

# Advancing Tennis Analytics: Comprehensive Modeling for Momentum Identification and Strategic Insights

Junhai Lin<sup>a, †</sup>, Pengjin Shao<sup>b, †</sup>, Qingyang Zhang<sup>c, †</sup>

School of Business Administration, Zhongnan University of Economics and Law, Wuhan, China

<sup>a</sup>2509615737@qq.com, <sup>b</sup>1642950575@qq.com, <sup>c</sup>1642950575@qq.com

<sup>†</sup>Junhai Lin, Pengjin Shao and Qingyang Zhang were co-first authors.

## ABSTRACT

The 2023 Wimbledon men's singles final featured a thrilling clash between the emerging Spanish talent Carlos Alcaraz and the seasoned Novak Djokovic. Djokovic initially exerted control over the game, but Alcaraz effectively turned the tables and emerged victorious, highlighting the significance of "momentum" in the sport. Despite frequent discussion, effectively quantifying and forecasting the impact of this phenomenon poses significant challenges. Question 1 investigates the methods of tracking the progress of a tennis match and identifying players who demonstrate exceptional performance at specific moments. We constructed a model that considers variables such as serve probability advantage, technical indications, physical fitness, and mental condition. The model use time series analysis to monitor the dynamics of the game. It presents the game progress and player performance through interactive charts, allowing for efficient tracking of shifts in game momentum. Question 2 examines the influence of "momentum" on the game's result. We employed the XGBoost machine learning technique to construct a model, coupled with the particle swarm optimization algorithm to fine-tune parameters, and examined the pivotal aspects that influence the game's momentum. The model's findings demonstrate a notable association between momentum and game performance, which contradicts the belief held by certain coaches that game flow is purely impacted by random events. Question three necessitates the anticipation of significant moments of change in momentum within the game. We developed a predictive model that uses the PSO-XGBoost forest algorithm to accurately forecast the timing of momentum transitions. This is achieved by identifying and analyzing important statistics. The findings demonstrate that the model has the capability to precisely anticipate moments of momentum change, hence offering strategic guidance for coaches and players. Question 4 assessed the adaptability of the model in various competitive settings and produced a suggestion report based on the study of the model's results. We examine the implementation of the concept in several game scenarios and offer coaches and athletes strategic counsel to enhance their readiness and reaction to pivotal occurrences throughout the game. In conclusion, our research offers a comprehensive viewpoint on the significance of momentum in tennis matches and demonstrates the potential of machine learning methods in forecasting and capitalizing on this phenomena. Our approach not only improves comprehension of match progression, but also offers direction on how to implement these observations in actual matches.

## KEYWORDS

Tennis match; Momentum; Time series analysis; XGBoos; Particle swarm optimization

# 1. INTRODUCTION

## 1.1. Context of the problem

The 2023 Wimbledon men's singles final featured a thrilling clash between the emerging Spanish prodigy Carlos Alcaraz and the seasoned veteran Novak Djokovic. Despite Djokovic's dominant performance in the early stages of the match, the momentum suddenly shifted as Alcaraz managed to stage a comeback to secure victory. This tennis match underscores the significance of "momentum" in the sport. Momentum refers to a change in the competitive advantage resulting from a sequence of successive events, frequently leading to a significant influence on the match's final result. Although momentum is frequently referenced in sports, the precise measurement and prediction of this phenomena pose significant difficulties.

## 1.2. Restatement of the problem

### **Question 1: Monitor the speed of the game and identify players in favourable situations.**

Our first objective was to construct a model capable of monitoring variations in the tempo of a tennis match and precisely recognising individuals who demonstrated exceptional performance during certain instances. The model must incorporate characteristics such as the server's scoring likelihood, technical indications, physical fitness, and mental condition.

### **Question 2: Assess the influence of game momentum.**

Our subsequent objective is to utilise the constructed model to investigate the genuine influence of "momentum" in a match and ascertain whether it has a substantial effect on the result of the match. Our main objective was to verify the concerns voiced by some coaches regarding crucial periods in games and the general level of consistency. The scoring of the game is solely controlled by random factors.

### **Question 3: Anticipate significant transitional phases during the game.**

The objective of this challenge is to develop a prediction model that use match data to properly forecast significant power shifts in a match, namely how momentum transitions from one side to the other. Models must incorporate crucial variables such as the difference in scores and the advantage in service.

### **Question 4: Assess the suitability of the model and provide recommendations.**

We must assess the suitability of the model in various competitive contexts and provide strategic suggestions based on the analysis of the model. This encompasses the model's capacity to forecast results in various match scenarios, encompassing diverse match formats and playing surfaces, along with providing strategic guidance for players and coaches while encountering distinct adversaries and match circumstances.

# 2. PROBLEM ANALYSIS

## 2.1. Analysis of Question One

To address this issue, our strategy involves constructing a model that will meticulously track the advancements made during a tennis match. This model will then ascertain which player's performance is more exceptional at any given point in time. In pursuit of this objective, we have established a set of significant performance metrics, such as scoring rate, service winning rate, break point utilisation rate, and so on. Subsequently, the following actions entail employing time series analysis to monitor the fluctuations of the game and integrating the potential probabilistic superiority of the server into the model. Ultimately, we include these indications into the model and present the game's progress

and each player's success at various phases using interactive charts or heat maps. This allows us to monitor shifts in momentum and individual player performance.

## **2.2. Analysis of Question Two**

The objective of this inquiry is to employ recognised models in order to examine the impact of "momentum" in games, particularly with regards to conjectures put out by a tennis coach. The coach expressed scepticism over whether the alterations in the game and the consecutive victories of the players were merely fortuitous occurrences. We examine the uncertainty of transitions in games by measuring "momentum," which refers to the rate of successive victories, and analysing game data using statistical methods such random walk tests. By conducting comparative study, our objective is to evaluate the probable influence of "momentum" on match results, thereby addressing coaches' inquiries and confirming the genuine function and impact of momentum in matches.

## **2.3. Analysis of Question Three**

The objective of this job is to create a prediction model that can accurately anticipate the periods in a match when one team gains an edge over the other. Initially, it is necessary to ascertain the pivotal elements that might potentially influence the variability of the game, including factors like score difference, serving side, break point conversion rate, and so on. Subsequently, we employ machine learning methods, such as decision trees or random forests, to construct a prediction model that assesses the likelihood of a transition taking place, relying on the present state of the system. Through the process of cross-validation and rigorous testing of the model using more game data, we may enhance the precision of forecasting crucial transition phases.

## **2.4. Analysis of Question Four**

The task at hand involves assessing the model's capacity to perform well in various competitive settings and formulating strategic suggestions based on the analysis results of the model. To assess the model's predictive capability, it is necessary to analyse its performance on several match data sets, encompassing different tournaments and court surfaces. This evaluation will indicate the model's ability to generalise. Through the identification of potential constraints or areas for enhancement of the model in particular circumstances, and the amalgamation of the outcomes of model analysis and generalisation testing, we can offer athletes strategic suggestions for effectively managing diverse adversaries and competitive scenarios, thereby assisting them in enhancing their competition readiness.

## **3. MODEL ASSUMPTIONS**

- i. Assumed that changes in the mental state between players have an important impact on the results of the game and can be reflected through model analysis.
- ii. Assumed that during the game, the change in momentum is directly related to the player's scoring ability, technical indicators and physical condition.
- iii. Assumed that the key performance indicators (KPIs) selected in the model can comprehensively reflect the players' game performance and the dynamic changes in the game process.

## **4. SYMBOL DESCRIPTION**

**Table 1.** Symbol Description

Symbol	description
SD	Score difference
SA	server scores the advantage
UE	Unforced error
$SCGR_i$	serve scoring rate
WP	Winning point

## 5. TASK 1&2 MOMENTUM DETECTION MODEL

### 5.1. Data preprocessing

For categorical data, we assign numerical codes to represent different categories. Due to the extensive range of categories, specifically 32 categories, associated with the terminal brand, encoding all of them will impede the speed of model training. Therefore, we only encode variables that do not pertain to these features. Enhances the accuracy of distance calculation.

While coding, the introduction of dummy variables is necessary to handle multi-category variables. The process of defining dummy variables involves turning a variable with many categories into several variables with two categories each. By incorporating dummy variables into the model, it is possible to enhance the model's ability to properly and intuitively capture the influence of various aspects of the independent variable on the dependent variable, hence boosting the model's accuracy.

### 5.2. Feature Engineering

Given the abundance of classification features in the original data set, it is necessary to process these features in order to consolidate many characteristics into a smaller set. This will enhance the computational efficiency of the model and lower its overall complexity.

Initially, we compute the feature intersection of the two data sets and remove the features that are not present in both data sets. The speech characteristics observed after removal are: Based on this foundation, we have established the subsequent novel characteristics. Only a subset of the features are displayed in this presentation:

#### 5.2.1. Winning score ( $WP$ ):

Winners are points acquired through a player's assertive gameplay, frequently involving shots that are beyond the opponent's reach. The magnitude of the victory points reflects the player's offensive capability and proactiveness.

#### 5.2.2. Efficient serving rate ( $ESR_i$ )

The Efficient Serve Rate measures the effectiveness of scoring during serving by determining the proportion of points won by a player through ace balls and successful service games out of the total number of service games played.

$$ESGR_i = \frac{ace_i + sg_{win_i}}{sg_{total_i}} \quad (1)$$

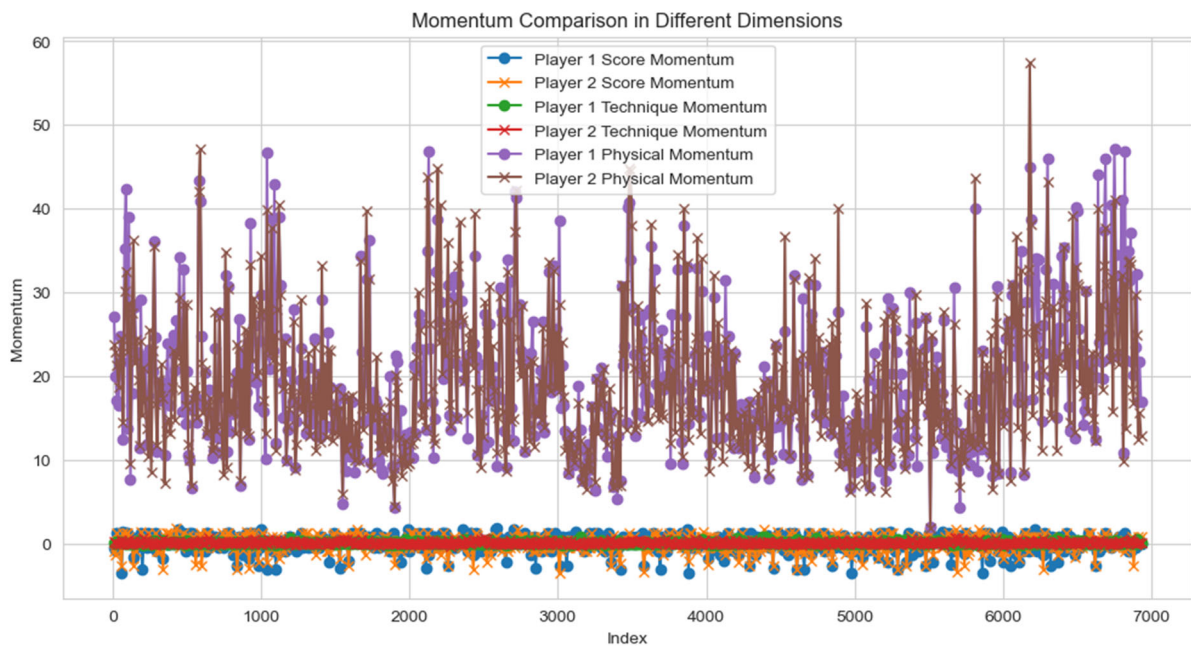
### 5.2.3. Service game scoring rate ( $SGR_i$ ):

The service game scoring rate is a metric that quantifies the percentage of points a player wins in their own service games, and it is closely linked to the efficiency of their serve. A higher scoring rate confers a bigger advantage to the player in their serving game.

$$SGR_i = \frac{sg_{win_i}}{sg_{total_i}} \quad (2)$$

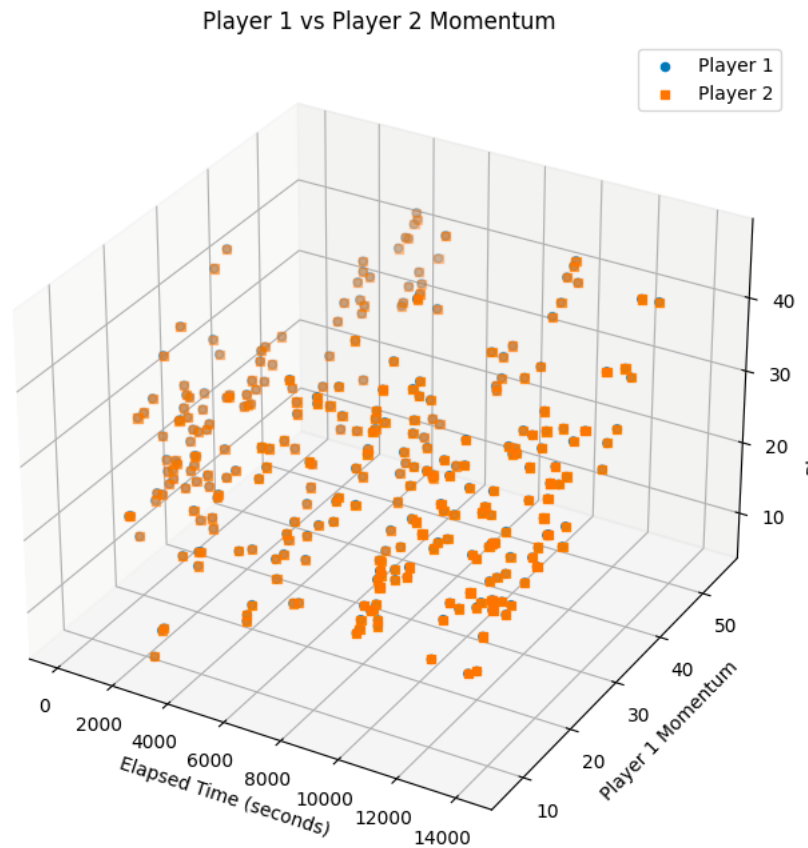
### 5.3. Task 1 Model solving

By following the aforementioned procedures for model creation, we utilize Python for programming, define our parameters, and compute our momentum outcomes, as depicted in the picture below:



**Figure 1.** Athletes' competition visualization

The image above depicts a comparison of the motivation levels of the two players across various parameters. The graphic displays six lines representing the scoring power, technical power, and physical power of both player one and player two. Every point corresponds to the power value at a certain moment. Player one demonstrates superior performance in scoring and technical prowess compared to player two, with minimal disparity in terms of physical strength. Player two exhibits a consistently low value in this category. Player one generally exhibits higher levels of motivation compared to player two in most circumstances, perhaps granting player one a competitive edge throughout gameplay.



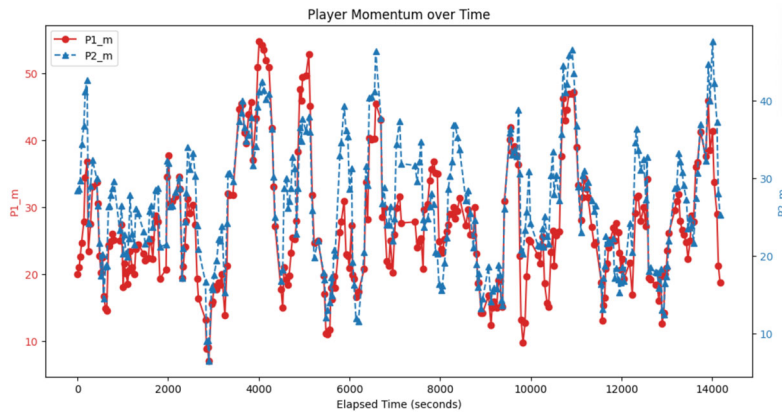
**Figure 2.** Plot of Player's Three-Dimensional Momentum

The graphic displays a three-dimensional scatter plot that examines the variations in momentum between two players throughout the game. The x-axis displays the game duration in seconds, while the y-axis indicates the momentum values of player 1 and player 2, respectively. The graphic displays the distribution of momentum values for Player 1 (blue dots) and Player 2 (orange dots) at various time points. The momentum of the two players varied during the game. Player 2 had a marginally greater momentum value compared to player 1 at specific intervals, implying that player 2 may hold an advantage in the game during these intervals. In general, the momentum values of the two players exhibit significant fluctuations, which indicate the shifting advantages between the two sides throughout the game and the intricate nature of the game process.

#### 5.4. Task 2 Model Solving

The graphic displays a three-dimensional scatter plot that compares the two players' momentum fluctuations throughout the game. The game time is displayed on the horizontal axis in seconds, while player 1's and player 2's momentum values are shown on the vertical

axis. The image clearly shows the distribution of the momentum values of Player 2 (orange dots) and Player 1 (blue dots) at various time points. Both players' momentum values changed as the game progressed. During several points in the game, Player 2 appeared to have a little more momentum than Player 1, which could indicate that Player 2 had the upper hand. It seems that the two players' momentum values shift a lot throughout the game, which is a reflection of the benefits and disadvantages that each side experiences and the intricacy of the game itself.



**Figure 3.** Player Momentum over time

Based on the data presented above, it is clear that the graph points indicate moments of momentum change for the athlete, and the frequency of these points also offers valuable insights into their performance. Athlete 2 has 65 momentum change points, while Athlete 1 has 59—a small decrease. Athlete 2 experiences more frequent changes in momentum during the game compared to athlete 1, which is reflected in this. One possible explanation for Athlete 1's outstanding result may have been that he kept a reasonable amount of momentum throughout the tournament. The final step was to employ t-tests to look for a connection between player momentum and the final score of the game. To determine whether momentum has a statistically significant effect on match results (wins or losses), we compute t-statistics and p-values. Below is the table displaying the output:

**Table 2.** t-test result table

	$P_1$	$P_2$
t-value	39.293	50.853
p-value	0.0000	0.0000

The t-test results show that the two players' momentum is significantly different from their game results. A t-statistic of 39.293 and a p-value of 0.0000 indicate that Player 1's momentum is significantly related to the game outcome. For the same reason, a t-statistic of 50.853 and a p-value of 0.0000 indicate that Player 2's momentum is significantly related to the game's outcome. With confidence, we may reject the null hypothesis that a player's momentum does not correlate with the outcome of the game, as the obtained p-value is significantly lower than the usual significance level of 0.05. That a player's momentum affects the game's outcome is supported by the fact that there is a direct correlation between the two. It also demonstrates that a shift in "momentum" is not a mere coincidence but rather a crucial factor in determining the game's ultimate result. For both coaches and athletes, this is crucial information.

## 6. TASK 3 & 4 FORECASTS AND RECOMMENDATIONS FOR ALTERING THE GAME

### 6.1. XGboost Forecast Model

For classification and regression issues, XGBoost (Extreme Gradient Boosting) is a scalable and efficient machine learning technique built on the gradient boosting framework. In order to construct an XGBoost model, one must use regularization, optimization of loss functions, and decision trees. Here is their model for addressing the elements that have an impact:

Step 1: Learner Base: XGBoost formulates its predictions using an additive model, which is a linear combination of base learners. Classification and Regression Trees, or CART, serves as the foundational learning here. This is the additive model:

$$f(x) = \sum_{k=1}^K h_k(x) \quad (3)$$

The output of the k-th decision tree, denoted as  $f(x)$ , is one example; K is the total number of decision trees.

Step 2: Function for loss: XGBoost calculates the margin of error between the actual value and the model's predictions using the loss function. For regression issues, the loss function can be squared error; for classification problems, it can be log loss; and so on. The loss function can be expressed mathematically in the following way:

$$\mathcal{L}(y, f(x)) \quad (4)$$

Among these,  $y$  represents the actual value and  $f(x)$  stands for the expected value.

Step 3: Function of the objective In XGBoost, the objective is to locate a model that reduces the loss function to its minimum. To further avoid model overfitting, XGBoost includes regularization terms. Here is the expression for the objective function:

$$\mathcal{J}(f) = \sum_{i=1}^n L(y_i, f(x_i)) + \Omega(f) \quad (5)$$

where  $n$  is the number of samples in the data set,  $\Omega(f)$  is the regularization term, and represents the objective function.

Step 4: The phrase for regularization A regularization term is used by XGBoost to penalize the model's complexity. Two components make up the regularization term: the first is the total number of leaf nodes, and the second is the sum of squares of their weights. The regularization term can be expressed mathematically in the following way:

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2 \quad (6)$$

Among them,  $\Omega(f)$  represents the regularization term,  $\gamma$  and  $\lambda$  are regularization parameters,  $T$  is the number of leaf nodes, and  $\omega_j$  is the weight of the  $j$ th leaf node.

## 6.2. Establishment of PSO optimization XGBOOST model

An intelligent optimization system called Particle Swarm Optimization (PSO) searches for the best solution by mimicking the actions of natural groupings like fish or birds in a flock. As an ensemble

learning technique, XGBoost (eXtreme Gradient Boosting) uses a number of different decision tree models to arrive at its predictions.

If we optimize the XGBoost model using the PSO algorithm, we can discover a better combination of parameters and the model's prediction accuracy will increase. The model's precise procedures are as follows:

Step 1: the XGBoost model's parameter space must be defined.

For example, you can change the learning rate, maximum depth, minimum child node weight, and many more parameters in the XGBoost model. A search range for every parameter must be first set. To illustrate:

learning\_rate (learning rate):

max\_depth (maximum depth):

min\_child\_weight (minimum child node weight):

Step 2: Particle Swarm Initialization

A particle swarm must be initialized before the PSO algorithm can be applied. Parameters for the XGBoost model are represented by each particle. Within the search space, the initial values of the particle's position and velocity will be chosen at random.

Step 3: create the fitness function

To determine how good a particle is, or what combinations of parameters in the XGBoost model are best, one uses the fitness function. Usually, fitness values can be determined using model performance measures like root mean square error (RMSE) or mean absolute error (MAE) that are derived by cross-validation. A possible expression for the fitness function is:

$$f(\eta, d, w) = -\text{RMSE}(\text{XGBoost}(\eta, d, w)) \quad (7)$$

In this example, we use negative RMSE as the fitness value because we want to minimize the RMSE.

Step 4: Update particle position and velocity

In each iteration, we will update the particle's position and velocity according to the following formula:

$$\begin{aligned} v_i(t+1) &= w_1 v_i(t) + w_2 c_1 r_1 (p_{best} - x_i(t)) + w_3 c_2 r_2 (g_{best} - x_i(t)) \\ x_i(t+1) &= x_i(t) + v_i(t+1) \end{aligned} \quad (8)$$

This includes the following variables:  $p_{best}$ , which stands for the particle's initial position,  $w_1$ , which stands for the particle's inertia weight,  $w_2$  and  $w_3$ , which stand for acceleration constants,  $c_1$  and  $c_2$ , which denote learning factors, and  $r_1$  and  $r_2$ , which denote random numbers. When we talk about  $i$ ,  $g_{best}$ , we're talking about the group's best position in history.

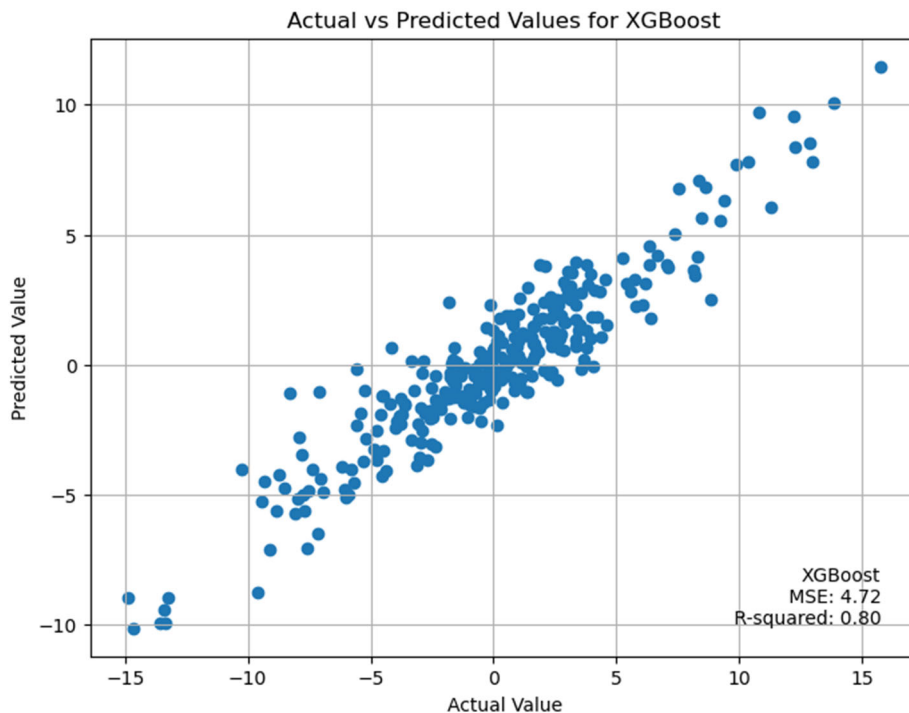
Step 5: Conditions for Termination

When the particle swarm's performance stops improving noticeably or when the maximum number of iterations is achieved, the method ends. As an example of an ideal parameter combination, we can look at the group's best performance in the past.

### 6.3. Model Solving

Python is the language of choice for executing the solution while we construct the XGBoost model. The XGBoost model's parameter values were initially determined. We optimized the model's performance after meticulously adjusting its settings, including the learning rate (learning rate=0.01), maximum tree depth (max\_depth=50), subsample ratio (subsample=0.2), column sampling ratio (colsample\_bytree=0.1), and more.

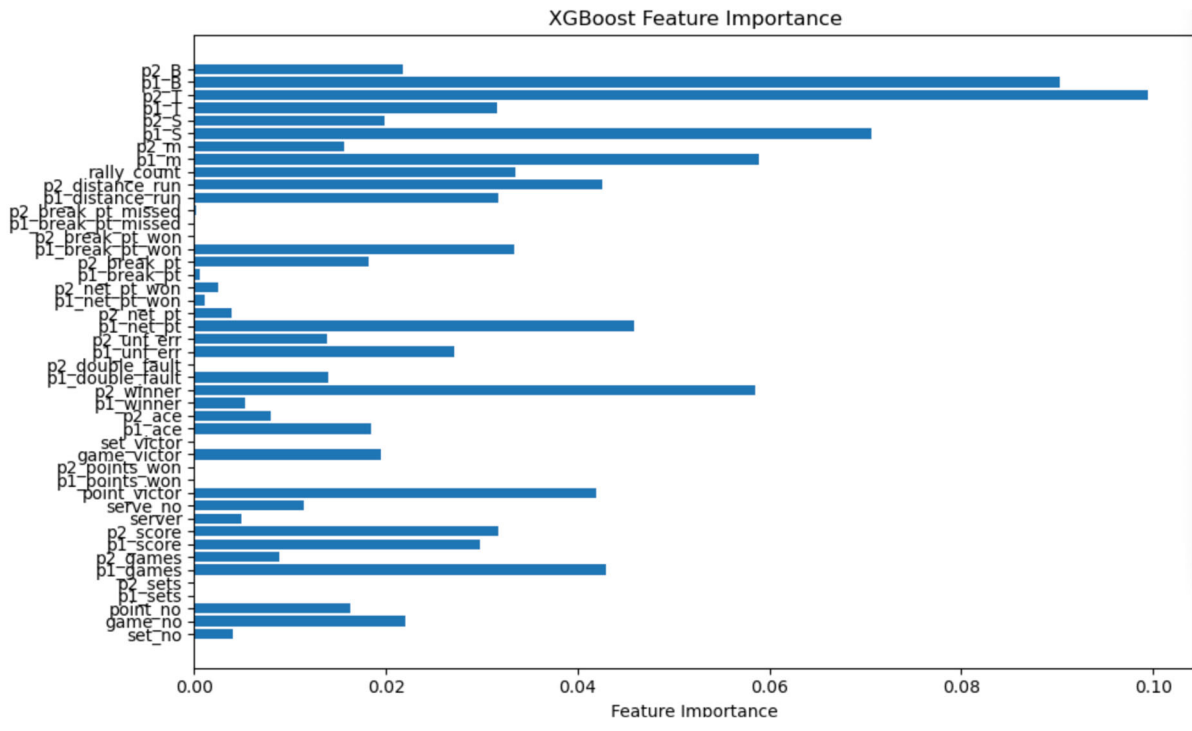
We train the model using XGBoost once we determine the parameters. Through the process of fitting and the input of segmented training data sets, the XGBoost algorithm iteratively learns and modifies the decision tree structure. We implemented an early halting mechanism during training to decrease overfitting and improve the model's generalizability. This means that after a specific number of iterations, training will cease unless there is a sufficient improvement in the test error.



**Figure 4.** XGboost regression result

Expected and true values cluster tightly around the contour  $y=x$  in the display of an XGBoost-based predictive model. This demonstrates that the model's predictions and the actual conditions are in good accord. We found that the model's coefficient of determination ( $R^2$ ) was 0.80 after performing a thorough statistical study. The model successfully accounts for a significant portion of the observed data fluctuation, as this number is near to the ideal fit threshold of 1. Along with that, the MSE (mean square error) is 4.72. The model's utility in capturing momentum shifts in tennis matches is demonstrated by lower mean square error (MSE) values, which suggest fewer prediction mistakes and higher forecast accuracy.

We used the XGBoost library's built-in capabilities to build feature importance charts so we could better understand how each feature affected the model's predictions. This graphic displays the forecast model's elements in descending order of relevance using bar graphs and other formats. This lets us zero in on the elements that matter most for predicting shifts in momentum.



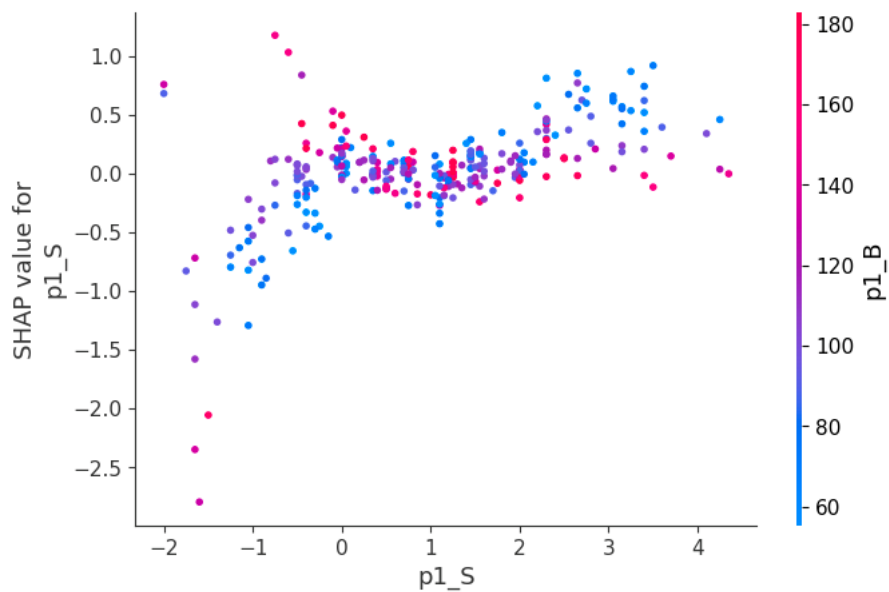
**Figure 5. Feature Importance**

To begin, determine the maximum and minimum values for the following parameters: sizepop=1000, particle dimension=402\*24, maximum number of iterations=100, inertia weights for individuals and groups  $c_1=0.8$ ,  $c_2=0.5$ , and  $c_3=0.5$ .

The XGBoost model's feature importance levels are displayed in the figure above. On one side, we have various features, and on the other, we have the impact that each feature has on the model's prediction outcomes. In terms of influence on the model, the features on the left side of the figure are more significant than those on the right. The fact that "p1\_T" is the most important variable in the model's predictions is supported by its high significance value. The second running distance of athlete 2 is represented by the variable "p2\_T" that follows immediately after.

**6.4. Model generalization analysis**

To discover the best values for the indicator parameters that correspond to p1 and p2, we utilize the optimization results to set the maximum value of  $f(x)$  and then optimize using the particle swarm technique. Above, you can see the specified parameters. In addition, we Here are the results: 'p1\_games', 'p1\_score', 'p1\_ace', 'p1\_break\_pt', 'p1\_distance\_run', 'p1\_m', 'p1\_T', 'p1\_B'. In addition, players 1 and 2's momentum fluctuations During the competition, it is recommended to adopt a more focused approach. Player 2 has some work to do in the areas of focus and consistency on court, serve receiving, and understanding crucial points. Meanwhile, Player 2 needs to work on his fitness and technique, up his offensive game, and apply more pressure to Player 1 on the court so that Player 1 makes more mistakes. From a strategic standpoint, Player 2 can outmaneuver Player 1 by slowing the game down, utilizing deflection balls, and prolonging the round duration to sap Player 1's energy. Player 2 can raise their winning percentage and keep their scoring edge with these techniques.



**Figure 6.** SHAP value (impact on model output)

One way to assess model predictions is by looking at Figure 9, which shows different SHAP values, or SHapley Additive exPlanations. The SHAP value is a representation of the feature values' contribution to the model output. One possible interpretation of this scatter plot is that player 1's SHAP value for their physical advantage (p1\_B) is represented vertically and player 1's scoring advantage (p1\_S) is represented horizontally. Variegated color gradients, with blue indicating lower values and red indicating higher values, can be used to indicate various p1\_B values.

From the graphic, we can deduce that there is a positive association between the SHAP value of p1\_B (physical advantage) and the score advantage (p1\_S), suggesting that physical advantage influences the model output in a score advantage-heavy scenario. Bigger as well. This highlights how a player's momentum can be influenced by their scoring advantage and physical edge in a game.

Key indications for player 1 according to the particle swarm optimization method results are 'p1\_games', 'p1\_score', 'p1\_ace', 'p1\_break\_pt', 'p1\_distance\_run', 'p1\_m', 'p1\_T', and 'p1\_B'. If Player 2 wants to keep up with Player 1 and increase his winning percentage, he needs to work on these areas of his game performance. If Player 2 wants to get the better of Player 1, he needs to focus more on the game, receive serves and important points better, get in better shape, and think of ways to use Player 1's strength against him.

## 6.5. Suggestion

The effect of "momentum" on tennis matches: findings and suggestions from the research.

Dear coach:

Following extensive research into the 2023 Wimbledon men's singles final, our group has developed a suite of models meant to provide light on the significance of "momentum" and foretell major changes in the game's trajectory. We believe that the following findings and recommendations will be useful to your team's training and game strategy as a result of our thorough analysis of aspects such as scoring advantage, technical statistics, and physical fitness.

Based on our findings, "momentum" is an important component of race results. A player's momentum changes the game's trajectory when they find themselves ahead in scoring. This occurrence is associated with the players' mental and physiological well-being as much as it is with their technical ability.

When it comes to momentum, there are a few things that should be prioritized throughout training:

Strengthen the player's capacity to serve efficiently, convert break points, and hit long balls and net volleys from the baseline with more solidity. To keep performing at a high level until the very end of the game, the players need work on their physical conditioning, particularly their speed and endurance. To assist players become more resilient under pressure and better understand game mechanics, we offer psychological counseling in addition to virtual game training.

Handling changes in the game:

In order to obtain a feel for the game's tempo, we recommend that the coaching staff study their opponents' playing styles and routines. Keep track of how the momentum of the players varies throughout the game, make tactical adjustments quickly, and capitalize on the momentum to your advantage. In order to help a player get back on track while their momentum is going down, tactical and psychological modifications are performed in a timely manner.

We have faith that your team's players can improve their understanding of the game's momentum and overall performance through the use of scientific training methods and meticulous game strategy. We hope that in future competitions, your team does very well.

I hope you have a great time training and that you win the competition!

## **7. EVALUATION OF THE MODEL**

### **7.1. Benefits of the model**

Our approach can identify shifts in momentum and crucial match turning points by integrating elements like scoring advantage, technical data, and physical fitness. To enhance the efficiency and accuracy of predictions, the model employs state-of-the-art machine learning techniques as XGBoost and particle swarm optimization algorithms. Also, the model can give coaches and players useful strategic advice that can improve their game.

### **7.2. Drawbacks of the model include**

Even while the model can give detailed game analysis, it uses statistics and data from the past, thus it might miss some unexpected events or sudden changes in player performance in real-time games. Further validation and adjustment at different levels of competition and under different conditions may be necessary if the model's generalizability is constrained by the data set's diversity and quality.

### **7.3. Enhancing the model**

More real-time game data would be great for capturing instant momentum changes and player performance. Adding more types of games and diverse specialized environmental conditions to the data set would improve the model's generalization ability. Applying deep learning to process complex data and capture more nuanced game dynamics would be another great step toward making the model more practical and accurate. These updates will allow the algorithm to analyze and forecast tennis matches with more precision and depth.

## **8. REFERENCE**

- [1] HOU Xianbiao, WANG Li. Reflections on the use of multi-ball training in college sports tennis teaching [J]. Z. Contemporary Sports Technology, 2021,21:137-140. (in Chinese). <https://www.zhangqiaokeyan.com/academic-journal-cn-contemporary-sports-technology-thesis/0201291082875.html>.
- [2] Fan Yahui, An Ping. An Analysis of the Application of "Monistic Training Theory" in the Teaching of Public Sports Tennis in Colleges and Universities[J]. Sporty style,2021,2:176-178.(in Chinese). [https://www.zhangqiaokeyan.com/academic-journal-cn\\_sport-style-thesis/0201287072601.html](https://www.zhangqiaokeyan.com/academic-journal-cn_sport-style-thesis/0201287072601.html)
- [3] Li Wanrong. An Analysis of the Ideological and Political Path of College Sports and Tennis Courses[J]. Journal of

Shanghai Second Polytechnic University, 2021,1:65-69.(in Chinese). [https://www.zhangqiaokeyan.com/academic-journal-cn\\_journal-shanghai-polytechnic-university\\_thesis/0201290354112.html](https://www.zhangqiaokeyan.com/academic-journal-cn_journal-shanghai-polytechnic-university_thesis/0201290354112.html).

- [4] Zhang Wei. Research on Error Technology Motion Recognition in Sports Tennis Based on Machine Vision Technology[J]. Journal of Qiqihar University (Natural Science Edition), 2021,6:26-30.(in Chinese). [https://www.zhangqiaokeyan.com/academic-journal-cn\\_journal-qiqihar-university-natural-science-edition\\_thesis/0201290682733.html](https://www.zhangqiaokeyan.com/academic-journal-cn_journal-qiqihar-university-natural-science-edition_thesis/0201290682733.html).