**安磊**

(+86)15952034742 | [anleial666@163.com](mailto:anleial666@163.com)

期望职位：NLP算法工程师 | 目标薪资：25k

**教育经历**

**哈尔滨工业大学**（焊接技术与工程专业） **本科** 2011-2015

**南京大学**（软件工程专业） **硕士** 2017-2019

在校期间学习了机器学习、自然语言处理、数据挖掘等课程，并在实验室中参与法院项目组的类案精准推荐项目的研发，增强了理论与动手能力。

**工作经历**

**南京晨光集团有限责任公司 工艺主管**  2015-2017

负责部分军工型号产品的焊接工艺指导

**星环信息科技（上海）有限公司 NLP算法工程师** 2018.5 -至今

负责知识图谱产品中的图谱构建、信息抽取等NLP算法研发；负责文本分类、情感分析、命名实体识别、关系抽取等相关工作

**项目经历**

**类案精准推荐系统 2017**

构建一个类案精准推荐系统，根据法官对案件的描述进行语义分析，从数据库中检索语义相关的历史案件及断案信息来辅助法官进行案件评审。

**Scrapy爬虫系统** **2018**

通过scrapy框架搭建爬虫系统，从百度百科、新浪财经、巨潮资讯、同花顺等金融网站爬取各种结构化、非结构化文本，为后面构建知识图谱提供数据保证，是构建知识图谱的pipline中的第一步。

**信息抽取任务 2018-2019**

在爬取非结构化文本数据后，先进行数据清洗，再将命名实体识别模型和关系抽取模型组合并结合Flask框架包装成信息抽取服务（ModelServer），在处理非结构化文本时调用信息抽取服务将文本信息转化成“实体-关系-实体”的三元组数据，形成构建知识图谱的知识库。

**平台NLP算子开发 2018-2019**

负责公司分布式人工智能平台中部分NLP算子开发，包括分词、文本摘要、情感分析、命名实体识别、文本分类、关系抽取等。其中使用Bi-LSTM+CNN+CRF结合各种领域的专业词汇实现垂直领域的中文分词，使用TextRank实现文本摘要算法，使用Bi-LSTM实现短文本情感分析等，最后用Scala编写成算子。

**Oncall问答机器人 2019**

根据日常Oncall整理的问题-答案对，通过词库匹配+实体识别识别新的提问中的实体信息，再对数据库中的问题进行实体匹配+语义相似度匹配来查询问题并返回对应答案。

**中行知识图谱项目 2019**

对中行内部的CTIS数据进行数据清洗、数据挖掘，根据行方的关注企业和相关交易建立企业关系图谱，从大量交易信息找到更多的关联企业并对各个企业的新闻舆情进行舆情分析来定制关联企业的预警功能。

**交行问答项目 2019**

对交行问答系统的标准问题库中相似问题进行合并：

1. 采用TF-IDF+相似度计算来计算问题的语义相似度；
2. 利用BERT获得问题的句向量，对句向量进行相似度计算获得TOP N的标准问题作为候选相似问题库，再对N个标准问题对应的答案进行主题分析，将答案主题相似的问题-答案对进行合并。

**技术栈**

熟悉Python、Tensorflow、sql、机器学习、深度学习

了解Java、Scala、分布式编程