

个人贡献声明

小组成员：22336216 陶宇卓、22336155 刘苇嘉、22336210 孙泽堃

为保障组内合作的高效与公平，我们在项目初期明确划分了任务分工，并在过程中通过 GitHub 版本控制和定期沟通协同完成各项工作。以下为每位成员的具体贡献说明：

陶宇卓

主要任务：

- 搭建组内 GitHub 仓库，制定协作规范；
- 数据预处理流程设计与实现；
- 随机森林与支持向量机模型的实现和调优；
- 设计 SHAP 分析特征重要性的相关函数；
- 实验报告中“第五部分：实验结果”和“第七部分：未来发展思考”的撰写。

具体贡献：

项目伊始，我主动承担了组内协作管理工作，创建了 GitHub 仓库并编写了初始版 README，以便组员高效同步代码。在数据处理阶段，我参考数据集仓库中的说明文档，制定了数据清洗与特征构造的流程，为后续建模奠定了基础。

在建模阶段，我负责实现随机森林与支持向量机两种模型，设计了统一的模型接口、输出日志规范和评价标准。在撰写实验报告时，我撰写了实验结果部分，结合 SHAP 分析、性能对比和混淆矩阵进行可视化展示，并进一步提出未来可行的优化方向。

挑战与解决方案：

初期我对原始数据的组织结构较为困惑，通过认真阅读开源仓库的说明并多次测试预处理流程，逐步厘清了字段间的逻辑关系。此外，在进行 SHAP 分析时曾多次遇到运行时错误，我主动查阅官方文档和社区问答，深入理解 SHAP 原理，最终成功解决了问题并生成可解释性图示。

心得体会：

本次项目锻炼了我在项目管理、模型实现和可解释性分析等方面的能力，也让我深刻体会到团队合作的力量。通过主动承担任务和跨模块配合，我不仅提升了技术能力，也增强了问题定位与解决的

能力。

刘苇嘉

主要任务：

- 神经网络与决策树模型的实现与调优；
- 撰写实验报告“第三部分：算法原理”与“第六部分：对比分析与结论”。

具体贡献：

我主要负责实现神经网络和决策树两种模型，并对比其在本任务中的适用性与性能表现。在神经网络部分，我基于标准库设计了模型架构，并通过调参提升其泛化能力。决策树模型中，我尝试使用不同分裂准则与深度限制，以防止过拟合。

在实验报告中，我系统性地总结了各算法的基本原理，并从准确性、鲁棒性、可解释性等多个维度对模型进行了深入对比，结合性能指标与 SHAP 分析结果撰写了结论部分。

挑战与解决方案：

无

心得体会：

通过本次项目，我深化了模式识别的理解，并提升了数据驱动分析与报告撰写的能力。与队友良好的沟通协作也让我受益匪浅，是一次非常有收获的团队实践经历。

孙泽堃

主要任务：

- 逻辑回归与 Kmeans 的实现与调优；
- 设计网格化调参相关代码；
- 设计 Compare.py 以进行系统化对比分析；
- 撰写实验报告“第一部分：算法原理”、“第二部分：数据集介绍”和“第四部分：实验设置及评估标准”。

具体贡献：

我负责实现逻辑回归与 KMeans 两种模型。逻辑回归部分注重于正则化策略与类别不平衡的处理；KMeans 聚类用于作为无监督对照方法，协助了解数据分布结构。

为了提升模型性能与实验效率，我设计了 GridSearch 自动调参的相关代码以便复用。在模型评估与对比阶段，我独立设计了 `Compare.py` 脚本，整合各模型的输出指标与可视化，使报告撰写更为高效直观。

此外，我还负责编写报告中的算法原理、数据集介绍和实验设置部分，从理论、数据背景及实验流程三方面为读者建立整体理解框架。

挑战与解决方案：

无

心得体会：

此次项目让我深刻体会到理论与实践的结合。通过负责基础模型与调参模块的实现，我对模式识别有了更系统的理解。同时，撰写报告前几部分的过程也加深了我对数据预处理、评估指标等基础环节重要性的认识。团队间分工明确、配合顺畅，使我们能够高效完成任务，是一次非常有价值的学习与合作体验。