

刘胜超

求职意向：数据分析、数据挖掘

年龄：2003.6.3(22岁)
联系电话：18677066909
微信：18677066909
居住地址：广西壮族自治区南宁市

身高体重：188cm\80kg
电子邮箱：2805093969@qq.com
意向地区：江苏、浙江、广东、成都、
重庆、广西、福建



教育背景

2021.09 - 2025.07 (25届) 哈尔滨信息工程学院 软件工程(本科)
主修课程:java程序设计、面向对象程序设计、SSM框架\SpringBoot框架软件工程、数据库原理及应用、Linux操作系统、数据库结构与算法分析、操作系统、人工智能基础、Python 程序设计

专业技能

数据分析：熟练使用numpy（数据计算）、pandas（数据清洗 / 分析）、matplotlib/seaborn（可视化图表制作）、scikit-learn（机器学习建模），能独立完成从原始数据预处理到模型输出的全流程
数据库：掌握 MySQL，熟悉 SQL 查询、分组聚合、窗口函数、表连接等
可视化工具：熟练使用 Excel（数据透视表、VLOOKUP/INDEX-MATCH 函数）、PowerBI（交互式仪表盘）、Tableau（动态可视化报告），能将分析结果转化为业务易懂的图表；
机器学习：了解常用算法（线性回归、逻辑回归、K-Means、决策树、朴素贝叶斯）及 Sklearn 等
文本分析：熟悉 Jieba 分词、词云、情感分析基础流程
爬虫技术：掌握 Requests、BeautifulSoup、XPath，具备基础网页数据采集能力
其他：熟悉 RFM 用户价值模型（用户分层）、AARRR 漏斗模型（用户生命周期分析）、漏斗转化模型（销售环节分析），了解电商核心指标（GMV、PV/UV、转化率、复购率）计算逻辑；掌握缺失值处理（均值 / 中位数填充、插值法）、异常值检测（IQR 法则、Z-score）、特征编码（One-Hot、Label Encoding）、特征选择（相关性分析、方差筛选）。

实习经验

2024.09 ~ 2024.11 哈尔滨云脉互联网科技有限公司 数据分析
核心工作:1: 协助处理公司电商业务数据，用 SQL 提取每日 GMV、转化率、用户留存率等 40 + 指标，制作日报 / 周报，确保数据准时交付（准确率 100%）;2: 指标监控与异常分析：参与“新用户首单转化”专项分析，用漏斗模型定位转化瓶颈（“加购后下单”环节流失率 40%），提出“加购后推送 5 元立减券”建议，推动后该环节转化率提升 10%；3: 可视化支持：用 Tableau 制作“用户增长仪表盘”，实时展示新用户注册量、渠道来源占比，帮助运营团队识别高转化渠道（抖音信息流渠道转化成本降低 15%）;
成果:优化 2 份核心报表生成流程，将原本 1.5 小时的报表制作时间缩短至 40 分钟；输出 1 份专项分析报告，被运营团队采纳落地。

项目经验

2023.03 ~ 2023.06 校园电商平台用户价值分层与精准运营 组长
项目介绍：针对校园“零食配送”电商小程序 1.5 万用户数据，解决“营销资源浪费、高价值用户流失”问题，为运营团队提供精准策略支持（课程实践项目，团队 3 人，负责数据分析与建模）。
职责: 1: 数据清洗：用pandas处理用户消费数据（订单金额、消费频次、最近消费时间），剔除异常值（如单笔订单超 500 元测试数据）、填充缺失值（占比 2.3%，用均值填充），最终得到有效数据 1.4 万条；
2: 模型构建：基于 RFM 模型定义 3 个核心指标（最近消费间隔 R、消费频次 F、累计消费额 M），用 K-Means 聚类将用户分为“高价值忠诚用户（15%）、潜力增长用户（30%）、一般维持用户（35%）、流失预警用户（20%）”4 类；
3: 策略输出：针对不同层级用户设计运营方案（如高价值用户赠送“免配送费券”，流失用户推送“满 20 减 8 券”），用 PowerBI 制作用户分层仪表盘，实时监控各层级转化情况；
成果:高价值用户复购率提升 18%，流失用户召回率提升 12%，运营成本降低 25%。
2023-10 ~ 2023-12 校园周边奶茶店销量预测与库存优化 个人
项目背景：为校园周边 2 家奶茶店提供未来 2 周销量预测，减少库存积压与断货风险，核心数据为店铺 6 个月销售记录（日销量、促销活动、天气、节假日）（个人课程设计项目）。

职责：1：特征工程：提取 20 + 特征（如 “是否周末” “是否促销” “当日气温” “节前 3 天” ），用 One-Hot 编码处理分类特征，通过 Pearson 相关性分析筛选出 12 个关键特征；

2：模型对比与优化：对比线性回归、随机森林、XGBoost 三种模型，以 “均方根误差（RMSE）” 为评估指标，最终选择随机森林模型（RMSE 较线性回归降低 30%），通过网格搜索调优超参数（树深度 = 7，叶子节点数 = 25）；

3：落地应用：根据模型预测结果，为每家店铺生成 “周销量预测表”，标注 “高需求日期”（如周末、考试周），建议提前备货 25%；

成果：店铺断货率下降 20%，库存周转率提升 16%，单店月均减少库存成本约 1000 元。

2024-03 ~ 2024-06 **竞品奶茶品牌评论情感分析** 组员

项目背景：爬取某外卖平台 “蜜雪冰城” “古茗” 2 个竞品的 6000 + 条用户评论，挖掘用户痛点，为校园奶茶店产品迭代提供方向（团队实践项目，负责数据爬取与文本分析）。

职责：1:数据爬取：使用Scrapy框架爬取竞品评论（含评分、评论内容、购买时间），设置请求间隔（2 秒 / 次）规避反爬，成功获取有效数据 5800 条；

2:文本分析：用jieba分词提取评论关键词（如 “甜度” “冰度” “包装” ），过滤停用词（如 “的” “了” ），通过朴素贝叶斯模型进行情感分类（正面 / 负面 / 中性），模型准确率达 90%；

3:痛点挖掘：通过 TF-IDF 分析负面评论，发现核心痛点为 “甜度不可调” “包装漏液”，生成词云图直观展示用户关注点，撰写《竞品评论分析报告》；

成果：校园奶茶店根据建议新增 “半糖 / 微糖” 选项、更换防漏包装，后续用户好评率提升 22%。

荣誉奖项

校三好生 国家励志奖学金 校优秀学生奖学金

自我评价

- 1：具备扎实的数据处理与分析能力，对数据敏感，逻辑清晰
- 2：学习能力强，能快速适应新工具与业务场景
- 3：具备良好的沟通能力和团队协作精神，能高效完成跨部门协作任务