

刘斌

☎(+86) 13127810651 | ✉liupoem@126.com | 📄liubinpoem | 📱liubinpoem

"Stay curious, stay passionate, life is a huge game, never cease to code and explore!"

专业技能

- 熟练: Spark、HIVE、Scala、Python、Java、TensorFlow、Keras、数据结构、机器学习、SQL、Docker、Linux
- 一般: Maven、Tomcat、Shell、GitHub/SVN、Visual Studio、Vim、Android Studio、IntelliJ IDEA
- 英语: CET6.530 (裸考), 基本交流无障碍

项目经历

汇兑账号挖掘打击

02/2019 - 至今

深圳/腾讯

- 项目目的: 国家对反洗钱管控越来越严格, 需要从账号中找出经营汇兑业务的账号, 净化支付环境
- 进行中

基于出入账文本的风险账号挖掘项目

12/2018 - 至今

深圳/腾讯

- 项目目的: 利用出入账文本信息, 挖掘赌博、诈骗、色情等风险账号, 净化支付环境
- 从真实数据库中, 选择赌博和微商账号构造数据集; 在 GPU Docker 中安装 TensorFlow 以及 Keras
- 利用全量账号的转账文本挖掘新词, 结合结巴分词的词典、funNLP 以及搜狗开源词库, 构造自定义词典; 利用结巴分词处理数据
- 利用 Keras 开发 TextCNN 模型, Gensim 库得到词向量和文本向量, 对比词向量和 TFIDF 方法作为向量化方法, 和 TextCNN 模型、Xgboost 模型、BERT 等几个分类模型
- 对比各个模型后得到最佳模型, 每周挖掘出高精度的 1w 的赌博账号, 然后写入账号系统进行打标、拦截和打击, 以及高风险赌博用户进行限制, 净化支付环境
- 相关技能: Python、TensorFlow、Docker、GPU、深度学习、NLP、BERT

大规模无监督算法异常发现项目

06/2018-11/2018

腾讯/深圳

- 项目目的: 有监督的方法需要大量的标注数据, 标注数据成本高且难以做到非常准确, 无监督方法可以避免这些弊端
- 利用 PrefixSpan 算法进行特征组合, 利用 KL 散度进行特征选择
- 利用 Spark 实现论文《Outlier Detection in Complex Categorical Data by Modeling the Feature Value Couplings》中属性共现算法, 衡量账号特征取值的异常度和整体账号的异常度
- 实现 Weighted MinHash 算法, 对计算得到异常度的账号进行聚类划分, 然后计算每个聚类的异常度, 把账号划分到异常度最高的聚类中去, 异常度挖掘算法确实有效
- 相关技能: Spark、Scala、Hive、HiveSQL、Hadoop

PipeML 自动化机器学习平台项目

04/2018-08/2018

腾讯/深圳

- 项目目的: 让团队成员无脑自动化建模, 并得到机器学习模型
- 利用 Spring+MyBatis 框架作为 Web 框架, jQuery+eCharts+Bootstrap 作为前端显示组件, MySQL 作为后台数据库存储数据
- 继承 Python Sklearn 库的类, 开发自定义的模型, 前台接收到数据请求后, 调起 Python 算法脚本, 进行自动化预处理、特征选择等操作
- 在前台输出每个用户该次训练的模型指标、特征重要性排序、IV 值柱状图, 选择最合适的模型和最重要的特征
- 利用 PMML 文件的方式和 pickle 文件的方式保存用户的模型, 得到可用的模型结果
- 相关技能: Python、Java、Sklearn、Spring、MyBatis、JavaScript、Bootstrap、MySQL

风险团伙可视化项目

06/2017-09/2017

腾讯/深圳

- 项目目的: 将利用 SparkX 和 Pregel 框架进行图算法挖掘出来的团伙进行可视化展示, 方便审核定性
- 以 Django 作为 Web 框架, 进行数据展示; MySQL 作为数据存储库; 采用 D3JS 技术展示团伙
- 在团伙展示中, 以丰富的数据信息使团伙尽可能变得可分, 如男女采用不同图标, 已标黑的账号以不同颜色展示, 以箭头展示资金流向, 边上标明数字表明资金流
- 相关技能: Scala、SparkX、Pregel、JS、Django、Python

教育经历

上海交通大学 电子信息与电气工程学院

09/2015 - 03/2018

硕士研究生 嵌入式体系结构实验室 (LOEA) 集成电路工程

上海

上海交通大学 电子信息与电气工程学院

09/2011 - 07/2015

本科 微电子系 微电子学

上海