Evaluating Momentum-Weighted LSTM Models for Predicting Tennis Match Outcomes

Heng Hua *, Wenqian Sun

School of Land Science and Technology, China University of Geosciences, Beijing, China, 100083

* Corresponding Author Email: qingw5261@gmail.com

Abstract. In the realm of sports competition, particularly in skill- and strategy-centric sports like tennis, momentum stands out as a pivotal factor influencing match outcomes. Conventional match analysis methods often fall short in capturing the nuanced shifts in momentum and their profound impact on players' psychological states and performances. To precisely forecast momentum changes and their repercussions, we introduce two integrated concepts: momentum and a player's point-winning probability. Weight calculations are optimized through cross-validation, and a model predicting match trends is formulated using the LSTM neural network theory. The results underscore a direct correlation between momentum swings and players' scores, significantly shaping match outcomes. Evaluation metrics reveal the model's high accuracy (87.06%) and minimal loss (0.6655) on the test set, attesting to its robust predictive capabilities. This research not only advances our understanding of momentum dynamics in sports but also showcases a methodological approach with potential applications across various disciplines, fostering a more nuanced comprehension of competitive dynamics.

Keywords: Long Short-Term Memory, Cross-Validation, Momentum, Tennis match prediction, Real-time performance.

1. Introduction

In the realm of sports, particularly in skill and strategy-intensive disciplines like tennis, momentum is widely acknowledged as a pivotal factor influencing match outcomes. Conventional sports match analysis typically relies on fundamental statistical data such as points won, games won, set victories, and the athlete's individual ranking [1, 2]. Regrettably, while these data provide valuable insights, they fall short of capturing the nuanced shifts in momentum during a match. These subtle changes can profoundly impact a player's psychological state, consequently influencing their performance in the game. Therefore, investigating player momentum during the course of a match becomes imperative.

In order to comprehensively investigate momentum changes and their impact on players, this study employed the Long Short-Term Memory (LSTM) algorithm. LSTM, proposed by Sepp Hochreiter and Jürgen Schmid-Huber in 1997 [3], represents a deep learning model specifically designed for processing sequential data. This methodology has found extensive application in various domains, including traditional sports event prediction [4], sports event data analysis [5], and emerging esports [6], consistently demonstrating significant efficacy. Consequently, this study posits the feasibility of applying the LSTM algorithm to tennis events.

In this study, detailed data analysis and preprocessing were conducted initially. Subsequently, based on the processed data, charts depicting playing time, hitting time, running speed, and running distance were created. Selected indicators were then utilized to formulate mathematical equations regarding momentum, and the weights for each indicator were determined through cross-validation [7]. Following this, the study, grounded in the principles of LSTM, established a momentum model based on the mathematical equations and further developed a real-time performance model for athletes. Rigorous testing was subsequently conducted to assess the accuracy of the models, confirming the precision of the results. Ultimately, utilizing the established models, this study achieved the visualization of the match process.

2. Model preparation

2.1. Data Processing

Data for this study was obtained from *www.comap.com*.The following steps were taken to address the problems in the data and to create a valid dataset: Study employ the technique of binary encoding to transform non-numeric features such as *serve_width*, *serve_depth*, and *return_depth* into numeric features. it addressed 316 cases of incorrect data with values from 1 to 9 in p1_score and p2_score by a reasonable interpolation method according to the process of the match. Simultaneously, for the 752 missing entries in the speed_mph column, this study applied Lagrange interpolation to handle the gaps. Regarding the 54 missing entries for serve_width, 54 for serve_depth, and 1309 for return_depth, it opted for a random sampling filling approach to address these gaps. These missing data points typically signify unknown information (NA data), and this paper's efforts through these methods aim to better preserve the integrity and availability of the data.

2.2. Data Analysis

In study's comprehensive examination of the preprocessed Wimbledon dataset, it embarked on an exploratory analysis focusing on data features that may be integral to understanding the dynamics of player momentum in professional tennis matches. This section delves into the insights garnered from the analysis of *elapsed_time*, *rally_count*, *speed_mph*, and *p1&2_distance_run*, as depicted in the Figure 1.

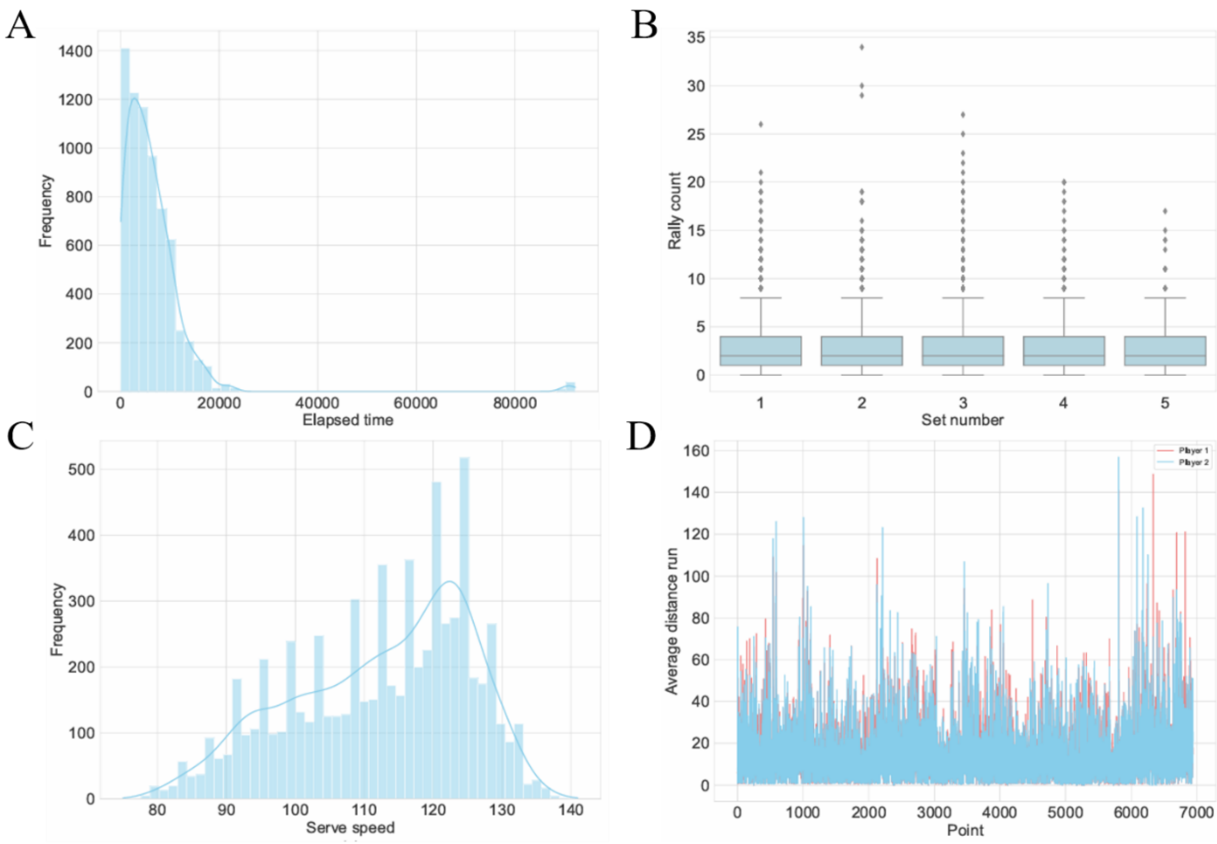

Figure 1. Partial data characterization charts

In subfigure 1A, histograms show a skew towards shorter matches, with a tail for longer durations, highlighting endurance in the sport. In subfigure 1B, boxplots reveal variability in rally lengths across sets, reflecting adaptability, fatigue, and tactical shifts, with median counts suggesting dynamic gameplay. In subfigure 1C, distributions center around medium serves speeds, with a spread towards both slower and faster serves. In subfigure 1D, line graphs of movement distances indicate physical exertion levels and court use, with fluctuations that may align with match pivotal moments.

3. Construction of Tennis Match Prediction Model

The model for this study considers each player's probability of winning each point at any given time as the flow of the game at the time the point occurs, and we use momentum to quantify a player's performance at the time of the in-game match, and determine a player's performance in real-time by comparing the trend of momentum curves over time. To achieve this objective, a neural network model is employed for prediction, taking specific features from the dataset as input and solving to obtain the predicted outcomes (as depicted in the Figure 2).

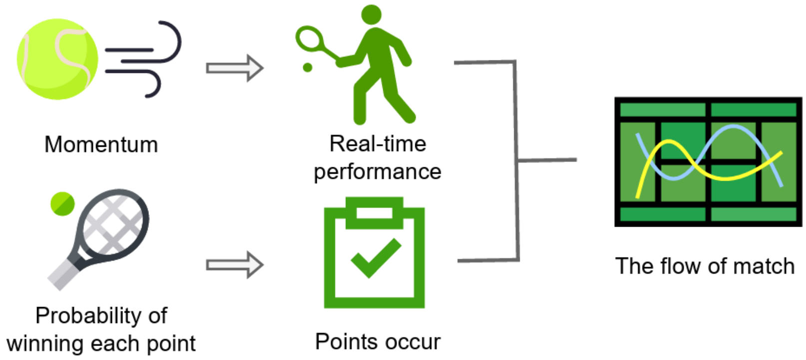

Figure 2. Schematic representation of key concepts and model solutions

3.1. Mathematical Modelling of Momentum and Real-time Performance

In this study, the concept of momentum in tennis matches is quantified through a composite measure reflecting a player's performance dynamics and potential to dominate phases of the match. Since in tennis the likelihood of a server winning a point is much higher, we will focus on this factor in our model. We will take into account the psychological and physical aspects of the players.

The proposed momentum metric, denoted as M, i.e. a player's performance in real-time, integrates various observable factors into a unified framework, as shown in Equation (1):

$$\mathrm{M} = \omega_1 \times \mathrm{SGD} + \omega_2 \times \mathrm{RGP} + \omega_3 \times \mathrm{M_{score}} + \omega_4 \times M_{serve} + \omega_5 \times M_{pressure} + \omega_6 \times M_{return} + \omega_7 \times M_{endurance} \tag{1}$$

Where $SGD = \frac{S_{win}}{TS} \times 100\%$, Serve Game Dominance, with S_{win} as held serves, TS as total serves. $RGP = \frac{R_{win}}{TR} \times 100\%$, Return Game Performance, with R_{win} as breaks, TR as total returns. $M_{score} = \sum_{i=1}^{n}(P_i - O_i)$, scoring dynamics, P_i and O_i are player's and opponent's scores in game i. $M_{serve} = \frac{A+EFS}{TS}$, serve efficiency, A as aces, EFS as effective first serves. $M_{pressure} = \frac{KPW}{KPT}$, pressure performance, KPW as key points won, KPT as total key points. $M_{return} = R_{depth} - R_{notdepth}$, return depth, R_{depth} and $R_{notdepth}$ for deep and non-deep returns. $M_{endurance} = S_{width} + S_{depth} + R_{depth} + D + S_{speed}$, endurance, incorporating serve and return depth, player movement(D), and serve speed (S_{speed}).

Since the probability of the serving side winning in tennis is much higher, study will consider in particular the relevant characteristics related to the serve, as shown in Figure 3.

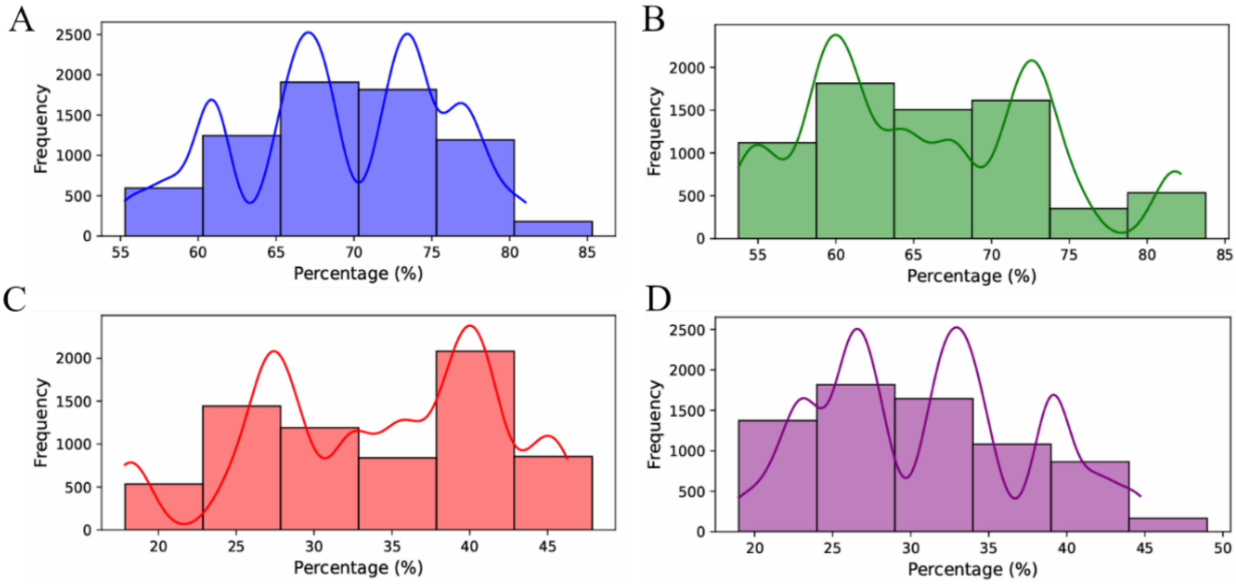

Figure 3. The distribution plots for SGD and RGP. SGD of player 1 corresponds to subgraph A, SGD of player2 corresponds to subgraph B, RGP of player 1 corresponds to subgraph C, and RGP of player 1 corresponds to subgraph D

The histograms for both players' SGD show a normal distribution, with Player 1 demonstrating a higher mean percentage (69.05%) compared to Player 2 (65.96%), indicating superior serve game performance. Conversely, the RGP distributions are right-skewed, suggesting variability in players' ability to win points on the return. Player 1's RGP mean is 34.04%, while Player 2's is slightly lower at 30.95%.

The probability of a player winning per point, i.e. the flow of play at the time points occurs, Pwin, in a tennis match can be calculated with the logistic model in Equation (2):

$$p_{win} = \frac{1}{1+\exp\left(-(\alpha\cdot\Delta Perf+\beta\cdot Prog+\gamma\cdot Dyn)\right)} \tag{2}$$

Where $\alpha = 0.5, \beta = 0.1, \gamma = 0.05$ are coefficients for the model. $\Delta Perf = f_{sets} + \delta \cdot f_{games}$ quantifies the differential in historical performance, with $f_{sets} = p1_{sets} - p2_{sets}$ and $f_{games} = p1_{games} - p2_{games}$, and $\delta = 0.3$. $Prog = \eta \cdot S_{no} + \theta \cdot G_{no}$ assess the match's progression, where S_{no} is the current set number, G_{no} is the game number, $\eta = 0.2$, and $\theta = 0.1$. $Dyn = \lambda \cdot \gamma_{time} + \mu \cdot RC + \nu \cdot V_{serve}$ captures the dynamics, with $\gamma_{time} = \frac{1}{elapsed_{time}+\varepsilon}$, RC as the rally count, V_{serve} the serve speed, and $\lambda = 0.01$, $\mu = 0.02, \nu = 0.03$.

3.2. Determining Weights for Momentum Calculation using Cross-Validation

The coefficients ω_1 through ω_8 are weighting factors that adjust the influence of each component on the overall momentum metric, tailored to reflect the nuances of match dynamics and player strategies. To determine the optimal weights, we employ a cross-validation approach on a dataset of match outcomes and performance indicators. The steps are as follows:

Step 1: Split the dataset into k folds.

Step 2: For each fold, use k −1 folds for training a predictive model and the remaining fold for validation.

Step 3: Optimize the weights in the training phase to minimize prediction error on the validation set.

Step 4: Repeat this process for each fold and average the weights across all folds. The following table presents the optimized weights obtained through cross-validation, shown below in Table 1:

Table 1. Optimized weights determined through cross-validation

Performance Indicator	Weight
SGD	0.15
RGP	0.20
$Mscore$	0.10
$Mserve$	0.25
$M_{pressure}$	0.15
$Mreturn$	0.05
$M_{endurance}$	0.10

So far, the input and output data of the neural network LSTM prediction model have been determined.

3.3. The Solution of LSTM

Long Short-Term Memory (LSTM) is a classical type of Recurrent Neural Network (RNN) model designed to handle sequential data [8, 9]. It exhibits superior performance in capturing long term dependencies and maintaining memory compared to traditional RNNs. LSTM finds applications in various domains such as time series prediction, natural language processing, and speech recognition [10]. The core of the LSTM model's ability to capture long-term dependencies in sequence data is its cell state update mechanism, which is governed by the following Equation (3):

$$C_t = f_t \odot C_{t-1} + i_t \widetilde{C_t} \tag{3}$$

Where Ct is the cell state at time step t. ft is the forget gate's activation at time step t, controlling the extent to which values in the cell state are kept or discarded. C_{t-1} is the cell state from the previous time step. i_t is the input gate's activation at time step t, determining how much of the new information is stored in the cell state. $\widetilde{C_t}$ is the cell candidate vector at time step t, representing the new Information that could be added to the state. $\odot$ denotes element-wise multiplication.

This study employs a two-layer LSTM network for tennis match player's real-time performance prediction, leveraging TensorFlow. The model is trained with Adam optimizer across 1000 epochs, incorporating weight decay and dropout to mitigate overfitting. Key training parameters are summarized below in Table 2:

Table 2. LSTM model training parameters

Parameter	Description	Value
MaxEpochs	Maximum epochs	1000
Batch Size	Samples per batch	32
Optimizer	Optimization algorithm	Adam
Loss Function	Prediction error measure	Categorical Crossentropy
Metrics	Evaluation metric	Accuracy
Dropout Rate	Input dropout fraction	0.2
GradientThreshold	Gradient clipping threshold	1
InitialLearnRate	Learning rate	0.01
LearnRateSchedule	Learning rate adjustment schedule	Steprdecayratr500repochs
L2 Regularization	Regularization factor	0.0001

The processed dataset will be randomly divided into 80% training set and 20% testing set. This study will normalize the test set, using the last time step of the training set and the target sequence of the training set as inputs to the network. After training is completed, it uses the output of the last time step in the training set as the first predicted value. Subsequently, this study iteratively predicts future time steps. it will use visualization techniques to illustrate the model's ability to capture the dynamics of the match process, facilitating a better understanding of the dynamic variations during the match. The training results of the model are shown below in Figure 4.

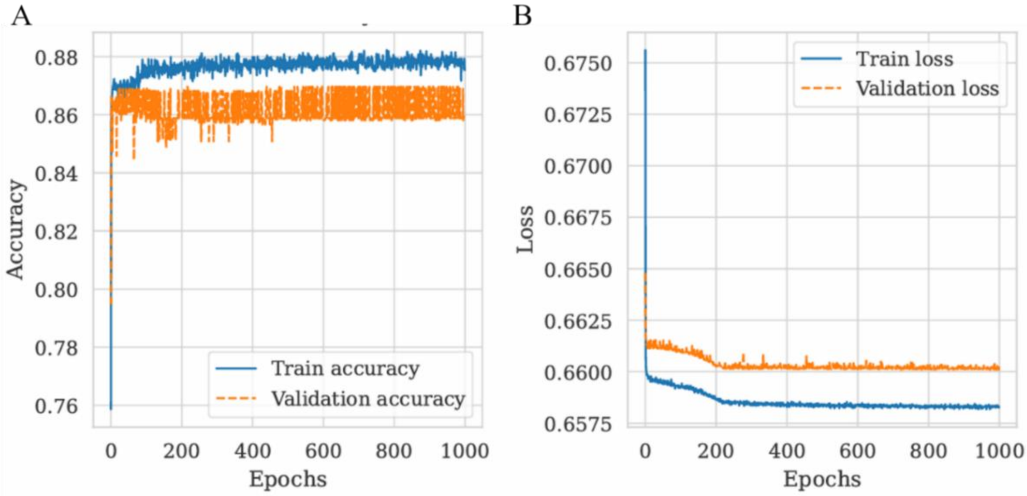

Figure 4. Schematic diagram of the LSTM neural network structure

The LSTM model's training results reveals a progression towards convergence over 1000 epochs. Initially, the model starts with a loss of 0.6765 and an accuracy of 75.81% on the training set, and a validation loss of 0.6655 with a validation accuracy of 79.39%. Over the epochs, both the training and validation loss decrease, albeit marginally, indicating a slight improvement in the model's ability to generalize. The training accuracy peaks at 88.17% in the 979th epoch, while the validation accuracy reaches a high of 87.06% in the 204th epoch.

Subsequent epochs do not show significant improvement in validation accuracy, oscillating around the 88% mark, which suggests the model has reached its performance capacity with the given data and architecture. The minimal decrease in validation loss from the first to the tenth epoch and the relatively stable validation accuracy indicate that the model is neither overfitting nor underfitting significantly, but may benefit from further hyperparameter tuning or additional training data to break the plateau in learning.

3.4. Visualisation of the Match Flow

In the Carlos Alcaraz vs. Nicolas Jarry match *(2023-Wimbledon-1301)*, study analyze momentum through an LSTM model, showcasing player's performance in real time in Figure 5, winning probabilities in Figure 6, and average serve speeds in Figure 7.

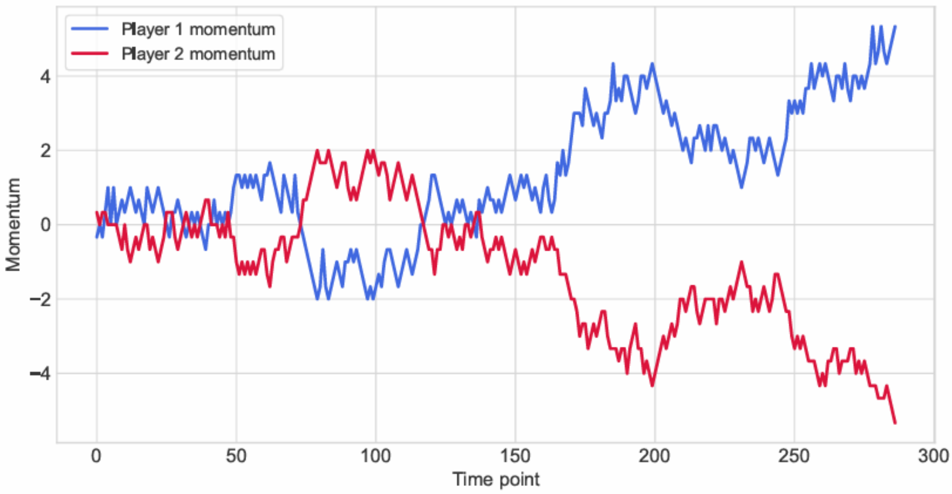

Figure 5. Match momentum for Alcaraz and Jarry

The momentum visual (Figure 5) reveals pivotal momentum swings and shows the flow of the game when scoring occurs, with Alcaraz gaining an upper hand at crucial moments, impacting the match's outcome. The area under the curve for player 1 and player 2 is 387.17 and -350.17. The match analysis reveals initial player parity, marked by a 3-3 score, before Alcaraz (Player 1) captured a pivotal 5-3 lead post the 50th time point, signaling his performance ascent. Jarry (Player 2) led early in the second set with a 3-0 score but was overtaken as Alcaraz reclaimed momentum, winning the set. Alcaraz maintained this pattern, notably during a tiebreak at the 143rd point, demonstrating his resilience. LSTM model data illustrates a direct link between momentum changes and Alcaraz's scoring streaks, culminating in a 3-1 victory.

Figure 6 delves into the winning probabilities (at any given time as the flow of the game at the time the score occurs), highlighting the dynamic nature of the match, with momentum swings aligning with shifts in winning probabilities.

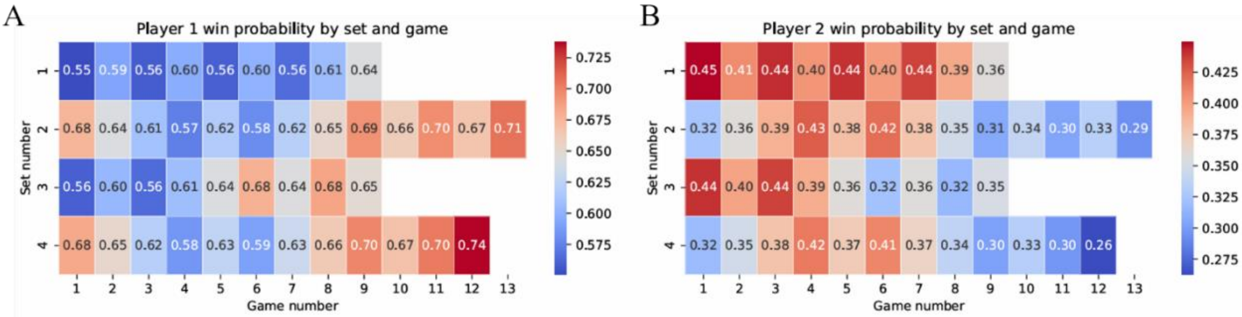

Figure 6. Winning probability heat map for Alcaraz and Jarry

Lastly, figure 7 examines serve speeds, offering insights into serve advantage and player fitness across the match.

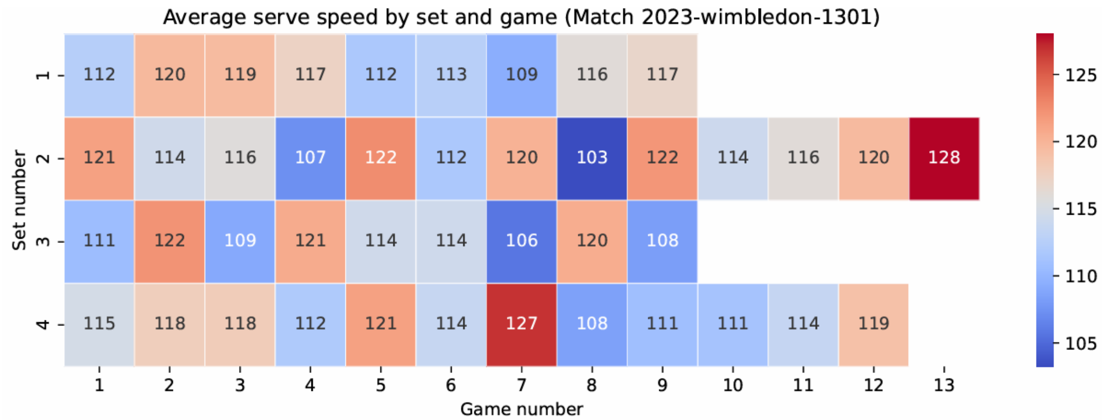

Figure 7. Heatmap of average serve speeds

The analyses underscore the correlation between momentum, scoring, psychosocial, and physical performance, culminating in Alcaraz's victory.

In summary, clear differences exist in the distribution of momentum and swing points across racket sports. Tennis matches demonstrate relatively stable momentum shifts, indicating consistent trends. Badminton matches show more variability and swing points, suggesting evenly matched and tense competitions. Table tennis matches display the most frequent and dramatic changes in momentum and swing points, indicative of fast-paced games with intense competition for the lead.

4. Conclusion

In this study, it proposed two integrated concepts of momentum and player's probability of winning each point, using cross-validation to calculate the weights. Study constructed a Long Short-Term Memory (LSTM) neural network model that predicts the match flows when the point occurs, and obtained a visualization of the momentum changes of both players to determine which player performed better at a given time. The momentum area is 387.17 and -350.17. The results show that there is a direct link between the momentum swing and the score of players, who prevailed at a crucial time and influenced the outcome of the match. The evaluation shows that model has an accuracy of 87.06% and a loss of 0.6655 on the test set, demonstrating a high level of confidence.

This study reveals a research framework for the fields of sports science and sports psychology, aiming to address the challenge of capturing subtle momentum changes in athletes during competitions. The study contributes to assisting coaches in formulating training strategies and aiding athletes in adjusting their states to secure victories.

References

[1] Fayomi, A., et al., Forecasting Tennis Match Results Using the Bradley-Terry Model. International Journal of Photoenergy, 2022. 2022.

[2] Wu, E. and H. Koike. Futurepong: Real-time table tennis trajectory forecasting using pose prediction network. in Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems. 2020.

[3] Hochreiter, S. and J. Schmid Huber, long short-term memory. Neural computation, 1997. 9 (8): p. 1735 - 1780.

[4] Zhang, Q., et al., Sports match prediction model for training and exercise using attention-based LSTM network. Digital Communications and Networks, 2022. 8 (4): p. 508 - 515.

[5] Zhao, Y., et al., Applying deep bidirectional LSTM and mixture density network for basketball trajectory prediction. Optik, 2018. 158: p. 266 - 272.

[6] Zhang, L., et al., Improved Dota2 lineup recommendation model based on a bidirectional LSTM. Tsinghua Science and Technology, 2020. 25 (6): p. 712 - 720.

[7] Bates, S., T. Hastie and R. Tibshirani, Cross-validation: what does it estimate and how well does it do it? Journal of the American Statistical Association, 2023: p. 1 - 12.

[8] Staudemeyer, R.C. and E.R. Morris, Understanding LSTM--a tutorial into long short-term memory recurrent neural networks. arXiv preprint arXiv:1909.09586, 2019.

[9] Yu, Y., et al., A review of recurrent neural networks: LSTM cells and network architectures. Neural computation, 2019. 31 (7): p. 1235 - 1270.

[10] Li, S., et al. Independently recurrent neural network (indrnn): Building a longer and deeper rnn. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.