

赖澄宇

✉ laichengyu@zju.edu.cn · 📞 18870756356 · 🌐 <https://github.com/cycl2018>

教育背景

浙江大学，软件工程，硕士	2022.9 - 2025.3
• 主要研究方向: 图神经网络, 图异常检测, 数据挖掘, 推荐系统	
南昌大学，数据科学与大数据技术，学士	2018.9 - 2022.6
• 专业排名第一 • 校级特等奖学金、校三好学生、优秀毕业生	

竞赛经历

天池 CIKM 2022 AnalytiCup Competition 冠军 (Rank 1/1746)	2022
Kaggle AI Village Capture the Flag @ DEFCON 亚军 (Rank 2/668)	2022
华为云 GaussDB 数据库挑战赛 亚军 (Rank 2/1123)	2021
CCF-BDCI 泛在感知数据关联融合计算 亚军 (Rank 2/1168)	2021
ACM-ICPC 亚洲区域赛南京站 银奖 (Rank 39/548)	2020
ACM-ICPC 江西省赛 金奖 (Rank 1/203)	2020

科研经历

Better Late Than Never: Formulating and Benchmarking Recommendation Editing	2024
• KDD2024(第一作者在投) 研究领域: 推荐系统 模型编辑 • 提出了一种全新的可编辑推荐任务, 旨在无需访问原始训练数据或重新训练模型的情况下, 从推荐系统中删除不恰当的推荐项。定义了三个基本目标: 严格纠正、协同纠正和集中纠正, 并设计了三种评估指标来定量衡量每个目标的完成度。针对推荐编辑目标提出了一个新颖的 Editing BPR Loss, 并通过一系列从相关领域借鉴的方法建立了一个综合基准。	
Frequency Self-Adaptation Graph Neural Network for Unsupervised Graph Anomaly Detection	2024
• IJCAI2024(第三作者在投) 研究领域: 图神经网络 异常检测 对比学习 • 针对图节点异常检测中异常节点导致的图信号频率偏移问题, 设计实现了一种全新的无监督频率自适应图神经网络 SAG, 通过将全通信号作为参考, 自适应地融合来自多个频带的信号。它以自监督的方式进行优化, 并为无监督的图异常检测产生有效表征, SAG 相比于现有最先进模型性能提升 5% 至 10%。	

项目经历

图联邦异质任务学习 CIKM CUP 🌐	2022.8 - 2022.9
• 技术栈: 图神经网络 联邦学习 自监督学习 • 背景: 本项目面向分子图的任务异质联邦场景, 某些参与者的目标是对分子的类型进行判断, 即分类任务, 另外部分参与者的目标是预测分子化学性质的强弱, 即回归任务。在这种任务场景下, 虽然参与者都要求训练得到的模型具有对分子图表征的理解能力, 但是其具体的学习目标是完全不同的, 相比数据分布的异质性更具挑战性。 • 由于不同客户端所拥有的图的节点属性、标签各不相同, 相关性较弱, 但所拥有的图结构是类似的甚至可能是相连接的, 具有较强的相关性。于是考虑只利用结构信息来进行联邦学习, 同时采样相同的自监督学习任务 (GraphMAE) 来统一不同客户端的任务, 采用 SGD 的方式在服务端依次训练公共模型获取图的结构表征。 • 利用公共模型获取的结构表征和客户端自身拥有的属性和标签, 在各个客户端内部训练私有模型进行预测。 • 结果: 作为第一贡献者取得比赛冠军, 性能领先第二名 1.6%	
无监督图异常检测算法库 DGLD 🌐	2022.6 - 至今
• 技术栈: 图神经网络 异常检测 无监督学习 • 背景: DGLD 是一个基于 PyTorch 和 DGL 的深度图异常检测开源库。它提供了流行的无监督图异常检测方法的统一接口, 包括数据加载器、数据扩充、模型训练和评估。 • 作为项目主要贡献者, 参与了项目搭建, 维护, 以及开源算法的贡献, 复现并集成了包括 GUIDE, GAAN, AAGNN 在内的多个图异常检测算法, 搭建项目环境, 设计项目主体框架, 后期维护等。	
广告曝光量精准预测 字节跳动夏令营	2021.8 - 2021.9
• 技术栈: xDeepFM LightGBM GRU • 背景: 广告的推荐与投放是一个复杂动态的系统, 因此对于给定广告的投放设置和投放环境等信息, 预估广告预期投放的效果, 无论对于投放系统还是广告主, 都有着非常重要的价值。 • 针对广告曝光量精准预估中的时序动态变化挑战, 采用 xDeepFM 自动提取交叉特征, GRU 提取时序特征, 采用先分类再回归的策略利用 LightGBM 建模最终的预测模型。针对冷启动挑战, 考虑到新旧广告的差异性, 对新老广告分开建模。 • 结果: id 粒度准确度提升 1.42%, 曝光量粒度准确度提升 1.36%	