

个人简历

我一直在努力成为更好的自己

基本信息

姓名：郭东杰 年龄：27岁
性别：男 民族：汉族
籍贯：广东广州 工作年限：5年经验
电话：15875111626 邮箱：tel15875111626@163.com

求职岗位

职位：AI算法工程师 城市：广州 期望薪资：面议 随时到岗

教育背景

2016-09 ~ 2020-07 广东白云学院 计算机应用（本科）

技能特长

- 基础工作能力**
具备扎实的 Python 基础知识和编程经验，熟悉Linux系统，具有良好的数据分析、挖掘能力
熟悉NLP算法、机械学习算法、Yolo目标检测算法等
- 大模型相关**
拥有RAG、知识图谱企业级开发经验，熟悉RAG优化检索召回方法
熟悉使用Langchain、transformer等大模型应用开发框架
熟悉FunctionCall和Crewai等agent常用框架
熟悉Lora微调，P-tuning等模型微调技术
熟悉模型优化技术，如正则化、超参数调优、混合精度等
熟悉知识蒸馏、量化、剪枝等推理加速技术
- 数据处理**
熟悉pandas、numpy、matplotlib等数据分析常用库
熟悉Spark分布式数据库以及PostgreSQL关系型数据库，拥有亿级企业数据处理经验
熟悉爬虫使用，拥有协程高效爬取网络数据的经验

荣誉证书

- 华为云人工智能HCCDP工作级开发者认证
- 阿里云Prompt提示词与RAG知识库专项技能认证
- 大学英语六级

工作经验

2023-04 ~ 2025-04 国电南瑞股份有限公司 NLP算法工程师

- 构建RAG和知识图谱，搭建企业级QA系统
- 熟悉大模型微调(如Lora、P-tuning等)、以及推理加速技术如知识蒸馏、量化等，让模型更好的适应项目需求
- 使用Function Call等技术进行agent系统的构建
- 采用Yolo目标检测算法以及NLP技术实现系统优化
- 进行数据挖掘，提升数据质量

2022-05 ~ 2023-04 深圳市蓝思网络技术有限公司 NLP算法工程师

- 使用NLP、机械学习等技术进行推荐、选品系统构建
- 进行多模型精度效果评估
- 进行模型精度、响应速度优化
- 进行数据挖掘，异构数据处理，提升数据质量

- 1、使用爬虫进行数据爬取
- 2、通过大数据技术(Spark、存储过程)生成APP、系统所需数据
- 3、使用kettle等工具进行数据库分层数据抽取
- 4、进行数据清洗、格式处理等，生成算法部门所需数据
- 5、对数据问题进行排查与修复

项目经验

2024-03 ~ 2025-04

南网QA项目

NLP算法工程师

项目背景：
该项目为深圳南网计量QA系统，为计量系统AI化升级项目之一，该项目采用RAG加知识图谱共同构建，供计量内部人员使用，旨在帮助供电局和班组人员提高工作效率，对于系统异常状况进行早发现、快速定位、及时解决，对于一些可能的业务问题，如偷电漏电、停电是否因为电压不稳、电量与表码不符等问题进行推测和回答。

- 主要职责：**
- 负责知识图谱的构建：
和业务专家确认每个模块的相关业务、主要问题、数据来源
将数据库现有数据导入知识图谱
将现有业务文档、图片抽取成文字，通过Chatglm进行三元组抽取，采用Lora微调，F1值87% -> 94.33%
知识融合，主要负责实体消歧，采用Qwen 2.5 7B处理，大量数据消歧采用TF-IDF+余弦相似度
 - RAG构建
将文档内容通过Qwen 2.5 7B模型进行语义分块
将数据向量化后导入Milvus向量数据库
 - 参与QA系统搭建
参与prompt的设计和构建
参与语义槽的构建和设计
参与数据库查询agent的构建

2023-04 ~ 2024-02

南网工单智能分类项目

NLP算法工程师

- 项目背景：
该项目为工单系统优化升级，南网的工单系统，不仅涉及系统本身的问题，还会牵扯到全市表码、电费的计算，因此该工单系统优化也是公司高层和南网高层高度重视的一个项目，目的在于提升工单运维速度和质量。
- 主要职责：
1、将长尾数据通过smoth还有英译中的方式进行加强
2、通过tf-idf+余弦相似度方式找出歧义文本(语义相似但标签不同)，通过专家进行校验
3、对较短文本进行外部特征拼接
4、Roberta模型进行多任务精细分类 F1值94%
5、用TinyBert(学生模型)根据Roberta(教师模型)进行知识蒸馏，F1值：92%
6、用Yolo v5进行图片关键信息定位

2022-10 ~ 2023-04

电商智能选品系统

NLP算法工程师

- 项目背景：
该项目从多个维度分析，挑选和预测最近销售量大增的商品，公司在跨境电商平台拥有多个账号，往常通过人工进行产品铺货，时间长、效率低，因此改用深度学习的方式去融合多种特征进行选品，选品系统每天按照最新的数据，将多种可能的爆款产品以及选品推荐因素呈现在选品师面前，极大的提高了公司的铺货效率与爆款概率。
- 主要职责：
1、异构数据数据处理，包括图片提取文字、不规范文字处理等
2、使用LSTM+Attention对新增购买数、同类商品数、同种商品数等进行评分
3、使用Bert模型进行评论情感分析,F1值92%
4、使用BiLSTM+CRF模型对商品特征进行提取，F1值94%

2022-05 ~ 2022-10

电商智能推荐系统

NLP算法工程师

- 项目背景：
该项目以用户为维度，向用户推荐合适的商品，公司在亚马逊、速卖通等跨境电商平台还有脸书等社交平台拥有多个账号、众多

订阅用户，智能推荐系统通过分析用户的购买历史、浏览记录、季节因素、喜好偏好等进行个性化商品推荐，精准找到目标客户，提高客户购买意愿，减少无效信息，带来更高的用户满意度和销量。

- 主要职责：
 - 1、进行数据清洗与数据结构处理，针对不同的算法，将数据解构成不同的输入格式
 - 2、进行用户分群与偏好识别，采用随机森林算法F1值83%，FastText算法F1值90%，RoBerta模型F1值93%
 - 3、评估商品推荐优先级，GBDT算法F1值76%，ResNet算法F1值88%，Bert模型F1值91.5%
 - 4、进行数据增强，如长尾类别进行smoth插值法，文本较短采用特征拼接等

2020-09 ~ 2022-05

美赞臣电商项目

大数据开发工程师

- 项目背景：

美赞臣是我国知名的奶粉企业，全国拥有NC、导购近万人，该项目主要是帮助美赞臣公司进行大数据管理的，此外也包括为美赞臣算法部门获取和提供已经清洗并格式处理好的相应训练数据。
- 主要职责：
 - 1、使用爬虫进行数据爬取所需数据，如社区评论数据、用户售后评价等
 - 2、将清洗过的数据按照特定格式进行数据处理
 - 3、通过大数据技术生成APP、系统所需数据
 - 4、使用kettle等工具进行数据库分层数据抽取
 - 5、对数据问题进行排查与修复

自我评价

工作认真负责，有良好的编码习惯，热爱不断学习新事物，喜欢开发带来的成就感, 日常热爱登山徒步，在职期间曾组织多次户外登山徒步活动，擅长团队活动策划，拍摄和宣传