|  |  |  |
| --- | --- | --- |
| **Problem Chosen**  C | **2024**  **MCM/ICM**  **Summary Sheet** | **Team Control Number**  1111111 |

**Your Paper's Title**

**Summary**

|  |
| --- |
| **说明：**  **本Word模版的版本编号是0.2版，是以美赛的标准创建的模版，也适用于其他用英文写作的数学建模比赛，例如亚太赛和小美赛。模版的使用方法可以查看下面这个视频：**  **数学建模清风——论文排版教程**  **https://www.bilibili.com/video/BV1Ci4y1c7Ld**  **未来当发现模版中的问题或者比赛要求有更新时，我会发布更新后的新版本。大家可以在公众号《数学建模学习交流》后台发送“美赛论文模版”获取最新的模版的信息。只要本文档的版本号和公众号后台最新的版本号一致，则说明你下载的是最新版本。**   * **不同的美赛特等奖论文写作风格差别较大，这里给的模版是提取出来的通用框架，大家可以根据自己的需要进行修改。** * **红色字体标注的内容是对这个部分的解释，看完后请删除。** * **黄色高亮的是需要大家修改的内容，大家修改后请取消高亮显示。**     这里写美赛的摘要，美赛中摘要格外重要！！官方对于摘要的说明：  *The summary is an essential part of your MCM/ICM paper and should appear as the first page of your solution report. The judges place considerable weight on the summary, and winning papers are often distinguished from other papers based on the quality of the summary.*  翻译：摘要是美赛论文的重要组成部分，应该放在论文的第一页展示。评委们对摘要相当重视，获奖论文和其他论文的区别往往就在于摘要的好坏！  论文摘要的写法可以参考这个视频：  <https://www.bilibili.com/video/BV1Na411w7c2/>  标题下面的Summary可要可不要，如果你觉得你摘要写的不够多，页面下方留白太大的话就加上，这样看起来稍微好看点。  注意：美赛要求英文写作，很多同学英文写作水平不够，可以先写成中文的论文，然后再进行翻译！翻译可以使用有道翻译或者谷歌翻译，翻译后一定要人工修改，机器翻译的很生硬，对专有名词的翻译也不是很准确。 |

**Keywords:** keyword1; keyword2; keyword3; keyword4

Contents最后记得更新整个目录

[1 Introduction 3](#_Toc720)

[1.1 Problem Background And Restatement 3](#_Toc25723)

[1.2 Our Work 4](#_Toc10437)

[2 Assumptions and Justifications 4](#_Toc21114)

[3 Notations 4](#_Toc6739)

[4 问题一 5](#_Toc14796)

[5.问题二 10](#_Toc25728)

[6.问题三 17](#_Toc15994)

[7.问题四 23](#_Toc510)

[8.备忘录 23](#_Toc25607)

[9.Sensitivity Analysis 25](#_Toc5468)

[10.Model Evaluation and Further Discussion 25](#_Toc9088)

[10.1Strengths 25](#_Toc22876)

[10.2Weaknesses 25](#_Toc729)

[10.3Further Discussion 25](#_Toc8543)

[11.Conclusion 26](#_Toc16459)

[References 27](#_Toc20931)

[Appendices 28](#_Toc31122)

# Introduction

## Problem Background And Restatement

In the 2023 Wimbledon men's singles final, 20-year-old Spanish rookie Carlos Alcarras defeated 36-year-old Novak Djokovic. The match was a compelling and highly reversed battle. Djokovic won the first set 6-1, but lost the second set 6-7 and the next set 1-6. Alcalas was in complete control at the start of the fourth game, but at one point the match changed direction and Djokovic won 6-3. Djokovic had the advantage at the start of the fifth, but at one point the tide turned and Alcalas won 6-4. The match data is contained in the "match\_id" of the dataset "2023-Wimbledon-1701". We usually think of the "momentum" factor as occurring when the seemingly dominant player's subsequent score and performance become abnormal.

In tennis, the most important thing is the mental factors and the physical effects of momentum to determine the direction of the match. Momentum, defined in the dictionary as "force gained by movement or series of events" that sometimes changes the rules of a game and sometimes defines a winner, can be quantified by calculating a player's scoring advantage over an opponent in a given window of time. A "momentum swing" is similar to a "turning point." By understanding the role that momentum can play in determining the direction of the situation and motivating the player, there are ways and strategies that can be adopted to ensure that the player is in control of momentum during the game.

Data is provided for each point after the first two rounds of the men's tournament at Wimbledon 2023, and you can add additional information or data at your discretion, as needed to record data sources. Combining the information and data, please solve the following problems:

1. Develop a model that captures the flow of a game when a score occurs, and apply it to one or more games. Be able to determine which player is doing better at a certain point in the race and how well they are doing. Provide visualizations based on your model to show the competition flow. (In tennis, servers are much more likely to win, and this is somehow factored into the model.)

2. A tennis coach was skeptical that "momentum" played any role in a match, assuming that situations in a match fluctuated and that a player's successive successes were random. Use your model/metrics to evaluate this claim.

3. Are there any indicators that can help determine when the trend of a game changes from one side to the other? Using the data provided for at least one match, build a model that can predict these fluctuations in a match. Which factors, if any, seem to be most relevant. Given the difference in "momentum" in past matches, how do you recommend a player to play in a new match against a different player?

4. Test your developed model in one or more other competitions. How well do you predict fluctuations in the game? If models sometimes perform poorly, can you identify factors that might need to be included in future models? How general is your model in other competitions (such as women's competitions), tournaments, venue types, and other sports (such as table tennis)?

5. Write a report of no more than 25 pages and attach a one - to two-page memo summarizing your findings. Coaches are advised on the role of "momentum" and how to prepare players for events that affect the course of a tennis match.

## Our Work

# Assumptions and Justifications

1. **假设给出的数据均为真实数据，真实有效。**

我们的数据来源于赛题和官方网站。

**2.假设对于一些较为异常的数据的出现具有一定的合理性，我们对数据中缺失值的处理方式不会对预测结果造成太大的误差。**

我们对缺失值利用bp神经网络进行补充，并进行了较好的拟合，且与平均水平较为接近。

**3.假设表现得分和动量得分均只与自己有关，与对手的不同无关。**

在考虑构建模型去评价一位选手的表现得分和动量得分时，其指标的选取不可能充分考虑到因对手的不同和表现而进行评价，因为一局比赛中对手的发挥和策略无法完全预估和评价。

**4.假设使用马尔可夫链进行动量预测在网球比赛中是合理的。**

在网球比赛中，当前的局势演变通常受到前几个局势的影响，而且当前的局势转变往往只与前一个或几个局势的情况有关，符合马尔可夫性质。如果系统的状态转移概率在短期内基本保持不变，那么马尔可夫链模型可以很好地预测未来的状态。在网球比赛中，虽然局势会因为球员的表现、策略等因素而发生变化，但在短期内通常不会发生剧烈的变化，因此马尔可夫链模型可以适用于这种情况。

**1. It is assumed that all the data given are real and valid.**

Our data comes from the contest and the official website.

**2. Assuming that the occurrence of some relatively abnormal data is reasonable, the way we deal with the missing value in the data will not cause too much error to the prediction result.**

We use bp neural network to supplement the missing value, and carry out a good fit, and is close to the average level.

**3. Assume that performance score and momentum score are related only to oneself, regardless of the difference of the opponent.**

When considering the construction of a model to evaluate a player's performance score and momentum score, the selection of its indicators cannot fully consider the evaluation due to the difference and performance of the opponent, because the opponent's play and strategy in a game cannot be fully estimated and evaluated.

**4. Assume that momentum prediction using Markov chains is reasonable in tennis matches.**

In tennis, the current evolution of the situation is usually influenced by the previous situations, and the current transformation of the situation is often only related to the situation of the previous situation or a few situations, which conforms to the Markov nature. If the state transition probability of the system remains basically constant in the short term, then the Markov chain model can predict the future state well. In tennis, although the situation will change due to the player's performance, strategy and other factors, it usually does not change drastically in the short term, so the Markov chain model can be applied to this situation.

# Notations

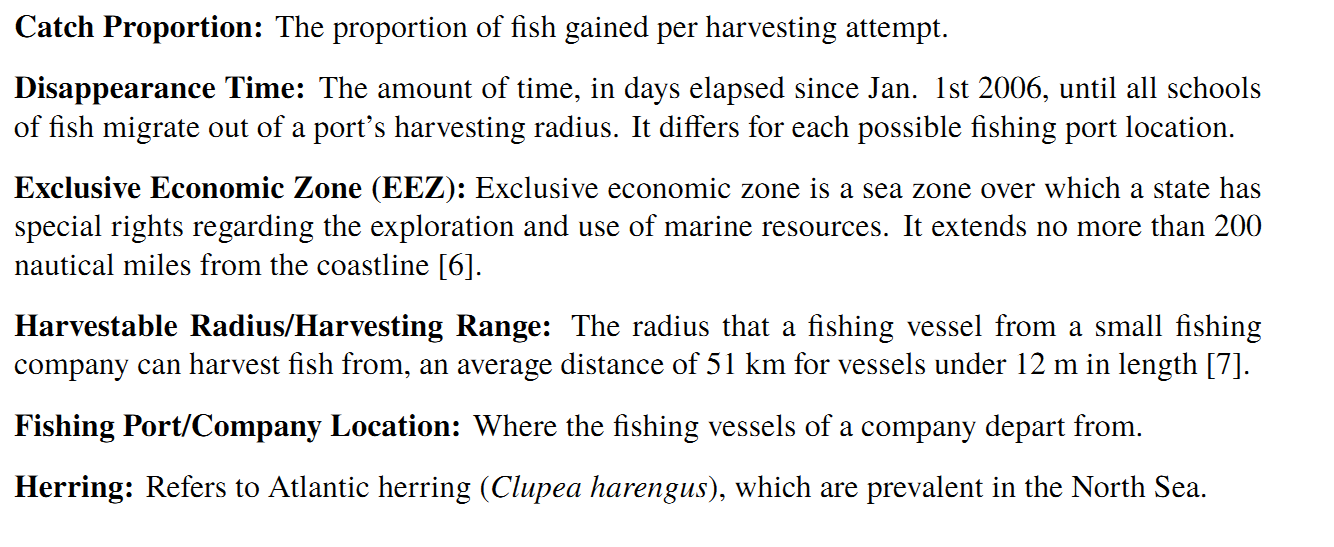
The key mathematical notations used in this paper are listed in Table 1.

Table 1: Notations used in this paper

|  |  |  |
| --- | --- | --- |
| **Symbol** | **Description** | **Unit** |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

本部分对应国赛论文中的符号（英文翻译为notation或者symbol）说明部分。

另外，少数论文有一个Deﬁnitions部分，这一部分主要是对文中出现的专有名词或者模糊的概念进行定义，这样可以帮助读者理解题目。



如果你需要写这一个部分的话，可以和Notations合并在一起变成一个大部分，该部分的命名为“Deﬁnitions and Notations”，写作时再把这个大的部分分成两个小的部分（用二级标题的方式）分别写即可。

# 问题一：表现得分的评价

首先，我们对数据集进行检查和清洗，发现数据表中rally\_count、return\_depth、speed\_mph、serve\_depth、serve\_width均存在缺失值，我们通过将数据集中的非缺失值作为训练集，将缺失值对应的特征作为目标值，构建BP神经网络模型，对缺失值进行插补。

一．特征提取

为了捕捉运动员表现随比赛流程时间的变化，提取数据特征如下：

|  |  |  |
| --- | --- | --- |
|  | 指标 | 取值 |
| X1（发球方、接球方均有） | 是否得分 | 0，1 |
| X2（仅发球 方有） | ACE | 0，1 |
| X3 （仅发球 方有） | 双误 | -1，0 |
| X4（发球方、接球方均有） | 非受迫性失误 | -1，0 |
| X5（发球方、接球方均有） | 上网得分 | 0，1 |
| X6 （仅发球 方有） | 发球宽度 | 见题给数据集 |
| X7 （仅发球 方有） | 发球深度 | 见题给数据集 |
| X8 （仅发球 方有） | 速度 | 见题给数据集 |
| X9（发球方、接球方均有） | 距离 | 见题给数据集 |
| X10（仅接球方有） | Winner | 0，1 |
| X11（仅接球方有） | 返回球深度 | 见题给数据集 |
| X12 （仅接球方有） | 是否破发点 | -1，0，1 |

对指标的描述：

1. 是否得分：反映了球员在比方上的优势，影响球员的心理状态。
2. ACE、发球深度和宽度：发球得分效率评估，代表球员的攻击力和技术优势，这些得分是对手无法回应的，对建立比赛势头非常重要。
3. 上网得分：往往需要较高的技术和敏捷性，显示球员在网前的主动性和技术实施能力。
4. 是否发球：在网球比赛中，发球方一般占据优势，记录每一分的发球者可以有助于我们理解比赛中的主动权转换。
5. 是否破发点：反映了球员在高压力情况下的表现。
6. 非受迫性失误，跑动距离：相持能力评估。

二．构建表现得分模型

发球表现得分模型：

接球表现得分模型：

三．指标权重

权重的确定对于评价指标的不同贡献至关重要。本文采用两种加权模型来计算权重向量。本文结合AHP法和critic法来确定我们评估模型的权重，这既增加了可信度，又反映了我们对每个因素的重视程度。

（1）AHP：

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：

一致性测试

对于上述的比较矩阵，我们得到了：发球方，n=9，CR=0.056<0.1；接球方，CR=0.005<0.1，因此比较矩阵是可以接受的。

一致性检验通过后，我们可以通过最大特征值对应的特征向量得到各主要因素的权重：

发球：[0.195,0.127,0.127,0.127,0.127,0.091,0.091,0.071,0.039]

接球：[0.327,0.123,0.123,0.045,0.123,0.123,0.132]

（2）Critic：

Critic权重法是一种客观赋权法，是基于评价指标的对比强度和指标之间的冲突性来综合衡量指标的客观权重。对比强度用标准差进行表示，如果数据标准差越大说明波动越大，权重会更高；冲突性用相关系数表示，如果指标之间相关系数值越大说明冲突性越小，那么其权重也就越低。

1. 对各个因素按照每个选项的数量进行归一化处理
2. 指标变异性

以标准差的形式来 ，表示第j个指标的标准差

在critic法中使用标准差来表示各指标的内取值的差异波动情况，标准差越大表示该指标的数值差异越大，越能反映出更多的信息，该指标本身的评价强度也就越强，应该给该指标分配更多的权重。

1. 指标冲突性

以相关系数的形式来表现，表示评价指标i和j之间的相关系数

使用相关系数来表示指标间的相关性，与其他指标的相关性越强，则该指标就与其他指标的冲突性越小，反映出相同的信息越多，所能体现的评价内容就越有重复之处，一定程度上也就削弱了该指标的评价强度，应该减少对该指标分配的权重。

1. 信息量

越大，第j个评价指标在整个评价指标体系中的作用越大，就应该给其分配更多的权重。

1. 权重

所以第j个指标的客观权重

得到指标权重如下：

发球：[0.077,0.109,0.109,0.109,0.109,0.127,0.127,0.125,0.103]

接球：[0.093,0.154,0.154,0.142,0.154,0.154,0.146]

（3）将两组权重组合起来。利用最小相对信息熵原理，建立优化模型，使两种决策方法下结果的相对偏差最小化。

采用拉格朗日乘数法求解上述优化问题，得到了最终权重：

发球：[0.405,0.377,0.377,0.377,0.377,0.312,0.312,0.239,0.111]

接球：[0.577,0.361,0.361,0.121,0.361,0.361,0.363]

则发球表现得分模型：

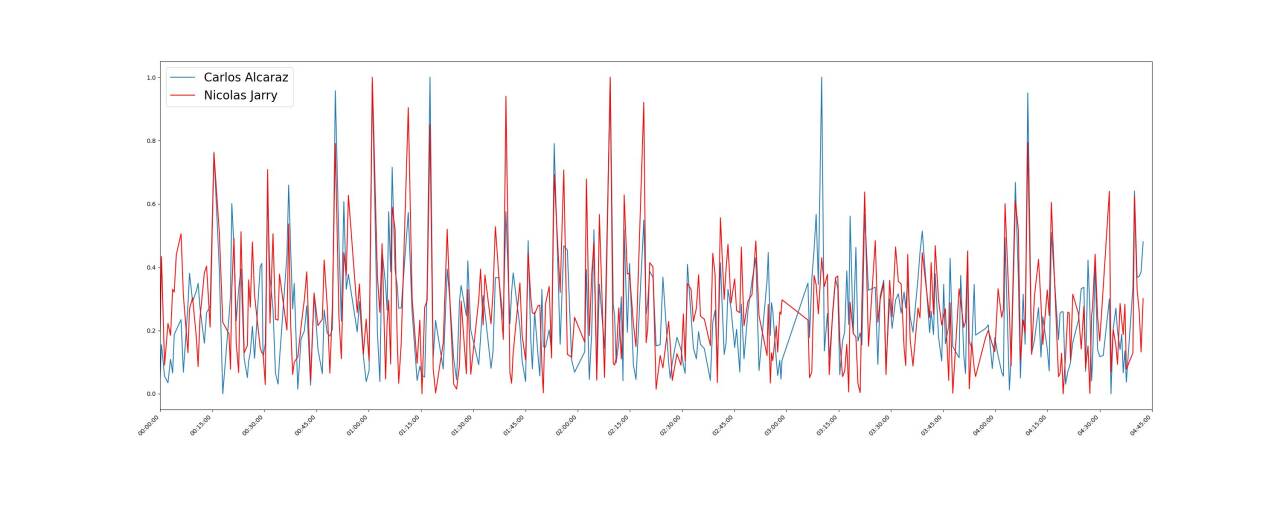
接球表现得分模型：

由此可见，在每一次对打时，上一次谁得分对本次对打双方影响最大。

对于发球方，ACE、双误、非受迫性失误、上网得分影响较大，发球宽度和发球深度对于是否能拿分也有一定影响，发球速度和距离的影响较小。

对于接球方，非受迫性失误、上网得分、winner、返回球深度、是否破发点影响较大，而距离影响较小。

四．将权重参数和“2023-Wimbledon-1701”这场比赛的数据带入表现得分公式，并将表现得分归一化，得到这场比赛两位选手表现得分随时间流程的变化情况如下



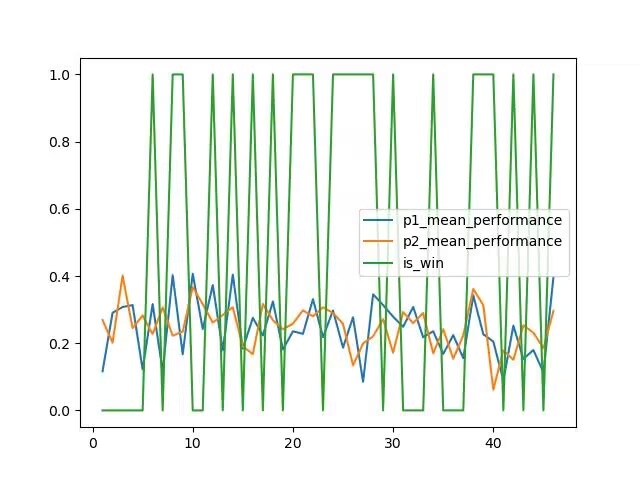
1.通过分析图，可以看出双方的表现得分为高峰值时均是发生在“自己为发球方”，“对手出现非受迫性失误”，“破发成功”，“保发成功”，“上网得分”；而表现得分为低谷时均发生在“自己出现失误”，“本局得分落后”，“双误”，“破发失败”，“保发失败”。

2.表现得分与真实比赛得分情况趋势一致，以00:00:00-00:15:00这段时间为例，首先是Nicolas Jarry得分高，接着是Carlos Alcaraz得分高，然后又是Nicolas Jarry得分高，与真实比赛数据一致，首先是Nicolas Jarry得分领先，接着是Carlos Alcaraz领先，最后是Nicolas Jarry领先。

将前六分钟的表现得分与真实比赛得分情况作比较如下，发现当某个选手赢球得分时，他相应的表现得分确实相较于对手要高，证明本文所建立模型的可靠性和正确性。

|  |  |  |
| --- | --- | --- |
|  | Carlos Alcaraz | Nicolas Jarry |
| 00:00:25 | 0.130 | 0.273（win） |
| 00:01:19 | 0.133（win）（破发） | 0.032 |
| 00:02:17 | 0.498（win） | 0.263 |
| 00:03:02 | 0.349 | 0.457（win） |
| 00:03:38 | 0.444（win）（保发） | 0.270 |
| 00:04:05 | 0.667 | 0.781（win） |

3.通过图也可以看出动量对不同的人影响也不同，Nicolas Jarry的波动幅度更大，说明动量对他的表现影响更大，而Carlos Alcaraz的波动则相对而言更小。



利用模型1计算出选手每局的表现得分，与实际结果对比可知，可以看出当选手1的表现得分高的时候，选手1赢得该局的概率更大。

# 5.问题二：动量对局势波动和连续成功的影响

首先我们对动量进行定义，“动量”在词典中的定义是“通过运动或一系列事件获得的力量”，它有时候会改变游戏规则，有时候会定义获胜者，它可以通过计算球员在特定时间窗口内相对于对手的得分优势来量化。

1. 以局来定义动量。

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。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。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。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。

1. 定义动量的指标

|  |  |  |
| --- | --- | --- |
| **连续得分** | wwin\_points | 在本局中最高的连续得分次数 |
| **连续赢的局数** | wwin\_games | 在本盘中到目前为止连续赢的最大局数 |
| **连续输的局数** | llose\_games | 在本盘中到目前为止连续输的最大局数 |
| **连续赢的盘数** | wwin\_sets | 在本次match中到目前为止连续赢的最大盘数 |
| **连续输的盘数** | llose\_sets | 在本次match中到目前为止连续输的最大盘数 |
| **发球优势** | sserve | 在网球比赛中，发球优势可以带来重要的竞技优势 |
| **连续失分** | llose\_points | 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。 |

3.定义动量计算公式：

对于动量指标权重的确定，沿用问题一中的critic法，计算权重为[0.107,0.105,0.118,0.147,0.145,0.057,0.321],

则

得到两位选手在“2023-Wimbledon-1701”这场比赛的动量得分随时间流程的情况如下



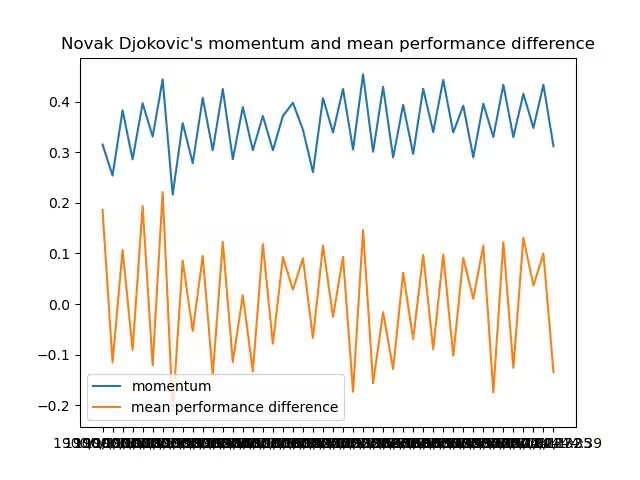
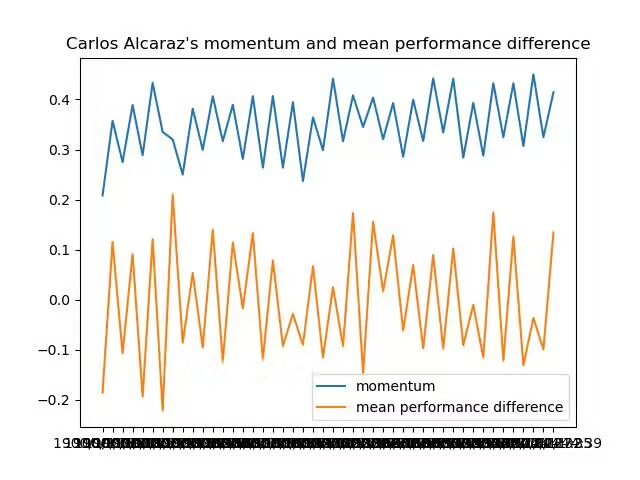
4.定义局势波动：

局势波动 =我当前的表现得分-对方的表现得分

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，可以捕捉到比赛中选手与对手之间的竞争情况。这有助于分析比赛中的关键时刻和转折点。局势波动的定义直观且易于理解，能够直接反映出选手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。

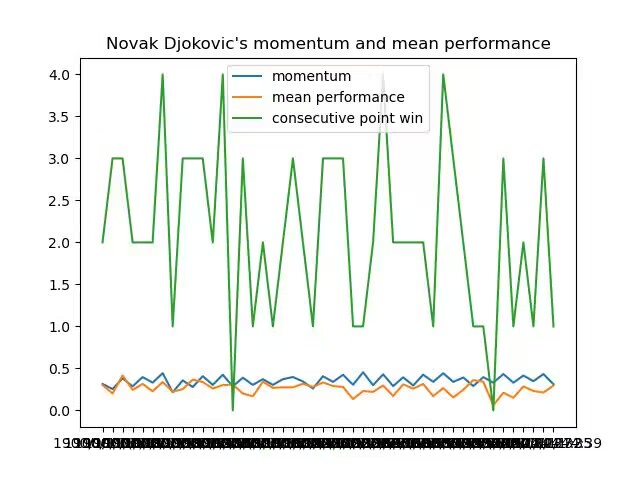
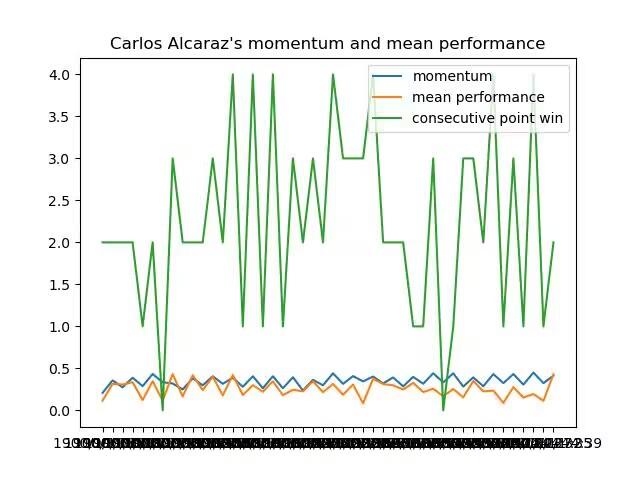
当选手的表现得分高于对手时，局势波动值会为正，表示走势偏向该选手，反之为负，这可以反映出比赛中的优劣势和局势的变化。

5.为了证明动量与局势波动有关联，我们将两位选手的动量情况和局势波动情况量化并可视化如下：



可以看出，动量和局势波动的转折点以及上升下降趋势大致保持一致，当某一位选手动量高涨时，局势也会向他那方偏移，且变化趋势和发展情况保持一致，证明动量对局势波动有影响。

6.为了证明动量对一个选手连续成功得分有影响，我们将两位选手的情况统计在如下图中，

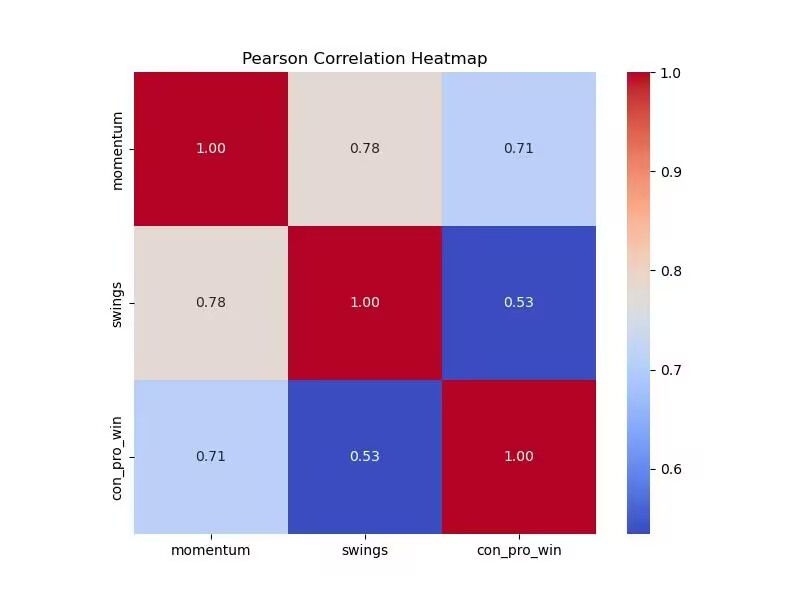


可以看出，当一个人的动量上升或者动量值高的时候，他更容易连续得分，；而一个人连续得分，他相应的动量值也就越高。两者是相辅相成的。

7.为了进一步说明动量与局势波动、连续成功之间的关系，我们建立回归模型来进行相关性分析，进行相关系数假设检验。

Pearson相关系数被广泛用来衡量两个变量之间的线性相关程度，其取值范围介于-1到1之间，其中正值表示正向相关，负值表示负向相关，而接近于0的值表示两个变量之间几乎没有线性相关性。计算公式如下：

求解结果如下



Pearson相关系数检验

对局势波动、连续成功和动量进行pearson相关系数检验，检验步骤如下：

Step1：提出假设

原假设：pearson相关系数R≠0

备择假设：pearson相关系数R=0

设定置信水平为95%

Step2：计算P值

利用Python进行相关系数检验，结果如下表所示：

|  |  |  |  |
| --- | --- | --- | --- |
|  | P值检验 | |  |
| Momentum | Swings | Con\_pro\_win |
| Momentum | 1 | 0.0 | 0.000594 |
| Swings | 0.0 | 1 | 0.000096 |
| Con\_pro\_win | 0.000594 | 0.000096 | 1 |

由表X可知，Swings和Con\_pro\_win的P值均小于置信水平0.05，拒绝原假设即Swings和Con\_pro\_win均与Momentum呈现显著的线性相关关系。

8.线性回归方程的验证与计算

由于Swings和Con\_pro\_win均与Momentum呈现显著的线性相关关系，因此可分别构建其线性回归方程：

Momentum与Swings：

|  |  |  |  |
| --- | --- | --- | --- |
|  | 标准化系数 | 标准误差 | P-value |
| 常数 | -0.504 | 0.047 | 8.54E-18 |
| Momentum | 1.432 | 0.131 | 3.93E-18 |

得到回归方程为

Momentum和Con\_pro\_win：

|  |  |  |  |
| --- | --- | --- | --- |
|  | 标准化系数 | 标准误差 | P-value |
| 常数 | -0.867 | 0.559 | 0.001 |
| Momentum | 8.649 | 1.565 | 3.15E-07 |

得到回归方程为

# 6.问题三：掌握局势的关键因素和策略

由问题二可知，局势波动是双方表现得分的差异，接下来我们将具体分析模型1中的每一个指标对局势波动的影响。

一．

我们用问题一中的指标与问题二中的走势波动一起构建随机森林模型，根据指标对局势波动进行预测，并计算出各个指标的重要性，选出与走势最相关的因素。

随机森林是一种强大的集成学习方法，它基于决策树构建，通过集成多棵决策树的预测结果来提高整体的预测性能，每棵决策树都是一个分类器。 随机森林利用“投票”机制来进行预测，即每棵决策树对输入样本进行预测，最终预测结果由多数决策树的投票决定。它的关键在于每棵决策树的随机性：它们在构建过程中使用了随机采样和随机特征选择，以增加模型的多样性和稳健性。

步骤：

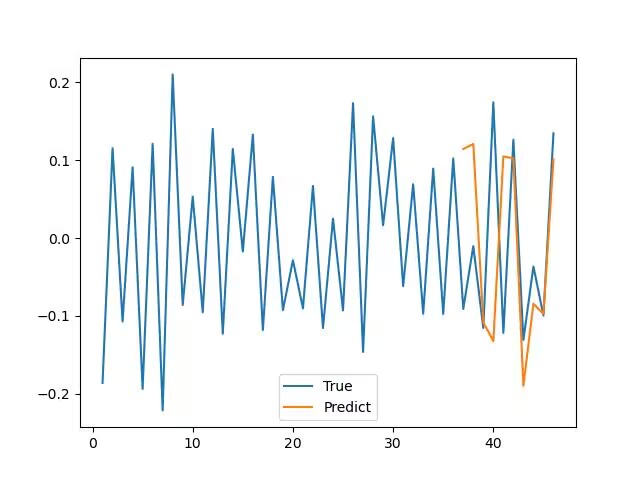
   1.随机采样（Bootstrap采样）：对“2023-Wimbledon-1701”这场比赛数据进行随机采样，有放回地从原始数据集中抽取相同大小的样本作为子数据集。这样每棵决策树的训练集都是随机抽样得到的，使得每棵决策树都是略有差异的。

   2.随机特征选择：对于每棵决策树的每个节点，在进行分裂时，只考虑一个随机子集的特征而不是所有特征。这样可以增加决策树之间的差异性，提高整体模型的泛化能力。

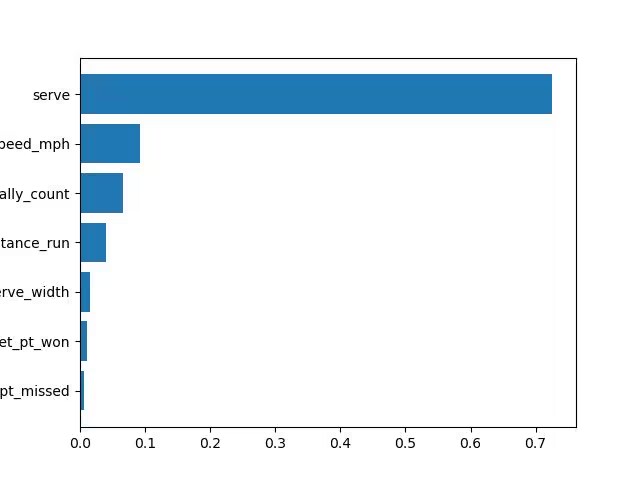
   3.决策树构建：基于上述随机采样和随机特征选择的训练集，构建多棵决策树。每棵决策树都是根据特征的信息增益或基尼系数等指标进行递归分裂，直到满足停止条件为止。

   4.集成预测：对于新样本的预测，通过让每棵决策树独立进行预测，然后根据投票原则或平均值来确定最终的预测结果。

局势波动的真实值与预测值如下图：



并得到各个指标对局势波动的影响大小如下：



可以看出不同指标对局势波动的影响程度不同，其中serve的权重最高，它对局势波动的影响最大

二．

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差异，预测未来的动量差异，为一名选手在与不同选手的新比赛中提供建议，从而使局势对我方更加有利， 动量差异的等级可以将比赛表现定量化，使得评估和比较不同比赛的结果更为方便。这有助于分析比赛结果，并从中获取更多的启示和经验。

首先，我们对动量的差异进行聚类，划分等级（为了根据不同的类别提供不同的建议），对已有的7个指标差异值进行聚类分析，聚为不同类 比如4类，用0，1，2，3代替 。

对于给定的一“2023-Wimbledon-1701”这场比赛含n个d维数据点的数据集X=,其中，以及要生成的数据子集的数目K，k-means聚类算法将数据对象组织为K个划分C=。每个划分代表一个类，每个类有一个类别中心。选取欧氏距离作为相似性和距离判断准则，计算该类内各点到聚类中心的距离平方和。

聚类目标是使各类的距离平方和J（C）=最小。

其中，。

显然，根据最小二乘法和拉格朗日原理，聚类中心应该取为类别类各数据点的平均值。

算法流程如下：

利用k值利用轮廓系数法确定为4，

|  |  |  |  |
| --- | --- | --- | --- |
| 类别 | 聚类中心 | 范围（结合聚类中心和过去比赛中动量差异范围得出） | 解释 |
| 0 | -0.128 | -7.5%~-15% | 动量比对方低较大水平 |
| 1 | -0.036 | -7.5%~0% | 动量比对方低较小水平 |
| 2 | 0.048 | 0%~7.5% | 动量比对方高较小水平 |
| 3 | 0.112 | 7.5%~15% | 动量比对方高较大水平 |

三．马尔可夫模型预测未来动量差异状态

马尔可夫链是一种基于状态转移概率的模型，可以用于预测系统在不同状态之间的转移情况。在这种情况下，我们将动量差异的不同等级视为状态，并利用历史数据来估计不同状态之间的转移概率。因为局势演变具有马尔可夫性质，在网球比赛中，当前的局势演变通常受到前几个局势的影响，而且当前的局势转变往往只与前一个或几个局势的情况有关并且在网球比赛中，虽然局势会因为球员的表现、策略等因素而发生变化，但在短期内通常不会发生剧烈的变化，因此马尔可夫链模型可以适用于这种情况。

客观的事物可能有,,…,共n种状态，其中每次只能处于一种状态，则每种状态都具有n个转向（包括转向自身），即

,,…,

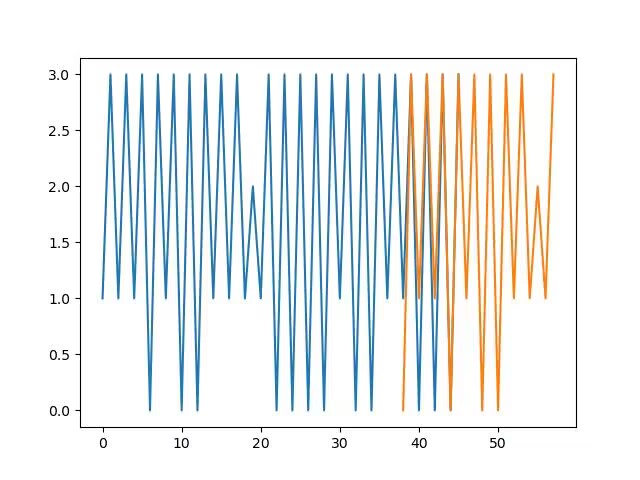
由于状态转移是随机的，因此，必须用概率来描述转移可能性的大小，将这种转移的可能性用概率描述，即为状态转移概率。

对于从状态转移到状态的概率，称为从i到j的转移概率，记为：

四类动量差异的转移状态矩阵收敛后的极限概率分布为

，

我们利用“2023-Wimbledon-1701”这场比赛数据来预测后20个game的动量差异



四．

（1）当根据预测模型的结果发现动量将有较大波动时，即处于0和3状态时，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。

5.赛前可以计算选手与对手的排名差异，为将要进行的比赛提供一个量化的预期难度，并且考虑球员近期的比赛数量和表现以及疲劳程度，做好赛前有针对性的计划和训练，并做好比赛策略，如对手的上网表现较弱则可以重点考虑这一点。

（2）当预测结果显示动量波动较小时，即处于1,2状态，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。排名较高的球员应该准备好承受压力，因为预期和心理压力会更高；相反排名较低的球员可以利用外界对他们的期望相对较低这一点来减轻压力。

# 7.问题四：所建立模型的普适性和应用价值

# 8.备忘录

MEMORANDUM

DATE:2月6号

TO:教练团队

FROM:某只建模团队

SUBJECT:势头在网球比赛中的影响和相应策略建议

尊敬的教练：

本报告是根据2023年温布尔登男子比赛的数据，以此来探究动量对网球比赛选手的作用以及影响因素。通过对比赛数据的分析，我们建立了相应模型来研究，得到以下关键发现：

比赛中确实存在势头现象，且势头在比赛中具有重要作用。

势头的形成因素：

1. 发球优势：发球方在比赛中通常具有更高的赢球概率。
2. 比赛心态：选手的比赛心态和紧张程度对势头的形成和维持有显著影响，积极的心态有利于球员在面临逆境时更好地利用势头。
3. 策略适用性：面对不同的对手准备好不同的策略，正确分析到对手的弱点有助于势头的增长。
4. 关键点得分：势头变化与比赛得分息息相关，特别是关键时刻，如破发、保发、抢七局或决胜局的表现优异有利于势头的积累，而势头的积累有利于提升选手的胜算。
5. 外界影响：观众的支持，教练的策略建议，以及战略性暂停可以在比赛关键时刻重新集结和调整球员的思路，有利于势头稳定。

策略建议：

1. 强化发球和接发球训练：鉴于发球局的重要性，选手需要提高发球水平，特别是提升发球质量和多样性，减少失误。
2. 心理韧性训练：教导球员管理压力，保持积极心态，特别是加强压力管理和情绪调节训练，提升球员的韧性、信心和应对压力的能力。培养球员在比赛中保持专注和集中注意力的能力，特别是在可能发生动量转换的关键时刻。实施心理训练练习，如想象、积极的自我对话和正念技巧，帮助球员在高压力下保持心理韧性和专注力。
3. 加强数据分析：利用以往比赛数据进行战术分析，根据不同对手建立不同的应对策略，识别对手的弱点和比赛中的势头转换点。根据球员反馈和比赛绩效数据的持续分析，定期审查和调整训练方案，重点关注如何利用动量转变的策略或减小其影响。

4.模拟训练：结合模拟训练，设计练习场景以模拟比赛中不同程度的动量变化，帮助球员培养韧性和适应能力。创建有竞争力的训练环境，让球员体验到真实的动量波动，并学会有效地应对。

总结：

势头的变化与比赛结果紧密相关，特别是关键节点，如抢七局和决胜盘，势头的积累有利于球员获得胜利。因其具有不容忽视的作用，在平时训练中，建立系统正确的训练模式和策略调整，可以有效地利用势头，提高比赛胜率。

敬请教练采纳上述分析和建议，衷心祝愿球员在接下来的每场比赛中都能取得优异成绩。

Dear coach,

This report is based on data from the 2023 Wimbledon Men's tournament to explore the effect of momentum on tennis players and the factors that influence them. Through the analysis of the competition data, we established the corresponding model to study, and obtained the following key findings:

Momentum does exist in the game, and momentum plays an important role in the game.

Momentum generators:

1. Service advantage: The server usually has a higher probability of winning the match.

2. Game mentality: Players' game mentality and intensity have a significant impact on the formation and maintenance of momentum, and a positive attitude helps players better use momentum in the face of adversity.

3. Strategy applicability: Prepare different strategies for different opponents and correctly analyze the weaknesses of opponents to help the momentum grow.

4. Key points: Momentum change is closely related to the score of the match, especially the excellent performance in key moments, such as break, save, tiebreak or decider, is conducive to the accumulation of momentum, and the accumulation of momentum is conducive to the improvement of the player's chances of victory.

5. Outside influence: The support of the crowd, the strategic advice of the coach, and the strategic timeout can regroup and adjust the players' thinking at the key moment of the game, which is conducive to the stability of the momentum.

Strategic Suggestions:

1. Strengthen service and receive training: In view of the importance of service games, players need to improve their service level, especially improve the quality and diversity of service, and reduce errors.

2. Mental toughness training: Teach players to manage pressure and maintain a positive attitude, especially strengthen pressure management and emotional regulation training to improve players' resilience, confidence and ability to cope with pressure. Develop a player's ability to stay focused and focused during a game, especially at critical moments when momentum shifts may occur. Implement mental training exercises such as visualization, positive self-talk and mindfulness techniques to help players maintain mental resilience and focus under high stress.

3. Strengthen data analysis: Make use of previous match data to conduct tactical analysis, establish different coping strategies according to different opponents, and identify opponents' weaknesses and momentum transition points in the match. Training programs are regularly reviewed and adjusted based on ongoing analysis of player feedback and match performance data, with a focus on how momentum shift strategies can be exploited or their impact reduced.

4. Simulation training: Combined with simulation training, practice scenarios are designed to simulate the momentum changes of different degrees in the game to help players develop resilience and adaptability. Create a competitive training environment where players experience real momentum swings and learn to cope effectively.

In summary:

The change in momentum is closely related to the outcome of the match, especially at key points, such as the tiebreak and the deciding set, the accumulation of momentum is conducive to the victory of the player. Because of its role can not be ignored, in peacetime training, the establishment of systematic and correct training mode and strategy adjustment, can effectively use the momentum, improve the winning rate of the game.

Please accept the above analysis and suggestions, and sincerely wish the players can achieve excellent results in every game in the future.

# 9.Sensitivity Analysis

在国赛论文写作视频中（<https://www.bilibili.com/video/BV1Na411w7c2/>），我介绍过怎么写模型的分析和检验这个部分：

模型的分析 ：在建模比赛中模型分析主要有两种，一个是灵敏度(性)分析，另一个是误差分析。灵敏度分析是研究与分析一个系统（或模型）的状态或输出变化对系统参数或周围条件变化的敏感程度的方法。其通用的步骤是：控制其他参数不变的情况下，改变模型中某个重要参数的值，然后观察模型的结果的变化情况。误差分析是指分析模型中的误差来源，或者估算模型中存在的误差，一般用于预测问题或者数值计算类问题。

模型的检验：模型检验可以分为两种，一种是使用模型之前应该进行的检验，例如层次分析法中一致性检验，灰色预测中的准指数规律的检验，这部分内容应该放在模型的建立部分；另一种是使用了模型后对模型的结果进行检验，数模中最常见的是稳定性检验，实际上这里的稳定性检验和前面的灵敏度分析非常类似，等会大家看到例子就明白了。

在美赛的写作中，写的最多的就是灵敏度分析（Sensitivity Analysis），因此这里我们的标题就直接取得是灵敏度分析；如果你既要写灵敏度分析，又要写误差分析（Error Analysis），那么你可以把标题改成： Sensitivity Analysis and Error Analysis

# 10.Model Evaluation and Further Discussion

## 10.1Strengths

1. 数据来源于真实比赛，为教练和球员提供精准的数据驱动的决策支持，使得模型建立和比赛策略更加科学和准确。
2. 本文对势头和局势波动的定义较为准确，模型能够识别并量化比赛中的势头变化。
3. 模型强调了发球和接发球的重要性，解释了比赛心态和情绪对势头的影响，为训练提供明确方向。
4. 对指标的选取和模型的选用均考虑了较为全面的因素，逻辑比较严谨。
5. 策略和建议从多方位进行考量和分析，较为全面，具有较好的稳定性。

1. Data comes from real matches, providing accurate data-driven decision support for coaches and players, making model building and game strategy more scientific and accurate.

2. The definition of momentum and situation fluctuation in this paper is more accurate, and the model can identify and quantify the momentum change in the match.

3. The model emphasizes the importance of serving and receiving, explains the influence of match mentality and emotion on momentum, and provides a clear direction for training.

4. The selection of indicators and models has taken into account relatively comprehensive factors, and the logic is relatively rigorous.

5. Strategies and suggestions are considered and analyzed from multiple aspects, which is more comprehensive and has good stability.

## 10.2Weaknesses

1. 模型高度依赖数据搜集的质量和完整性，对模型的建立的准确性和可靠性有一定影响。
2. 模型是基于特定比赛数据进行分析的，泛化能力到其他比赛可能有局限。
3. 心理和情绪的影响因素较多，指标的选取可能无法充分捕捉到。

1. The model is highly dependent on the quality and completeness of data collection, which has a certain impact on the accuracy and reliability of model establishment.

2. The model is analyzed based on specific competition data, and the generalization ability to other competitions may be limited.

3. There are many influencing factors of psychology and emotion, and the selection of indicators may not be able to fully capture them.

## 10.3Further Discussion

进行进一步的讨论，这里可以写模型的改进和拓展：

模型的改进：主要是针对模型中缺点有哪些可以改进的地方；

模型的拓展：将原题的要求进行扩展，进一步讨论模型的实用性和可行性。

# 11.Conclusion

结论部分，这个部分在国赛论文很少见到，但在美赛中出现的频率很高。

这个部分可以是论文中心思想的重申、研究结果或主要观点的归纳，也可以是某些启示性的解释或考虑。

有些论文把“Model Evaluation and Further Discussion”的内容放到了结论部分，这也是可以的，大家可以灵活调整。

# References

参考文献：所有引用他人或公开资料(包括网上资料)的成果必须按照科技论文的规范列出参考文献，并在正文引用处予以标注。

一般新起一页列出参考文献，如果上一个部分的下面有很多空白，那么就不用新起一页了。

美赛中不要出现中文，如果引用中文文献请翻译过来。

# Appendices

|  |
| --- |
| Appendix 1 |
| Introduce: 这里放上附录1的介绍 |
|  |

|  |
| --- |
| Appendix 2 |
| Introduce: 这里放上附录2的介绍 |
|  |

本部分是附录部分，美赛对于附录不是特别看重，今年还限制了论文的页数（从第二页开始编号，不能超过25页）。

一般新起一页列出附录。

在不超过页数限制的条件下，附录中可以包括下面内容：

* 你们写的代码；
* 某一问题的详细证明或求解过程；
* 自己在网上找到的数据；
* 比较大的流程图；
* 较繁杂的图表或计算结果。