

柴文军

面试岗位：python 应用开发

个人资料

年龄：26

邮箱：1757808781@qq.com

学 历：本科

性别：男

电话：18326025526

工作时长：2 年

个人能力

- 1) 精通 python 语言，熟悉 Scrapy 爬虫框架；
- 2) 熟悉线性回归、逻辑回归、SVM、HMM 等机器学习算法；
- 3) 熟悉深度学习 CNN、RNN、LSTM 有关算法的原理；
- 4) 熟悉 sklearn 机器学习及 TensorFlow 深度学习框架的使用；
- 5) 熟悉 NumPy、pandas、matplotlib 进行数据分析和可视化展示；
- 6) 熟悉 MySQL、redis 等数据库；
- 7) 熟悉 Linux 环境；

项目经验

项目名称：用能设备命名实体识别

项目时间：2018.06-2018.12

项目描述：该项目是对上一个项目升级的实验项目，对每个类别的文档建立一个模型，识别出文档中设备的信息：型号，功率，能耗等信息。目的是提取出文档中的关键信息，方便查看。尝试利用 TensorFlow 深度学习框架构建 BiLSTM-CRF 模型来完成这个任务。

项目职责：

- 8) 文本数据预处理和清洗，制作数据集；
- 9) 对文本进行 jieba 分词，并把分词信息 embedding 到其中；
- 10) 利用 TensorFlow 深度学习框架构建 BiLSTM-CRF 算法模型进行训练；
- 11) 利用 Viterbi 算法对模型预测结果进行解码；
- 12) 模型评估：精确率、召回率、f1 参数分析；

项目名称：基于 NLP 自然语言构建的文档自动分类系统

项目时间：2019.03-2019.07

项目描述：对公司文档(电脑)进行自动分类，方便管理。

项目职责：

- 13) 获取公司内部和网上的电脑文档
- 14) 根据文档信息，使用 Label Encoder 编码实现文章类别标签化
- 15) 通过 jieba 对每篇文章进行分词处理，并去除停用词
- 16) 使用 TFIDF-Vectorizer 计算 TF-IDF 词袋模型、对所有文章构建 word2vec 模型
- 17) 分别对 TFIDF 词袋模型、word2vec 词向量模型构建模型分析
- 18) 模型评估：精确率、召回率、f1 参数分析

项目名称：无人银行智脑技术创新研发与应用系统项目

项目时间：2019.10-2020.3

项目描述：对公司业务和系统所提出的问题给出相应的回复(问答机器人)

项目职责：

- 19) 理解业务：分为知识图谱处理公司业务和生成式模型处理闲聊
- 20) 知识图谱：将已经整理好的 excel 业务表导入 neo4j，根据所问问题去图谱中搜寻答案
- 21) 生成式模型：使用 seq2seq 模型对已经处理好的数据进行训练得到相应的模型
- 22) 生成式模型答案评估，得分不足的用检索式模型或固定模板来补充
- 23) 图谱与模型的结合，确保问答速度能够秒回
- 24) 参与项目的上线

工作经历

工作时间：2017.12-2019.7

公司名称：微盟电子(昆山)有限公司

所任职位：算法工程师

工作内容：

- 25) 基于 python 脚本语言，使用 NumPy、pandas 等库对数据进行清洗、预处理、特征提取、特征转换等；
- 26) 熟悉常用机器学习算法模型评估及优化方法，对算法结果进行评估；
- 27) 理解 TensorFlow 框架下深度学习的 CNN、RNN 等算法的代码，并尝试使用这些算法解决手中的工作；
- 28) 在自然语言处理方面，对于语言模型如 LSTM、seq2seq 等都有自己的理解；

工作时间：2019.10-2020.3

公司名称：上海文思海辉技术有限公司

所任职位：设计开发工程师

工作内容：

- 29) 设计问答系统（知识图谱与模型结合）；
- 30) 将所设计的框架用代码实现；
- 31) 系统回复的实用性（在图谱和模型之外加入检索和模板，以保证系统回复实际应用效果）

自我评价

- 32) 有较强的沟通能力，团队协作能力和快速学习能力
- 33) 热爱编程和算法，乐于提出、分析和解决问题
- 34) 思路清晰，具备较强的表达能力，工作主动积极
- 35) 具有较强的适应能力，并且具有很强的团队精神跟责任心，善于与人沟通

教育经历

2014.09-2018.06 安徽建筑大学