Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 5
AIDML 2024

A Quantitative Analysis of the Effect of Pitch Sports on Players' Potential Energy

Weiyu Tang *, #, Xuanzeng Yi #, Huimin Zhao #

College of Transport and Communications, Shanghai Maritime University, Shanghai, China, 201306

* Corresponding author: 202110613066@stu.shmtu.edu.cn

#These authors contributed equally.

Abstract. This paper introduces a research framework for sports science, particularly in competitive tennis, focusing on momentum as a crucial factor impacting match outcomes. It addresses controversies surrounding momentum's definition, quantification methods, and its influence on results, highlighting the lack of research on momentum strategies. The study proposes a paradigm for calculating potential energy and employs the XGBOOST model optimized with DA to predict players' potential energy values with high accuracy (R2=0.9836). Through SHAP analysis, it delves into potential energy fluctuations, identifying factors such as continuous court running, winning shot execution frequency, stroke distance, intensity, and ball speed as significant influencers, varying among players. Furthermore, the paper utilizes GA for potential energy optimization, offering tailored training strategies. It concludes by discussing the model's universality, emphasizing its research significance. This framework aids in pinpointing match shortcomings in individual players and devising personalized training plans accordingly.

Keywords: Momentum; Prediction Model; XGBOOST; Dragonfly algorithm; Genetic algorithms.

1. Introduction

In sports competitions, players' performances are influenced by various factors on the field, thereby affecting the direction and outcome of the match. It is observed that seemingly dominant players may experience unexpected changes in their performances due to changes in the environment or psychological factors.[1] M. Dietl, H., and Nesseler, C. analyzed similar matches and found that players benefit from controlling certain positive or negative momentum, leading to changes in the outcome of the game.[2] It is evident that the dynamics of momentum in a match are significant and cannot be overlooked. Therefore, conducting strategic research based on the fluctuations of player momentum holds practical significance.

In pursuit of our research objectives, we have employed the XGBOOST model, the Dragonfly optimization algorithm, the SHAP model, and genetic algorithms. The XGBOOST model iteratively trains to minimize the error between predicted values and actual values. Ji utilized this model to enhance sales forecasting accuracy, which is superior to other models.[3] Gu improved the precision of land subsidence prediction using an enhanced XGBOOST model.[4] Mirjalili introduced the Dragonfly Algorithm in 2016, which is designed to solve optimization problems swiftly and accurately.[5] Wang utilized this algorithm to optimize air pollution forecasting, addressing the deficiency of traditional methods in predicting pollution fluctuations.[6] Nordin employed the SHAP model to interpret the characteristics of individual suicide attempts.[7] Ziakopoulos utilized the SHAP model to understand the influencing factors of using mobile phones during driving processes.[8] Chen utilized genetic algorithms to optimize personalized travel route planning algorithms, resulting in more targeted recommended routes.[9] Ren improved flood control scheduling for reservoirs by optimizing it using a genetic algorithm.[10] Li obtained a more optimized financial investment portfolio based on genetic algorithms.[11] In conclusion, the aforementioned algorithms have been applied in a wide range of practical scenarios, demonstrating their feasibility for tasks involving prediction and optimization.

In this study, a specific match dataset was selected for analysis, and the XGBOOST predictive model was enhanced with the DA algorithm for parameter optimization. After establishing model evaluation metrics, both subjective and objective factors in the match were chosen as analysis targets. Using the SHAP model, the feature importance of different factors on the predictive results was visualized. For post-match strategy selection, genetic algorithms were employed along with the XGBOOST model to explore various combinations of subjective factors. By comparing the potential energy levels under different parameter combinations, the optimal solution for maximizing potential energy was determined. Additionally, visual analysis was conducted to investigate players' competitive strategies.

2. Calculation of Momentum value

Establishing a momentum prediction model requires us to first quantify and define the players' momentum. During the game, a player's momentum can be described by their scoring ability at different time points. Therefore, it is necessary to select indicators that can reflect the player's performance in various aspects of the game to determine their scoring ability and level, thus quantifying the player's momentum. Based on the characteristics of tennis matches, we have selected a total of 6 indicators to gain insight into the player's ability changes, as detailed in Table 1.

Table 1. Indicators of Player’s Ability and Mathematical Interpretation

Index	Player’s Ability	Formula	Description
Winner	Mastery of the Game	$M_W = W_{P1} - W_{P2}$	Subtracting winning points for two players
Break Points	Strategic Adaptability	$M_{BrPts} = BrPts_{P1} - BrPts_{P2}$	Subtracting the number of points won by two players as receivers.
Serve Advantages	Scoring Ability when in Lead	$M_{S-A} = SWon \times 0.1$	Multiply points a player wins as a server by the serve advantage weighting
Unforced Error	Stability & Mental Quality	$M_{Unf-Err} = -UnfErr$	Negative value of players' unforced errors
Points Advantages	Scoring & Performing Ability	$M_{Pts-A} = Pts_{P1} - Pts_{P2}$	Difference between the scores of the two players
Set Won/ Game Won	Understanding of the Game	$M_{St-Won} = StWon_{P1} - StWon_{P2}$ $M_{G-Won} = GWon_{P1} - GWon_{P2}$	Subtracting the number of games/sets won by two players

The impact of each indicator on the player's momentum varies. To determine the correlation between a player's performance and their abilities, we utilize the entropy weight method to calculate the weight of each indicator. A higher entropy value indicates lower dispersion of the indicator, resulting in a smaller weight, implying a lesser impact of that indicator on the player's performance. Since no subjective assignment of weights is required, this method is more objective and accurate compared to other weighting methods.

Calculating the entropy of indicators first requires all indicator data to be on the same scale. Therefore, it is necessary to normalize the corresponding indicator values for each time point of the player. The normalized value x'_{ij} of indicator x_{ij} at time point i is given by Eq. (1):

$$
\begin{cases}
x'_{ij} = \dfrac{x_{ij} - \min(\sum x_j)}{\max(\sum x_j) - \min(\sum x_j)} & (x_j > 0) \\
x'_{ij} = \dfrac{\max(\sum x_j) - x_{ij}}{\max(\sum x_j) - \min(\sum x_j)} & (x_j < 0)
\end{cases}
\tag{1}
$$

After standardizing the indicators, the entropy value e_j of each indicator can be calculated, as shown in Eq. (2):

$$
e_j = -\frac{1}{\ln(n)} \sum_{i=1}^{n} p(x'_{ij}) \log(p(x'_{ij}))
\tag{2}
$$

In the equation, n represents the total number of time points for indicator j, and p represents the probability of its occurrence. After obtaining the entropy of each indicator, the weights of each indicator can be calculated, as shown in Eq. (3):

$$
w_j = \frac{1 - e_j}{\sum_{j=1}^{m} (1 - e_j)}
\tag{3}
$$

Where m is the total number of indicators, 1-e_j represents the redundancy of the indicator j. The comprehensive score obtained by multiplying the indicator weights at each time point by their corresponding values and summing them up represents the player's momentum at that time point, as shown in Eq. (4):

$$
M = \sum_{j=1}^{m} w_j \cdot x_j
\tag{4}
$$

By calculating the comprehensive score at all time points using Eq. (4), we can determine the momentum changes for each player throughout the entire match. A higher score indicates higher momentum for the player, increasing their probability of winning. Conversely, a lower score suggests an unfavorable situation for the player, necessitating timely strategic adjustments.

3. Optimizing Tennis Performance Prediction

3.1. Data Acquisition

For this study, we will focus on analyzing the men's singles matches from the 2023 Wimbledon Championships. The required dataset can be obtained from www.mathmodels.org, specifically the file named [Wimbledon_featured_matches.csv] [12].

3.2. Data Preprocessing

This article selects the 'match_id' index to analyze all data corresponding to the match '2023-wimbled-1301'. The players involved in the match are 'Carlos Alcaraz' (hereinafter referred to as 'CA') and 'Nicolas Jarry' (hereinafter referred to as 'NJ').

To assess the feasibility and accuracy of the predictive model, the following data processing steps are required:

3.2.1. Data type query.

Based on the Jupyter Notebook, the following index columns of non-continuous numerical values can be obtained from the original data, where various fields have significantly different data types, including multiple text-based data which are often not conducive to analysis, as shown in Table 2.

Table 2. Index query

Index Column	Data Type	Index Column	Data Type
match_id	object	p1_score	object
player1	object	p2_score	object
player2	object	winner_shot_type	object
return_depth	object	serve_depth	object
elapsed_time	object	serve_width	object

3.2.2. Missing value handling.

Based on the Jupyter Notebook, a query was conducted to identify missing values in all index columns. The results indicate that there are 11 missing values in 'speed_mph', one missing value in 'serve_width', and one missing value in 'serve_depth'. However, 'return_depth' has 48 missing values, which are of text type and are difficult to effectively analyze. Therefore, this column will be removed from the analysis. Missing values in the remaining index columns will be filled with 0.

3.2.3. One-hot encoding processing.

Based on the data query, it is found that there are multiple categorical variables. To facilitate subsequent model training, it is necessary to perform one-hot encoding processing on them. One-hot encoding is a commonly used method in data processing to convert categorical variables into binary vectors. This method can transform multiple abstract values under a single feature into a single concrete value (e.g., 1, with all other values being 0) and ultimately obtain a binary sequence.

For example, in the "p2_score" column, there are 11 categorical values such as "15", "30", "40", etc. Encoding processing will create 11 columns for it. In each column, the corresponding value will be transformed to "1", while all other values not in the column will be "0". Finally, the binary columns will be arranged in the original order.

The same method will be applied to the "p1_score", "winner_shot_type", "serve_width", and "serve_depth" index columns. The encoded index data will then be integrated with the retained data from the source file.

3.2.4. Incremental Difference Sequence.

Taking "CA" as an example, to effectively quantify his on-court performance, this article constructs a differential sequence that represents the momentum increment and decrement.

With a step length of l, i.e., taking the difference between the momentum value at the current moment and the previous moment, the resulting value represents the current moment's momentum increment or decrement, where a positive value indicates growth and a negative value indicates weakening. Following the chronological order, a complete sequence of momentum fluctuations is obtained sequentially. This sequence will serve as the independent variable Y for model training and prediction. Moreover, this sequence can continuously reflect the player's momentum fluctuations on the court. The formula for lagged features is as follows:

$$\Delta X_t = X_t - X_{t-l} \tag{5}$$

$\triangle X_t$ represents the current moment's energy-momentum, where $X_t$is the energy value at the current moment, and X_{t-l} is the energy value at the moment l units ago with a step length of l being 1 unit.

3.2.5. Selection of Adjustable Metrics under Optimized Competitive Strategy.

To ensure that players can effectively respond to various types of matches and achieve outstanding results, it is necessary to develop a real-time adaptable game strategy that continuously adjusts oneself to unleash maximum momentum and gain a competitive advantage in the match.

After one-hot encoding, the index columns have expanded from the original 42 to 69. Using Jupyter Notebook, all index column information is filtered, removing 32 objective factors, and selecting the remaining 9 subjective factors. These factors represent information that "CA" or "NJ" can control on the court by adjusting themselves, as shown in Table 3.

Table 3. Controllable Variable of “CA”

Ace	Winner	Double_fault	Unf_eff	Net_pt	Net_pt_won	Break_pt	Break_pt_won	Break_pt_missed
a1	a2	a3	a4	a5	a6	a7	a8	a9

3.3. Energy Change Prediction and Interpretable Model

3.3.1. Establishment of XGBOOST Model after DA Optimization.

This article employs the Dragonfly Algorithm (DA)[5] to optimize parameters for the XGBOOST model. Known for its strong global search capability and fast convergence speed, DA is well-suited for a variety of optimization problems, particularly continuous ones. The algorithm mimics the behavior of a population of dragonflies searching for prey. After fine-tuning, key parameters such as "max_depth," "min_weight," "learning_rate," and "n_estimators" are optimized within a range over a maximum of 100 iterations. Optimal parameters obtained using the DA algorithm include "max_depth" of 6, "min_weight" of 3, "learning_rate" of 0.1, and "n_estimators" of 100. These optimized parameters enhance the performance of the XGBOOST model, which is further integrated with the differential sequence model for improved efficacy, represented by the following formula:

$$Yt = f(\Delta X_1, \Delta X_2, \Delta X_3, \ldots, \Delta X_t) \tag{6}$$

In this context, f represents the mapping function of the machine learning model; Ytrepresents the dependent variable of the model output, i.e., the potential energy value at future time steps; ΔX_t represents the potential state at the current time step.

3.3.2. Data Set Splitting.

The data set will be divided into a 7:3 ratio, with 70% of the data used as the training set and the remaining 30% for the test set.

3.3.3. Selection of Evaluation Metrics.

To evaluate model performance, we choose the following metrics:

- Root means square error (RMSE): used to measure the deviation between the predicted value and the actual value.
- Mean absolute error (MAE): measures the average absolute value of the prediction error.
- R2 score: Evaluates the model’s ability to explain price changes.

3.3.4. Construction of Global Feature Importance.

Based on the research needs of competitive strategy, this study aims to determine which factors have the greatest impact on a match. To achieve this, a SHAP model is established to interpret the input parameters of the XGBOOST model.

The main or objective factors from "2023-Wimbledon-1301" are selected as the subjects of analysis, with each sample's global feature importance denoted as S_j. Within each sample, there are n feature values, denoted as S_{ij}. The global feature importance can be calculated using the following formula:

$$S_j = \frac{\sum_{i=1}^{n} | S_{ij} |}{n} \tag{7}$$

Taking the absolute values of each feature value, summing them up, and then averaging them provides an effective measure to assess the contribution of the j factor to the prediction results. This approach disregards the positive or negative effects, thereby mitigating the instability introduced by directly using the XGBOOST model for feature selection, i.e., the instability caused by "feature_importances".

3.4. Dynamic Adjustment Strategy Model for Optimal Momentum

This study aims to maximize the competitive advantage of players in the future by exploring how competition strategies should be adjusted or optimized based on the difference between maximum momentum and original momentum. Therefore, utilizing nine subjective indicators that can be adjusted [12], with "CA" as the subject, an objective programming problem is established and solved. Combined with the momentum prediction model, the optimal momentum values are repeatedly calculated to design a set of strategies for this player to face "NJ" in matches. The relevant calculations are as follows:

3.4.1. Definition of Unit Time Adjustment of Momentum($\boldsymbol{M}$).

Players can compensate for disadvantages during matches by adjusting their strategies, offensive and defensive tactics, focusing on hitting, and other series of on-court strategies. The unit time momentum will be composed of nine subjective factors through some functional relationship, expressed as follows:

$$M = f(a1, a2, a3, a4, a5, a6, a7, a8, a9) \tag{8}$$

Among them, a1: number of shots; a2: number of winning points; a3: number of errors; a4: number of unforced errors; a5: attempts at net points; a6: successful net points; a7: number of breakpoints; a8: number of successful breakpoints; a9: number of failed break points.

3.4.2. Definition of Maximum Momentum per Unit Time($\boldsymbol{M_{max}}$).

The optimal competitive strategy stems from the player's maximum burst of momentum at the current moment. This momentum value can be regarded as the optimal function combination of the nine subjective factors. Based on the calculation formula of unit time momentum(M_{max}), an objective programming model can be established to ultimately obtain a parameter structure for when the momentum is maximized.

$$\begin{cases} S.T.M_{\max} = f(a1, a2, a3, a4, a5, a6, a7, a8, a9) \\ a1, a2, a3, a4, a5, a6, a7, a8, a9 \geq 0 \end{cases} \tag{9}$$

3.4.3. Unit Time Maximum Momentum Prediction Model Based on Genetic Algorithm (GA).

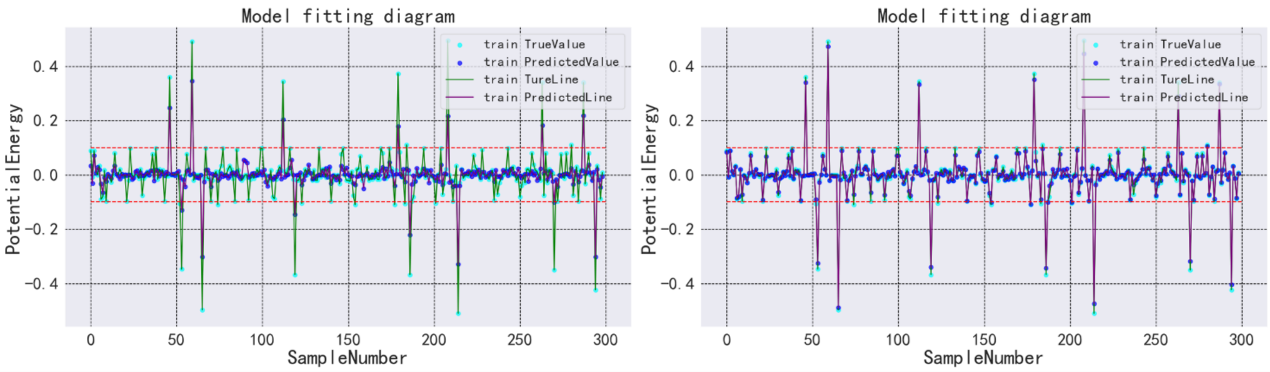
The selection of the nine subjective factors involves a large number of parameter combinations. This study aims to find a combination from this vast set that maximizes the current momentum, termed as the optimal combination. Genetic Algorithm (GA) is widely used for such problems and has demonstrated good solving performance, making it effective for addressing this problem.

The current momentum is predicted by the established XGBOOST model. By combining the XGBOOST model with the GA model, a composite model is formed. This composite model compares momentum values for multiple input combination parameters. When the momentum is maximized, denoted as M_{max}, the nine input parameters represent the optimal combination. The composite model exhibits both the rapid convergence of the GA algorithm and the good momentum prediction ability.

4. Strategies for Enhancing Athlete Potential Based on Prediction Model Insights

4.1. Model Performance

4.1.1. Comparison of Individual Models.



XGBOOST **DA+XGBOOST**

Figure 1. Comparison between Formal and Lateral Model

For the prediction of momentum values for player "CA," the initial performance of the XGBOOST model was poor during testing on the test set. However, after optimization using the Differential Evolution (DA) algorithm, the model's performance improved significantly, as shown in Figure 1. Based on evaluation metrics, it can be observed that the model's performance increased from 0.6886 to 0.9836, indicating a substantial improvement in both generalization and learning abilities, as shown in Table 4.

Table 4. Comparison of Evaluation Indicators

Model.	RMSE.	MAE.	R2.
DA+XGB	0.9391	0.0215	0.9836
XGB	0.6210	0.0609	0.6886

4.1.2. Comparison of Multiple Models.

Continuing with the evaluation of the test set, comparisons were made between the optimized XGBOOST model and other models including Random Forest, Multilayer Perceptron, and Linear Regression. By comparing the metrics data of multiple models, a visual representation was obtained to observe the performance of each model.

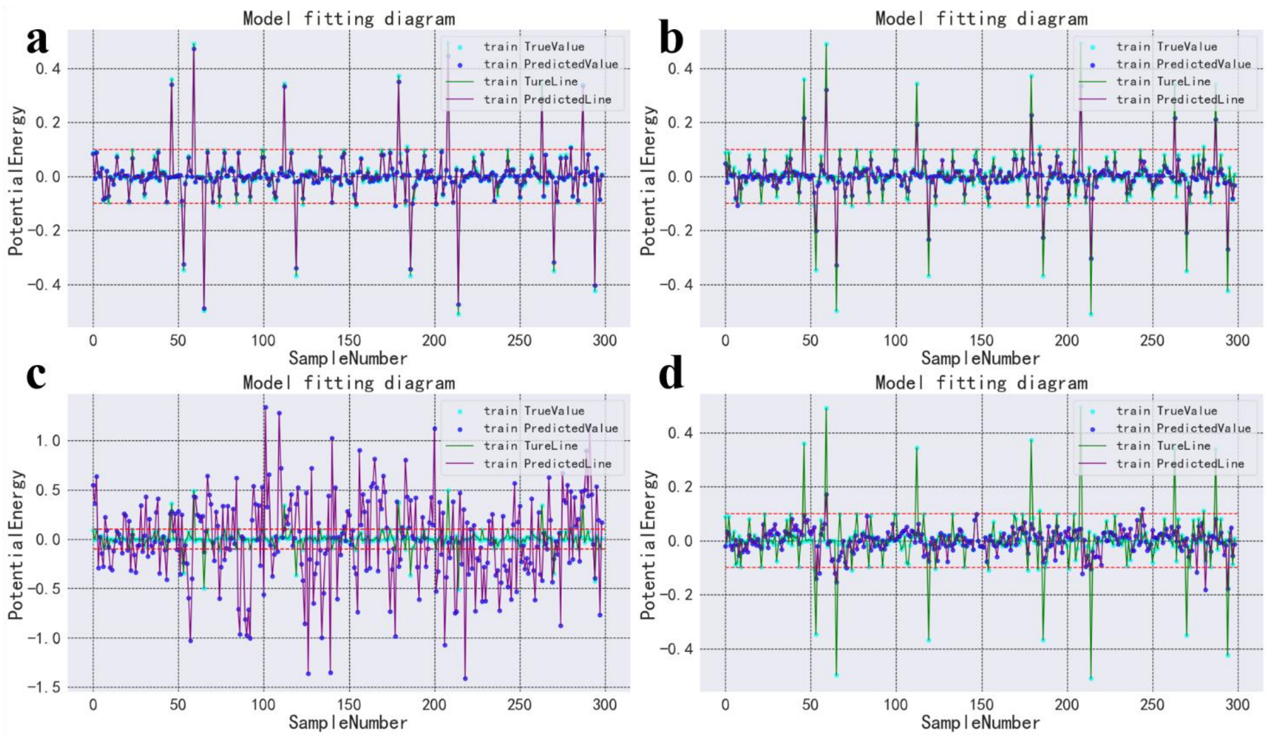


Figure 2. Model Performance of Different Algorithm.

(a. Extreme Gradient Boosting (XGBOOST). b. Random Forest (RF). c Linear Regression (LR). d. Multilayer Perceptron (MLP).)

Based on Figure 2, the following observations can be made: the light blue dots represent the positions of the original data, the green lines represent the distribution of the data, and the purple lines represent the results obtained from fitting the model. The predicted purple dots can be compared with the positions of the original data points. It can be observed that after optimization using the algorithm proposed in this study, the purple lines essentially overlap with the green lines, indicating a close fit between the predicted results and the data distribution. However, other models exhibit relatively poorer performance compared to this model, as shown in Table 5.

Table 5. Comparison of Evaluation Indicators

Model.	RMSE.	MAE.	R2.
DA+XGB	0.9391	0.0215	0.9836
RFR	0.7562	0.0586	0.8185
LR	0.3834	0.0891	0.2123
MLP	0.1867	0.1568	-2.9849

By establishing the RMSE, MAE, and R2 indicators, the performance of the four models can be more intuitively evaluated. The accuracy of the model proposed in this study reached 0.9836, which is the highest score.

In summary, the optimized XGBOOST model performs excellently in terms of both prediction error and accuracy, demonstrating a high degree of fit to the observed values. Therefore, it is evident that this model is suitable for predicting and analyzing the momentum of player "CA."

4.2. Explanation Using SHAP Model

To understand the contribution and relevance of different factors to the performance of players on the field, this study selected a total of 69 factors related to "CA" and "NJ" as samples and calculated their global feature importance. These factors were then sorted based on their values. Figure 4 displays the top 20 factors contributing the most to the momentum of "CA" on the field.

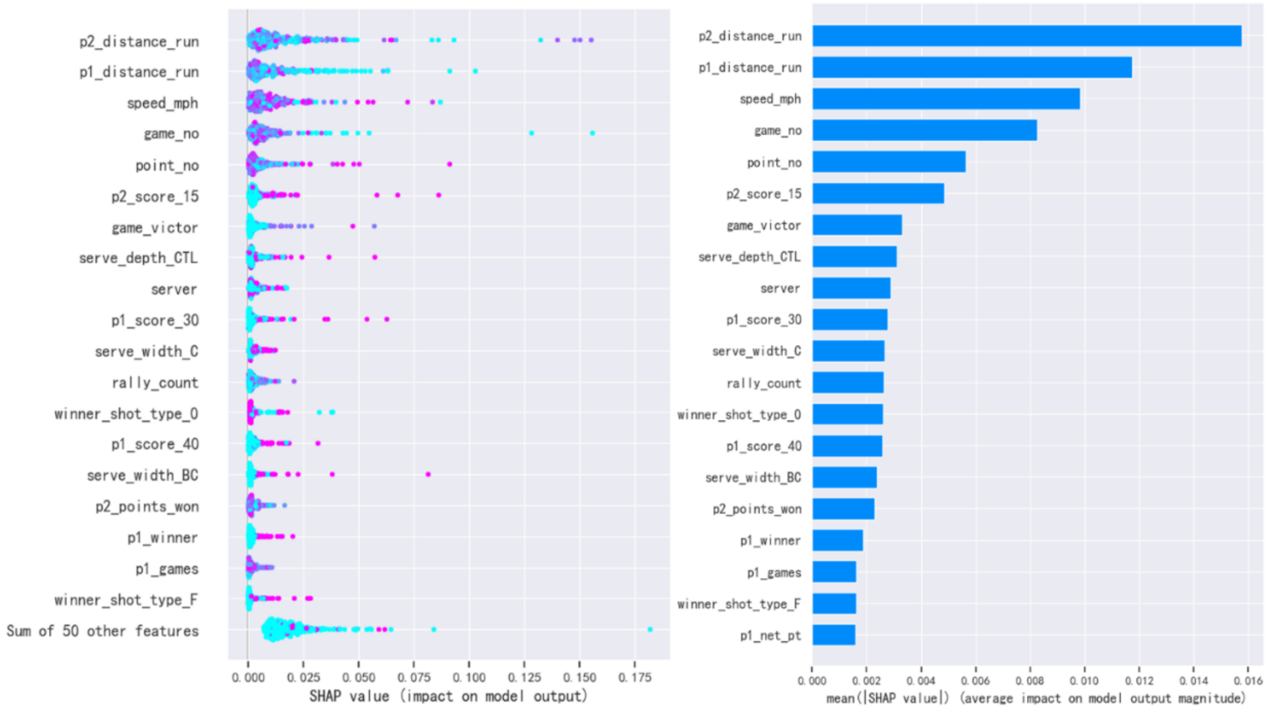


Figure 3. Global Feature Importance Ranking

As shown in **Figure 3**, the larger the global feature importance of a sample, the higher its contribution. It can be observed that the fluctuation in "CA's" momentum is not only related to their performance but also to the situation of the opponent and the choice of strategies. For the sake of clarity, this paper will discuss these 20 factors divided into subjective and objective categories.

Table 6. SHAP Values of Subjective Factors

P1_distance_run	Speed_mph	Serve_width	P1_score
0.011499	0.008425	0.007324	0.005845
P1_winner	P1_games	P1_net_pt	Serve_depth
0.002669	0.002667	0.001847	0.001569

Based on Table 6, among the top four subjective factors, the highest is "p1_distance_run," with a feature importance of 0.011499, followed by the feature importance of ball speed, with a value of 0.008425, "CA's" serve width, with a value of 0.007324, and finally the feature importance of "p1_score," with a value of 0.005845. This indicates that "CA's" contribution to momentum on the field is most significant in terms of running distance, hitting intensity, hitting technique, and cumulative score. Therefore, it can be seen that "CA" possesses good physical fitness, which is his advantage. Additionally, he excels in employing effective hitting techniques and can control the hitting intensity appropriately, demonstrating strong scoring ability. Moreover, "CA's" number of winning points, number of games won in a set, whether "CA" serves into the net, and "CA's" serving distance all contribute to his momentum to some extent.

Table 7. SHAP Values of Objective Factors

Game_Victor	P2_distance_run	Game_no	P2_score
0.02594	0.01368	0.008869	0.008511
Point_no	Rally_count	server	Point_victor
0.007063	0.004345	0.002879	0.002284
P2_Break_pt_missed	P2_winner	Winner_Shot_Type	P2_unf_err
0.001875	0.001884	0.001681	0.001506

In objective factors, as shown in Table 7, the winning player of the current set, the opponent's running distance, the number of games in a set, and the opponent's cumulative score have a significant impact on the player's performance. This suggests that "CA" experiences greater pressure and shows less stable performance when he is at a disadvantage or when there are many games in a set. The cumulative score of all games in a set, the number of consecutive rallies, the serving side, and the player winning the points also have a certain contribution. However, factors such as the opponent's missed opportunities to win points, whether the opponent hits winning shots, and the opponent's unforced errors have a relatively smaller impact.

From both perspectives, it can be concluded that the player "CA" possesses strong endurance and scoring ability. However, he also needs to improve his mental toughness and tactical skills to adapt to matches more quickly or to create more attacking opportunities.

5. Model discussion

The strategy choice is heavily influenced by the selected player, "CA," as evident from this study's analysis process. Results vary with changes in the subject of analysis, highlighting the player-specific nature of offensive and defensive strategies. Additionally, environmental unpredictability adds complexity to strategy selection. The study's significance lies in its ability to develop personalized strategy plans for individual players through its analysis methods. To broaden the application of the optimization strategy model established here across different sports or competitions, validation can be approached from two perspectives:

- Discussion on the universality of model design framework: The modeling process of this study is primarily derived from various scientific theories and mathematical methods, such as the XGBOOST algorithm, to accomplish the task of momentum prediction. Therefore, not only tennis matches but also table tennis can utilize the same modeling approach for problem-solving. Moreover, the SHAP model, as an analysis method not restricted by the playing field, yields highly reliable analysis results.
- Discussion on the universality of model training: Since the dataset used in this study lacks universality, the model built based on this data is also not universally applicable. However, because various sports have their unique competitive characteristics, models need to analyze each feature of each sport specifically, making the modeling process unique. This makes it challenging to apply well-trained models to other domains. However, the research approach and process of this study are universal, allowing it to serve as a reference for exploring strategies in other ball sports.

6. Conclusion

Six indicators reflecting athletes' competition abilities were selected, and their weights were determined using the entropy weight method. These weights were aggregated to formulate the momentum calculation formula. Employing quantitative definitions, we utilized the XGBOOST model optimized with Differential Evolution (DA) to develop a momentum prediction model. Various models, including Random Forest, were also assessed and compared with the optimized XGBOOST model. Results demonstrated that our proposed model achieved superior prediction accuracy, confirming its feasibility and effectiveness. Using the SHAP model, this study analyzed 68 indicators for two athletes to interpret momentum values. Taking the "CA" athlete as an example, the top 20 indicators contributing most to the game were subjectively and objectively analyzed. Findings revealed that the athlete's momentum relies heavily on continuous field running, winning shot production and frequency, appropriate hitting distance, intensity, and ball speed. Additionally, the impact of these factors varies among athletes due to individual differences, suggesting athletes should not only focus on common influencing factors but also consider the unique influence of specific factors on their performance when evaluating their on-field performance.

In conclusion, this study enables athletes to grasp their momentum shifts, pinpoint performance shortcomings in competitions, and craft tailored training plans to enhance their competitiveness and future winning prospects.

References

[1] Ramsay, A. (2023, July 16). Alcaraz Ends the Djokovic Run. Wimbledon. https://www.wimbledon.com/en_GB/news/articles/2023-07-16/alcaraz_ends_the_djokovic_run.html.

[2] Dietl, H. M., Nesseler, C. (2017). Momentum in Tennis: Controlling the Match. *International Journal of Sport Psychology*, 48 (4), 459 – 471. https://doi.org/10.7352/IJSP.2017.48.459.

[3] Ji, S. W., Wang, X. J., et al. (2019). An Application of a Three-Stage XGBoost-Based Model to Sales Forecasting of a Cross-Border E-Commerce Enterprise. *Mathematical Problems in Engineering*, 2019. https: //doi.org/10.1155/2019/8503252.

[4] Gu, Z. Y., Cao, M. C. (2022). Research on Mining Maximum Subsidence Prediction Based on Genetic Algorithm Combined with XGBoost Model. *Sustainability*, 14 (16). https://doi.org/10.3390/su141610421.

[5] Mirjalili, S. (2016). Dragonfly Algorithm: A New Meta-Heuristic Optimization Technique for Solving Single-Objective, Discrete, and Multi-Objective Problems. *Neural Computing & Applications*, 27 (4), 1053 – 1073. https: //doi.org/10.1007/s00521 - 015 - 1920 - 1.

[6] Wang, J. Y., Li, J. R., Li, Z. W. (2022). Prediction of Air Pollution Interval Based on Data Preprocessing and Multi-Objective Dragonfly Optimization Algorithm. *Frontiers in Ecology and Evolution*, 10. https://doi.org/10.3389/fevo.2022.855606.

[7] Nordin, N., Zainol, Z., et al. (2022). An Explainable Predictive Model for Suicide Attempt Risk Using an Ensemble Learning and Shapley Additive Explanations (SHAP) Approach. *Asian Journal of Psychiatry*, 79. https://doi.org/10.1016/j.ajp.2022.103316.

[8] Ziakopoulos, A., Kontaxi, A., & Yannis, G. (2023). Analysis of Mobile Phone Use Engagement during Naturalistic Driving through Explainable Imbalanced Machine Learning. *Accident Analysis and Prevention*, 181. https://doi.org/10.1016/j.aap.2022.106936.

[9] Chen, C. M., Zhang, S. G., et al. (2021). Personalized Travel Route Recommendation Algorithm Based on Improved Genetic Algorithm. *Journal of Intelligent & Fuzzy Systems*, 40(3), 4407 – 4423. https://doi.org/10.3233/JIFS-201218.

[10] Ren, M. L., & Zhang, Q., et al. (2022). Research and Application of Reservoir Flood Control Optimal Operation Based on Improved Genetic Algorithm. *Water,* 14 (8). https://doi.org/10.3390/w14081272.

[11] Li, H., Shi, N. Y. (2022). Application of Genetic Optimization Algorithm in Financial Portfolio Problem. *Computational Intelligence and Neuroscience*. https://doi.org/10.1155/2022/5246309.

[12] Problem Title: Momentum in Tennis. (2024, February 1). Mathmodels. https://www.mathmodels.org/Problems/2024/MCM-C/index.html.