

男 / 1994.2

重庆邮电大学 · 软件工程

本科 / 2017年毕业

github.com/haibincoderhaibincoder@outlook.com

18883880458

工作经历

2018.07 ~ 至今 中国平安人寿保险股份有限公司

工作

人工智能研发团队 NLP策略算法

2017.07 ~ 2018.07 海能达通信股份有限公司

工作

指挥与调度产品线

2016.06 ~ 2016.10 德勤全球交付中心

实习

数字咨询-Consulting Digital

专业技能

- 主要研究NLP方向智慧客服、舆情分析、知识图谱等，对推荐、搜索等领域有所了解。参与多个项目0-1落地，良好的工程能力。
- 多轮任务对话：主要研究 意图识别、槽位提取、对话管理、对话生成技术，对 DM模型、时序多分类、End-to-end 方法有研究，有实际落地项目。
- 问答系统：主要研究 FAQ、KBQA、知识图谱等，对 Elastic Search、Lucene、Annoy 等召回有研究，对 DSSM、Siamase Network 语义相似度方法有研究和项目落地。
- 算法基础：了解常用分词、词性标注、实体识别、分类方法，有 纠错模型、指代消解 实际应用，对 TextCNN、Lstm、CRF、FastText、StarSpace 等常用模型有应用。有 数据挖掘、舆情分析 实际项目经验。
- 工程方向：主要使用 Spring boot、Spring Cloud，熟悉 MQ、Zookeeper、Redis、Mysql 等常用技术。

项目经验

- 2019.1-今 AI视频柜员 [Link](#)
负责算法落地 意图识别 多路召回重排
参与剧本生成、意图识别、对话管理模块设计和上线，主要使用 Star Space FastText 等实现意图分类，参与 通用分词、实体识别、实体对齐、纠错模型、指代消解 等预处理模块相关工作，工程部分负责对接多个task bot、QA bot、chat bot。

项目经验

2018.7-今 金管家智慧客服 [Link](#)

任务型对话 FAQ KBQA

负责算法落地和工程开发，策略部分负责任务型bot和QA模块优化，工程部分负责spring cloud开发和上线。

Task Bot主要负责研究 分类模型 时序分类 匹配分类 等方法，有限状态机 对话管理模型 强化学习 等对话管理方法，模板生成 seq2seq 等对话生成方法。

QA Bot主要研究 Lucene ES 等字面召回和 LSTM DSSM Siamase 等语义相似度计算方法。负责维护原有KBQA系统。接手项目后，Task负责每日支撑上亿流水任务办理，FAQ部分NLP准确率90%，KBQA部分维护十万级实体-关系对。

2017.12-2018.7 Hytera ICC 2.0

深度学习 自然语言处理 命名实体 文本分类

智能接处警系统，负责自动填单功能，实现流程自动化助手，参与客户端开发，对句子文本进行分类和实体识别，系统采用 Python Django WPF 搭建，使用 Tensorflow 实现 Text-cnn 并优化网络，使用 pyltp jieba SVD k-means word2vec 进行数据预处理，熟悉 数据预处理。7个一级分类f1分数可达96%，67个二级分类平均f1分数可达86%，独立实现地址输入联想等功能。

2017.3-2017.6 两会舆情监控项目 [Link](#)

数据挖掘 聚类 分类 机器学习 舆情分析

重庆社科院项目，分析两会政策对不同年龄、不同阶级的影响。负责头条新闻抓取以及数据分析，使用 python 爬取数据50w+，爬取效率约300条/分钟，使用 request 实现模拟登陆，使用 BeautifulSoup4 解析网页，使用 编辑距离 奇异值分解 K-Means 等算法进行数据分析。通过聚类对爬取新闻进行分类以及标注，输出两会政策对不同年龄、不同阶级的影响报告。

2017.2-2017.5 自媒体舆情分析及监控 [Link](#)

非监督学习 数据挖掘 微博、微信 舆情分析

使用 Java 爬取微博、微信平台数据100w+，使用 selenium 模拟登陆，使用 jsoup 解析网页，搭建 Hadoop 平台并使用 MapReduce 进行数据分析。构建词向量，根据关键词对自媒体数据进行情感标注。

2016.6-2016.10 Deloitte Knowledge Management

德勤GDC ERP 项目，全程参与项目开发，负责 数据挖掘、搜索算法、推荐系统 开发及优化。