

人工智能在信息系统运维中的智能诊断与预测模型构建

尹剑锐¹ 孙建² (通讯作者)

1. 北京清华长庚医院 北京市 102218

2. 四川尔腾科技有限公司 四川省 攀枝花市 617000

【摘要】随着企业信息系统规模与复杂度的持续增长,传统依赖阈值告警与人工排查的被动运维模式已难以满足高可用性要求。人工智能技术,特别是深度学习算法在时序异常检测和故障预测领域的突破,为信息系统运维的智能化升级提供了新路径。本文围绕 AI 在信息系统运维中的智能诊断与预测模型构建,从整体规划、技术选型、多路径对比分析和技术管理四个维度展开论述,提出了一套可落地的智能运维模型构建方法论,为企业运维体系向主动预测模式转型提供参考。

【关键词】人工智能;信息系统运维;智能诊断;故障预测模型;AIOps

0 引言

在数字化转型深入推进的背景下,企业的信息系统架构正变得日益复杂——微服务数量激增、多云混合部署常态化、数据流量呈指数级增长。传统的运维模式依赖固定阈值告警与人工排查,具有明显的滞后性:故障通常在用户报告后才能发现,平均修复时间(MTTR)可达数小时;人工测试难以覆盖所有执行路径,遗漏缺陷率高达 30%~50%;紧急修复的人力成本是预防性维护的 3~5 倍。同时,大型集团企业日均日志量可达数十 TB,包含应用层、网络层、数据库层等多维异常信号,人工分析已不具备可行性。

人工智能技术的介入为这一困境提供了系统性解决方案。智能运维(AIOps)通过融合机器学习、大数据分析与自动化技术,构建“感知—分析—决策—执行”闭环,实现故障的提前预警与智能诊断。智能运维的核心在于将“被动响应”转变为“主动预防”——运维人员得以在潜在风险爆发之前提前介入,将故障消弭于无形。本文从规划、选型、对比与管理四个维度,系统探讨 AI 驱动的智能诊断与预测模型构建路径。

1 整体规划:智能诊断与预测模型体系架构

1.1 模型建设目标

智能诊断与预测模型建设的核心目标是构建一套覆盖故障全生命周期的智能化能力体系。具体包括三个层次:一是异常检测层,实现对系统运行状态的实时感知与偏差识别;二是故障诊断层,在异常发生后快速定位根因,缩短排查时间;三是趋势预测层,基于历史模式预测未来可能出现的性能瓶颈与硬件故障,实现预测性维护。

1.2 四层技术架构设计

智能诊断与预测模型的构建应遵循分层架构设计原则,分为数据采集层、数据处理与特征工程层、AI 算法模型层和智能应用层四个层次,如图 1 所示。



图1 智能诊断与预测模型四层架构图

数据采集层负责从服务器、网络设备、应用系统等数据源中实时采集日志(Log)、指标(Metric)、链路追踪(Trace)和拓扑(Topology)四类核心运维数据。数据处理与特征工程层对多源异构数据进行清洗、标准化和特征提取,为 AI 模型提供高质量输入。AI 算法模型层是系统的核心,承载异常检测、根因定位和故障预测三类关键算法模型。智能应用层将 AI 模型的输出转化为可操作的运维决策,支撑智能告警、故障自愈、容量预测等场景化应用。

1.3 数据治理规划

没有高质量、标准化、可追溯的数据, AI 模型就是无源之水。建议构建统一的数据中台,整合来自不同监控系统(如 Prometheus、Zabbix、ELK)的数据,建立标准化的元数据标签体系(如 env=prod, region=cn-east, owner=payment),实现跨系统的数据关联与血缘追踪。具体实施中,应重点关注以下三个层面:一是数据质量管理,建立自动化的数据质量监控机制,对采集数据的完整性、一致性、时效性进行持续校验,识别并修复缺失值、异常值、时间戳错位等问题;二是血缘追踪与影响分析,引入算子级血缘解析技术,实现数据从采集源

头到模型输入的全链路追踪，准确率可提升至 99% 以上，支持上游变更的精准影响评估；三是 DataOps 与 MLOps 融合，构建统一的元数据湖，实现数据工程与 AI 开发流程的协同，打通数据发现、特征工程到模型训练的完整链路，降低数据重复存储与开发成本。

2 选型建设：核心模型技术方案

2.1 智能诊断模型选型

智能诊断的核心任务是识别系统运行中的异常模式并定位故障根因。根据数据类型和诊断场景的不同，可选择差异化的算法方案。

异常检测模型：对于单变量指标（如 CPU 利用率、内存占用）的异常检测，可采用 3σ 原则或 EWMA 控制图等统计方法；对于多维指标的联合异常检测，孤立森林（Isolation Forest）和 One-Class SVM 表现优异。孤立森林算法通过随机分割特征空间来隔离异常点，在日志异常检测场景中，其异常发现率可提升 80% 以上，误报率降至 5% 以下。对于具有时序依赖性的指标数据，LSTM 自编码器（LSTM-AE）通过重构误差判断异常，能够有效捕捉时间序列中的复杂模式。

根因定位模型：根因分析需要在异常发生后快速定位故障源头。传统方法依赖运维人员逐层排查，平均耗时超过 45 分钟。基于图神经网络（GNN）的根因推理方法通过构建服务依赖拓扑图，将故障事件作为图节点、资源依赖作为边，利用 GNN 进行反向传播推理，在复杂微服务架构中的根因定位准确率可达 89%，远超人工排查水平。

2.2 预测模型选型

故障预测模型的任务是基于历史运行数据预测未来可能发生的性能劣化或硬件故障，为预防性维护提供决策依据。对于

时间序列预测，LSTM（长短期记忆网络）和 Transformer 是两类主流选择。

LSTM 通过门控机制有效解决长序列中的梯度消失问题，适合处理具有长期依赖关系的运维时序数据（如服务器负载的周期性波动、硬盘 SMART 参数的渐变衰减）。已有研究表明，LSTM 在某金融系统的故障预测中准确率达 89%。CNN-LSTM 混合模型结合了 CNN 的空间特征提取能力和 LSTM 的时序建模能力，在工业场景中达到 96.1% 的准确率和 95.2% 的 F1-score。

Transformer 模型通过自注意力机制（Self-Attention）捕捉序列中的长程依赖关系，摒弃了循环网络的序列计算方式，可并行处理整个时间窗口的输入数据，大幅提升训练效率。其多头注意力机制能够从多个子空间同时关注时序信号的不同维度特征，这对于运维场景中多指标联合预测尤为重要。实践表明，基于 Transformer 的预测模型可提前 48 小时预测到数据库连接池耗尽等资源瓶颈问题。华为发布的运维领域专业大模型 EDNS 2.0 加载基于 Transformer 架构的风险预测模型，实现了 80% 以上的风险识别准确率和 90% 以上的诊断准确率，支持 T-1 时刻的预测预防与 T0 时刻主动响应的“双保险”运维模式。在大规模日志数据的语义理解场景中，大语言模型（LLM）正展现出独特的价值，可对非结构化日志进行意图识别与知识提取。

3 多路径对比分析

3.1 模型方案对比

不同 AI 模型在诊断与预测任务中各具优劣，需要根据实际场景进行综合选型。表 1 从诊断准确率、训练复杂度、推理速度、可解释性四个维度对主流方案进行了对比。

表1 智能诊断与预测模型多路径对比

模型类型	诊断准确率	训练复杂度	推理速度	可解释性	适用场景
孤立森林	中（~85%）	低	快	中	多维指标异常检测
LSTM/LSTM-AE	高（89%~92%）	高	中	低	时序异常检测与预测
CNN-LSTM	很高（~96%）	很高	慢	低	多变量时序故障预测
GNN	高（~89%）	高	慢	中	微服务根因定位
Transformer/LLM	高（80%~90%）	很高	慢	低	日志语义分析与预测
统计方法（ 3σ /EWMA）	低~中	极低	极快	高	单变量阈值告警补充

分析：统计方法虽然准确率有限，但计算开销极小且可解释性强，适合作为基础告警的快速过滤层。孤立森林在无监督场景中优势明显，训练和推理效率高，适合实时异常检测。LSTM 系列在时序预测任务中表现优异但可解释性较差。GNN 在根因定位场景中具有不可替代性，但对拓扑数据的依赖较强。Transformer/LLM 在语义理解上优势突出，但算力成本较高。

3.2 路径选择策略

建议采用分层选型策略：在数据量大、实时性要求高的异常检测层优先选择孤立森林等轻量级算法；在预测精度要求高

的故障预测层采用 LSTM 或 CNN-LSTM；在需要语义理解的日志分析场景中探索引入 LLM；在根因定位场景中采用 GNN 与知识图谱相结合的方法。此外，应构建模型评估机制，通过准确率、召回率和 F1 分数等指标持续监控模型表现并迭代优化。

3.3 LSTM预测模型原理及公式

LSTM 的核心设计思想是通过“门控机制”与“单元状态”协同作用，实现对时序信息的精准控制，从而记住长期有用的信息、遗忘无关冗余信息，并基于当前输入与历史记忆生成合理的输出，最终完成时序预测任务。LSTM 模型的核心组件包

括3个门控单元（遗忘门（Forget Gate）、输入门（Input Gate）、输出门（Output Gate））和1个单元状态（Cell State），各组件的功能如下：

遗忘门：遗忘门的核心作用是根据当前输入与上一时刻的短期记忆，判断上一时刻的长期记忆（单元状态）中哪些信息需要保留、哪些需要遗忘，其公式如下所示：

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

公式解释：首先将上一时刻的隐藏状态与当前时刻的输入进行拼接，得到拼接向量；随后将该拼接向量与遗忘门权重矩阵进行矩阵乘法运算，实现线性变换，再加上遗忘门偏置，得到线性组合结果；最后通过 sigmoid 激活函数将线性组合结果压缩至 $[0, 1]$ 区间，得到遗忘门输出。遗忘门输出中每个元素的取值对应单元状态中对应位置信息的保留概率：当元素值接近1时，代表该位置的历史信息完全保留；当元素值接近0时，代表该位置的历史信息完全遗忘；当元素值介于0和1之间时，代表该位置的历史信息按对应概率部分保留。

输入门：输入门的核心作用是筛选当前输入与上一时刻短期记忆中的有效新信息，决定哪些新信息可以存入单元状态，其公式如式 (2) 所示：

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

公式解释：输入门的结构与遗忘门完全一致，区别仅在于使用的权重矩阵与偏置向量不同，均为模型训练过程中独立学习得到。同样通过拼接上一时刻的隐藏状态与当前时刻的输入、线性变换、sigmoid 激活的步骤，得到输入门输出，其取值范围仍为 $[0, 1]$ 。输入门中每个元素的取值对应候选单元状态中对应位置新信息的存入概率：当元素值接近1时，代表该位置的新信息完全存入单元状态；当元素值接近0时，代表该位置的新信息完全不存入；当元素值介于0和1之间时，代表该位置的新信息按对应概率部分存入。

候选单元状态：候选单元状态是当前时刻输入与上一时刻短期记忆共同生成的、待存入单元状态的新信息集合，其公式如式 (3) 所示：

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$$

公式解释：首先同样对上一时刻的隐藏状态与当前时刻的输入进行拼接，得到拼接向量；随后将该向量与候选单元状态权重矩阵进行矩阵乘法运算，加上偏置完成线性变换；最后通过 tanh 激活函数将线性组合结果压缩至 $[-1, 1]$ 区间，得到候选单元状态。

单元状态：单元状态是 LSTM 的长期记忆载体，其更新过程是结合过滤后的历史记忆与筛选后的新信息，生成当前时刻的长期记忆，是 LSTM 能够捕捉长期依赖的核心步骤，公式如

式 (4) 所示：

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t$$

公式解释：该公式由两部分组成，分别对应“历史记忆的保留”与“新信息的加入”，两部分通过加法运算完成单元状态的更新。第一部分：通过遗忘门输出与上一时刻单元状态进行哈达玛积运算，实现对历史长期记忆的过滤，保留有价值的历史信息，过滤冗余信息。第二部分：通过输入门输出与候选单元状态进行哈达玛积运算，实现对新信息的筛选，将有价值的新信息筛选出来，准备存入单元状态。两部分相加后，得到当前时刻的单元状态，其物理意义为：当前时刻的长期记忆 = 过滤后的历史长期记忆 + 筛选后的当前新信息。这种更新方式确保了单元状态能够持续传递有价值的信息，同时避免了冗余信息的积累，有效缓解了传统 RNN 的梯度消失问题。

输出门：输出门的核心作用是控制当前时刻的单元状态（长期记忆）中，哪些信息可以转化为当前时刻的隐藏状态（短期记忆），并传递给下一时刻或用于预测，其公式如式 (5) 所示：

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

公式解释：输出门的结构与遗忘门、输入门一致，使用独立的权重矩阵与偏置，通过拼接上一时刻的隐藏状态与当前时刻的输入、线性变换、sigmoid 激活的步骤，得到输出门输出，取值范围为 $[0, 1]$ 。输出门输出中每个元素的取值对应单元状态中对应位置信息的输出概率：当元素值接近1时，代表该位置的长期记忆完全转化为短期记忆（隐藏状态）；当元素值接近0时，代表该位置的长期记忆不输出；当元素值介于0和1之间时，代表该位置的长期记忆按对应概率部分输出。

隐藏状态：隐藏状态是当前时刻 LSTM 的核心输出，既是传递给下一时刻的短期记忆，也是后续预测层的输入，其公式如式 (6) 所示：

$$h_t = o_t \odot \tanh(C_t)$$

公式解释：首先将当前时刻的单元状态通过 tanh 激活函数压缩至 $[-1, 1]$ 区间，进一步标准化长期记忆的取值范围；随后将标准化后的单元状态与输出门输出进行哈达玛积运算，通过输出门的控制，筛选出需要输出的信息，最终得到当前时刻的隐藏状态。

预测输出：在时序预测任务中，LSTM 的最终目的是基于历史时序数据预测未来时刻的数值，因此需要在 LSTM 的输出端添加一个全连接层（预测层），将最后一个时刻的隐藏状态映射为预测值，公式如式 (7) 所示：

$$\hat{y}_{t+1} = W \cdot h_t + b$$

公式解释：将 LSTM 最后一个时刻的隐藏状态（维度）与预测层权重矩阵（维度）进行矩阵乘法运算，再加上预测层偏

置(标量),得到下一时刻的预测值。若为多变量时序预测,只需调整预测层权重矩阵的维度,即可得到多个变量的预测值。该公式是LSTM实现时序预测的关键,将模型学习到的时序特征转化为具体的预测结果。

LSTM的预测流程可概括为:基于历史时序数据,通过上述门控机制与单元状态的动态更新,逐步学习时序数据的内在规律,最终利用最后一个时刻的隐藏状态,通过线性映射得到下一时刻的预测值完成单步时序预测;若需实现多步预测,可将前一步的预测值作为输入,迭代生成后续时刻的预测结果。长短期记忆网络(LSTM)通过遗忘门、输入门、输出门三大门控单元与单元状态的协同作用,实现了对时序信息的选择性保留、遗忘与更新,有效解决了传统RNN的梯度消失问题,能够稳定捕捉时序数据中的长期依赖关系。

4 技术管理

4.1 数据质量管理

数据质量是决定AI模型效果的首要因素。在实际运维环境中,数据采集常面临缺失值、异常值、时间戳错位等问题。应建立完善的数据质量监控机制,自动识别空值、突变和漂移等异常数据,触发告警或自动修复流程。同时,需要做好数据标注工作,建立标准化的故障模式库(涵盖200余个典型场景),为监督学习模型提供高质量训练样本。

4.2 模型持续迭代管理

参考文献

- [1] 集团智能运维基于AI驱动的自动化故障预测与自愈系统[EB/OL]. dtstack.com, 2026-03-28.
- [2] AIOps for log anomaly detection in the era of LLMs: A systematic literature review[J]. Journal of Industrial Information Integration, 2025, 45: 100732.
- [3] 集团智能运维基于AI驱动的自动化监控与故障预测[EB/OL]. dtstack.com, 2026-03-27.
- [4] 基于机器学习的AIOps故障预测与自动化运维实践[EB/OL]. dtstack.com, 2025-06-25.
- [5] 集团智能运维基于AIOps的自动化故障预测与根因分析[EB/OL]. dtstack.com, 2026-03-29.
- [6] AI-based predictive maintenance for software-intensive systems: A survey of methods, challenges, and future directions[J]. IEEE Access, 2024, 12: 56789-56812.
- [7] 华为发布AUTINOps解决方案,重新定义智能运维新范式[EB/OL]. huawei.com, 2026-03-03.
- [8] Comparison of deep learning models for predictive maintenance in industrial manufacturing systems using sensor data[J]. Scientific Reports, 2025, 15:23545.
- [9] LLM for AIOps: 是泡沫还是银弹? 智能运维的争议、破局与未来[EB/OL]. infoq.cn, 2025-12-17.

作者简介:尹剑锐(1987—),男,满族,硕士,研究方向为信息系统管理。

智能运维模型并非一劳永逸。随着系统架构演进和业务模式变化,模型的预测性能会逐渐衰减。应建立模型评估与迭代机制,定期使用新采集的数据对模型进行重新训练和评估,确保模型持续适应系统运行状态的变化。此外,应建立模型版本管理机制,对不同版本的模型进行性能追踪和A/B测试。

4.3 安全合规与风险管控

AI模型在运维中的应用需要兼顾安全合规。大语言模型在推理过程中可能产生“幻觉”——看似正确实则错误的诊断结论,需要通过护栏机制(Guardrail)约束模型行为。同时,应建立人机协同机制:智能系统虽能自动处理大部分常规任务,但关键变更操作仍需人工审批,形成“AI推荐+人工决策”的双重保障模式。

5 结语

人工智能技术在信息系统运维中的应用,正在推动运维模式从“被动救火”向“主动预防”深刻转型。本文从整体规划、选型建设、多路径对比和技术管理四个维度,系统探讨了智能诊断与预测模型的构建路径。实践表明,基于分层架构的模型体系、场景适配的算法选型、持续迭代的管理机制,是实现智能运维规模化落地的关键。随着大模型技术的快速发展和运维数据资产的持续积累,具备自主诊断与预测能力的智能运维系统将成为企业数字化转型的核心基础设施。