

投递编号	申请职位	期望实习城市	投递时间
19030325377208	1 算法研发工程师（当前查看职位）	北京	2019-03-03 22:00:29

# 叶成



男 25岁 上海 硕士 | 在校生  
18018590558 2838065383@qq.com  
身份证 32072319940918385X  
籍贯：江苏连云港 政治面貌：中共党员

## 教育经历

2017.09 - 2020.03 硕士 上海大学 控制工程 专业排名1%~5%  
2013.09 - 2017.06 本科 淮阴工学院 自动化 专业排名前1%

## 工作经历

2018.09 - 2018.12 百度（中国）有限公司 测试开发实习生  
内容：1、基于django自动化测试平台工具开发，提升测试效率与质量  
2、执行测试用例，负责输入法语音识别结果测试  
业绩：● 增加前端按钮支持语音识别结果的删除与编辑，比之前只能进行操作数据库提升4~5倍效率；  
● jenkins 自动编译输入法apk完善并支持输入参数以完成版本需求，无需每次使用vim编辑配置文件；  
● 支持 android 测试工具上传音频到数据库，省去每次都要整理一堆音频再统一使用django API上传；  
● 人工编辑语音识别的结果会相应的保存到数据库里，以后出现相同问题会自动修改结果；  
● 前端支持测试项目参数修改、评测维度中途新增与删除、音频下载、评测结果生成表格下载；  
2018.06 - 2018.09 上海寰融信息技术有限公司 自然语言处理实习生  
内容：1、阅读前沿NLP相关英文文献，实现该文献算法模型；并将算法应用到公司数据库；  
2、比较各种NLP 任务下如relation extraction不同模型的性能；  
业绩：● 利用pytorch对金融研报建立问答系统，并对研报总结摘要，关系提取，语义解析。主要运用到了BiLSTM Encoding、BiLSTM Decoding、Self-Attention、Dual Coattention、Multi-head Decoder Attention、Context and Question Attention、Multi-Pointer-Generator 关键技术。  
● 完成knowledge graph 中的entity and relationship extraction部分，提取两个结点的relation作为它们的边；利用tensorflow对knowledge 和文本联合表示学习并建立相互关注机制，提取所给的结点之间的所有关系；

## 项目经历

2018.08 - 2018.09 智慧中国杯之达观数据竞赛 算法工程师  
描述：● 目的：建立模型通过长文本数据正文article，预测文本对应的类别class  
● 建立ML模型，利用gensim提取文本的tfidf，LDA等，sklearn孤立森林去除离群值，SVM进行特征选择，mlxtend 融合LR、SVM、KNN、random forest、Xgboost、Lightgbm单模型，融合方式为stacking和voting；  
● 建立DL模型，参考Tsinghua University于2016发表的论文，gensim生成word2vec词向量，theano工具构建分层次LSTM模型生成句子和文档级别表示，添加局部attention机制，生成最终文档的表示，最后输入softmax层；  
业绩：模型平均f1-score 为 0.79，最高分为0.80；  
项目链接：<https://github.com/cheng-Ye/daguan-competition-master>  
2018.04 - 2018.05 基于人工智能的药物分子筛选 算法工程师

描述：●目的：预测出致病蛋白与小分子的亲和力值；

●对数据进行分析，使用随机森林填充缺失值，构建蛋白质序列词袋向量；

●分别使用lightgbm、xgboost、随机森林、支持向量机、神经网络以及K邻近预测亲和力值，最后采用stacking融合使用模型；

业绩：●优胜奖（第10名/成功提交共920支队伍）

---

## ■ 获奖情况

2018.11 “华为杯”中国研究生数学建模竞赛全国三等奖

---

## ■ 个人技能

英语技能：英语四、六级

语言技能：Python、C++；

专业技能：熟悉MySQL、Docker，熟练使用Linux命令、git命令，熟练使用

TensorFlow，Pytorch，Keras，NLTK，Gensim，Sklearn，pandas等nlp相关Python库；

理论知识：熟悉NLP、机器学习相关算法理论，如Xgboost、SVM、Seq2Seq、LSTM、CNN、各种attention等；