

个人简历

给我一个支点，我可以撬起整个地球。

Give me a place to stand on, and I can move the earth.



基本信息 Basic information

姓名：陈家乐
籍贯：广东惠州
电话：13790780960

性别：男
工作经验：2.5 年
邮箱：961313212@qq.com



教育背景 Education

2016.09-2020.06 福州大学至诚学院 计算机科学与技术（数据服务方向）
主修课程：python 爬虫，大数据分析挖掘，大数据技术基础、大数据信息采集、高级数据库技术、计算机组成原理、计算机操作系统、计算机网络、数据库原理及应用
2023.04-2024.03 和阳日语学院 JLPTN2 证书



自我评价 Self evaluation

个人标签： 学习能力强，做实事、靠谱、
自我描述：

1. 熟练掌握 python 数据分析库 Pandas、Numpy、Matplotlib 等库的使用，以及掌握 sklearn, Tensorflow 机器学习，神经网络的使用，掌握爬虫的过程，并有过类似经验。
2. 有过多次数据的挖掘的项目经验，如利用神经网络预测股票未来 5 日走势，机器学习通过选取特征对航空用户进行聚类分析，爬取豆瓣网 TOP100 电影排名，还有网上毕设。
3. 讲究高效率的工作方式，如通过 python 的将每日工作的报表自动化合成，减少重复操作。
4. 工作认真负责有进取心，具备团队合作精神，踏实，敬业，能够自主解决难题，态度端正。
5. 个人性格开朗，能够较快融入集体和调动氛围。



工作经历 Work experience

2020.07-2021.6 深圳市微微玛科技有限公司 数据分析师
• 根据销售的需要输出每日报表（库存表和精品数据表），月度报表等等，精通 Excel。
• 根据销售的数据，沟通采购，仓储，物流，销售各部门计划商品数量，控制亚马逊库存 KPI
• 通过八爪鱼采集商品数据，价格，星级，评论等。
• 公司各类会议主持、活动组织，如年会主持、读书会主持、团队团建组织。
2021.07-2023.3 深圳市法本信息科技有限公司 数据分析师
• 根据基金指标书写 python 代码，实现数据计算到展示到系统前端。
• 根据业务需求通过 python 爬虫手段定期获取一些监管机构发布的基金信息并存储数据库。
• 负责基金的日报、月报、季报年报的数据生成和报送。

在校情况 Self evaluation

校内职务：在校担任副班长一职

在校荣誉：2020/06 优秀毕业生以及优秀毕业论文

2016-2020 每学期荣获奖学金（一等 2 次，二等 5 次）

2016-2020 福州大学 “优秀学生干部” 三次，福州大学 “三好学生”



网站前端部署阿里云服务器

项目描述: 通过 python 的 Django 修改美食网站并通过部署在阿里云上的, 通过外网可以访问并跳转, 以及信息的留言。该项目主要了解云服务器上 Webservice 的使用以及 Django 框架搭建网站

电商集数据清洗及 RFM 模型客户分类

项目描述: 从 Kaggle 获取到一份电商数据, 通过利用 pandas, numpy 等库进行了数据的清洗处理, 然后进行 RFM 模型的数据整个, 划分 8 种客户, 后续通过导出 EXCEL 通过可视化, 函数, 透视表的整合, 处理成报表形式。

LSTM 预测股票趋势 (毕设)

项目描述: 通过 Tushare 金融接口, 获取股票的历史数据, 通过 KDJ 和 MACD 技术指标可视化进行分析, 并通过函数计算出研判点后的未来涨跌幅的准确率, 其次搭建神经网络 LSTM&GRU 预测模型对数据进行训练, 以收盘价为标签, 其他处理的后的数据作为特征, 以过去 20 天的数据预测未来的 5 天收盘的方式, 预测未来的收盘价的趋势, 并通过调整神经层和优化器的选择的方式优化, 得出最终的模型 LSTM 采用 Adagrad 优化器, GRU 模型采用 SGD 优化器。

航空客户群分析

项目描述: 主要针对航空公司客户数据进行客户群细分和预测客户流失两个方面分别进行建模分析, 先通过数据集进行探索, 在聚类中, 以 RFM 模型为基础建立客户价值细分评判模型, 后续构建了 LRFMWT 模型 (即入会时间、总里程、飞行次数、平均每公里票价、时间间隔差值、平均折扣率) 然后分为五个客户群最为合适, 其次利用其它信息构建了机器学习预测的模型, 其次明确目的是客户流失类别预测, 最后 XGBOOST 的效果最好。

Scrapy 框架爬取猎云网

项目描述: Scrapy 框架爬取猎云网存储在本地 mysql 数据库, 利用 LinkExtractor 和 CrawlSpider 匹配过滤 URL, 采用 Xpath 匹配原则, 使用 twisted 的 adbapi 来异步连接池, 爬取内容大致为标题, 时间, 内容, 用户名等。过程调整了 settings 中的 DOWNLOAD_DELAY, 设置了 header 的, 没有遇到反爬虫

根据 APNEA 心电数据集判断是否存在睡眠呼吸暂停是

项目描述: 睡眠呼吸暂停是一种潜在的严重睡眠障碍, 其中呼吸反复停止和开始。如果整夜睡眠后大声打 and 并感到疲倦, 则可能有睡眠呼吸暂停。通过对心电信号的小波降噪处理, 提取时域特征, 线性特征等, 通过原本数据集的标签对心电信号进行切割, 采用随机森林, SVM 多种机器学习进行训练, 并通过后期调参等到最优模型, 同时通过迁移学习借助 Kaggle 平台的 GPU 算力, 进行了 LSTM、CNN 两种神经网络模型的搭建及训练, 准确率好的模型为 89%。

根据各种基金指标计算出相应的数据并实现自动化

项目描述: 负责基金的定期报告及业务需要的指标数据的自动化实现, 使用 python 从数据抽取数据, 然后经过数据处理后, 将数据封装成接口提供到业务端使用, 定期报告部分包含证监会, 人行监管部门的所需的定期报告, 根据返回的结果, 调整数据的计算逻辑, 并重新输送。