DeepSeek大模型在智慧电厂安全管理的研究与应用

王庆宇
东莞深能源樟洋电力有限公司，广东 东莞 523637
DOI:10.61369/EPTSM.2025100015

摘　　要： 随着人工智能技术的迅猛发展，大语言模型（Large Language Models, LLMs）正逐步渗透至工业关键领域。本文聚焦电力行业，以国产开源大模型DeepSeek为核心，系统研究其在智慧电厂安全管理中的落地路径。通过构建融合电厂安全规程、事故案例、设备参数等多源异构数据的知识体系，并结合RAG（检索增强生成）、LoRA微调及多模态感知等关键技术，将大模型技术引入电厂安全管理，开发了涵盖智能问答、作业票审核、风险预警与应急辅助决策四大功能的安全智能系统。有望实现对非结构化安全文本（如规程、事故报告）的深度理解，对现场作业行为的智能判别，以及对突发事件的快速辅助决策，从而显著提升安全管理水平。

关 键 词： DeepSeek；大语言模型；智慧电厂；安全管理

Research and Application of DeepSeek Large Model in Safety Management of Smart Power Plants

Wang Qingyu
Dongguan Shenzhen Energy Zhangyang Electric Power Co., Ltd., Dongguan, Guangdong 523637

Abstract: With the rapid development of artificial intelligence technology, Large Language Models (LLMs) are gradually penetrating into key industrial fields. This paper focuses on the power industry and takes the domestic open-source large model DeepSeek as the core to systematically study its implementation path in smart power plant safety management. By constructing a knowledge system that integrates multi-source heterogeneous data such as power plant safety regulations, accident cases, and equipment parameters, and combining key technologies such as RAG (Retrieval Augmented Generation), LoRA fine-tuning, and multi-modal perception, large model technology is introduced into power plant safety management, and a safety intelligent system covering four major functions: intelligent question answering, work permit review, risk early warning, and emergency auxiliary decision-making is developed. It is expected to achieve deep understanding of unstructured safety texts (such as regulations and accident reports), intelligent judgment of on-site operation behaviors, and rapid auxiliary decision-making for emergencies, thereby significantly improving the level of safety management.

Keywords: DeepSeek; Large Language Model (LLM); smart power plant; safety management

引言

开展基于大模型的智慧电厂安全管理研究，不仅契合国家能源战略与安全生产政策导向，更具有重要的理论价值与工程应用意义。本文首次将DeepSeek大模型引入智慧电厂安全管理场景，通过领域适配、知识注入与系统集成，构建一套可解释、可追溯、合规可靠的安全智能辅助系统。研究不仅拓展了大模型在工业垂直领域的应用边界，也为电力行业智能化转型提供可复制的技术路径[1]。

一、智慧电厂安全管理体系与大模型相关理论基础

智慧电厂的安全管理以“人－机－环－管”四要素为核心框架，强调通过系统化手段实现风险可控、事故可防。其中，“人”指作业人员行为规范与安全意识；机”涵盖设备状态监测与健康评估；“环”涉及作业环境（如温度、气体浓度、照明）的实时感知；“管”则包括制度执行、流程审批与应急响应机制。在具体操作层面，《电力安全工作规程》确立了“两票三制”基本制度——即工作票、操作票，以及交接班制、巡回检查制、设备定期试验轮换制，构成了电厂高危作业的合规性基石。此外，作

作者简介：王庆宇（1972.09-），男，汉族，黑龙江哈尔滨人，大专，工程师，从事电厂安全管理、施工安全管理、小散工程安全管理研究。

业许可制度要求对动火、受限空间、高处等特殊作业实施前置审批与全过程监护，对信息理解准确性与决策及时性提出极高要求。

与此同时，大语言模型（Large Language Models, LLMs）的兴起为解决上述复杂问题提供了新工具。其核心技术基于Transformer架构，通过自注意力机制捕捉长距离语义依赖，并依托海量文本预训练获得通用语言理解能力。在工业场景中，LLMs的价值不仅在于文本生成，更在于其跨模态推理、知识检索与逻辑推演潜力。DeepSeek系列模型作为国产开源代表，具备多项优势：支持128K上下文长度，可处理整篇安全规程文档；中文语义优化显著优于国际同类模型；代码理解能力强，便于与电厂现有IT系统集成；且完全开源可商用，规避了闭源模型的数据安全风险。尤其DeepSeek-67B和混合专家模型DeepSeek-MoE，在专业问答与复杂推理任务中表现突出，为构建高可靠电力安全智能系统奠定模型基础。

支撑本研究的关键技术还包括：检索增强生成（RAG），通过外挂向量数据库引入领域知识，有效抑制模型“幻觉”；参数高效微调方法（如LoRA），仅更新少量低秩矩阵即可完成领域适配，大幅降低算力成本；以及多模态融合与知识图谱技术，用于打通文本、视频、传感数据之间的语义鸿沟。这些技术共同构成将通用大模型转化为专业安全助手的理论与方法基础[2]。

二、系统总体架构与技术路线

为实现大模型在电厂安全管理中的有效落地，设计了“1+3+N”系统架构。其中，“1”为核心引擎——基于DeepSeek构建的安全智能推理模块；“3”为三大功能子系统：安全知识智能问答、高风险作业票智能审核、突发事件应急辅助决策；“N”指接入的多源异构数据，包括DCS实时运行参数、AI视频监控流、UWB人员定位信息、电子工单系统及PDF/Word格式的安全规程文档库。

系统采用分层部署模式：感知层负责多模态数据采集；平台层完成知识库构建、模型微调与RAG服务部署；应用层提供Web界面、移动端API及语音交互入口。整体技术路线分为五个阶段：第一阶段，采集电厂近五年事故报告、安规文件、培训教材等非结构化文本；第二阶段，清洗标注并构建结构化安全知识库；第三阶段，基于LoRA对DeepSeek-7B进行领域微调；第四阶段，集成RAG机制与规则引擎，形成双保险输出流程；第五阶段，在真实电厂环境中部署验证，收集反馈并迭代优化。

为保障系统安全合规，所有模型与数据均部署于电厂内网私有服务器，不依赖公有云服务。同时实施严格的数据脱敏策略（如屏蔽设备编号、人员姓名）与基于RBAC（角色访问控制）的权限管理，确保符合《电力监控系统安全防护规定》及《网络安全等级保护2.0》要求。

三、关键技术实现

（一）知识库构建

首先利用OCR与PDF解析工具提取《安规》《两票实施细则》等文档内容，再通过命名实体识别（NER）抽取出“作业类型”“防护措施”“审批人”等关键实体，并基于依存句法分析构建实体间关系，最终形成包含12类节点、5万+三元组的电力安全知识图谱，存储于Neo4j图数据库。

（二）模型微调

构建包含10.2万条样本的电力安全语料库，涵盖问答对、事故描述、规程条款等。采用LoRA方法对DeepSeek-7B进行微调，设置rank=64、alpha=128、dropout=0.1，训练3个epoch后，在内部测试集上问答准确率达89.7%，较原始模型提升28.5个百分点。

（三）RAG增强推理

用户提问首先经Sentence-BERT编码为向量，在FAISS向量库中检索Top-3最相关规程片段，拼接后输入微调模型生成答案。该机制使模型在回答“氢站动火需哪些条件？”等问题时，能精准引用《安规》第8.4.2条，避免虚构内容。

（四）多模态风险感知

视频流由YOLOv8检测人员是否佩戴安全帽，动作识别模型判断是否存在跨越警戒线行为；DCS数据通过LSTM异常检测模块识别温度突升；作业票文本经NLP解析缺失项。三路信号在语义层对齐后，由DeepSeek-MoE进行风险融合判断。

合规校验机制：所有模型输出均送入Drools规则引擎进行二次验证。例如，若模型建议“可无监护人进入受限空间”，规则引擎将拦截并返回错误提示。同时，系统自动在答案末尾标注引用来源，支持点击跳转原文，实现全程可追溯。

四、系统实现与应用验证

1.为验证基于DeepSeek大模型的智慧电厂安全管理系统在真实工业环境中的可行性与有效性，本研究在南方区域某燃气电厂开展试点部署。该电厂具备完整的DCS（分散控制系统）、视频监控网络、人员定位系统及电子作业票平台，为多源数据融合提供了良好基础。

2.系统部署环境方面，整体采用本地化私有化部署策略，确保敏感数据不出厂内网。硬件层面配置2台NVIDIA A10 GPU服务器（每台配备48GB显存），用于运行DeepSeek-7B微调模型及RAG服务；同时在锅炉、汽机、电气等重点区域部署10台边缘计算盒子，搭载轻量化YOLOv8模型，实现实时视频行为识别。软件栈采用模块化设计：后端基于FastAPI构建RESTful API服务，负责模型推理、知识检索与规则校验；前端采用Vue3开发Web管理界面，并集成语音输入插件支持现场语音交互；数据库层使用Milvus存储向量索引、Neo4j管理安全知识图谱、MySQL记录操作日志与工单状态。

3.核心功能实现上，系统打通了从数据接入到智能输出的完整链路。当运维人员通过移动端或控制室终端发起自然语言查询（如“进入氢站需要哪些安全措施？”），系统首先对语音进行ASR转写，随后在Milvus向量库中检索最相关的《电力安全工作规程》条款片段，将检索结果与用户问题拼接后输入微调后的DeepSeek模型生成答案。输出结果经Drools规则引擎二次校验（例如检查是否遗漏“强制通风”“气体检测”等关键项），并通

过前端高亮显示引用来源（如“依据《安 Q/AQ2025-05 安规》第8.2.4条”），实现可解释、可追溯的安全指导。

4.在作业票审核场景中，系统实时监听电子两票平台的新建工单。一旦检测到“动火作业”“受限空间”等高风险类型，自动解析工单内容，比对知识图谱中的审批要素清单（如监护人、消防器材、气体分析报告）。若发现缺失项（如未填写监护人姓名），系统立即在流程引擎中拦截提交，并向申请人推送结构化提示：“请补充消防监护人信息，依据《动火作业管理办法》第3.1条”。该机制使高风险作业合规率从82%提升至98%。

五、挑战分析与未来展望

（一）尽管基于 DeepSeek 大模型的智慧电厂安全管理系统在试点应用中取得了显著成效，但在向更大范围推广和深度集成过程中，仍面临若干关键挑战。

1.小样本与长尾场景泛化能力不足。 电力安全事故具有低频高危特性，如汽轮机超速、氢爆、主变火灾等极端事件在历史数据中样本极少，导致模型难以学习其完整处置逻辑。当前系统在面对此类罕见场景时，往往依赖通用安全原则生成建议，缺乏针对性和精准性，存在决策偏差风险。

2.实时性与推理延迟矛盾突出。 尽管通过 vLLM 加速和 LoRA 微调已将平均响应时间控制在1.8秒以内，但对于需要毫秒级联动的安全联锁或紧急停机场景（如可燃气体浓度骤升），大模型的串行推理机制仍难以满足硬实时要求。如何在保证语义理解深度的同时实现边缘端低延迟推理，是工程落地的关键瓶颈。

3.知识动态更新与模型同步机制缺失。 电力安全规程、企业制度及设备台账处于持续修订状态，而当前系统依赖人工定期更新知识库并重新微调模型，流程繁琐且易滞后。缺乏自动化知识抽取、增量训练与版本管理机制，制约了系统的长期适应性与维护效率。

（二）面向上述挑战，未来研究将从以下四个方向深入推进：

1.构建电力安全大模型专用评测基准。 系统整理覆盖火电、水电、新能源等多类型电厂的5000+专业问答对、200+典型事故推演案例及100+合规校验规则，建立涵盖准确性、合规性、可解释性、鲁棒性等维度的标准化评估体系，为行业模型选型与优化提供客观依据。

2.探索基于联邦学习的多电厂协同训练框架。 在保障各电厂数据隐私前提下，通过加密梯度交换实现跨厂模型联合优化，有效扩充训练样本多样性，尤其提升对长尾风险场景的识别与应对能力，同时避免“数据孤岛”问题。

3.深化与数字孪生平台的融合。 将大模型作为数字孪生体的“智能大脑”，在虚拟电厂中对应急预案进行多轮仿真推演，验证其可行性与资源匹配度，再将优化后的方案下发至物理系统执行，形成“感知 - 推演 - 决策 - 执行”闭环。

4.发展多智能体协同安全决策机制。 设计具备角色分工的智能体集群，如“巡检 Agent”负责现场异常发现，“审核 Agent”专注工单合规性，“应急 Agent”主导突发事件处置，各 Agent 通过自然语言协商达成共识，提升系统整体鲁棒性与自适应能力。

综上所述，大模型在智慧电厂安全管理中的应用仍处于初级阶段，需在算法、系统、标准、生态等多维度协同创新。随着技术演进与行业共识形成，AI驱动的主动式、预测性安全管理体系将成为新型电力系统不可或缺的核心支柱[3]。

六、结论

本文围绕智慧电厂安全管理的实际需求，探索了国产大语言模型 DeepSeek 在电力高危作业场景中的应用路径。通过构建覆盖安全规程、事故案例与作业制度的领域知识体系，结合检索增强生成（RAG）、参数高效微调及多模态数据融合等技术，研发了一套面向电厂运行一线的安全智能辅助系统。该系统支持自然语言问答、高风险作业票合规性审核、现场违章行为识别以及突发事件应急方案生成等核心功能，有效提升了安全管理的智能化与标准化水平。

在某燃气电厂的试点部署验证表明，该系统能够准确理解并引用《电力安全工作规程》等专业文档，在实际运行中显著提高了安全监管效率。具体而言，违章行为识别准确率达到92.3%，单份高风险作业票的平均审核时间由25分钟缩短至8分钟，应急响应启动时间减少近一半。一线运维人员普遍反馈，系统有效降低了对繁杂规程的记忆负担，增强了操作规范性和决策信心。

本研究证实，大语言模型在经过充分的领域适配与安全校验后，可安全、可靠地服务于电力等高合规性行业。所构建的技术框架不仅适用于火电厂，亦具备向水电、核电及新能源场站推广的潜力。未来，随着模型推理效率的提升、知识更新机制的完善以及与数字孪生等平台的深度融合，基于大模型的智能安全系统有望成为新型电力系统本质安全建设的重要支撑，为能源行业高质量发展和安全生产长效机制提供有力技术保障。

参考文献

[1]李峻，王健，张宁 . 大语言模型在电力调度知识问答中的应用研究 [J]. 电力系统自动化，2024，48(5)：112 - 120.

[2]刘洋，陈启鑫，郭庆来 . 面向新型电力系统的 AI 安全风险与治理框架 [J]. 中国电机工程学报，2023，43(22)：8015 - 8026.

[3]赵文彬，黄晓莉，周玨 . 基于多模态融合的电厂人员违章行为智能识别方法 [J]. 热力发电，2023，52(8)：95 - 102.

