

张炜浩

13820952112 | zhang.weiha098@gmail.com

教育经历

- 圣路易斯华盛顿大学** 海外QS前100 2021年08月 - 2022年12月
信息系统管理 硕士 圣路易斯
GPA: 3.83/4.0
- 美国大学** 海外QS前100 2017年08月 - 2021年05月
计算机科学 本科 华盛顿特区

实习经历

- PolicyEngine** 2023年03月 - 2023年12月
数据分析师-模型开发 华盛顿特区
- 对亚利桑那州和肯塔基州的法规进行了彻底的研究和分析，使用 PolicyEngine API、Python 开发了一个用于模拟税收优惠计算的基线模型，并使用专为税收优惠系统定制的内部构建的 Python 包将该模型定制为独特的政府数据集。
 - 对政府文件进行全面研究，以了解税收优惠规则和法规并使用YAML 严格验证模型功能，并经常监控参数，确保持续与最新政策保持一致，提高模型的可靠性和准确性
 - 通过 Git 版本控制与其他贡献者合作开发开源软件，使个人、政策制定者和组织能够作出明智的决策，从而增加每月的活跃用户。
 - 审核同事的每周拉取请求，解决格式问题，识别逻辑和语法错误，并提供建设性反馈以提高代码质量。

- 中邮信科** 2022年10月 - 2022年12月
数据分析实习生 北京

- 项目名称：《邮储银行风险客户识别模型构建》
- 利用邮储银行提供的部分客户交易流水数据、相关客户信息，结合优化升级版的评分卡模型框架，建立风险客户识别模型，在评分置信度阈值 0.55 的基础上整体识别准确率为 88.6%。
- 本人负责业务逻辑知识梳理、整体模型设计开发、版本迭代优化工作技术要点：
 - 数据层：整合邮储银行部分客户交易流水数据库和客户静态信息库，确保全面获取相关数据。
 - 特征工程：创新性地对客户身份信息编码特征、对方交易账户信息编码特征、交易流水统计信息特征以及交易时序统计特征的处理，以提高模型对风险客户的敏感度。
 - 模型层：采用特征WOE编码，结合特征组合筛选机制，以提高模型的稳定性和泛化能力。基于LightGBM构建风险客户评分模型，充分发挥其在大规模数据集上的高效性和准确性。

- RallyCry Ventures** 2020年08月 - 2020年12月
风险投资数据分析师实习生 波士顿

- 在产品投资团队中，全面负责调研项目加速器、孵化器、大学生创业项目以及其他创业活动。通过对比行业调研、商业模式、市场准入标准等投资基准，成功创建了40多种创业投资指标。
- 运用SQL和Python进行数据清理和异常数据整合，产出可视化报告，为投资团队提供决策支持。
- 深入分析行业投资潜力和发展趋势，为投资团队尽职调查环节提供数据支持。
- 独立编写投资调查报告，通过行业规模分析，精选最具投资价值的客户。

项目经历

- 房屋销售预测和分类模型选择** 2022年08月 - 2022年12月
组长 圣路易斯
- 领导并参与房屋销售预测和分类模型的选择项目，通过对最近售出的物业价格区间和地点的收集，结合Excel评估，确定和传播各种复杂数据集中的趋势或模式。利用精确率和F1值测试六种不同的模型，并代入完成预测，最终选择最优模型。
 - 运用数据挖掘技术，发现数据中的关键特征和潜在模式，为后续模型选择提供基础。
 - 评估和测试六种不同的预测模型，考察其在房屋销售预测和分类任务中的性能表现。

- US CDC死亡率分析** 2021年08月 - 2021年12月
组员 圣路易斯

- 使用python 解析CDC 从 2005 年到 2015 年的数据，开发基于逻辑回归模型的预测模型进行数据训练，生成完全符合逻辑回归模型的预测以及用于数据训练的广义线性模型进行管理控制。
- 创建多个表和视图，同时通过使用 Python 准确描述视图。及时有效地识别并解决了与数据流入数据库的完整性有关的复杂问题。
- 成果：通过数据可视化，得到影响死亡率最高的 10个因素，并根据结论提出预防措施和未来的研究目标。

技能/证书及其他

- 技能**：数据分析，Office，PowerPoint，Excel，Word，Python，SQL，R，Tableau，前端HTML
- 语言**：普通话（母语），英语（流利），日语（初级）