

# 张振飞-大模型应用开发

年龄：30

学历：本科

毕业院校：上海电机学院

工作年限：7 年

籍贯：河南郑州

电话：13838371945

性别：男

邮箱：[13838371945@163.com](mailto:13838371945@163.com)

## 工作经历

时间	公司名称	岗位
2023.03-至今	志恒融兴有限公司	数据开发，大模型应用
2021.09-2023.01	神州新桥科技有限公司	数据开发工程师
2019.02-2021.09	上海信天翁信息科技有限公司	数仓工程师

## 项目经历

### 基于RAG的知识库问答系统开发 —— 2024.03-至今

背景：基于开源项目 RAGFlow，构建企业级 RAG（检索增强生成）应用，实现文档智能解析、知识检索和 Agent 智能助手，提升企业知识管理和决策效率。

核心职责与成果

- 主导 RAGFlow 核心模块重构，将文档解析和检索模块独立化，提升系统可维护性
- 设计并实现企业知识库构建流程，支持多格式文档解析和向量化存储
- 多模态RAG开发，充分利用图表信息，减少信息缺失
- 基于vllm, xinference部署开源大模型，可视化监控

### 基于LoRA大模型微调 —— 2024.08-至今

背景：针对企业内部私有知识回答准确性不足的问题，基于公司内部知识库数据，采用LoRA技术对开源大模型进行微调，提升模型对企业特定领域知识的理解和回答能力。

核心职责与成果

1. 基于qwen2-7b模型架构，设计并实施LoRA微调技术方案，降低训练成本的同时保持模型性能
2. 构建企业内部知识训练数据集，整理和标注集团内部8000条高质量QA对数据
3. 优化模型训练流程和参数配置，确保微调效果和训练效率的平衡
4. 建立模型评估体系，内部数据集验证微调效果，回答准确率提升7%
5. 解决RAG系统检索信息不准确的核心问题，显著提升企业知识问答的可靠性

## Multi-agent 研报助手开发 —— 2025.01-2025.03

背景：基于LangGraph框架构建多智能体研报分析系统，通过协同工作模式实现研报自动化分析和点评生成，帮助研究员快速获取关键信息并生成专业分析报告。

核心职责与成果

1. 设计多智能体协作架构，包含监督者Agent和专业功能Agent，实现任务分工和协调管理
2. 开发金融问答Agent，集成内部咨询API接口和RAG系统，支持智能图表生成和专业问答
3. 构建研报点评Agent，支持季报、中报、年报等多模板选择，实现维度化分析和信息搜集
4. 基于LangGraph框架实现Agent间的状态管理和 workflow 编排，确保任务执行的可靠性
5. 建立报告生成、审查和反思机制，显著提升研究员信息搜集和分析效率

## 实时风控系统 & 基于图谱数据的风险数据挖掘 —— 2022.03-2023.01

背景:构建基于规则引擎的实时风控系统，实现用户行为实时监控和风险预警，支持规则动态调整，无需停机即可完成规则更新。

项目架构:基于 Flink 实时计算 + Drools 规则引擎 + Kafka 消息队列 + HBase/Redis/ClickHouse 多级存储架构

核心职责与成果:

1. 设计并实现规则引擎对接方案，支持不停机情况下风控规则动态调整，确保系统稳定运行
2. 开发核心风控逻辑，实现用户画像和行为事件的实时规则匹配
3. 设计缓存管理机制，优化 ClickHouse 查询性能，降低系统延迟
4. 实现定时规则监控，支持关联事件的后续行为追踪
5. 基于知识图谱进行风险数据挖掘，通过直接构图、属性增强构图、复杂异构网络、相似构图等方法进行团伙挖掘，支持风控规则的持续优化改进

## 清结算离线数仓 —— 2021.09-2022.03

构建银行卡清算业务数据基础设施，实现客户数据深度洞察，支持运营策略优化和精细化经营，为业务决策提供数据支持。项目架构基于 Hive 离线数仓 + Kylin 多维分析 + Oozie 调度 + Kafka 实时同步 + MySQL/SQL Server 多源数据集成

## 供应链与交易服务分析平台 —— 2019.06-2021.09

构建供应链智能分析平台和交易服务中心数据驾驶舱，实现热销商品预测、库存调拨优化、供应商评估，同时支持服务质量分析和薪酬管理决策，全面提升运营效率。

### 专业技能（Professional Skills）

---

#### 机器学习：

- 熟练掌握深度学习模型方法，熟悉Pytorch深度学习框架，熟悉MLP、RNN、Transformer、BERT、GPT等深度学习模型，熟练掌握vllm推理引擎工作原理，熟悉大模型预训练和Lora高效微调。

#### 开发与运维

- 熟练使用 Python，Java 进行后端项目开发
- 熟练使用大数据技术栈hadoop，hive，kafka，flink，Hbase等
- 熟练掌握 Docker、Kubernetes 容器化技术，能够独立完成项目部署和运维
- 熟练使用 Linux 系统管理，具备服务器运维和性能调优能力

#### 其他技能

- 熟悉 Git 版本控制，具备良好的代码管理习惯
- 了解敏捷开发流程，具备团队协作能力
- 具备良好的问题分析和解决能力