摘要

奥运会是一项全球性的体育活动，各国在这个盛大的比赛中综合较量其体育实力。对奥运会.因此奖牌榜的预测一直是体育研究领域和大众关注的热点之一.奥运会奖牌榜结果会受到诸多因素的影响,.著名的教练对国家获奖情况的影响也不容忽视,预测是一项具有挑战性的工作,为此我们建立了两个模型来完成任务,LightGBM 模型和多元线性回归模型.

对于任务1，我们使用了LightGBM 模型。首先我们考虑了NOC，Sport,Event，athletes\_number,Year,Host,这几个特征对预测结果的影响。把题目提供的数据进行相应的处理,并计算每个特征与各个奖牌数量的相关性系数.筛选了这些特征对模型进行训练,并计算出对金牌数,银牌数,铜牌数,总牌数的R^2分别为0.9885,0.9872,0.9804,0.9916.然后根据2028年奥运会项目的调整对模型输入数据进行调整以及东道主国家情况,使用模型预测出奖牌榜。接着我们设定置信区间为60%,计算出了预测区间的上下限。我们对不同奖牌赋予分数，计算国家2024年和2028年奖牌总得分,作差比较得出进退步国家。我们根据2028年预测结果统计可能获得首枚奖牌的国家，再根据预测奖牌数量分区估计胜算概率。最后我们使用了我们使用AHP 层次分析法对金牌数，银牌数，铜牌数计算权重。根据国家在某个项目上的奖牌数量计算加权得分，分析不同国家最为重要的体育项目。

对于任务2，我们使用了多元线性回归模型。首先我们根据找到的数据对筛选出可能出现”great coach”的国家和对应Event,然后计算这些国家奖牌的得分,考虑教练执教的滞后性加入了短期的（1 - 2 届）coach\_lag1-2,和长期的(3 - 5 届）coach\_lag3-5作为特征,一共选取了10个特征,使用多元线性回归模型训练了短期模型和长期模型.分别计算短期和长期收到 “名帅效应”国家对应Event的得分变化,得出“名帅效应”对国家奖牌得分的影响。计算得模型=0.828.最后我们选择了美国，日本和英国三个国家，通过对某个项目上历史平均得分和近三届奥运会得分的比较,制定了一个标准,得出在哪些体育项目上建议聘请名帅。

对于任务3，我们从两个模型中挖掘与奥运会奖牌榜相关的独特发现，得出奖牌分布与项目参与度，奖牌榜与经济发展水平的关系，奖牌背后的训练体系与科研支持这三条见解，并进行解释，给出奥委会参考建议。

最后，对我们对LightGBM 模型year和total\_event两个参数进行了敏感性分析，这些参数敏感性高有助于捕捉细微变化对奖牌数量的潜在影响，有助于更准确的预测。

此外，我们评估了该模型的优缺点。