

钱晨

159-7152-2439 | qianchen2019@163.com



教育经历

弗吉尼亚大学	2019.08 – 2021.01
信息管理与信息系统 硕士	GPA: 3.60/4.00
中国地质大学（武汉）	2015.09 – 2019.06
遥感科学与技术 本科	GPA: 3.54/4.00

实习经历

浙江西谷数字技术股份有限公司	2018.07 – 2018.08
数据分析实习生	嘉兴
<ul style="list-style-type: none">收集并整理智能锁市场销售数据，分析用户评价，辅助技术部门制定研发策略搭建Linux服务器爬取电商平台智能锁销量及用户评价，使用snowlp库进行初步情感分析进行分词及停用词过滤等操作，利用gensim库构建LDA主题分类模型，制作分析报告并提供智能锁研发方向建议	
武汉蚁图时空科技有限公司	2018.01 – 2018.02
软件开发实习生	武汉
<ul style="list-style-type: none">国家第三次土地利用情况调查期间，负责开发无人机自动拍摄取样功能调用高德及大疆提供的SDK，设计并添加自动围绕兴趣点拍摄和多次悬停拍摄等功能集成该功能模块至土地利用情况调查及记录应用程序中，解决了外业调查中地形复杂难以拍摄的困难	

项目经历

音乐推荐系统	2019.11 – 2020.01
<ul style="list-style-type: none">整理网易云原始歌单数据并进行数据预处理，分别提取歌单和歌曲维度的信息针对小批量数据，运用协同过滤算法建立歌曲推荐预测模型，采用网格搜索交叉验证法评估结果并调节参数，分别针对歌曲和用户进行歌曲推荐预测迁移word2vec建立歌曲序列模型，探究歌曲之间的相似性，针对大批量数据，采用Tensorflow框架和Spark分布式系统搭建预测模型	
互联网金融风控预测模型	2019.09 – 2019.10
<ul style="list-style-type: none">清洗金融贷款数据并过滤无用特征，替换字符型数据为数值型数据并剔除相似特征调节正负样本权重解决样本不平衡问题，建立逻辑回归和随机森林等模型预测是否发放贷款，采用k折交叉验证法评估预测结果选取不同特征组合方式，多次调节参数，比较不同模型的表现	
基于深度学习的城市职住混合模式分析	2018.12 – 2019.05
<ul style="list-style-type: none">引进深度学习进行高分辨率遥感影像分类，探究城市居民职住混合情况预处理来自两千多个社区超过六十万的用户数据，根据数据分布情况进行分级并标注遥感影像结合过采样和多尺度随机采样方法解决了样本不平衡问题，多尺度问题和样本集过小产生的过拟合问题，应用Pytorch框架实现卷积神经网络BagNet，最终整体分类精度超过85%	

技术与其他

- 编程语言：Python, SQL, R
- 相关工具：MySQL, PostgreSQL, MongoDB, Tableau, Pandas, Numpy, AWS, Git
- 框架：Scikit-learn, Pytorch, Tensorflow, Keras