

支龙

(+86) 155 0465 3947 ◇ 1427979142@qq.com

教育经历

哈尔滨工业大学 (学士学位)

2015 年 9 月 - 2019 年 6 月

核心课程: 算法设计与分析 (89) / 操作系统 (91) / 计算机组成原理 (82) / 数据库系统 (91)

计算机网络 (82) / 高级语言程序设计 (100)

专业: 计算机科学与技术 专业绩点: 3.24/4

复旦大学 (硕士学位)

2019 年 9 月 - 至今

专业: 计算机技术

方向: 自然语言处理, 半监督学习

CET-6

语言和技能

C/Java/Python, Git, Linux

项目经历

Rasa 中文智能问答平台 (Python)

1 月 2020 - 4 月 2020

- 该项目基于西门子公司提供的燃机材料数据以及业务知识, 为西门子提供一套智能的业务问答系统。
- 主要负责中文 Rasa 开源框架的改进与开发工作。开发了新的中文意图识别组件, 改进了新的文本学习方式, 将意图识别准确率从 40% 提升到了 90% 以上。针对西门子业务需求, 提供了识别复杂专有名词的识别功能, 修复了中文实体识别错误的问题。

商铺地址爬虫 (Python)

11 月 2019 - 1 月 2020

- 该工作用以实验室地址图谱构建项目的数据库支持, 用以获取一个生活服务平台上各省市商铺具体地址, 经纬度等信息。负责爬虫的开发, 为项目提供了数十万条商铺地址数据。
- 基于 Python 和 requests 库, 通过多线程与代理请求的方式解决了点评 IP 检测封锁的问题, 提高爬取效率。破解了 SVG 矢量图文字加密, 经纬度加密等反爬虫机制。提供了爬虫重启恢复, 错误日志记录方案以优化爬虫效率。

西门子知识图谱搭建系统 (Java, SpringBoot)

10 月 2019 - 12 月 2020

- 该项目主要用来为西门子的业务文档进行知识抽取并自动化地搭建知识图谱, 并且为西门子开发了一个基于 SpringBoot 开发的 Web 客户端, 提供了知识图谱展示, 知识查找与文档上传解析等功能。
- 主要负责 Web 客户端的协助开发。使用了 Bootstrap 作为前端框架, 使用 Ajax 异步请求的方式展示知识图谱动态矢量图, 可以在同一页面完成图谱的更新展示。改善了知识查找模块, 对不同的答案做出不一样的响应, 结果展示更加人性化。

获奖经历

研究生一等奖学金

2020 年度