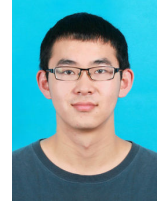


莫致良

5年工作经验 | 硕士 | 32岁 | 男
所在地: 新加坡(Employment Pass)
mzl9039@gmail.com +65 80321171



教育经历

浙江大学 地理信息系统 本科 2010.8 - 2014.6
浙江大学 地图学与地理信息系统 硕士 2014.9 - 2017.3

工作经历

Shopee 高级数据科学家 2021.10 - 2022.9

地图 Geocoding 服务: 2022.1 - 2022.9

- 项目介绍: GC服务为外卖、快递等业务提供地理位置查询定位支持。负责服务的线上稳定性及准确率提升(QPS: 90, p95: 205ms), 通过召回、排序两阶段优化, 服务准确率提升7%-9%。
- 召回优化: 针对地图POI数据特点, 优化BM25, 召回准确率提升 0.8%-3.1%, 准确率提升1.0%。
- 排序优化: 从训练数据、特征工程、重排序模型、后处理过滤等角度优化结果, 准确率提升 6%-8%。
- 工具: python, LightGBM, Elasticsearch, pyspark, CI/CD, GCP, Grafana, Docker

杭州海康威视 - 海康威视研究院 算法工程师 2019.1 - 2021.8

科技治超项目: 2020.11- 2021.6

- 项目介绍: 为政府超载货车治理项目提供轨迹算法支持。提供地图匹配、轨迹分类算法, 并聚类生成热门路线以供超载货车治理。一期地图匹配落地部署; 二期轨迹分类模型尚未上线。
- 超载状态分类: 训练分类模型, 预测货车是否超载(AUC 0.96)。包括轨迹特征, 车辆特征, 位置特征等
- 工具: python, LightGBM, Map Matching, Java, Word2vec

轨迹预处理项目: 2019.1-2019.8

- 项目介绍: 研究性项目, 实现轨迹预处理能力。基于HMM实现地图匹配算法并优化准确率, 处理百万级路网图数据, 最终评价指标(cmp, 类似准确率) 达到89%。
- 地图匹配算法: 旨在将GPS轨迹与路网路段匹配。基于HMM模型, 添加方向信息优化准确率; 借助空间索引, A*算法和Viterbi算法加速搜索与计算。
- 工具: Java, Hbase, Graph, Map Matching, Hidden Markov Chain, Viterbi, A*, Spatial Index

数字化还原项目: 2019.9-2020.6

- 项目介绍: 室内场景数字化还原, 即建筑物建模及跨相机下的运动目标轨迹还原。
- 轨迹匹配模型: 结合运动目标轨迹、图像face & body的embedding构建轨迹匹配模型。
- 工具: python, LightGBM, IndoorGML

杭州海康威视 - 海康威视研究院

大数据开发工程师

2017.4 - 2018.12

园区跨镜追踪项目: 2018.9-2018.12

- **轨迹生成:** 构建项目原型系统, 通过标记多目标的人脸、人体的身份ID, 还原目标的运动轨迹。

视频大数据开发与维护: 2017.4-2018.4

- **CI/CD:** 基于jenkins实现应用的自动化测试
- **性能优化:** 优化应用的子模块, 实现子模块50%的性能提升
- **工具:** Java, Scala, Shell, Spark, Hbase, Elasticsearch, Zookeeper, Jenkins

竞赛

Kaggle - Foursquare Location Matching

32/1079(top 3%)

2022.4 - 2022.7

- **项目介绍:** 将地理点(Location, 包括坐标及文本等信息)聚类为POI, 使用 IOU 指标评价聚类结果。
- **完成 pipeline 及准确率优化:** 文本预处理, 距离+文本召回, 粗排序, 二分类
- **召回:** 基于 ES 进行文本+距离召回及纯文本召回, max IOU: 0.9892
- **特征工程:** 距离特征, 文本距离特征, word2vec 特征, TF-IDF 特征, 图结构特征, 其它特征等 400+ 特征
- **粗排序 & 二分类:** 基于 LightGBM 训练粗排及二分类模型, IOU: 0.90985
- **工具:** python, Elasticsearch, LightGBM, Gensim, Word2vec, pytorch & pytorch-geometry

其它

语言技能: 英语六级, 读写流利, 听说中等

个人介绍: 坚持阅读相关领域论文, 跟进最新的研究成果。

专业技能: 熟悉python及基本的java开发, 大数据组件如spark/Elasticsearch/hadoop/hbase, 掌握基本的Linux、git等工具。