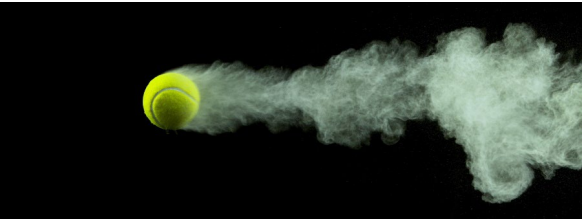
2024 MCM

Problem C: 网球运动的动力



2023年温网绅士决赛，20岁的西班牙新星卡洛斯·阿尔卡拉斯击败了36岁的诺瓦克·德约科维奇。 这是德约科维奇自2013年以来首次在温网失利并结束了大满贯历史上最伟大球员之一的非凡表现。

这场比赛本身就是一场非凡的战斗。[1] 德约科维奇似乎注定会轻松获胜，因为他

以6-1统治首盘（7场比赛中赢下6场）。 然而第二盘比赛气氛紧张最终阿尔卡雷斯在决胜局中以7-6获胜。第三盘与第一盘相反，阿尔卡拉斯6-1 轻松获胜。第四盘开始时，这位年轻的西班牙人似乎完全掌控了局面，但是不知何故，比赛再次改变了方向，德约科维奇完全掌控了比赛并赢得了6– 3. 第五盘也是最后一盘，德约科维奇在第四盘中占据优势，但再次方向发生改变，阿尔卡拉斯控制了局面，以 6-4 获胜。本次比赛的数据比赛在提供的数据集中，“match\_id”为“2023-wimbledon-1701”。 你可以看到所有当德约科维奇使用“set\_no”列等于 1 的优势时，第一盘的得分。令人难以置信的波动，有时是很多分甚至是比赛，发生在球员身上似乎拥有优势的人常常将其归因于“势头”。

字典中对动量的定义是“通过运动或一系列运动获得的力量或力量”事件。”[2] 在体育运动中，团队或运动员可能会觉得他们有动力，或“力量/力量”在比赛/比赛期间，但很难衡量这种现象。 此外，也不容易明显地看到比赛期间的各种事件如何产生或改变动力（如果存在）.

提供 2023 年温网男子比赛前 2 场比赛后所有得分的数据回合。 您可以自行决定选择包含其他玩家信息或其他数据，但您必须完整记录来源。 使用数据：

* 开发一个模型，捕捉得分发生时的比赛流程，并将其应用于一场或多场比赛。 您的模型应该确定哪位球员在比赛的给定时间表现更好，以及他们的表现有多好。 根据您的模型提供可视化来描述比赛流程。

注意：在网球比赛中，发球的球员赢得分数/比赛的概率要高得多。 您可能希望以某种方式将其纳入您的模型中。

* 一位网球教练对“势头”在比赛中发挥的作用表示怀疑。 相反，他假设一名球员的比赛波动和成功是随机的。 使用您的模型/指标来评估此声明。
* 教练们很想知道是否有指标可以帮助确定比赛流程何时从有利于一名球员转向另一名球员。
* 使用至少一场比赛提供的数据，开发一个模型来预测比赛中的这些波动。 哪些因素似乎最相关（如果有的话）？
* 考虑到过去比赛“势头”波动的差异，您如何建议一名球员与另一名球员进行新的比赛？
* 在一场或多场其他比赛中测试您开发的模型。 您对比赛走势的预测能力如何？ 如果模型有时表现不佳，您能否确定未来模型中可能需要包含的任何因素？ 您的模型对其他比赛（例如女子比赛）、锦标赛、球场和其他运动（例如乒乓球）的通用性如何。
* 制作一份不超过 25 页的报告，其中包含您的发现，并包括一份一到两页的备忘录，总结您的结果，并为教练提供有关“势头”作用的建议，以及如何让球员做好准备应对影响流程的事件 网球比赛期间的比赛。

您的 PDF 解决方案总页数不超过 25 页，应包括：

• 一页摘要表。

• 目录。

• 您的完整解决方案。

• 一到两页的备忘录。

• 参考文献列表。

• AI 使用报告（如果使用，则不计入 25 页限制。）

注意：完整的 MCM 提交没有具体要求的最小页面长度。 你您的所有解决方案工作以及您想要的任何其他信息总共最多可使用 25 页包括（例如：图纸、图表、计算、表格）。 部分解决方案被接受。我们允许谨慎使用人工智能，例如 ChatGPT，尽管没有必要创建一个这个问题的解决方案。 如果您选择使用生成式 AI，则必须遵循 COMAP人工智能使用政策。 这将产生一份额外的人工智能使用报告，您必须将其添加到您的报告末尾PDF 解决方案文件，不计入解决方案 25 页的总页数限制。

提供的文件：

• Wimbledon\_featured\_matches.csv – 2023 年温布尔登绅士赛的数据集

第二轮之后的单打比赛。

• data\_dictionary.csv – 数据集的描述。

• data\_examples – 帮助理解所提供数据的示例

**词汇表**

大满贯（Grand Slam）：网球大满贯是指在一个日历年内赢得一个项目的全部四个大满贯冠军的成就。 四大满贯赛事分别是澳大利亚网球公开赛、法国网球公开赛、温布尔登网球公开赛和美国网球公开赛，每项比赛都为期两周以上

**关键术语/概念词汇表：**

- **得分(Scoring)**：[3]

o 比赛(Match)：五盘三胜（温布尔登的绅士比赛）

o Set：游戏集合； 6局获胜一盘，但在进行决胜局时，选手必须以两局获胜，直到盘数为6-6平（见下文）

o 游戏(Game)：收集积分； 玩家达到 4 分时获胜，但必须以 2 分获胜。 请参阅下面的“比赛得分”。

**比赛得分**：[3]

o 0 分 = Love

o 1 分 = 15

o 2 分 = 30

o 3 分 = 40

o 并列分数(Tied score) = All（例如，“全部 30 分”）

o 40 – 40 = 平局（玩家赢得相同点数，每人至少 3 分）

o 发球手赢得平分（Server wins a deuce point） = Ad-in（或“优势”）

o 接发球手赢得平分（Receiver wins a deuce point） = Ad-out(优势出局)

**发球**：球员轮流担任“发球手”（击中一分的初始击球的球员）和“接发球手”。 在职业网球比赛中，发球者往往具有很大的优势。 球员在每个点上有两次发球机会将球投入比赛（进入“发球区”）。

在比赛中两次尝试未能击中发球将被视为“双误”，返回的球员将获得该分。

o **破发（**Breaking serve**）** – 当返回的球员赢得比赛时。

o **破发点（**Break point**）**——如果回攻者获胜，他们将赢得比赛。

o **保持发球（**Holding serve**）** – 当发球选手赢得比赛时。

**决胜局（**Tie-breakers**）**：当一方选手赢得 6 局比赛且至少领先两局（即 6 - 4）时，每盘结束。 如果没有，比赛将继续进行，直至打成 6-6 平局。 此时进行决胜局。 在温布尔登，决胜局首先获得 7 分（必须以2 分），但在比赛的第 5 盘率先达到 10 分时除外（必须以 2 分获胜）

**休息时间/球场两侧（**Rest breaks/sides of court**）**：球员在第一场比赛后以及每两场比赛后交换球场两侧。 从第三场比赛开始，每次换边都允许有 90 秒的休息时间。 在决胜局中，球员每六分就换边。 每局比赛结束后，选手们也休息至少2分钟。 允许医疗暂停和上厕所一次

References: [1] Braidwood, J. (2023), Novak Djokovic has created a unique rival – is Wimbledon defeat the beginning of the end, The Independent, https://www.independent.co.uk/sport/tennis/novak-djokovic-wimbledon-final-carlos-alcarazb2376600.html.

[2] <https://www.merriam-webster.com/dictionary/momentum>

[3] Rivera, J. (2023), Tennis scoring, explained: A guide to understanding the rules terms & point system at Wimbledon, The Sporting News, https://www.sportingnews.com/us/tennis/news/tennis-scoring-explained-rules-system-pointsterms/7uzp2evdhbd11obdd59p3p1cx