基于人工智能的电力信息系统异常数据检测研究

史上乐

（国网四川省电力公司成都供电公司，四川 成都 610041）

摘 要 本文以人工智能技术为支撑，面向电力信息系统异常数据检测领域展开系统性研究，立足于电力信息系统异常数据检测场景下的人工智能技术应用体系，分别梳理不同关键技术的落地应用要点，包括基于机器学习的电力信息系统基础运行类异常数据检测、基于深度学习的电力信息系统高维时序与设备监测类异常数据检测、基于混合智能检测技术的电力系统信息多源交互场景下复合型异常数据检测等，在此基础上研判人工智能技术在该领域的未来发展趋势，以期为后续相关研究提供参考。

关键词 电力信息系统；异常数据检测；人工智能技术；机器学习；深度学习

中图分类号：TM73；TP393.08；TP18 **文献标志码**：A **DOI**：10.3969/j.issn.2097-3365.2026.27.001

0 引言

电力信息系统主要负责电流、电压和功率等关键数据的采集、存储以及传输，其运行效果会对电网调度、设备运维以及能源优化等产生直接影响。如果电力信息系统中存在异常数据，将会使调度决策出现失误或者是设备故障漏判，严重时甚至会引起一系列连锁事故，对电网安全性造成不良影响。针对异常数据，传统检测方法主要以人工阈值设定和规则匹配等方式实现，在自适应能力、检测精度和复杂数据处理中都存在弊端，无法和现代电力信息系统实际运维需求相适配。因此，引入先进的人工智能技术，为电力信息系统的异常数据检测提供支持具有重要意义。

1 电力信息系统异常数据检测中的人工智能技术体系

1.1 机器学习检测技术

机器学习检测技术的主要算法有三种：一是支持向量机，该算法能够在高维特征空间中寻找到最优分类超平面，以此来精准划分正常和异常数据，具有较强抗干扰能力和优异的泛化性能；二是孤立森林。通过多棵孤立树的合理构建，随机进行特征选择以及阈值分割，让偏离正常范围的异常数据被快速孤立出来，这种算法不需要正态分布假设，具有更高效率，和电力系统的实时检测需求相适配；三是 K 近邻算法，它是将“物以类聚”作为基本思想，对被检测数据和样本集中的数据距离进行计算，根据近邻样本类别来判断数据是否存在异常，这种算法对数据分布没有严格要求，能够对电力数据动态变化做出快速适应，在短期电力信息系统异常数据检测中非常适用[1]。

1.2 深度学习检测技术

深度学习检测技术的主要支持技术有三种：一是长短期记忆网络，该技术主要利用输入门、遗忘门以及输出门之间的协同作用，对电力时序数据长短期依赖关系进行记忆，以此来捕捉电流和电压等时序数据，明确其趋势异常，它的时序建模能力和抗噪声干扰能力都比较强，能有效适配电力信息系统的连续运行模式；二是自编码器，该技术主要利用编码器来压缩输入数据，使其转变成低维特征向量，再借助解码器对原始数据加以重构，并根据重构误差来判断数据有无异常，该技术不需要标注数据就能够实现正常模式的自主学习，在电力异常数据未标注的检测场景下适用；三是卷积神经网络，利用卷积层和池化层对电力数据全局和局部特征进行提取，以此来精准识别异常特征，它有着特征提取能力强和高维度数据处理等优势，在电力设备和多元传感等异常数据监测中都非常适用。

1.3 混合智能检测技术

混合智能检测技术是将机器学习和深度学习优势混合在一起形成的复合型技术体系，借助多种技术的协同作用，能够使单一技术局限得到有效弥补，进一

作者简介：史上乐（1994-），男，硕士研究生，工程师，研究方向：信息技术及数据管理。

步提高异常数据检测精度及其适应性。它的主要优点是能够使检测效率和精度得到全面兼顾，有着更强的自适应能力，可以有效适配电力信息系统不同场景与类型下的数据检测需求[2]。

2 电力信息系统异常数据检测中的人工智能技术应用要点

2.1 基于机器学习技术的电力信息系统基础运行异常数据检测

在电力信息系统中，基础类运行异常数据有着较低维度，样本分布也比较清晰，与机器学习算法最典型的轻量化检测逻辑相适配[3]。处理时的主要流程有数据预处理、模型训练、异常识别和结果检验等，基本适配情况如表1所示。

表1 三种机器学习算法在电力信息系统异常数据检测中的基本适配情况

序号	算法类型	异常数据检测类型
1	支持向量机算法	小样本隐蔽运行异常数据检测
2	孤立森林算法	海量实时运行异常数据检测
3	K近邻算法	短期时序运行异常数据检测

支持向量机主要能检测小样本隐藏类异常数据，如设备故障、初级产生的微小电压偏差或者隐性功率波动等。它的主要分析步骤有四个：

1. 剔除原始运行数据中的冗余数据，插值补充缺失数据，滤除其中的高斯噪声，采取以下归一化处理法将量纲差异消除，将数据映射到[0.1]区间。

$$x_{morm}=\frac{x-x_{\min}}{x_{\max}-x_{\min}} \tag{1}$$

式（1）中，x_{morm}代表归一化处理后映射到[0.1]区间内的标准化数据；x为电力信息系统的原始运行数据；$x_{\min}$为原始运行数据中的最小值；$x_{\max}$为原始运行数据中的最大值。通过求解这个标准化数据，就能够将量纲差异有效消除。

2. 将电流波动幅度、电压偏差数据、有功功率波动等选作核心特征，完成低维度特征向量建立。

3. 将径向基函数选作核函数，采取5折交叉法对惩罚系数以及核函数参数进行校验优化，让分类误差达到最小，以下是核函数公式：

$$K(x_i,y_i)=\exp(-\frac{|x_i-x_j|^2}{2\sigma^2} \tag{2}$$

式（2）中，$K(x_i,y_i)$为核函数；K是惩罚系数；(x_i,y_i)是不同样本特征向量；$\|x_i-y_i\|$是向量的欧式距离；σ是核函数参数。

4. 在模型中输入检测数据，由模型将分类结果输出，如果判定为异常数据，则需要根据电网实际运行规程对具体异常类型加以验证。

孤立森林主要用来检测海量实时异常数据，如传输链路短时干扰导致的随机异常、电网调度端瞬时采集的数据异常等，它的主要检测步骤有四个：

第一，从每个批次实时数据中随机抽出200个～500个作为样本，以此来合理减轻计算负担。

第二，设置随机决策树100棵～200棵，以此来对数据样本实施递归分割处理。

第三，通过以下公式对样本的孤立分数进行计算，分数越高代表数据有着越大的异常概率，并将基础阈值设为0.7，如果超出这个阈值，就直接判定成异常数据。

$$s(x,n)=2^{-\frac{E[h(x)]}{C(n)}} \tag{3}$$

式（3）中，$s(x,n)$代表样本的孤立分数；$E[h(x)]$代表样本路径平均长度；$C(n)$代表n个样本路径的标准长度。

第四，根据历史实时异常数据，对阈值进行微调处理，以此来降低误报率。

K近邻可以检测短期时序异常数据，比如短时间内的功率间歇性波动或者电流连续超出额定范围等情况，处理时先合理确定最优K值，一般在5～15之间。再对监测数据和样本数据之间的欧式距离进行计算，根据计算机筛选出K个近邻样本，如果异常样本的占比在50%以上，则可以判定为异常数据，并进一步通过KD树来提高计算效率。

2.2 基于深度学习技术的电力信息系统高维时序和设备监测异常数据检测

现代电力信息系统高维复杂类异常数据主要包括持续连续运行异常和设备监测异常这两种。前者属于非线性数据，后者有着多特征耦合以及高维特点[4]。

具体检测时，长短期记忆网络可以对长时序连续运行异常数据进行检测。先将1 h确定为时间窗口，对时序样本进行划分，通过以下Z-score标准化公式将量纲消除：

$$z=\frac{x-\mu}{\sigma} \tag{4}$$

式（4）中，z代表标准化处理后的时序数据；x代表单一时序数据；μ代表时序数据平均值；σ代表时序数据标准差。然后搭建长短期记忆网络模型，使输入层与时间窗口长度对应，隐藏层按1～2层设置，每层神经元数量设置为32个～64个，采用三门结构来记忆时序的依赖关系。之后将均方差用作损失函数，通过Adam优化其对模型进行迭代训练，基本损失函数公式如下：

$$L_{MSE}=\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 \quad (5)$$

式（5）中，L_{MSE} 表示损失函数值；n 表示样本总量；y_i 表示原始时序数据；$\hat{y}_i$ 表示模型重构数据。最后将计算出的重构误差和既定动态阈值进行对比，如发现误差超标，则可以判定为时序异常数据。

自编码器主要用来对无标注设备的监测异常数据进行检测，先将高维数据压缩成低维特征，通过解码器对原始数据进行重构处理，使其和正常数据的误差小到可以忽略不计，如果发现重构数据出现激增情况，激增幅度超过正常误差的 3 倍，则可以将对应数据判定为异常数据。

通过卷积神经网络可以对多元设备监测异常数据进行检测，先利用卷积层对局部异常特征进行提取，通过池化层进行降维处理，再利用全连接层进行特征整合，最后将异常概率输出。如果发现输出的异常概率在 0.8 以上，就可以将相应数据判定为异常数据，从而实现各类异常数据源头的精准定位。

2.3 基于混合智能检测技术的电力信息系统多源交互和复合型异常数据检测

针对复合型多元交互异常数据，单一的人工智能算法往往不能让检测效率和精度得到全面兼顾。因此，在具体检测过程中，电力单位需要将混合智能技术作为依托，通过两种或多种智能技术结合的方式进行协同检测[5]。将快速初筛和精准复核作为核心检测逻辑，以此来弥补单一智能检测技术的应用短板。

其中最主流的混合智能检测技术是随机森林和长短期记忆网络协同检测。先通过孤立森林算法快速筛选海量的多元交互数据，将疑似异常数据子集筛选出来，这样就能够大幅度降低后续检测中的数据量，提高精准检测效率。之后再将疑似异常的数据子集输入长短期记忆网络模型中，利用该模型的时序建模能力，对数据之间的依赖关系进行深度挖掘，从而使真实复合型异常以及噪声干扰等情况得到精准区分。最终将误判数据剔除，实现复合型异常情况的精准定位。

3 电力信息系统异常数据检测中的人工智能技术主要发展趋势

在当今电力系统的分布式、数字化和智能化转型过程中，电力信息系统的数据体量越来越大，涉及的异常数据类型也更加丰富且复杂。在该背景下，人工智能检测技术应该朝着更加轻量化、自适应化、协同化和在线迭代化的方向发展，各发展方向及其适配的异常数据检测类型如表 2 所示。

表 2 电力信息系统异常数据检测中的人工智能技术基本发展方向汇总

序号	发展方向	适配异常检测	优化要点
1	轻量化边缘智能监测	终端实时采集异常数据、短时传输异常数据等	算法轻量化、模型压缩，与边缘终端相适配
2	小样本自适应学习	新型未知异常数据、稀有隐性异常数据等	优化迁移学习模型，降低其对标注样本的依赖性
3	多模态数据协同检测	多源交互复合型异常数据	多源数据特征融合，进一步提高关联分析能力
4	实时在线迭代优化	突发异常数据、长时序波动异常数据等	动态化模型更新，使其和系统数据的动态变化相适配

4 结束语

将人工智能技术合理应用到电力信息系统异常数据检测过程中，能够及时、准确地发现异常情况，明确异常类型，为后续数据校准和调节等创造有利条件。因此，电力企业及相关从业者一定要重视人工智能技术的合理应用，根据实际情况和检测需求等，合理引入人工智能算法，通过机器学习和深度学习算法有机结合的方式，对信息系统异常数据进行检测，以此来提高检测效率和精度，满足电力信息系统的实际运维需求，促进其智能化发展。

参考文献：

[1] 李天慧,谢云澄,车荣花,等.基于DRSN-BiLSTM的电力信息网络入侵检测模型[J].电力信息与通信技术,2023,21(09):30-37.
[2] 孙泽锋.AI在电力信息化系统网络安全防护中的应用研究[J].中国高新科技,2025(12):104-105,108.
[3] 余梦.人工智能技术在电力营销信息系统中的创新应用研究[J].模型世界,2025(10):113-115.
[4] 杨兆威,孙玉龙.人工智能技术在电力营销信息系统中的应用[J].消费电子,2025(13):158-160.
[5] 高僮.人工智能技术在电力营销信息系统中的创新应用[J].高科技与产业化,2024,30(12):82-84.