

投递编号	申请职位	投递时间
18082081159304	2 搜狐视频-算法开发工程师	2018-08-20 16:55:49
18082064634250	1 智能平台-NLP算法工程师	2018-08-20 16:55:49

郭彤

男 25岁 西安 硕士 | 在校生
18706806085 740138238@qq.com



教育经历

2016.09 - 2019.04 硕士 西北工业大学 计算机科学与技术 专业排名前1%
2012.09 - 2016.06 本科 西北工业大学 计算机科学与技术 专业排名5%~20%

校内经历

2018.07 - 至今 IFLYTEK 科大讯飞 校园大使
信息收集，校内宣传，现场执行等
作为联络人负责活动组织策划及宣传
2017.11 - 至今 中国计算机学会(CCF) 高校宣传员
主要职责是在本校宣传CCF各类活动、通过微信群及QQ群对CCF进行宣传。
2016.09 - 至今 西北工业大学计算机学院 就业助管
策划和组织就业相关活动40余场，同时负责统计分析当年就业数据
2014.06 - 2014.06 志愿者活动 志愿者
参与ACM-ICPC全国邀请赛活动并获评“优秀志愿者”

项目经历

2018.03 - 2018.04 IJCAI'18 阿里巴巴国际广告算法大赛 个人参赛
项目内容：本次比赛以阿里电商广告为研究对象，提供了淘宝平台的海量真实交易数据。要求参赛选手构建预测模型预估用户的购买意向，即给定广告点击相关的用户(user)、广告商品(ad)、检索词(query)、上下文(context)、店铺(shop)等信息的条件下预测广告产生购买行为的概率(pCVR)。形式化定义：
$$pCVR = P(\text{conversion}=1 \mid \text{query}, \text{user}, \text{ad}, \text{context}, \text{shop})$$

实施方案：基于给定的比赛数据，经过数据清洗和预处理，围绕用户、商品、店铺以及上下文等四个实体来提取特征，得到包括各个实体的基础特征，交叉特征以及排序特征。通过 GBDT 模型 (Gradient Boosting Decision Tree) 进行特征变换，最后利用 LR 模型(Logistic Regression)作为预测模型
项目成果：初赛排名 Top 8% (共 5200 余支参赛队)
2017.09 - 至今 研究生期间合作项目成果 合作完成
《Modeling & Forecasting the Popularity Evolution of Mobile Apps: A Multivariate Hawkes Process Approach》学生二作 · ACM UbiComp'19 · CCF-A 类 (under revision)
《Enhancing Mobile App User Understanding & Marketing with Heterogeneous Crowdsourced Data》学生二作 · IEEE Transactions on Big Data · SCI 检索期刊论文 (under revision)

《Mining and Analyzing User Feedback from App Reviews: An Econometric Approach》 第一作者 · IEEE UIC'18 · EI 检索会议论文

2017.02 - 2017.05 面向连锁药店的会员管理方案研究-高校横向预研项目 学生负责人

合作单位：上海科瓴(CareLinker)医疗科技有限公司

项目内容：利用连锁药店提供的数据(包括会员信息、购药记录、药品数据等)进行挖掘建模，预测新会员转化为重要会员的可能性。

实施方案：鉴于传统 RFM 模型的缺陷，该项目利用会员相关数据构建有效的用户表示。具体来说，将会员相关数据转化为文档，文档内容主要由用户个人信息和历史购药记录组成，其中每条购药记录转换为对应的药品类别，最终将原问题转化为文本分类任务

项目成果：将该模型在 5 家试点药店中进行实地验证，半个月平均客频次增加了 30%左右，有效提高药店盈利能力，帮助药店实现精准营销

2016.11 - 2017.07 基于旅游社交数据的分析挖掘研究-国家自然科学基金 项目核心成员

项目内容：随着人们旅游出行需求的高速增长，国内的众多旅游网站(蚂蜂窝，百度旅游等)产生了大量旅游相关数据，主要包括用户评论和游记。通过挖掘旅行评论和游记为用户提供准确的旅行参考，同时进一步丰富旅游网站的现有内容

实施方案：编写爬虫程序获取国内热门景点的数据(包括评论、游记和景点基本信息)，利用“景观-特征刻画”模式从多个角度对景点进行细粒度的刻画，挖掘用户关注热点，进而通过序列模式挖掘算法从历史游记数据中得到多数用户偏好的旅游路线，最后基于评论文本和图像上下文实现跨媒体信息关联。

关键技术：网络爬虫，文本分析挖掘，图像聚类

项目成果：

《CrowdTravel: Leveraging Heterogeneous Crowdsourced Data for Scenic Spot Profiling and Recommendation》 第一作者 · PCM 2016 · EI检索会议论文

《多源社交数据融合的多角度旅游信息感知》 第一作者 · 浙江大学学报 · EI检索期刊论文

《CrowdTravel: Scenic Spot Profiling by Using Heterogeneous Crowdsourced Data》 第一作者 JAIHC · SCI检索期刊论文

“多源社交数据融合的多角度旅游信息感知与智能推荐方法” 第一学生发明人 · 发明专利

■ 获奖情况

2017.11 国家奖学金

2017.10 西北工业大学研究生一等奖学金

2017.10 西北工业大学 校“优秀研究生”

2017.07 中国研究生智慧城市技术与创意设计大赛校内赛 一等奖

2016.09 西北工业大学研究生一等奖学金

2016.06 本科期间 国家励志奖学金(1次)，校“优秀学生”(1次)，院“优秀学生”(2次)，C语言程序竞赛一等奖

■ 个人技能

熟练掌握 Python、C、Java 等编程语言，掌握常用数据结构、算法

熟悉机器学习基本算法模型，熟练使用 scikit-learn 学习框架，对深度学习有基本了解

熟悉文本处理方法，熟练使用文本挖掘工具

英语水平 CET-6，具有良好的英文文献阅读及写作能力，发表多篇英文学术论文并做口头报告

■ 论文发表

《CrowdTravel: Leveraging Heterogeneous Crowdsourced Data for Scenic Spot Profiling and

Recommendation》 第一作者 · PCM 2016 · EI 检索会议论文

《多源社交数据融合的多角度旅游信息感知》 第一作者 · 浙江大学学报 · EI 检索期刊论文

《CrowdTravel: Scenic Spot Profiling by Using Heterogeneous Crowdsourced Data》 第一作者

JAIHC · SCI 检索期刊论文

《Modeling & Forecasting the Popularity Evolution of Mobile Apps: A Multivariate Hawkes Process Approach》

学生二作 · ACM UbiComp'19 · CCF-A 类 (under revision)

《Enhancing Mobile App User Understanding & Marketing with Heterogeneous Crowdsourced Data》 学生二作 · IEEE Transactions on Big Data · SCI 检索期刊论文 (under revision)

《Mining and Analyzing User Feedback from App Reviews: An Econometric Approach》

第一作者 · IEEE UIC'18 · EI 检索会议论文

“多源社交数据融合的多角度旅游信息感知与智能推荐方法” 第一学生发明人 · 发明专利

■ 附件作品

[简历.pdf](#)