

邮箱: asdwsq@foxmail.com | 毕业时间: 2018.03

教育背景

北京邮电大学	网络技术研究院	2015.09 – 至今
• 交换与智能控制国家重点实验室	计算机科学与技术专业	工学硕士 GPA: 3.37/4.00
北京交通大学	电子信息工程学院	2011.09 - 2015.07
• 通信工程专业	学士学位 校级二等奖学金(2次)	GPA: 3.54/4.00 (排名 27/312)

专业技能

- 熟悉 C/C++, Python 编程语言, 了解 Java 和 NodeJS, 熟悉 UNIX/Linux 编程环境
- 熟悉基本的数据结构与算法
- 了解常见的机器学习和深度学习算法, 如决策树, SVM, 朴素贝叶斯, CNN, RNN, LSTM 等
- 有 Elasticsearch, Druid.io 等海量数据仓库的实践经验, 了解 LevelDB, Redis 等 NoSQL 数据库
- 具备较好的英文读写译能力, 通过 CET-6 (517/710)

实习经历

Hulu Inc. 北京研发中心	2017.05 - 至今
<ul style="list-style-type: none">• 协助构建流视频播放器的日志分析系统, 采用 Elastic Stack 和 Druid.io 作为数据驱动, Kafka 作为中间件实现数据的分布式流处理• 借助 Elasticsearch 的数据聚合功能实现 PB 级数据的压缩和存储, 压缩率在 95% 以上• 设计采用 LevelDB 和 Redis 数据库作为 Elasticsearch 缓存中间件, 经测算 ES 集群的压力整体可下降到原来的 20%• 参与构建异常数据监控系统, 采用 LSTM 实现时间序列的自动告警和异常检测	
百度公司 (互联网数据研发部)	FEED 流通用转码项目 2016.09 - 2017.03
<ul style="list-style-type: none">• 完成网页特征的提取和过滤, 优化网页分块与结构化数据提取等策略• 参与中文网页分类项目, 使用最大熵模型对离线网页进行多标签分类, 使用 C++ 完成开发工作• 通过挖掘内容页链接和翻页动作, 完成网页翻页块和图集页面的内容识别与抽取工作, 使用 C++ 和 Python 实现;• 个人负责的图集识别的线上准确率由原先的 83.8% 提升到 95.7%, 并已上线运行	

项目经历

腾讯社交广告算法大赛, 广告转化率预估	初赛148/1000, 复赛162/200	2017.05 - 2017.07
<ul style="list-style-type: none">• 初期通过多线程提取点击率、转化率、组合比例等统计特征, 结合滑动窗口避免避免信息泄露问题; 采用 xgboost 训练预测模型, 采用交叉验证进行模型调参• 中期针对转化率特征进行贝叶斯平滑以减少预测误差; 采用 stacking 的方式提升模型的泛化能力和鲁棒性; 采取 logistic 平均对多模型的预测结果进行融合• 后期尝试采用 xgboost 提取特征, 结合 FM 或 FFM 构建预测模型		
FY4 气象卫星应用测试与优化分析		2016.03 - 2017.05
<ul style="list-style-type: none">• 实验室与国家气象局横向项目, 为风云4号气象监测系统提供系统性能评估与异常预警等功能• 使用隐马尔科夫链模型构建时间序列隐特征, 并融合统计特征, 利用层次聚类的方式实现时间序列的无监督标注; 采用决策树和 CNN 实现时间序列的分类		

竞赛与奖励

• ACM-ICPC 亚洲区预选赛上海赛区	铜奖	2014年
• ACM-ICPC 西安邀请赛	银奖	2014年
• 北京邮电大学 ACM 程序设计竞赛	银奖 (银奖第一)	2016年
• 全国大学生数学建模竞赛	北京市二等奖	2013年