



求职意向：数据研发工程师

教育背景

- 2015.09-至今 ◆ 北京大学 (前 5%) 硕士 软件工程 (数据挖掘) 北京大学三好学生、五四奖学金
- 2011.09-2015.07 ◆ 东北师范大学 (前 5%) 学士 计算机科学与技术 优秀毕业生、连续四年校长奖学金

实习经历

- 2017.04-至今 ◆ 微软亚洲研究院 创新工程组 研发实习生
- 1、微软小英是 msra 发布的英语口语学习软件。随着用户量增加到 90W+, 负责微软小英后台数据工作, 包括将 Azure 数据源导入分布式存储平台(HDInsight), 使用 Hive 进行数据分析, 为产品提供决策依据;
- 2、为小英建立基于 Azure 的 Data Factory, 实现从 Azure Blob Storage 到 HDInsight 的用户排名计算功能, 引入 Redis 内存数据库提升 pipeline 流程;
- 3、参加微软 hackathon 比赛, 完成小英测试公众号的“看图聊天”功能: 接收用户图片, 基于 cognitive service 识别情绪, 对识别结果进行图片或声音匹配, 以对话形式返回给用户;
- 4、实现微软小英第三方分享接口以及维护“看脸起名”、“消消乐”等模块。
- 2016.09-2017.02 ◆ 猿辅导 Pipeline 数据组 数据分析实习生
- 1、针对每日活跃用户, 进行用户画像, 具体实现使用 Pig 获取平台活动数据, 根据用户字典筛选目标用户, 完成进一步的推荐课程;
- 2、使用 HiveQL 获取营销活动的 pv, uv 以及课程数据, 为产品运营提供数据支撑。
- 2016.07-2016.08 ◆ 京东集团 京锐夏令营 实习生
- 1、通过数据挖掘提升服饰家居部门优惠券的使用情况, 挽回流失用户。针对平台消费者数据, 提取相关特征, 采用决策树模型定位用户, 经过 A/B test 提高到 0.2%的使用效果;
- 2、为客服部门设计小游戏提高用户调研的访问效率, 单个用户的调研时间从每人 10min 降到 3min。

比赛经历

- 2016.10-2016.12 ◆ 阿里天池大数据竞赛: O2O 优惠券使用预测
- 主要负责数据预处理以及算法设计, 根据用户三个半月内的消费记录, 提取用户、优惠券以及商家特征, 初赛使用 xgboost 算法, 复赛在 odps 平台上使用 GBDT 模型和数据规律相结合, 预测准确率达到 73%, 在 1505 支队伍获得第 45 名 (团队共 2 人)。
- 2017.04-2017.05 ◆ 华为软件精英挑战赛: 大视频时代布局问题
- 独立负责算法设计, 将服务器选址问题转化为图论中的 Capacitated Fixed-Charge Network 问题, 使用最小费用最大流以及启发式算法, 在京津赛区 500 多支队伍取得 66 名成绩。

兴趣爱好

热爱游泳、骑行、阅读 (年阅读量 50+) 资深海贼迷
通过大学英语四级 (CET4) 英语六级 (CET6)