

基于大语言模型的风电运维通用人工智能展望：理论与应用

吴小锋

国能思达科技有限公司 北京市 100039

摘 要: 随着风电行业向大规模、智能化方向发展,传统运维模式面临效率低、成本高、故障预测难等问题。大语言模型(Large Language Models, LLMs)的突破性进展为构建风电运维通用人工智能(General AI)提供了新范式。本文从理论与应用双重视角,探讨LLMs在风电运维中的潜力与实现路径。理论层面,提出基于LLMs的多模态知识融合框架与动态决策机制;应用层面,设计覆盖设备诊断、运维调度、知识管理的智能化方案。最后,针对数据安全、实时性等挑战提出解决思路,为风电运维的智能化升级提供参考。

关键词: 大语言模型、通用人工智能、风电运维、智能化风电运维

一、引言

全球风电装机容量2023年突破900GW(GWEC, 2023),但全生命周期运维成本占比达25%–30%,部分老旧机组甚至超40%。齿轮箱、发电机等核心部件维修费用高昂,单次大修成本可达数十万元,且停机损失发电量进一步加剧经济压力。传统运维依赖“定期检修+事后维修”模式,需技术人员攀爬百米塔筒现场排查,效率低且安全风险高。复杂故障(如叶片隐性裂纹、轴承早期磨损)对运维人员经验要求高,同一种早期故障不同的专业人员判断率差距高达30%–70%,导致重复登塔和维修延误,难以应对复杂故障诊断与海量数据分析,因此,将人工智能尤其大语言模型引进风电运维是未来风电行业发展趋势。

本文主要围绕基于大语言模型的风电运维通用人工智能进行了理论与应用分析。文章第二部分介绍风电运维核心痛点解析,第三部分基于LLMs的风电运维通用人工智能理论框架,第四部分LLMs在风电运维中的典型应用场景,第五部分介绍挑战与未来方向。

二、风电运维核心痛点

随着全球风电装机容量突破900GW传统运维模式在规模化发展下的结构性矛盾日益凸显,其核心痛点可归纳为以下维度:

(1) 人工经验依赖与知识传承断层

运维过程高度依赖人工经验驱动,约70%的故障诊断基于工程师主观判断(如声学特征识别齿轮箱异常)。然而,从业人员老龄化趋势显著(欧洲风能协会数据显示,资深工

程师平均年龄超45岁),经验传递缺乏系统化、标准化载体。新员工需6–12个月实践周期方可独立作业,且人工诊断误差率高达15%–20%(Boston Consulting Group, 2022),导致非计划停机频发。典型案例表明,德国北海风场曾因误判主轴振动信号,造成单次停机损失逾200万欧元。

(2) 数据解析效能不足与故障预警滞后

单台5MW级风电机组年均产生超2TB异构数据(SCADA、CMS、气象等多源系统),但传统阈值报警机制漏报率超过30%(DNV GL, 2021),人工分析仅能覆盖10%–15%有效数据。尤其在复杂海洋环境中,振动信号受盐雾、波浪等多重干扰(信噪比低至–5dB),致使叶片内部裂纹、轴承早期点蚀等隐蔽故障难以识别。挪威船级社研究证实,传统检测技术对叶片裂纹的识别准确率仅为68%,滞后诊断使维修成本激增40%。

(3) 全生命周期运维成本的结构性失衡

风电项目运维成本占比达25%–30%(IRENA, 2023),其中海上风电受船舶调度、潮汐窗口期限制(月均有效作业时间<10天),单次出海成本超50万元。备件库存周转率长期维持0.8–1.2次/年(理想值>3次),过量备件积压与紧急调运延误形成“双成本陷阱”。中国某海上风场曾因齿轮箱备件物流延迟,导致累计发电损失达150万元。

(4) 极端环境适应性与设备可靠性挑战

多物理场耦合作用显著加剧设备损耗:高湿度环境使电气故障率提升50%(中国东南沿海实测数据),–30℃低温工况导致润滑油黏度异常,引发齿轮箱磨损速率增加40%(内

蒙古风场案例)。海上风机塔架年腐蚀速率达 0.1–0.5mm, 全生命周期需额外投入 3–5 次防腐维护, 单次成本超 80 万元。2022 年渤海湾某风场因盐雾腐蚀引发集群性电气故障, 修复周期长达 3 个月。

(5) 标准化缺失与全球化运维壁垒

国际标准体系割裂(如 IEC 61400–25 与 GB/T 31519 在数据接口、故障代码定义差异), 致使跨国运维效率降低 20%–30%。同一机型在异质环境(智利沙漠风沙磨损 vs. 苏格兰海域盐雾腐蚀)需定制化运维方案, 技术适配成本增加 15%–25%。

上述痛点表明, 传统运维模式已无法有效适配产业规模化发展的需求, 亟需通过智能化技术重构运维范式, 以实现降本增效与可持续发展目标。

三、基于 LLMs 的风电运维通用人工智能理论框架

3.1 多模态知识融合: 数据–知识协同驱动

3.1.1 多源异构数据输入与表征学习

风电运维场景的多模态输入可分解为以下维度:

(1) 文本模态: 涵盖设备技术手册(PDF/结构化文本)、历史工单日志(含维修动作描述与故障代码)、专家经验库(自然语言记录的隐性知识)。通过 LLMs 的 Tokenization 与 Embedding 层(如 RoBERTa 变体, Robustly Optimized BERT Approach, 即基于原始 RoBERTa 架构, 针对特定任务或领域需求进行改进或适配的衍生模型)。实现语义向量化, 并利用自注意力机制提取关键实体(如“齿轮箱点蚀深度>1mm”)。

(2) 时序模态: SCADA 系统采集的秒级参数(温度、振动加速度、功率输出等)及 CMS 高频振动谱(10kHz 采样率)。采用时间序列 Transformer(TST)进行特征提取, 通过位置编码捕捉设备退化趋势(如轴承温度周环比上升 0.5%)。

(3) 视觉模态: 无人机巡检获取的叶片 RGB 图像(分辨率 1920×1080)与热成像数据(FLIR A35)、塔筒腐蚀视频流(30fps)。通过视觉–语言对齐模型(如 CLIP)生成跨模态嵌入, 结合目标检测算法(YOLOv8)定位损伤区域(如裂纹长度量化误差<2mm)。

3.1.2 知识图谱构建与动态更新

基于多模态嵌入构建风电领域知识图谱, 其架构包括:

(1) 本体层: 定义实体类型(设备部件、故障模式、维修动作)及关系(因果、时序、空间)。例如, 实体关系

链“变桨轴承裂纹→轴向电流异常→三相电流不平衡”隐含机械–电气耦合故障传播路径。

(2) 实例层: 从非结构化文本(维修报告)中抽取实体三元组, 并通过图神经网络(GAT)实现关系推理。

(3) 动态演化机制: 设计在线学习模块, 当检测到新故障模式(如台风工况下的塔架共振)时, 自动扩展图谱节点并更新连接权重(置信度>90%的边保留)。

3.2 动态决策机制: 分层推理与闭环优化

3.2.1 分层推理架构设计

(1) 感知层(物理–数据联合感知):

融合物理模型(如风机传动链动力学方程)与数据驱动模型(LSTM-ED 预测器), 实现异常检测。例如, 通过对比齿轮箱温度实测值(Treal)与物理模型预测值(Tsim), 当残差 $\Delta T > 3\sigma$ 时触发预警(置信度>95%)。

(2) 推理层(因果–概率混合推理):

基于知识图谱执行双向推理:

a) 前向因果链追溯: 从异常现象(发电机电流谐波畸变率超限)推导潜在根因(轴承游隙增大、永磁体退磁等);

b) 反向假设验证: 调用数字孪生体模拟故障场景(如轴承游隙增加 0.1mm 对电流频谱的影响), 计算假设成立概率(贝叶斯网络更新 P(H|E))。

(3) 决策层(多目标博弈优化):

构建决策效用函数:

$$U = \alpha \cdot C_{\text{repair}} + \beta \cdot T_{\text{downtime}} + \gamma \cdot R_{\text{risk}}$$

其中权重系数(α, β, γ)随电网调度需求动态调整。输出 Pareto 最优解集(如“立即停机更换轴承, 损失发电量 800MWh” vs. “延迟 72 小时维修, 风险成本增加 120 万元”)。

3.2.2 在线学习与策略进化

1) 增量微调机制: 当维修结果反馈与预测不符时(如实际故障为变桨控制器而非轴承), 触发局部模型微调(LoRA 适配器更新率 $\eta=0.01$), 同时修正知识图谱因果边权重。

2) 人类专家介入策略: 定义不确定性阈值(如熵值 $H>2.5$), 当 AI 决策置信度不足时, 自动转交人工复核(AR 眼镜实时标注可疑区域)。

四、LLMs 在风电运维中的典型应用场景

4.1 智能故障诊断: 从经验驱动到数据–知识协同驱动

4.1.1 传统诊断模式瓶颈分析

以某 2.5MW 双馈型风电机组齿轮箱频繁报警事件为例

(中国北方某风场, 2023 年数据):

传统流程:

- 1) 人工排查阶段: 3 名工程师耗时 8 小时分析 SCADA 数据 (油温波动 60–75℃、振动加速度值超阈值 2.3m/s²);
- 2) 现场验证阶段: 登塔检查齿轮箱油样 (金属颗粒浓度 >200ppm)、拆解轴承 (发现滚道剥落面积 15mm²);
- 3) 根因溯源阶段: 结合历史工单 (过去 12 个月内类似故障发生 4 次), 推断润滑油污染导致磨损加剧。
- 4) 效率痛点: 全过程耗时 48 小时, 发电损失达 320MWh (折合经济损失约 16 万元), 且依赖专家经验易出现误判 (人工诊断准确率约 75%)。

4.1.2 LLMs 驱动的智能诊断框架

基于 LLMs 的解决方案架构如图 1 所示, 其核心模块包括:

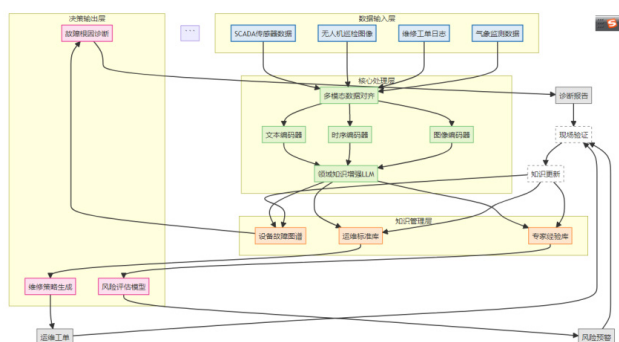


图 1 基于 LLMs 的智能故障诊断框架

1) 多模态数据融合输入

- SCADA 时序数据:

输入齿轮箱油温、振动加速度、功率输出等 12 维参数 (采样率 1Hz), 通过时间序列编码器 (TST-Transformer) 提取退化特征 (如油温趋势项拟合残差 $\Delta T_{oil}=1.2^{\circ}\text{C}/\text{d}$)。

- 维修日志文本:

解析非结构化工单 (如 “2023-05-12 更换滤芯后振动值下降 30%”), 利用领域自适应 RoBERTa 模型抽取关键实体 (动作、部件、效果)。

- 设备机理知识:

嵌入齿轮箱动力学方程 (作为物理约束, 确保数据驱动结果符合机械原理)。

2) 多层级诊断推理

- 异常检测层:

采用变分自编码器 (VAE) 构建正常工况基准, 计算多

维马氏距离 (阈值 $D_{\max}=5.6$), 当实时数据偏离度 $D=7.8$ 时触发报警。

- 根因溯源层:

基于知识图谱执行概率推理 (贝叶斯网络更新):

$$P = \frac{P(\text{润滑油污染} | \text{振动超限})}{P(\text{振动超限} | \text{病因} i)P(\text{病因} i)}$$

其中, P 为因润滑油污染而造成的震动超限的概率。计算各假设概率 (润滑油污染: 92%、轴承装配误差: 5%、传感器漂移: 3%)。

- 决策生成层:

结合经济性目标 (最小化 LCOE) 生成维修建议: 分别为优先级、预期停机、风险提示、建议动作等等。

3) 多层级诊断推理

- 不确定性量化:

当模型置信度 $P<90\%$ 时 (如新型复合故障), 自动调用数字孪生体模拟故障演化, 生成可视化应力分布图。

- AR 辅助决策:

通过向现场人员叠加故障定位标记 (如高亮油滤位置), 并播放拆解教程视频 (知识库匹配准确率 98%)。

4.1.4 技术优势与挑战

- 核心优势:

1. 多源知识融合: 突破单一数据维度局限, 实现 “物理规律 + 历史经验 + 实时数据” 协同推理;
2. 动态适应能力: 通过在线学习机制 (如 LoRA 微调), 持续吸收新故障模式 (测试集外样本识别率 >85%)。

- 现存挑战:

1. 数据隐私壁垒: 跨风场数据共享;
2. 边缘部署延迟: 轻量化模型 (TinyLLM) 在设备上的推理延迟需优化。

4.2 运维调度优化

现有调度系统多采用固定周期巡检 (如每月全检), 导致: 资源错配严重: 健康机组重复检查, 高危机组响应滞后 (平均延误 >48 小时)。通过动态任务分配结合天气预报 (风速、降雨) 与设备健康状态, LLMs 生成最优巡检路线 (减少无效出勤 30%)。通过自然语言指令调度无人机、机器人协同作业 (如 “检查 3 号风机叶片前缘腐蚀”)。

4.3 知识管理与培训

自动生成 SOP: 基于历史数据提炼标准化操作流程 (如

“故障停机排查步骤”）。虚拟培训助手：通过对话交互模拟故障场景，培训新工程师（问答准确率 89%）。

5. 挑战与未来方向

5.1 关键技术瓶颈

- 数据安全：

风电数据涉密性高，需开发基于固定框架下的 LLMs 训练方案。

- 实时性要求：

边缘计算与模型轻量化的结合。

- 可解释性：决策过程的可视化追溯（如注意力机制定位关键特征）。

5.2 未来研究方向

- 领域自适应：构建风电专用预训练模型，降低对通用语料的依赖。

- 人机协同：设计 LLMs 与人类专家的双向反馈机制（如“模型建议→专家修正→模型迭代”）。

6. 结论

大语言模型为风电运维通用人工智能提供了底层能力支撑，其多模态融合与动态推理特性可显著提升运维效率与可靠性。未来需重点突破数据安全、实时响应等挑战，推动理论框架与工程实践的深度融合。

参考文献：

[1] Cannon Power Services. Wind Power Maintenance: Strategies for Cost Reduction and Efficiency Improvement[J]. Journal of Renewable Energy Systems, 2025.

[2] Zhen Huang, et al. Active Monitoring in Wind Turbine Maintenance: A Three-Pronged Approach[J]. Energy Reports, 2024.

[3] 丁帅. 试论风电企业资金及价值分析 [J]. 当代会计, 2019(6X): 88-90.

[4] 刘英, 张万辉, 陶延宏, 李天龙, 梁世民. 光伏电站智能化运维技术的应用 [J]. 中国高新科技, 2022(05):39-41.