Problem Chosen	2024 MCM/ICM Summary Sheet	Team Control Number
C		2429101

Wimbledon is a 7-round, best-of-5 sets tournament played every year at the All England Lawn Tennis and Croquet Club. Given the prize total exceeding $40,000,000 last year, it is important for players to take any opportunity they have to optimize their style of play against uncertainty.

We begin by identifying flaws in the dataset, particularly in the elapsed time aspect of the data, and correcting them. We then propose and implement a novel metric for determining the lead that one player possesses by computing the probability that they will win the match from the given score and current server. We then used **Single Exponential Smoothing** to temper our volatile data and have a model that uses the previous points to inspire future values. From this metric, we took the derivative of this function to compute our momentum value at every given point. To determine if this value carried any meaning beyond random noise, we calculated the expected performance for future points and compared this to the actual outcome. We found that when a player had positive momentum, their ratio of outperforming expectations with momentum vs without it was over $1.4\times$ when considering 4 points in advance, and with this trend growing stronger as the tournament proceeded. We also discovered that momentum has a tendency to last longer periods, as outperformance ratio tended to increase when evaluating longer horizons with similar expectation thresholds. We classified points as momentum swings if the curve crossed the x-axis and determined the direction of the change in order to implement our next model.

After developing a model for the lead one player has over another and quantifying the changes in momentum over the course of a match, we develop and train a **Gradient Boosting Classifier** to identify likely points of swing. This model performs moderately well, with an F1 score of **51%** when predicting positive momentum swings for a player and an F1 score of **52%** when predicting negative momentum swings for a player. We also give a short discussion of the reason for using F1 score over accuracy or other metrics in our evaluation of the model. These performance values guide our conclusion that it is in fact possible to predict when a swing in momentum might occur during a match and whether this swing will be in a positive or negative direction for a given player.

We then analyze the relative feature importances of the various predictor variables that are input into our model. We conclude that 3 metrics in particular are the most important for a player to swing momentum in their favor (in addition to/aside from winning the current point): their **current momentum**, the **distance traveled** by both players, and the **speed** at which the ball is served.

We then analyze one particular match in depth, the 5th round match between Daniil Medvedev and Christopher Eubanks. On this match, our predictive model achieves F1 scores of over **71%** and **72%** on predicting positive and negative shifts in momentum, respectively.

We finally analyze the sensitivity of our models. For the lead model, we conclude that the model is highly resilient to reasonable perturbations in its only parameter (namely, the probability of winning a serve point). We also conclude that with the exception of a few highly essential features, our Gradient Boosting Classifier is relatively resilient (unsensitive) to the loss of single predictor variables.

Our model carries a strong benefit in its **high generalizability**, as it is possible to use this method of tracking lead in different sports to calculate momentum and swings in a similar way. We also note that this analysis has implications for other Grand Slams which are played on different surfaces and it can also easily be applied to women's tennis, as the most notable difference is a reduced number of sets. With these conclusions about matches between pros, we speculate that on an amateur or recreational level, the effects of momentum will be even more pronounced.

We finally conclude by writing a letter to tennis coaches regarding how to predict when swings in momentum may occur and how best to coach their players to keep momentum in their favor, swing momentum in their favor, and in the end, win the match.

Contents

1 Introduction

1.1 Problem Restatement

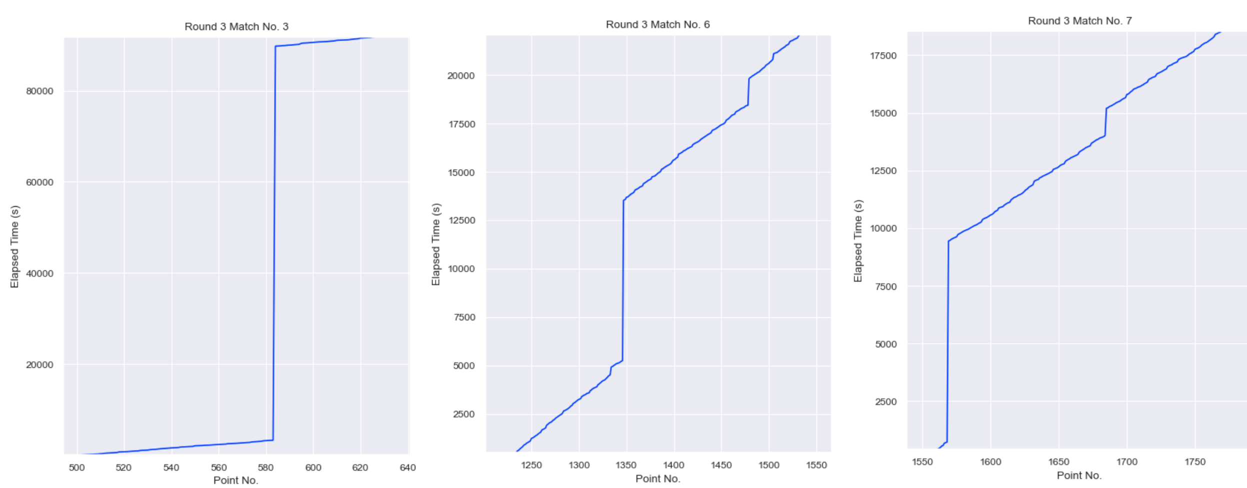
With these ideas about tennis scoring in mind, we aim to address the following problems in our paper:

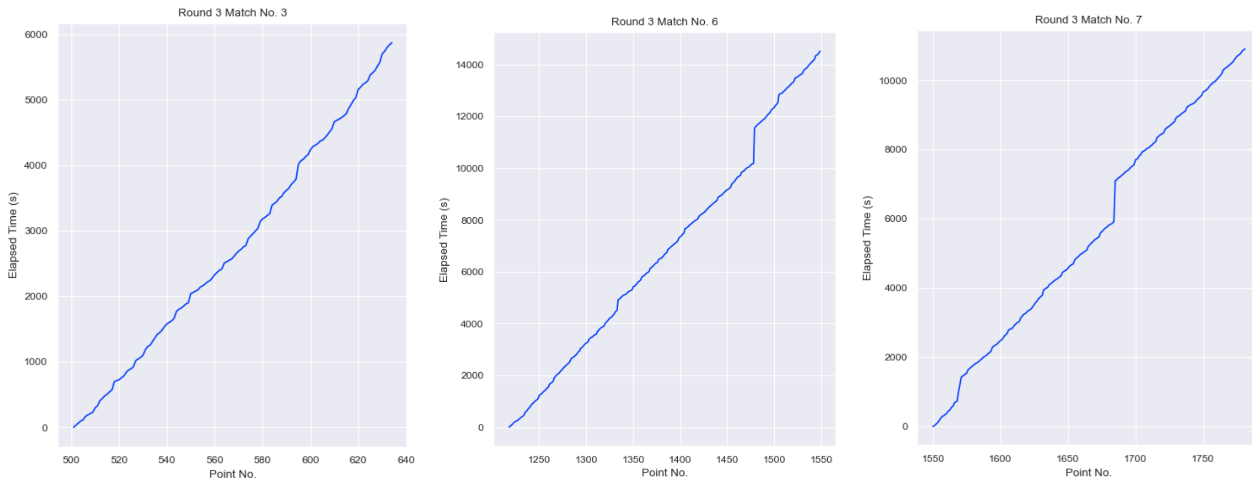
- **Task 1:** Create a model to quantify how far ahead a player is based solely on the score and who is serving. This allows us to evaluate the flow of the match over time.
- **Task 2:** Address the skeptical coach, i.e. determine how much of an effect momentum has on the outcome of a match and if it can be explained by random swings.
- **Task 3:** Systematically identify momentum swings and develop a model to evaluate the causes of these swings. Additionally, determine if there is a way for players to observe these effects in action and how to react.
- **Task 4:** Test the extent to which our modeling is applicable to other tennis tournaments and to other racquet sports.
- **Task 5:** Write a letter to tennis coaches regarding the ways to keep momentum in their players favor and how to determine if a swing is about to occur.

2 Data

2.1 Data cleaning

While examining the provided dataset, there were a few problems that we encountered regarding games that had a delay. Two of the games in the tournament had a rain delay while another game had a jump of 24 hours in the data. In order to fix these issues, we wrote a python program to clean these jumps in timings. We graphed our results to showcase the changes:





After doing so, we then had to preprocess the data as there were a handful of columns in the csv file that were categorical data (i.e., columns labelled: serve_width' , serve_depth, return_depth). We opted to use one hot encoding to transform these columns into data points that can be used for our machine learning models.

3 Assumptions

- **Assumption 1:** When evaluating the distance a given player is ahead, we only factor in the score (e.g. 1-0 set score, 3-2 game score, 15-0 point score) and if they are currently serving.

  Reason 1: For tasks 1 and 2, we approach this problem from the perspective of the skeptical coach, that is to say that we do not know if momentum exists and if it has a nontrivial effect on a point, game, set, or even match. Thus, when computing how much of a lead a player has, we do not want to introduce confounding variables such as ace probability or winner probability, since these would be a circular assumption about the existence of momentum. In sections 5.2 and 5.3, we evaluate the effects of momentum and will make conclusions that influence the remainder of the paper, but initially our model must remain agnostic of factors beyond score and serve.

- **Assumption 2:** The probability that a player wins a service point is treated as constant for all players and for all matches.

  Reason 2: In the same vein as the above assumption, the probability that a player begins to have a weaker service percentage, slower serve, lower ace percentage, etc. is intuitively correlated with momentum, thus using this data in our initial evaluation would serve as circular reasoning. However, using the same serve for all players during all rounds of the tournament is a limitation of our model that could be rectified if we had access to more data from more Grand Slams. We will evaluate the effects that momentum has on serve performance in this paper and perform sensitivity analysis on our chosen constant probability.

- **Assumption 3:** We will ignore the existence of match to match momentum, i.e. every player starts with neutral momentum at the start of every match.

  Reason 3: It is certainly true from intuition and personal experience that a player can accumulate momentum over the course a tournament, especially if they are a bottom-seed or top-seed (underdog belief and top-dog confidence respectively), and this could certainly play a role in matches. However, with the 31 matches that we are provided it is difficult to determine how much of an effect this has, though a more complete model than ours with a larger dataset may find results about this.

- **Assumption 4:** We will ignore knowledge about a player's past performance and previous head-to-head matchups between players.

  Reason 4: Similarly to assumption 3, it tends to be true that a player who played a 5-set match in the semi-final tends to perform worse in the final than a player who played an easy 3-setter. Additionally, previous matches between players can have a psychological effect going into a match even at the professional level. We believe it to be outside the scope of the tasks above to determine the effects of this, though in our discussion of our strengths and weaknesses we will provide ideas on how to approach these effects.

4 Notation

Symbol	Definition
p_s	Probability of winning a service point
$L(t)$	Lead of player 1 at point t
$m(t)$	Momentum of player 1 at point t
$\tau_i(t)$	Total number of points won by player i after t points
α	Smoothing factor for exponential smoothing
$\hat{L}(t)$	Smoothed lead curve of player 1 at point t
$t = \mathbf{0}$	The first point of the match
δ_T	Threshold value for actual minus expected

Table 1: Notations

In addition, when representing tennis scores in writing, they will be denoted

$$t(\mathbb{S}, \mathbb{G}, \mathbb{P}, a) = [(S_1, S_2), (G_1, G_2), (P_1, P_2)\ (a)]$$

where S_i, G_i, and P_i are the sets, games, and points held by Player i respectively, and a is the number of the player serving. For instance $t((1, 0), (5, 3), (40 - 0), 2)$ corresponds to the case when player 1 won the first set, leads $5 - 3$ in the second, and has three chances to break player 2's serve.

5 Quantifying Flow

5.1 Lead in Tennis

Unlike many sports, tennis has a layered scoring system and thus the degree to which one has an advantage at any given point is more complex to quantify. In basketball, for instance, the lead one team has over another is essentially captured by the number of points they are ahead, while in tennis it is a bit more complex to capture how much one is ahead when the score is $2-0, 0-5, 0-40$ (2).

To address this, we developed a novel model for measuring the lead one player has on another through the following thought experiment: suppose at a given point in the match (e.g. when the score is $2-0, 0-5, 0-40$ (2)), Player 1 and Player 2 are substituted out for "replacement-level" players (i.e. players that perform averagely compared to all other players at that level of play) named $R1$ and $R2$. Clearly, without loss of generality, $R1$ will stand a much better chance of winning the match if Player 1 left him with a far bigger lead than if not. Hence, the absolute lead Player 1 has over Player 2 can be measured by the probability of Player 1's replacement player beating Player 2's replacement player.

The use of replacement-level players in the analysis of sports is not new. Baseball famously uses a metric called WAR (wins above replacement) to determine the value in raw wins of a player compared to a league replacement [3]. Furthermore, a breakthrough paper by Yurko, Ventura, and Horowitz uses the notion of replacement-level players in football to accurately estimate the contributions of each player on the gridiron to a particular play, and by extension, to the outcome of the game [2].

We formalize this notion of lead by computing the probability of a given point being won by a replacement-level player in this section.

To begin to understand how to address these concerns, we can create a diagram to determine the probability that a player wins a service game, which will allow us to extrapolate:

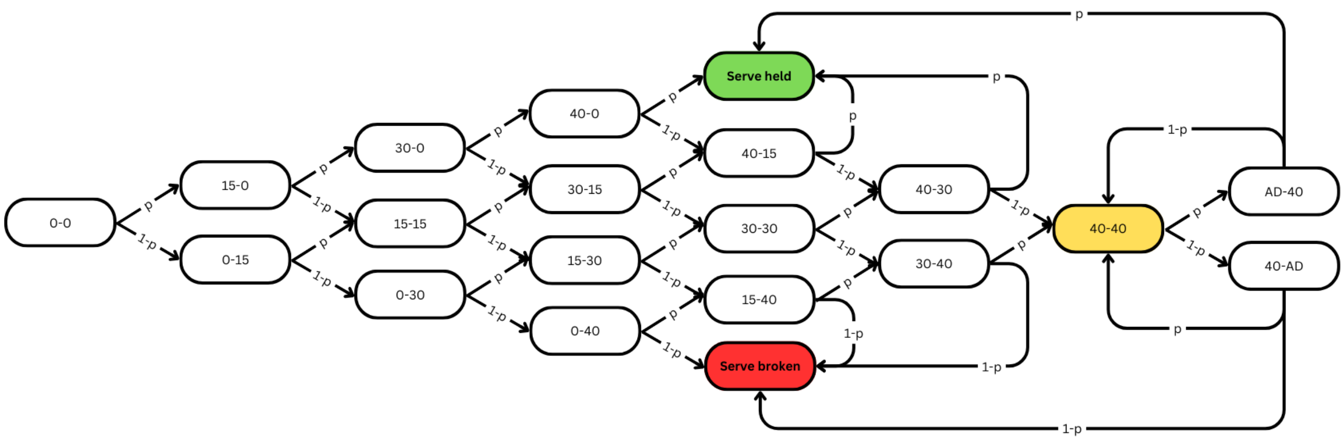


Figure 1: Flow diagram for the outcomes of a service game, with $p = p_s$

Without loss of generality, consider player 1. Then define p_c as the probability that player 1 wins the current game from position $\mathbb{P}(a)$. Then we have

$$L(t(\mathbb{S}, \mathbb{G}, \mathbb{P}, a)) = p_c L(t(\mathbb{S}, (G_1 + 1, G_2), (0, 0), a')) \\ + (1 - p_c) L(t(\mathbb{S}, (G_1, G_2 + 1), (0, 0), a')),$$

where a' is the next player to serve. To rephrase, we can recursively compute the lead that a player has by looking into the future, since it is possible to compute p_c from any given point above by using the chart. With some careful attention to edge cases like tiebreaks (10pt for the final set, 7pt otherwise), we can compute the lead value across various matches in the tournament:

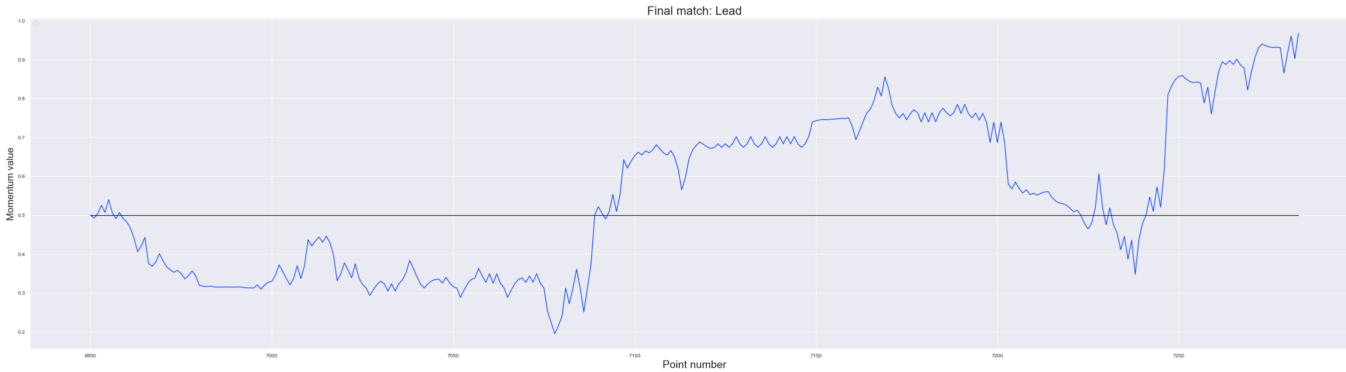


Figure 2: Chart with $L(t)$ for the finals

One key aspect of this chart is the selection of the value for p_s. We found that the average across this tournament in particular was roughly 0.67, as evidenced in the figure below, but this can be very different depending on the tournament. The effects of this are discussed at length in sections 7 and 9.

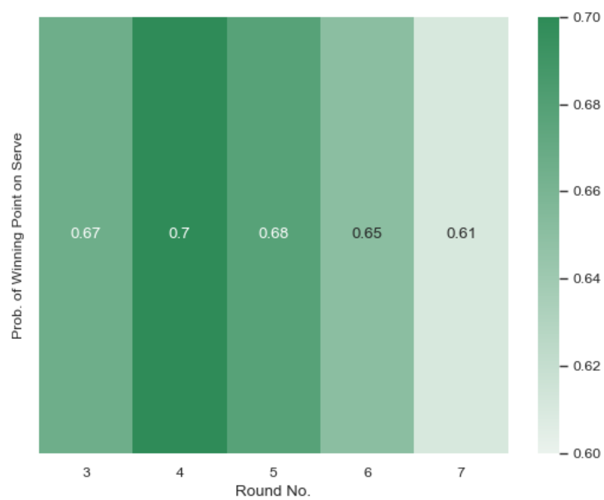


Figure 3: Probabilities of winning a serve at various rounds of the tournament.

As was expected, the charts for the lead are very unsmooth due to the nature of the sport. In the later sets and in tiebreaks especially, there can be wild fluctuations in the lead that a player has. In order to perform an accurate analysis, we will need to address this.

5.2 Defining Momentum

All athletes have intuition for the way that momentum impacts a match: a streak of favorable exchanges cause an increase in confidence which lead to better performance. While we will not attempt to quantify confidence in this paper (see section 9), we are able to look at how the lead changes over time. Conceptually, an prolonged increase in lead will correspond to that player having positive momentum.

Therefore we are justified in writing

$$m(t) = \frac{dL}{dt} \tag{1}$$

However, naively taking a derivative of our lead graphs will create problems, since we see that momentum is changing almost every point which breaks away from our intuitive understanding. To account for this, we can use Single Exponential Smoothing to allow past lead values to influence the present:

$$\hat{L}(\mathbf{0}) = L(\mathbf{0})$$
$$\hat{L}(t) = \alpha L(t) + (1 - \alpha)\hat{L}(t - 1)$$

By performing this past-influenced smoothing, we allow the data to reflect trends of leads rather than quick changes, but we still want to allow huge swings to occur. We found via experimentation that $\alpha = 0.2$ was the most successful for incorporating both the past and the future. We also experimented with double-exponential smoothing, but our data does not have sufficiently consistent trends to allow this to work well. Thus, we saw

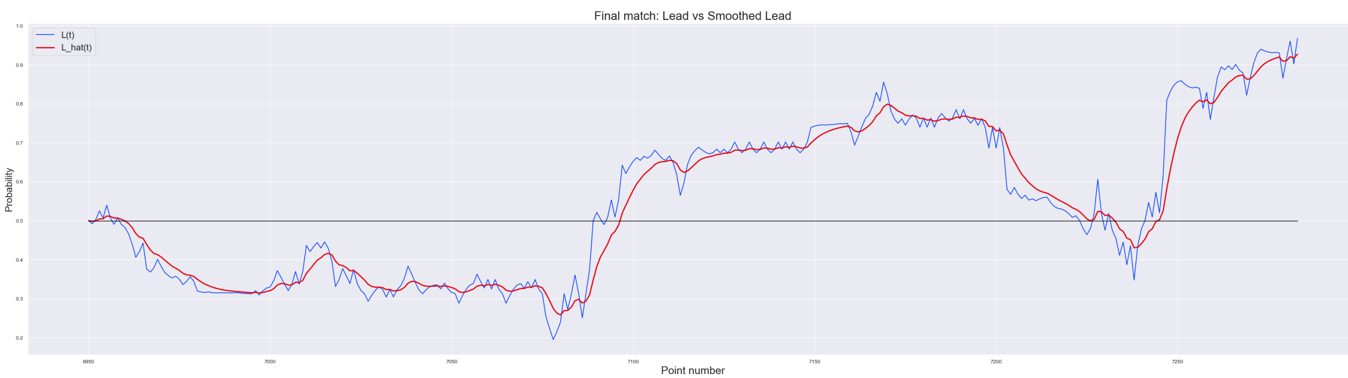


Figure 4: Chart with $L(t)$ vs $\hat{L}$ for the finals

We now have an understanding of the flow of a match. At any given point, $\hat{L}(t) > 0.5$ implies that player 1 is doing well, while $\hat{L}(t) < 0.5$ implies that player 2 is doing well. The value $|\hat{L}(t) - 0.5|$ corresponds to the magnitude of the advantage that the leading player possesses. Additionally, with these new smoother charts, we can now generate our derivative $\hat{m}(t)$ for the final match:

With these charts of momentum, we now have a theoretical idea of a players performance and expected performance for the next series of points.

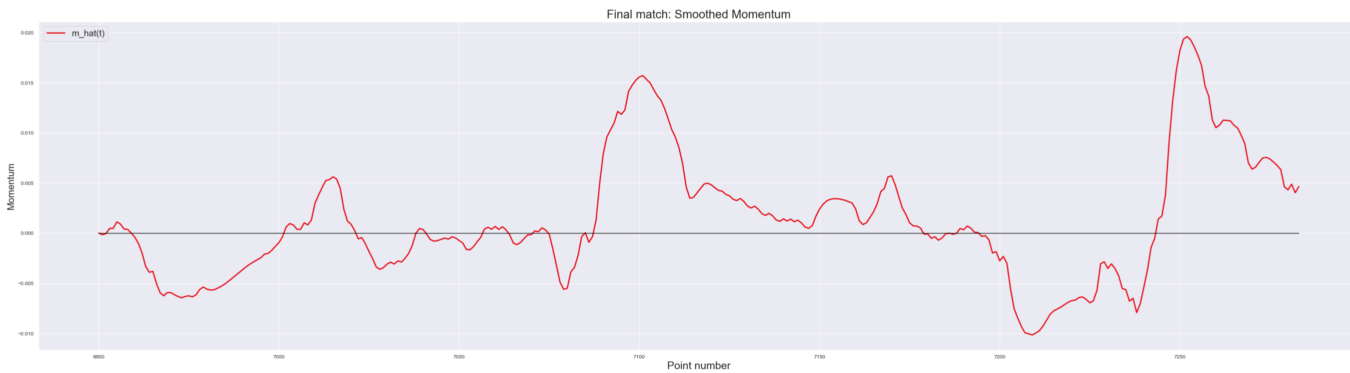


Figure 5: Chart with $\hat{m}(t)$ for the finals

5.3 Outperforming Expectations

We now address the skeptical coach: we may have charts for momentum but do they mean anything? Does a player having momentum have an impact on the match, or are strings of points and success simply random?

Our approach to these questions focuses on performance. If a player has positive momentum, then over the course of the next few points they should outperform expectations with respect to the number of points scored. Using the same methodology as in the previous section, we can compute the expected value over the next k points as

$$\mathbb{E}_i(t_0, k) = \sum_{j=0}^{k-1} |t_j[a] - i|(1 - p_s) - (|t_j[a] - i| - 1)p_s$$

where t_j is the j^{th} point after t_0, $i = 1$ or 2 depending on which player we are analyzing, and $t_j[a]$ is the number of the player serving the j^{th} point. For instance, if a player serves each of the next 3 points, then their expected number of points won is $3p_s$. We can then compute the difference between their actual performance and the expectation as

$$\Delta_i(t_0, k) = (\tau_i(t_{k-1}) - \tau_i(t_0)) - \mathbb{E}_i(t_0, k)$$

where $\tau_i(t)$ is the total number of points won by a player when reaching point t. We can now characterize these Δ_i's to evaluate player performance:

Behavior	**Interpretation**
$\Delta_1(t) > \delta_t, \hat{m}(t) > 0$	Player 1 outperforms expectations with momentum
$\Delta_1(t) > \delta_t, \hat{m}(t) < 0$	Player 1 outperforms expectations without momentum
$\Delta_2(t) > \delta_t, \hat{m}(t) > 0$	Player 2 outperforms expectations with momentum
$\Delta_2(t) > \delta_t, \hat{m}(t) < 0$	Player 2 outperforms expectations without momentum

Table 2: Player's Performance and Momentum

We can now plot these values for various matches in the tournament:

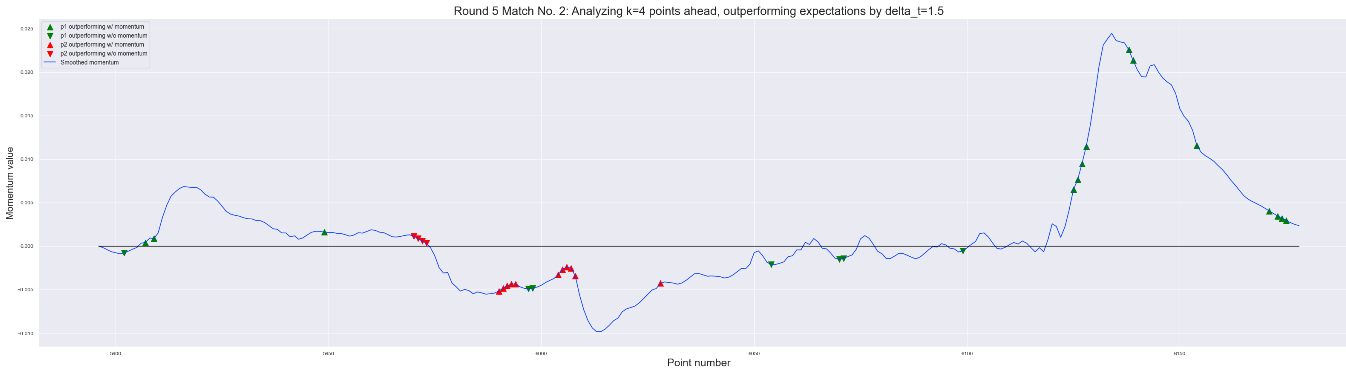


Figure 6: Players outperforming expectation in Round 5, Match 2

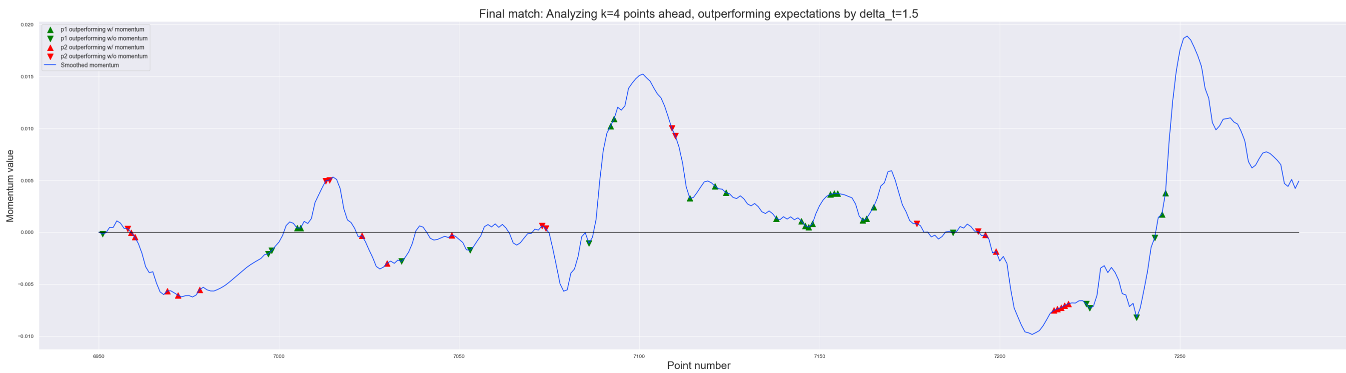


Figure 7: Players outperforming expectation in Finals

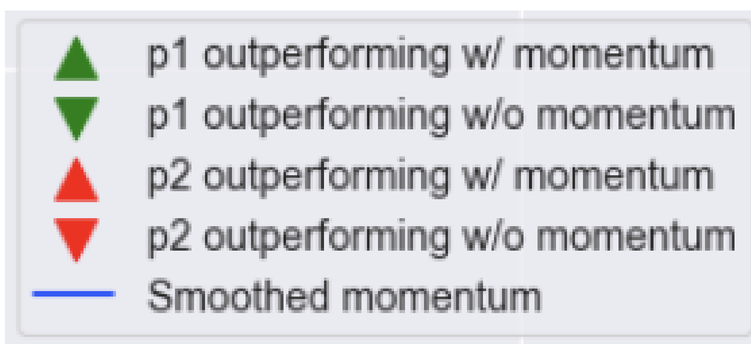


Figure 8: Enlarged legend.

In the match displayed in Figure 6, player 1 outperforms expectations twice as often with momentum compared to without, and the same can be said for player 2. In the final match displayed in Figure 7 we see that Player 1 outperformed expectations almost twice as often when he had momentum compared to when he didn't (20 times with vs. 11 times without) while Player 2 outperformed with momentum more modestly (15 times with vs 11 times without). If we define outperformance ratio as the total number of times a player outperforms with momentum divided by the times a player outperforms without momentum, we see

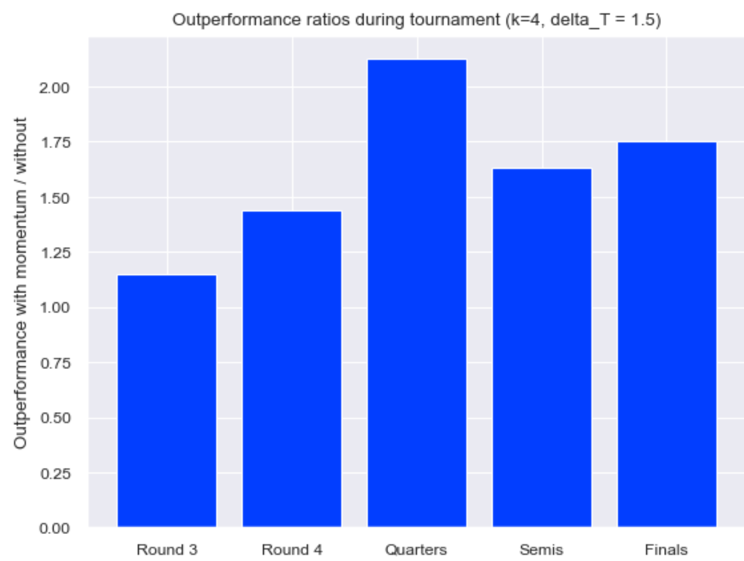


Figure 9: With $k = 4$, $\delta_T = 1.5$, outperformance ratio increases with round

This is no fluke, as we can observe in the figure below that players consistently outperform while having momentum regardless of the number of points into the future that we are looking: All values above

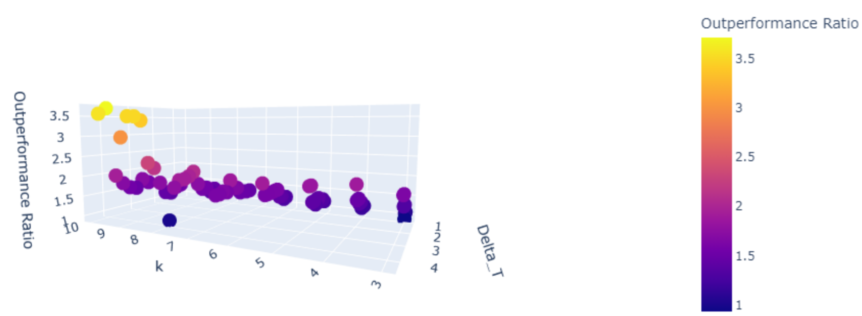


Figure 10: Values of outperformance ratio for $3 \leq k \leq 10$ and $\delta_T = jk$, with $j \in [\frac{1}{4}, \frac{1}{2}]$

exceed 1.0, meaning that regardless of the horizon, we see our momentum values tend to correlate strongly to increased performance.

5.4 Discussion

To answer the questions initially posed in this section, we believe that the data displayed in the above figures is sufficient evidence to suggest that momentum is an influential factor during this particular tournament. We further claim that outperformance ratio trends upwards later in the tournament, which has the possible explanation that players who are able to capitalize on their lead are generally able to win more often than others, leading to them progressing farther in the tournament.

As an interesting note on the figure directly above, we observe that the farther out we are measuring, the more likely players are to outperform when the have momentum. This confirms the intuitive hypothesis that there are only a few momentum swings in a given set, and that momentum generally carries forward for longer periods.

In the next sections we will discuss how the occurrence of swings in play are not fully random either, there may be specific moments in matches that lead to these shifts.

6 Causes of Swings

6.1 Model Architecture

We fit a Gradient Boosting Classifier to the point-by-point Wimbledon dataset to identify likely points of swing. The fundamental idea powering boosting methods is the concept that many predictive models are better than just one. Hundreds of weak predictive models (often shallow decision trees known as "stumps") are sequentially trained, each one correcting the errors made by the previous stumps. In each stage of the training process, the gradient of the loss function (in this case, log loss) is computed with respect to the parameters of the stumps. The direction of the gradient indicates how to update the model parameters to minimize the loss, hence the name *Gradient* boosting.

Gradient boosting is a powerful nonlinear ensemble method that often provides a significant improvement over vanilla decision trees and transparency regarding the relative importances of the various features used, which is why we found this to be the best model for this use case.

6.2 F1 Score

Note that we will evaluate our classifier using the F1 metric, rather than raw accuracy, due to the imbalance of datapoints. Namely, there are understandably far fewer momentum-swing points than non-momentum-swing points. Evaluating our model based on accuracy

The F1 score applied separately to predicting positive swing points and negative swing points is a much better metric for evaluating the performance of our model, as it is the harmonic mean of the precision and recall of our predictor:

$$F1 = \frac{TP}{TP + \frac{1}{2}(FP + FN)},$$

where TP is the number of true positives, FP is the number of false positives, and FN is the number of false negatives.

The *precision* of our momentum classifier can be thought of as the answer to the question "If the model claims a point is a point of momentum swing, what is the probability that this is actually the case?", whereas the *recall* of our classifier can be interpreted as the answer to "Of all the momentum swing points in the data, how many did our classifier correctly identify?". We wish to take the answers to both of these questions into account when evaluating our model, and, since the answer to these questions are best expressed as rates, the harmonic mean of these is a natural way of doing so.

6.3 Model Fitting and Performance

The model is fit to a dataset of the following format. We include columns of the dataset regarding contextual information about the point (e.g. lead, break point, etc.), the result of the serve (e.g. ace, speed, etc.), and other notable information about the flow of the game during a given point (e.g. distance run by players, rally count, etc.). We do not explicitly include data regarding the current score of the game, as that is all encapsulated in our lead metric. To classify points as swings, we find points at which $\hat{m}(t)$ crosses the x-axis and determine the direction of the movement.

We train the Gradient Boosting Classifier on all 5707 points that occur in rounds 3 and 4 of the tournament (24 total matches), and test the classifier on the 1577 points that occur in rounds 5, 6, and 7 of the tournament (7 total matches). We achieve an F1 score of **51%** when predicting positive momentum swings for the first player and an F1 score of **52%** when predicting negative momentum swings for the second player. These performance values, while relatively low, indicate that it is in fact possible to predict when a swing may occur, and furthermore, predicting whether a swing will be in a certain player's favor is in fact achievable.

6.4 Predicting Swings

We now provide an example of the performance of our model on one match in particular.

For the first example, we consider the match between Daniil Medvedev and Christopher Eubanks in Round 5, whose lead chart and momentum chart are pictured below. The thin multicolored lines surrounding the main blue line represent potential uncertainty in our chosen value of p_s, which will be addressed in section 8 of this paper.

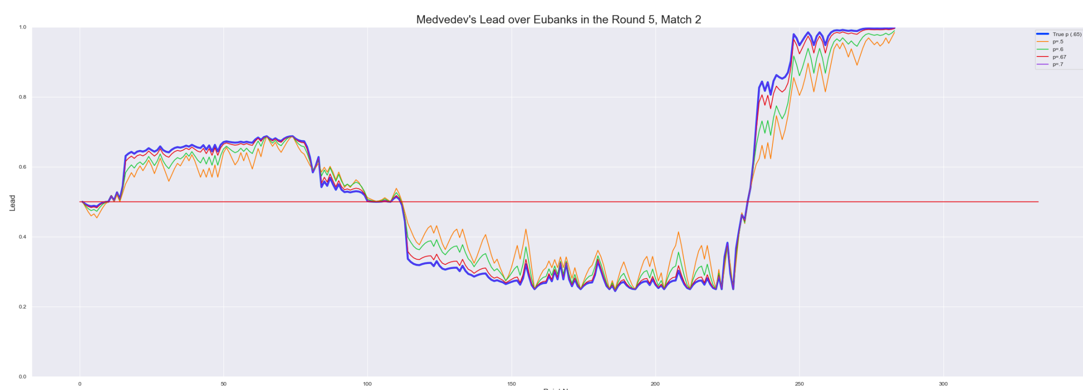


Figure 11: Medvedev's Lead over Eubanks in the 5th round.

In these graphs, we observe three significant lead changes, as well as a few momentum changes over the course of the match. Our model is able to achieve an F1 score of **71.4%** and **72.7%** on predicting positive and negative swings in momentum, respectively, for Medvedev.

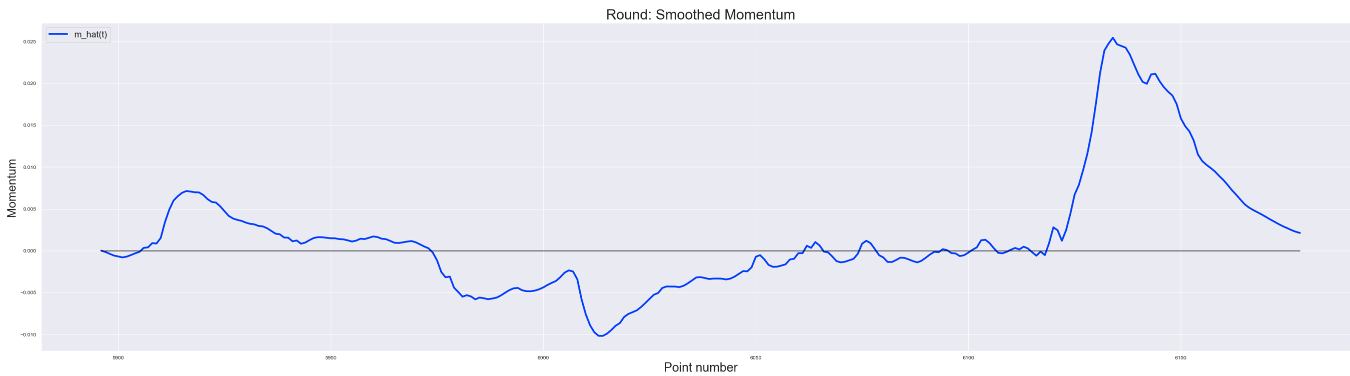


Figure 12: Chart with $\hat{m}(t)$ for the 2nd match of the 5th round

6.5 Advice to a Player

Below we show a table of the relative importance of various factors to our model in making its momentum predictions.

Metric	Weight
Momentum	0.09
p1_dist_run	0.02
p2_dist_run	0.02
speed_mph	0.02
p2_winner	0.009
p1_winner	0.006
serve_width	0.005
rally_count	0.002
p1_brk_pt_missed	0.001
return_depth	0.0006
serve_depth	0.0003
p1_brk_pt_won	0.0002
p2_brk_pt_won	0.0002
p2_brk_pt_missed	0.0002
p1_dbl_fault	0.00005
...	

Table 3: Sorted Metrics by Weight (values are relative)

As is evident, outside of momentum, the distance run by both players and the speed of the serve are the most influential factors in predicting whether a shift in momentum will occur. Therefore, based on this evidence, we would advise that a player heading into a match against a new opponent be mindful of these factors. In particular, if he seems to be losing momentum, serving harder (i.e. faster speed) and forcing their opponent to run further distances, while trying to set themselves up in positions that require less movement, are all strategies the player should employ to shift the momentum back in his favor. These strategies all also work to maintain or increase momentum that is already in his favor.

7 Generalizability

7.1 Beyond Grass

Note the Wimbledon is only played on grass. This plays a role on momentum as grass courts are known for their fast ball speed and low bounce which gives an advantage to the play serving. Clay courts, are slower and produce a higher bounce. Hard courts have medium speed and produce the highest bounce out of all three surfaces.

We will now see what would happen if we used our model on tennis matches with clay courts and hard courts. It's important to note that our model puts heavy emphasis on the distance the players run and the speed of the serve. These are the two factors that are heavily impacted by court surface (according to table 3).

Firstly, we will consider how the model runs on a clay court. We can expect the distance the players run to increase as the rally size will increase and speed of the serve will decrease. With players running more, there is more likely to be a momentum swing. When the speed of the serve decreases, it's easier to return serves and the rally count will increase. Furthermore, many pros have more trouble holding serve on clay courts due to the factors noted above, which means that it is harder to hold onto a lead, as being up $5-4$ preparing to serve for the set is less of a guaranteed win compared to other surfaces. Thus we anticipate more momentum swings, generally occurring after long rallies.

Now, consider a game that's played on a hard court. Serve speed is high, but rallies also tend to be longer. It would be hard to estimate the effects that this would have on the number of momentum swings without data, but we would guess that hard courts are similar to grass.

Fortunately, if provided with a dataset for these different surfaces with similar features, we are confident that our model would perform well in predicting the swings.

7.2 Women's Tennis

Women's tennis consists of matches that are 3 sets while men's consists of matches of 5. Having lesser sets will in turn increase the chance of momentum swings. Although our model doesn't directly take into account the number of sets that are played, the total rallies won by each play will be significantly less. This means that every time a rally is won, the swing in momentum is much greater. For an extreme example, consider a tennis match that is played to 1 set. Each rally won is weighted significantly meaning momentum swings would be much greater. However in a 7 set tennis match, the benefit of winning an individual rally is weighted much less. The overall momentum graph would be much more smooth with many less swings.

7.3 Other Sports

Comparing how our model will run in different sports, consider a volleyball match. One of the main differences to note is that in volleyball, the winner keeps serving unlike tennis. This would lead to momentum being steady with major momentum swings when there is a turnover and the opposing team gets to serve. Furthermore, a team can start to dictate to flow of the match if they maintain possession for a prolonged period of time. In this example, the momentum curve would be much more smooth with very little swings. Overall, the model can implemented in other sports.

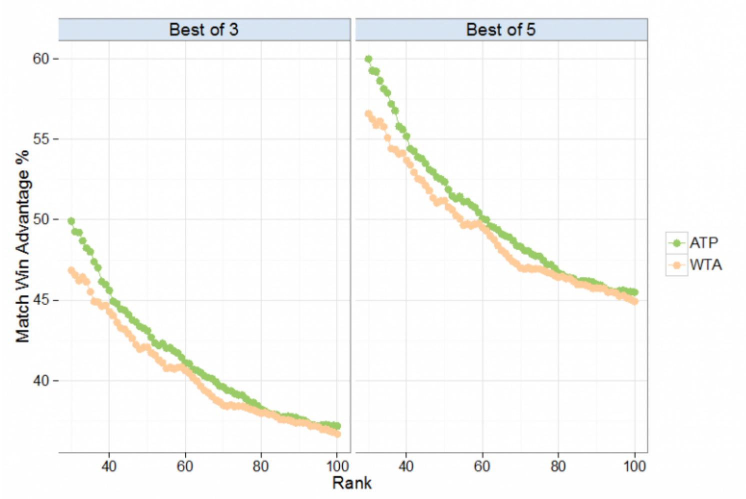


Figure 13: Women vs Mens Tennis Matches

8 Sensitivity Analysis

We now return to the figures presented in section 5.1 for an analysis of the sensitivity of the lead model for different values of p_s. Note that for values of p closer to 0.5, we obtain lead charts with a greater variability from point to point. Furthermore, we see far more erratic patterns for the momentum graph late in the match the lower the value of p_s (until it reaches 0.5).

Note that for small deviations in p_s (such as, for example, the deviation between the 0.67 value computed by averaging over the Wimbledon point-by-point dataset and 0.65), the graph of the lead is relatively unchanged. Thus, our model is not overly sensitive to choices of p_s that deviate a little from our accepted value of 0.67, and so we confidently assert that our lead model passes the sensitivity test with respect to p_s.

We now turn to evaluate the sensitivity of our Gradient Boosting classifier to perturbations in the input variables. To accomplish this, we systematically remove one predictor variable at a time from the training set, retrain the model on the depleted dataset, and then evaluate the F1 score of this new model. The results of this process are depicted in the radar charts below. Each label represents the column that was removed, and the consequent F1 score of the model. Recall that the F1 score of the fully trained model (i.e. without any predictor variables removed) was 0.51 for predicting positive swings in momentum and 0.52 for predicting negative swings in momentum for a given player. This is reflected in that the maximum values on each of the radar charts shown represents the F1 score of the model when considering every independent variable (i.e. no predictor variable left out).

In both instances, the model achieves F1 scores very close to the optimal F1 score for any removal of a single predictor column. The only notable exceptions are those for the current momentum and distance run variables. The heavy reliance of the model on the current momentum is justifiable, as a swing in momentum can only really occur if there is somewhere to swing from. Furthermore, the model's performance suffers noticeably when a distance run metric is excluded, especially when predicting negative swings in momentum. This also makes sense, as a player who expends a great deal of physical effort during a point, only to lose the point, can suffer from a temporary decline in morale, affecting his performance on the subsequent points.

Therefore, while our model may not pass the sensitivity test with respect to the current momentum

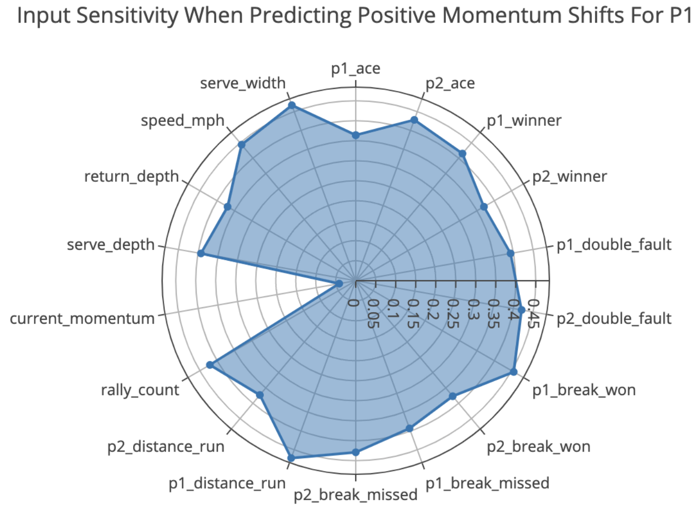


Figure 14: Sensitivity of model to input variables when predicting positive shifts in momentum.

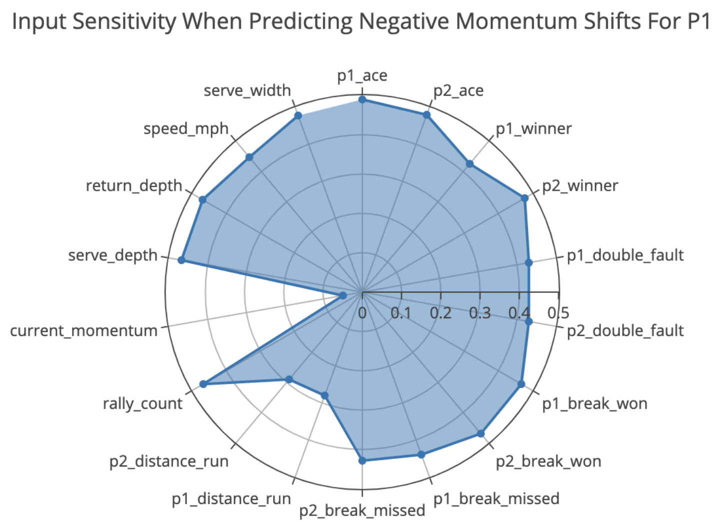


Figure 15: Sensitivity of model to input variables when predicting negative shifts in momentum.

input variable, it does pass with respect to the other predictors. We conclude that the current momentum and distance run by both players are essential inclusions to the model, should our model be adopted by tennis coaches and players to assess their performance.

9 Strengths and Weaknesses

9.1 Strengths

A major strength of our approach is the extent to which our model can be applied to contexts other than tennis at a Grand Slam. In any sport where a function $L(t)$ can be computed, it is possible to use the same methods. For instance, in basketball, a game with discrete scoring, it may be possible that dunks are highly correlated to momentum swings due to the psychological impact. Or in football, it may be the case that goals scored near the end of a half have a huge impact. While these sports are more complex due to substitutions, our approach still holds some effect in team sports due to morale being a social phenomenon.

Another strength of our model is that it is not exceedingly sensitive to inputs despite us taking almost all of the given data into account. This means that our results are generally resistant to small changes or errors in the dataset as described above.

9.2 Weaknesses

As mentioned in the previous section, the model will likely suffer significant performance issues if data regarding the distance each player runs during a particular point is not available. This made finding other similar datasets a challenging task which we were not able to complete.

As noted in our assumptions, the idea of a constant p_s is not representative of professional play. Some players, such as John Isner, are serve specialists who have a much higher likelihood of aces and higher p_s as a whole. However, we believe that for the purposes of understanding momentum, treating this value as constant is not terribly detrimental. Given the confounding variables of nerves, service confidence, and tiredness, we believe that in order for a model to factor in p_s and be robust, it must also track match-to-match momentum and head-to-head matchups between players.

Another variable that was not considered in section 5 is the converse of our analysis: namely how often does a player underperform when they have momentum vs without it. Due to resource constraints, we were not able to investigate this specifically, but per our other findings we believe that our results would still hold, namely that players would underform more often when momentum is against them. An interesting point in the same vein that we were not able to evaluate is how to measure game-to-game momentum and set-to-set momentum. Since we were working with a dataset of only 7300 points, roughly 1000 games, and just over 100 sets, we were not able to generate any strong conclusions about these larger scale momentum patterns. However we believe that our apporach would generalize, namely taking the average momentum value across games and finding trends. With a larger dataset, we believe that our approach would yield strong findings given our conclusions in 5.4 (i.e. that momentum tends to be a longer-term trend rather than something that flips every few points).

10 Letter to Coaches

Dear Tennis Coaches,

We have spent the last few days analyzing every point the last five rounds of the 2023 Wimbledon tournament and wish to present our results and conclusions to you. Per our analysis, we believe that momentum is not just an idea, it is a quantifiable phenomenon that occurs even in matches between players of the highest caliber. We found that players at this level experience roughly three momentum swings per set in their matches, with more swings occurring as matches progress. We also found that the most successful players had the best ability to capitalize on moments when momentum went in their favor, some exceeding expectations by 50% over the course of the next 10 points! With this in mind, it is critical to consider the role that psychology plays, and be ready to react when things are not looking up.

The critical factors that influence momentum swings that we found were distance run by both players and the speed of the serve, with these metrics and other being able to predict momentum swings with an accuracy over 70% for some matches. We also found that momentum tended to accrue over time, so multiple consistent successful points tend to cause success. In a game as volatile as tennis, it is critical not to allow one to get in ones own head about unforced errors or unlucky moments.

In order to prepare your players to better respond to these scenarios, specific drilling would be our recommendation. Have them play against an opponent of similar strength starting down 0 – 4 in a set. When playing these drills, encourage them to focus on making their opponent move and keeping in mind their cues for their first and second serves. It is critical that their serve speed and consistency does not decrease as they get tired or nervous, as these are the keys to controlling the start of the point and ultimately making their opponent make mistakes.

Also, with our findings indicating the importance of winning long, hard-fought points, we advise doing drills that force your players into keeping the ball in play without conceding position too much. Tell them to focus on staying in the middle of the court at the baseline and making their opponent run.

If a player feels that momentum is swinging against them (either intuitively or using the cues noted above), it is important to stay calm. Particularly in a best-of-5 match like at Wimbledon, there are many opportunities for a comeback as there are plenty of instances where players are unable to hold on to a 2 set lead. If they feel that they are losing long rallies consistently or just being forced to run around the court too much, then it is time to mix up their strategy. Try serve-and-volley, try focusing on lobs, just try something different than what is currently failing.

Ultimately, tennis matches at a high-level come down to holding serve and break points. If your player is confident in their serve, then they can afford to take risks on their returning games. Go for hard winners, apply pressure to the opponents second serve, whatever it takes to cause momentum to shift back in their direction. With our findings, while we do see that the chance of winning massively increases after the first set, but we see that momentum swings continue to occur well after this point. As noted before, momentum is almost guaranteed to fluctuate back and forth, it is the player that is able to capitalize on the critical moments that tends to take home the prize.

We hope that these finds serve you and your players well and we wish you the best of luck in any tournaments!

Best,
Team 2429101

References

[1] American Statistical Association. "Analysis of men's and women's tennis performance finds differences from match format." *ScienceDaily*, 9 August 2015. [Online]. Available: `https://www.sciencedaily.com/releases/2015/08/150809170256.htm`.

[2] Yurko, Ronald, Ventura, Samuel, and Horowitz, Maksim. "nflWAR: A Reproducible Method for Offensive Player Evaluation in Football." *Journal of Quantitative Analysis in Sports*, vol. 15, no. 3, 2019, pp. 163-183. *De Gruyter*. [Online]. Available: `https://doi.org/10.1515/jqas-2018-0010`.

[3] MLB. "Wins Above Replacement (WAR): Glossary." *MLB.com*. [Online]. Available: `https://www.mlb.com/glossary/advanced-stats/wins-above-replacement`. Accessed 5 Feb. 2024.

[4] Frantisek Vaverka, Jiri Nykodym, Jan Hendl, Jiri Zhanel David Zahradnik (2018) "Association between serve speed and court surface in tennis." *International Journal of Performance Analysis in Sport*, vol. 18, no. 2, 2018, pp. 262-272. DOI: `doi.org/10.1080/24748668.2018.1467995`.