王家威

求职意向:数据分析师

三 汉族

育 群众

○ 江苏省宿迁市

34 岁

17352322715

M 1817219421@qq.com



◇ 教育背景

2024-07 - 至今 中南财经政法大学 | 经济统计学专业 | 硕士

主修课程:时间序列分析、数据挖掘与机器学习、国民核算与宏观经济统计分析

2020.09 - 2024.06 常州大学 | 金融工程专业 | 学士 | 专业成绩: GPA 3.66/4 (专业前 5%)

主修课程:金融计量、统计学、财务会计学、管理学原理、衍生金融工具、Python 实训、Matlab 实训等。

[©] 项目经历

2024.08 - 至今 基于财付通数据的深圳消费测算与分析项目 | 项目核心成员

项目概述:作为项目核心成员,参与了基于**财付通数据**对深圳消费进行测算与分析的项目。该项目通过分析 2018 年 1 月至 2024 年 6 月的零售商户数据及 2018 年 1 月至 2024 年 9 月的零售和餐饮商户数据,探索**了第三方 支付数据**在消费统计中的应用,为政府决策提供了数据支持。

关键成就:

- **数据处理与清洗**: 负责采集和清洗 2018 年 1 月至 2024 年 6 月的零售数据,以及 2018 年 1 月至 2024 年 9 月的零售和餐饮商户数据,使用 **Python** 进行数据匹配、整合和异常数据剔除,确保数据质量满足模型构建需求。
- 数据分析与可视化:进行数据可视化,测算商户交易额相关指标,并总结分析结果。对比财付通数据与传统数据的特征,构建限下零售业商户底册,筛选出327667个样本进行深入分析,揭示消费模式和趋势。
- 模型构建与预测:利用统计模型 VAR 和机器学习模型 LSTM 对深圳社消零进行预测分析,优化模型参数,并对比分析结果,为深圳消费市场发展提供科学的数据支持。
- 技术应用与提升:熟练运用 Python 进行数据处理和分析,掌握机器学习算法在实际问题中的应用,以及数据可视化技术,有效提升数据分析的准确性和效率。
- **决策支持与建议**:通过构建的相关模型和底册,为政府提供决策支持,推动深圳消费市场的发展,提出了基于数据分析的政策建议,助力消费市场的可持续增长。

技能应用:

- 数据处理与分析: 熟练使用 Python 进行数据采集、清洗和分析,确保数据的准确性和可靠性。
- **数据可视化**:掌握数据可视化技术,能够清晰展示数据分析结果,帮助决策者直观理解数据。
- 机器学习应用:运用统计模型 (如 VAR) 和机器学习算法 (如 LSTM) 进行复杂的数据分析和预测,提升预测的准确性。
- 问题解决能力: 具备良好的数据处理能力和分析问题、解决问题的能力,能够在数据分析中发现潜在问题并提出解决方案。

项目成果:

● 项目成功将第三方支付数据应用于消费统计,为深圳消费市场提供深入的分析和预测,帮助政府制定更有效的政策,推动消费市场的发展,促进经济增长。

2023.09 - 2023-09 美国大学生数学建模竞赛——预测 Wordle 游戏结果 | 项目组长

项目概述:作为团队核心成员,负责开发和应用统计和机器学习模型,对全球流行的猜词游戏 Wordle 的结果进行分析和预测。

关键成就:

- **数据处理与预处理**:负责执行数据清洗和预处理,包括**异常值检测、数据转换**和**标准化处理**,确保数据质量满足模型构建需求。
- 模型构建与预测:
 - **问题 1:** 利用 ARIMA **模型**和**二次曲线拟合模型**预测报告结果数量变化,通过比较 R²和 RMSE,选定 ARIMA(1,1,0)模型进行预测,成功预测 2023 年 3 月 1 日的报告结果区间。
 - **问题 2:** 采用**中心对数比转换**处理百分比数据,通过**特征筛选**和 Adaboost **模型**预测未来日期的报告结果分布,模型评估显示 R²值高于 0.9,验证了模型的合理性。
 - **问题 3:** 基于**熵权法**的 **TOPSIS 评价模型**对单词难度进行分类,计算单词的综合得分,并准确评估特定单词 EERIE 的难度。

● 数据分析与洞察:

- 对数据集进行深入分析,发现难度模式下报告结果的百分比与选择难度模式的人数呈相反趋势。
- 揭示了超过 90%的玩家能够猜出超过 95.51%的单词,对 Wordle 游戏的长期吸引力提出了悲观预测。
- 模型评估与敏感性分析:
 - 对 TOPSIS 模型进行敏感性分析,调整关键指标权重,验证模型的准确性和稳健性。
 - 总结模型的优势和局限性,提出模型改进的可能方向。

技能应用:

- 熟练运用:ARIMA、Adaboost、TOPSIS 等统计和机器学习模型进行数据分析和预测。
- **掌握**:数据预处理、特征工程和模型调优的技能。
- 能够进行:复杂的数据分析,并将结果转化为商业洞察。

项目成果:项目不仅提高了对 Wordle 游戏动态的理解,而且通过提出的预测模型和数据分析,为《纽约时报》的谜题编辑提供了实际的业务改进建议,具有重要的商业价值和实践意义。

2023.03 - 2023.05 北京移动用户体验影响因素研究项目 | 项目组长

项目概述:作为项目核心成员,负责运用数据分析技术深入研究北京移动用户满意度的影响因素,通过构建客户特征集成模型和评价模型,成功识别关键因素并提出优化方案。

关键成就:

- 数据分析与特征工程: 运用随机森林和 XGBoost 算法建立客户重要特征模型, 精准识别影响用户满意度的主要因素。
- 模型优化与验证:通过改进的 K-means++算法构建客户特征集成模型,提升模型的 DBI 指标,确保模型的准确性和合理性。
- 评价模型创新:结合矩估计组合权重的 TOPSIS 评价模型,重新评估客户满意度,通过方差分析和 Spearman 相关性分析筛选合理客户数据,确保评分客观性。
- **预测模型开发:** 应用 **Stacking 集成学习**技术,构建回归预测融合模型,实现测试集拟合可决系数高于 0.9 的 高效预测。
- 算法改进与创新:提出首创的基于 K-means++改进的 K-medoids 算法,展现项目先进性和创新性。
- **非技术沟通与报告撰写**:设计简洁的非技术报告,将复杂数据分析结果有效地传达给中国移动北京公司。 技能应用:
- 数据分析与建模: 熟练运用随机森林、XGBoost、K-medoids 等算法进行深入数据分析和模型构建。
- **评价方法创新**:结合**矩估计理论**优化 TOPSIS **评价模型**,提升评价科学性和准确性。
- 集成学习应用:成功应用 Stacking 集成学习技术,提升预测模型性能。
- **非技术沟通**:将技术结果转化为易于理解的非技术报告,展现出色的沟通和表达能力。

项目成果:项目深化了对北京移动用户体验影响因素的理解,提出的优化方案为中国移动北京公司提供了业务 改进的实际建议,具有显著的商业价值和实践意义。

○ 技能荣誉

专业技能: 熟练掌握 SQL、EXCEL、Python、Matlab 等数据分析技术

语言能力:大学英语四六级,具备良好的英语读写能力和基本的沟通技巧,能够阅读和理解英文技术文档。

荣誉奖项: ●连续获得过十八、十九、二十届江苏省高等数学竞赛一等奖

- ●获得过十五届全国大学生数学竞赛二等奖
- ●获得过 2022 年全国大学生数据分析大赛二等奖
- ●2022 年 MathorCup 高校数学建模挑战赛——大数据竞赛本科生组一等奖
- ●数维杯国际大学生数学建模竞赛 Meritorious 奖
- ●美国大学生数学建模竞赛 Honorable Mention 奖
- ●第九届全国大学生统计建模大赛江苏省赛区三等奖

② 自我评价

个人背景:拥有金融工程学士和经济统计学硕士,具有扎实的统计学、数据挖掘与机器学习知识背景,为数据分析工作提供坚实的理论支持。

工作能力:熟练运用 Python、SQL、ARIMA、Adaboost 等数据分析工具和算法,具备丰富的项目经验,包括预测模型构建、数据预处理、特征工程和结果可视化,能够独立承担并完成复杂的数据分析任务。

综合素养:具备出色的逻辑分析能力和问题解决能力,工作细致负责,能够快速适应新挑战,良好的英语沟通 技巧和团队合作精神,能够高效地在多文化背景下进行工作和沟通。