

黄力恒

liheng.huang@outlook.com | 广州 | 个人主页



教育背景

香港城市大学 生物统计学硕士	2025.09 – 2026.10
GPA: 3.83/4.3	
华东交通大学 数据科学与大数据技术学士	2020.09 – 2024.07
平均分: 84.3/100; 优秀毕业论文奖 (学院前 1%)	

工作经历

康景堂有限公司 实习生 (业务系统开发与报表分析)	2026.05 – 2026.08
- 基于开源系统定制诊所收支管理工具, 关联患者、服务项目、收支明细与次卡账户, 构建按日、月、年及服务项目汇总的经营报表, 支持账务核对。 - 设计次卡发放、核销与余额同步规则, 处理重复导入及同名客户识别; 实现导入前备份、撤回与调整原因记录, 支持数据恢复和账务变更追溯。	
深圳前海新心数字科技有限公司 数据分析实习生	2024.01 – 2024.06
- 使用 PostgreSQL 编写客户、交易及信贷数据的多表查询, 整理业务分析数据并生成统计报表。 - 使用逻辑回归、XGBoost 及 WOE/IV 方法参与客户特征分析与优先级排序, 为短信营销客群筛选提供支持。	

项目经历

金融业务指标核验与自动化报表 [GitHub] [看板]	2026.07 – 2026.09
- 基于合成信贷数据, 使用 Python、PostgreSQL 定义审批率、放款额、贷款余额及逾期指标, 区分期间发生额与时点余额; 通过 Streamlit 展示信贷漏斗、贷款批次及渠道表现。 - 处理重复事件、延迟到达与历史更正, 建立明细与汇总对账, 完成 248 个报告日期的处理与核验; 既有验收包含 19 项测试, 并验证重复发布一致性和异常事务回滚。	
预算约束下的营销策略建模与评价 [GitHub]	2026.07 – 2026.09
- 围绕 30% 短信发送预算, 分析约 20 万名客户的历史数据, 构建 RFM、消费趋势、积分使用等 53 项特征, 区分客户购买倾向与短信带来的增量购买潜力。 - 设计正则化 DR-learner, 通过限制树复杂度控制过拟合, 按正增益排序分配发送预算; 比较八个候选策略, 依据验证集 AIPW 价值选定 C1, 分析发送名单重合与客户画像差异。 - 在约 4 万名客户的历史评价样本中, 相对同预算随机发送和响应预测, 分别估计每千名客户增加 4.89 和 7.17 名购买客户。	
纵向临床数据与生存风险分析 [GitHub]	2025.04 – 2025.08
- 基于 5,368 名乳腺癌患者的纵向记录, 使用 SQLite 整合诊断、治疗与疾病进展事件, 构建 Landmark 样本, 采用患者级划分及训练侧预处理控制信息泄漏。 - 结合 Transformer 与 Cox 生存损失估计不同随访时点的后续风险, 使用 SHAP 和事件扰动分析模型对疾病进展信息的敏感性, 区分预测关联与因果效应。	

更多项目请见我的个人主页

技能

数据分析: Python、SQL、R、PostgreSQL、Excel、多表查询、数据清洗
统计与实验: 回归分析、假设检验、抽样与区间估计、模型评价
建模与可视化: scikit-learn、XGBoost、特征工程、SHAP、Streamlit、Git
语言能力: 普通话、粤语、英语 (雅思 7.0、英语六级)