

黄真川

✉ hzhchuan@gmail.com

🌐 chuan92.com

教育背景

- 2014.9–2017.6 硕士, 南京大学计算机系, 机器学习与数据挖掘研究所 (LAMDA).
2010.9–2014.6 学士, 南京大学计算机系, 保研, GPA Top20% .

实习经历

- 2016.8–至今 阿里巴巴, 搜索事业部搜索算法团队, 算法工程师.
* 分析手淘日志里不同点击序列后的 PV, CTR, 为商品展示比例控制提供参考
* 基于 ODPS GRAPH (面向迭代的图计算处理框架) 设计实现一种排序学习算法, 将商品点击转化率和价格因素综合考虑, 优化提升 GMV
- 2016.4–2016.7 腾讯, 互动娱乐运营部数据挖掘算法组, 数据挖掘工程师.
* 在 Spark 平台上设计实现一种多示例多标记学习算法用于游戏道具推荐, 重点优化顶部位置推荐的准确率; 算法高效且准确率相对于人气推荐有一倍提升;
- 2013.7–2013.10 eBay, StubHub DevOps Team, Software Engineer Intern.
* 开发内部网络 (scp, curl) 以及追踪 Maven 仓库依赖变动等的运维工具

科研经历

- 2015.12–2016.4 多模态多标记学习研究.
将在校大学生的奖学金、助学金评定建模成一种多模态多标记学习问题。将多个来源的信息看作不同的模态, 提取过程有效保护隐私, 通过正则项引入标记之间的关系, 在半监督的框架下预测奖学金等级。相关论文 *College Student Scholarships and Subsidies Granting: A Multi-Modal Multi-Label Approach* 已被 *IEEE International Conference on Data Mining 2016* 录用
- 2015.7–2015.12 基于近邻重构的多示例学习研究.
多示例学习是一种半监督学习, 其输入为一系列被标注的“包”, 每个“包”都包括许多示例。设计并实现了一种基于近邻重构的多示例学习算法, 可用于文本情感识别、图像感兴趣区域检测等。论文发表于 *2016 International Conference on Intelligence Science and Big Data Engineering*.

竞赛经历

- 2015.12–2016.1 Prudential Life Insurance Assessment, Kaggle Competition, 116th/2619.
* 使用 GBDT, DNN 作为基学习器, 通过线性回归融合, 解决生命保险评级的问题.
- 2014.12–2015.1 Click-Through Rate Prediction, Kaggle Competition, 108th/1604.
* 移动设备广告点击率预测问题, 主要使用 FTRL 和 FM 模型

社区

- Blog chuan92.com.
使用 Jekyll 在 DigitalOcean VPS 上搭建的个人博客, 整理技术方面的一些思考和书评
- GitHub github.com/chuan92.
热爱开源, 为 breeze 数值计算库, python-mode 等项目提过 issues 和 pull requests

荣誉奖项

南京大学优秀毕业生, 南京大学计算机系优秀学生