

罗杰

电话/微信: 17823896729 邮箱: luojie16@ctbu.edu.cn

地址: 重庆市南岸区

重庆工商大学 - 大数据管理与应用专业 | 本科

2022年9月 - 2026年6月

核心课程: 数据库原理、Hadoop技术栈、Python编程、机器学习、统计学、Tableau可视化分析



专业技能

SQL Hive Spark Python 数据仓库 Tableau

- 熟练掌握SQL，能够编写复杂查询和多表关联，熟悉Hive/GBase等数据库的使用和优化
- 熟练使用Python进行数据处理分析，掌握Pandas、NumPy、Scikit-learn等库在机器学习项目中的应用
- 具备数据仓库分层建模实践经验，理解数据清洗、特征工程在真实业务场景中的实施
- 能够使用Tableau、Power BI、Streamlit等工具进行数据可视化分析和看板开发

项目经验

艾瑞数智 - 数据开发实习

2025年5月 - 2025年10月

- 负责重庆移动存量运营数据处理，涵盖用户属性、业务类型等多维度数据，支撑业务分析需求
- 编写复杂SQL语句，通过多表关联、分组聚合等操作提取业务指标，遵循数据分区规范优化查询性能
- 基于数据中台进行模型分层开发，主要在DW层和ST层构建数据处理流程，确保标准化和可维护性
- 配置数据交换任务，实现跨数据库表同步和数据分发，支撑前端标签系统和业务看板
- 参与用户画像标签体系建设，持续更新和维护用户特征维度

信用卡欺诈检测系统

2025年3月

- 基于高度不平衡的信用卡交易数据(欺诈样本0.172%)，构建欺诈检测机器学习模型
- 实施特征工程：重构时间特征，对交易金额进行对数变换和标准化处理，提升特征判别能力
- 设计Stacking集成学习框架，融合多种树模型优势，通过五折交叉验证确保模型稳健性
- 应用SMOTE过采样技术解决类别不平衡问题，模型达到AUC 0.9768，召回率81.63%
- 完成模型工程化准备，实现预处理流水线和预测模型的持久化，支持灵活阈值调节

评论情感分析实时系统

2025年4月

- 基于Kafka+Redis构建实时数据处理流水线，实现评论数据从采集、处理到可视化的全链路覆盖
- 使用Kafka实现系统组件解耦，通过Redis存储实时统计结果，优化系统查询性能
- 采用Docker容器化部署大数据组件，解决环境依赖问题，提升部署效率

- ▶ 基于Streamlit开发交互式可视化看板，支持实时情感分析结果展示和用户反馈收集

个人总结

具备扎实的数据开发与机器学习技能，拥有数据仓库分层开发、用户画像构建、数据可视化看板开发经验。对数据与业务的关联有深入思考，善于通过数据驱动业务决策。对数据工作充满热情，期待在数据开发领域持续深耕，创造业务价值。