

林嘉希

手机/微信: +86 18825645323 | 邮箱: althealam0531@gmail.com

个人网站: [Althea Lam's Website](#) (仍在开发中...)



教育经历

- 康奈尔大学 工学院 运筹学与信息工程-数据分析专业 (已有 Uncon-Offer 且延期一年入学) 2025.08-2026.06 (预期)
- 剑桥大学 2023 年剑桥大学暑期研习项目-人工智能项目组 2023.07-2023.09
- 武汉大学 数学与统计学院 数据科学与大数据技术专业 2020.09-2024.06
- 荣誉奖项:** 数学与统计学院优秀学生丙等奖学金, 武汉大学“优秀学生”称号, 宝钢教育基金会优秀学生奖学金, 珞珈交换优秀学生奖学金, 港澳及华侨优秀学生奖学金一等奖, “FLTRP·ETIC”全国大学生英语写作大赛二等奖, 全国大学生数学竞赛湖北省二等奖, 美国大学生数学建模比赛 H 奖 [Work Link](#)
 - 平均成绩:** 87.98/100 (排名 15%)
 - 相关课程:** 机器学习, 数据结构与算法, 数据库技术, 优化理论与算法, Python 语言程序设计, 数据科学引论, 多元统计分析, 统计计算, 数值分析, 概率论与数理统计, 数值分析

实习经历

- 字节跳动 Global E-Commerce – Data Science | 数据分析实习生 2024.08-至今
- Python, HiveQL, ClickHouse, Aeulous 上海, 中国
- UK 月报数据集开发: 该专项旨在构建数据集为 UK 市场的运营团队提供关键指标数据支持, 数据集包含 Longtail Order Rate, Create to Delivery Hrs, Trace Ontime Rate, DNR 等物流关键指标
 - 具体开发任务: 使用 Hive 构建月报所需指标的数据仓库, 并用 ClickHouse 优化查询性能, 完成了从数据预处理到性能调优的全流程开发
 - 周报自动化程序编写: 该专项旨在帮助 SEA 运营成员实现周报数据的自动化提取和更新, 并实现每周一自动运行程序, 更新周报中的关键指标数据, 帮助减轻团队成员取数负担
 - 具体开发任务: 利用 Python 编写脚本自动从多看板中提取所需的数据, 包含 SEA 五个国家 (新加坡、马来西亚、泰国、菲律宾、印尼), 并独立对接印尼团队并完成特殊周报改造需求
 - 月报、季报异动归因: 该专项旨在撰写 US、UK 和 SEA 的月报和季报, 并分析 US、UK 和 SEA 在 Lobal 与 Global Selling 履约时效的变化趋势, 针对优化和恶化的因素进行详细归因
 - 具体任务: 利用 SQL 提取数据编写月报和季报, 对异动的指标根据两种情况进行拆解分析: 第一, 变动情况 (连续性、非连续性); 第二, 服务商/快递员/买家/卖家/产品端
- 数能觉醒有限公司 | 数据工程实习生 2024.04-2024.08
- Python, HTML, Jinja, MySQL, Agent, Git, GitLab, 机器学习 深圳, 中国
- 周搜索量预测: 该专项旨在为广告域数据集新增有关周搜索量的指标, 帮助用户了解不同 ASIN 的搜索量情况
 - 具体开发任务: 利用亚马逊 ABA 的历史搜索量数据构建回归模型
 - 底层数据搭建及看板复刻: 该专项旨在获取亚马逊平台数据并复刻相关看板
 - 具体开发任务: 通过亚马逊平台的 API 拉取数据, 编写脚本定期拉取产品信息, 包括价格、评价、销售排名等关键指标, 使用 JSON 和 XML 解析库堆获取的数据进行格式化处理, 并复刻与广告&库存域相关的看板
 - 巡检 Agent 配置: 该专项旨在为客户的店铺提供告警功能, 当店铺达到用户所设定的告警条件时, 将会通过两种方式实现告警: 第一, 在平台发送告警通知; 第二, 通过邮件发送告警信息
 - 具体开发任务: 在短时间内掌握 HTML 与 Jinja 模版语法并实现销量巡检告警 Agent 的开发应用, 巡检效果获得了种子客户的一致好评, 同时持续跟进巡检规则的实现、落地与反馈, 完成了完整项目的闭环

项目经历

- 阿里云天池-新闻推荐竞赛 [GitHub Link](#)
- 该项目旨在为用户推荐个性化的新闻内容, 提高用户的点击率和满意度。其中, 具体开发任务为: 多路召回部分中复现了物品协同过滤与用户协同过滤模型, 并设计了冷启动召回策略
- 基于 GBDT 与 LR 预估 CTR [GitHub Link](#)
- 该项目旨在构建和优化 GBDT 模型以捕捉特征之间的复杂关系, 同时实现逻辑回归模型, 对 GBDT 输出进行最终 CTR 预测
- 基于 Transformers 微调常见 NLP 任务 [GitHub Link](#)
- 该项目旨在通过利用 Hugging Face 的数据集对 Transformer 相关模型 (如 BERT) 进行微调以实现常见 NLP 任务。其中, 涉及到的任务为: 问答系统、Quora 问题对、句子相似度、文本分类、标记分类、文本生成、摘要和翻译
- 模型平均方法在乳腺癌生存时间预测中的应用 [GitHub Link](#)
- 该项目旨在对乳腺癌患者的生存时间进行预测, 其中构建了三种生存分析模型 (Cox 比例风险模型、Weibull 模型与基于对数正态分布的加速失效时间模型) 与三种模型平均方法 (贝叶斯模型平均、Bagging、加权模型平均), 并对比较权重选择方法 (AIC、BIC、Mallows、Jackknife) 的效果

专业技能

- **编程语言:** Python, SQL, R, Shell, C, MATLAB, Lingo, HTML
- **Web Dev Tools:** PyCharm, VSCode, Linux, RStudio, DataGrip, XCode, MySQLWorkBench
- **Control Version:** Git, GitLab, GitHub
- **可视化:** SuperSet、Tableau、FineBI
- **语言:** 英语 (IELTS: 7, CET-6, CET-4), 普通话 (二甲), 粤语 (能够听懂)