

1 A 题：资源可用性和性别比例

1.1 赛题翻译



虽然一些动物物种存在于通常的雄性或雌性之外，但大多数物种基本上都是雄性或雌性。虽然许多物种在出生时表现出 1:1 的性别比例，但其他物种偏离了均匀的性别比例。这被称为适应性性别比变异。例如，美洲鳄孵化卵的巢穴温度会影响出生时的性别比。

七鳃鳗的作用是复杂的。在一些湖泊栖息地，它们被视为对生态系统有重大影响的寄生虫，而七鳃鳗也是世界上一些地区的食物来源，如斯堪的纳维亚半岛、波罗的海地区，以及北美太平洋西北地区的一些土著民族。

七鳃鳗的性别比例可以根据外部环境而变化。七鳃鳗在幼虫阶段的发育速度决定了它们的性别是雄性还是雌性。这些幼虫的生长速度受到食物供应的影响。在食物可得性较低的环境中，生长速度会较低，雄性的比例可达到种群的 78% 左右。在食物更容易获得的环境中，观察到雄性的比例约占人口的 56%。

我们关注的是性别比例问题及其对当地条件的依赖，特别是对于海七鳃鳗。七鳃鳗生活在湖泊或海洋栖息地，并向上游洄游产卵。这项任务是检查一个物种根据资源可用性改变其性别比例的能力的利弊。你的团队需要开发和检查一个模型，以提供对生态系统中由此产生的相互作用的见解。

需要检查的问题包括：

- 当七鳃鳗的种群可以改变其性别比例时，对更大的生态系统有什么影响？
- 七鳃鳗种群的优势和劣势是什么？
- 鉴于七鳃鳗性别比例的变化，对生态系统稳定性的影响是什么？
- 七鳃鳗种群性别比例变化的生态系统能否为生态系统中的其他物种(如寄生虫)提供优势？

1.2 赛题要求

您的 PDF 解决方案总页数不超过 25 页，应包括：

- 一页总结表。
- 目录表。
- 完整的解决方案。• 参考文献。
- AI 使用报告(如果使用，不计入 25 页的限制)

注意:对于完整的 MCM 提交，没有特定的最低页数要求。您可以使用最多 25 页的总页数来完成所有解决方案工作和您想要包含的任何其他信息(例如:图纸，图表，计算，表格)。部分解决方案是可以接受的。我们允许谨慎地使用 AI，

如 ChatGPT，尽管没有必要为这个问题创建一个解决方案。如果您选择使用生成式 AI，则必须遵循 COMAP AI 使用策略。这将导致额外的 AI 使用报告，您必须将其添加到 PDF 解决方案文件的末尾，并且不计入解决方案的总页面限制 25 页。

1.3 术语表

七鳃鳗(有时被不准确地称为**七鳃鳗**)是一种古老的无颌鱼类，属于**岩螯目**。成年七鳃鳗的特征是长有牙齿、漏斗状的吸吮嘴。七鳃鳗主要生活在沿海和淡水中，在大多数温带地区都有发现。

1.4 使用 AI 工具声明

在 COMAP 竞赛中使用大型语言模型和生成式 AI 工具

这一政策的动机是大型语言模型(LLMs)和生成 AI 辅助技术的兴起。该政策旨在为团队、顾问和评委提供更大的透明度和指导。这项政策适用于学生工作的各个方面，从模型的研究和开发(包括代码创建)到书面报告。由于这些新兴技术正在迅速发展,COMAP 将适当地完善这一策略。

团队必须公开和诚实地使用 AI 工具。一个团队及其提交的内容越透明，他们的工作就越有可能得到他人的充分信任、赞赏和正确使用。这些披露有助于理解智力工作的发展和对贡献的适当承认。如果没有对 AI 工具作用的公开和清晰的引用和参考，那么有问题的段落和工作更有可能被认定为抄袭并被取消资格。

解决这些问题不需要使用 AI 工具，尽管允许负责任地使用它们。COMAP 认识到 LLMs 和生成 AI 作为生产力工具的价值，可以帮助团队准备提交;例如，为一个结构产生初步的想法，或者在总结、释义、语言润色等时。在模型开发的许多任务中，人类的创造力和团队合作是必不可少的，对 AI 工具的依赖会带来风险。因此，我们建议在将这些技术用于模型选择和构建、协助创建代码、解释模型的数据和结果以及得出科学结论等任务时要谨慎。

值得注意的是，LLMs 和生成式 AI 有局限性，无法取代人类的创造力和批判性思维。COMAP 建议团队在选择使用 LLMs 时要意识到这些风险：

- 客观性:LLMs 生成的文本中可能出现先前发表的包含种族主义、性别歧视或其他偏见的内容，一些重要观点可能未被代表。

- 准确性:LLMs 可能会产生“幻觉”，即产生虚假内容，特别是在他们的领域之外使用或处理复杂或模棱两可的主题时。他们可以生成语言上但科学上不合理的内容，他们可以错误地获取事实，并且他们已经被证明可以生成不存在的引用。一些 LLMs 只接受特定日期之前发布的内容的培训，因此呈现的是不完整的画面。

- 语境理解:LLMs 不能将人类的理解应用到一篇文章的语境中，特别是在处理习惯用语、讽刺、幽默或隐喻语言时。这可能会导致生成的内容出现错误或误解。

- 训练数据:LLMs 需要大量高质量的训练数据来达到最佳性能。然而，在某些领域或语言中，这样的数据可能并不容易获得，从而限制了任何输出的有用性。

1.5 团队指南

参赛队伍需要：

1. 在**报告中明确指出使用了 LLMs 或其他 AI 工具**，包括使用了哪个模型以及用于什么目的。请使用内联引文和参考文献部分。在你的 25 页解决方案之

后，还要附上 AI 使用报告(如下所述)。

2.验证内容的**准确性**、有效性和**适当性**以及由语言模型生成的任何引用，并纠正任何错误或不一致之处。

3.提供引用和参考文献，遵循这里提供的指导。仔细检查引文，以确保它们是准确的，并被正确引用。

4.要注意**抄袭的可能性**，因为 LLMs 可能会复制其他来源的大量文本。检查原始来源，以确保你没有抄袭别人的作品。

当 COMAP 确定提交的内容未公开使用了这些工具，我们会采取相应的行动。

1.6 引文与参考文献说明

仔细考虑如何记录和引用团队可能选择使用的任何工具。各种风格指南开始纳入引用和参考人工智能工具的政策。在你的 25 页解决方案的参考部分，使用内联引用并列出所有使用的人工智能工具。

无论团队是否选择使用人工智能工具，主要解决方案报告仍然限制在 25 页。如果一个团队选择使用人工智能，在你的报告结束后，添加一个名为人工智能使用报告的新部分。这个新章节没有页数限制，不会被计入 25 页的解决方案中。

例子(这不是详尽的-根据你的情况调整这些例子):

人工智能使用报告

1. OpenAI ChatGPT (Nov 5, 2023 version, ChatGPT-4)Query1: <insert the exact wording you input into the AI tool>Output: <insert the complete output from the AI tool>

2. OpenAI Ernie (Nov 5, 2023 version, Ernie 4.0)

Query1: <insert the exact wording of any subsequent input into the AI tool>

Output: <insert the complete output from the second query>

3. Github CoPilot (Feb 3, 2024 version)

Query1: <insert the exact wording you input into the AI tool>Output: <insert the complete output from the AI tool>

4. Google Bard (Feb 2, 2024 version) Query: <insert the exact wording of your query> Output: <insert the complete output from the AI tool>

2 B 题：搜寻潜水器

2.1 赛题翻译



MCMS 是一家总部位于希腊的公司，该公司制造的潜水器能够将人类带到海洋的最深处。一艘潜水器被移动到该地点，并与主船不受束缚地部署。MCMS 现在希望用他们的潜水器带领游客冒险探索爱奥尼亚海的海底沉船残骸。然而，在此之前，他们需要通过制定安全程序来获得监管机构的批准，以防与主船失去通信，以及可能出现的机械缺陷，包括潜水器失去推进力。特别是，他们希望你能开发一个模型来预测潜水器随时间的位置。与在陆地或海面上进行的典型搜救不同，有缺陷的潜水器可能会发现自己位于海底或水下某个中性浮力点。它的位置可能会进一步受到洋流、海洋中不同密度和/或海底地理位置的影响。你们的任务是：

- **定位**—开发一个模型，预测潜水器随时间的位置。与这些预测相关的不确定性是什么？在事故发生前，潜水器可以定期向主船发送哪些信息来减少这些不确定性？潜水器需要什么样的设备才能做到这一点？
- **准备**—如果有的话，你会建议公司在主船上携带哪些额外的搜索设备，以便在必要时部署？您可以考虑不同类型的设备，但也必须考虑与该设备的可用性、维护、准备和使用相关的成本。如果有必要，救援船可能需要携带哪些额外的设备来协助？
- **搜索**—开发一个模型，该模型将使用您的位置模型中的信息来推荐设备的初始部署点和搜索模式，从而最大限度地减少找到失踪潜水器的时间。确定找到潜水器的概率作为时间和累积搜索结果的函数。
- **探索**—你的模型如何扩展到其他旅游目的地，如加勒比海？你的模型将如何改变，以考虑在同一地区移动的多个潜水器？

2.2 赛题要求

准备一份不超过 25 页的报告，提供你的计划的细节。包括一份给希腊政府的两页报告备忘录，以帮助获得批准。

不超过 25 页的 PDF 解决方案应包括：

- 一页总结表。
- 目录表。
- 完整的解决方案。
- 两页备忘录。
- 参考书目。
- AI 使用报告(如果使用，不计入 25 页的限制)

注意：对于完整的 MCM 提交，没有特定的最低页数要求。您可以使用最多

25 页的总页数来完成所有解决方案工作和您想要包含的任何其他信息(例如:图纸, 图表, 计算, 表格)。部分解决方案是可以接受的。我们允许谨慎地使用 AI, 如 ChatGPT, 尽管没有必要为这个问题创建一个解决方案。这将导致额外的 AI 使用报告, 您必须将其添加到 PDF 解决方案文件的末尾, 并且不计入解决方案的总页数限制。

2.3 术语表

潜水器:潜水器是一种需要由更大的船只或平台运输和支撑的水下交通工具。这是潜水器与潜艇的区别, 后者是自给自足的, 能够在海上长时间独立作业。

中性浮力发生在物体的平均密度等于其浸入流体的密度时, 导致浮力平衡重力, 否则重力会导致物体下沉(如果物体的密度大于浸入流体的密度)或上升(如果小于)。具有中性浮力的物体既不会下沉也不会上升。

2.4 使用 AI 工具声明

在 COMAP 竞赛中使用大型语言模型和生成式 AI 工具

这一政策的动机是大型语言模型(LLMs)和生成 AI 辅助技术的兴起。该政策旨在为团队、顾问和评委提供更大的透明度和指导。这项政策适用于学生工作的各个方面, 从模型的研究和开发(包括代码创建)到书面报告。由于这些新兴技术正在迅速发展, COMAP 将适当地完善这一策略。

团队必须公开和诚实地使用 AI 工具。一个团队及其提交的内容越透明, 他们的工作就越有可能得到他人的充分信任、赞赏和正确使用。这些披露有助于理解智力工作的发展和对贡献的适当承认。如果没有对 AI 工具作用的公开和清晰的引用和参考, 那么有问题的段落和工作更有可能被认定为抄袭并被取消资格。

解决这些问题不需要使用 AI 工具, 尽管允许负责任地使用它们。COMAP 认识到 LLMs 和生成 AI 作为生产力工具的价值, 可以帮助团队准备提交;例如, 为一个结构产生初步的想法, 或者在总结、释义、语言润色等时。在模型开发的许多任务中, 人类的创造力和团队合作是必不可少的, 对 AI 工具的依赖会带来风险。因此, 我们建议在将这些技术用于模型选择和构建、协助创建代码、解释模型的数据和结果以及得出科学结论等任务时要谨慎。

值得注意的是, LLMs 和生成式 AI 有局限性, 无法取代人类的创造力和批判性思维。COMAP 建议团队在选择使用 LLMs 时要意识到这些风险:

- 客观性: LLMs 生成的文本中可能出现先前发表的包含种族主义、性别歧视或其他偏见的内容, 一些重要观点可能未被代表。

- 准确性: LLMs 可能会产生“幻觉”, 即产生虚假内容, 特别是在他们的领域之外使用或处理复杂或模棱两可的主题时。他们可以生成语言上但科学上不合理的内容, 他们可以错误地获取事实, 并且他们已经被证明可以生成不存在的引用。一些 LLMs 只接受特定日期之前发布的内容的培训, 因此呈现的是不完整的画面。

- 语境理解: LLMs 不能将人类的理解应用到一篇文章的语境中, 特别是在处理习惯用语、讽刺、幽默或隐喻语言时。这可能会导致生成的内容出现错误或误解。

- 训练数据: LLMs 需要大量高质量的训练数据来达到最佳性能。然而, 在某些领域或语言中, 这样的数据可能并不容易获得, 从而限制了任何输出的有用性。

2.5 团队指南

参赛队伍需要:

2. 在**报告**中明确指出使用了 **LLMs 或其他 AI 工具**，包括使用了哪个模型以及用于什么目的。请使用内联引文和参考文献部分。在你的 25 页解决方案之后，还要附上 AI 使用报告(如下所述)。

2. **验证内容的准确性、有效性和适当性**以及由语言模型生成的任何引用，并纠正任何错误或不一致之处。

3. **提供引用和参考文献**，遵循这里提供的指导。仔细检查引文，以确保它们是准确的，并被正确引用。

4. **要注意抄袭的可能性**，因为 LLMs 可能会复制其他来源的大量文本。检查原始来源，以确保你没有抄袭别人的作品。

当 **COMAP** 确定提交的内容未公开使用了这些工具，我们会采取相应的行动。

2.6 引文与参考文献说明

仔细考虑如何记录和引用团队可能选择使用的任何工具。各种风格指南开始纳入引用和参考人工智能工具的政策。在你的 25 页解决方案的参考部分，使用内联引用并列所有使用的人工智能工具。

无论团队是否选择使用人工智能工具，主要解决方案报告仍然限制在 25 页。如果一个团队选择使用人工智能，在你的报告结束后，添加一个名为人工智能使用报告的新部分。这个新章节没有页数限制，不会被计入 25 页的解决方案中。

例子(这不是详尽的-根据你的情况调整这些例子):

人工智能使用报告

1. OpenAI ChatGPT (Nov 5, 2023 version, ChatGPT-4)Query1: <insert the exact wording you input into the AI tool>Output: <insert the complete output from the AI tool>

2. OpenAI Ernie (Nov 5, 2023 version, Ernie 4.0)

Query1: <insert the exact wording of any subsequent input into the AI tool>
Output: <insert the complete output from the second query>

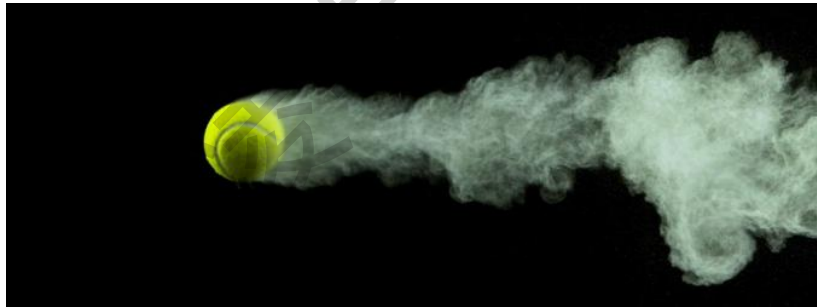
3. Github CoPilot (Feb 3, 2024 version)

Query1: <insert the exact wording you input into the AI tool>Output: <insert the complete output from the AI tool>

4. Google Bard (Feb 2, 2024 version) Query: <insert the exact wording of your query> Output: <insert the complete output from the AI tool>

3 C 题：网球的动量

3.1 赛题翻译



在 2023 年温布尔登男单决赛中，20 岁的西班牙新星卡洛斯·阿尔卡拉兹击败了 36 岁的诺瓦克·德约科维奇。这是德约科维奇自 2013 年以来在温布尔登的首次失利，也终结了这位大满贯史上最伟大球员之一的辉煌战绩。

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要轻松获胜，因为他以 6 比 1 的比分控制了第一局(7 局中赢了 6 局)。然而第二局气氛紧张，最终阿尔卡拉兹在抢七局中以 7-6 获胜。第三局与第一局相反，阿尔卡拉兹以 6-1 轻松获胜。这位年轻的西班牙人在第四盘开始时似乎完全控制了比赛，但不知何故，比赛再次改变了方向，德约科维奇完全控制了比赛，以 6 比 3 赢得了比赛。第五盘也是最后一盘，德约科维奇从第四盘开始保持优势，但再次改变方向，阿尔卡拉兹控制局面，以 6-4 获胜。本场比赛数据在提供的数据集中，“2023-温布尔登-1701”的“match_id”。你可以看到德约科维奇在第一盘领先时的所有得分，使用“set_no”列等于 1。出现在看似有优势的球员身上的不可思议的挥拍，有时是多分甚至多局，往往被归结为“冲劲”。[2]在体育运动中，一支球队或一名球员可能会在比赛/比赛中感觉到他们有动量，或“力量/力量”，但这种现象很难测量。此外，如果动量存在的话，比赛中的各种事件是如何产生或改变动量的，这一点也不容易弄清楚。

2023 年温布尔登男单前两轮之后的每一分数据。您可以自行选择加入额外的球员信息或其他数据，但您必须完整地记录来源。使用这些数据：

- 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。你的模型应该识别出哪位球员在比赛的特定时间表现得更好，以及他们的表现有多好。你可能希望以某种方式把这个因素考虑到你的模型中。
- 一位网球教练对“动量”在比赛中起作用持怀疑态度。相反，他认为一个球员在比赛中的摇摆和成功是随机的。用你的模型/指标来评估这一说法。
- 教练们很想知道是否有指标可以帮助确定何时
- 比赛流程即将从有利于一名球员转变为另一名球员。
- 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)?
- 考虑到过去比赛中“动量”波动的差异，你如何建议一名球员进入一场与不同球员的新比赛?
- 在一场或多场其他比赛中测试你开发的模型。你对比赛中挥杆的预测有

多好?如果模型有时表现不佳,你能识别出任何可能需要纳入未来模型的因素吗?你的模型对其他比赛(如女子比赛)、锦标赛、球场表面和其他运动(如乒乓球)的泛化程度如何?

- 用你的发现制作一份不超过 25 页的报告,并包括一到两页的备忘录,总结你的结果,并就“势头”的作用向教练提出建议,以及如何让球员准备好应对影响网球比赛过程的事件。

3.2 赛题要求

不超过 25 页的 PDF 解决方案应包括:

- 一页总结表。
- 目录表。
- 完整的解决方案。
- 一到两页备忘录。
- 参考书目。
- AI 使用报告(如果使用,不计入 25 页的限制)

注意:对于完整的 MCM 提交,没有特定的最低页数要求。您可以使用最多 25 页的总页数来完成所有解决方案工作和您想要包含的任何其他信息(例如:图纸,图表,计算,表格)。部分解决方案是可以接受的。我们允许谨慎地使用 AI,如 ChatGPT,尽管没有必要为这个问题创建一个解决方案。这将导致额外的 AI 使用报告,您必须将其添加到 PDF 解决方案文件的末尾,并且不计入解决方案的总页数限制。

3.3 提供文件

• **Wimbledon_featured_matches.csv** -温布尔登 2023 绅士单打第二轮后的比赛数据集。

• **data_dictionary.csv** -数据集的描述。

• **data_examples** -帮助理解所提供数据的示例。

3.4 术语表

大满贯:网球的大满贯是在一个日历年内赢得所有四个大满贯的成就。四项大满贯赛事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,每项比赛都持续两周。

得分:[3]

• **比赛:**五局四胜制(温布尔登绅士赛)

• **盘:**比赛合集;6 局赢一局,但选手必须先赢两局,直到决胜局以 6 比 6 打平(见下文)

• **局:**收点;玩家达到 4 分时获胜,但必须赢 2 分。见下文“一局得分”。

一场得分:[3]

• 0 分=LOVE

• 1 分=15 分

• 2 分=30 分

• 3 分=40 分

• 平局=所有(例如,“30 All”)

• 40 -40 =平分(玩家获得相同的分数,每人至少 3 分)

- 服务器赢得一个平分点= Ad-in(或 “advantage in”)

- 接发球方赢得一分= Ad-out

发球:玩家交替担任“发球者”(击球第一个点的玩家)和“接发球者”。在职业网球比赛中,发球者往往有很大的优势。在每个点上,球员有两次发球机会将球送入比赛(送入“发球箱”)。在两次发球尝试中发球失败是“双误”,回发球的球员得分。

- 破发球——当回击球员赢得一场比赛时。

- 破发点——在这个点上,如果接发球者获胜,他们将赢得比赛。o 守发球——当发球球员赢得比赛时。

- 抢七:每一局结束时,只要一名球员赢了6局,只要他们领先至少2局(即6-4)。如果没有,继续比赛,直到6比6打成平手。此时进行决胜局。在温网比赛中,抢七是先比7分(必须胜2分),但在5th局比赛中,先比10分(必须胜2分)除外。

- 休息时间/场边:球员在第1场比赛后,然后每两场比赛后更换场边。从第3场比赛开始,每次换边时允许有90秒的休息时间。在决胜局中,球员每6分换边。每盘结束后,球员也要休息至少2分钟。允许医疗暂停和一次洗手间休息。

引用:

- [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-alcaraz-b2376600.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-points-terms/7uzp2evdhbd11obdd59p3p1cx>.

3.5 帮助理解数据集的例子

Example 1: row 5

Column(s)	Value(s)	Description
<i>match_id</i>	"2023-wimbledon-1301"	The 3 in "1301" indicates a round 3 match and the "01" indicates the first match listed from that round.
<i>elapsed_time</i>	"0:01:31"	The point begins with a serve 1 minute and thirty-one seconds after the start of the first point of the match.
<i>point_no, game_no, set_no</i> (“no” is an abbreviation for number)	4, 1, 1	The point played is the 4 th point of the 1 st game of the 1 st set of the match.
<i>p1_sets, p2_sets, p1_games, p2_games</i>	0, 0, 0, 0	Since this is the first game of the match neither player has won a game or set yet.
<i>p1_score, p2_score</i>	15, 30	The score when the point is played is 15 (player 1), to 30 (player 2). Thus, player 1 won one of the previous points and player 2 won two points.
<i>server</i>	1	Player 1 (Alcaraz) is serving on this point.
<i>serve_no</i>	1	The point was played on the first serve meaning Alcaraz hit his first serve in play.
<i>point_victor</i>	1	Alcaraz wins this point (player 1).
<i>p1_points_won, p2_points_won</i>	2, 2	Player 1 (Alcaraz) is the point victor so his total is now 2 for the match (it was previously 1). For player 2 the value remains 2 since player 2 lost the point.
<i>game_victor, set_victor</i>	0, 0	Alcaraz winning the point makes the score in the game 30 – 30 (2 points each) so neither a game or set was won by either player on this point (both = 0).
Columns U – AC		Allow us to determine how the point was won:
<i>p1_winner</i>	1	Alcaraz won the point by hitting an “untouchable” shot.
<i>p1_ace</i>	0	The shot was not a serve (since = 0).
<i>winner_shot_type</i>	F	The shot was a forehand (as opposed to a backhand).
<i>p2_net_pt</i>	1	Player 2 (Jarry) positioned himself near the net somewhere during the point.
<i>p2_net_pt_won</i>	0	Since Alcaraz won the point, although Jarry was at the net during the point this value is 0.
Columns AH – AM		Even had player 2 won the point, the game would not have been over so the point was not a “break point” and these are all 0.
<i>p1_distance_run, p2_distance_run</i>	51.108, 75.631	The distance each player ran (in meters) on this point.
<i>rally_count</i>	13	Number of shots hit during the point by both players combined.
<i>speed_mph, serve_width, serve_depth, return_depth</i>	130, BW, CTL, D	Alcaraz (the server) hit a 130 serve “Body/Wide” of the returner (we saw it was a first serve previously) and close to the line denoting in or out of play. Jarry (the returner) returned the ball “Deep” in the court (so near the other end of the court).

Example 2: rows 8 – 12

The final four points of the first game illustrate the concept of tied score (“deuce”) and advantage (“ad”). Each row is a subsequent point in time in the match.

Row	Column(s)	Value(s)	Description
Row 8	<i>p1_score</i> , <i>p2_score</i>	40, 40	The score is 40 – 40 meaning each player has won 3 previous points (this is also called “deuce”).
	<i>point victor</i>	1	Alcaraz wins point 7 (in row 8).
Row 9	<i>p1_score</i> , <i>p2_score</i>	AD, 40	Since Alcaraz won the previous point (point 7) the score on point 8 is now “AD” for Alcaraz and “40” for Jarry meaning Alcaraz has won one more point and could win the game on the next point.
	<i>point victor</i>	2	Jarry (player 2) wins point 8 (in row 9).
Row 10	<i>p1_score</i> , <i>p2_score</i>	40, 40	The score returns to 40 – 40 (“deuce”) meaning each player has won the same number of previous points although now it is 4 points each.
	<i>point victor</i>	1	Alcaraz wins point 9 (in row 10).
Row 11	<i>p1_score</i> , <i>p2_score</i>	AD, 40	Alcaraz again has the advantage having won point 9.
	<i>point victor</i>	1	Alcaraz wins point 10 (in row 11) which means he has won the game (has score 2 more points now).
Row 12	<i>game_no</i>	2	This is now the first point of game 2.
	<i>p1_games</i>	1	Alcaraz won game 1.

Example 3: row 51

The 51st point of the match illustrates “break points” – points where the player not serving (the player who is returning serve) has an opportunity to win the game.

Row	Column(s)	Value(s)	Description
Row 51	<i>p1_score</i> , <i>p2_score</i>	40, 30	The score is 40 – 30 meaning player 1 (Alcaraz) is ahead.
	<i>server</i>	2	Jarry (player 2) is serving.
	<i>p1_break_pt</i>	1	If Alcaraz wins the point he will win the game; since he is not serving this is a “break point.”
	<i>point victor</i>	1	Alcaraz wins the point (and therefore the game).
	<i>p1_break_pt won</i>	1	Alcaraz won the game and was not serving on the point.

3.6 使用 AI 工具声明

在 COMAP 竞赛中使用大型语言模型和生成式 AI 工具

这一政策的动机是大型语言模型(LLMs)和生成 AI 辅助技术的兴起。该政策旨在为团队、顾问和评委提供更大的透明度和指导。这项政策适用于学生工作的各个方面，从模型的研究和开发(包括代码创建)到书面报告。由于这些新兴技术正在迅速发展,COMAP 将适当地完善这一策略。

团队必须公开和诚实地使用 AI 工具。一个团队及其提交的内容越透明，他们的工作就越有可能得到他人的充分信任、赞赏和正确使用。这些披露有助于理解智力工作的发展和对贡献的适当承认。如果没有对 AI 工具作用的公开和清晰的引用和参考，那么有问题的段落和工作更有可能被认定为抄袭并被取消资格。

解决这些问题不需要使用 AI 工具，尽管允许负责任地使用它们。COMAP 认识到 LLMs 和生成 AI 作为生产力工具的价值，可以帮助团队准备提交;例如，为一个结构产生初步的想法，或者在总结、释义、语言润色等时。在模型开发的许多任务中，人类的创造力和团队合作是必不可少的，对 AI 工具的依赖会带来风险。因此，我们建议在将这些技术用于模型选择和构建、协助创建代码、解释模型的数据和结果以及得出科学结论等任务时要谨慎。

值得注意的是，LLMs 和生成式 AI 有局限性，无法取代人类的创造力和批判性思维。COMAP 建议团队在选择使用 LLMs 时要意识到这些风险：

- 客观性:LLMs 生成的文本中可能出现先前发表的包含种族主义、性别歧视

或其他偏见的内容，一些重要观点可能未被代表。

·**准确性**:LLMs 可能会产生“幻觉”，即产生虚假内容，特别是在他们的领域之外使用或处理复杂或模棱两可的主题时。他们可以生成语言上但科学上不合理的内容，他们可以错误地获取事实，并且他们已经被证明可以生成不存在的引用。一些 LLMs 只接受特定日期之前发布的内容的培训，因此呈现的是不完整的画面。

·**语境理解**:LLMs 不能将人类的理解应用到一篇文章的语境中，特别是在处理习惯用语、讽刺、幽默或隐喻语言时。这可能会导致生成的内容出现错误或误解。

·**训练数据**:LLMs 需要大量高质量的训练数据来达到最佳性能。然而，在某些领域或语言中，这样的数据可能并不容易获得，从而限制了任何输出的有用性。

3.7 团队指南

参赛队伍需要:

3. 在**报告中明确指出使用了 LLMs 或其他 AI 工具**，包括使用了哪个模型以及用于什么目的。请使用内联引文和参考文献部分。在你的 25 页解决方案之后，还要附上 AI 使用报告(如下所述)。

2.**验证内容的准确性、有效性和适当性**以及由语言模型生成的任何引用，并纠正任何错误或不一致之处。

3.**提供引用和参考文献，遵循这里提供的指导**。仔细检查引文，以确保它们是准确的，并被正确引用。

4.**要注意抄袭的可能性**，因为 LLMs 可能会复制其他来源的大量文本。检查原始来源，以确保你没有抄袭别人的作品。

当 **COMAP** 确定提交的内容未公开使用了这些工具，我们会采取相应的行动。

3.8 引文与参考文献说明

仔细考虑如何记录和引用团队可能选择使用的任何工具。各种风格指南开始纳入引用和参考人工智能工具的政策。在你的 25 页解决方案的参考部分，使用内联引用并列所有使用的人工智能工具。

无论团队是否选择使用人工智能工具，主要解决方案报告仍然限制在 25 页。如果一个团队选择使用人工智能，在你的报告结束后，添加一个名为人工智能使用报告的新部分。这个新章节没有页数限制，不会被计入 25 页的解决方案中。

例子(这不是详尽的-根据你的情况调整这些例子):

人工智能使用报告

1. OpenAI ChatGPT (Nov 5, 2023 version, ChatGPT-4)Query1: <insert the exact wording you input into the AI tool>Output: <insert the complete output from the AI tool>

2. OpenAI Ernie (Nov 5, 2023 version, Ernie 4.0)

Query1: <insert the exact wording of any subsequent input into the AI tool>
Output: <insert the complete output from the second query>

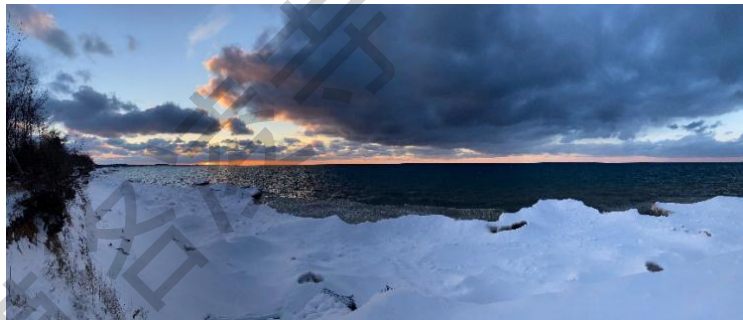
3. Github CoPilot (Feb 3, 2024 version)

Query1: <insert the exact wording you input into the AI tool>Output: <insert the complete output from the AI tool>

4. Google Bard (Feb 2, 2024 version) Query: <insert the exact wording of your query> Output: <insert the complete output from the AI tool>

4 D 题：五大湖的水问题

4.1 赛题翻译



美国和加拿大的五大湖是世界上最大的淡水湖群。这五个湖泊和相连的水道构成了一个巨大的流域，其中包含了这两个国家的许多大城市，气候和当地的天气条件各不相同。

湖区的水有多种用途(捕鱼、娱乐、发电、饮用、航运、动物和鱼类栖息地、建筑、灌溉等)。因此，各种各样的利益相关者都对流入和流出湖泊的水的管理感兴趣。特别是，如果从湖泊排出或蒸发的水太少，那么可能会发生洪水，沿岸的家庭和企业受到影响；如果排水过多，那么大型船只就无法通过水路运送补给，支持当地经济。主要问题是调节水位，使所有利益相关者都能受益。

每个湖泊的水位是由进出湖泊的水量决定的。这些水位是温度、风、潮汐、降水、蒸发、测深(湖底形状)、河流流量和径流、水库政策、季节周期和长期气候变化等复杂相互作用的结果。在五大湖系统的水流中有两种主要的控制机制：苏河水闸补偿工程。玛丽(三个水力发电厂，五个航行船闸和一个在激流顶端的闸门大坝)和康沃尔的摩西-桑德斯大坝，如附录所示。

虽然这两座控制水坝、许多渠道和运河以及流域水库可能是由人类控制的，但降雨、蒸发、侵蚀、冰塞和其他水流现象的速率是人类无法控制的。地方政府的政策可能会产生与预期不同的影响，流域的季节和环境变化也可能产生不同的影响。这些变化反过来又会影响该地区的生态系统，从而影响湖泊内外动植物的健康以及生活在水盆中的居民。尽管五大湖似乎有一个规律的年度模式，但水位从正常水平的 2 到 3 英尺的变化会极大地影响一些利益相关者。

这种动态的网络流量问题是“邪恶的”——由于相互依赖、复杂的要求和固有的不确定性，解决起来异常具有挑战性。对于湖泊的问题，我们有不断变化的动态和利益相关者的利益冲突。有关附加信息，请参阅问题 D 附录。

要求：

国际联合委员会(IJC)请求贵公司国际网络控制建模师(icm)提供支持，协助管理和建模直接影响五大湖水网水位的控制机制(附录中所示的两座水坝-补偿工程和摩西-桑德斯大坝)。你的 ICM 主管已经让你的团队领导开发模型和实施模型的管理计划。你的导师指出，有几个考虑因素可能有助于实现这一目标，首先是为五大湖建立一个网络模型，并将从苏必利尔湖到大西洋的河流连接起来。你的导师提到的其他一些可选的考虑因素或问题是：

- 考虑到各个利益相关者的愿望(每个利益相关者的成本和收益可能不同)，确定五大湖区在一年中任何时候的最佳水位。
- 根据五大湖的流入和流出数据，建立算法以维持五大湖的最佳水位。

- 了解您的控制算法对两个控制坝的流出的敏感性。考虑到 2017 年的数据，对于各利益相关者来说，你的新控制方法是否会使当年的实际记录水位令人满意或更好？
- 你们的算法对环境条件(例如，降水、冬季积雪、冰塞)的变化有多敏感？
- 将您的广泛分析集中在影响安大略湖的利益相关者和因素上，因为最近对该湖的水位管理有更多的关注。
- IJC 还对您使用什么历史数据来通知您的模型和建立参数感兴趣，因为他们很想比较您的管理和控制策略与以前的模型的比较。向 IJC 领导层提供一页备忘录，说明您的模型的关键特性，以说服他们选择您的模型。

4.2 赛题要求

准备一份不超过 25 页的报告，提供你的计划的细节。包括一份给希腊政府的两页报告备忘录，以帮助获得批准。

不超过 25 页的 PDF 解决方案应包括：

- 一页总结表。
- 目录表。
- 完整的解决方案。
- 一页备忘录。
- 参考书目。
- AI 使用报告(如果使用，不计入 25 页的限制)

注意:对于完整的 MCM 提交，没有特定的最低页数要求。您可以使用最多 25 页的总页数来完成所有解决方案工作和您想要包含的任何其他信息(例如:图纸，图表，计算，表格)。部分解决方案是可以接受的。我们允许谨慎地使用 AI，如 ChatGPT，尽管没有必要为这个问题创建一个解决方案。这将导致额外的 AI 使用报告，您必须将其添加到 PDF 解决方案文件的末尾，并且不计入解决方案的总页数限制。

4.3 参考资料

4.3.1 提供的文件

- 问题 D 附录-附加背景信息。
- 数据示例 - 这些是数据的可能来源。其中一些用于填充 **Problem_D_Great_Lakes.xlsx** 数据集。注:这些例子不是成功制定解决方案所必需的。
- **Problem_D_Great_Lakes.xlsx** -湖泊流入，流出和水位的数据。

4.3.2 参考资料(除了包含的背景数据文件)

1. Explanation of the IJC's Efforts to Manage the Great Lakes Basin: National Research Council; The Royal Society of Canada. (2006). Lawrence River Studies. Washington D.C.: National Research Council of the National Academies. Retrieved from <https://nap.nationalacademies.org/catalog/11481/review-of-the-lake-ontario-st-lawrence-river-studies>

2. Description of the Great Lakes Navigation Systems: Great Lakes Seaway NavigationSystem. (2023). Retrieved from American Great Lakes Ports Association:<https://www.greatlakesports.org/industry-overview/the-great-lakes-seaway-navigation-system/#:~:text=Lake%20Erie%20drains%20into%20Lake,in%20elevation%20approximately%20600%20feet>

4.4 使用 AI 工具声明

在 COMAP 竞赛中使用大型语言模型和生成式 AI 工具

这一政策的动机是大型语言模型(LLMs)和生成 AI 辅助技术的兴起。该政策旨在为团队、顾问和评委提供更大的透明度和指导。这项政策适用于学生工作的各个方面,从模型的研究和开发(包括代码创建)到书面报告。由于这些新兴技术正在迅速发展,COMAP 将适当地完善这一策略。

团队必须公开和诚实地使用 AI 工具。一个团队及其提交的内容越透明,他们的工作就越有可能得到他人的充分信任、赞赏和正确使用。这些披露有助于理解智力工作的发展和对贡献的适当承认。如果没有对 AI 工具作用的公开和清晰的引用和参考,那么有问题的段落和工作更有可能被认定为抄袭并被取消资格。

解决这些问题不需要使用 AI 工具,尽管允许负责任地使用它们。COMAP 认识到 LLMs 和生成 AI 作为生产力工具的价值,可以帮助团队准备提交;例如,为一个结构产生初步的想法,或者在总结、释义、语言润色等时。在模型开发的许多任务中,人类的创造力和团队合作是必不可少的,对 AI 工具的依赖会带来风险。因此,我们建议在将这些技术用于模型选择和构建、协助创建代码、解释模型的数据和结果以及得出科学结论等任务时要谨慎。

值得注意的是,LLMs 和生成式 AI 有局限性,无法取代人类的创造力和批判性思维。COMAP 建议团队在选择使用 LLMs 时要意识到这些风险:

- 客观性:LLMs 生成的文本中可能出现先前发表的包含种族主义、性别歧视或其他偏见的内容,一些重要观点可能未被代表。

- 准确性:LLMs 可能会产生“幻觉”,即产生虚假内容,特别是在他们的领域之外使用或处理复杂或模棱两可的主题时。他们可以生成语言上但科学上不合理的内容,他们可以错误地获取事实,并且他们已经被证明可以生成不存在的引用。一些 LLMs 只接受特定日期之前发布的内容的培训,因此呈现的是不完整的画面。

- 语境理解:LLMs 不能将人类的理解应用到一篇文章的语境中,特别是在处理习惯用语、讽刺、幽默或隐喻语言时。这可能会导致生成的内容出现错误或误解。

- 训练数据:LLMs 需要大量高质量的训练数据来达到最佳性能。然而,在某些领域或语言中,这样的数据可能并不容易获得,从而限制了任何输出的有用性。

4.5 团队指南

参赛队伍需要:

4. 在报告中明确指出使用了 LLMs 或其他 AI 工具,包括使用了哪个模型以及用于什么目的。请使用内联引文和参考文献部分。在你的 25 页解决方案之后,还要附上 AI 使用报告(如下所述)。

- 2.验证内容的准确性、有效性和适当性以及由语言模型生成的任何引用,并纠正任何错误或不一致之处。

- 3.提供引用和参考文献,遵循这里提供的指导。仔细检查引文,以确保它们

是准确的，并被正确引用。

4.要注意抄袭的可能性，因为 LLMs 可能会复制其他来源的大量文本。检查原始来源，以确保你没有抄袭别人的作品。

当 COMAP 确定提交的内容未公开使用了这些工具，我们会采取相应的行动。

4.6 引文与参考文献说明

仔细考虑如何记录和引用团队可能选择使用的任何工具。各种风格指南开始纳入引用和参考人工智能工具的政策。在你的 25 页解决方案的参考部分，使用内联引用并列所有使用的人工智能工具。

无论团队是否选择使用人工智能工具，主要解决方案报告仍然限制在 25 页。如果一个团队选择使用人工智能，在你的报告结束后，添加一个名为人工智能使用报告的新部分。这个新章节没有页数限制，不会被计入 25 页的解决方案中。

例子(这不是详尽的-根据你的情况调整这些例子):

人工智能使用报告

1. OpenAI ChatGPT (Nov 5, 2023 version, ChatGPT-4)Query1: <insert the exact wording you input into the AI tool>Output: <insert the complete output from the AI tool>

2. OpenAI Ernie (Nov 5, 2023 version, Ernie 4.0)

Query1: <insert the exact wording of any subsequent input into the AI tool>

Output: <insert the complete output from the second query>

3. Github CoPilot (Feb 3, 2024 version)

Query1: <insert the exact wording you input into the AI tool>Output: <insert the complete output from the AI tool>

4. Google Bard (Feb 2, 2024 version) Query: <insert the exact

wording of your query> Output: <insert the complete output from the AI tool>

5 E 题：财产保险的可持续性

5.1 赛题翻译

极端天气事件正在成为业主和保险公司的危机。“近年来，全球因 1000 多起极端天气事件而遭受了超过 1 万亿美元的损失。”与 30 年的平均水平相比，保险业在 2022 年的自然灾害索赔增加了 115%。“由于洪水、飓风、气旋、干旱和野火等与天气有关的严重事件造成的损失可能会增加，预计未来状况将变得更糟。保险费用正在迅速上涨，预计到 2040 年，气候变化将推动保费增长 30-60%[1]。

随着保险公司改变**承保方式（how）**和**承保地点（where）**，财产保险不仅变得更加昂贵，而且更难找到。在不同的地区，与天气有关的事件会推高财产保险的保费成本。此外，**全球保险保障缺口**平均为 57%，而且还在不断扩大[2]。这凸显了保险业的困境——保险公司的盈利能力和业主的负担能力正在出现危机。

- COMAP 的灾害保险模型师（Insurance of Catastrophes Modelers **ICM**）对财产保险行业的可持续性很感兴趣。由于气候变化增加了更严重的天气和自然灾害的可能性，**ICM** 希望确定如何**最优配置财产保险**，以使系统具有弹性，以覆盖未来索赔的成本，同时确保保险公司的**长期健康**。如果保险公司在很多情况下不愿意承保，他们就会因为客户太少而倒闭。相反，如果他们承保的保单风险太大，他们可能会支付过多的索赔。**保险公司应在什么情况下承保？什么时候他们应该选择冒险？业主能做些什么来影响这个决定吗？**为保险公司**建立一个模型**，以确定他们是否应该在极端天气事件不断增加的地区承保。在经历过极端天气事件的两个大洲上（两个不同地区）演示你们的模型。
- 展望未来，社区和房地产开发商需要考虑**如何建设和在哪里发展**。随着保险业格局的变化，未来的房地产决策必须确保房地产更具弹性，并精心建造，包括为**不断增长的社区和人口**提供适当服务的可行性。**如何调整**你的保险模型，以评估在哪里、如何以及是否在某些站点上建立？
- 在某些社区，你的保险模式可能不建议承保当前或未来的财产保险政策。这可能导致社区领导人面对**具有文化或社区意义的财产**的艰难决定。例如，为了保护这座历史悠久的灯塔以及以它为中心的当地旅游业，哈特拉斯角灯塔被移到了北卡罗来纳州的外岸[3]。作为社区领导人，您如何确定社区中的建筑应因其**文化、历史、经济或社区意义**而被保存和保护？为社区领导人制定一个保护模式，以确定他们应该采取的措施的范围，以保护其社区内的建筑。
- 选择一个历史地标，不能是哈特拉斯角灯塔，因为它在一个经历极端天气事件的地方。运用你的保险和保护模型来**评估这个地标的价值**。考虑到您从您的保险和保存模型的结果中获得的见解，撰写一页给社区的信，推荐一个计划、时间表和成本提案，为他们珍视的地标的未来提供建议。

5.2 赛题要求

不超过 25 页的 PDF 解决方案应该包括：

- 一页的摘要，清楚地描述了你解决问题的方法，以及从问题的背景中分析出的最重要的结论
- 目录

- 完整的解决方案。
- 一页社区信
- AI 使用报告(如果使用)

注意:对于一个完整的 ICM 提交,没有明确要求的最小页面长度。您可以使用多达 25 页的全部解决方案工作和任何额外的信息,您希望包括(例如:图纸,图表,计算,表)。部分解决方案是可以接受的。我们允许谨慎地使用 AI,如 ChatGPT,尽管没有必要创建这个问题的解决方案。如果你选择使用生成式 AI,你必须遵循 COMAP AI 的使用策略。这将导致一个额外的 AI 使用报告,你必须添加到你的 PDF 解决方案文件的结尾,不计入总 25 页的限制为你的解决方案

5.3 词汇表

保险保障缺口(Insurance Protection Gap):自然灾害造成的经济损失与所承保的经济损失之间的保护范围差异。

承保(Underwrite):在发生损失或损坏时承担赔偿责任,从而保证赔偿。

6 F 题：减少非法野生动物贸易

6.1 赛题翻译

非法野生动物贸易对我们的环境产生负面影响，并威胁到全球生物多样性。据估计，每年涉及的金额高达 265 亿美元，被认为是全球第四大非法贸易[1]。你将开发一个数据驱动的 5 年项目，旨在显著减少非法野生动物贸易。你的目标是说服客户执行你的项目。要做到这一点，你必须选择一个客户和一个合适的项目。

你的工作应该探索以下问题：

● 谁是你的客户？这个客户实际能做什么？（换句话说，你的客户应该拥有实施你提议的项目所需的权力、资源和兴趣。）

● 解释为什么你开发的项目适合这个客户。从已发表的文献和你自己的分析中，有哪些研究支持你选择这个项目？使用数据驱动的分析，你将如何说服你的客户，这是一个他们应该承担的项目？

● 你的客户需要哪些额外的权力和资源来执行这个项目？（记住要使用假设，但也要尽可能将你的工作建立在现实基础上。）

● 如果项目实施了，会发生什么？换句话说，对非法野生动物贸易的可衡量影响是什么？你做了什么分析来确定这一点？

● 项目达到预期目标的可能性有多大？此外，根据情境化的敏感性分析，是否存在可能不成比例地帮助或损害项目实现目标的能力的条件或事件？

虽然您可以限制您对非法野生动物贸易的方法，但您也可以将非法野生动物贸易视为更大的复杂系统的一部分。具体来说，你可以考虑其他领域的其他全球努力，例如，减少其他形式的贩运的努力或减少气候变化的努力以及减少非法野生动物贸易的努力，如何成为一个复杂系统的一部分。这可能会为这一领域中意想不到的行动者创造协同机会。

如果您选择在解决方案中利用复杂性框架，请务必通过讨论此建模决策的优点和缺点来证明您的选择是正确的。

此外，您的团队必须提交一份 1 页的备忘录，其中包含您的客户的关键点，突出您的 5 年项目提案以及为什么该项目适合他们作为客户（例如，访问资源，他们任务的一部分，与他们的使命声明保持一致等）。

评委们将特别寻找客户选择的创造力，以及在整个分析过程中使用的适当建模过程的选择和论证。他们还将寻找阐述 (1) 在客户和拟议项目之间建立牢固的联系，(2) 在数据分析和拟议项目的设计之间建立清晰而直接的联系。

6.2 赛题要求

准备一份不超过 25 页的报告，提供你的计划的细节。包括一份给希腊政府的两页报告备忘录，以帮助获得批准。

不超过 25 页的 PDF 解决方案应包括：

- 一页总结表。
- 目录表。
- 完整的解决方案。
- 一页备忘录。
- 参考书目。

- AI 使用报告(如果使用, 不计入 25 页的限制)

注意:对于完整的 MCM 提交, 没有特定的最低页数要求。您可以使用最多 25 页的总页数来完成所有解决方案工作和您想要包含的任何其他信息(例如:图纸, 图表, 计算, 表格)。部分解决方案是可以接受的。我们允许谨慎地使用 AI, 如 ChatGPT, 尽管没有必要为这个问题创建一个解决方案。这将导致额外的 AI 使用报告, 您必须将其添加到 PDF 解决方案文件的末尾, 并且不计入解决方案的总页数限制。

6.3 参考文献

摘自 <https://wildlifetrade.wcs.org/Wildlife-Trade/Why-should-we-care.aspx>

6.4 术语表

客户:将要实施拟议项目的演员。他们可能是官方行为者(政府或准政府)或非官方行为者(非政府组织)。

非法野生动物贸易:走私、偷猎、捕获或收集濒危物种、受保护野生动物或这些物种的衍生物/产品

6.5 使用 AI 工具声明

在 COMAP 竞赛中使用大型语言模型和生成式 AI 工具

这一政策的动机是大型语言模型(LLMs)和生成 AI 辅助技术的兴起。该政策旨在为团队、顾问和评委提供更大的透明度和指导。这项政策适用于学生工作的各个方面, 从模型的研究和开发(包括代码创建)到书面报告。由于这些新兴技术正在迅速发展, COMAP 将适当地完善这一策略。

团队必须公开和诚实地使用 AI 工具。一个团队及其提交的内容越透明, 他们的工作就越有可能得到他人的充分信任、赞赏和正确使用。这些披露有助于理解智力工作的发展和对贡献的适当承认。如果没有对 AI 工具作用的公开和清晰的引用和参考, 那么有问题的段落和工作更有可能被认定为抄袭并被取消资格。

解决这些问题不需要使用 AI 工具, 尽管允许负责任地使用它们。COMAP 认识到 LLMs 和生成 AI 作为生产力工具的价值, 可以帮助团队准备提交;例如, 为一个结构产生初步的想法, 或者在总结、释义、语言润色等时。在模型开发的许多任务中, 人类的创造力和团队合作是必不可少的, 对 AI 工具的依赖会带来风险。因此, 我们建议在将这些技术用于模型选择和构建、协助创建代码、解释模型的数据和结果以及得出科学结论等任务时要谨慎。

值得注意的是, LLMs 和生成式 AI 有局限性, 无法取代人类的创造力和批判性思维。COMAP 建议团队在选择使用 LLMs 时要意识到这些风险:

- 客观性:LLMs 生成的文本中可能出现先前发表的包含种族主义、性别歧视或其他偏见的内容, 一些重要观点可能未被代表。

- 准确性:LLMs 可能会产生“幻觉”, 即产生虚假内容, 特别是在他们的领域之外使用或处理复杂或模棱两可的主题时。他们可以生成语言上但科学上不合理的内容, 他们可以错误地获取事实, 并且他们已经被证明可以生成不存在的引用。一些 LLMs 只接受特定日期之前发布的内容的培训, 因此呈现的是不完整的画面。

- 语境理解:LLMs 不能将人类的理解应用到一篇文章的语境中, 特别是在处理习惯用语、讽刺、幽默或隐喻语言时。这可能会导致生成的内容出现错误或误

解。

·训练数据:LLMs 需要大量高质量的训练数据来达到最佳性能。然而,在某些领域或语言中,这样的数据可能并不容易获得,从而限制了任何输出的有用性。

6.6 团队指南

参赛队伍需要:

5. 在**报告中明确指出使用了 LLMs 或其他 AI 工具**,包括使用了哪个模型以及用于什么目的。请使用内联引文和参考文献部分。在你的 25 页解决方案之后,还要附上 AI 使用报告(如下所述)。

2.**验证内容的准确性、有效性和适当性**以及由语言模型生成的任何引用,并纠正任何错误或不一致之处。

3.**提供引用和参考文献,遵循这里提供的指导**。仔细检查引文,以确保它们是准确的,并被正确引用。

4.**要注意抄袭的可能性**,因为 LLMs 可能会复制其他来源的大量文本。检查原始来源,以确保你没有抄袭别人的作品。

当 COMAP 确定提交的内容未公开使用了这些工具,我们会采取相应的行动。

6.7 引文与参考文献说明

仔细考虑如何记录和引用团队可能选择使用的任何工具。各种风格指南开始纳入引用和参考人工智能工具的政策。在你的 25 页解决方案的参考部分,使用内联引用并列出生所有使用的人工智能工具。

无论团队是否选择使用人工智能工具,主要解决方案报告仍然限制在 25 页。如果一个团队选择使用人工智能,在你的报告结束后,添加一个名为人工智能使用报告的新部分。这个新章节没有页数限制,不会被计入 25 页的解决方案中。

例子(这不是详尽的-根据你的情况调整这些例子):

人工智能使用报告

1. OpenAI ChatGPT (Nov 5, 2023 version, ChatGPT-4)Query1: <insert the exact wording you input into the AI tool>Output: <insert the complete output from the AI tool>
2. OpenAI Ernie (Nov 5, 2023 version, Ernie 4.0)
Query1: <insert the exact wording of any subsequent input into the AI tool>
Output: <insert the complete output from the second query>
3. Github CoPilot (Feb 3, 2024 version)
Query1: <insert the exact wording you input into the AI tool>Output: <insert the complete output from the AI tool>
4. Google Bard (Feb 2, 2024 version) Query: <insert the exact wording of your query> Output: <insert the complete output from the AI tool>