



吴新华

15001839216

alicewoo_hi@163.com

女|26 岁 (1992/08/07)|现居住上海-闵行区|1.5 年工作经验

最近工作

最高学历/学位

职位:	nlp 算法工程师	专业:	计算机科学与技术
公司:	数库(上海)科技有限公司	学校:	上海海事大学
行业:	计算机软件	学历/学位:	硕士

自我评价:

1. 熟练使用 Python 编程开发, 常用 Git 代码管理工具以及 MySQL 数据库;
2. 掌握自然语言处理领域的基础理论和方法, 有分词、文本分类、信息抽取的项目经验;
3. 对机器学习深度学习有了解, 积累了一定的文本数据处理方面项目开发经验。

工作经验

2017/10-2019/2 **nlp 算法工程师**|nlp
数库(上海)科技有限公司 [1 年 4 个月]
计算机软件|外资(非欧美)

工作描述:

1. 运用自然语言处理技术解决具体业务场景中的信息抽取, 文本分类等实际需求;
2. 运用机器学习等算法对新闻资讯进行不同维度的标签分类;
3. 负责负责语料库构建与维护, 金融领域文本挖掘算法、自然语言处理相关技术和方法的研究

2017/7-2017/10 **开发工程师(数据分析方向)**|战略中心
上海德拓信息技术股份有限公司 [3 个月]
计算机软件|150-500 人|民营企业

工作描述:

- 1.通过 Python 编程调用 sklearn 的机器学习方法对海量的招标数据按照行业进行分类,对每个招标信息预测行业标签.为业务部门提供有价值的数据;
- 2.对公司人事管理系统 itbasic 进行关键词分析。抽取 outbasic 数据库中相应的表和字段数据, 对数据进行分词并统计词频, 实现词云展示。

项目经验

2018/11-至今 **嘉实基金-定期报告运营指标提取**

所属公司：数库（上海）科技有限公司

项目描述：提取定期报告中各项运营指标数据。

对解析好的文本数据进行数据预处理，用标注语料训练一个分词标注器，对文本进行分词并给词打上角色标签，根据不同的角色标签从而提取出句子中的每个数值对应的指标，单位，时间，公司以及产品等信息。

2018/7-2018/10 **同名简历对比**

所属公司：数库（上海）科技有限公司

项目描述：对同名的简历进行相似度对比以及通过命名实体识别出简历中人物曾任职公司和职位，综合多方因素判断同名简历是否为同一个人所有，方便业务人员维护高管简历数据库

2018/3-2018/6 **资讯平台项目**

所属公司：数库（上海）科技有限公司

项目描述：新闻资讯主题、行业和产品标签分类。

训练分类模型，对新闻资讯从主题、行业、产品等维度进行分类，打上不同维度的标签，从而为用户推荐他需要关注的新闻，过滤无用资讯。

2017/10-2018/2 **国泰君安 AI 文档项目**

所属公司：数库（上海）科技有限公司

项目描述：主要任务是从招股说明书和债券募集书中自动化的提取出特定的 25 个财务科目的具体年份和对应的值等信息。其中数据包含文本和表格中的数据，针对文本和表格设计不同的提取算法，训练语义角色标注模型，建立相应提取关系从而输出标准化结果，最后实现相同科目相同报告期的数据自动核对校验，进而发现可能的错误数据。

教育经历

2015/9-2017/7	上海海事大学	研究生硕士	计算机科学与技术
2011/9-2015/7	鞍山师范学院	本科学士	电子科学与技术

技能特长

Python、Git、MySQL、MongoDB