| **评估名称** | **课程作业-数据科学研究** | **权重** | **75%** |
| --- | --- | --- | --- |
| 描述与提交物 | 本作业要求3人小组合作完成。需运用所学数据科学步骤分析数据集，构建并比较训练模型。需以联合学术论文形式撰写研究（含参考文献和图表共6-8页），使用IEEE模板排版为PDF提交。代码需以单个Jupyter Notebook提交，并注明各成员贡献。同时需通过Moodle单独提交对组员的同行评估。 |  |  |
| 发布时间 | 2025年2月12日（星期三） |  |  |
| 提交截止时间 | 2025年5月2日（星期五）15:00 |  |  |
| 迟交政策 | 诺丁汉大学默认政策适用：每逾期1个工作日扣5分（总分100分）。小组项目需整体按时提交，特殊情况学生可提交修订版说明个人贡献。最晚迟交日期为2024年5月9日15:00，此后仅接受特殊情况申请。 |  |  |
| 反馈机制 | 通过Moodle在15个工作日内反馈 |  |  |

2. 任务说明

小组需完成完整数据科学项目，包括：  
(1) 选择数据集  
(2) 确定研究问题  
(3) 构建机器学习模型  
(4) 设计验证实验  
(5) 结果分析  
(6) 按IEEE格式撰写6-8页学术论文

需提交：  
(a) 联合学术论文（PDF）  
(b) 可复现的Jupyter Notebook代码  
(c) 个人反思文档（含300字项目说明、贡献描述、生成式AI使用说明及同行评估）

2.1 主报告结构

标题

摘要

数据集与研究问题介绍（含探索性分析）

文献综述

方法论（含方法依据）

结果（数据分析、预处理与预测）

讨论（与文献对比）

结论与建议

参考文献

贡献声明（使用CRediT分类）

2.2 代码要求

需注明成员分工，确保在A32实验室/Virtual Desktop环境可复现结果。需完整展示数据科学全流程。

2.3 个人反思

包含：  
● 项目过程描述（含未成功尝试）  
● 个人贡献说明（附证据）  
● 生成式AI使用声明  
● 组员评估（努力程度、会议参与度、团队精神）

3. 评分标准

报告与代码（95%）

| **部分** | **权重** | **关键评估点** |
| --- | --- | --- |
| 项目介绍 | 15% | 标题/摘要准确性，数据集统计描述与研究问题相关性 |
| 文献综述 | 10% | 文献覆盖广度与关键研究识别 |
| 方法论 | 20% | 方法合理性及证据支持 |
| 结果与分析 | 25% | 结果展示清晰度与深度分析 |
| 讨论与结论 | 20% | 结果解释的洞察力与建议可行性 |
| 格式与呈现 | 5% | IEEE规范与图表质量 |
| 代码质量 | 10% | 可复现性、注释完整性与流程完整性 |

个人反思（5%）  
完整覆盖四项要求即可满分。

小组完成数据科学全流程项目（数据选择→模型构建→实验验证→论文撰写）

提交：IEEE格式论文（6-8页）、可复现代码、个人反思文档

发布时间：2025.2.12

提交截止：2025.5.2 15:00

最晚迟交：2024.5.9 15:00（特殊情况需申请）

论文结构要求: 需包含10个标准部分，重点关注：

方法论需提供方法选择的证据支持

讨论部分需对比已有文献

必须使用CRediT声明贡献分工

代码规范

单文件Jupyter Notebook

明确标注成员分工代码段

确保在指定计算环境中可复现

​个人反思重点

需说明生成式AI的使用场景（如使用需声明）

组员评估需包含努力程度、会议参与等维度

需附实证材料说明个人贡献

​评分权重

结果与分析部分占比最高（25%）

方法论和讨论共占40%

代码质量占10%（侧重可复现性）

格式规范占5%（严格遵循IEEE模板）

数据探索的深度、方法选择的合理性、结果与文献的对比分析，并提前测试代码运行环境。