

王显富

男 | 24岁 | 籍贯：吉安 | 共产党员 | 17870124538 | 2140717632@qq.com

求职意向：AI工程师 | 期望城市：深圳



教育经历

内蒙古科技大学 本科 软件工程

2020-2024

绩点：3.85/5

专业排名：5/90

主修课程：Python基础、C语言、Java程序设计、MySQL数据库、数据结构、操作系统、计算机组成原理、计算机网络、网页开发高级技术、软件测试技术、软件设计与体系结构、ACM等

个人优势

语言能力 英语CET6

编程语言

- 熟悉Python、C语言、Java基础语法。
- 熟练掌握numpy进行高效的数值计算，使用pandas进行数据分析和处理以及利用matplotlib进行简单的数据库可视化。

算法基础：

- 熟悉KNN、线性回归、逻辑回归、决策树、随机森林、SVM以及K-Means 和DBScan等聚类的机器学习算法流程。
- 掌握神经网络（DNN、CNN、RNN、LSTM、Transformer）、激活函数（ReLU、Sigmoid、Tanh）、损失函数（均方误差、交叉熵）、优化器（Adam、SGD）、正则化（Dropout、L1/L2）等。
- 了解经典视觉神经网络，如LeNet、ALexNet、ResNet、VGG、GoogleNet、MobileNet等。
- 熟悉Sklearn、Pytorch、TensorFlow等框架搭建、训练和调优深度网络模型。

大模型理论基础：

- 了解大模型微调和部署。
- 掌握Ollama本地化部署，以及熟练使用百炼平台、LangChain和扣子等agent workflows构建智能体及应用。
- 熟悉RAG技术，了解大模型结合向量数据库进行检索增强和生成的过程。
- 设计并调优大模型提示词。

工程化技术：

- 熟悉FastAPI构建现代 Web 应用或数据接口，便于处理业务逻辑。
- 熟悉使用Streamlit、vue等前端技术构建基本的响应式页面。
- 熟悉Docker，能够使用Docker容器化技术部署和管理应用。
- 熟悉在云服务器上进行项目的部署，并且可直接通过域名访问。
- 熟练使用git完成代码版本控制和分支管理。个人项目GitHub网址：<https://github.com/wxf20010608>

荣誉奖项

- 阿里云天池比赛第2名（20000+支参赛队伍）。
- 蓝桥杯内蒙古赛区Web应用开发组大学组三等奖（省三）。

工作经历

内容:

- 制定文本分类/图像分割标注规则，熟悉使用labelimg等平台 and 工具对文本、图像和语音进行收集和标注。
- 利用 pandas 清洗文本、图像、语音原始数据（处理缺失值），结合 numpy 进行特征提取与向量化（文本词向量/语音频谱），构建结构化标注数据集。
- 搭建FAQ知识库（覆盖客户高频问题）、推荐商品库（10万+SKU，整合种类/销量/用户偏好），基于LangChain Agent实现语义检索，采用RAG技术增强大模型，将知识库检索结果注入通义千问提示词，推荐准确率提升。
- 使用langchain设计与实现agent工作流和设计RAG（检索增强生成）架构。

业绩:

- 通过模型优化和工具优化，进一步提升模型的泛化能力以及客户问题解决率和满意度。
- 通过数据标注和数据增强，为模型训练提供了高质量的数据支持，支撑大模型意图识别准确率和问题回答准确率有所提高。

项目经历

在线教育平台智能答疑系统

大模型应用工程师

2024.07-2024.12

内容:

项目描述

本项目是为一家在线教育平台开发的智能答疑系统，旨在利用先进的RAG（检索增强生成）技术为学生提供24小时随时可用的智能解答服务，并减少教师的重复性劳动。系统可以自动识别和解析学生问题，从知识库中高效地检索相关信息，并结合自然语言处理技术生成清晰、准确的回答。

技术栈

Python、OpenAI GPT-3.5 API、word-embedding、MathJax、OpenCV、Axios、Jieba、MySQL、Redis、Docker、Nginx、Docker Compose、FastAPI、Vue.js、OCR、numpy、pandas、云服务器等。

项目职责:

负责 10万+条教育知识数据的清洗、结构化与向量化，检索增强生成（RAG）架构设计。

业绩:

- 设计混合检索策略（关键词+向量搜索），通过Jieba分词结合Word-embedding实现语义检索，提升关键词匹配准确率。
- 利用 OCR+ 文本嵌入（word-embedding）处理教材/讲义非结构化数据，知识库扩容至10万+条，覆盖95%课程知识点。
- 基于 FastAPI+Docker 构建微服务架构，支持横向扩展，通过 Nginx负载均衡实现高数量的学生请求，服务可用性高。

多模态智能客服agent平台

大模型应用工程师

2025.03-2025.06

内容:

项目描述:

本项目是基于langchain快速构建的支持文本、图像、语音输入的Agent多模态客服系统。通过集成大模型进行意图识别和智能回复，系统能够处理订单查询、问题咨询和产品推荐等多种业务场景，并且能够高效解答问题与推荐，实现全方位智能客服服务。

技术栈:

Python、OpenAI API、通义千问、云服务器、LangChain、LangGraph、OCR、Web搜索、Excel处理、MySQL、Redis、Docker、Nginx、FastAPI、Streamlit等。

项目职责:

负责数据标注，数据预处理和Agent工作流设计。

业绩:

- 构建了支持3种模态输入的智能客服系统，可以成功处理多种复杂业务场景，支持多种常见客服操作。
- 客户问题解决率和客户满意度增加，人工介入率、客服成本和平均响应时间稍微有所降低。