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

Analysis of Momentum in Tennis Matches Based on Ensemble Learning

Huanyu Gong [1], Xuanxuan Guo [2], Zhuohui Gao [3, *]

[1] School of Management, Xi'an Jiaotong University, Xi'an, China, 710049

[2] School of Economics and Finance, Xi'an Jiaotong University, Xi'an, China, 710049

[3] Faculty of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, China, 710049

* Corresponding author: 2211410888@stu.xjtu.edu.cn

Abstract. In tennis, player performance is often attributed to "momentum," pivotal for game control. We analyzed tennis rules, extracting performance indicators such as point differentials, dynamic scoring, and serving advantages. Utilizing a Random Forest model with player performance as the dependent variable, we mapped features to momentum and determined feature importance, achieving an AUC of 0.8796. We then introduced the Comprehensive Ability Volatility Index (CVI) based on factors like serve win rate and breakpoint success rate, identifying match-turning points. To capture these turning points, we integrated momentum, player fitness, and overall strength into an XGBoost model, which yielded an AUC of 0.8716. The model's feature importance output guided concise recommendations for strategic decision-making. Our research offers comprehensive strategies for athletes and coaches to consider momentum, physical fitness, and other match factors, enhancing guidance for daily training and match adjustments.

Keywords: Momentum; Random Forest; CVI; XGBoost.

1. Introduction

Tennis, originating in France in the 12th-13th centuries, has a rich historical background. Evolving into a structured sport with diverse competition rules and court types, it hosts international events such as the Grand Slams, ATP, and WTA tournaments. The Grand Slams—Australian Open, Wimbledon, US Open, and French Open—serve as premier showcases, drawing elite players globally and demanding exceptional technical, tactical, and physical prowess. Coaches play a pivotal role in matches, closely monitoring court dynamics to offer timely guidance and adjustments. They may highlight opponent movements or suggest tactical tweaks, requiring rapid decision-making skills, especially in critical moments. Various factors influence matches, including strength, momentum, and psychological aspects. Understanding these elements aids in predicting and responding to fluctuations, enabling coaches to tailor strategies effectively for athletes.

In tennis, Depken (2022) based the judgment on if "the tendency of an outcome to be followed by a similar outcome is caused by any strategic incentives of the player [1]." Based on this, Momentum is mainly divided into two categories: strategic momentum and psychological momentum, and it is concluded that both strategic momentum and psychological momentum contribute to the outcomes of the best-of-three tennis contest. Elia and Simcha, while criticizing the concept of momentum, argue that the term psychophysiological momentum is fitting in the context of sports competitions [2].

Methods to study factors affecting momentum mainly focus on linear models. Philippe Meier (2019) used a linear probability model (LPM) to test for the existence of momentum, and their results suggest that interruptions terminate the momentum effect [3]. Christoph (2019) gave descriptive statistics and results from their econometric models including robustness checks when studying the impact of psychological traits on performance in sequential tournaments [4]. Ötting (2021) investigated the potential occurrence of change points—commonly referred to as "momentum shifts"—in the dynamics of football matches [5].

We mainly employ the random forest model and XGBoost model to explore momentum and the factors that affect it. Random forest gets the attention while working with the fused datasets [6], and according to the definition of momentum, using the random forest model can better solve related problems. Sagi (2021) presented a novel method for transforming a decision forest of any kind into an interpretable decision tree, showing that in some cases the generated tree can approximate the predictive performance of a XGBoost model while enabling better transparency of the outputs [7].

In reviewing prior literature, we observed a scarcity of considerations for momentum concepts, making it difficult to encompass various indicators of tennis matches comprehensively. Furthermore, there has been limited progress in quantifying momentum due to the complex and nonlinear relationships among different momentum features and their interactions. Conventional linear models struggle to provide accurate descriptions of momentum. Hence, we employed Ensemble Learning models for modeling and description purposes. By constructing a Random Forest model, we could better capture these complexities and correlations. Lastly, considering other factors influencing match trajectories, we offer predictions for match-turning points.

2. The basic introduction to two Ensemble Learning models

2.1. The Structure of Random Forest

Ensemble Learning, by amalgamating predictions from multiple base models, enables better capture of underlying patterns within data, thereby enhancing model generalization. It mitigates the risk of overfitting associated with individual models, reducing both bias and variance and consequently improving overall prediction accuracy. Ensemble Learning has garnered widespread adoption in practical applications.

The final prediction result is determined through a voting mechanism (classification problem) or averaging (regression problem). Assume that there is a weak learner $g_t(x)$, and $G(x)$ is an integrated learner, which satisfies:

$$\mathrm{G(x)} = \mathrm{uniform}\{\mathrm{g_t(x)}\},\ \mathrm{t} = 1,2,\cdots,\mathrm{T} \tag{1}$$

Then the expectation of the variable x satisfies:

$$Avg\big(E_{exp}(g_t)\big) = \mathrm{Avg}(\exp(\mathrm{g}_t - \mathrm{G(x)})^2) + \mathrm{E}_{exp}(\mathrm{G(x)}) \geq \mathrm{E}_{exp}(\mathrm{G(x)}) \tag{2}$$

It can be concluded that ensemble learning can significantly reduce the actual performance error of a single learner.

Random Forest is an Ensemble Learning model. It is built based on Decision Trees and consists of multiple Decision Trees. An important feature of Random Forest is that it can reduce the over-fitting of Decision Trees due to over-fitting data, thus improving the performance and stability of the model.

Decision Trees serve as the base learners for the Random Forest model. A Decision Tree is a tree-like discriminative model, named so because the branching decision-making process resembles the branches of a tree. Decision Trees exhibit excellent performance in capturing complex relationships such as non-linearity. The decision tree algorithm adopts a tree structure and uses layer-by-layer reasoning to realize the final classification. A Decision Tree consists of the following elements:

(1) Root node: a complete set containing samples

(2) Internal nodes: corresponding feature attribute test

(3) Leaf node: represents the result of decision.

When forecasting, a certain attribute value is used to judge at the internal node of the tree, and according to the judgment result, it is decided which branch node to enter until it reaches the leaf node, and the classification result is obtained. An example of a Decision Tree is shown in Figure 1.

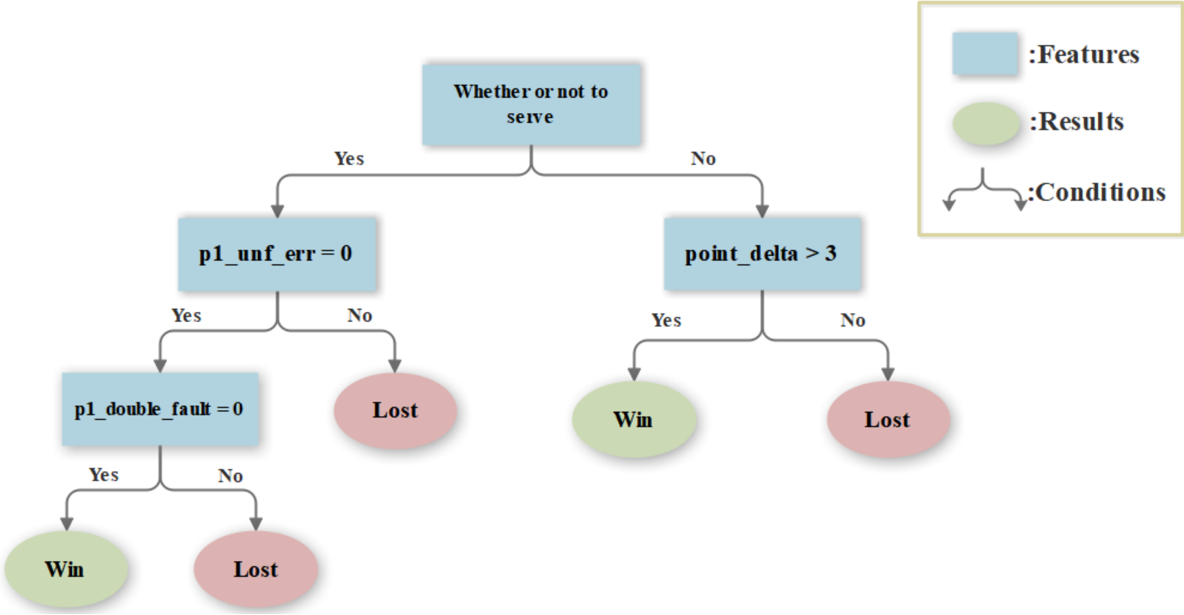

Figure 1. Decision Tree

Random Forest is composed of many Decision Trees, and there is no correlation between different decision trees. When we carry out the classification task, as a new input sample enters, each Decision Tree in the forest will be judged and classified separately, and each decision tree will get its classification result. Finally, the final prediction result will be determined by voting mechanism (classification problem) or average value (regression problem). A schematic diagram of a Random Forest is shown in Figure 2.

In Random Forest, the importance of each feature variable is measured by observing the split contribution of each feature on the Decision Tree node:

Firstly, for each Decision Tree, the contribution of each feature in the tree splitting is calculated by measuring the degree to which each feature reduces the impurity (for example, Gini index or information gain) at the splitting node.

Then, in the Random Forest, the overall importance of each feature can be obtained by averaging or weighted averaging the contributions of all trees.

Finally, the calculated importance values are standardized to get the weights of each feature variable.

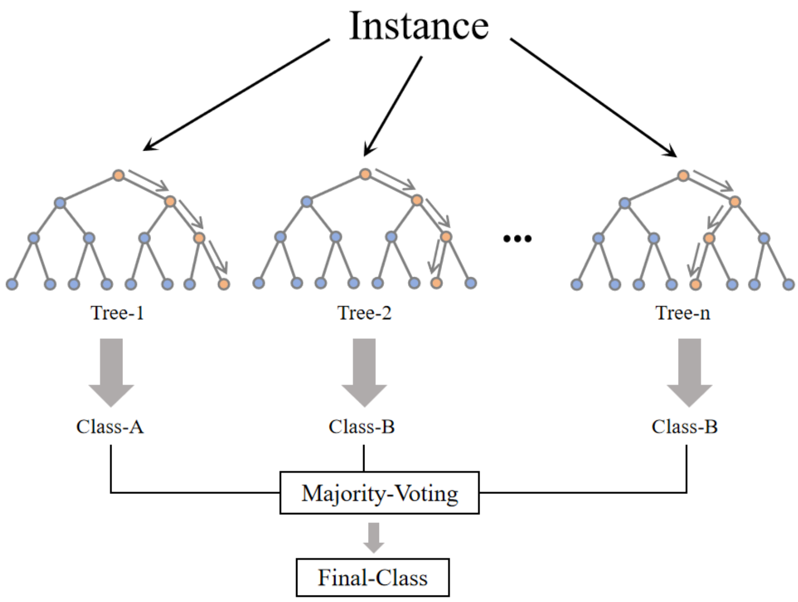

Figure 2. Random Forest

Overall, as an Ensemble Learning model, Random Forest demonstrates excellent performance, making it excel in many machine learning tasks. By integrating specific model interpretation methods, the Fandom Forest can further investigate the relationship between features and labels.

After a thorough understanding of tennis match rules and related metrics, we have identified the following quantities that can reflect the momentum status of players during the match. While there is no universally accepted formula for momentum in a tennis match, we can still incorporate these metrics into the definition of momentum, as follows:

(1) Point difference: The difference in points won by two players within a certain period.

(2) Break points saved and won: The ability to win points when at a disadvantage or seize opportunities.

(3) Advantages in serving: The winning rate of serving reflects the efficiency of serving.

(4) Return game performance: The percentage of winning return games reflects the effectiveness of the player's return.

(5) Winning streak: The number of consecutive winning points, games, or sets, showing the player's current advantage over the opponent.

(6) Ace: The number of serving aces, which can provide insight into the serving performance and pressure handling.

(7) Past: Whether the player scored in the last point

(8) Serve: Whether the player is the serving side on the court

(9) Double fault: the serving player makes two mistakes in the same point

(10) Unforced error: A basic error committed in a tactical game

These indicators will help to establish a comprehensive momentum assessment model, enabling us to more accurately quantify and understand a player's performance in the match. We provide specific calculations for each indicator as follows.

These indicators will help to establish a comprehensive momentum assessment model, enabling us to more accurately quantify and understand a player's performance in the match. We provide specific calculations for each indicator as follows.

Point difference (PD): the difference in points won by two players in a certain period of time can be used to measure the advantage in the game.

$$PD = PointsWonByPlayer1 - PointsWonByPlayer2 \tag{3}$$

Guaranteed serve and break: the ability of a player to save the break point in the face of being broken by an opponent (BPS), the ability of a player to win points in the opponent's serve and successfully break.

Advantages in serving: the percentage of a player winning their own service games, reflecting their serving efficiency.

$$SGD = \frac{SGW}{TSG} \times 100\% \tag{4}$$

Where SGW is ServiceGamesWonbyPlayer and TSG denotes TotalServiceGames.

Performance in return games: the percentage of the opponent's service games won by the player, indicating the effectiveness of the player's return of serve.

$$RGP = \frac{RGWP}{TRG} \times 100\% \tag{5}$$

Where RGW is ReturnGamesWonbyPlayer and TRG represents TotalReturnGames.

To mitigate potential overfitting risks, following discussions and consultations with domain experts, we have incorporated the variables Past, PD, Serve, Double fault, and Unforced error into the model's feature set, thereby defining momentum (M) as follows:

$$M = \mathcal{M}(\mathrm{Past}, \mathrm{PD}, \mathrm{Serve}, \mathrm{Doublefault_i}, \mathrm{Unforcederror_i}), \qquad i = 1,2 \tag{6}$$

We utilized the variables as the independent variables for the model input, while the outcome of winning the match was designated as the dependent variable, thus formulating a binary classification problem based on the random forest model.

2.2. The Structure of XGBoost

As another Ensemble Learning algorithm, XGBoost adopts a Boosting algorithm, which differs from the Bagging algorithm used by Random Forest. The Boosting algorithm is an iterative Ensemble Learning method that sequentially trains a series of weak learners (in this case, Decision Trees), with each iteration attempting to correct the errors of the previous iterations [7]. In each iteration, the Boosting algorithm adjusts the weights of the samples based on the performance of the previous models, placing more emphasis on samples that were previously misclassified, thereby continuously improving the model performance. Both Boosting and Bagging algorithms can reduce the variance of the model and prevent overfitting by combining the predictions of multiple models, thus enhancing the stability and generalization ability of the model. A schematic diagram of XGBoost is shown in Figure 3.

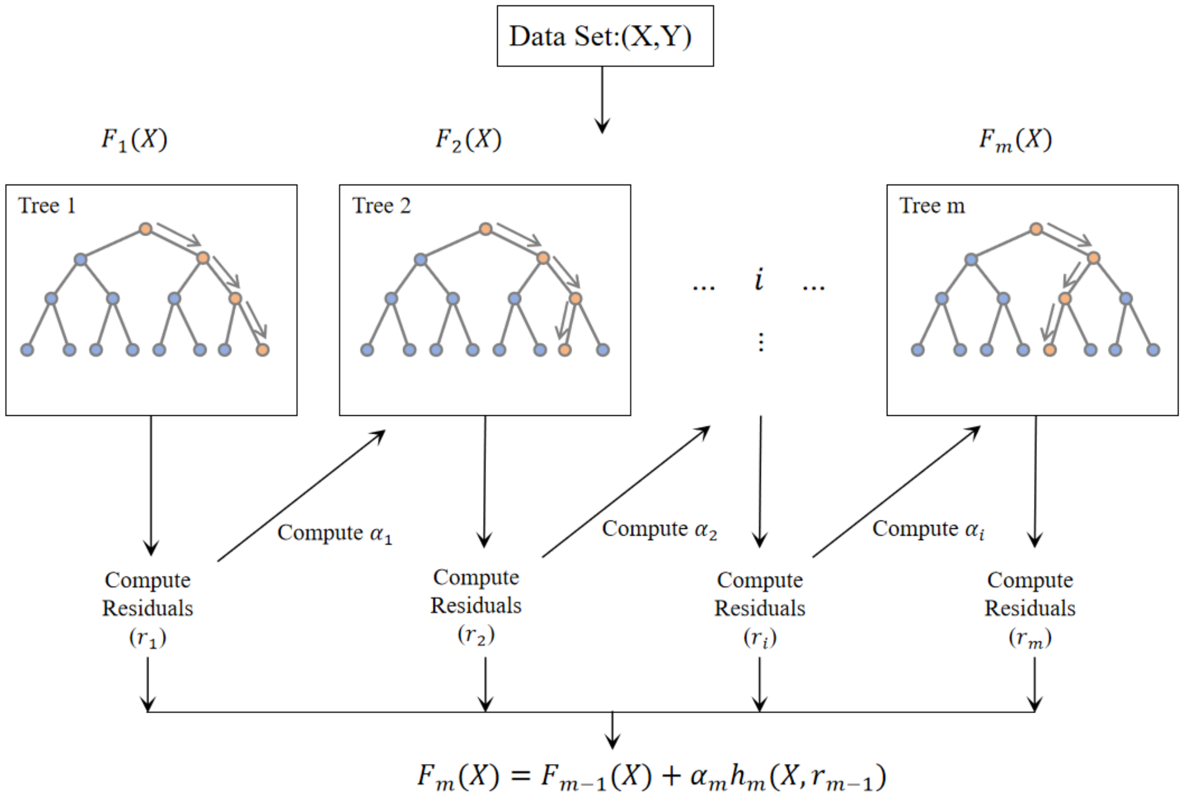

Figure 3. XGBoost

As one of the most representative models in the field of Ensemble Learning, XGBoost is an Ensemble Learning algorithm based on the gradient boosting framework. XGBoost employs second-order Taylor expansion on the cost function to obtain first-order and second-order derivative information, which is crucial for optimizing the training process, enhancing model performance, and improving convergence speed. Moreover, XGBoost introduces regularization terms into the cost function to control the model's complexity, which, from the perspective of balancing variance and bias, reduces the model's variance and prevents overfitting. Techniques such as shrinkage further contribute to the powerful training and prediction performance of the XGBoost model. In each iteration, XGBoost trains a new decision tree, attempting to correct the residuals (the difference between the actual values and the current model's predicted values) from the previous iteration of the model. Through multiple iterations, multiple decision trees are gradually combined, ultimately forming a powerful ensemble model adept at handling structured data and feature engineering. Additionally, XGBoost can provide an assessment of the importance of each feature in the model, allowing us to better understand which features play a critical role in the model's performance.

The given objective function of XGBoost is as follows:

$$\mathcal{L}^{(t)} = \sum_{j=1}^{T} [G_j w_j + \frac{1}{2}(H_j + \lambda) {w_j}^2] + \gamma T \tag{7}$$

The objective function for each leaf node j is then:

$$f(\mathrm{w_j}) = \mathrm{G_j w_j} + \frac{1}{2}(H_j + \lambda) {w_j}^2 \tag{8}$$

It is a quadratic function of $\mathrm{w_j}$.

The definition of tree complexity is:

$$\Omega(f_t) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^{\,2} \quad (9)$$

Where G_j is the sum of the first-order partial derivatives with respect to the leaf node j, considering all samples it contains, is a constant, H_j denotes the sum of the second-order partial derivatives with respect to the leaf node j, considering all samples it contains, is a constant, and w_j: represents the weight vector of leaf node j.

3. Results

3.1. Analysis of Momentum Based on Random Forest

Through the construction of decision trees, we established a mapping from independent variables to the dependent variable, and utilized the resultant tree-based model as the definition of momentum. Subsequently, upon obtaining the model, we invoked commands to display the feature importance and evaluation metrics of the model (accuracy, recall, precision, F1 score, and AUC value) as shown below:

Using the random forest model, we obtained the feature importance rank as shown in Table 1.

Table 1. Feature Importance

Future	Past	PD	Serve	DF1	DF2	UE1	UE2
Importance	0.0217	0.1773	0.2464	0.0169	0.0155	0.2421	0.2800

In Table 1, PD, Serve and unforced errors are more important in the process of quantifying momentum. The feature importance reflects the magnitude of the impact of independent variables on the dependent variable. From the model's outcomes, variables such as PD and UE exhibit the greatest influence on winning. Generally, if a player gains a lead in past matches, it is likely to generate a stronger momentum. Additionally, unforced errors by both oneself and the opponent are critical factors influencing the match's trajectory. For professional players, technical mishaps by oneself or the opponent tend to be amplified during the match. Minor errors can lead to significant fluctuations in both the player's and the opponent's mental states. The ranking of feature importance reminds us that variables such as PD carry substantial weight in the function mapping created by our tree model, while the influence of other variables is relatively minor.

The evaluation indicators of the model are listed in Table 2.

Table 2. Evaluation Indicators

Evaluation index	Accuracy	Precision	Recall	F1 Score	ROC AUC
Value	0.8202	0.8040	0.7680	0.7856	0.8796

The various evaluation metrics of the model are shown in Table 2. Each evaluation index is significantly greater than 0.5 and relatively close to 1, indicating a well-fitted model. This suggests that our defined momentum can effectively predict the outcome of the match, ensuring the validity of the model. On the one hand, the model demonstrates excellent generalization performance on the partitioned test set. The outstanding performance of the model ensures its effectiveness, indicating the reasonableness of our feature selection and modeling processes [10]. On the other hand, we derive significant indicators influencing player performance through feature importance ranking, laying the groundwork for further research into the importance of features in match turning points.

3.2. Comprehensive consideration of various indicators demonstrates XGBoost's predictive capability for game-turning points

Firstly, we will present our definition of game-turning points.

To determine whether a certain point is a turning, we need to consider the relative value of the volatility of a player's performance.

To quantify the volatility of player performance and formulate it into a standard mathematical form, here is a mathematical description of each of the previously discussed metrics:

Assume that a certain indicator X (such as winning rate, score difference, etc.) of a player in N match has a value of X_i in the ith point, and the volatility of this indicator is V_x.

$$V_x = X_i - \bar{X}, \tag{10}$$

Where $\bar{X}$ is the average value of X_i.

Specific indicators are shown in Table 3.

Table 3. Indicator Symbol Description

Indicator	Symbol
service scoring rate volatility	V_{serve}
break success rate volatility	V_{break}
double fault volatility	V_{faults}
direct score volatility	V_{aces}
first serve success rate volatility	$V_{\text{first serve}}$
opponent level volatility	$V_{\text{opponent level}}$
win point volatility	$V_{\text{winning points}}$
losing point volatility	$V_{\text{losing points}}$

Assuming we have m different performance indicators, each with volatility V_1, $V_2, \dots, V_m$, the composite volatility index (CVI) can be calculated in the following formula:

$$\text{CVI} = \frac{\sum_{j=1}^{M} V_j}{M}, \tag{11}$$

Where V_j represents the volatility of the jth performance indicator.

CVI provides a comprehensive view that quantifies the stability of a player's performance.

Assume that when both players perform stably, that is, when there is no fluctuation (CVI=0), the flow of play does not exist and is quantified as 0. When fluctuations occur, if a player's CVI>0, it means that his performance is moving towards Fluctuations go in the better direction, and in the opposite direction the worse. In order to determine whether a given point is a turning point, we need to calculate the $\text{CVI}_l (l = 1,2)$ of the two players respectively, and make a difference between them. The difference between the two is noted as $d = \text{CVI}_1 - \text{CVI}_2$. If $d > 0$, it indicates that the flow of the play now favors player1; if $d < 0$, then it indicates that the flow of the play favors player 2.

To better illustrate the fluctuations in the match and demonstrate the match-turning points as defined by us, we simulated the match fluctuations between two players, as illustrated in Figure 4. In Figure 4, the two curves represent the fluctuations exhibited by the players during the match, as determined by the CVI. Finally, we identified the match-turning points using the intersection of the two performance curves. In Figure 4, the red shaded area represents the difference in performance between the two players, providing supplemental clarification for our definition of match-turning points.

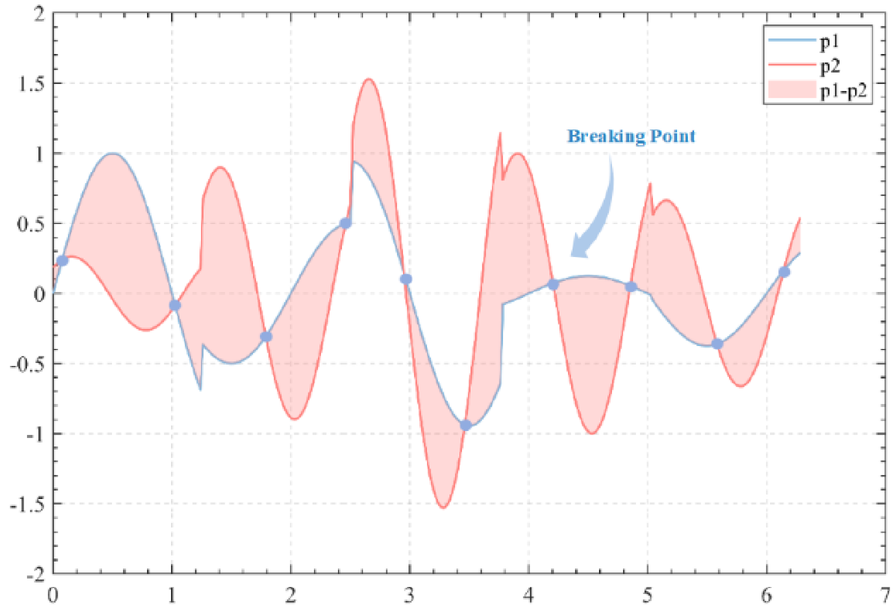

Figure 4. Turning Judgment

After modeling, to evaluate the performance of the classification model, we plotted the ROC curve and calculated the AUC value, as shown in Figure 5. The area under the ROC curve (AUC) is the integral area below the ROC curve, where a larger AUC value indicates better model performance to some extent [8] [9]. We observed that the model achieved an AUC value of 0.84, and the ROC curve exhibits a nearly horizontal trend when the false positive rate exceeds 0.3, indicating that our model performs well.

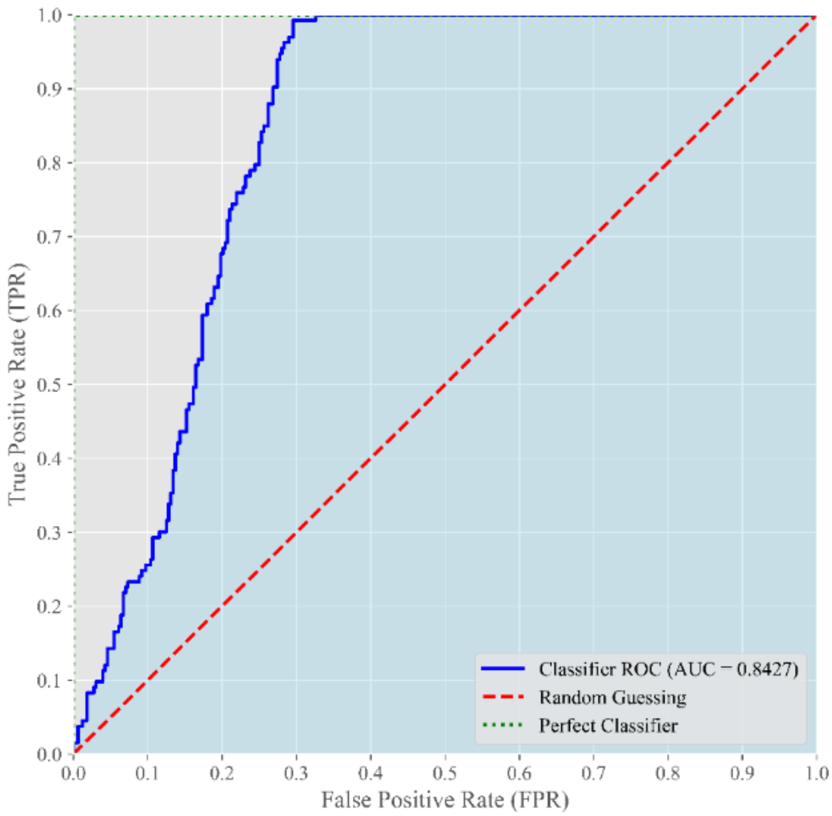

Figure 5. ROC Curve

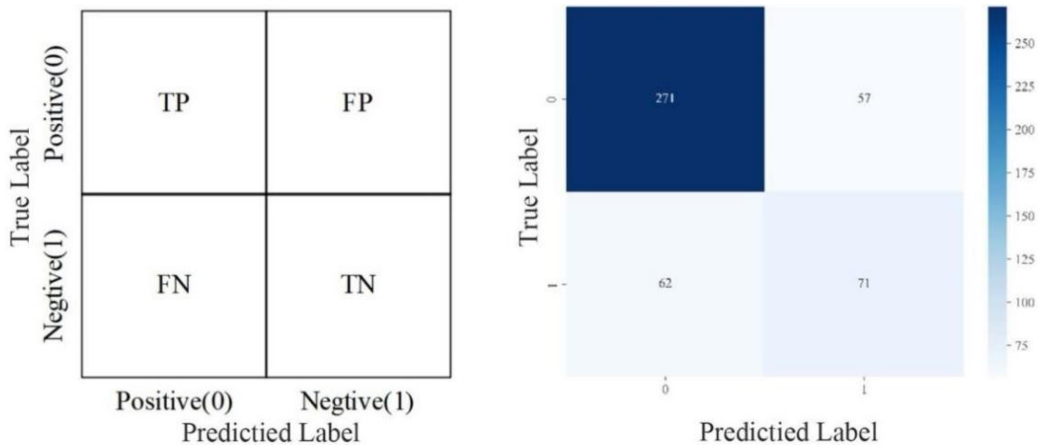

Figure 6. Confusion Matrix

In predictive analytics, a confusion matrix is a table with two rows and two columns that reports the number of true positives, false negatives, false positives, and true negatives. To gain a clearer insight into the classification performance of the model, we also plotted the confusion matrix, as depicted in Figure 6. The matrix indicates that the proportion of TP and FN accounts for approximately 74.1%, implying a correctness rate of 74.1% for the model. This suggests that our model has achieved a satisfactory performance, although there is still room for improvement.

4. Conclusions and outlooks

In our study, to analyze the complex definition of player momentum in tennis matches and subsequently predict match volatility, we adopted an Ensemble Learning approach to establish a mathematical analysis model. For the definition of momentum, we initially selected factors that could potentially influence momentum from existing research and indicators. We then established a random forest model with the players' actual performance as the dependent variable, mapping factors such as historical scores and unforced errors to momentum. The model's AUC value was 0.8796, accuracy was 0.8202, and the F1 score and other metrics also indicated good model performance. Furthermore, we provided the feature importance ranking of the model, revealing that factors such as historical scores carried significant weight in player momentum. To capture match volatility, we first defined match turning points using the CVI index and established an XGBoost model. Factors such as player momentum, individual strength, and physical fitness were incorporated into the model. The AUC value of the model was 0.8427. Additionally, we presented the confusion matrix of the model, illustrating the distribution of correctly predicted and misclassified samples. Overall, our conclusion provides an important reference for pre-match preparation and in-match strategy for players and coaches, as it employs tree models rather than conventional linear models to define momentum. This contributes to our understanding of the concept of momentum in sports competitions.

References

[1] Depken C A, Gandar J M, Shapiro D A. Set-level strategic and psychological momentum in best-of-three-set professional tennis matches[J]. Journal of Sports Economics, 2022, 23 (5): 598 - 623.

[2] Morgulev E, Avugos S. Beyond heuristics, biases and misperceptions: the biological foundations of momentum (hot hand) [J]. International Review of Sport and Exercise Psychology, 2023, 16 (1): 155 - 175.

[3] Meier P, Flepp R, Rüdisser M, et al. Investigating the conditions for psychological momentum in the field: Evidence from men’s professional tennis [J]. Available at SSRN 3438123, 2019.

[4] Bühren C, Steinberg P J. The impact of psychological traits on performance in sequential tournaments: Evidence from a tennis field experiment [J]. Journal of Economic Psychology, 2019, 72: 12 - 29.

[5] Ötting M, Langrock R, Maruotti A. A copula-based multivariate hidden Markov model for modelling momentum in football [J]. AStA Advances in Statistical Analysis, 2021: 1 - 19.

[6] Sheykhmousa M, Mahdianpari M, Ghanbari H, et al. Support vector machine versus random forest for remote sensing image classification: A meta-analysis and systematic review [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 13: 6308-6325.

[7] Sagi O, Rokach L. Approximating XGBoost with an interpretable decision tree [J]. Information sciences, 2021, 572: 522 - 542.

[8] Janssens A C J W, Martens F K. Reflection on modern methods: Revisiting the area under the ROC Curve [J]. International journal of epidemiology, 2020, 49 (4): 1397 - 1403.

[9] Muschelli III J. ROC and AUC with a binary predictor: a potentially misleading metric [J]. Journal of classification, 2020, 37 (3): 696 - 708.

[10] Zhang Y, Liao Q V, Bellamy R K E. Effect of confidence and explanation on accuracy and trust calibration in AI-assisted decision-making [C]//Proceedings of the 2020 conference on fairness, accountability, and transparency. 2020: 295 - 305.