Problem Chosen	2024	Team Control Number
C	MCM/ICM Summary Sheet	2406905

Momentum: The Ethereal Arsenal Above the Skies, Guiding Triumph

Summary

In tennis, a player's performance can fluctuate over multiple points or even games, and the player with the upper hand in power may not be able to win. The aim of this paper is to create a momentum-centred system that explores the relationship between momentum and the match, and provides accurate dynamic predictions of match situations.

For the first part, we built a **momentum index model** based on the **G1-Entropy -Independence** Weight integrated evaluation method. With this integrated evaluation method, we identified different weights for different data. Subsequently, we calculated the momentum indices of the players before the start of the match to assess the quality of their performance and the degree of their dominance. Next, we qualitatively compared the momentum indices with score point curve and game point curve to visualise trends in tennis matches.

For the second part, we establish a matching system between the momentum index and player performance based on **DTW**. We represent continuous scores in terms of continuous highlight performances. Subsequently, we match the player momentum index curve, the point score curve and the games score curve by the dynamic time regularisation algorithm to get the match between the momentum index curve and the two curves. Eventually, it was found that the matching degree of momentum index curve with point score curve and games score curve reached 88.635% and 91.156%, respectively. This suggests that momentum plays a key role in games and that the coach's claim is either incorrect or not rigorous.

For the third part, we applied the **LSTM** model in deep learning and compared it with the **ARIMA** prediction model, and found that the LSTM model is more accurate and better. At the same time, we applied the LSTM model for prediction, and the accuracy of the validation set reached 81.32%, and the loss rate was reduced to 0.07, which is already a very high accuracy. Not only that, we also gave constructive advice to the athletes on how to prepare for the game when facing new opponents from the three dimensions of opponent, player's self, and on-site factors.

For the fourth part, we chose data from the 2023 NBA Eastern Conference Finals game between the Heat and the Celtics. We included pre-game player psychological pressure, the course of the first six games as considerations in the momentum index model developed. We then matched the momentum index curve to the scoring uptick curve of the game and found a match rate of 65.918%. Based on the momentum index model, we conclude that the Heat will win, which also matches the final result.

Finally, we provided a non-technical report and conducted **sensitivity analysis** of the evaluation model, demonstrating the robustness and accuracy of the model.

Keywords: G1-Entropy-Independence DTW ARIMA LSTM

Contents

1 Introduction

1.1 Problem background

In the 2023 Wimbledon final, 36-year-old Novak Djokovic lost for the first time at Wimbledon since 2013 when he was defeated 2:3 by 20-year-old Alcaraz. In the first set, Djokovic crushed to a 6:1 win. But Alcaraz came out of nowhere to win 3:2 in the following game. It was a landmark match, no doubt about it. But at the same time, there was a lot of discussion about how Alcaraz was able to regain his confidence and win the match after losing the first set. One element that was brought up was the player's momentum, which is affected by all aspects of the player's game, but can be countered by the player's new game, and whose magnitude is a key factor in whether or not the player is able to win.

1.2 Restatement of the Problem

Consider the background information, we need to complete the following tasks in sequence:

- Build a momentum index model to determine which player is performing better at a given moment and visualize the changing game situation.
- Establish the matching system of momentum and victory, evaluate the coach's opinion, and make a quantitative evaluation of whether the momentum is real.
- Build a match trend prediction model based on players' form to identify turning points in matches and give players advice on how to prepare for matches.
- We need to transfer the momentum index model of tennis matches to other matches to see if the model is still valid and whether the model is widely used.

1.3 Literature Review

This subject focuses on simulating player momentum and investigating the flow of events throughout a tennis match. After doing a thorough literature search, our team discovered the following three main strategies:

- First, deep learning with neural networks[1] and machine learning-based modeling using popular techniques like logistic regression[2], random forests[3], and so on.
- Another strategy that uses extensive simulations instead of computing methods is Monte Carlo-based simulations[4].
- In the end, it is based on the methods of evaluation, and we discovered G1[5], the independence weight method[6], the entropy weight method[7], etc.

The specific method is shown below:

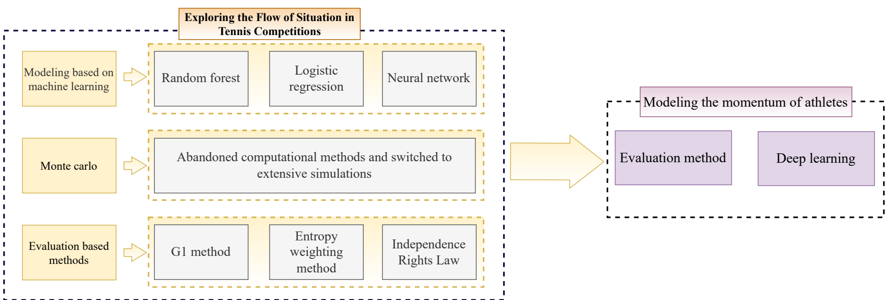


Figure 1: Literature Review Framework

After considering many trade-offs, our team decided to integrate deep learning modeling with evaluation techniques.

1.4 Our work

The entire modeling process is as follows:

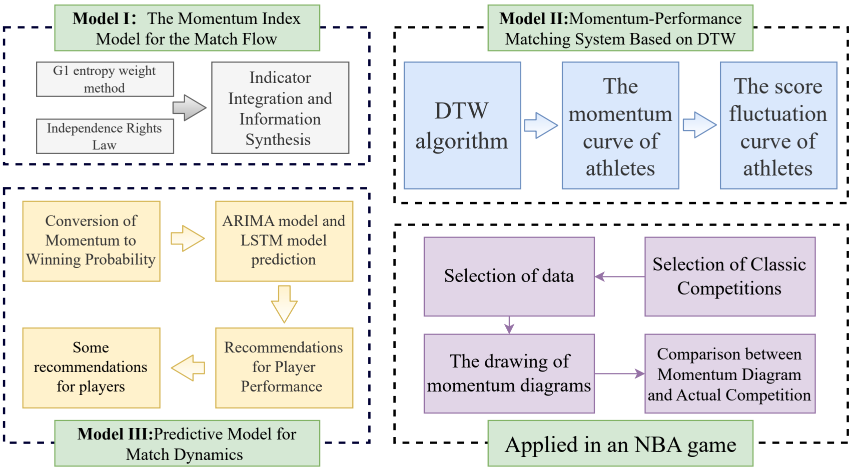


Figure 2: Model framework

2 Assumptions and Justifications

The following basic assumptions are made to simplify problems.

Assumption 1: We assume that external factors such as weather, spectators, etc. do not affect the player's performance, so that the model and its results are fully analysed from the available data.

Justification 1: In fact players' play can be affected by factors such as wind speed, sunlight, and whether or not the spectators are disciplined in the game, and disregarding these can help us simplify the model.

Assumption2: We assume that the players are all in the same physical condition before starting the match.

Justification 2: Wimbledon is a knockout format, and player fitness is theoretically affected by the previous matches, but since we don't know a player's fitness reserves or original fitness status, it is assumed that the two players before the match were in the same physical condition.

Assumption 3: We assume that players' psychological factors are not affected by their record against their opponents.

Justification 3: In actual matches, players' psychology and strategy will be affected by past performance, which is too difficult to be considered due to the statistics.

3 Notations

The primary notations used in this paper are listed in Table 1.

Table 1: Notations

Symbol	Description
n	Number of data points
(x_i, y_i)	Represents the coordinates of the i-th data point
$x_i^{'}$	The i-th data point after minimum maximum normalization
$x_{\max}$	Maximum value in the original dataset
$x_{\min}$	The minimum value in the original dataset
b_i	Indicator set
r_j^a	The a-th person judges the weight ratio between adjacent indicators b_j and b_f
w_n	The final weight of indicator b_n
B	Non negative standard matrix
H_j	The entropy value of the j-th evaluation indicator b_j
p_{ij}	The proportion of data for the i-th time under the j-th indicator
W_j	The comprehensive weight vector of all indicators
x_i	Represent whether i scored at a specific time point
y_i	Momentum index at time point i
D	Cost matrix
S	Matching score
M_i	Player i's momentum
P_i	Player i's winning rate
f_t	Forgetting Gate
i_t	input Gate
c_t	Control the candidate state at the current moment
o_t	Output gate

4 Model I: The Momentum Index Model for the Match Flow

4.1 Indicator Integration and Information Synthesis

We believe that momentum can help us capture the flow of the game during a match, so building a model that captures the flow is equivalent to building a momentum index model.

As momentum is related to various conditions and data during the game, we categorized and summarized the data for each set of variables detected during the game. It was classified into five dimensions: course of the match, match pressure, offensive ability, physical condition and defensive ability. Each dimension is characterized by two to nine specific indicator variables, and the structural relationship between each dimension and its internal indicators is shown in Figure 3:

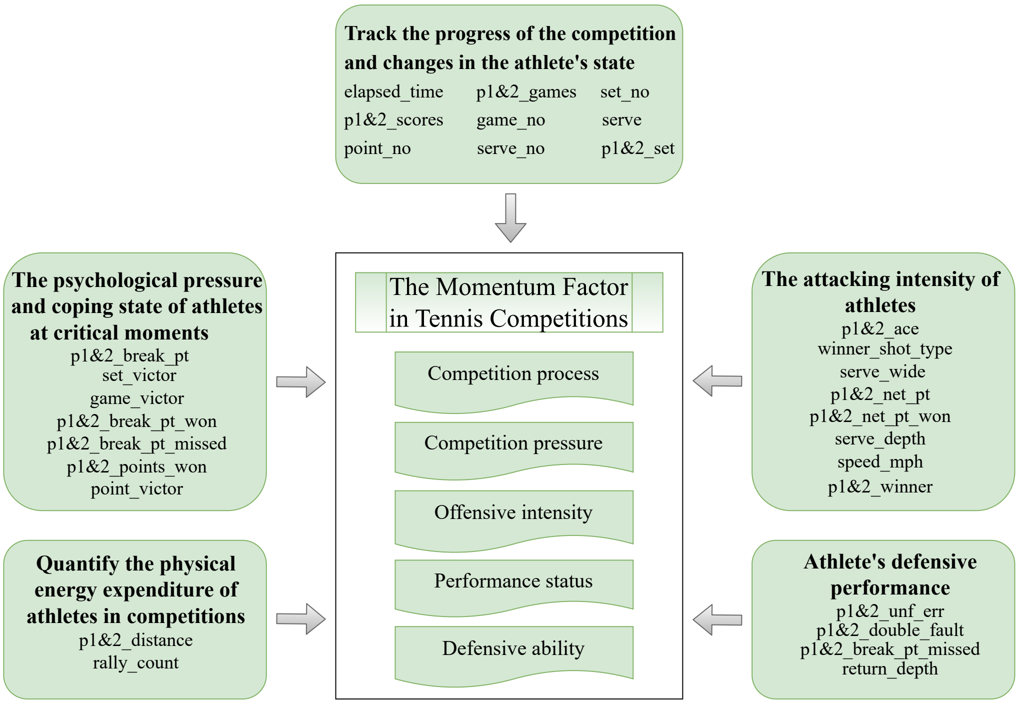


Figure 3: Classification chart of indicator information

4.2 The G1-Entropy-Independence Weight Comprehensive Evaluation Method

The weights of different types of indicators can reflect the amount of information of different levels of importance contained in different types of indicators. However, due to the wide variety of assessment indicators, there is inevitably information redundancy between indicators. Therefore, this paper proposes a comprehensive assessment method called G1-Entropy-Independence Weight (G1-Entropy-Independence Weight).

The method calculates the composite weight of each momentum indicator of the athlete by combining subjective weighting methods (ordinal analysis method (G1 method)) with objective weighting methods (entropy weighting and independent weighting methods). A certain amount of human subjectivity can be mitigated and changes in objective situations can be reflected by combining subjective and objective weighting techniques. The use of the independent weights method can improve the problem of over-sensitivity in the entropy weights method, reduce the problem of redundant information between indicators, and supplement the subjective weight determination method to finally assess the athlete's condition.

The process flow of this method is as follows:

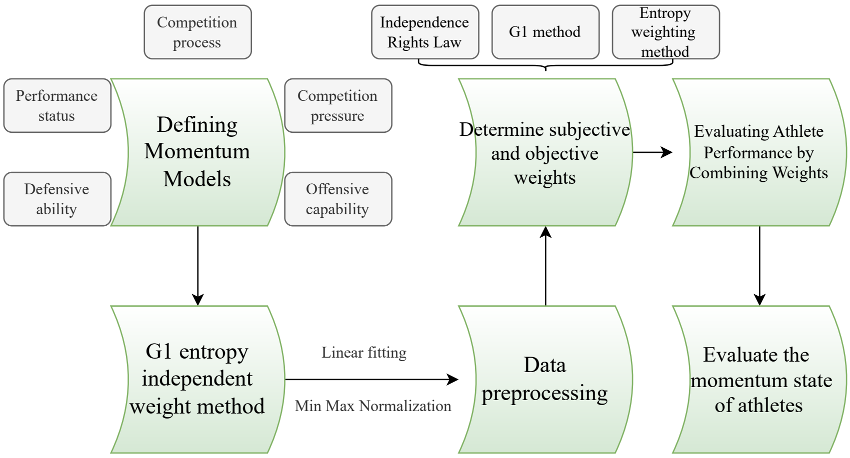


Figure 4: Evaluation method flowchart

4.2.1 Data Preprocessing

Linear Fit Filling

Missing values may lead to biases in the dataset that may mislead the final conclusions, such as missing values in ball speed. Therefore, it is necessary to perform a linear fit to the attribute columns with missing values to fill them. The mathematical principle used in linear fitting is the method of least squares[8], which requires finding a line that minimises the sum of the distances from all data points to the line, i.e. solving the following equation:

$$\min_{a,b} \sum_{i=1}^{n} (y_i - (ax_i + b))^2 \tag{1}$$

where n is the number of data points and (x_i, y_i) represents the coordinates of the i th data point. The method involves summing the error squares of the n data points, i.e., the sum of the squares of the distances from each data point to the fitted line. The least squares approach aims to minimize this sum of squares by determining the values of a and b.

Min-Max Normalization

Throughout the Excel dataset, there are a number of tennis-related metrics, such as depth, ball speed, and so on. Different metrics have different units and dimensions, and their value ranges, sizes and units vary significantly. This may lead to bias and errors in the results of data analysis. In order to eliminate the influence of dimensions and value ranges between data metrics on the results and to make different metrics comparable and analyzable[9], we scaled the raw data to the range [0,1] using min-max normalization. This process ensures that each indicator is comparable. The formula for normalization is given below:

$$x_i' = x_{min} + \frac{x_j - x_{min}}{x_{max} - x_{min}} \tag{2}$$

where x_i' is the i^{th} data point after min-max normalization, and x_{max} and x_{min} are the maximum and minimum values in the original dataset, respectively.

4.2.2 Utilizing the G1 Method to Determine Subjective Weights

The advantage of the G1 method[10] over hierarchical analysis is that it does not require consistency checks. In addition, it can overcome inaccuracies due to the hesitation of experts in making judgements when many indicators are present. It can establish the subjective weights rapidly and precisely, giving experts assessments a solid platform.

For each set of indicators $b_1, b_2, \cdots, b_n$, the specific steps for using the G1 weighting method are as follows:

Table 2: The G1 Weighting Method for Indicator Sets

Algorithm 1: The G1 Weighting Method for Indicator Sets

Input: Indicator set $\{b_1, b_2, \ldots, b_n\}$, number of indicators n, number of experts p, weight ratio values $\{1.0, 1.2, 1.4, 1.6, 1.8\}$

Step 1: For each expert a, identify the least important indicator b_n, then the second least important b_{n-1}, and so on, until a unique ordering is obtained.

Step 2: Calculate the weight of the least important indicator w_n^a using Equation (3):

$$w_n^a = \frac{1}{1 + \sum_{k=1}^{n-1} \prod_{j=k}^{n-1} (r_j^a - 1)} \tag{3}$$

Step 3: Calculate the weights of other indicators using the recursive formula (4):

$$w_{k-1}^a = \frac{w_k^a}{r_{k-1}^a} \tag{4}$$

Step 4: Average the weights of the n^{th} indicator obtained from all experts to calculate the final weight w_n using Equation (5):

$$w_n = \frac{\sum_{a=1}^{p} w_n^a}{p} \tag{5}$$

Output: Final weight for each indicator $\{w_n\}$.

4.2.3 Determining Objective Weights Using the Entropy Weight Method

An objective way for determining the weights of different indicators based on the quantity of information they carry is the entropy weighting method[11]. An indicator of the quantity of information is entropy. An indicator's entropy decreases with more information and variability in a complex decision-making scenario. Therefore, the more weight the indicator has, the more important it is in weighing possibilities[12]. This approach minimizes the impact of subjective opinions and has straightforward computation procedures, increased impartiality, and efficient data utilization.

Since the original data matrix has been processed into a dimensionless, non-negative standard matrix $B = (b_{ij})_{m\times n}$:

$$B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} \tag{6}$$

In the evaluation of an athlete's momentum over m time points and n evaluation indicators, the entropy value of the j^{th} evaluation indicator b_j is defined as:

$$H_j = -K \sum_{i=1}^{m} p_{ij} \ln p_{ij}, \quad j = 1, 2, \ldots, n \tag{7}$$

Where

$$\begin{cases} K = \frac{1}{\ln m} \\ p_{ij} = \dfrac{b_{ij}}{\sum_{i=1}^{m} b_{ij}} \end{cases} \tag{8}$$

In the formula, p_{ij} is the proportion of the data at time i for indicator j. It is also stipulated that: when $p_{ij} = 0$, $p_{ij} \ln p_{ij} = 0$.

Then, the entropy weight of the j^{th} indicator is defined as:

$$w'_j = \frac{1 - H_j}{\sum_{j=1}^{n} (1 - H_j)} \tag{9}$$

4.2.4 Determining Objective Weights Using the Independence Weight Method

While the entropy weighting method can effectively use data to minimize the impact of subjective expert opinion, it determines weights based solely on the amount of information contained in each individual indicator, without taking into account the overlap of information between indicators. If these indicators are used directly for assessment, information redundancy may occur, thus reducing the accuracy of the assessment[13]. Therefore, this paper adopts the independent weight method to consider the interrelationship between indicators. Combined with the entropy weight method, a composite objective weight of the indicators is formed, making the assessment results more accurate.

The independent weight coefficient method determines weights by calculating the coefficient of partial correlation between indicators[14]. The degree of redundant correlation can be assessed by calculating the partial correlation coefficient, which is a measure of the correlation between a dependent variable and a group of independent variables. The higher the partial correlation coefficient, the more redundant the information and the lower the weight. Using the inverse of the partial correlation coefficient as the weighting factor reflects the amount of information in the assessment indicators that is significantly different from the other indicators, effectively addressing the issue of redundant information between indicators.

In assessing the momentum of athletes, each group of indicators contains $\{b_1, b_2, \ldots, b_n\}$. If the indicator b_1 has a stronger linear relationship with the other indicators, this means that b_1 can be represented by a linear combination of the other indicators; from another point of view, this means that there is more redundant information, and therefore the weight of this indicator should be smaller. The formula for calculating the partial correlation coefficient R_j of the indicator b_j is as follows:

$$R_j = \frac{\sum_{j=1}^{n}(b_j - \bar{b})(\tilde{b} - \bar{b})}{\sqrt{\sum_{j=1}^{n}(b_j - \bar{b})^2 \sum_{j=1}^{n}(\tilde{b} - \bar{b})^2}} \tag{10}$$

In the formula, $\tilde{b}$ is the matrix obtained from matrix B by removing the column corresponding to b_j; $\bar{b}$ is the average value of the matrix B.

Since there is an inverse relationship between R_j and the weight w''_j, by selecting the reciprocal of R_j and then normalizing it, the weight values for each indicator can be determined. The final formula for calculating the weights is:

$$\begin{cases} R = \left(\frac{1}{R_1}, \frac{1}{R_2}, \ldots, \frac{1}{R_n}\right) \\ w''_j = \frac{\frac{1}{R_j}}{\sum_{j=1}^{n} R_j} \end{cases} \tag{11}$$

4.2.5 Determination of Composite Weights

After determining the weights $\bar{w_j}$, w'_j, and w''_j using the G1 method, the independence weight coefficient method and the entropy weight method, respectively, the composite weight for the j^{th} indicator can be obtained using the product method in the combination of weights as:

$$w_j = \frac{\bar{w}_j w'_j w''_j}{\sum_{j=1}^{n} \bar{w}_j w'_j w''_j} \tag{12}$$

Then the composite weight vector W_j for all indicators is:

$$W_j = [w_1, w_2, \ldots, w_n]^T \tag{13}$$

4.2.6 Determination of the Momentum Formula

At this point, the momentum formula can be determined as follows:

$$P = \sum_{i=1}^{n} w_i x_i \tag{14}$$

After incorporating the match data of Djokovic and Alcaraz, we have drawn a chart showing the momentum trend changes of the two players:

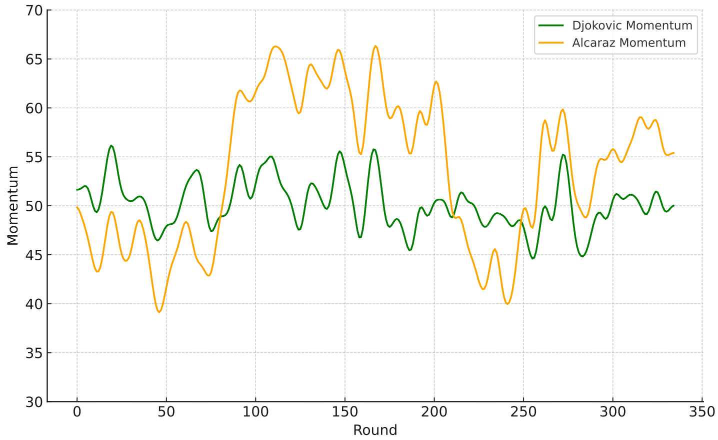
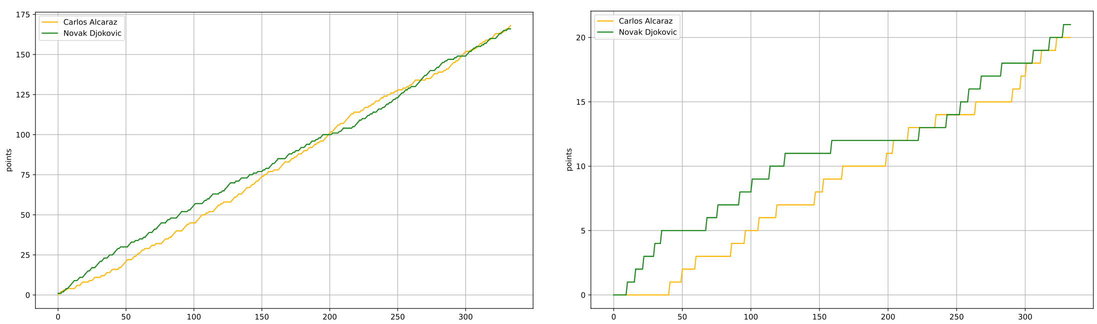


Figure 5: Trend chart of player momentum indicators

Not only that, we also drew the scoring curves of the players and the point scoring curves of the games, facilitating a comprehensive assessment of the flow of the game situation:



(a) (b)

Figure 6: Trend chart of player point score and game score

The momentum trend chart shows that Djokovic took the lead at the start of the match and won the first set. In fact, Alcaraz chose to strategically give up some sets when he was clearly behind in the game. In this way, Alcaraz ended up winning the match by virtue of his leading disc score, despite scoring lower in the total number of discs. The momentum trend graph also shows the specific performances of both players - although Alcaraz lost some games and points, he did so strategically and did not affect his momentum. Despite fluctuating momentum and alternating leads during the match, Alcaraz still performed better.

5 Model II:Momentum-Performance Matching System Based on DTW

In problem 2, a coach questions the existence of the "momentum" factor, arguing that players' performance fluctuations and scoring streaks are random. We need to develop an index and create a calculation system to evaluate coaches' statements about the randomness of players' performance fluctuations and scoring streaks, and to explore the relationship between the momentum index and players' wins and losses, as well as their scoring streaks.

For this purpose, we use the Dynamic Time Warping (DTW) algorithm[15] to build a DTW-based matching system between momentum indices and player performances, where consecutive scores are regarded as a kind of performance fluctuation, i.e., consecutive highlight performances. The first step is data preparation to obtain momentum change curves and players' scoring curves. Next is the initialisation of the cost matrix to ensure that the algorithm has a clear boundary. The calculation of the cost matrix is then performed, which efficiently identifies and quantifies the similarity between two time series while dealing with changes and inconsistencies over time. The DTW distance is then calculated, also to quantify the similarity between two time series. Subsequently, the DTW distance is converted into a matching degree which is used to present the results clearly. The closer the matching degree is to 1, the better the match between the two curves, indicating a higher correlation between them[16]. The steps of the DTW algorithm are shown in Figure 7.

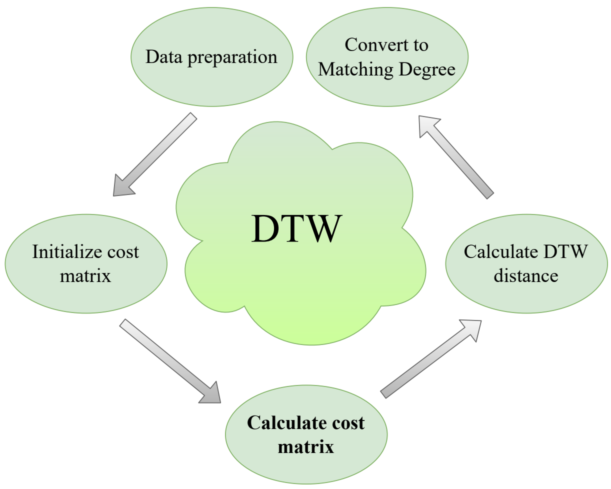


Figure 7: DTW algorithm flowchart

5.1 Data Preparation

In applying the Dynamic Time Warping (DTW) algorithm, we first need to prepare two sets of time-series data: one is the player's momentum index curve and the other is the player's scoring curve.

We represent the scoring curve as a vector: $X = \{x_1, x_2, \dots, x_N\}$. This curve represents the athlete's scoring record over time. Each x_i represents whether or not a score was scored at a particular point in time i. Also, we use a vector to represent the momentum index curve: $Y = \{y_1, y_2, \dots, y_M\}$. This curve reflects the change in a player's momentum index. Each y_i represents the momentum index at the point in time i.

Here, N represents the length of the score curve and M represents the length of the momentum index curve.

5.2 Initializing the Cost Matrix

In the course of the Dynamic Time Warping (DTW) algorithm, initialising the cost matrix involves creating a cost matrix D. This is a two-dimensional array of size $(N+1) \times (M+1)$. Its purpose is to record the distance and alignment cost between the score curve X and the momentum index curve Y. The smallest cumulative distance between the first i elements of the sequence X and the first j elements of the sequence Y is represented in the matrix by $D[i][j]$.

This matrix is initialised by setting $D[0][0]$ to 0, which represents a starting point, i.e. the state where there are no matching sequence elements, and the rest of the matrix is set to infinity (∞).

5.3 Calculating the Cost Matrix

Calculating the cost matrix$D$ is the core stage of the Dynamic Time Warping (DTW) algorithm[17]. This process aims to accurately quantify the similarity between the score curve$X$ and the momentum index curve$Y$.

For each pair of elements x_i and y_j in the sequences, we first calculate the distance between them as $d(x_i, y_j) = (x_i - y_j)^2$. This distance measure considers the square of the value difference between the two sequences at corresponding points, which is a common method for assessing the similarity between two points.

$$D[i][j] = d(x_i, y_j) + \min(D[i-1][j], D[i][j-1], D[i-1][j-1]) \tag{15}$$

Thus, each $D[i][j]$ ultimately contains the minimum cumulative distance from the beginning of the sequence to the point (i, j), allowing for temporal scaling and variation.

This process of padding begins at $D[1][1]$ and continues through to $D[N][M]$, ensuring that the cost matrix comprehensively reflects the cost of all possible alignments between the two sequences. In this way, the DTW algorithm can flexibly align time series and perform accurate calculations of the similarity of non-aligned curves.

Calculating the DTW Distance

In the DTW algorithm, the final element of the cost matrix $D[N][M]$ determines the DTW distance between the sequences X and Y once the cost matrix D is fully populated. The similarity between the two time series is measured by this value, which is the least cumulative distance of all feasible alignments from the start to the finish of the sequence.

Converting to Match Score

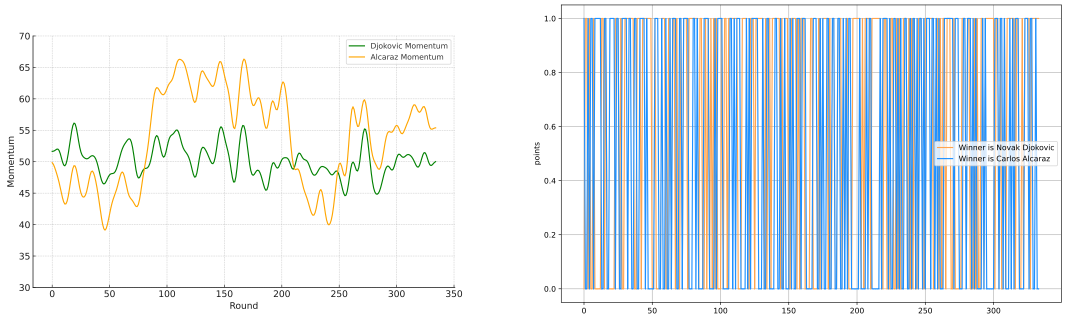
To show the similarity between the curves more visually, we use maximum distance normalisation to show the matching scores between the score curve and the momentum curve. In this method, the matching score$S$ is calculated by subtracting from 1 the ratio of the DTW distance between the sequences X and Y $D[N][M]$ to the calculated maximum possible DTW distance $D_{\max}$ as given in equation:

$$S = 1 - \frac{D[N][N]}{D_{\max}} \tag{16}$$

This method allows the distance values to be converted to a score between 0 and 1, where 0 indicates a complete mismatch, 1 a perfect match, and 0.5 and above a relative match.

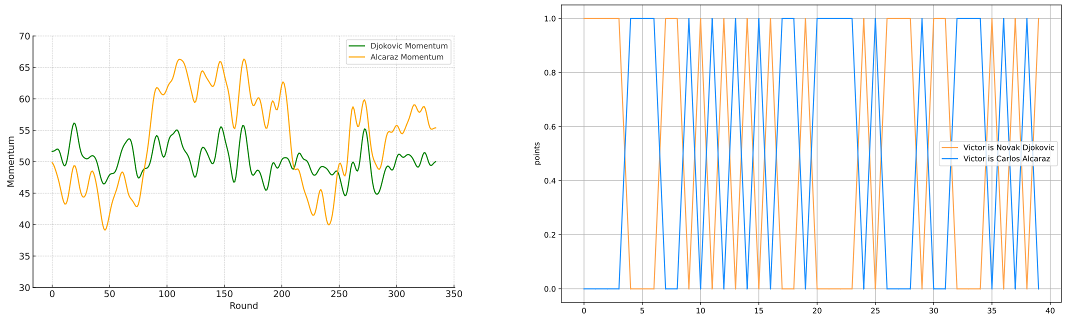
Result analysis

We have matched the momentum index curve with the player's point scoring curve and the game scoring curve. The criteria for judging are as follows: if the momentum index is higher during a certain period, it is more likely to score during that time, which means the scoring curve will show an upward trend. Specific curve pairs are shown in Figures 8 and 9.



(a) (b)

Figure 8: Momentum trend and point winner comparison chart



(a) (b)

Figure 9: Momentum trend and game winner comparison chart

In the end, we obtained a match of 88.635% between the momentum index curve and the player score curve, as well as a match of 91.156 % between the momentum index curve and the game score curve. This confirms that there is indeed a genuine correlation between the momentum index curve and the player score curve as well as that the concept of momentum does exist. Momentum is affected by a player's performance in a game and in turn affects the player's performance in the next game. Our model can effectively disprove the idea of coaching.

6 Model III:Predictive Model for Match Dynamics

6.1 Conversion of Momentum to Winning Probability

Starting from the definition of momentum obtained in the first problem, we assume that the momentum of two players in a match is M_1 and M_2. We can define the winning percentage of each player by normalising these two momentum values. The specific formula is as follows:

$$\begin{cases} P_1 = \dfrac{M_1}{M_1 + M_2} \\ P_2 = \dfrac{M_2}{M_1 + M_2} \end{cases} \tag{17}$$

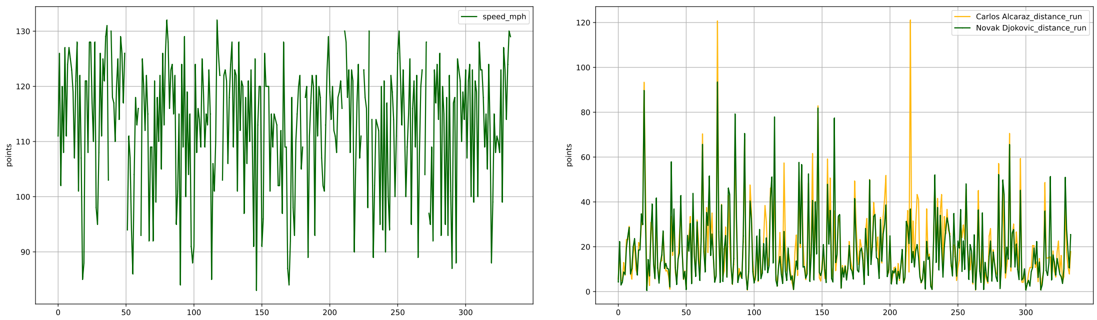
Where P_1 and P_2 represent the winning percentage of the first and second player respectively. In this way, we can ensure that the sum of the winning percentage of the two players is 1.

6.2 ARIMA Modeling for Forcasting Match Dynamics

Time series analysis allows forecasting based on the periodicity or seasonality of the data. Here, we have chosen the Autoregressive Integrated Moving Average (ARIMA) model, which combines the features of the Autoregressive (AR) model and the Moving Average (MA) model and introduces differencing to stabilise the data[18].

The ARIMA(p,d,q) model first differentiates the data and then transforms it into a stable time series for modelling.

Here, we imported the match data of Carlos Alcaraz and Novak Djokovic and using time series graphs, we found that many of the players' metrics showed long-term trends and volatility over the course of the match. The time series graphs for some of the metrics are shown below:



(a) (b)

Figure 10: Partial variable timing chart

Therefore, without being subject to any coercive factors, we predict the momentum change curves of the two players in subsequent matches based on their actual momentum in the first 200 matches.

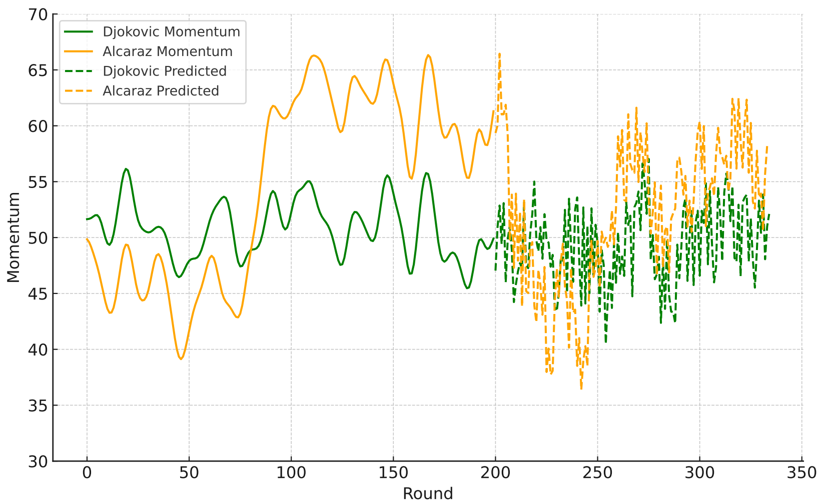


Figure 11: ARIMA model prediction chart

We observe that the momentum predicted by ARIMA differs from the actual momentum. This difference may be due to the fact that ARIMA is suitable for predicting the linear part of the time series. Therefore, we used the Long Short-Term Memory (LSTM) model to predict athletes' momentum more objectively and accurately.

Recurrent Neural Networks (RNNs) pass the output of a model at one moment in time as an input to the next moment in time[19], making them suitable for time series analysis. This approach allows for the prediction of future momentum based on an athlete's past momentum. The figure below illustrates how the "store" function retains the contents of the hidden layer, thus preserving the momentum information of the athlete from previous moments.

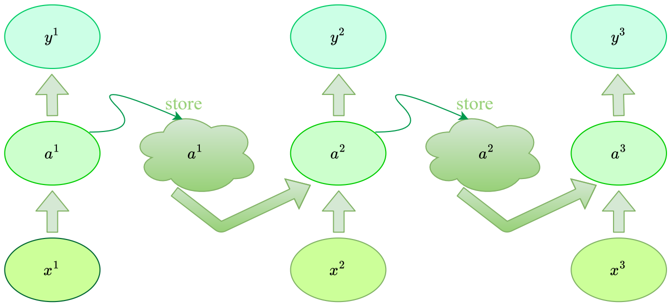


Figure 12: Schematic diagram of RNNs model

Here, we utilise the LSTM model, which employs a gating mechanism to control the output of each element in the state, effectively overcoming the limitations of simple Recurrent Neural Networks (RNNs)[20].

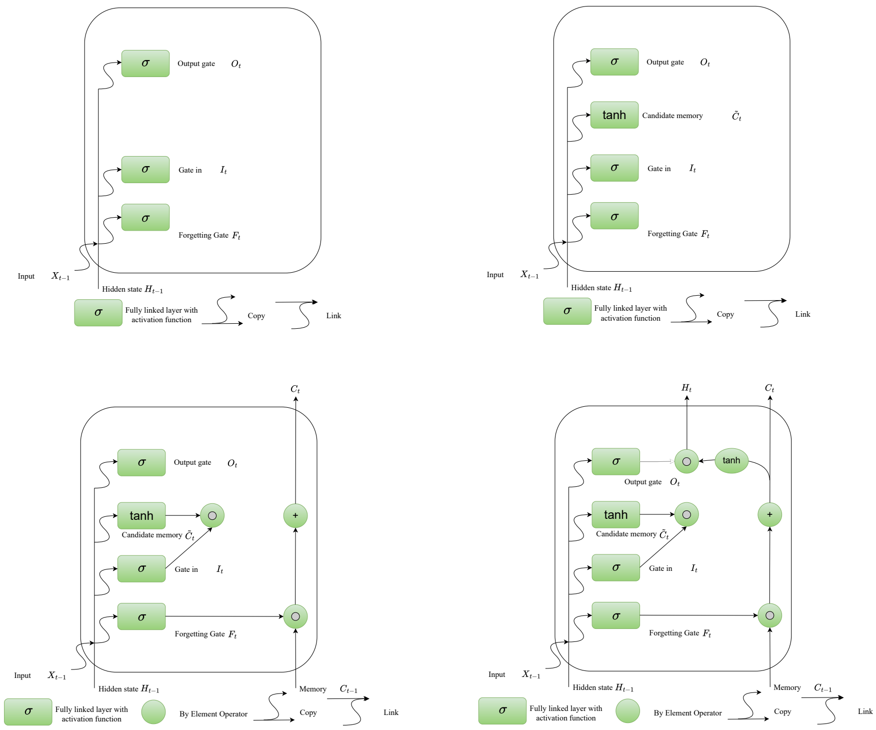


Figure 13: LSTM operating diagram

As previously mentioned, the forget gate is represented by f_t, which determines how much data from the prior internal state must be forgotten, and the input gate is represented by i_t, which determines how much data must be saved from the current candidate state. An output gate called o_t regulates the amount of data that must be output from the internal state that is currently in the external state following nonlinear activation. The following is its formula:

$$\begin{cases} f_t = \sigma(W_f^t x_t + U_f^T h_{t-1} + b_f) \\ i_t = \sigma(W_i^t x_t + U_i^T h_{t-1} + b_i) \\ o_t = \sigma(W_o^t x_t + U_o^T h_{t-1} + b_o) \end{cases} \tag{18}$$

The input is denoted by $x_t \in \mathbb{R}^{n\times d}$ in the formulas, and the hidden state from the previous time step is represented by $h_{(t-1)} \in \mathbb{R}^{n\times d}$. The bias terms for the forget gate, input gate, and output gate are b_f, b_i, and b_o, respectively, whereas the weights for the forget gate, input gate, and output gate are W_f^t, W_i^t, W_o^t.

The formula for calculating candidate memory cells is similar to the gates described earlier (input gate, forget gate, and output gate), but it uses the tanh function as the activation function. the output of the tanh function has an output range of $(-1,1)$, which helps to keep the memory cell values within a reasonable range and thus avoids the problem of vanishing gradients or gradient explosion. The following is the formula for calculating a candidate storage unit for the time step t:

$$\tilde{C}_t = \tanh(X_t W_{xc} + H_{t-1} W_{hc} + b_c) \tag{19}$$

In this equation, the candidate memory cell at time step t is represented by C_t in this equation, the input for the current time step is X_t, the hidden state from the previous time step is H_{t-1}, the weight matrix from the inputs to the candidate memory cells is W_{xc}, and the weight matrix from the hidden state to the candidate memory cells matrix is W_{hc}. The hyperbolic tangent activation function tanh transfers the weighted inputs to a range of values of $(-1,1)$, and b_c is the bias term of the candidate memory cell.

Based on the previously constructed formula for evaluating the athlete's momentum, we calculate the athlete's momentum value at each moment and construct an LSTM neural network with this label.

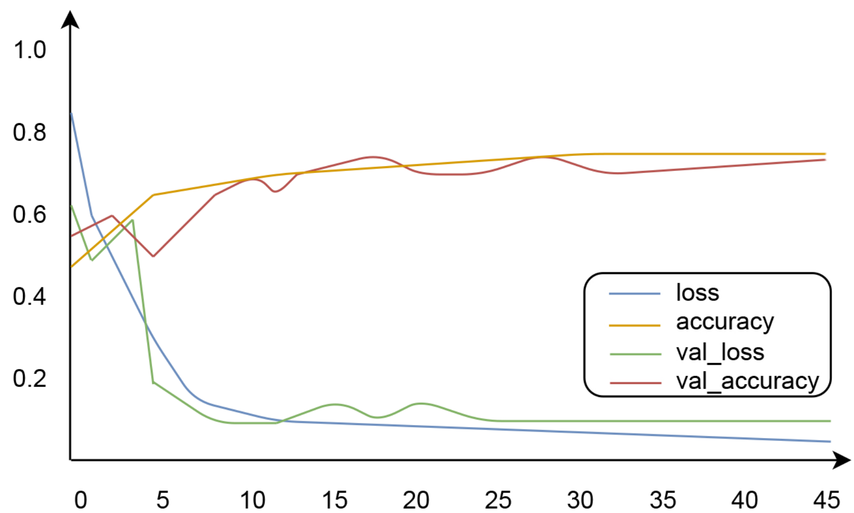


Figure 14: LSTM operating diagram

After 100 rounds of training, the model achieved an accuracy of 81.32% and the loss was reduced to 0.07. The learning curves shown in Fig. 2 indicate that the loss for both training and validation decreased overall over time and that the accuracy of the validation set was similar to that of the training set, which suggests that the model is not overfitting. The LSTM was successful in predicting the athlete's momentum at every instant in time, as seen by the comparison curve between the observed and predicted values of the athlete's momentum in the following figure.

We plot the winning probability curves for both athletes, as shown in Figure 15. It can be observed that the intersection points of the curves serve as turning points in the match.

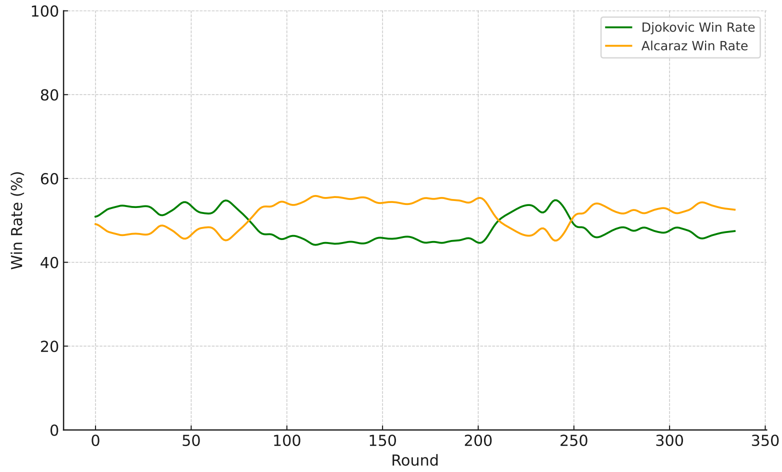


Figure 15: Player win rate curve

6.3 Recommendations for Player Performance

First, we simplify the model by defining that each athlete has a series of actions, each with a corresponding winning probability:

$$\begin{cases} S = \{s_1, s_2, \cdots, s_n\} \\ P = \{p_1, p_2, \cdots, p_n\} \end{cases} \tag{20}$$

The winning probabilities corresponding to each of these actions can be derived from the momentum statistics mentioned previously. When an opponent hits a ball in a certain way, their probability of winning is p_1 and the probability of winning our counter action is p_2. Then the probability of winning this action exceeds some threshold ε:

$$\begin{cases} p = p_2 \times (1 - p_1) \\ p > \varepsilon \end{cases} \tag{21}$$

At this point, the corresponding p_2 is:

$$p_2 > \frac{\varepsilon}{1 - p_1} \tag{22}$$

Therefore, actions with a probability of winning greater than p_2 should be selected from the sequence of actions in the counterattack.

For example, the depth of the second serves of Carlos Alcaraz and Novak Djokovic in a match is shown below:

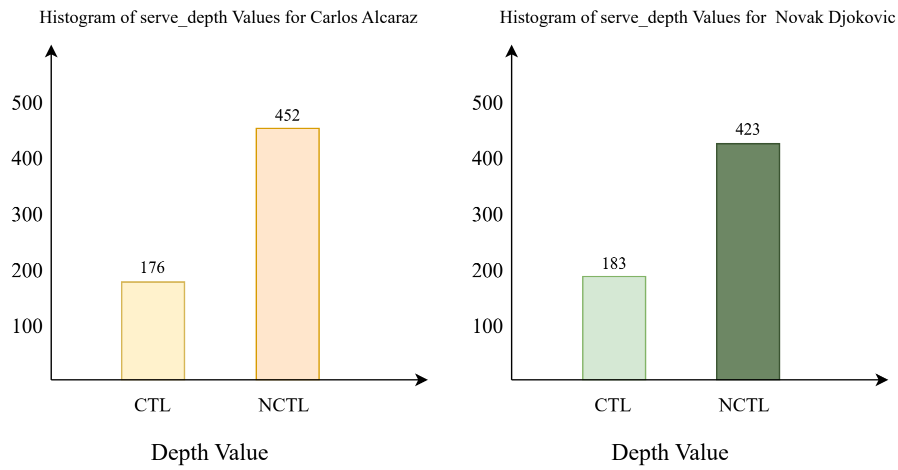


Figure 16: Player win rate curve

When we set the threshold to $\varepsilon = 0.8$, assuming that Carlos Alcaraz has a higher probability of having a shallow depth of serve on his second serves, then Novak Djokovic should consider shifting his position back a little to improve the quality of his returns.

6.4 Some recommendations for players

- **Opponent**
  a. **Data Analysis**
  It is important to thoroughly analyse your opponent before a match. Use historical video data and other resources to gain insight into your opponent's performance and win percentages in past matches. Analyse how your opponent reacts in a given situation.
  b. **Simulation Training**
  Based on the opponent's data characteristics, conduct pre-match simulations and formulate tactics in advance.
- **Player's Self**
  a. **Data Analysis**
  It is essential to know yourself thoroughly before a match. Identify the situations you excel in and your winning percentage, analyse your moments of high or low momentum and work on strengthening your weaknesses.
  b. **Psychological Factors**
  Mental factors are often the most influential factors on the outcome of a game, apart from an athlete's physical ability. Good athletes have good psychological qualities, which enable them to remain calm under pressure and turn pressure into motivation, thus achieving excellent results.
- **On-site Factors**
  a. **Strategic Adjustments**
  During a match, players should have a strong grasp of the situation on the field and know exactly when they have the upper hand and when they are at a disadvantage. Athletes need to adjust their mindset and strategy in a timely manner, accurately manage their physical condition and adopt different strategies at different stages of the game.

b. **Coaching Factors**

Athletes should communicate with their coaches during breaks. Coaches should use the model to analyse the game, provide key guidance to athletes during short breaks, address technical issues and point out weaknesses in their opponents. At the same time, coaches should also aim to stimulate momentum in their athletes to improve their performance.

7 Momentum Index Modelling on Other Races

The Heat and Celtics meet again in the 2023 NBA Eastern Conference Finals in a best-of-seven series. Butler's three-point shutout fell short, but he vowed to come back for the kill in the same situation. However, this time the Heat beat the Celtics 103-84 on the road to eliminate their rivals 4-3 and advance to the Finals, where they will face the Denver Nuggets.

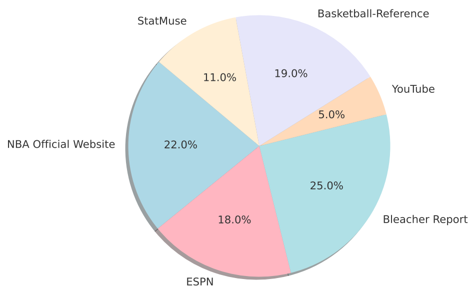


Figure 17: Pie Chart of Data Sources

Table 3: Website of Data Source

Database Websites
https://www.youtube.com
https://www.bleacherreport.com
https://www.espn.nl
https://www.basketball-reference.com
https://global.nba.com/scores/#!/

During the game, both teams shot a low percentage from the field, with the Heat hitting only two shots halfway through the first quarter and the Celtics missing all of their first 12 threes. This led to a low momentum index graph at the beginning of the game with small swings for both teams. However, the Heat's Momentum Index gradually grew.

Mazzulla pep-talked his team before the game, stating that the game might be long, but the players should enjoy it. However, the Celtics were slow to find their form in the steal. In the first quarter, the Heat took the initiative to call a timeout when they were ahead, only to be overtaken by the Heat's 11-2 offensive wave.

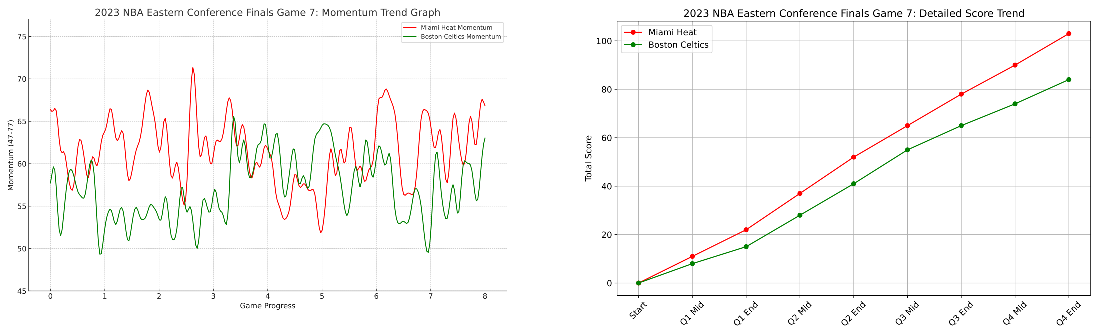
The Celtics scored a total of five points on 2-of-15 shooting in the second seven minutes of the first quarter, leading to a 22-15 lead for the Heat at the end of the first quarter, the same as the Celtics' lowest scoring first quarter of the season. The Celtics' sluggish offence did not improve in the second quarter, as the Heat started the second quarter with a Robinson three-pointer and a Hay-Smith steal of Tatum with an assist to Vincent for a layup, leading to the Heat calling another game. The Heat extended their deficit to as many as 18 points during the second quarter. Though the Celtics were able to chase down consecutive points, led by Tatum, they still trailed the Heat 41-52 at halftime. This is consistent with the Momentum Chart performance, where the Heat led in scoring throughout the first half, and both teams' Momentum Indexes were on the rise. Additionally, the Celtics struggled from three-point range in the first half, hitting only four of their 21 attempts for a 19% clip.

It's worth noting that on the Celtics' only off day, when they were down 0-3, Horford suggested that Mazzulla forgo the video lessons and gather the team to play golf to relax. Gervais recalled the occasion this way: "We skipped the video session altogether, we got away from basketball and focused on each other, trying to get our camaraderie back. We kind of lost touch under the pressure, but we were able to enjoy time with each other again."

As it happens, the module of game pressure is also considered in our momentum index model, and at one point in the post-break day game, the Celtics overtook the Heat in the momentum index, and since then, the Celtics have pulled out three straight games Obviously, it's obvious that for the Celtics, it's how to handle the pressure that's the team's obsession.

When the Celtics will be close to the score, this section, Martin hit a number of difficult shots, a single section of 4 shots to get 9 points, the end of the three sections of 14 shots 10 contributed 23 points and 8 rebounds. Greatly encouraging the morale, the Heat's momentum once again achieved rapid growth, successfully overtook the Celtics. In the subsequent fourth quarter, the Heat continued to team bloom, once the difference will be opened to more than 20 points, the NBA East final this suspense end.

After collecting various data about this game, we substituted them to calculate the momentum of both teams and plotted the momentum index curve against the total goals scored in this game as shown below:



(a) (b)

Figure 18: Momentum Index and Section Score Data for Two Teams

The two curves matched by 65.918%.

8 Sensitive analysis

The sensitivity of the model is crucial to its stability, and it is related to whether the model will be disturbed by some other factors. In order to test the sensitivity of the model, we increased and decreased speed_mph by 0.5% respectively, and then input it into lstm to observe the change of accuracy when lstm predicts. The prediction results for these three cases are shown below:

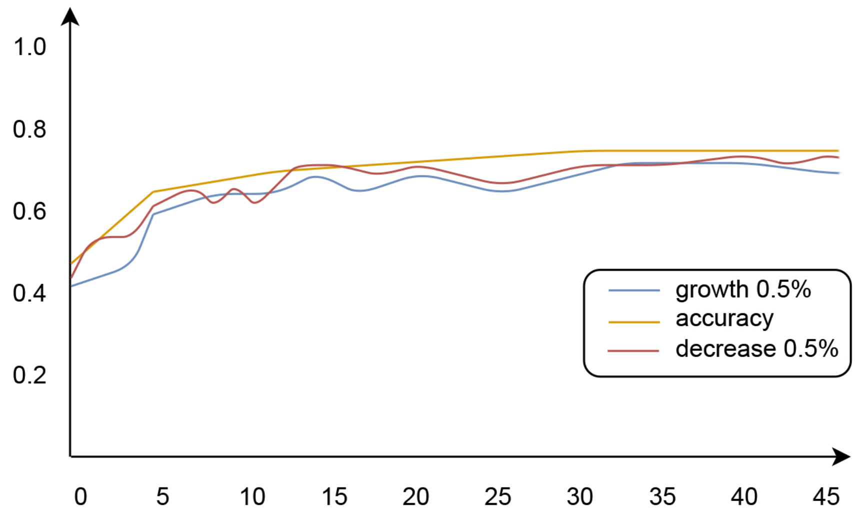


Figure 19: Sensitivity test curve

As can be seen from the figure, when increasing or decreasing the speed_mph by 0.5% for speed_mph, the trend of the accuracy level does not change and stays within the acceptable range. Therefore, we can conclude that our model is stable and capable of solving practical problems.

9 Strengths and Weaknesses

9.1 Strengths

- Our model provides a more objective and comprehensive quantitative portrayal of player momentum, with G1-entropy weights-independence weights integrating the subjective and objective aspects of momentum, and this approach is not only applicable to tennis, but also relevant to the game in general.
- The DTW model reasonably demonstrates that momentum exists and more accurately compares the correlation between momentum and fluctuations in athlete levels.
- The lstm model can predict complex nonlinear problems and addresses the shortcomings of the traditional ARIMA model.

9.2 Weaknesses

- The predictions of the model are based on the metrics of the given dataset, but the presence of metrics that influence the predictions but are not in the dataset may result in a less accurate model.
- The lstm model requires a large dataset, however limited data can lead to model overfitting.

10 Conclusion

In summary, this paper firstly establishes a G1-entropy weight-independence weight evaluation model, which comprehensively evaluates the momentum from subjective and objective aspects respectively. We determined the weights of each index, and made a quantitative mathematical assessment of the player's momentum.

We then used the DTW model to determine the similarity between the momentum curve and the player's score curve, thus proving that momentum exists, and showing that there is indeed a momentum factor that determines the movement of the outcome of a tennis match.

Then, we used ARIMA model to predict the momentum of the athlete, and found that its matching degree was not very accurate, we analysed that this might be related to the fact that ARIMA is good at predicting the linear part. So we used the lstm model to predict the momentum of this complex non-linear problem, and found that the results are better. We defined the winning percentage with the help of momentum, made a winning percentage graph and found that the intersection of the winning percentage curves of two athletes is the turn of the match. We have analysed the game briefly using game theory and have given some suggestions to the athletes.

Finally, a sensitivity analysis of the lstm model was done and it was found that when increasing or decreasing the speed_mph by 0.5% for the speed_mph, the accuracy of the model varied within acceptable limits.

11 The future and outlook of modelling

Our momentum index model has done an excellent job in that problem, being able to correctly determine how high or low a player's momentum is during a match, and thus how well the player is performing. However, since there is more to actual matches than just the data provided, the model is not accurate enough and does not generalise well enough in new tennis tournaments as well as other tournaments, which is an obvious problem.

Therefore, we have the following outlook on the model and anticipate that it will be updated and further refined in the future in order to increase the model's accuracy and generalizability. Firstly, the sample categories of the dataset should be richer, and factors such as weather and players' head-to-head record should be included. Secondly, the number of data samples should be more huge, only when the number of samples goes up, the model is not easy to be overfitted and the generalisation can be improved. Then, our model is rough, and the matching degree only considers the DTW algorithm, which should be considered with more methods, or even a combination of different methods to make the results more accurate. Finally, our model lacks comparison, similar to logistic regression, random forests and other methods we have not modelled and calculated, resulting in a lack of side-by-side comparisons of our model, it is better, but not necessarily the best model.

Not only that, in model III, we also only consider ARIMA and LSTM two models, the use of models is not broad enough at the same time, but also shallow, if the other time series methods or ARIMA and LSTM for the integrated use of the model, perhaps the model will be more accurate, there will be better models and algorithms used in the LSTM.

In the limited time we had, we built an excellent model, but clearly not the best one. We hope that in the future we can improve our model by more efforts, and we also hope that our model can play an innovative role in solving this kind of problems and lay a solid foundation for future generations. We hope that our model has a bright future.

MOMENTUM OF THE ATHLETE

Using momentum to predict the direction of a game

1.What is an athlete's momentum?

In sports, many athletes feel that they have momentum, a magical fluctuation, and that the situation may proceed with this fluctuation during the course of the game. In physics, momentum is a physical quantity that describes the state of motion of an object and is related to both the mass and speed of the object. [1]In fact, in many match events, "momentum"is implicitly determining the direction of the match. Momentum is also used in a number of sports. [2]Therefore, our team investigated a comprehensive G1-Entropy-Independence evaluation method to quantify the momentum of athletes by studying and analysing a large amount of data.

Figure 1

2.Does momentum exist?

A tennis coach once questioned the existence of "momentum"in a player, arguing that a player's performance fluctuations and scoring streak were random. In response, we developed DTW's Momentum-Performance Matching System and found that the fluctuation curve of a player's performance is highly correlated with the momentum curve, which shows that momentum can be a rough measure of a player's performance. For example, if a certain player scores consecutive points, the corresponding momentum curve also shows higher values. The table below shows the correlation between the player's volatility curve and the momentum curve for seven tennis matches, and it is found that the similarity between the two curves is above 0.9 for almost every match.

Table 1

Match	Player1	Player2
1	0.912	0.924
2	0.879	0.906
3	0.941	0.927
4	0.938	0.913
5	0.886	0.912

3.How to predict?

Our team first used the arima model to predict the momentum in the match, but found that the actual momentum may not match well with the momentum obtained from the prediction, which may be related to the fact that the ARIMA model is better at predicting the linear part. Therefore, our team went on to try the lstm model and found that it worked better with an accuracy of 76%. Then we defined the winning percentage of the match with the help of momentum. As shown below, the intersection of the winning percentage of two athletes is the turning point of the race.

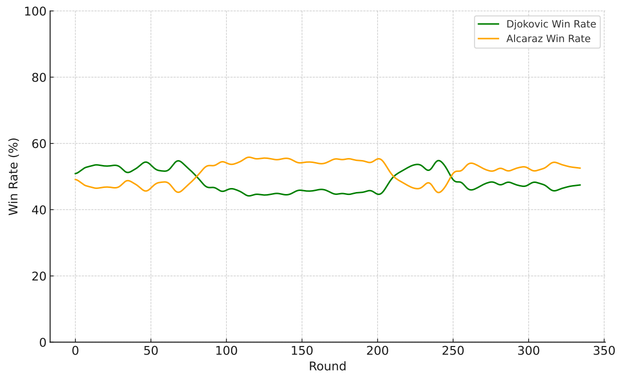
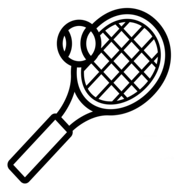


Figure 2

4.What it says about athletes?

After we have quantified momentum, we can firstly analyse all aspects of ourselves and our opponents. Players can fully analyse their opponents before the match to formulate a strategy, analysing the situations in which their opponents excel versus their weaknesses in catching and multihitting. In this process, we use game theory to find ways to maximise the player's gains. In addition, the amount of momentum is also related to the psychological state of the players, good psychological quality and strong mental strength can bring high momentum for the players, which helps them to win the game. At the same time, this momentum also gives the athletes a message: adjust the mentality without nervousness, and use their advantages to take the game.



[1]Fitts, P. M., & Widrick, J. J. (1996). Force-velocity and impulse-momentum relationships: Implications for resistance training. *Journal of Sports Sciences*, 14(3), 265-275.
[2]https://www.merriam-webster.com/dictionary/momentum

References

[1] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735-1780.

[2] Maalouf, M. (Year). Logistic regression in data analysis: An overview. Journal/Publisher.

[3] Author(s). (Year). Title of the article. Journal/Conference.

[4] https://medium.com/draftkings-engineering/building-a-tennis-simulation-d6afdaa97d19

[5] Pan, Ouyang, Wu, Ecaterina, Ion, Jiao et al. (Year). Title of the Article. *Frontiers in [specific field or journal section]*.

[6] Shannon, C.E. (1948). A Mathematical Theory of Communication. Bell System Technical Journal, 27(3), 379-423.

[7] Shannon, C.E. (1948). "A Mathematical Theory of Communication". *Bell System Technical Journal*, 27(3), 379-423, 623-656.

[8] Little, R.J.A., & Rubin, D.B. (2014). This work is a foundational reference on missing data imputation.

[9] Han, J., Pei, J., & Kamber, M. (2011). *Data Mining: Concepts and Techniques* (3rd ed.). Morgan Kaufmann.

[10] https://www.infona.pl/resource/bwmeta1.element.ieee-art-000006115088

[11] Zeleny, M. (1982). "Multiple Criteria Decision Making". *McGraw-Hill*.

[12] Li, X., & Liu, L. (2015). "A novel method to determine the weights of indicators in a combined evaluation model based on entropy weight". *Journal of Computational and Applied Mathematics*, 277, 73-80.

[13] Wang, Z., & Zeng, S. (2009). "A multi-criteria decision-making method based on a cross-entropy with interval numbers". *Knowledge-Based Systems*, 22(8), 586-589.

[14] Li, R., & Li, D. (2005). The Empirical Research on the Comprehensive Evaluation of the Regional Economic Competitiveness Based on Entropy Method. Systems Engineering, 23(2), 22-26.

[15] Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(1), 43-49.

[16] Berndt, D. J., & Clifford, J. (1994). Using dynamic time warping to find patterns in time series. In *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining (KDD-94)* (pp. 359-370).

[17] Keogh, E., & Ratanamahatana, C. A. (2005). Exact indexing of dynamic time warping. *Knowledge and Information Systems*, 7(3), 358-386.

[18] Box, G.E.P., & Jenkins, G.M. (1970). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.

[19] Graves, A. (2012). Supervised sequence labelling with recurrent neural networks. In *Studies in Computational Intelligence* (Vol. 385). Springer, Berlin, Heidelberg.

[20] Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. *Neural Computation*, 12(10), 2451-2471.

Report on Use of AI

1. Chatgpt 4.0

Query1: Can DTW algorithm be used to calculate curve matching degree?

Output: Yes, the Dynamic Time Warping (DTW) algorithm can be used to compute the similarity between curves, effectively measuring the degree of match between them. DTW is a powerful technique that allows for the comparison of two temporal sequences which may vary in time or speed. In the context of curve matching, DTW can align points on two curves in a way that minimizes the overall distance between them, taking into account the temporal distortion between the curves. This makes DTW particularly useful in applications where the shapes of the curves are similar but they are not perfectly aligned in the time axis, such as in time series analysis, speech recognition, and in the analysis of any data that can be represented as a curve or a sequence over time.

2. Chatgpt 4.0

Query2: What are the basic steps of DTW?

Output: The basic steps of Dynamic Time Warping (DTW) include:

- creating a cost matrix to store distances between elements of two time series,
- using dynamic programming to fill this matrix with the aim of finding the minimum cumulative distance, and
- tracing back from the bottom right corner of the matrix to find the optimal path that aligns the two series. This process allows for the comparison of two sequences that may vary in time, by optimally aligning them despite their differences in speed or timing.

3. Chatgpt 4.0

Query3: Whether LSTM can be used for time series prediction?

Output: Yes, Long Short-Term Memory (LSTM) is a type of deep learning model that is suitable for time series prediction. LSTMs are a variant of recurrent neural networks (RNNs) designed to handle and learn long-term dependencies in time series data.

Time series data exhibits temporal patterns with data points changing over time. LSTM models, with their internal memory units, can capture and learn long-term dependencies in sequences, making them well-suited for time series prediction tasks.

When using LSTM for time series prediction, it is common to split the data into training and testing sets. The model is trained on the training set, and then its performance is evaluated on the testing set. The model takes past time steps as input and predicts future time steps.
In summary, LSTM is a powerful tool for time series prediction, but careful tuning of model parameters, selection of appropriate features and data preprocessing methods, and thorough model evaluation and optimization are essential for achieving optimal performance in specific applications.

4. Chatgpt 4.0

Query4: What are the disadvantages of LSTM?

Output: While LSTM is a powerful deep learning model, it does have some drawbacks, including:

- High Computational Cost: LSTM models often require significant computational resources, especially when dealing with large datasets or complex model structures. This can result in long training times and a need for powerful hardware.
- Complex Hyperparameter Tuning: LSTMs have many hyperparameters to tune, such as learning rate, the number of layers, and the number of hidden units. Choosing appropriate hyperparameters may require extensive experimentation and experience.
- Risk of Overfitting: On small datasets or without proper regularization techniques, LSTMs may be prone to overfitting, leading to poor performance on unseen data.
- Difficult to Explain: Deep learning models, including LSTMs, are often considered black-box models. Explaining the decision-making process and how specific patterns are learned can be challenging.
- Vanishing and Exploding Gradients: Despite being designed to address the vanishing gradient problem, LSTMs may still encounter issues of vanishing or exploding gradients in certain situations, affecting training stability.
- Data Intensity: For more complex tasks, LSTMs typically require a large amount of training data to fully leverage their advantages. In situations with scarce data, overfitting or insufficient learning of complex patterns may occur.

Despite these drawbacks, LSTM remains a preferred model for many time series and sequence modeling tasks. However, its advantages and limitations should be carefully considered and balanced based on specific application requirements.