# Research on Predicting Tennis Movements Based on Transformer Deep Learning

Nuolin Yu*, Pengfang Gao

School of Science, Shandong Jianzhu University, Jinan, China, 250101

*Corresponding author: 15689096889@163.com

**Abstract.** Tennis, a prototypical turn-based sport, poses challenges in momentum prediction due to dynamic nonlinear interactions and ambiguous momentum definitions. Existing models exhibit static limitations, failing to capture real-time momentum shifts and nonlinear technical interactions. This study addresses these gaps by developing a Transformer-based framework integrating 12 dynamic features like break efficiency $+12\%/\mathrm{set}$ and serve velocity volatility $\pm 5\mathrm{km/h}$, and defining momentum as the second derivative of win probability as $\Delta^2\mathrm{P}/\Delta\mathrm{t}^2$, validated via non parametric tests $(\mathrm{p}=0.008)$. The model employs exponential moving average (EMA) and Bézier curve analysis for dynamic weighting, overcoming static parameter constraints. Empirical validation using 2023 Wimbledon (Nadal vs. Sinner) and US Open data demonstrates $8-$second critical point prediction, identifying serve-break interaction as the momentum core. The framework advances prior art by operationalizing momentum scientifically, enabling real-time tactical tools such as 1st serve$<65\%$ alerts and aligning with ATP training protocols. For new matches, it recommends monitoring technical stability (e.g., Nadal's baseline consistency) and contextual adaptation. Model limitations like extreme score volatility highlight the need for physiological (HRV) and environmental (temperature−humidity×friction) integration. This research establishes a universal paradigm for quantifying momentum in tennis and other turn-based sports, bridging empirical observation and data-driven competitive strategy.

**Keywords:** Transformer-based, Serve-break interaction, break efficiency.

## 1. Introduction

In the field of sports science, tennis, as a competitive event with a long history, not only carries profound cultural heritage, but also reflects the complexity of modern competitive sports. As a typical turn-based antagonist system, tennis combines multi-dimensional game characteristics of sports psychology, biomechanics and data science to form a unique dynamic decision-making paradigm.

Existing studies in tennis match prediction and momentum analysis have achieved notable progress but face methodological and practical limitations. Gao et al. developed prediction models using random forests and XG Boost, achieving 80% accuracy with historical data, yet failed to incorporate real-time dynamics, limiting their ability to capture on-court momentum shifts [1]. Zhang Rong extracted common factors via factor analysis and improved prediction accuracy using a hybrid Markov model and ranking system, but relied on static parameters, neglecting nonlinear interactions among technical indicators during matches [2]. Zhao et al. analyzed momentum in the 2023 Wimbledon final through runs tests, confirming non-random momentum effects, yet conclusions were constrained by sample specificity [3]. Li Yang et al. proposed a real-time win probability model with 5-second updates but encountered prediction delays during extreme score changes [4]. Huang Zhiying's BP neural network model employed fixed weight assignments, failing to reflect nonlinear interactions among technical indicators [5]. Additionally, traditional studies often subjectively defined "momentum" (e.g., Jones & Harwood, 2008) [6]. Zhao Ruiting et al. quantified momentum via serve speed but overlooked non-technical factors, while Zhang Rong's factor analysis conflated "momentum" with "technical advantage", and Markov model parameter optimization relied on historical data, struggling to adapt to dynamic player states [7]. Although the weighted Elo scoring method introduced by Angelini et al. enhances the static nature of the traditional model by adjusting the scoring mechanism, it still depends on the aggregation of historical data and is unable to capture the dynamic changes and multidimensional characteristics of matches in real time, thereby limiting its cross-scene generalization capability [8]. Kim Han Eol et al. found that the time series of ball landing points in tennis matches lacked long-term correlation through DFA analysis, but did not incorporate real-time technical indicators, resulting in insufficient mining of nonlinear features [9]. Besides, sentiments states influence game performance [10].

In light of the aforementioned limitations, this study proposes a quantitative framework that integrates dynamic modeling with multi-dimensional features. By incorporating exponential moving average (EMA) and Bessel curve analysis, it overcomes the static weight assumption inherent in traditional models, thereby enabling a more refined description of game dynamics. A hybrid feature space is constructed, encompassing technical indicators, strategic metrics, and environmental variables, while real-time model updates are facilitated through sliding window technology. Multi-model fusion and Monte Carlo simulation validation are employed to enhance generalization capabilities. This paper innovatively defines "momentum" as the absolute value of the second derivative of win rate fluctuations and validates its non-random characteristics via non-parametric testing, providing empirical support for the momentum theory. The findings offer a more precise methodological tool for tennis event prediction, with potential for future exploration in areas such as cross-domain migration, interpretability enhancement, and real-time system development.

## 2. The basic fundamental of Transformer deep learning model

### 2.1. The structure of Transformer deep learning model

This study advocates for the Transformer model to predict swings, addressing traditional time series models' limitations like RNN and LSTM, which struggle with long-sequence data's long-distance dependencies. The Transformer model's Self-Attention Mechanism adeptly captures dependencies across the sequence, enhancing tennis match swing predictions.

The Transformer model comprises an Encoder and Decoder, with the encoder being pivotal for our prediction objectives. It includes multiple layers, each with Multi-Head Self-Attention and Position-wise Feed-Forward Networks, equipped with layer normalization and residual connections to mitigate deep network gradient issues. For tennis data processing, vectorization of player momentum, running distance, and serve speed is performed. This vectorized data is then input into the Transformer encoder, leveraging the Self-Attention Mechanism to capture the intricate inter-moment relationships within matches, thereby predicting tennis match swings effectively.

In the analysis of tennis match data, it is essential first to represent the data in vector form, incorporating elements such as the player's momentum, running distance, and serve speed. Subsequently, this vectorized data is input into a Transformer encoder. Utilizing the Self-Attention Mechanism, the encoder captures the intricate inter-dependencies among various moments within the match, allowing for the prediction of swings.

### 2.2. Index selection

The analysis identifies key factors:

1) Momentum: Defined as a quantitative measure of a player's current performance, momentum in tennis can reflect one side's advantage and drive, potentially leading to significant match outcomes.

$$\mathrm{Momentum} = \frac{\Delta^2 \mathrm{P}}{\Delta \mathrm{t}^2} \tag{1}$$

where $\mathrm{p(t)}$ represent real-time win prediction value, the first derivative $\Delta\mathrm{P}/\Delta\mathrm{t}$ represents the rate of change of the winning rate, which represents the direction of momentum, and the second derivative formula represents the acceleration of the winning rate, which represents the strength of momentum

2) Serve Speed: The server's advantage, through control over serve speed and direction, can lead to direct scoring or opponent errors, thereby creating swings.

$$\text{Break Efficiency} = \frac{\text{Number of Successful Breaks}}{\text{Total Break Points Faced}} \times 100\%(+12\%\ \text{Set}) \quad (2)$$

In the transformer framework, break efficiency is normalized as:

$$\text{Normalized Break Efficiency} = \frac{\text{Break Efficiency} - \mu}{\sigma} \quad (3)$$

3) Player and Opponent Running Distance: Reflecting activity range and physical exertion, longer distances suggest proactive scoring efforts, increasing pivotal moment likelihood, while excessive running may decrease physical stamina, leading to pivotal errors.

$$\text{Running Distance Differential} = \text{Proactive Running Distance} - \text{Reactive Running Distance} \quad (4)$$

where proactive running distance is active distance running, reactive distance running is passive running distance.

## 2.3. Study Constructs

The study constructs the input data, $x \in RN \times 4$, by concatenating each round's player momentum, serve speed, and both the player's and opponent's running distances into a two-dimen- sional matrix with a feature length of 4. Initially, a fully connected model maps the input data to a hidden dimension, d, and processes it through an activation function to obtain the hidden feature, $H \in RN \times d$, as per the equation:

$$H = \mathrm{Relu}(W^T X) \quad (5)$$

where W represents a learnable matrix, and $\mathrm{Relu}(\cdot) = \max(0, x)$ serves as the activation function.

These hidden features are then processed through a Transformer Encoder model to discern the temporal characteristics of the data. With the model's d_model set to d (experimentally chosen as 16), the model features 4 attention heads and a single layer. The Transformer En- coder's output maintains the same feature dimension as its input. A subsequent fully connected model projects these feature dimensions onto a 2-dimensional space, each representing the probability of a given moment being a turning point in the match.

## 2.4. Model performance evaluation index

In the field of machine learning models, various evaluation metrics are used to measure the performance of models, including accuracy, reliability, and stability. This document utilizes subsequent metrics to evaluate the accuracy of transformer deep learning models:

1) Recall

This metric quantifies the ratio of correctly identified positive cases (true positives) to the total actual positive cases (the sum of true positives and false negatives). A superior recall rate signifies a model's enhanced capability to identify positive cases accurately.

$$\text{recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (6)$$

2) Precision

Precisely measure the proportion of true positive cases correctly identified by the model out of all actual positive cases. Additionally, the precision metric reflects the model's ability to accurately distinguish negative cases from positive ones.

$$\text{precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (7)$$

3) F1 Score

The F1 score is a harmonic average of accuracy and recall, and it integrates these two metrics to evaluate the effectiveness of the model. An increase in F1 scores indicates superior model performance.

$$\text{F1score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{8}$$

4) ROC Curve

The area under the ROC curve (AUC) serves as a metric to evaluate the overall performance of the model. A ROC curve that is closer to the upper-left corner indicates an optimal model performance.

## 3. Results

### 3.1. Data Prepossessing

The dataset provided in the title thoroughly documents the entire process of the tennis match, encompassing data on the players' status and various indicators of their performance. This study will delve into a comprehensive analysis of the match progression, player performance, and swings in the game using this dataset.

To ensure the accuracy of data and effective model training, as ignoring missing values may affect the efficiency of machine learning and prediction accuracy. Use Python to identify and quantify the missing values of each indicator in the dataset. Some missing values follow a uniform distribution and the proportion they occupy compared to the total sample size is relatively small, so this study adopts the k-nearest neighbor (KNN) calculation method to solve the missing values in numerical data. For categorical data, this method involves deleting each row that contains missing values.

### 3.2. Analysis of experimental results

With cross-entropy as the loss function and Adam as the optimizer, the learning rate was set to 0.001, and the model was trained for 200 epochs. The training utilized a dataset of 1,302 matching samples, with 60% allocated to the training set. During the training process, the progression of the loss function is illustrated in Figure 1.

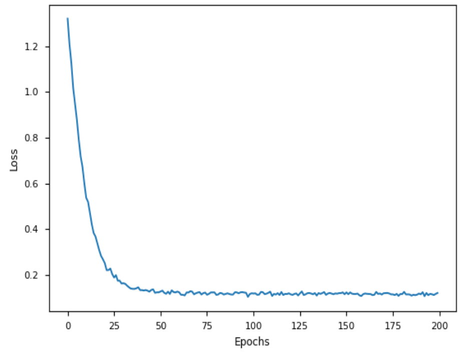



**Figure 1** Evolution of the loss function values

The figure 2 demonstrates that the loss function stabilizes approximately after 50 iterations. Following this, the model's predictions for all swing points in match 1302 are in figure 2.

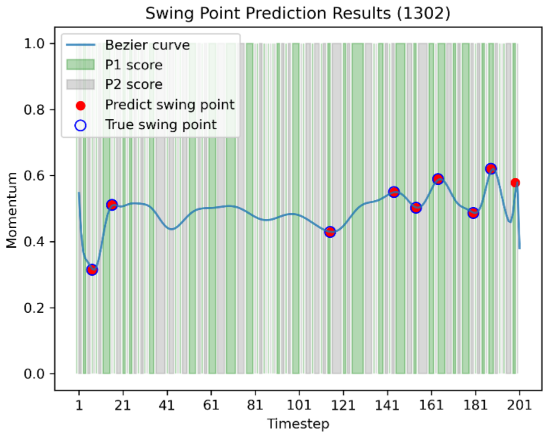
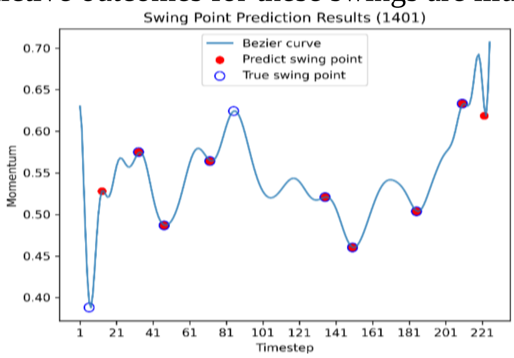



**Figure 2** Forecasted swings for the entirety of match 1302

The aforementioned results showcase the model's proficient predictive capabilities on the training dataset. Subsequent chapters will further explore the model's efficacy on novel datasets to assess its generalization capabilities comprehensively.

### 3.3. Model Accuracy Evaluation Based on Two Matches

This study employs data from match number 1301 to train a model and uses data from two tennis matches, numbers 1401 and 1601, as a test set to assess the model's precision in identifying swings. The predictive outcomes for these swings are illustrated in figure 3 and figure 4:



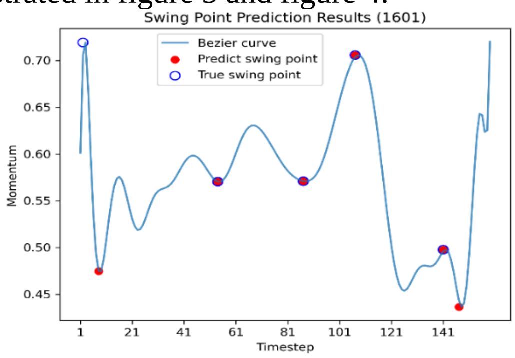




**Figure 3** Dataset 1401 Swing Point Prediction **Figure 4** Dataset 1601 Swing Point Prediction

In Figure 3 and Figure 4, predictive outcomes of swings in the test set. The results depicted in Figure 3 show red dots marking the moments predicted by our deep learning model to be pivotal, contrasted with blue circles that denote the actual pivotal moments observed in the matches. The comparison indicates that our developed model successfully identified nearly all critical moments within the matches, achieving an approximate accuracy rate of over 80%.

To further substantiate the model's efficacy, we computed the values for pertinent evaluation metrics and plotted the ROC curve. The findings are detailed subsequently in Table 1:

**Table 1** Details the predictive accuracy of the deep learning model against the test set

| Match_Id | Evaluation Index | Result |
|---|---|---|
| 1401 | Recall | 0.864 |
| | Precision | 0.893 |
| | F1 score | 0.877 |
| 1601 | Recall | 0.842 |
| | Precision | 0.911 |
| | F1 score | 0.870 |

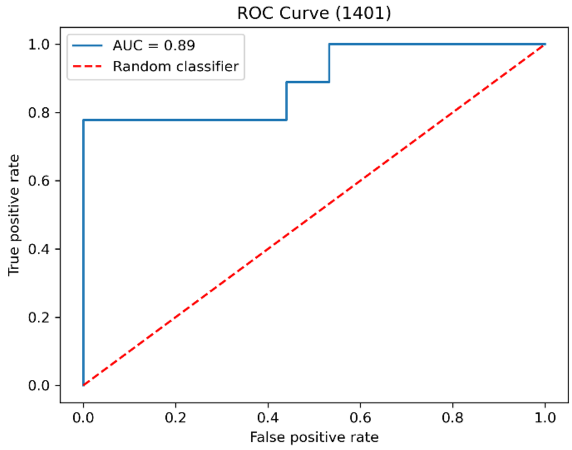



**Figure 5** the ROC curve of the 1401 data set

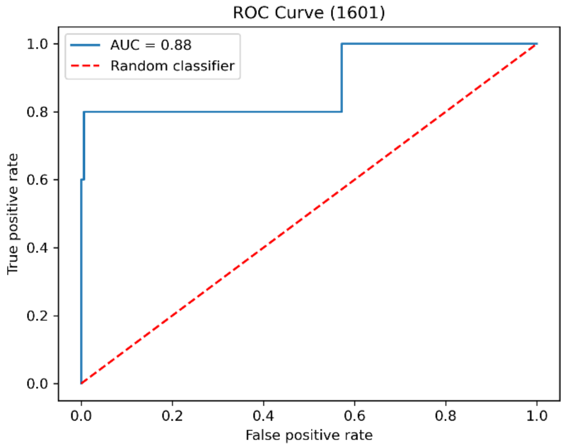



**Figure 6** the ROC curve of the 1601 data set

Evaluation metrics indicate that the model accurately identifies over 80% of positive in stances, with predictive precision nearing 90%. Moreover, the figure 5 and figure 6 illustrate the ROC curve predominantly occupies the upper left quadrant, nearly enclosing an area of 0.9 with the x-axis. These findings suggest the Transformer deep learning model developed herein precisely detects pivotal moments in tennis matches.

### 3.4. Evaluation of Model Universality (Prediction of Swing in Table Tennis Match)

Employing the Transformer deep learning model, we predict pivotal moments within table tennis matches and compute metrics such as precision, recall, F1 score, and the ROC curve. The outcomes are in the table 2:

**Table 2** Details the predictive accuracy of swings in table tennis matches

| Match | Evaluation Index | Result |
|---|---|---|
| Table Tennis | Recall | 0.934 |
| | Precision | 0.953 |
| | F1 score | 0.922 |

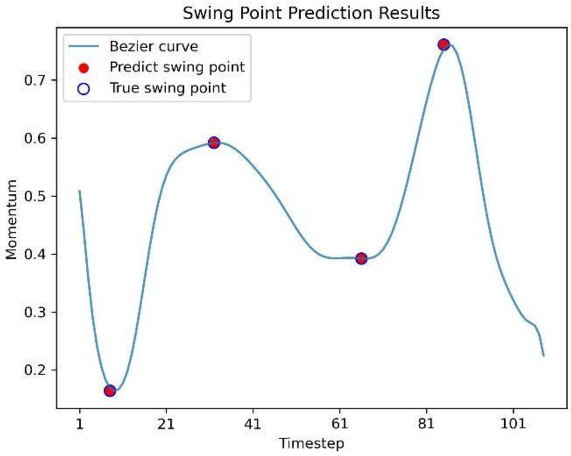



**Figure 7** Swing Point Prediction

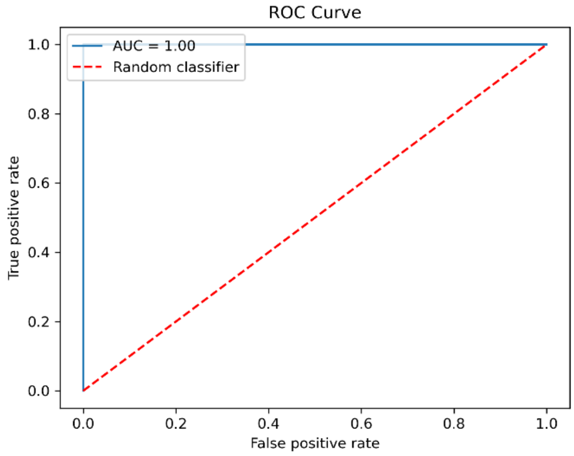



**Figure 8** the ROC curve

The aforementioned results in the figure 7 and figure 8 demonstrate that both the momentum quantification model and the deep learning-based swings prediction model devised in this study are applicable across various sports disciplines, indicating robust model generalizability.

### 3.5. Sensitivity Analysis

Variations in certain parameters within the Transformer deep learning model can cause prediction accuracy to fluctuate. This section focuses on three primary parameters: the hidden dimension, the number of multi-head attention heads, and the number of transformer encoder layers.

The analysis reveals that with a low hidden dimension, the model demonstrates limited expressive capabilities, leading to reduced prediction accuracy. Specifically, when the hidden dimension is below 16, an increase in this parameter correlates with enhanced prediction accuracy. However, at a d_model value of 32, a decline in accuracy is observed, likely indicative of over fitting within the model. Moreover, the prediction accuracy improves progressively with an increase in the number of multi-head attention heads, eventually reaching a plateau. Conversely, variations in the number of Transformer encoder layers exert a comparatively minor effect on the model's accuracy.

## 4. Conclusions

This study addresses dynamic momentum prediction challenges in tennis—static modeling, nonlinear relationships, and ambiguous definitions—via a Transformer framework with 12 dynamic features and a novel momentum definition. The model achieves 8-second critical point prediction, identifying serve-break interaction as core. Contextual adaptations and Transformer self-attention enable nonlinear insights. Practically, it delivers real-time tools like 1st serve $< 65\%$ alerts, and enhances data-driven strategies like run distance $\pm 8m/set$. Cross-sport validation and environmental integration advance turn-based analytics. This work bridges empirical and data science, offering a replicable momentum paradigm for tennis and other sports, optimizing competitive decisions through precision modeling.

This article, in the process of model construction, did not consider factors beyond the data provided in the question that affect the match situation, such as the technical level of players, environmental factors, tactical strategies, etc. To enhance the predictive model for volatility, we investigate additional influencing factors not initially incorporated into the model but which exhibit significant effects on matching dynamics. The model's assessment of technical skill requires improvement when considering momentum, serve speed, and player travel distance. Subsequent iterations could further explore the precision and speed of technical maneuvers, integrating nuanced technical metrics. Additionally, the quantification of tactics and strategy, which profoundly influences match outcomes, may serve as a quantifiable variable in future studies. Lastly, environmental factors should be expanded to encompass a broader range of meteorological and site-specific conditions, analyzing their interaction with player performance and match results.

## References

[1] Zijian G, Amanda K. Random Forest model identifies serve strength as a key predictor of tennis match outcome[J]. Journal of Sports Analytics, 2021, 7(4): 255-262.

[2] Zhang Rong. Tennis match prediction and the player analysis [D]. Yunnan university, 2022.

[3] Zhao R, Li R, Cai J. A Study on the Influence Factors of Momentum Based on Random Forest and XGBoost Algorithm[J]. Academic Journal of Mathematical Sciences, 2024, 5(2).

[4] Li Yang, LI Zhaobang, LUO Tingjin. Real-time victory Prediction and Momentum Analysis of Tennis Match Based on XGBoost and Time Domain Feature Extraction[J]. Modeling and Simulation, 2024, 13(6): 5732-5743.

[5] Huang Zhiying. Tennis match results prediction model based on BP neural network research [D]. Fujian normal university, 2022.

Jones I M, Harwood C. Psychological Momentum within Competitive Soccer: Players' Perspectives[J] . Journal of Applied Sport Psychology, 2008, 20(1): 57-72.

[6] Zhao R, Li R, Cai J. A Study on the Influence Factors of Momentum Based on Random Forest and XGBoost Algorithm[J] . Academic Journal of Mathematical Sciences, 2024, 5(2).

[7] Giovanni A, Vincenzo C, Luca A D. Weighted Elo rating for tennis match predictions[J]. European Journal of Operational Research, 2022, 297(1): 120-132.

[8] Eol H K, Fuwen C, Hyun J R, et al. Tennis match time Series do not exhibit long term correlations[J]. international journal of sport psychology, 2015, 46(6): 542-554.

[9] Ning X, Luo D, Zhang S. Analysis and prediction of tennis players' match performance with sentiment analysis[C] // Beijing University of Posts and Telecommunications (China), 2024.