

刘子扬

联系方式: 15951917230 电子邮箱: 691653736@qq.com 意向岗位: 数据分析/数据挖掘/异常检测
年龄: 23 户籍: 江苏 毕业时间: 2020 年 4 月



教育背景

2013-2017	南京邮电大学	光电工程学院	工学学士	光电信息科学与工程	排名前 30%
2017-2020	南京邮电大学	计算机学院	工学硕士	计算机技术	排名前 20%

专业技能

计算机技能:	熟练掌握开发语言(Python), 掌握算法与数据结构、Linux、数据库、计算机网络等知识, 通过 CCIE(数据中心)认证;
机器学习:	熟练掌握传统机器学习算法理论, 熟悉相关工具包 scikit-learn、pandas;
异常检测:	熟悉异常检测相关算法, 例如: Isolation Forest、One-class SVM;对异常检测在信用卡反欺诈领域有深入研究; 了解信用卡评分模型。

项目经历

2017.9-2018.4	基于 Gaussian Mixture Model 的增量聚类	工具: Python
<ul style="list-style-type: none">任务描述: 在数据流情景下的聚类问题实际上是增量聚类的问题。因为数据流连续不断的特点, 对算法的处理的效率要求较高, 需要针对新数据的不断流入, 动态地调整和更新聚类的结果, 以此真是反应数据流的聚类形态。独立设计一种针对数据流的增量聚类算法, 通过对新旧数据建立 Gaussian Mixture Model, 模型的求解过程中采用 EM 算法, 并根据 BIC 选择最佳的 k 值; 采用 Kullback-Leibler divergence 对新旧混合模型中的各个高斯分布进行相似性计算, 并对相似性较高的两个高斯分布进行合并, 从而达到模型的更新。		
学术成果: 申请专利		

2018.09-2019.01	基于 Dirichlet Process Mixture Model 的信用卡反欺诈	工具: Python
<ul style="list-style-type: none">任务描述: 采用 Dirichlet Process Mixture Model 对正常数据进行建模, 并在此基础上采用 One-class SVM 实现对欺诈交易的检测, 且实现了对模型的增量更新。该模型主要解决了以下问题: 欺诈样本几乎无穷: 有监督方法大都是用标记好的正常数据和异常数据进行训练, 得出模型检测未知数据的类型, 所以这种方法检测不出新类型的异常; 欺诈样本获取代价过高: 信用卡异常数据获取代价过高, 仅存的一些数据仅能用于对模型的测试; 严重的类别不平衡问题: 训练集中正常数据往往比异常数据多的多, 这种不平衡会导致误分类, 以至于分类精度不高。		
学术成果: 申请专利		

比赛经历

2018.04-2018.09	蚂蚁金服算法大赛-支付风险识别	工具: Python	开源框架: Xgboost
<ul style="list-style-type: none">完成前期数据清洗、数据分析、特征提取等工作;针对训练集中的灰样本, 提出了多种解决方案, 其中基于 Isolation Forest 和 k-means 的解决方案对初赛成绩有较大提升。			

实习经历

2018.09-2019.01	南京邮电大学盐城大数据研究院	数据挖掘研究员
<ul style="list-style-type: none">调研多种异常检测算法, 尝试对其进行增量式更新的改进。		
2016.03-2016.05	贝塔斯曼欧唯特信息系统公司	产品调研员
<ul style="list-style-type: none">任职期间主要任务是研究国外、国内主流信息安全厂商产品的功能特点、部署方式和环境, 以及防护效果, 可能包括但不限于: 下一代防火墙、IPS、DDoS 防护产品、漏洞管理产品和网络流量可视化监控产品等。		

曾获奖励

- 本科在校期间获得 1 次三等奖学金、1 次社会工作奖;
- 研究生期间获得 1 次二等奖学金、1 次三等奖学金。