases dramatically and is susceptible to extreme values, while the gray correlation analysis method has relatively low data requirements and is highly adaptable. The dataset we used to build model 2 is "Wimbledon_featured_matches", which has a moderate sample size and is a more robust choice for gray correlation analysis.

3. Gray correlation is an analysis of the degree of similarity between the changes in two time series, and our treatment of the dataset happens to be time-based. Therefore, the choice of gray correlation analysis makes our model more comprehensive and reliable.

4. GSRF: Random Forest classification based on grid search combines two powerful techniques: random forest and grid search.GSRF uses the optimal combination of hyperparameters to train and predict the Random Forest model, preventing overfitting of the training data and improving the model's generalization ability.

9.2 Weaknesses And Further Discussion

1. Other external disturbances such as weather conditions, site conditions, external stimuli, etc. are not taken into account.

2. We use grid search in the buildup of model III and it may lead to increased computational overhead in large-scale parameter space. We can consider using more efficient tuning methods such as random search to find better parameter configurations in less time.

3. Due to the time factor, we did not take into account the personal factors of both sides of the match, such as the players' rankings in the tennis world, so our predictions may be less comprehensive. In the future, we can train with more evaluation factor datasets.

10 Memo

To: coaches
From: coaches
Subject: Results of the analysis of "Momentum" and advice for coaches

MEMORANDUM

Date: February 6, 2024

In sports, teams or players may feel they have momentum, or ”power,” but it is difficult to measure this phenomenon during a game, and it is not always obvious how events in a game create or change momentum (if at all). According to the requirements of the topic, we combined with the comprehensive evaluation system of entropy weight method to make a model that can capture the situation flow during the match. Using the correlation analysis model based on the grey correlation degree, we obtained that the grey correlation degree between the momentum of players and the outcome of the match reached 0.789, and the correlation coefficient was close to 1. It shows that the correlation between the two is strong, and the model can accurately judge which player performs better at a certain moment during the game and what factors affect his performance. At the same time, we also established a binary classification model based on gsrf algorithm, and the importance degree of each feature obtained through random forest training. the top three characteristics were serve depth nctl (serve depth not Close To Line) of 0.119, nt pt(made it to the net) of 0.108, and double fault(missed both serves) is 0.085. These three main factors have the greatest impact on the situation of the game, and further affirm the existence of momentum, momentum will affect the change of the situation of the game. In view of the influence of ”momentum” on the game situation, we make suggestions to the coach from the following aspects: 1.Strengthen daily physical training, so that players can have physical strength to cope with the next round even if they move a long distance during the game. 2.More random double singles training, so that players can be more flexible to deal with different characteristics of opponents. 3.Usually pay attention to letting players train more service related skills to increase, such as serving direction, serving distance from the line, serving speed, etc., try not to make service mistakes. 4.Players with the same matching ability make more contacts, improve the level of players returning the ball, and improve the probability of playing an untouchable winning ball. It can be seen from the reference [6] that not only a certain psychological process plays a role in the performance of athletes at a specific time and place, but also involves a variety of psychological processes and project psychological characteristics. Moreover, these psychological processes and individual psychological characteristics interact with each other to form a certain structural state. Namely: psychological state, thus affecting sports performance. Therefore, the coach should consider how to deal with the mental state of training when training players, how to solve the problem of mental state before, during and after the game, which has obvious practical significance for improving the level of sports training and competition.

11 References

[1] https://www.merriam-webster.com/dictionary/momentum

[2] https://zhuanlan.zhihu.com/p/641967744

[3] LIN Y J,KRUGER U,ZHANG J P,et al. Seasonal analysis and prediction of wind energy using random forests and ARX model structuresTJ1. IEEE transactions on control systems technology,2015,23(5):1994-2002.

[4] XU L N.SHEN Y B,FENG Z,et al. Applicability research and integrated application of various corection methods for wind speed forecast in a wind farm[J].Arid land geography,2022,45(4):1114-1124.

[5] WANG X J.SHI G PZHOU Z Q,et al. Revision of solar radiation product ERA5 based on random forest algorithmlJj. Remote sensing for natural resources.2022.34(2):105-111.

[6] YANG S S. An experimental study on the validity of competition psychological anxiety on serve scoring in tennis[J]. China Science and Education Innovation Magazine, 2011, (22): 21.