

魏翔

北北京市海淀区学院路 37 号北京航空航天大学 (地址)
100083 (邮编)

18525425662 (电话)
weixiang@buaa.edu.cn (邮件)
github.com/spaces-x (GitHub)



教育经历

- **北京航空航天大学** 计算机学院 · 计算机应用与技术 2018.09 – 至今 硕士在读
– 矩阵理论 (88), 算法设计与分析 (91), 学术英语 (87), 信息系统集成技术 (93) 等
- **大连理工大学** 软件学院 · 网络工程 2014.09 – 2018.06 学士学位
– 计算机网络, UNIX 网络编程, 数据结构与算法, 计算机组成原理, 操作系统, 离散数学等

技能专长

- **英语:** 六级(565)
- **语言:** Python, Java
- **工具:** VS code, Pycharm, Eclipse 具有良好的编写代码习惯
- **项目/框架:** Hadoop, PySpark, NumPy, pandas, Django, scikit-learn
- **算法:** 熟悉机器学习相关算法 SVM, 朴素贝叶斯算法, kmeans 聚类算法, 决策树生成算法等
- **操作系统:** 熟悉 Linux 下的基本命令, 具备在 Linux 系统下搭建分布式集群的经验

项目经历

- **高速交通大数据分析** 路网中心 2018.11 – 2019.01
– 搭建 hadoop 分布式集群存储环境, 对路网中心上传的月数据量达 500G 全国收费站数据进行存储管理
– 在交易数据记录中按属性进行聚类, 检测出离群点, 并将这部分离群点数据的信息完整杜进行评估, 过滤出存在逃费的异常行为记录
– 利用 spark 工具, 对交易数据进行分析, 生成不同省份的出行基本特征、区域运输特征以及通行费用基本特征的画像。根据分析结果, 给出合理的高速收费标准, 并针对不同的费率模拟推演计算盈亏数量
- **高速路链数据拓扑分析** Java/路网中心 2018.11 – 2019.01
– 对“路链”对象进行抽象封装, 采用双向链表的数据结构
– 读取 mapinfo 文件生成路链拓扑结构, 依据经纬度信息, 将某省所有收费站映射到路链
– 依据路链拓扑信息, 采用 Dijkstra 计算不同高速收费站间最短路链集合、距离及行驶时间
– 依据路链拓扑信息和站间流量, 通过 Map、Reduce 计算得到每个路链上承载的货运量和通行量, 并对这些路链进行聚类以标识高速关键路段
- **数据可视化网站搭建** Python/后端 2019.01 – 2019.02
– 基于高速交通大数据分析的结果完成数据可视化
– 后端采用 Django 框架, 通过对数据分析结果数据库的增删查为前端提供数据请求接口
– 前端采用 Echarts 框架, 对后台的数据加载渲染

自我评价

- 熟悉 Spark Hadoop 等大数据分析工具, 理解 Hadoop 工作原理
- 了解 Hadoop2 资源分配管理器 YARN 框架
- 热爱学习, 有较强的文档 (文献) 阅读能力, 做事认真负责, 有强烈团队意识和责任感