

# 吴臻函

150-4814-0526 | 15048140526@163.com  
WZH15048140526



## 教育经历

新加坡国立大学 海外QS前100

2023年08月 - 2025年02月  
新加坡

统计学 硕士

GPA: 4.3/5.0

主要课程: 应用回归分析、非参数回归、随机过程、时间序列分析、深度学习、有限样本抽样

东南大学 985

2019年08月 - 2023年06月  
江苏南京

信息与计算科学 本科 数学学院

GPA: 85/100

主要课程: 数学分析、高等代数、C++程序设计、数据结构与算法、机器学习、数据库系统原理、概率论、数理统计、运筹学

## 实习经历

美团

2024年09月 - 至今  
北京

商业分析实习生 核心本地商业/美团平台/用户增长与运营分析中心

- 效果广告专题分析专题: 利用因果推断模型PSM、CI等, 基于不同效果广告项目发展, 如医保买药效果广告、啤酒节效果广告等, 协助完成相关专题报告, 帮助业务发现问题并提出潜在解决方案, 产出有应用性的数据结论;
- 品牌广告专题分析专题: 协助完成美团平台不同业务的品牌营销项目分析, 通过使用不同的因果推断方法, 评估品牌广告对美团用户的行为影响;
- 平台级大促和常规促销分析专题: 协助平台级大促和日常促销的分析工作, 主要针对促销的效果进行深入分析, 提出迭代策略和方案;
- 增长经营分析专题: 协助完成经营分析报告, 以周/月维度对重点营销项目进行数据解读, 对项目的数据波动进行归因;

哈啰普惠科技有限公司

2024年05月 - 2024年08月  
上海

数据分析实习生 两轮出行事业部/大盘收入用户分析组

- 主要参与共享单车及助力车骑行收入分析工作, 负责单车和持卡骑行相关 AB 实验复盘, 协助周月报单均指标体系构建及拆解归因, 推进单车及卡等业务新种类的上架、扩量及迭代
- AB实验复盘: 协助完善单车和助力车单车及卡两类AB实验复盘逻辑及指标体系, 涉及因果推断等优惠算法调价、位置嗅探网格调价、APP前端页面模块调整等细分种类, 输出分析报告14次, 制作看板协助业务进行长期效果监控4个
- 收入数据指标构建及拆解归因: 协助完善单车和助力车大盘收入侧周月报的指标监控和异动归因, 主要关注单均价、流水、均摊等指标; 归因单车连续包月卡续卡率连续四周下降问题, 完成周报1次
- 模型建设: 协助完成2024年上半年200+城市单车和300+城市助力车单车骑行调价准实验场景复盘, 通过均摊及订单综合指标判定调价效果, 迭代单车调价城市特征模型, 结合XGboost、ElasticNet等统计学习模型输出新指标27个, 提升均摊3%
- 独立负责业务新卡种“储值卡”的上架、扩量及迭代的潜在用户识别及触达方面数据分析工作, 挖掘用户特征提炼核心指标4个, 提高用户转化率5%, 达成5城月收入100000+

国联证券

2023年03月 - 2023年05月

行业研究, 医疗信息化

- 深度报告

补充撰写医疗信息化行业深度报告: 收集医疗信息化行业相关政策; 分析国内医疗信息化行业竞争格局及发展空间; 更新数据, 分析行业重点公司卫宁健康、创业惠康等业务构成、2023年营收情况及优劣势

- 数据收集

利用Python从企查查等数据源对申万计算机200+标的进行股权穿透分析, 收集企业前十大股东中国央企基本面信息, 为投资决策提供数据支持

## 项目经历

车载动力电池健康状态算法流数据漂移检测

- 将长度为360001s的20组电池信息时序数据集, 按变量类型分割; 进行特征工程, 提取16个主特征进行分析
- 进行置换检验。针对每个特征计算Wasserstein测试统计量; 重复抽样多次, 利用抽样分布计算p值, 进行假设检验并建立置信区间
- 利用测试统计量结果和置信区间进行数据可视化

基于深度学习的股票价格预测

2022年01月 - 2023年01月

大学生创新创业训练 (SRTP) 校级重大项目

- 选取 Yahoo Finance 10只股票近十年数据, 计算18个股票技术 (STIs) 指标作为自动编码器 (Autoencoder) 训练输入; 基于 PyTorch 结合已训练的自动编码器进行输入数据降维, 随后利用一维稠密连接网络 (1D-DenseNet) 进行训练预测股票收盘价格
- 将18个股票技术 (STIs) 指标和股票数据集共同输入自动编码器 (Autoencoder) 进行数据降维; 基于 PyTorch 搭建 LSTM 循环神经网络进行训练预测股票收盘价格
- 利用 Python 可视化工具库 Scikit-plot、Seaborn 和 Pandas 进行数据可视化, 呈现股票价格的预测趋势
- 与随机森林、GDBT 和 XGBoost 算法相比, 稠密连接网络的 MAPE 和 MAE 值分别为 0.41 和 8.12, 下降超过80%; LSTM 的 MAPE 和 MAE 值分别为 0.63 和 14.19, 下降超过60%

知识图谱评测: 商品常识知识显著性推理

- 评估电商常识三元组显著性, 文本类型数据是否显著的二分类任务
- 基于 BERT-base 进行微调, 在输入表示上用特定的特殊分类标记区分三元文本对输入序列分类和文本对间隔; 在输出层前加入一个额外的基于多层感知机 (MLP) 架构
- 模型结果准确率71.4, 召回率58.2, F1分数64.7, 比 BERT-base 提高10个百分点

## 专业技能

- 技能: 编程语言: Python、C++; 工具: MySQL、R、Matlab
- 语言: 雅思6.5(6), 流利的英语口语能力, 擅长英语书面工作