Predicting Model——Ideas

In our initial inquiries, we developed a momentum score framework predicated on specific metrics observed during gameplay, enabling the computation of each competitor's momentum at the conclusion of every round within a given match. The objective of discerning the match's pivotal juncture – the round where the dynamics notably shift – necessitates an understanding of the on-court realities and the strategies employed by both the offensive and defensive players that contribute to this shift. This logically extends from our primary question, directing our attention to the intersection of the two momentum trajectories. Yet, this approach surfaces two pertinent challenges:

1. The foundational model, as established in our primary query, is anchored in assumptions that are straightforward and intuitively resonant. For the third query, where the focus is on pinpointing match turning points, our predictive model must integrate a broader array of features pertinent to serving and returning dynamics (such as 'p1\_ace', 'p2\_ace', 'p1\_double\_fault', 'p2\_double\_fault', 'serve\_width\_B', and others). This necessitates employing machine learning techniques to assimilate these feature weights effectively into the momentum curve modeling.
2. Our initial model presupposes that momentum scores are contingent on recent round performances, interpreting 'momentum' as an expression of a player's arousal state. Given the fluctuating fortunes in elite tennis matches, momentum shifts are expected to be frequent. This results in an overabundance of intersection points on the momentum curves. To address this, it becomes imperative to develop methodologies that filter out insignificant, repetitive stalemate points, thereby concentrating on rounds that signify a transition into or out of a deadlock.

Predicting Model——Approaches

For the first issue, we employed a random forest regressor to train our model using data processed through the momentum score model established in the first question. The momentum score serves as a crucial label in the supervised learning process. We used the R² score as the evaluation metric and observed the importance of features post-training. To prevent model overfitting and enhance the generalizability of our predictive model, it's essential to select primary features for further training.

Regarding the second issue, we first identify all intersection points where the change in momentum is not less than a certain threshold. Then, we select those points that do not have any other intersection points in their vicinity, meaning there are no further momentum crosses in several rounds before or after these points. We locate these turning points and, in conjunction with the specific on-court situations, analyze the rationality of our predictions.

Predicting Model——Results

The following chart shows the ranking of feature importance after the initial training.图表

描述已自动生成

Select the top 10 features with the highest importance for retraining. Then, obtain the actual and fitted momentum curves for Player1 and Player2.图表

描述已自动生成

In the process of pinpointing crucial junctures and evaluating the dynamics of the game, let's consider the match labeled as 1302. The players involved in this encounter are Alexander Zverev, designated as Player 1 (P1), and Matteo Berrettini, referred to as Player 2 (P2).

图表

描述已自动生成

1. Intersection 1: (x=26, y=0.53) - Change in Momentum 1: 0.65, Change in Momentum 2: -0.71

大比分——0：0

小比分——1：2

当前——40: 15（1发球）

拍数：1

跑动距离：5.6m-6.9m

发球速度：133mph

结果：1一发保发成功，小比分2：2

模型预测：此球后，2优1劣->胶着

1. Intersection 2: (x=44, y=0.64) - Change in Momentum 1: -0.47, Change in Momentum 2: 0.64

大比分——0：0

小比分——3：4

当前——15：30（1发球）

拍数：2

跑动距离：7.2m-8.8m

发球速度：112mph

结果：2二发得分，1非受迫性失误，15：40

模型预测：此球后，2优势进一步扩大

1. Intersection 3: (x=64, y=0.72) - Change in Momentum 1: 0.36, Change in Momentum 2: -0.44

大比分——0：1

小比分——1：1

当前——40：0（1发球）

拍数：1

跑动距离：0.6m-0.9m

发球速度：126mph

结果：1一发ace保发成功，小比分2：1

模型预测：此球后，1一转颓势，重新将比赛打入胶着状态

1. Intersection 4: (x=112, y=0.58) - Change in Momentum 1: -0.81, Change in Momentum 2: 0.82

大比分——0：1

小比分——6：6

当前——15：40（1发球）

拍数：1

跑动距离：0.8m-0.9m

发球速度：134mph

结果：1一发ace得分，30：40

模型预测：此球后，2再次优势

1. Intersection 5: (x=135, y=0.73) - Change in Momentum 1: 2.16, Change in Momentum 2: -3.30

大比分——0：2

小比分——1：1

当前——30：40（2发球）

拍数：5

跑动距离：26.5m-16.1m

发球速度：127mph

结果：2一发保发成功，小比分1：2

模型预测：此球后，2重新将比赛拉入胶着

Although there is ambiguity in the increase or decrease of momentum for individual points, considering the trend of the momentum curve and the actual events on the court, we believe our predictive model has successfully identified the key turning points that significantly affect the momentum of the match. Once these turning points are identified, we can extract them separately and analyze actual game strategies such as the types of serves and returns, the number of strokes, running distance, etc.

When Player1 plays server:

Coefficient of features:

Feature Coefficient

0 p1\_sets -0.660716

1 p2\_sets 1.079027

2 p1\_games 0.523895

3 p2\_games -0.461889

4 p1\_double\_fault 0.360961

5 p1\_unf\_err -0.310627

6 p2\_unf\_err -0.924556

7 rally\_count 0.404886

8 speed\_mph -0.641765

9 serve\_width\_B -0.342782

10 serve\_width\_BC -0.237314

11 serve\_width\_BW -0.029538

12 serve\_width\_C 0.361789

13 serve\_width\_W 0.111184

14 serve\_depth\_CTL -0.393256

15 serve\_depth\_NCTL 0.391367

16 distance\_diff -0.314234

17 p2\_distance\_run -0.570710

18 p1\_distance\_run -0.515582

Accuracy of model: 0.75

When Player1 plays servee:

Coefficient of features:

Feature Coefficient

0 p1\_sets 0.124868

1 p2\_sets -0.149697

2 p1\_games 0.519961

3 p2\_games -0.155793

4 p2\_double\_fault 0.643224

5 p1\_unf\_err -0.136075

6 p2\_unf\_err -0.087853

7 rally\_count -0.305280

8 speed\_mph 0.205257

9 return\_depth\_D 0.259512

10 return\_depth\_ND 0.372136

11 distance\_diff -0.093827

12 p2\_distance\_run 0.399639

13 p1\_distance\_run 0.333224

Accuracy of model: 0.782608695652174

Predicting Model——Advices

**Recommendations for the Serving Player:**

1. Enhance First Serve Success: Increasing the success rate of first serves generally correlates with a stronger offensive position and reduced pressure.
2. Diversify Serve Tactics: Utilizing a variety of serve trajectories (B, BC, BW, C, W) and depths (CTL, NCTL), servers are advised to dynamically alter their serving strategies to bewilder their opponents.
3. Dictate Match Tempo: The rally\_count's positive impact implies that prolonging rallies may contribute to building momentum, likely linked to patient play, capitalizing on opponent errors, or generating scoring opportunities.
4. Limit Unforced Errors: The significance of minimizing unforced errors, as indicated by the negative impact of p1\_unf\_err, highlights the need for consistency and error reduction.

**Guidance for the Receiving Player:**

1. Emphasize Defense and Counterplay: With their data reflecting the opponent's serve, receivers should prioritize defensive and responsive strategies.
2. Strengthen Second Serve Returns: The coefficient related to p2\_double\_fault suggests an opportunity to increase aggression on second serve returns, targeting the opponent's weaknesses.
3. Modify Return Techniques: Based on return\_depth\_D and return\_depth\_ND, receivers should adapt their positioning and return tactics in response to ongoing match conditions.
4. Optimize Movement Efficiency: By effectively managing movement on the court, receivers can reduce unnecessary exertion, preserving energy and focus.