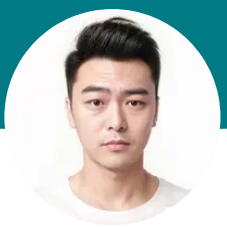

张高峰

男 | 40岁 18937999037 250145682@qq.com

17年工作经验 | 求职意向：大模型算法 | 期望薪资：15-25K | 期望城市：上海

个人优势

1. 基于大模型搭建高效，可扩展的智能体框架，整合知识库、工具调用、记忆模块等能力。优化智能体的交互逻辑、意图识别、上下文管理及错误处理机制，提升用户体验。精通智能体（Agent）开发和RAG开发，熟悉langchain,langgraph,langserve,langsmith联合开发部署。熟悉llamaindex的各个组件，和milvus的联合开发。包含embedding生成，向量检索，rerank算法等关键模块优化，有rag知识库实际项目经验，对知识库调优有成熟方法

2. 具备DIFY和COZE系统的二次开发能力，以及DIFY插件的开发经验。熟悉 docker

3.熟悉大模型微调：量化，剪枝，微调（sft,lora,p-tune,指令微调），熟悉知识蒸馏，分布式多机多卡大模型微调，数据集的构建

4. 熟练进行Python的fastapi，熟悉numpy,pandas等深度学习库。利用影刀RPA技术成功实施多个项目

5. 具备良好的英语能力（CET-6级别），能够阅读和理解常规英文资料，并进行有效英语沟通。

6. 熟悉深度学习中的CNN、RNN算法，掌握YOLO、BERT、Transformer等模型的实际部署与简单开发，熟练使用PyTorch框架。

工作经历

洛阳日辰集团 大模型算法工程师 2024.02-2025.04

内容:

多轮对话简历问答智能问答系统搭建。rpa和大模型配合生成高质量小红书文案，dify 对接 crm 系统，大模型压缩：量化，剪枝，微调（sft,lora,p-tune,指令微调）

业绩:

1、职责： 研发公司的客服系统

业绩：主导大模型算法在智能客服领域的研发与优化工作，大幅提升客户问题识别准确率，大幅提高客服响应效率。响应效率提高200%

2、职责：帮助人力建立智能简历筛选

业绩：参与设计智能简历筛选系统的核心算法，运用自然语言处理技术实现候选人资质精准匹配，筛选效率提升50%。

3、 职责：公司内部 crm 系统智能化改造，并且实现业务 ai 自动出题考试系统

业绩：基于 dify 插件开发对接公司 crm 系统，显著提升系统智能化水平，节约操作时间达 80%。

4、 职责：智能生成小红书文案

业绩：利用 agent 自动生成小红书文案及图片，内容生产效率提升 5 倍，满足商业化需求。

5、职责：提升 dify 商业应用

业绩：优化 dify 并发量性能，成功将并发能力提升至原水平的 5 倍，支持高负载商业场景应用。

6、微调通义千问模型进行微调，损失降到 0.5 以下。并且做企业级部署

7、自己知乎写的文章关于 chroma 数据库使用
https://zhuanlan.zhihu.com/p/658217843

洛阳幻影网络科技有限公司 技术管理 2014.10-2024.01

运用magento搭建网站和服务器,外贸独立站建站推广后期运营全流程,上百关键词出现在谷歌首页,能让关键词短时间提升到首页。用langchain+chroma做企业的智能问答系统客服。其中调用了openai的api,并且对chatpgt3.5进行了prompt微调
业绩:
用了自己研发的智能问答系统客服,从此夜间客户回复率达到100%

河南君兰动画股份有限公司 PHP 2012.11-2014.10

内容:
负责中国国际动漫网的开发和维护,做好网站功能改进和改版。
业绩:
提高网站客户使用,使加载时间提高3秒,网站使用jquery重新添加很多新的功能。
和ucenter打通用户沟通桥梁

阿里巴巴（中国）有限公司 网站备案专员 2011.10-2012.09

网站备案客服

深圳飞达数码 php开发工程师 2008.07-2011.09

网站开发维护

项目经历

大模型微调和企业级部署 实施 2025.01-2025.03

内容:

下载 qwen2.5 模型做微调然后做企业级部署，模型 loss 收敛到 0.5 以下，部署完可以完美运行

业绩:

1. 负责模型的微调工作，确保模型在特定业务场景下的性能达到最优。
2. 通过精细调整模型参数，成功将模型的loss降至0.5以下，显著提升了模型的准确率和稳定性。
3. 主导了模型在企业级环境的部署工作，保障了模型的高可用性，满足了企业运营的实际需求。
4. 针对企业级部署的需求，优化了模型的计算效率，确保了模型在处理大规模数据时的高效性能。
5. 实施了全面的测试流程，包括单元测试、集成测试及性能测试，确保了模型部署的质量和可靠性。
6. 协同团队成员，针对部署过程中遇到的技术挑战提出创新解决方案，有效提升了项目的整体进度。
7. 通过本项目的成功实施，为公司在同行业中树立了技术领先的标杆，增强了企业的市场竞争力。

yolo在视频领域应用 Python 2023.10-2023.12

内容:
利用python把视频切分成不同帧,然后用labelimg软件标注出自己的数据集。再把数据集放到yolov5里面训练,最后把模型使用gradio做成web应用。
业绩:
使用tensorRT来加速推理。推理速度提升50%左右。自定义数据集的训练mAP50达到90.成功实现自己所选数据集的目标检测

新闻分类自动识别 数据标注/AI训练师 2023.01-2023.09

内容:

用pytorch的RNN和LSTM模型,对15000条新闻,做10类别的新闻分类。

业绩:

通过pandas数据清洗,清洗掉没用的数据。利用腾讯的embedding来完成语料表的embedding。通过数据删除一些数据,让模型准确率从平均80%达到90%

教育经历

安阳师范学院 本科 信息管理与信息系统 2004-2008

资格证书

大学英语六级 大学英语四级 英语6级 agent开发者证书