

引文格式:王金浩,李瑞,韩肖清,等. 基于改进XGBoost的电力系统暂态稳定评估方法及其可解释分析[J]. 太原理工大学学报, 2026, 57(2): 264-277.

WANG Jinhao, LI Rui, HAN Xiaoqing, et al. Power system transient stability assessment method based on improved XGBoost and its interpretable analysis[J]. Journal of Taiyuan University of Technology, 2026, 57(2): 264-277.

基于改进XGBoost的电力系统暂态稳定评估方法及其可解释分析

王金浩¹, 李 瑞¹, 韩肖清², 曲 莹¹, 常泽州¹, 薄利明¹, 牛哲文²

(1. 国网山西省电力公司 电力科学研究院, 山西 太原; 2. 太原理工大学 电气与动力工程学院, 山西 太原)

摘 要:【目的】随着新能源和电力电子设备的大规模并网, 新型电力系统安全稳定特性面临严峻挑战, 而数据驱动方法为建立暂态稳定评估模型提供了可行思路, 但由于模型自身的黑箱性决定了其决策依据的不可知, 成为限制其在线应用的关键因素。【方法】针对以上问题, 本文提出基于改进XGBoost的电力系统暂态稳定评估模型及其可解释性方法。一方面, 通过提升XGBoost模型对决策边界样本的学习能力, 在满足评估速度要求的前提下显著提升模型的暂态稳定评估精度。另一方面, 为提高评估结果的可解释性, 基于沙普利加性原理提出特征—样本可解释方法, 对模型的评估结果进行归因分析。【结果】在IEEE 39节点系统和某省级电网仿真结果表明, 所提方法相较于传统模型具有更高的评估精度, 同时具备良好的可解释能力。

关键词: 暂态稳定评估; 机器学习; 可解释性; 沙普利加性

中图分类号: TM712; TP18 **文献标识码:** A

DOI: 10.16355/j.tyut.1007-9432.20240591

文章编号: 1007-9432(2026)02-0264-14

Power System Transient Stability Assessment Method Based on Improved XGBoost and Its Interpretable Analysis

WANG Jinhao¹, LI Rui¹, HAN Xiaoqing², QU Ying¹, CHANG Zezhou¹, BO Liming¹, NIU Zhewen²

(1. Electric Power Science Research Institute of State Grid Shanxi Electric Power Company, Taiyuan, Shanxi, China;

2. College of Electrical and Power Engineering, Taiyuan University of Technology, Taiyuan, Shanxi, China)

Abstract: 【Purposes】With the large-scale grid-connection of new energy and power electronic devices, the security and stability characteristics of new power systems are facing serious challenges, and data-driven methods provide feasible ideas for the establishment of transient stability assessment models. However, the black-box nature of the models themselves determines their decision-making basis being not known, which has become a key factor limiting their online application. 【Methods】To address this issue, a power system transient stability assessment model based on improved XGBoost and its interpretability method were proposed. 【Results】On one hand, by improving the learning ability of the machine learning model on the decision boundary samples, the transient stability assessment accuracy of the model is significantly improved under the premise of meeting the assessment speed re-

收稿日期: 2024-07-24; 修回日期: 2024-10-25

基金项目: 国网山西省电力公司科技项目(52053023000B)

第一作者: 王金浩, 教授级高工, 主要从事电力系统分析研究, (E-mail) liuxinyuan1208@163.com

通信作者: 牛哲文, 博士, 讲师, 主要从事电力系统运行与控制、人工智能在电力系统中的应用, (E-mail) niuzhewen@tyut.edu.cn

quirement. On the other hand, in order to improve the interpretability of the assessment results, attribution analyses of the model's assessment results are carried out based on the Shapley additivity principle from the perspectives of global features and local samples, separately. **【Conclusions】**The simulation results at IEEE 39 nodes and a provincial power grid show that the proposed method has higher assessment accuracy than that of traditional models, and also has good interpretability.

Keywords: transient stability assessment; machine learning; interpretability; Shapley additive explanation

电力系统的暂态稳定分析是保障电力系统正常运行的基础。“双碳”目标下,随着高比例可再生能源的并网及大规模电网互联^[1],电源结构发生了重大变化,电网安全稳定面临着严峻考验^[2]。传统的暂态稳定评估(transient stability assessment, TSA)主要依托机理模型,但随着电网规模的增大,新型电力系统高维非线性的特性使其计算复杂度和时间成指数增长^[3],且当运行方式或拓扑结构改变时,须重新建模仿真,因此亟需引入能兼具快速性、准确性和普适性的暂态稳定评估方法^[4]。

人工智能技术和相量测量单元(phasor measurement unit, PMU)的广泛应用,为基于数据驱动的电力系统暂态稳定评估研究提供了技术支撑^[5-7]。由于实际电网失稳样本数量远少于稳定样本,样本的不平衡性限制了数据驱动模型在提升综合性能和增强泛化能力方面的潜力^[8]。为了应对这一挑战,现有研究主要从过采样、引入约束、集成学习、拟合少数样本等方面进行优化和改进。文献[9]提出改进合成少数过采样方法,对少数类样本进行线性插值以实现样本扩增,减少不平衡数据集的影响。文献[10]引入梯度范数均值比,量化了数据集不平衡度,提出基于代价敏感法的损失函数改进方法,仿真结果表明,对模型评估具有良好的修正效果。文献[11]通过Easy Ensemble算法进行数据采样,并基于集成学习策略提高分类器精度。文献[12]为辅助模型更好地学习数据分布特征,提出双生成对抗网络,通过交替训练生成器和判别器,批量生成符合真实分布的少数类样本,实验结果表明所提方法能显著提高样本多样性。上述研究主要从数据层面进行少数样本拟合,然而,电力系统运行方式和拓扑结构处于动态变化中,无法从根本上提升模型在多场景下的泛化能力。

另一方面,机器学习模型自身的黑箱性质使其可解释性差,降低了TSA评估的可靠性^[13]。从解释方法看,主要存在三种方法^[14]:事前解释、自解释和事后解释。然而,前两种方法在迁移能力上表现不

佳,并且可能会降低评估结果的精确度,这与电网在线安全评估所需的高精度标准不相符^[15]。因此,目前的研究趋势主要集中在事后解释方法上,采用了代理模型的策略。具体来说,文献[16]用线性模型作为深度神经网络在单个数据点附近的代理,以此来解释输入变量的敏感性和贡献度。文献[17]将决策树作为GRU神经网络的正则项,通过可视化决策树进行决策规则方面的解释。文献[18]基于贪婪优化决策树,并以树的最大深度作为可解释的度量。文献[19]从特征映射层面基于XGBoost模型计算特征重要性分数得出针对特定样本的特征重要性排序。文献[20]运用集成学习思想,针对不同事故分别训练决策树,提取IF-THEN规则指导预防控制措施,但随着特征维度的增长,规则将成倍增长,解释结果因复杂性过高仍然难以理解。

为解决现有TSA研究存在模型泛化能力低、可解释性差的问题,本文提出基于改进XGBoost的暂态稳定评估模型及其可解释性分析方法。一方面,通过改进模型的焦点损失函数,提高模型对临界态样本的关注度,增强了模型应用于实际电网多种不平衡样本工况下适应能力。另一方面,针对评估模型的黑箱性质,基于沙普利加性原理(Shapley additive explanations, SHAP)训练广义加权线性模型,分别从全局特征和局部样本角度得出影响模型评估结果的特征贡献度及模型做出预测的原因,提高了数据驱动在线TSA的准确率和可靠性。

1 改进XGBoost模型

1.1 含正则项的XGBoost算法

XGBoost是由若干基模型并行运算的集成算法^[21],其目标函数 O_{bj} 由模型的损失函数 L 与抑制模型复杂度的正则项组成,如公式(1)所示:

$$O_{bj} = \sum_{i=1}^n L(\hat{y}_i, y_i) + \sum_{i=1}^k \Omega(f_i). \quad (1)$$

式中: i 表示样本, t 表示基模型, $\hat{y}_i$ 表示第 i 个样本的预

测值, y_i 表示第 i 个样本的真实值, L 为损失函数, n 为样本容量, k 为迭代次数, Ω 为各基础模型的复杂度。

通常取损失函数为平方损失或逻辑损失。将基模型取为决策树时, 模型的复杂度由叶子数及叶子节点的权重决定。将损失函数取为平方损失并通过泰勒展开式进行适当简化, 可得到目标函数为:

$$O_{bj} = -\frac{1}{2} \sum_{j=1}^T \frac{G_j^2}{H_j + \lambda} + \gamma T. \quad (2)$$

$$G_j = \sum_{i \in I_j} g_i = \sum_{i \in I_j} \partial_{\hat{y}^{(t-1)}} L(y_i, \hat{y}^{(t-1)}). \quad (3)$$

$$H_j = \sum_{i \in I_j} h_i = \sum_{i \in I_j} \partial_{\hat{y}^{(t-1)}}^2 L(y_i, \hat{y}^{(t-1)}). \quad (4)$$

式中, T 为树模型的个数, λ 和 γ 是为了防止模型过拟合的正则化系数。

求每个节点每个样本的一阶导数 g_i 和二阶导数 h_i , 然后对每个节点所含样本求和得到 G_i 和 H_i , 最后遍历决策树的节点即可得到目标函数。为了评价不同树状结构的优劣, 使用精确贪心算法(exact greedy, EG)。在树的节点进行分裂时, 依据公式(5)来进行判断。

$$G_a = L_L + L_R - L_F. \quad (5)$$

式中, L_L 表示分裂后左侧节点的损失值, L_R 表示分裂后右侧节点的损失值, L_F 表示不进行分裂, 直接通过父节点得到的损失值。根据这一标准, 如果 G_a 的值大于 0, 则应继续对当前节点进行分裂; 如果 G_a 的值不大于 0, 则应停止对该节点的进一步分裂。

1.2 基于不平衡样本的损失函数改进

交叉熵(cross entropy, CE)是机器学习领域最常用的损失函数之一^[22], 在分类任务中常被用于指导模型训练, 其公式为:

$$I_{CE}(Y, \hat{Y}) = -\sum_i [y_i \cdot \ln(\hat{y}_i)]. \quad (6)$$

由公式(6)可知, 当 $\hat{y}_i$ 与 y_i 的距离越远时, $I_{CE}(Y, \hat{Y})$ 的值越大, 对模型设置的惩罚项越大。在电力系统的实际操作中, 暂态失稳的案例数量远低于稳定的案例, 这导致模型在预测时更倾向于识别稳定样本, 增加了对不稳定样本的误判风险。面对这种样本不平衡的问题, 目前主要有两种应对策略^[23]: 一种是在数据层面采取措施, 比如使用超采样、欠采样或数据增强等手段; 另一种是在算法层面进行调整, 比如应用加权损失函数等方法。本文通过引入焦点损失函数对 I_{CE} 进行改进, 公式如下:

$$I_{FL}(Y, \hat{Y}) = \begin{cases} -\alpha(1-\hat{y}_i)^\beta \cdot \log \hat{y}_i, & y_i = 1 \\ -(1-\alpha) \cdot \hat{y}_i^\beta \cdot \log(1-\hat{y}_i), & y_i = 0 \end{cases}. \quad (7)$$

式中, $\alpha \in (0, 1)$ 为平衡因子, 其作用是调整多数类与少数类样本的相对数量; $\beta \in [0, +\infty]$ 为调制因子, 这个调节参数可以降低多数类样本在损失函数中的权重, 以此增加少数类样本与多数类样本在损失函数中的差异, 进而提升模型对少数类样本的重视程度; 其他参数的定义与公式(1)中的保持一致。

2 基于 SHAP 的模型可解释性分析

2.1 全局特征贡献度解释

下文将从全局角度解释哪些输入特征对模型预测具有突出的影响, 以及特征对不同目标变量的贡献度差异。

根据沙普利值的加性理论, 广义加权线性模型通过一种特殊的训练方法来拟合一个分类器。该模型对单个样本的预测结果等于所有样本预测值的平均值再加上每个特征的 SHAP 值之和, 如公式(8)所示:

$$F(x) = G(x) = \beta_0 + \sum_{i=1}^M \beta_i. \quad (8)$$

式中, β_0 代表模型对样本的预测基础值, β_i 反映了模型对任意样本的预测结果的期望值。 β_i 表示样本的第 i 维度特征的 SHAP 值。SHAP 值衡量了在不同特征组合下, 样本 x 中各个特征的平均边际贡献。如公式(9)所示:

$$\begin{cases} \beta_i = \sum_S \frac{|S|!(M-|S|-1)!}{M!} [F_x(S \cup \{x_{(i)}\}) - F_x(S)] \\ S \subseteq \{x_{(1)}, x_{(2)}, \dots, x_{(M)}\} - \{x_{(i)}\} \end{cases}. \quad (9)$$

其中, $\{x_{(1)}, x_{(2)}, \dots, x_{(M)}\}$ 表示所有特征集合, 而 S 是指不包含特定特征 i 的特征子集。SHAP 值的绝对值大小衡量了该特征对模型预测结果的影响力, 即绝对值越大, 该特征对模型预测的贡献就越显著。此外, SHAP 值的符号指示了特征对模型输出的影响方向: 正值表示特征增加模型输出, 负值表示特征减少模型输出。

2.2 特征映射规律分析

为深入分析关键特征对预测目标的影响, 采用累积局部效应图(accumulated local effects plot, ALE)。该图通过局部效应分析, 排除了特征间相互关联的干扰, 可以更清晰地看到每个特征对目标变量的独立影响。此外, ALE 图还能评估具有强相关性的变量对预测目标的联合影响, 相较于部分依赖图, 它提供了一种更为有效且无偏的分析方法。ALE 图的构建过程涉及对预测变化的累积平均的

计算,并将这些变化累加到特定的网格上,如公式(10)所示:

$$f_{j,ALE}(x_s,x_c)=\frac{\partial f(x_s,x_c)}{\partial x_s}.$$
(10)

其中 x_s 为待解释的特征, x_c 代表除了待解释特征之外的所有其他特征集。ALE 图直观地显示了特征值变化对预测结果的影响,为预测结果提供了精确的分析。

2.3 样本评估结果溯源

为解释深度学习模型对给定实例的预测原因及其置信度,采用局部代理模型提供解释^[24]。首先在目标样本的邻域内进行扰动采样,生成新的样本,并基于这些新样本与原始样本之间的距离,赋予它们相应的权重。随后,利用这些加权的新样本集来训练代理模型,目的是获得对原始模型在局部区域内的准确近似。

所构建的目标函数如式(11)所示,其核心是在损失函数和模型复杂度之间实现最优平衡,确保局部代理模型的预测性能与全局评估模型高度一致,同时避免过度拟合。

$$\text{Exp}(x)=\arg \min [L(f, g, \pi_x)+\Omega(g)].$$
(11)

式中, L 表示局部代理模型的均方误差,它反映了局部代理模型 g 与全局评估模型 f 之间的拟合紧密度。邻近度参数 π_x 限定了在样本 x 周围的扰动采样范围,而 $\Omega(g)$ 则量化了模型的复杂性,其数值与所选代理模型及特征个数有关,代理模型结构越复杂,拟合保真度越高,但可理解性越差。因此,本文将广义加性模型作为代理模型,可以使其模型复杂度达到最低,并且可以通过特征被赋予的权重解释模型做出预测的原因。

3 基于 XGBoost 的可解释性暂态稳定评估

3.1 输入特征选择

本文在借鉴前人经验的基础上^[25-27],综合选取了暂态过程中母线(电压)、发电机(功角、角速度、有功功率、无功功率、机械功率)、交流线(有功功率、无功功率)、负荷(电压、电流、有功功率、无功功率)作为原始特征,选取经过简单数学运算后具有物理可解释性的复合特征作为数据集,以保证在全面涵盖暂态信息的前提下减少输入维度。选取的复合特征及其物理意义如表 1 所示。

3.2 标签构建

现有研究通常是将系统暂稳状态划分为稳定

表 1 选取的复合特征及其物理意义

Table 1 Selected composite features and their physical significance

复合特征	物理意义
母线电压最大值	故障后母线电压转移情况
母线电压最小值	故障处母线电压是否越限
负荷有功功率之和	与暂态稳定性呈负相关
发电机输入机械功率平均值	各发电机的平均发电水平
发电机转子动能最大值	故障过程中发电机最大能力聚集
发电机转子初始加速度最大值	受扰最严重的发电机失稳趋势
发电机初始加速度均方根误差	各发电机转子初始加速度离散程度
发电机初始加速功率的均值	发电机在故障时供求失衡平均程度
发电机转子动能平均值	所有发电机总动能增量
具有最大动能发电机的转子角度	具有最大动能发电机在故障切除后的减速能力
具有最大加速度发电机初始相对转子角度	具有最大动能发电机在故障切除后的减速能力
具有最大转子角度发电机的转子动能	转子角度最领先的发电机失稳趋势
发电机初始加速度的均值	各发电机平均受扰程度
发电机间最大功角差	领先机组与滞后机组的最大失谐差值
发电机所受最大有功冲击	受扰最严重发电机的影响程度
系统总能量调整	故障对系统整体的影响
系统总有功冲击	故障对系统功率振荡的影响

和不稳定两种,忽略了暂态过程的中间状态,难以及时发现系统存在的隐患。本文根据不同的运行条件,将电力系统运行状态分为安全态、警戒态、紧急态和崩溃态,图 1 为运行状态以及相应的转换关系图,其中安全态、警戒态合称正常状态,紧急态和崩溃态合称非正常状态。

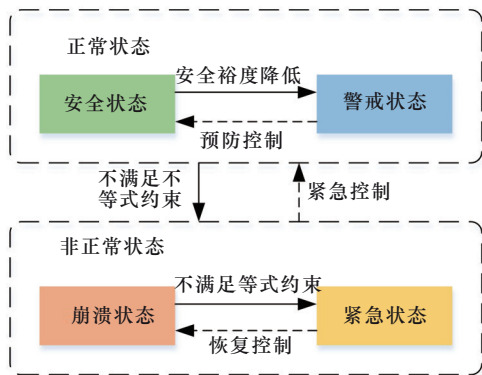


图 1 运行状态划分及其转换关系

Fig.1 System transient state division and transformation relationship

本文首先通过发电机间最大功角差是否大于 360° 判断系统处于正常状态还是非正常状态。在正常状态中,设置全网 $N-1$ 故障,若存在一种故障使系统不满足功角判据,则当前状态为警戒态,否则

为安全态。在非正常状态中,通过判断继电保护装置动作后系统能否稳定运行区分紧急态和崩溃态,即保护装置动作,故障切除后系统满足功角判据,则当前状态为紧急态,否则为崩溃态。

3.3 评估指标

准确率通常被用作评估数据驱动电力系统暂态稳定性模型效能的常规指标,但其未能特别关注对重点类别预测误差的影响。本文在电力系统运行状态分析中,加入混淆矩阵来评估模型对重点类别样本的预测准确性,如表2所示。

表2中列出的变量是通过模型对测试集中的样本进行预测,并将预测结果与实际结果进行比较后,逐一累加得出的。 T 、 F 分别记录了模型预测正确与错误的样本数量。参数下标的第一列索引代表系统的真实状态,第二列索引则表示模型的预测状态。进一步地,从准确率、漏检率和误检率这三个关键维度来评估模型的性能,定义如下。

1)准确率:作为评价模型整体拟合能力的指标,其值反映了模型对所有状态样本预测的精确度,如公式(12)所示:

$$A_{cc} = \frac{T_{11} + T_{22} + T_{33} + T_{44}}{N} \quad (12)$$

其中, N 为测试集样本总数,准确率越高,则对未知

表 2 暂态稳定评估的混淆矩阵
Table 2 Confusion matrix for transient stability assessment

实际状态	预测状态			
	安全态	警戒态	紧急态	崩溃态
安全态	T_{11}	F_{12}	F_{13}	F_{14}
警戒态	F_{21}	T_{22}	F_{23}	F_{24}
紧急态	F_{31}	F_{32}	T_{33}	F_{34}
崩溃态	F_{41}	F_{42}	F_{43}	T_{44}

状态样本预测正确的把握越大。

2)漏检率:用于衡量模型未能正确识别异常状态的错误率。在电网运行过程中,未能检测到故障信息可能导致严重的连锁反应,如公式(13)所示:

$$M_{JR} = \frac{F_{31} + F_{32} + F_{41} + F_{42}}{N} \quad (13)$$

3)误检率:用于衡量模型错误地将正常状态样本识别为异常状态的误差,在实际应用中这种误判可能导致误报警,从而引起保护设备的不必要动作,如公式(14)所示:

$$F_{JR} = \frac{F_{13} + F_{14} + F_{23} + F_{24}}{N} \quad (14)$$

3.4 评估流程

本文所提基于改进XGBoost的暂态稳定评估流程由三部分组成,分别为离线训练、可解释性分析和在线应用,如图2所示。

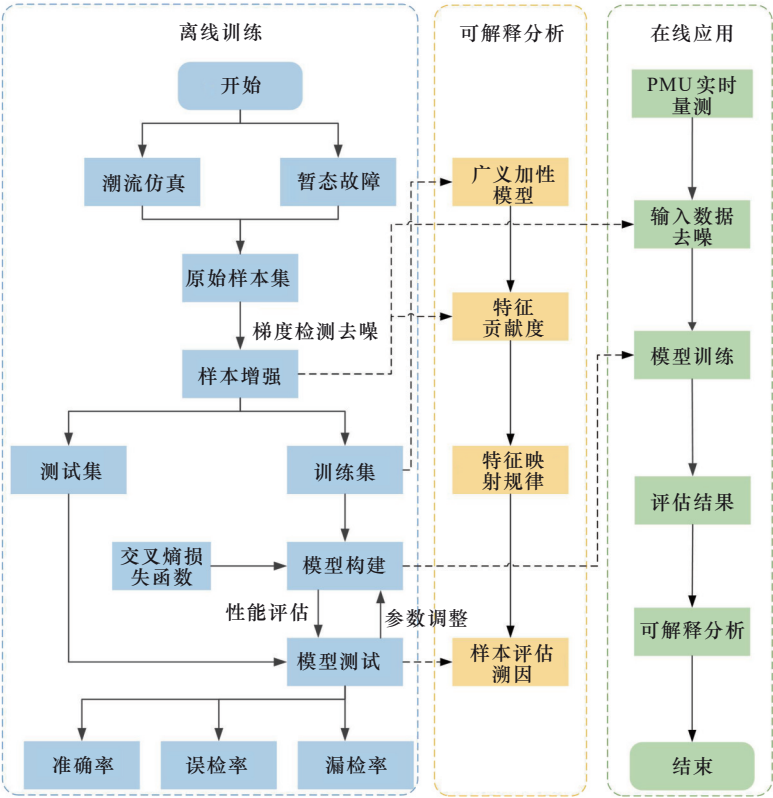


图2 数据驱动的暂态稳定评估及其可解释分析流程

Fig.2 Data-driven transient stability assessment and its interpretable analysis process

在离线训练阶段,首先对构建的电力系统进行潮流设置和暂态故障设置,按照预想特征集进行特征提取,并对每个样本标注状态标签构成原始数据集。经过滚动标准差法检测去噪后,按照预定的比例将数据集分为训练集和测试集。训练集用于 XGBoost 模型的训练,测试集用于评估训练后模型的性能,并进行参数调整以优化模型的拟合效果,使其达到最佳。接着,以电力系统运行状态为导向,基于 SHAP 原理量化各输入特征对模型预测的贡献度,针对重要特征分析其映射规律,并结合实例样本溯因模型的评估依据。在线应用时,对于 PMU 采集的实时数据,经数据增强后调入预训练好的模型中,并对模型评估结果给出合理解释。

模型的复杂度与模型准确性相关联,又与模型的可解释性相对立。本文从全局特征和局部样本的角度对模型预测作了解释。其中全局解释是以模型评估结果为依据,通过计算特征对预测输出结果的贡献来解释模型如何基于其输入特征做出决策。局部解释是一种与模型无关的解释,只研究某数据点和该点周围特征空间中的局部区域,并基于该区域解释该样本的模型决策,进而提升“黑箱”模型在线应用的可靠性。

4 算例分析

4.1 样本构造

所提方法的有效性验证,采用新英格兰 IEEE 39 节点系统作为算例,仿真平台为 PSASP。负荷水平为 70%~120%,发电机出力为 80%~100%,故障类型为三相短路故障,故障点分别设置在各线路的 1%、50% 和 99% 处,以模拟不同位置的故障影响。故障持续时间为 0.1 s、0.15 s 和 0.2 s。按 3.1 节所示选择母线、发电机、线路、负荷 185 维运行特征和 17 维复合特征,共计 202 维特征作为输入变量。采用 3.2 节方法对各样本进行状态判别,分别将安全、警戒、紧急和崩溃状态标记为 1、2、3、4。对故障前、中、后时间断面以 0.01 s 为步长采样,通过仿真共得到 12 750 个样本。从样本中随机抽取 70% 用于构建训练集,剩余的 30% 保留作为测试集。

4.2 数据预处理

PMU 提取的时间序列常包含离群值和异常值,本文采用梯度检测法,滑动时间窗口大小设置为 5,可以在每段中筛选出梯度变化幅度过大的样本点,避免极端数据对样本预测的干扰,如图 3

所示。

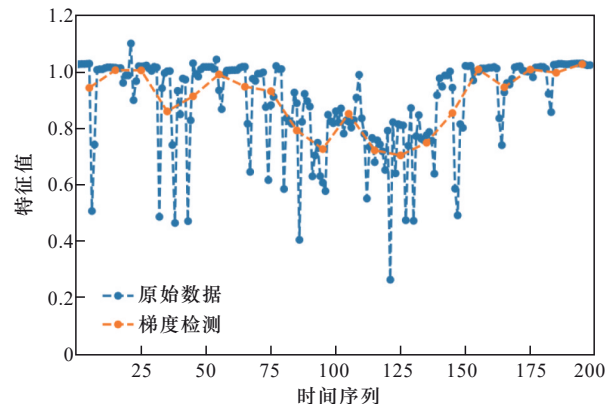


图 3 基于梯度检测的样本去噪
Fig.3 Sample denoising based on gradient detection

4.3 模型性能对比

为了验证本文所提改进 XGBoost 模型的性能,在相同数据集下将其与常用机器学习模型支持向量机 (support vector machine, SVM)、K-近邻 (k-nearest neighbor, KNN)、堆叠自动编码器 (stack autoencoder, SAE)、随机森林 (random forest, RF)、递归神经网络 (recurrent neural networks, RNN) 进行对比。SVM、KNN、RF 基于 Python 的 sklearn 库构建,SAE、RNN 基于 Pytorch 框架实现,Batchsize 为 400,epoch 为 500,初始学习率为 0.005,L2 的惩罚因子为 0.01,dropout 为 0.5,其余超参数按照网格寻优。为避免测试结果的偶然性,随机打乱训练集与测试集,对各模型进行 10 次独立非重复实验,各项指标取测试集的平均值,结果如表 3 所示。

表 3 不同模型的性能对比
Table 3 Performance comparison of different models

评估模型	$A_{cc}/\%$	$M_{JR}/\%$	$F_{JR}/\%$
SVM	97.76	0.62	0.68
KNN	97.48	0.77	0.95
SAE	96.81	1.03	1.15
RF	98.92	0.34	0.42
RNN	99.13	0.32	0.23
改进 XGBoost	99.25	0.20	0.21

由测试结果可知,SAE、KNN 对高维时序特征的映射能力有限,SVM 对于时序数据的分类可能受维数灾难的影响,测试集上的准确率不足 98%。RF 和 RNN 能够捕捉长时间序列的全局依赖关系,性能稳定且准确率较高,但漏检率超过 0.3%,在实际应用中因未能及时识别故障而引起连锁事故的概率较大,仍有待进一步改进。本文所提 XGBoost 模型具有更强的数据挖掘能力,将误检率、漏检率降到 0.25% 以下,与其他的评估模型相比,综合性

能表现更为优异。

4.4 模型鲁棒性测试

实际电力系统发生故障的概率较低,导致稳定样本远多于失稳样本,对模型抗样本失衡能力提出了较高要求^[11]。为了测试不平衡样本对各种深度学习模型性能的影响,以及所提改进焦点损失函数的优势,定义样本不平衡度 C 如下:

$$C = \frac{N_{\max} - N_{\min}}{N_{\text{all}}} \quad (15)$$

式中, $N_{\max}$ 、 $N_{\min}$ 分别为测试集上最多、最少类样本的数量, N_{all} 为测试集上样本总数。

在 $C=0.1、0.15、0.2、0.25$ 的较均衡、稍不均衡、不均衡、极不均衡四种数据集环境下,将传统 XGboost、RNN、KNN 和所提改进 XGBoost 模型分别进行 16 次独立重复实验,准确率结果如图 4 所示。

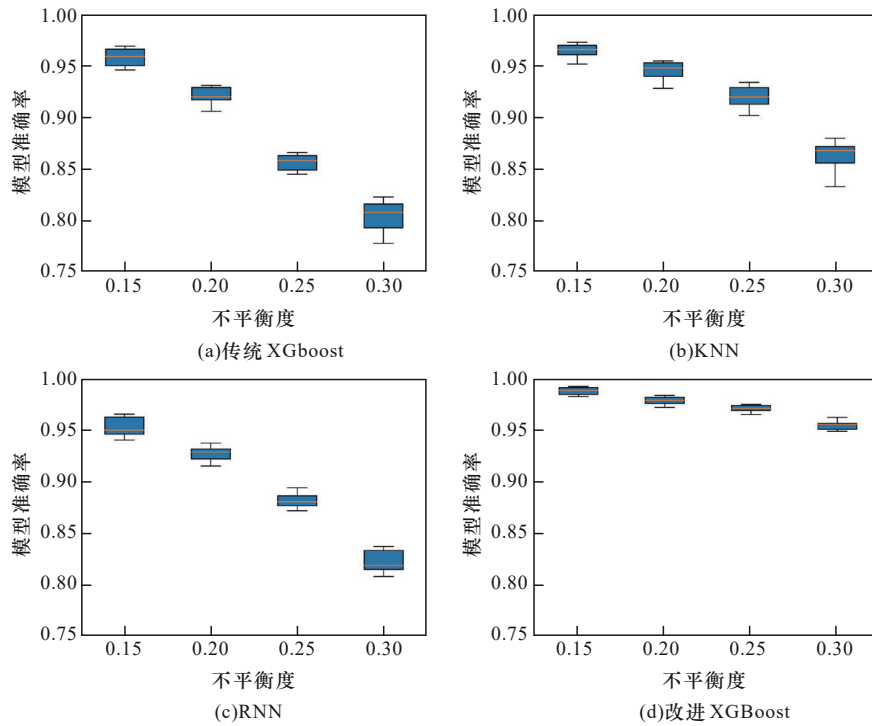


图 4 抗噪性能及去噪方法对比

Fig. 4 Comparison of anti-noise performanc of different denoising methods

图中箱型展示了各模型在测试集上准确率的最小值、下四分位数、中位数、上四分位数及最大值。实验结果表明,KNN 数据分布较为均匀,最容易受样本不均衡性干扰,特别是当 $C=0.3$ 时,准确率的最小值低于 0.8,表明其在极端不均衡环境下模型性能受限;传统 XGbosst 在不均衡测试集上性能较好,但 C 过大时,准确率有明显突降;而本文所提改进 XGBoost 在各种样本场景下的准确率都高于其他机器学习模型,在 $C=0.3$ 的数据集中,准确率最低仍位于 0.95 左右,表明通过交点损失函数改进的 XGbosst 模型具有优良的鲁棒性能力。

4.5 可解释分析

4.5.1 全局特征解释

考虑电力系统运行方式差异,分析各输入特征在不同状态下的 SHAP 值,并将这些变量根据对模型预测总贡献度进行可视化展示,如图 5 所示。其

中横轴代表 SHAP 值,不同颜色则代表不同的状态标签。由上述分析,可得以下结论:

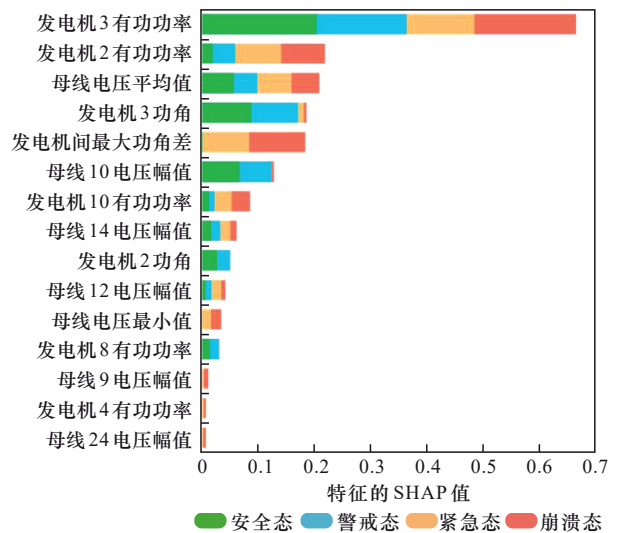


图 5 基于 SHAP 的全局特征重要性

Fig.5 Global feature importance based on SHAP

1)从总体特征贡献度看,发电机 3 的有功功率对模型评估的影响最大,它在预测所有运行状态时都显示出较高的边际影响力,表现出较好的平衡性。

2)具体到某一特征,发电机的最大功角差和无功功率在预测紧急态和崩溃态时贡献最为显著,这些特征可以作为预防控制或故障切除后恢复控制的重要参考。而母线电压最小值则在预测正常状态时更为有效,对非正常状态的预测影响较小,可以作为故障切除后恢复阶段的检验指标。

3)从特征节点来看,关键节点如发电机 3 对数据驱动模型评估有显著影响,其功率变化与整个系

统的安全性密切相关。同时,发电机 2、母线 10 等节点也应予以重点保护和监控。

4)从物理特征看,发电机的有功功率、功角和母线电压是系统暂态过程的主要物理信息载体,这些特征在模型的预测和评估中起到了核心作用。

对部分重要特征做出平行坐标图如图 6 所示。各特征的中轴线表示每个特征的取值,每条折线表示一组样本,颜色代表模型对样本的预测状态标签。可以看出,这些特征能将相同颜色的折线近似集中在一定区域,不同颜色之间有较大差距,表明这些特征对于预测标签类别有较大作用,且有以下规律。

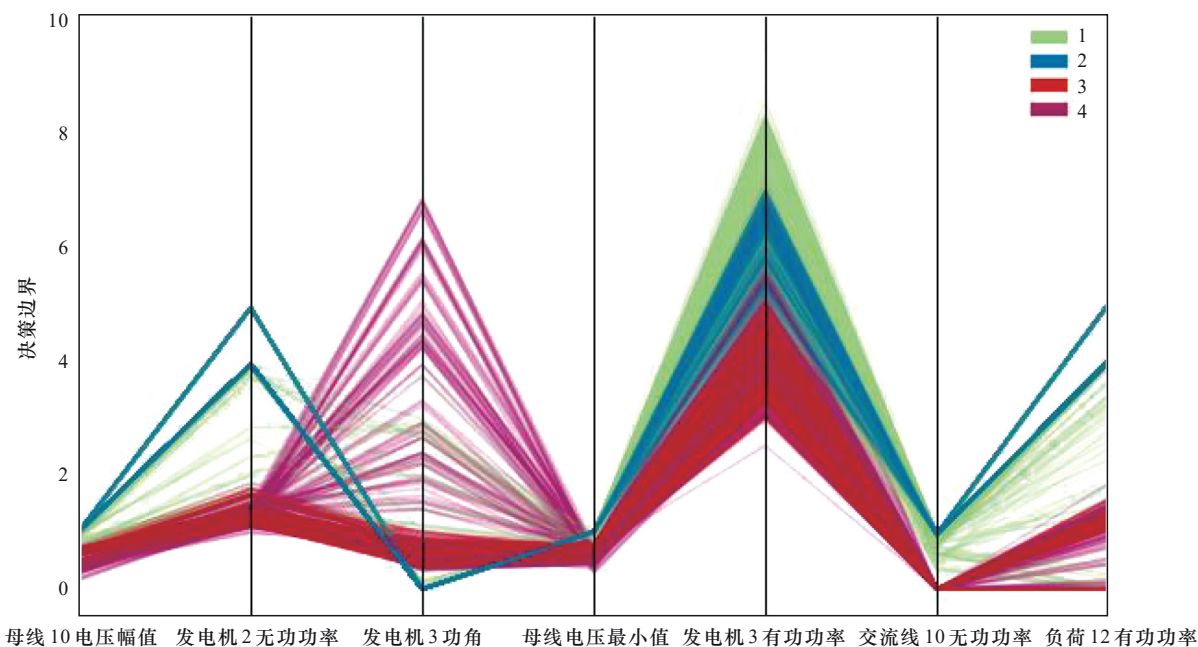


图 6 部分重要特征的平行坐标图
Fig.6 Parallel coordinates of some important features

1)从标签分布看,对绝大多数特征,正常状态与非正常状态有明显界限,而紧急态和崩溃态的决策边界较为接近。

2)从输入特征看,发电机 3 有功功率将电力系统运行状态划分为 4 个集中带状区域,具有较强的状态区分度;发电机功角可以放大紧急状态和崩溃状态的决策边界,提高模型对非正常状态的辨识能力。

为验证解释结果的正确性,按照元件拓扑结构及潮流关系,可将 IEEE 39 节点系统结构分为图 7 所示的 3 个分区,并按稳态潮流方向对功率流向进行标注。其中 I 区作为 II、III 分区的连接纽带,处于拓扑结构的核心区域,对系统的稳定性起最关键作用。特别地,在各分区边界部分的元件,如母线 10、11、14,负荷 4、16 和发电机 2、3,其作为区域间功

率传输的重要节点,应当采取重点保护与监视。

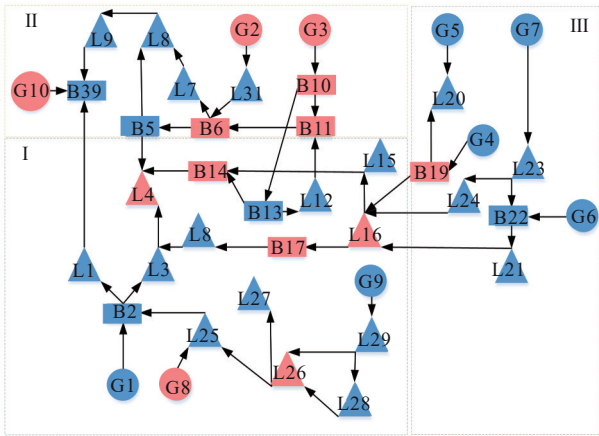


图 7 特征贡献度与拓扑结构
Fig.7 Feature contribution and topology

4.5.2 样本溯因解释

为深入解释模型对样本的评估依据,针对特定故障进行的暂态稳定性分析如下。

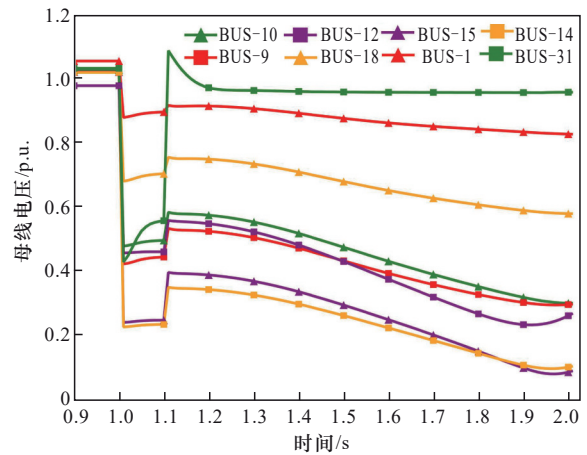
场景一:在PSASP中设定一种潮流情况,所有发电机均以额定水平运行,负荷功率设定为75%,在支路14-15上发生了持续时间为0.15 s的三相短路故障。各发电机间功角差及部分母线电压随时间的动态变化如图8所示。仿真结果表明,母线电压的大幅下降,发电机3的功角发生振荡并失步,系统进入了崩溃状态。

对该故障仿真生成的样本构建了一个局部加权线性模型。利用该模型对样本状态进行评估,并进行归因分析,如图9所示。图中红色和绿色分别用以标识模型预测的崩溃态和紧急态的特征,特征越宽表示该特征所占预测权重越高。

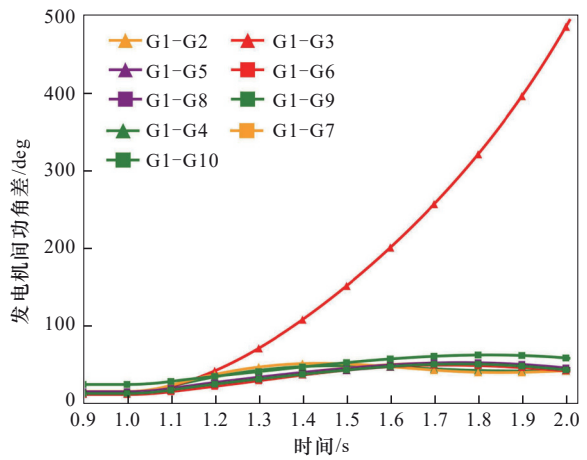
由图9可知,母线14的电压显著降低和发电机3的功率低于设定阈值是样本失稳的关键因素,导致功角振荡增加,最终引发发电机失步。尽管母线电压的平均水平和其他因素在紧急状态的范围,但它们对预测的影响有限。有98%的置信度可以确认样本处于崩溃状态,建议立即执行切机等措施,以避免系统进一步解体。

场景二:此时运行状态为发电机的输出功率维持在90%,负荷的有功功率达到100%。当支路7-8遭遇三相短路故障,且故障持续时间为0.1 s,观察各发电机之间的功角差和关键母线电压随时间变化的趋势,如图10所示。

基于SHAP原理解释模型对该样本的预测原因,如图11所示。图中红色矩形代表正SHAP值的特征,而蓝色部分则代表负SHAP值的特征,矩形面积的大小反映了特征SHAP绝对值的大小,即变量对模型决策的影响程度。模型对于该样本的预



(a)系统母线电压变化曲线



(b)发电机功角变化曲线

图8 故障场景一下部分母线电压及发电机功角变化曲线

Fig.8 Partial bus voltage and generator power angle change curves for fault scenario 1

测结果为所有样本的平均基准值特征加该样本所有特征的SHAP值总和。由图9与11对比可知,不同实例样本中导致模型做出决策原因的变量有一定差异,这是由运行工况和故障类型决定的,但具有全局重要贡献度的特征普遍对各样本的预测均有较为重要的影响,具有共性指导意义。对该样本

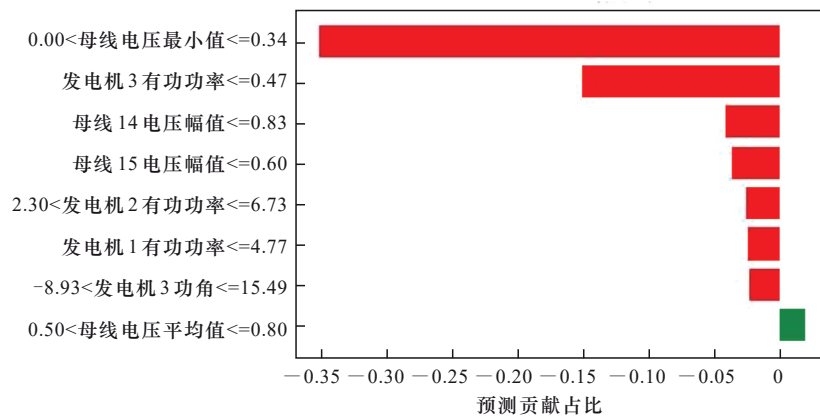


图9 样本预测归因分析

Fig.9 Sample predictive attribution analysis

而言,由于母线电压最小值、发电机 3 功角、有功功率均处于正常范围内,且所占权重最高,因此模型

评估该样本为稳定状态。

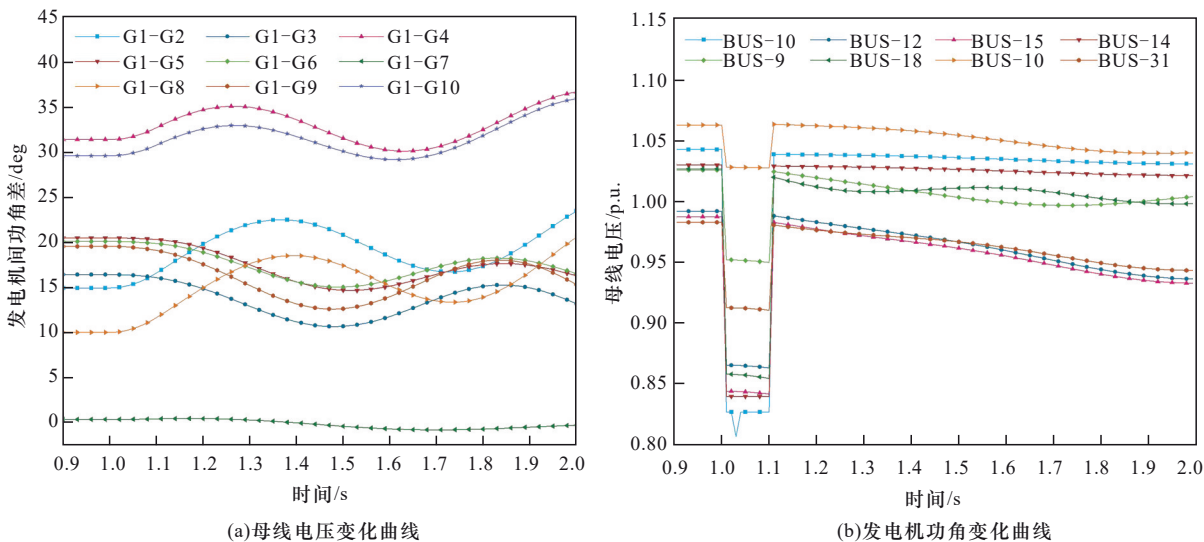


图 10 故障场景二下部分母线电压及发电机功角变化曲线
Fig. 10 Partial bus voltage and generator power angle change curves for fault scenario 2

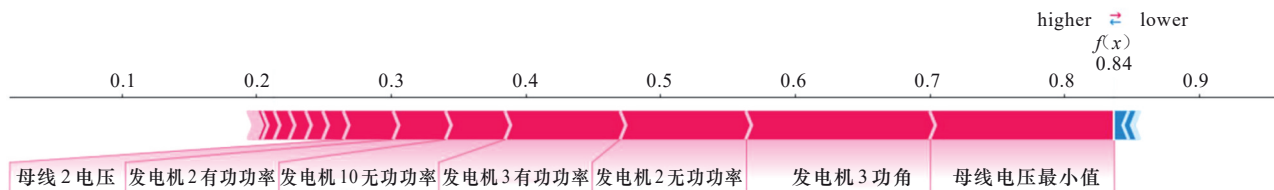


图 11 场景二样本的预测结果归因
Fig.11 Prediction result attribution for scenario 2 samples

4.6 省级电网算例分析

以 2023 年某省电网的运行数据为依据,建立仿真计算的案例。该省调装机以火电为主,占比 60%,风电占比 24%,光伏占比 12%,运行方式主要分夏季和冬季。夏季电网运行的基础方式为 500 kV 系统大开机,220 kV 系统小开机,光伏机组、风电机组 0 出力或光伏机组 80% 出力、风电机组 30% 出力。

考虑到该省电网风电具有明显的区域特性,分析其同时率,对于研究长时间尺度(日、周、月、年)下风电空间相关性,以及电网规划中的电力平衡和安全稳定具有指导意义。本文在基础运行方式下将该电网按区域和电力传输容量分为 A、B、C、D 四个片区,调整区域风电、光伏及火电出力比率,构造不同运行方式,在 PSASP 中调整不同区域发电比例来模拟机组开停,保持负荷水平为该区域原始数据不变。各区域装机容量如表 4 所示。

为充分考虑新能源机组出力情况,选择风电机组出力为 0~100%、光伏机组出力为 0~80%,设置不同区域风电和光伏机组的出力变化,相应改变其他区域的火电机组出力,以实现电力潮流的运行配

表 4 某省级电网分区装机容量表
Table 4 Installed capacity of a provincial power grid region (MW)

分区	风电	光伏	火电
A	36.46	27.73	54.9
B	102.34	13.5	189.78
C	59.53	33.41	148.9
D	48.97	25.95	152.7

置。考虑电力系统正常运行方式、特殊运行方式、负荷波动,以及各分区火电和新能源出力比例变化,共计生成 500 种潮流运行方式,经潮流计算,有效潮流运行方式 468 种。

考虑重要线路 N-1、N-2 作为故障事件,仿真输出信息包括 500 kV 母线电压幅值,火电机组发电机功角、机端电压、电流、有功、无功,负荷有功、无功,共计 1 152 维特征,在 PSASP 中进行仿真运算。采用过采样技术,共生成 20 000 个样本,其中安全态样本 14 000 个、警戒态样本 2 750 个、紧急态样本 2 750 个和崩溃态样本 500 个。经过特征降维与样本去噪,筛选得到 228 维输入变量。在相同测试集上分别对 SVM、RF、RNN 和本文所提 XGBoost 的

性能对比,混淆矩阵如图 12 所示。

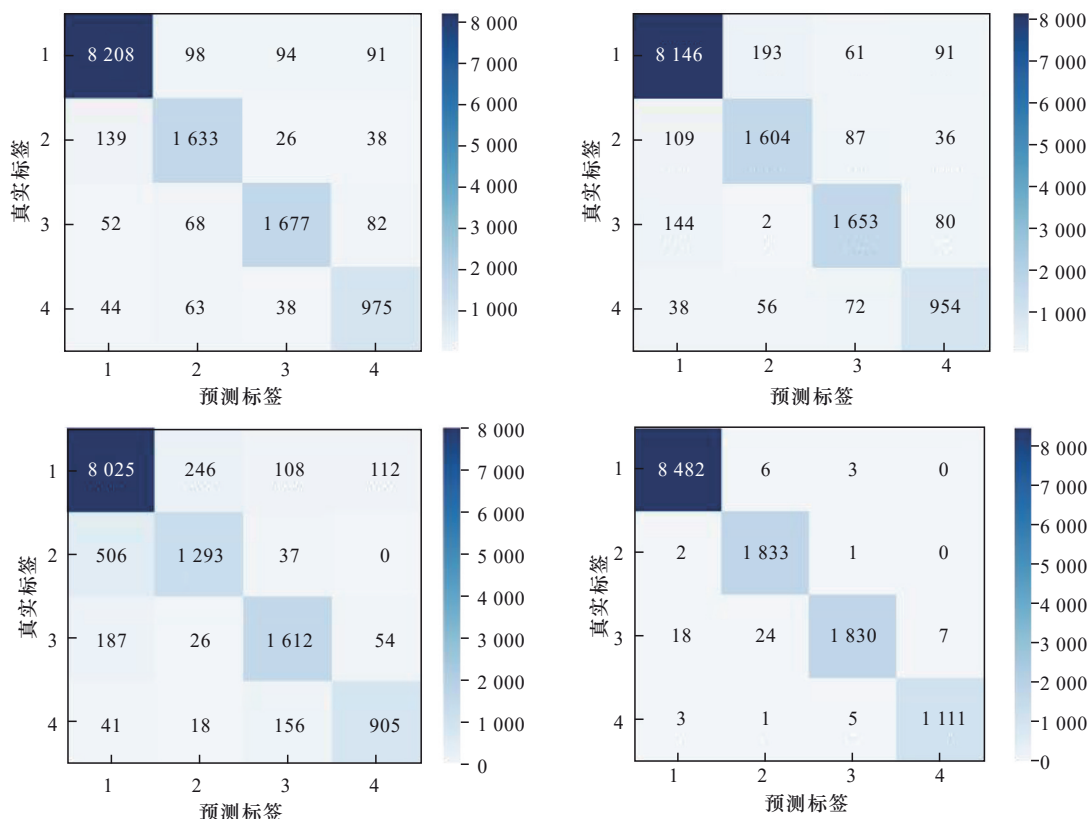


图 12 各模型 TSA 混淆矩阵

Fig. 12 TSA confusion matrix for each model

由图 12 可知, SVM 和 RNN 的总体准确率较低, 特别是误评率较高, 容易导致保护装置误动作; RF 的准确率较高, 但仍存在较大程度的漏检风险; 所提改进 XGBoost 模型的准确率可达 99.4%, 且漏检率、误检率可以控制在 0.25% 以下, 能够满足实际应用需求。

做出全数据集上特征重要性排序如图 13 所示, 其中横轴代表特征的 SHAP 值, 纵轴代表特征的实际值。由图 13 可知, 母线电压最小值具有最高的全局贡献度, 且其实际值越低, 对模型预测值提升得越多, 系统越容易暂态失稳; 而发电机间最大功角差主要在低值时对模型评估起重要贡献, 降低模型预测值, 进而提升系统暂态稳定性。

对给定样本做预测归因如图 14 所示, 其中 $f(x)$ 为预测的 SHAP 平均值, 蓝色正向预测表示特征对样本预测为正常状态且有重要贡献, 红色反向预测表示有非正常状态的倾向。由于发电机 A1 的有功功率、功角和发电机 A3、B2 功角均偏离正常状态范围, 且对预测的贡献最大, 因此有 99.3% 的置信度认为该样本处于崩溃状态, 且该样本主要是由于发电机功角甩开而失稳, 同时预测 SHAP 值的期

望为 0, 表明其失稳程度较大。

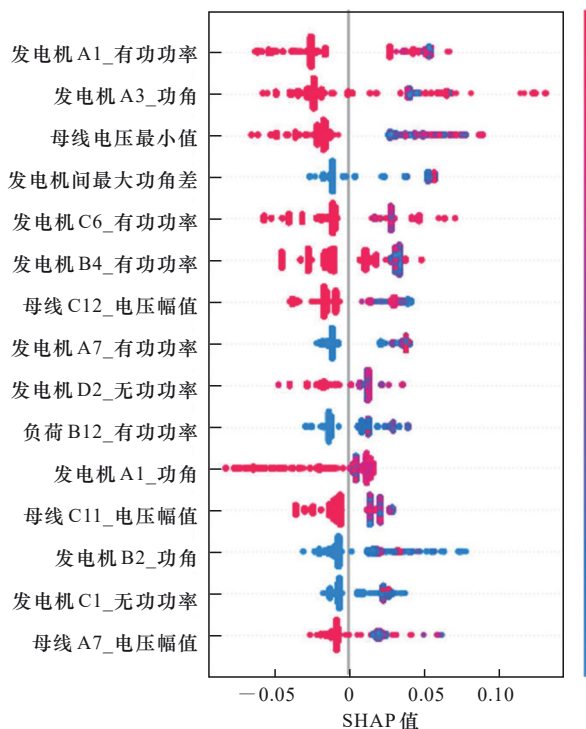


图 13 特征全局贡献度

Fig. 13 Global contributions of features

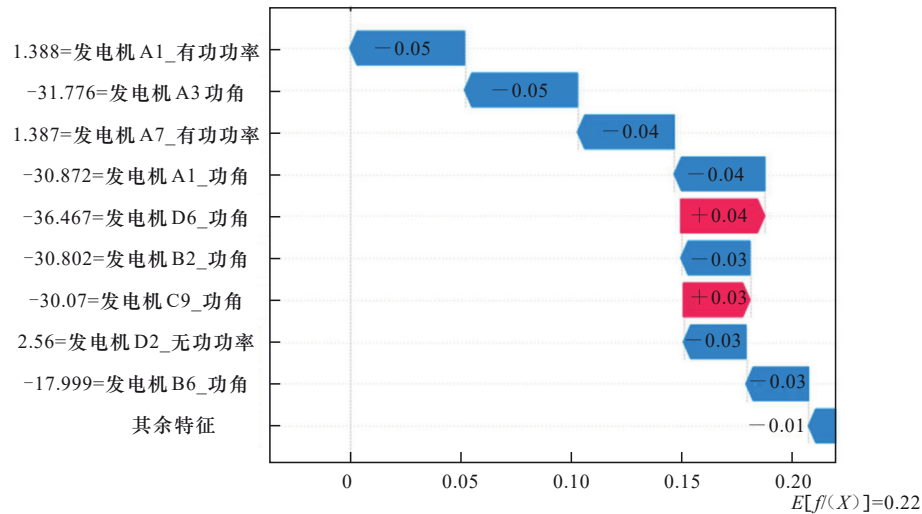


图14 实例样本的预测归因分析

Fig.14 Predictive attribution analysis on instance sample

5 总结

本文针对现有数据驱动 TSA 研究存在评估精度有待提高、可靠性差的问题,提出了基于改进 XGBoost 的暂态稳定评估模型及其可解释性分析方法,主要贡献如下:

1)考虑不同样本对模型评估的贡献度差异,以焦点损失函数作为 XGBoost 的损失函数,提高模型对少数类样本的关注度,在 IEEE 39 节点系统和某

省级电网仿真结果表明,所提模型评估准确率在多种电力系统运行场景下均可维持在 95% 以上,显著提升了数据驱动模型的泛化能力;

2)考虑机器学习的黑箱性质,提出基于 SHAP 的特征一样本解释方法,量化了考虑运行状态差异性的全局特征贡献度,分析了重要的特征映射规律,并给出了实例样本的溯因依据,提高了调度人员对模型决策依据的可理解度。

参考文献:

- [1] 武宇翔,韩肖清,牛哲文,等.融合多注意力深度神经网络的可解释光伏功率区间预测[J].电网技术,2024,48(7):2928-2939.
WU Y X, HAN X Q, NIU Z W, et al. Interpretable photovoltaic power interval prediction using multi-attention deep neural networks[J]. Power System Technology, 2024, 48(7): 2928-2939.
- [2] ZHANG Z G, KANG C Q. Challenges and prospects of building a new power system under the goal of carbon neutrality [J]. Journal of CSEE, 2022, 42(8): 2806-2819.
- [3] SUN H D, XU T, GUO Q, et al. Analysis of the "8.9" blackout in the UK and its enlightenment to China's power grid [J]. Journal of CSEE, 2019, 39(21): 6183-6192.
- [4] HUANG D, CHEN S Y, GRACE W W. Research on power system transient stability control technology based on wide-area measurement information [J]. Power Generation Technology, 2018, 39(5): 389-396.
- [5] ZHU L P, HILL D J. Data/Model jointly driven high-quality case generation for power system dynamic stability assessment [J]. IEEE Transactions on Industrial Informatics, 2022, 18(8): 5055-5066.
- [6] 张亮,安军,周毅博.基于时间卷积和图注意力网络的电力系统暂态稳定评估[J].电力系统自动化,2023,47(7):114-122.
ZHANG L, AN J, ZHOU Y B. Transient stability assessment of power system based on temporal convolution and graph attention network[J]. Automation of Electric Power Systems, 2023, 47(7): 114-122.
- [7] ZHU L P, WEN W J, LI J Y, et al. Integrated data-driven power system transient stability monitoring and enhancement[J]. IEEE Transactions on Power Systems, 2024, 39(1): 1797-1809.
- [8] TAN B, YANG J, TANG Y, et al. A deep imbalanced learning framework for transient stability assessment of power system [J]. IEEE Access, 2019, 99(7): 81759-81769.
- [9] 刘颂凯,党喜,崔梓琪,等.针对样本类不平衡的深度残差网络电力系统暂态稳定评估方法[J].智慧电力,2024,52(1):116-123.

- LIU S K, DANG X, CUI Z Q, et al. Transient stability evaluation method for power systems with deep residual network considering class imbalance of samples[J]. Smart Power, 2024, 52(1): 116-123.
- [10] 胡力涛, 王怀远, 党然, 等. 基于梯度范数的暂态稳定评估模型的不平衡修正方法[J]. 电力自动化设备, 2024, 44(4): 158-163, 177.
- HU L T, WANG H Y, DANG R, et al. Imbalance correction method of transient stability assessment model based on gradient norm[J]. Electric Power Automation Equipment, 2024, 44(4): 158-163, 177.
- [11] 李嘉敏, 杨红英, 闫莉萍, 等. 样本不平衡情况下的电力系统暂态稳定集成评估方法[J]. 电力系统自动化, 2021, 45(10): 34-41.
- LI J M, YANG H Y, YAN L P, et al. Integrated assessment method for transient stability of power system under sample imbalance[J]. Automation of Electric Power Systems, 2021, 45(10): 34-41.
- [12] 杨东升, 吉明佳, 周博文, 等. 基于双生成器生成对抗网络的电力系统暂态稳定评估方法[J]. 电网技术, 2021, 45(8): 2934-2945.
- YANG D S, JI M J, ZHOU Bowen, et al. Transient stability assessment of power system based on DGL-GAN[J]. Power System Technology, 2021, 45(8): 2934-2944.
- [13] 武宇翔, 韩肖清, 牛哲文, 等. 基于变权重随机森林的暂态稳定评估方法及其可解释性分析[J]. 电力系统自动化, 2023, 47(14): 93-104.
- WU Y X, HAN X Q, NIU Z W, et al. Transient stability assessment method based on variable weight random forest and its interpretability analysis[J]. Automation of Electric Power Systems, 2023, 47(14): 93-104.
- [14] 纪守领, 李进锋, 杜天宇, 等. 机器学习模型可解释性方法、应用与安全研究综述[J]. 计算机研究与发展, 2019, 56(10): 2071-2096.
- Ji S L, Li J F, Du T Y, et al. Survey on techniques, applications and security of machine learning interpretability[J]. Journal of Computer Research and Development, 2019, 56(10): 2071-2096.
- [15] DENG H T. Interpreting tree ensembles with intrees[J]. International Journal of Data Science and Analytics, 2019, 7(4): 277-28.
- [16] 韩天森, 陈金富, 李银红, 等. 电力系统稳定评估机器学习可解释代理模型研究[J]. 中国电机工程学报, 2020, 40(13): 4122-4131.
- HAN T S, CHEN J F, LI Y H, et al. Study on interpretable surrogate model for power system stability evaluation machine learning[J]. Proceedings of the CSEE, 2020, 40(13): 4122-4130.
- [17] REN C, XU Y, ZHANG R. An interpretable deep learning method for power system transient stability assessment via Tree Regularization[J]. IEEE Transactions on Power Systems, 2022, 37(5): 3359-3369.
- [18] CREMER J L, KONSTANTELOS I, STRBAC G. From optimization-based machine learning to interpretable security rules for operation[J]. IEEE Transactions on Power Systems, 2019, 34(5): 3826-3836.
- [19] ZHANG S, ZHANG D, QIAO J, et al. Preventive control for power system transient security based on XGBoost and DCOPF with consideration of model interpretability[J]. CSEE Journal of Power and Energy Systems, 2021, 7(2): 279-294.
- [20] LIU C. A systematic approach for dynamic security assessment and the corresponding preventive control scheme based on decision trees[J]. IEEE Transactions on Power Systems, 2014, 29(2): 717-730.
- [21] 岳有军, 吴明沅, 王红君, 等. 基于CNN-GRU-ISSA-XGBoost的短期光伏功率预测[J]. 南京信息工程大学学报(自然科学版), 2024, 16(2): 231-238.
- YUE Y J, WU M Y, WANG H J, et al. Short term photovoltaic power prediction based on CNN-GRU-ISSA-XGBoost[J]. Journal of Nanjing University of Information Science & Technology, 2024, 16(2): 231-238.
- [22] 孙辉, 孙宝硕, 高正男, 等. 基于量测数据时序谱分布特性的电力系统暂态功角稳定在线判别[J]. 电网技术, 2023, 47(3): 1107-1116.
- SUN H, SUN B S, GAO Z N, et al. Online identification for power system transient angle stability based on time-series spectral distribution characteristics of measured data[J]. Power System Technology, 2023, 47(3): 1107-1116.
- [23] 邵美阳, 吴俊勇, 李宝琴, 等. 基于两阶段集成深度置信网络的电力系统暂态稳定评估[J]. 电网技术, 2020, 44(5): 1776-1787.
- SHAO M Y, WU J Y, LI B Q, et al. Transient stability assessment of power system based on two-stage ensemble deep belief network[J]. Power System Technology, 2020, 44(5): 1776-1787.
- [24] 周挺, 杨军, 詹祥澎, 等. 一种数据驱动的暂态电压稳定评估方法及其可解释性研究[J]. 电网技术, 2021, 45(11): 4416-4425.
- ZHOU T, YANG J, ZHAN X P, et al. Data-driven method and interpretability analysis for transient voltage stability assessment[J]. Power System Technology, 2021, 45(11): 4416-4425.

- [25] 甄永赞,阮程. 基于强化学习的混合元启发式暂态电压稳定特征选择方法及可解释性研究[J]. 电网技术,2024,48(4):1519-1532.
- ZHEN Y Z, RUAN C. Reinforcement learning-based hybrid element heuristic transient voltage stability feature selection and Its interpretability[J]. Power System Technology, 2024, 48(4): 1519-1531.
- [26] 李鹏,董鑫剑,孟庆伟,等. 基于 Fisher Score 特征选择的电力系统暂态稳定评估方法[J]. 电力自动化设备,2023,43(7):117-123.
- LI P, DONG X J, MENG Q W. Transient stability assessment method for power system based on Fisher Score feature selection[J]. Electric Power Automation Equipment, 2023, 43(7): 117-123.
- [27] 姜新雄,徐箭,廖思阳,等. 基于增强进化计算的电力系统暂态稳定关键特征智能选择方法[J]. 中国电机工程学报,2023,43(14):5358-5372.
- JIANG X X, XU J, LIAO S Y, et al. An intelligent selection method for power system transient security key features based on enhanced evolutionary computation[J]. Proceedings of the CSEE, 2023, 43(14): 5358-5371.

(编辑:赵树庆)