假设：考虑表现得分时不考虑对手表现的影响

数据处理

异常值处理

数据转换

特征提取

计算各场比赛的优胜比例

连续得分

连续失分

问题一

（1）筛选建模指标

|  |  |  |
| --- | --- | --- |
|  |  |  |
| X1 | Point\_victor | 是否得分 0.，1 |
| X2发球 |  | ACE 0，1 |
| X3 发球 |  | 双误 -1，0 |
| X4 |  | 非受迫性失误-1，0 |
| X5 |  | 上网得分 0，1 |
| X6 发球 |  | 发球宽度 |
| X7 发球 |  | 发球深度 |
| X8 发球 |  | 速度 |
| X9 |  | 距离 |
| X10接球 |  | Winner 0，1 |
| X11接球 |  | 返回球深度 |
| X12 接球 |  | 是否破发点 -1，0，1 |

(2)构建表现得分模型

发球表现得分模型：

接球表现得分模型：

采用AHP-CRITIC综合方法计算权重

AHP步骤

发球判断矩阵（由于考虑到发球优势较明显，所以在设置发球判断矩阵时将X1相对于其他指标的重要性适当减弱了）：

接球判断矩阵：

判断矩阵设计时考虑到了发球优势：由于发球方胜率为0.65，

CRITIC步骤：

AHP-CRITIC组合权重

（3）可视化

Pl和p2的表现得分可视化

从图中可以看出，xxx的表现得分总体高于xxx，符合比赛结果，说明模型准确性较高

（4）与历史数据结果与趋势对比，说明模型的准确性（可小波分析表现得分趋势 后面有时间补） 并分析在p1发球的5球内，p1的发球得分是否高于p2，说明发球优势

发球和接球

（5）预测模型

问题二：

模拟随机数据，检验说明非随机，接着进一步说明其与动量的关联 （待补充）

（1）计算动量

解释为何以局为单位来构造动量 合理性

1. \*\*局的重要性\*\*：在网球比赛中，每个局都是赢得比赛的基本组成部分。因此，球员在每个局中的表现对于整场比赛的结果至关重要。

2. \*\*局内的动量\*\*：在每个局内，球员的表现会受到前几个球的影响，比如发球、接发球、得分和失误等。这些事件的连续发生会构成局内的动量变化。

3. \*\*局间的转折点\*\*：当一个球员赢得一个局时，他们可能会在心理上获得优势，并在下个局中保持动量。相反，输掉一个局可能会影响球员的信心和表现。因此，局与局之间的胜负关系可以作为动量变化的重要指标。

4. \*\*动量的传递\*\*：在网球比赛中，局内的胜利通常会传递到下个局中，从而影响到整个比赛的走势。因此，局内的动量变化可以帮助我们理解比赛中的关键时刻和转折点。

通过以局为单位构造动量，我们可以更好地理解球员在比赛中的表现，并分析他们在不同局中的动态变化。这有助于教练和分析员更好地制定战术和策略，以应对比赛中的不同情况。

定义动量，说明为何选取该指标

动量 连续得分win\_points 连续失分 lose\_points 连续赢的局数win\_games 连续输的局数lose\_games 连续赢的盘数win\_sets,连续输的盘数lose\_sets, 发球优势serve

CRITIC法赋权

连续赢的局数:在本盘中到目前为止连续赢的最大局数

连续赢的盘数：在本次match中到目前为止连续赢的最大盘数

连续得分：在本局中最高的连续得分次数

题目中提到动量被定义为“力量或通过运动或一系列事件获得的力量”，并指出在比赛中球队或球员可能会感觉到他们拥有动量，但很难对这种现象进行量化和衡量。因此，在构建动量模型时，我们应该尽可能地接近这个定义，并尝试从比赛中可测量的因素中推导出动量的概念。

考虑到这一定义，我们的动量模型应该能够反映出球员在比赛中的“力量”或“势头”的变化，以及这种变化是如何由一系列事件或表现所驱动的。因此，我们选择的因素和指标应该能够捕捉到球员在比赛中的表现变化，并将其转化为动量的量化度量。

1. \*\*表现得分的斜率\*\*：考虑到斜率可以反映球员表现的变化速度，因此将其纳入动量的计算中是合理的。较大的正斜率可能表示球员表现的快速改善，而较大的负斜率可能表示球员表现的迅速恶化。

2. \*\*连续赢的局和盘\*\*：连续赢得局和盘是一个重要的动量指标，因为它们可以反映球员在比赛中建立的优势和动力。连续赢得局和盘的球员往往会在心理上更有信心，同时也会对对手产生压力。

3. \*\*发球优势\*\*：在网球比赛中，发球优势可以带来重要的竞技优势，因为发球方有更大的控制权和更高的获胜概率。因此，将发球优势纳入动量的考虑是合理的。

4. \*\*连续得失分\*\*：连续得失分可以反映比赛中的连续性，这对于捕捉到比赛中的动态变化和转折点是很重要的。连续得分可能表示球员的持续优势，而连续失分可能表示球员的快速下滑。

计算动量

CRITIC法赋权：CRITIC法是一种常用的赋权方法，它考虑了多个指标之间的相对重要性，并根据其对比值来赋予权重。这种方法能够有效地平衡各个指标之间的影响，从而更准确地反映出动量的变化。

划分种类 （为了根据不同的类别提供不同的建议）

（2）计算局势波动 =我当前的表现得分-对方的表现得分 和连续成功次数

为正表示走势偏向我，为负表示比赛偏向对方

将局势波动定义为你的表现得分与对手的表现得分之差是一个合理的方法，因为这可以捕捉到比赛中你与对手之间的竞争情况。当你的表现得分高于对手时，局势波动值会为正，反之为负，这可以反映出比赛中的优劣势和局势的变化。

为了说明这种定义的合理性，你可以考虑以下几点：

1. \*\*竞争性体现\*\*：局势波动的定义直接反映了比赛中你与对手之间的竞争情况。正值表示你处于优势地位，负值表示对手处于优势地位。

2. \*\*动态性\*\*：局势波动随着比赛的进行而动态变化，能够捕捉到比赛中局势的变化趋势。这有助于分析比赛中的关键时刻和转折点。

3. \*\*量化指标\*\*：局势波动的定义提供了一个量化的指标来衡量比赛中的局势变化。你可以通过比较局势波动的大小和方向来评估比赛中的优劣势。

4. \*\*可解释性\*\*：局势波动的定义直观且易于理解，能够直接反映出你与对手之间的得分差异。这样的定义有助于深入分析比赛中的各种情况和策略。

综上所述，将局势波动定义为你的表现得分与对手的表现得分之差是合理的，因为它能够全面地反映比赛中的竞争情况和局势变化，为比赛分析提供了有力支持。

用小波变换提取局势波动和连续成功的特征，分析趋势

求导数，如果导数趋于一致，说明二者有很大的相关性

（3）画局势波动、连续成功与动量的折线图，判断是否随机 应为不随机 可视化分析

例如：动量高的时候，局势对我也是有利的，连续成功次数也是高的（仅供参考！！！请自己撰写）

（4）相关性分析 进行相关系数假设检验 进一步说明有关联 待补充

问题三

（1）联系问题一

接下来我们具体分析动量中的每一个指标对局势波动的影响。

在模型一的基础上，我们可以计算出两个球员的表现得分，并计算相应走势

接下来，我们可以用问题一中的指标与走势一起构建随机森林模型，对局势波动进行预测，并根据变量的重要性选出与走势最相关的因素

（2）问题二中已经证明动量与局势波动有关联，我们可以根据已有的动量，预测未来的动量，在模型中，考虑包括连续得分、发球优势等因素，以提高预测的准确性。

对球员1来说，

双方动量差异==

对已有的5个指标值进行聚类分析，聚为不同类 比如4类，用1，2，3，4代替

划分种类 （为了根据不同的类别提供不同的建议）

并得到聚类中心 然后根据聚类中心的值计算相应的动量 k-means聚类

然后用马尔可夫链（or 其他预测模型） 预测未来动量状态 未来动量离哪个近就是哪一类

计算动量状态的差值

不同种类采取不同的战术，提供不同建议

【建议参考】

当根据预测模型的结果发现动量将有较大波动时，我们可以制定如下建议策略：

1. \*\*调整战术\*\*：根据预测结果，分析动量的变化趋势，有针对性地调整战术。例如，在动量有利的时候，可以选择更进攻性的战术，积极主动地发起进攻；而在动量处于劣势时，则可以采取更保守的战术，稳健地守住局面。

2. \*\*加强心理素质\*\*：面对动量的波动，球员需要保持良好的心理素质，不受外界因素的影响。可以通过心理训练、专业心理咨询等方式，提高球员的心理抗压能力，保持冷静、沉着的心态。

3. \*\*调整发球策略\*\*：发球是影响比赛走势的重要因素之一，根据预测结果，可以调整发球策略。例如，在动量有利的情况下，可以选择更具攻击性的发球策略，以增加得分的机会；而在动量劣势时，则可以选择更为保守的发球策略，减少失误的可能性。

4. \*\*及时反应变化\*\*：根据模型预测的动量变化趋势，球员和教练团队需要及时反应，灵活调整战术和策略。保持对比赛的敏锐观察和分析，根据实际情况及时做出调整，以最大程度地利用动量的优势。

当预测结果显示动量波动较小时，我们可以建议球员继续保持稳定的表现，并保持当前的战术和策略不变。此时的重点在于稳定表现，保持比赛的掌控力，避免因过度调整而引起的不必要风险。同时，球员也应保持对局势的关注，随时准备应对可能的变化。

问题四

1. 在其他场比赛中测试开发模型

用问题三中已训练好的随机森林模型预测未来的局势波动并于实际数据比赛结果进行对比，说明预测效果

考虑准确率，召回率

1. 未来因素 场外

敏感性分析

具体分析每一个指标变化对局势波动的影响，提出建议

随机森林 调参查看MAPE（均方误差）或其他指标是否有错