

教育背景

北京邮电大学	本科	计算机学院	计算机科学与技术	2014-2018
北京邮电大学	硕士	网络与交换技术国家重点实验室	计算机科学与技术	2018-至今

个人技能

- ◆ 使用Java编程语言，了解JVM的内存模型与GC机制的基本原理，具有Web开发框架的使用经验
- ◆ 熟悉Redis基本操作，理解持久化（RDB、AOF）、主从复制、哨兵模式
- ◆ 掌握数据结构知识，熟悉常用算法。研究方向是深度学习但是更习惯做工程架构设计
- ◆ 有深度学习框架(python)、消息队列、kv数据库、RPC框架的使用经验
- ◆ 英语：六级 508

实习经历

阿里巴巴-淘系技术部	后台开发实习生(Java)	2020.06~2020.08
------------	---------------	-----------------

职能 实习于导购营销团队，致力于在搜索、推荐、短信召回等业务场景下，增加商品的曝光率与拍下率。

主要参与需求

- ◆ 算法对接类：将线上实时数据或者原本由工程化方法实时推荐的内容沉积到算法侧，提升线上推荐算法的性能。涉及内容包括：行为日志埋点，实时链路异步化改造，定时任务数据补齐/更新等。
- ◆ 运营工具开发：设计并实现通用型运营工具例如配置管理后台、保真打标后台等，解决现有场景业务痛点，提升团队开发效率

项目经历

短信智能分析平台	负责算法模块与服务接口的研发	2017.07 – 2018.08
----------	----------------	-------------------

平台提供实时和离线两种服务，实时服务对短信文本应用机器学习模型分类审核、过滤后中转发送；离线服务针对累积数据进行模板提取、信息抽取以及文本分类。

- ◆ 实现以ActiveMQ为消息中间件、Redis为缓存的离线分析任务处理历史的批量短信数据。
- ◆ 实现了基于句频统计和字典树的短信模板提取方法，作为人工标注的输入，有效减少标注时耗
- ◆ 实现了基于XGBoost的短信行业分类与风险分类模块，附带基于词典与匹配的信息提取功能
- ◆ 使用Spring Boot开发实时服务接口，对一条或多条短信文本提供文本分类、模板匹配、敏感词过滤等服务，并追加了LRU缓存机制。

智能商标审核系统	后端模块研发	2018.09 – 2018.12
----------	--------	-------------------

系统对用户提交的商标名称提供包括近似检索、精确检索、审核通过率评分等服务。

- ◆ 负责商标近似检索模块的研发，以Redis中间件为主要数据缓存，依照多项业务需求进行筛选排查
- ◆ 以历史商标与申请商标的多项近似度为主要特征，训练XGBoost机器学习模型对商标近似度进行打分

金融文本信息抽取系统	平台后端研发、系统核心算法研发	2019.04 – 至今
------------	-----------------	--------------

解析上市公司的招股说明书、年报等pdf文件，通过命名实体识别、关系提取，生成结构化结果进行存储。同时提供数据审核平台方便人工操作，迭代产生数据

- ◆ 审核平台后端采用Spring Boot+Jpa+Mysql研发，支持对提交的标注数据进行检错与自动校正。
- ◆ 使用Tensorflow实现命名实体识别模块，并进行工程化的调整和改动
- ◆ 实现具有文本提取、数据对齐与清洗能力的PDF解析模块，串接信息抽取模块，解析结果自动存为备审数据，结合审核平台快速迭代产生新数据集

荣誉奖励

◆	第十届北京邮电大学程序设计竞赛	金奖	2016.04
◆	CCPC中国大学生程序设计竞赛（合肥）	银奖	2016.10
◆	北京邮电大学一等奖学金		2018.10