

基于随机森林和多任务学习的电力设备异常状态预警研究

郭红玉¹ 殷 刚¹
GUO Hongyu YIN Gang

摘 要

在电力设备运行状态监测中,多源异构数据的高维耦合特征与设备退化过程中故障模式的弱关联性,导致计算电力设备运行数据特征的基尼重要性评分不准确,无法提取运行数据重要特征;同时,传统的单任务学习方法难以有效整合多维度信息(如振动、温度、电流电压等),导致异常评分不够精准,致使其异常状态预警性能较差。基于此,提出基于随机森林和多任务学习的电力设备异常状态预警方法。利用随机森林算法计算电力设备运行数据特征的基尼重要性评分,筛选出重要性评分大于预设阈值的关键特征,筛选运行数据重要特征。这些筛选出的重要特征作为多任务学习模型的输入,构建多任务学习模型,通过主任务(电力设备整体异常概率预测)与辅助任务(振动分析、温度监测、电流电压稳定性评估)的联合训练机制,充分整合电力设备运行状态的多维度信息,从而生成更精准的异常评分。以异常评分为依据,划分电力设备异常状态预警级别(不预警、低级别预警、中级别预警与高级别预警),实现异常状态的精细化分级预警。实验结果显示:设计方法应用后特征重要性评分计算误差最小值达到了0.3%,多任务共性特征提取结果与实际共性特征相同,电力设备运行状态异常评分与实际异常评分趋于一致。

关键词

多任务学习模型;异常状态;随机森林算法;电力设备;预警机制制定

doi: 10.3969/j.issn.1672-9528.2025.08.027

0 引言

电力系统规模的不断扩大和智能化水平的提升,电力设备运行状态的实时监测与异常预警成为保障电网安全稳定运行的关键。然而,电力设备运行数据来源多样、结构复杂,且设备退化过程中故障特征与运行状态之间的关联性较弱,导致传统方法难以从海量数据中有效提取关键特征;同时,单一任务的学习模型难以全面捕捉设备运行的多维度信息,导致异常评分的准确性和全面性不足,预警效果不理想。

其中,文献[1]利用时间序列自回归模型和自组织映射神经网络,将连续数据离散化,并通过计算状态转移概率和聚类算法实现快速异常检测。设计方法依赖于大量的历史数据和精确的时间序列模型。该方法无法从高维数据中筛选出对异常状态敏感的关键特征,导致特征提取过程效率低下,难以满足精准预警的需求。文献[2]采用独立成分分析方法降噪,提升运算效率;构建深度学习前馈神经网络,实现时间序列的异常检测。但该方法主要聚焦于时间序列数据,缺乏对电力设备运行状态多维度信息(如振动、温度、电流电压等)的有效整合,导致异常评分的全面性和精准性不足。

文献[3]基于宽频振动传感器进行电力设备监测,其能够同时捕捉1 Hz~500 kHz的振动信号,有效监测变压器、电容器等状态。设计方法在实际应用中可能面临复杂电磁环境干扰,影响振动与超声信号的准确捕捉,从而影响电力设备监测的准确性。文献[4]集成4项核心技术——传感采集、信号处理、数据分析及信息传输,设计了新型电力系统设备状态监测与故障诊断传感芯片,以此为基础,实现了电力关键设备的状态监测与故障诊断。该方法特征提取过程效率不高,难以满足精准预警的需求。

针对上述挑战,提出一种基于随机森林和多任务学习的电力设备异常状态预警方法。

1 电力设备异常状态预警方法

1.1 电力设备运行数据重要特征筛选

由于数据来源多样、结构复杂,且设备退化过程中故障特征与运行状态之间的关联性较弱,传统方法难以从海量数据中有效提取关键特征,导致异常状态预警的精准性和全面性不足^[5]。针对这一问题,通过随机森林的特征重要性评估机制,从高维数据中有效分离出对异常状态敏感的重要特征,解决了高维耦合特征下关键特征子集提取困难的问题,为后续异常状态预警模型提供了高质量的特征输入,从而提升了

1. 朔州职业技术学院能源与资源工程系 山西朔州 036002

预警的精准性和可靠性。

设定原始电力设备运行数据为 $X = \{x_1, x_2, \dots, x_p, \dots, x_n\}$ ，其中， x_i 代表第 i 个电力设备运行数据， n 代表电力设备运行数据的总数量^[6]。通过统计技术、傅里叶变换算法、小波包分解算法等提取电力设备运行数据的时域特征、频域特征、时频域特征与电力设备物理特性构造特征，获取电力设备运行数据特征集合 $F = \{f_1, f_2, \dots, f_j, \dots, f_m\}$ ，其中， f_j 代表第 j 个电力设备运行数据特征， m 代表电力设备运行数据特征的总数量。

基于随机森林算法对电力设备运行数据的重要特征进行筛选，具体过程为：

步骤 1：构建随机森林模型

使用提取的电力设备运行数据特征集合 $F = \{f_1, f_2, \dots, f_j, \dots, f_m\}$ 和目标变量（电力设备状态标签——正常或者异常）训练随机森林模型。需要注意的是，每棵决策树在构建过程中随机选择部分特征进行节点分裂。

步骤 2：计算特征重要性评分

对于每个特征 f_j ，计算其基尼重要性，即特征重要性评分^[7]。基尼重要性主要是衡量特征在节点分裂时对数据集纯度的提升程度，计算公式为：

$$\alpha(f_j) = \sum_{t \in T} \Delta \alpha(t, f_j) \cdot \frac{N_t}{N} \quad (1)$$

式中： $\alpha(f_j)$ 代表特征 f_j 的基尼重要性，即特征重要性评分； t 代表某棵决策树中的一个节点； T 代表所有决策树的集合； $\Delta \alpha(t, f_j)$ 代表在节点 t 使用特征 f_j 进行分裂后，基尼系数的减少量； N_t 代表节点 t 中的样本数量； N 代表总样本数量。

步骤 3：特征排序与重要特征筛选

根据特征重要性评分 $\alpha(f_j)$ 对所有特征排序，选择重要性评分大于阈值 $\hat{\alpha}$ （需要根据电力设备异常状态预警需求进行设置）的特征作为重要特征，对其整合处理，获得电力设备运行数据重要特征集合 $\hat{F} = \{f_1, f_2, \dots, f_j, \dots, f_k\}$ ， k 代表重要特征的总数量^[8]。

步骤 4：验证重要特征筛选结果

使用筛选后的重要特征集合 $\hat{F}$ 重新训练模型，并评估其性能。如果性能显著下降，则调整筛选阈值或者重新选择特征。

基于随机森林算法从高维电力设备运行数据特征中提取出对异常状态敏感的重要特征，为后续异常状态预警模型提供高质量的特征输入。

1.2 基于多任务学习的电力设备运行状态异常评分生成

传统的单任务学习方法通常只关注某一特定指标，而忽略了电力设备运行状态的多维度特性，致使其异常状态预警性能较差。而多任务学习通过同时学习多个相关任

务，能够整合振动分析、温度监测、电流电压稳定性评估等多种评估任务，从而全面描述电力设备的运行状态，并生成更准确的异常评分，可以最大限度地提升异常状态预警性能。因此，以筛选出来的电力设备运行数据重要特征 $\hat{F} = \{f_1, f_2, \dots, f_j, \dots, f_k\}$ 为基础，构建多任务学习模型。

在电力设备异常状态预警中，多任务学习模型主要包括几个任务：

一是主任务：电力设备整体异常概率预测。

二是辅助任务：

(1) 振动分析：评估电力设备振动信号的异常程度；

(2) 温度监测：预测电力设备关键部位的温度变化；

(3) 电流电压稳定性评估：分析电流电压波动的稳定性。

三是通过联合优化多个任务，构建模型能够更好地捕捉电力设备运行状态的多维度信息，从而生成更全面、更精准的异常评分^[9]。

多任务学习模型应用过程为：

Step 1：共享层

采用卷积层提取多个任务之间的共性特征，表达式为：

$$h_0 = W_0 \cdot \hat{F} + b_0 \quad (2)$$

式中： h_0 代表共享特征； W_0 与 b_0 代表共享特征提取过程中的权重矩阵与偏置项。

Step 2：任务特定层

多任务学习模型中，每个辅助任务有独立的网络分支，用于学习辅助任务特定的特征，为后续的任务执行提供特征数据支撑，表达式为：

$$h_i = \frac{\zeta[W_i \cdot \hat{F} + b_i]}{\beta^e} \quad (3)$$

式中： h_i 代表第 i 个辅助任务对应的特定特征。 i 取值为 1~3； $\zeta[\cdot]$ 代表特定特征提取函数； W_i 与 b_i 代表第 i 个辅助任务特定特征提取过程中的权重矩阵与偏置项； β^e 代表特定特征标准化因子。

Step 3：输出层

依据 Sigmoid 激活函数计算辅助任务的评分：

$$s_i = \exp\left[\frac{h_0 + h_i \pm \varepsilon_i}{\chi_i}\right] \quad (4)$$

式中： s_i 代表第 i 个辅助任务的评分； $\exp[\cdot]$ 代表 Sigmoid 激活函数； χ_i 代表第 i 个辅助任务评分的计算常数，取值范围为 0~1； ε_i 代表误差调整项，决定着辅助任务评分的精准性。

将辅助任务评分进行加权求和，即可获得电力设备运行状态异常评分（即主任务输出结果）：

$$S = \sum_{i=1}^3 \omega_i \times s_i \quad (5)$$

式中： S 代表电力设备运行状态异常评分； ω_i 代表辅助任务

评分 s_i 对应的权重系数。

通过上述步骤，获得了电力设备运行状态异常评分，为最终研究目标实现提供助力。

1.3 异常状态预警机制制定

以异常评分 S 为依据，划分电力设备异常状态预警级别（不预警、低级别预警、中级别预警与高级别预警），实现了异常状态的精细化分级预警。通过分级预警机制，能够更精确地反映设备运行状态的异常程度，为运维人员提供差异化的决策支持，从而优化资源配置，提升故障处理的效率，并有效降低设备故障带来的风险和损失。

- （1）当电力设备运行状态异常评分 S 在 $[0, 0.3)$ 范围内时，认定电力设备运行状态为正常，不触发预警机制，即预警灯处于关闭状态。
- （2）当电力设备运行状态异常评分 S 在 $[0.3, 0.5)$ 范围内时，认定电力设备运行状态为轻微异常，触发低级别预警机制，即预警灯显示为绿色。
- （3）当电力设备运行状态异常评分 S 在 $[0.5, 0.8)$ 范围内时，认定电力设备运行状态为一般异常，触发中级别预警机制，即预警灯显示为黄色。
- （4）当电力设备运行状态异常评分 S 在 $[0.8, 1]$ 范围内时，认定电力设备运行状态为显著异常，触发高级别预警机制，即预警灯显示为红色^[10]。

综上所述，在随机森林算法与多任务学习模型的联合应用下，实现了电力设备异常状态的精准预警。

2 实验与结果分析

2.1 实验环境搭建

依据实验需求，搭建实验环境，如图 1 所示。

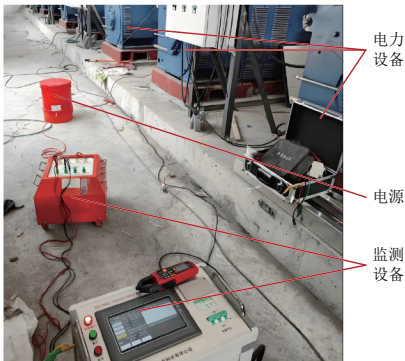


图 1 实验环境示意图

选取多个电力设备作为实验对象，获取其运行状态数据，对其进行预处理，并将其随机划分为 6 个实验组别，为后续实验顺利进行提供一定的便利，具体如表 1 所示。需要注意的是，每个实验组别进行 10 次测试，并且获取相应的实验数据，通过此方式可以大幅提升实验结论的可信度。

表 1 实验数据表

实验组别编号	电力设备数量 / 个	数据体量 / MB
1	15	56.23
2	20	64.12
3	11	50.11
4	24	75.02
5	21	50.12
6	19	53.62

2.2 多任务学习模型训练

在实验进行之前，对多任务学习模型进行训练：

- （1）初始化参数：随机初始化共享层和任务特定层的参数。
- （2）前向传播：计算共享特征 h_0 和任务特定特征 h_i 。
- （3）损失计算：根据损失函数计算总损失：

$$L_{\text{total}} = \sum_{i=1}^3 \kappa_i \times L_i$$

(6)

式中： L_{total} 代表总损失； L_i 代表第 i 个辅助任务的损失； κ_i 代表第 i 个辅助任务损失的权重系数，用于平衡不同任务的重要性。

- （4）反向传播：通过梯度下降法更新模型参数。
- （5）迭代优化：重复上述步骤直到模型收敛为止。

2.3 特征重要性评分计算误差分析

应用设计方法、对比方法 1（基于时间序列和神经网络的电力设备异常状态预警方法）与对比方法 2（基于 ICA-FNN 的电力设备异常状态预警方法）获取特征重要性评分，并计算其与实际评分之间的误差，如图 2 所示。

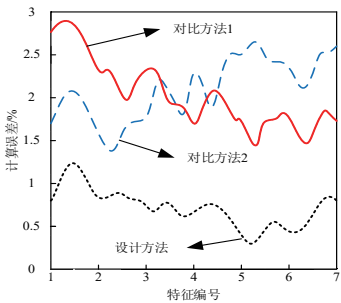


图 2 特征重要性评分计算误差示意图

相较于对比方法 1 与对比方法 2 来看，设计方法应用后特征重要性评分计算误差显著较小，最小值达到了 0.3%。这主要是因为设计方法引入了随机森林算法对特征重要性评分进行计算，其不仅能够有效处理高维数据，还能减少过拟合的风险，从而在特征筛选过程中提供更加稳健和准确的评分。

2.4 多任务共性特征提取性能分析

应用设计方法、对比方法 1 与对比方法 2 提取多任务共性特征，其是电力设备运行状态异常评分生成的关键依据之

一,如图3所示。

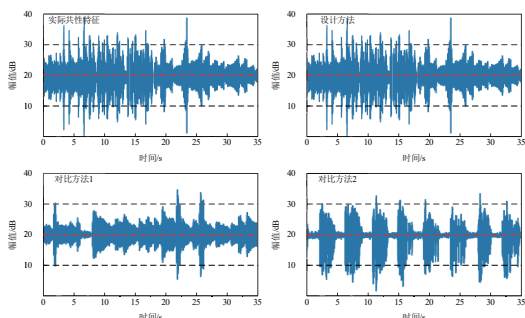


图3 多任务共性特征提取结果示意图

设计方法提取的多任务共性特征与实际共性特征相同,而对比方法1与对比方法2提取的多任务共性特征与实际共性特征差异性较大,主要是因为设计方法应用了多任务学习模型,其具备较强的关键特征提取功能,从而大幅度提升了设计方法多任务共性特征的提取性能。

2.5 异常状态预警结果分析

应用设计方法、对比方法1与对比方法2进行电力设备异常状态预警实验,获取电力设备运行状态异常评分,如图4所示。

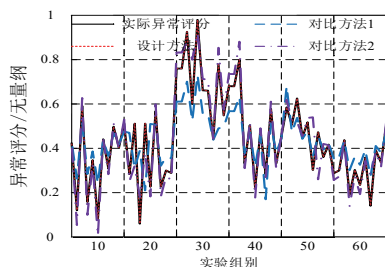


图4 电力设备运行状态异常评分示意图

设计方法应用后获得的电力设备运行状态异常评分与实际异常评分趋于一致,而对比方法1与对比方法2应用后获得的电力设备运行状态异常评分与实际异常评分存在着较大的偏差。这主要是因为设计方法引入了多任务学习模型计算电力设备运行状态异常评分,其核心优势在于其能够同时处理多个相关任务,通过共享层来捕捉不同任务间的潜在关联信息,能与之相关的其他任务中提取有用信息,从而更全面、深入地理解设备的运行状态。

3 结语

在电力设备的运行周期内,随着使用环境的变化、部件的老化以及操作维护的不当,电力设备性能会逐渐下降,异常状态悄然滋生。异常状态预警方法通过持续监测设备运行数据,运用先进的算法模型分析数据中的异常波动,能够在故障发生前捕捉到预警信号,为维修人员争取到宝贵的抢修时间,有效避免或减轻故障带来的后果。因此,提出基于随

机森林和多任务学习的电力设备异常状态预警研究。实验结果显示:设计方法大幅降低了特征重要性评分计算误差,提升了多任务共性特征提取性能与电力设备运行状态异常评分的精度。

参考文献:

- [1] 丁江桥,文屹,吕黔苏,等.基于时间序列和神经网络的电力设备状态异常检测方法[J].电测与仪表,2024,61(2):185-190.
- [2] 闻宇,陈艳霞,李菁,等.一种基于ICA-FNN的多模型高压网络保护设备异常状态风险预警方法[J].电力科学与技术学报,2024,39(4):78-83.
- [3] 相晨萌,高树国,刘宏亮,等.基于宽频振动传感器的电力设备状态监测方法可行性分析[J].高压电器,2024,60(4):83-91.
- [4] 秦潘昊,陈威宇,胡秦然,等.新型电力系统设备状态监测与故障诊断传感芯片关键技术与展望[J].电力系统自动化,2024,48(6):83-95.
- [5] 黄旭锐,于丰源,杨波,等.基于Transformer网络和多任务学习的园区综合能源系统电-热短期负荷预测方法[J].南方电网技术,2023,17(1):152-160.
- [6] 王文森,贺馨仪,杨晓西,等.基于多参量数据回归分析的电力变压器状态监测方法[J].电网与清洁能源,2023,39(4):83-90.
- [7] 计蓉,侯慧娟,盛戈皞,等.基于组合赋权法和模糊综合评价的电力设备状态数据质量评估[J].高电压技术,2024,50(1):274-281.
- [8] 谢庆,汲胜昌,张晓星,等.数字化技术在输变电设备状态评估中的应用专题特约主编寄语[J].电工技术学报,2023,38(1):13-14.
- [9] 赵金币,琚理.基于多任务学习的文本信息关联性抽取仿真[J].计算机仿真,2023,40(1):315-318.
- [10] 秦潘昊,陈威宇,胡秦然,等.新型电力系统设备状态监测与故障诊断传感芯片关键技术与展望[J].电力系统自动化,2024,48(6):83-95.

【作者简介】

郭红玉(1982—),女,山西朔州人,硕士,讲师,研究方向:电工电子技术、发电厂及电力系统。

殷刚(1979—),男,山西孝义人,硕士,讲师,研究方向:电路学、发电厂及电力系统。

(收稿日期:2025-03-05 修回日期:2025-07-29)