

陈文钰

电话: 13950068454 | 邮箱: wenyvchen@gmail.com | 个人网站: wenyvchen.github.io



教育经历

哥伦比亚大学

纽约

应用分析硕士, GPA: 3.83

2022.09-2023.12

- 相关课程: 统计理论与科学研究, 机器学习理论与应用, 数据分析与可视化, 数据仓库与管理等

厦门大学

厦门

金融学学士, GPA: 3.75

2018.09-2022.06

- 相关课程: 数理统计, 计量经济学, 数学分析, 金融工程等
- 荣誉: 厦门大学优秀毕业生, 校级奖学金, 经济学院学业优秀奖学金, 厦门大学优秀学生干部等

实习经历

美团-点评事业部笔记产品部内容生态组

上海

内容治理策略分析实习生

2023.05-2023.08

- 日常数据监控:** 针对点评首页双列信息流负向曝光情况和治理专项的目标, 使用 SQL 和 Python 提取、分析内容信息、曝光、审核记录等数据指标, 监控数据的表现和趋势, 发现内容生态治理策略中存在的缺陷, 例如未能屏蔽“左下角”引流内容等问题, 并协同研发、审核团队, 通过完善模型、进行培训等方式改善治理效果, 最终达到使负向标签漏放率下降 60% 的效果
- 治理专项分析:** 依据日常对信息流高热高曝内容的跟踪以及团队策略方向, 诊断潜在的问题或者可优化策略, 跟进违规营销治理、优质画风产品等项目。其中包括前期对影响面、消费情况等数据以及案例方面的现状下探, 而后协同风控、产品等团队, 推进相应识别和预测模型的建设、创作者教育、标签体系完善等工作, 并构建用于监控治理效果的指标体系, 最后根据结果探索后期的治理方向

得物

上海

鞋类品牌运营实习生

2021.07-2021.09

- 数据分析:** 对 Jordan, New Balance 等品牌站内的 GMV、均单价、UV、转化率等销售和流量数据进行提取及分析, 包括对子类目、用户群、客单价等不同属性商品之间指标进行对比, 从而协助如何优化资源位分配、寻找如爆款补货等利益点的业务决策
- 品牌日常运营:** 追踪站内外品牌鞋款趋势, 确定各品牌卖点、调性等方面的运营策略, 并以此为根据确定 banner 位等页面的设计元素、进行页面前侧主推款及榜单选品、会场搭建、主题页建立与维护, 最终共搭建会场、榜单各 10+ 个, 主题页 30+ 个
- 活动跟进:** 参与某次 New Balance 新品发售活动, 过程包括前期通过协助流量分发刺激商家货量倾斜、端外流量投放、布局站内社区和开屏资源位分配等运营手段, 推动活动的上线, 以及后期完成活动效果的复盘; 最终发售售罄率达 100%, 品牌 GMV 当日环比+52.6%

项目经历

餐厅评价快速搜索器—基于 Yelp 数据

2023.04-2023.05

- ETL:** 利用 Apache Spark 读取、清洗、连接 Yelp 公开数据集中的店铺基本信息、星级、所获的评价等多表数据, 将清洗后的 JSON 格式数据储存至 MongoDB
- UI 及 API 搭建:** 利用 HTML, CSS, JavaScript 搭建搜索引擎的 UI 界面, 并使用 Flask 创建 API, 用于数据库与前端的通信; 最终实现在用户输入店名、所在地、星级、评分中至少一项后, 返回所有符合条件店铺的基本信息以及最有用的 3 条评论

因性格而异的营销策略—由用户所发表评论预测其 MBTI

2023.04

- 描述性分析:** 通过柱状图、饼图、词云等工具分析、呈现某论坛用户评论数据集中用户 MBTI 类型的分布情况、各个 MBTI 的用户在评论中最常使用的词语
- 文字清洗及模型搭建:** 对评论数据进行清洗、分词; 利用随机森林以及 XGBoost 搭建模型, 实现通过用户所发表的评论来预测其 MBTI 的功能, 最终模型在测试集上的准确率均达到 80% 以上, 从而可帮助根据用户不同性格类型制定营销方案的策略实现

基于歌曲特征预测其流行度

2022.11-2022.12

<https://www.kaggle.com/competitions/lalasongs22/overview>

- 数据预处理:** 运用 R 对包含 18 个变量 19485 个观测值的训练集进行数据清洗及描述性统计分析; 利用已有信息建立新的预测因子, 例如描述歌曲是否由多个艺术家合作演绎的分类变量; 通过主成分分析完成数据降维
- 模型搭建:** 利用 ranger 包, 搭建随机森林模型, 从而利用歌曲的节奏、流派等特点预测其受欢迎程度; 并基于均方根误差(RMSE), 进行超参数调优, 最终 RMSE 为 14.23

技能

- 数据处理、管理、可视化:** Python, R, SQL, Tableau, Excel
- 研究分析:** 假设检验, 实验设计, 回归分析
- 机器学习算法:** 决策树, 聚类分析, 文本挖掘, 时间序列分析, 神经网络