

基本信息

姓名：朱卫涛
联系方式：16630599183
政治面貌：中共党员

意向岗位：大模型算法、AI 应用开发
电子邮件：16630599183@163.com
籍贯：河南商丘



教育背景及服役经历

2024.09~至今，东北大学（985 推免），信息科学与工程学院控制工程专业，副班长，成绩排名 11%（21/194）连续两年获校级一等奖学金。

2018.09~2019.06、2021.09~2024.06，燕山大学，电气工程学院，自动化专业，GPA 排名 4.055/4.5（2/68）河北省优秀毕业生、燕山大学优秀毕业生、燕山大学北京校友会奖学金、国家励志奖学金、燕山大学三好学生（2019、2022、2023），电气工程学院优秀共青团员，校级奖学金七次，计算机三级证书。

2019.09~2021.09，服役于武警北京总队某部，荣获“四有”优秀士兵、嘉奖，参与完成二十余项重大安保任务。

项目经历

2024.10~至今 东北大学钢铁金相组织垂直大模型智能平台搭建 项目负责人

项目描述：面向钢铁金相分析场景，搭建本地化部署的垂直大模型系统。负责领域数据集构建、金相图谱特征提取、基于 LoRA 的大模型微调、Ollama+Dify 私有化部署、知识库与多模态 RAG 系统开发，并实现基于余弦相似度的相似金相图谱检索与网页交互应用，提升专业问答、材料信息查询及图谱对比分析效率。

- 基于专属数据构建钢铁金相领域垂直大模型，完成从数据清洗、LoRA 微调、RAG 检索到本地部署的全链路落地。
- 通过 Qwen 基座模型+LoRA 微调，增强模型对金相组织、材料属性和热处理工艺知识的理解与生成能力。
- 结合 Embedding、Reranker、分层检索构建 RAG 知识库，提升复杂专业问答场景下的召回准确率与回答可信度。
- 融合图像特征向量与文本知识，实现多模态的图谱上传、特征提取、向量检索、结果排序与图文联动展示。
- 依托 Ollama+Dify 完成大模型本地私有化部署与应用编排，满足企业对数据安全、权限隔离和离线可用性的要求。

2025.05~2025.12 基于 MiniMind 大语言模型的轻量级离线车载智能助手 项目负责人

项目描述：基于 MiniMind 模型实现的离线车载智能助手，覆盖从基础语言建模、指令跟随到对齐优化的全链路闭环。实现车载离线智能助手对低延时、低功耗与多模块协同的要求，支撑车辆功能问答、车载知识查询与整机协同任务。

- 从底层实现并验证类 LLaMA 架构关键组件，包括 RoPE、RMSNorm、SwiGLU、Causal Self-Attention 等模块，完成模型结构搭建、前向传播、损失计算与自回归生成逻辑实现，系统掌握 Decoder-only 架构核心原理。
- 构建多阶段数据处理线，完成中文预训练语料清洗、指令数据整理，分别支撑 Pre-train、SFT 与 DPO 三阶段任务。
- 实现从无监督预训练到监督微调再到偏好对齐的完整训练闭环，形成对 LLM 训练范式与对齐机制的系统性理解。
- 面向离线车载助手场景，围绕车辆功能问答、车载手册检索、自然语言控制指令解析等典型任务开展应用验证。
- 综合考虑推理延迟、内存占用等多模块协同运行时的系统资源开销，验证小参数模型在车载端边缘部署的可行性。

2025.12~至今 面向多场景的 LLM-Agent 舆情监测智能分析系统 项目负责人

项目描述：面向多场景风险事件监测需求，以自然语言监测需求为输入，结合大模型与 Agent 执行链路实现舆情监测智能分析系统。实现从任务解析、信息获取、内容分析到自动预警的闭环流程。

- 基于 Qwen API 与提示词工程构建监测任务解析链路，将用户需求映射为结构化表示，为后续任务提供统一输入。
- 面向多源异构信息场景，基于多阶段 LangGraph 工作流，根据监测目标动态设计平台选择策略及内容获取工具。
- 构建候选内容的结构化清洗流程，对抓取结果进行筛选过滤，统一转换为可供模型分析的 JSON 数据格式。
- 基于大模型对结构化内容进行风险判定与舆情生成，形成从语义理解到风险判定再到预警输出的闭环分析能力。
- 基于 Flask+HTML 搭建 Web 交互页面，对 Agent 执行链路中进行可视化展示，增强系统原型的可交互性。

研究课题

- **研究课题：**基于多模态信息融合的脑部疾病诊断算法研究
- **项目内容：**联合 fMRI 脑影像数据与表型信息，基于图卷积网络开展多模态信息融合，联合跨半球交互信息与全脑拓扑结构特征，实现跨疾病下的统一识别。相比基线方法，在公开脑疾病数据集上模型识别准确率平均提升 3%。
- **论文成果：**《A graph convolutional network framework for ASD classification and model-level analysis of candidate ASD-related regions》，状态：under second-round review
- **专利成果：**一种基于关键点检测和目标识别的驾考安全员违规操作检测方法（202511267838.1）状态：实质审查

专业技能知识

- 熟悉 PyTorch，能够完成模型搭建、训练、微调、评估与推理部署，具备良好的算法实现与工程调试能力。
- 熟悉 Transformer、LLaMA、Decoder-only 架构，掌握 Pre-train、SFT、DPO 等 LLM 全链路训练流程。
- 具备 Linux 服务器使用与维修、Jetson Nano 边缘部署、数据清洗、Prompt 设计、实验分析等工程实践能力。
- 曾参与国家级重点研发项目申报，具备科研论文写作能力和专利申请经验。