

罗嘉濠 (Edward)

基本资料

年龄: 29岁

手机号码: 18718752526

邮箱: jiahao.luo2014@gmail.com

教育经历

Georgia Institute of Technology (Atlanta, USA)

08/2015 – 05/2017

- 硕士 / 计算机科学与工程 GPA: 3.4/4.0
- 核心课程: Machine Learning, Artificial Intelligence, Computational Data Analysis and Modeling and Simulation.

华南理工大学

09/2010 – 07/2014

- 本科 / 土木工程 GPA: 3.5/4.0
- 优秀毕业论文 (典型城市风气候分析及其在建筑抗风中的应用) / 国家励志奖学金 / 利海奖学金

主要技能

- 编程语言: Python / Scala
- 深度学习框架: 了解 PyTorch / TensorFlow
- 大数据处理框架: 掌握 Apache Spark / Hive SQL
- 其他: Git / Shell / Docker

工作经验

人工智能与机器学习工程师 – 顺丰科技有限公司

07/2017 – 至今

- 对海量的物流数据进行数据提取、清洗和转换, 建立各种指标和报表。
- 根据不同的业务应用场景需求, 深入分析与洞察, 提炼和发现业务规律, 指导模型特征构建, 定位产品相关数据问题及分析优化。
- 利用机器学习等技术挖掘海量物流运单数据, 开展理赔特征分析与洞察, 构建理赔风控预警模型。
- 开展理赔成因评估, 构建和优化客户画像, 物流网点画像以及快递小哥画像, 并进行检验评价。

IT Specialist – (SSP) Southeastern Security Professionals, LLC

12/2014 – 07/2015

- 网络管理、数据库维护与管理, 远程或现场给产品提供技术支持;
- 场地平面图测量与绘制, 设备线路图设计等。

比赛经验

技术嘉年华算法挑战赛 (冠军 1/53) – 顺丰科技有限公司

10/2019 – 12/2019

- **顺丰产品推荐 (初赛):** 基于顺丰亿级运单信息, 根据历史记录判断并推荐给客户应采用的顺丰速运产品。主要负责特征工程, 采用了 LightGBM 模型训练, 以及相关的样本平衡方法, 最后部署 docker。
- **顺丰客服对话意图识别 (决赛):** 根据客服对话 (文字), 通过自然语言处理技术, 识别出客户来电意图。主要负责对话文本多方式增强, 采用了 BERT-base 模型训练, 以及多模型融合, 最后部署接收 API 请求的 client 端。
- 除了作为比赛的主要参与者, 还承担了团队负责人角色, 承担起最初的团队组建, 任务制定, 以及后续团队力量凝结并推进项目前进的责任。

项目经历

物流理赔风险预警项目 (Pyspark / XGBoost – Random Forest / 词嵌入)

07/2017 – 至今

项目背景

由于导致快递件发生理赔的原因可能出现在快递件运输的各个阶段, 为此, 基于快递件理赔大数据, 对快递在 揽件阶段、运输阶段 以及 理赔阶段 整个流程进行分析, 并对各阶段的主要实体进行画像构建。(收寄件客户画像、托寄物画像、运输中转网点画像、快递员画像以及理赔专员画像)

项目描述

- 基于快递件理赔大数据, 对快递在 揽件阶段 进行分析, 并以收寄件客户为实体进行画像构建。根据业务方需求, 在 理赔阶段 的快递件采取关键字规则匹配, 建立高风险客户黑名单。

- 建模阶段，针对托寄物的品类对个别地区采取分地区建模，如生鲜水果大省海南、山东等，大闸蟹产地江苏苏州、江苏泰州等，分地区建模能有效提高模型的效果。
- 根据模型的理赔风险预警值，在揽件阶段进行开箱拍照、身份证校验等操作，从而降低理赔风险，减少理赔成本。

项目职责

- Hive 大数据平台上进行数据提取、分析、准备与清理，提取基本特征。
- 基于 Spark/Scala 对预处理后的数据进行特征工程，丰富特征。主要以时间维度、空间维度以及地址文本进行特征扩充，时间维度包括按不同时间窗口对目标特征进行统计，空间维度包括对客户收件地址变换统计等，并利用 XGBoost 和随机森林算法进行建模，得出理赔风险预警值。
- 结合业务需求，对理赔风险预警值进行合理划分，平衡召回率和精确率，后续模型上线后，特征不断优化和改进。

项目成果

- 目前项目快递员揽件阶段以及理赔员理赔阶段建模已上线运行，并发表了以下两项专利。
- 专利号 201810067080.0 《一种基于 XGBOOST 算法的快件理赔预警技术》，专利号 201810067068.X 《基于随机森林算法的理赔预警模型》

客户差异化管理 (SQL / Pandas)

03/2020 – 至今

项目描述

- 理赔客户画像：根据客户相关的收件、理赔、常寄托寄物等数据，构建理赔客户画像，并对客户的理赔风险进行评分评级，目前应用于辅助理赔人员判断该客户是否符合快速理赔标准，同时也可以为相关业务提供支持。
- 智能化谈赔：根据相关的收件信息、理赔信息以及历史案例等，为理赔工单测算并提供理赔历史指导价格，辅助谈赔。

项目职责

- 梳理业务逻辑，构建理赔风险客户画像基表，定时自动化输出各项指标，通过评分卡模型，对客户风险进行评分评级，通过数据挖掘为地区理赔人员决策提供数据支持；
- 基于搜集到的 3C 产品新品价格、维修价格、以及历史理赔谈赔的数据，通过 Elasticsearch 构建历史理赔案例索引，利用相似性匹配和机器学习算法分别给出市场指导价和理赔指导价，辅助理赔客服人员进行谈赔。

相似工单推荐 (SQL / Pandas / fasttext / jieba)

08/2020 – 至今

项目描述

- 根据理赔员当前处理的工单特征，从历史结案的工单中，推送经过筛选的历史标准工单，以此来作为谈赔案例，从而辅助理赔员更好地谈赔。

项目职责

- 前期通过采用权重求和的方式，对相似工单涉及的有效特征（电话号码、保价金额、索赔金额、托寄物名称等）进行计分排序，其中涉及到 jieba 分词、fasttext 词嵌入等技术，同时采集理赔员对推荐结果的质量反馈；
- 后期会基于采集到的反馈结果，作为标签，采用 Learning to Rank 方法，对最终的计分排序结果优化，提高推荐工单的相似性，排序的合理性。

大数据物流图谱 (Python / Apache Spark – GraphX / JavaScript – D3.js)

03/2018 – 12/2018

项目描述

- 基于历年物流数据，以不同的寄件和收件方为节点，收件关系为边，收件数量、占比、金额等为权重，构建物流大数据图谱，研究探索不同公司客户之间的物流关系（供应链关系、金融关系）。

项目职责

- 提取并清洗历史物流数据，建立 Neo4j 图数据库；

- 基于 NetworkX Python 库对提取数据进行物流图谱构建，计算图谱相关结构特性，如节点的 PageRank 值、经过每个节点的三角形数目等；
- 基于 SIR 理论模型，建立图谱节点间的风险传播模型，模拟随机风险对图谱中其余节点的影响，探索不同公司客户间风险传播的影响及抗风险能力；
- 利用 D3.js，对提取建立的物流图谱进行可视化。