

# 光伏阵列故障诊断的可解释性智能集成方法

陈泽<sup>1</sup>, 刘文泽<sup>1</sup>, 王康德<sup>1</sup>, 余涛<sup>1,2</sup>, 黄展鸿<sup>1,2</sup>

(1. 华南理工大学 电力学院, 广东 广州 510641; 2. 广东省电网智能量测与先进计量企业重点实验室, 广东 广州 510641)

**摘要:**针对现有光伏阵列故障检测和诊断智能方法存在的泛化性不强、可解释性差的问题,提出了一种可解释性智能集成方法。对采集的光伏阵列输出时序电压、电流波形进行特征挖掘,并将多个已成熟应用于光伏故障诊断的智能算法作为不同基学习与元学习器,构建结合不同智能算法优势且更具泛化性的Stacking集成学习模型;以沙普利可加性特征解释方法为总框架,并结合局部近似可解释性方法,对模型训练过程与结果进行解释分析,通过获取各特征的贡献、分析该集成模型的决策机制,并了解其如何进行诊断,提高其可靠度和可信度。算例实验结果表明,所提可解释性智能集成方法在不同规模数据集的测试中均实现了高精度的故障诊断,模型的可解释性结果表明由该智能集成模型建立的故障特征和诊断结果的映射遵循物理见解,增强了智能方法的可信度和透明性。

**关键词:**光伏阵列;故障诊断;Stacking集成;可解释性智能方法;沙普利可加性特征解释方法

**中图分类号:**TM615

**文献标志码:**A

**DOI:**10.16081/j.epae.202401006

## 0 引言

在我国“双碳”目标以及建设新型电力系统的战略指引下,光伏发电因具有可用性高、安装速度快、维护成本低等特点获得长足发展,已成为最常用的可再生和清洁能源之一<sup>[1]</sup>。而随着全球光伏装机容量迅速提升,其可靠性问题不容忽视,且光伏阵列安装在室外环境中,工作环境复杂,容易发生光伏组件老化、接地、短路、开路、局部遮阴等故障问题,降低光伏发电效率,缩短光伏阵列使用寿命<sup>[2-3]</sup>。并且目前研究大多采用具有“黑箱”性质的人工智能方法,其合理性与决策容易引发质疑。因此,研究快速、准确、可解释的光伏阵列故障诊断方法对光伏发电的良性发展具有重要意义。

国内外学者针对光伏故障检测与分类进行了大量的研究,大致可以将其分为如下2类:视觉和热处理方法(visual and thermal methods,VTM)以及基于电气的方法<sup>[4]</sup>。VTM通过红外热成像、电致光成像、视觉检测等技术研究和分析故障模组的物理特性,简单、快速地实现故障检测,但其难以有效实现故障分类,且维护成本高,存在局限性<sup>[5]</sup>。现有研究大多认为基于电气的方法使用更多信息提供了更全面的分析,大部分的常见故障可被识别与诊断。依据故障特征的选择,基于电气的方法可以分为 $I-V$ 特性分析方法、功率损失分析方法、电压与电流测量方法、电

压/电流谱分析方法等<sup>[5]</sup>。而在这些检测方法中,依托电气信息所获得的特征进行故障诊断,往往依赖以机器学习为代表的人工智能技术,其无须投入更多成本,也不依赖数学模型,因此被广泛应用于光伏阵列故障检测与诊断中。文献[6]提出了一种贝叶斯优化的 $K$ 近邻( $K$ -nearest neighbor, KNN)算法,用于检测线对线故障和开路故障。有研究以光伏系统的 $I-V$ 特性作为输入,使用遗传算法优化的支持向量机(support vector machine, SVM)检测光伏系统中的线对线故障<sup>[7]</sup>。文献[8]使用随机森林(random forest, RF)检测光伏阵列的开路、短路、热点、老化和部分阴影故障,有较高的准确率。基于梯度提升决策树(gradient Boost decision tree, GBDT)算法框架的改进方法,如极度梯度提升(extreme gradient Boosting, XGBoost)算法<sup>[9]</sup>、CatBoost(categorical Boosting)算法<sup>[10]</sup>, 通过从 $I-V$ 曲线中提取包括开路电压、短路电流、最大功率点工作电压和电流等特征,在光伏阵列故障检测与诊断中收获好的效果。除了以上多种分类器的应用,神经网络方法也已经被广泛应用于光伏系统的故障检测和分类,其通过自动提取非线性特征,提升了检测效率,其中包括反向传播(back propagation, BP)神经网络、卷积神经网络(convolutional neural network, CNN)、人工神经网络(artificial neural network, ANN)等<sup>[11-13]</sup>。但是,现有方法多以光伏组件 $I-V$ 输出特性为原始输入特征,其需要专业的测试设备,且获取存在一定的延时,导致无法实现实时故障检测,故障将在阵列中存在一段时间,进而造成一定的安全风险<sup>[4]</sup>。同时,单一人工智能模型依赖大量有标签的数据进行训练,在不同数据规模、不同类型的故障检测任务中也难以保证其性能稳定。此外,最关键的一点是,以人工智能技术

收稿日期:2023-04-23;修回日期:2023-12-20

在线出版日期:2024-01-12

基金项目:国家自然科学基金委员会-国家电网公司智能电网联合基金资助项目(U2066212)

Project supported by the National Natural Science Foundation of China-State Grid Joint Fund for Smart Grid(U2066212)

为代表的驱动方法本质上是个黑匣子,研究可解释的智能方法可以在识别影响故障的关键因素、分析不同故障特征对各类故障的贡献程度、增加“黑箱”模型的透明度、澄清模型的决策过程、提高结果的可信度等方面发挥重要作用<sup>[14]</sup>。

针对上述问题,本文提出了一种以在线测量的光伏阵列的电压、电流为输入特征,运用可解释性的Stacking集成智能模型对光伏阵列故障进行诊断的方法。首先,通过考虑不同的故障情况下光伏阵列输出电流、电压时序波形与变化规律,基于统计分析挖掘故障特征;其次,基于Stacking集成框架,构建由KNN、SVM、RF、XGBoost组成的异质集成模型,对故障进行诊断与检测;然后,提出可解释性分析方法,帮助理解该模型如何进行诊断,解释各故障特征的贡献程度,提高模型诊断结果的可靠性与可信度;最后,基于仿真数据进行算例分析,对线对线、开路、遮光等故障进行诊断,并通过对比各模型与所提模型,验证所提方法的有效性与准确性。

## 1 光伏故障特性分析与特征选择

### 1.1 光伏发电系统

典型的并网光伏发电系统一般由光伏阵列、汇流箱、逆变器、变压器、相关的保护控制设施等组成<sup>[15]</sup>。光伏阵列由多块光伏组件串并联组成,每块光伏组件由串联的光伏电池片和并联在电池片两侧的旁路二极管组成,电路图见附录A图A1。

考虑以双二极管模型描述光伏电池和光伏组件的伏安特性,建立电路等效模型,见附录A图A2。光伏组件的基本电流方程如式(1)所示,由于光生电流受辐射度影响,由式(1)可知,光伏组件的输出不仅受到内部参数的影响,还与环境温度和辐射度有关。

$$I_L = I_{ph} - I_{d1} \left[ \exp \left( \frac{V_L + R_s I_L}{A_1} \right) - 1 \right] - I_{d2} \left[ \exp \left( \frac{V_L + R_s I_L}{A_2} \right) - 1 \right] - \frac{V_L + R_s I_L}{R_{sh}} \quad (1)$$

式中: $I_L$ 、 $V_L$ 分别为光伏组件输出电流和电压; $I_{ph}$ 为光生电流; $I_{d1}$ 、 $I_{d2}$ 分别为二极管 $D_1$ 、 $D_2$ 的反向饱和电流; $R_s$ 、 $R_{sh}$ 分别为串、并联电阻; $A_i$  ( $i=1, 2$ )为二极管修正品质因子,  $A_1 = n_1 k^* T/q$ ,  $A_2 = n_2 k^* T/q$ ,  $T$ 为温度,  $n_1$ 、 $n_2$ 分别为扩散、复合二极管的理想因子,  $q$ 为电子电荷,  $k^*$ 为波兹曼常数。

### 1.2 光伏阵列的故障特性分析

光伏阵列最常见的故障有线对地、线对线、开路、遮光、光伏组件退化等故障<sup>[10]</sup>,其中光伏组件退化是个长期变化过程,难以通过短暂的时序电流、电压变化识别,因此本文不做讨论。同时,光伏发电系

统通常配备过电流保护装置和接地保护装置来保护设备,很容易识别接地故障和故障程度较为严重的故障情况并提供保护,因此本文主要针对以下轻型故障情况进行具体分析。

1)线对线故障指阵列中某些位置的意外短路,使一个或多个模块短路并导致功率损失。光伏阵列中的线对线故障可能是由于载流导体的意外短路、连接器之间的绝缘不良、直流箱串、运行过程中造成的损坏以及绝缘失效而发生的<sup>[16]</sup>。这种情况可能发生在同一组串内或2个相邻组串之间,如图A1中的 $F_{11}$ 与 $F_{12}$ 所示。

2)开路故障指系统中某个点断开连接,导致光伏阵列某条组串电流中断。光伏阵列的开路故障主要发生在破裂的光伏电池和模块上或模块间互连中,影响到光伏组件的输出电流,如图A1中的 $F_2$ 所示。

3)遮光故障指光伏阵列的单个组件或多个组件被遮挡,导致阵列输出功率受损、阵列失配的情况。其中:硬遮光故障由需要人工清理污垢或物体之类的遮挡物造成的,并且会造成遮蔽的组件发热,进而产生热斑效应,如图A1中的 $F_{31}$ 所示;暂时性遮光由云层、树荫之类暂时性的遮挡物遮挡造成的,其阴影面积和阴影程度会随时间变化,导致阵列出现具有时变性的局部阴影,如图A1中的 $F_{32}$ 所示。

为了实现快速的故障诊断,最优的选择是通过分析光伏阵列的输出电压与电流在不同故障状态下的时序波形变化,深度分析不同故障场景之间的运行特征差异,进而挖掘电参数的变化共性与异性。

各运行状态下的时序电压、电流如图1所示,图中电压、电流为标准化值,设置在指定区间内。正常运行情况下,阵列电压与电流平缓波动,几乎不发生突变,如图1(a)所示;当发生线对线故障时,故障串的电电压小于正常串的电电压,这导致阵列电流突然下降,输出功率也突然下降,之后,为了满足最大功率点跟踪的最大功率点优化,阵列的电流逐渐上升,电压逐渐下降,如图1(b)所示;当发生开路故障时,由于1条支路断路,阵列电流下降,电压短暂下降后基本保持不变,如图1(c)所示;当光伏阵列出现硬遮光故障时,光伏阵列电流序列呈现连续下降趋势,电压出现下降,直到从原始最大功率工作点过渡到新的最大功率工作点,电压电流开始稳定,如图1(d)所示;当出现暂时性的部分遮光故障时,会持续影响阵列的输出电流,造成电流和电压波动,出现多个最大功率工作点,如图1(e)所示。

### 1.3 故障特征挖掘与选择

故障特征的挖掘和选择对于准确识别光伏阵列故障具有关键作用。如1.2节所述,不同状态下的时序电压、电流的波形变化规律与趋势各有不同,因此需要从中选择能表达故障信息、表征波形特征的参



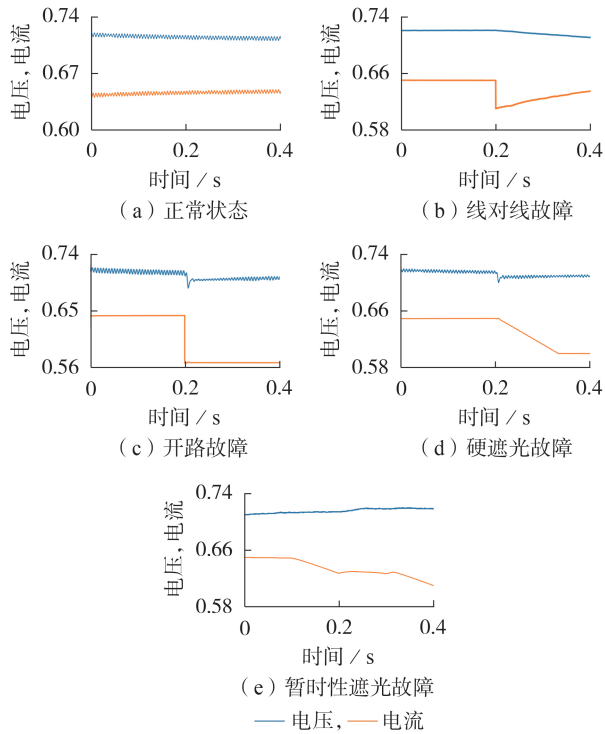


图1 各运行状态下的时序电压、电流  
Fig.1 Time series voltage and current under each operating state

数作为故障特征。参考相关研究<sup>[15,17]</sup>与时序电压、电流波形的变化特征,基于统计分析选取直流侧光伏阵列输出电流 $I_{PV}$ 与电压 $V_{DC}$ 的均值、方差、最大最小差、峰值系数作为故障特征,各统计特征的表达式分别为:

$$X_{\text{mean}} = \frac{1}{N} \sum_{k=1}^N x_k \quad (2)$$

$$X_{\text{Var}} = \frac{1}{N-1} \sum_{k=1}^N (x_k - X_{\text{mean}})^2 \quad (3)$$

$$X_{\text{Range}} = \max_{k=1,2,\dots,N} \{x_k\} - \min_{k=1,2,\dots,N} \{x_k\} \quad (4)$$

$$X_{\text{CF}} = X_{\text{Range}} / \sqrt{\frac{1}{N} \sum_{k=1}^N x_k^2} \quad (5)$$

式中: $X_{\text{mean}}$ 、 $X_{\text{Var}}$ 、 $X_{\text{Range}}$ 、 $X_{\text{CF}}$ 分别为均值、方差、最大最小差、峰值系数; $N$ 为采集时间序列样本中的数据总数; $x_k$ 为 $k$ 时刻电气量的测量值。

如1.1节所述,由于光伏阵列的输出电压、电流还与外界环境因素有关,考虑对不同温度和辐射度下所发生的故障进行诊断,将光伏阵列处所测得的辐射度与温度也列为故障特征。综上,所选取的特征及描述如附录B表B1所示。

## 2 基于可解释性 Stacking 集成方法的光伏阵列故障诊断

### 2.1 Stacking 集成学习方法

Stacking 是一种分层模型集成框架,通常设计为

2层:第1层包括称为基学习器的若干算法,而第2层称为 Stacking 算法的元学习器<sup>[18]</sup>。这种集成的优点是首先通过使用原始数据集训练基学习器得到预测结果,然后将每个学习器的输出组合成一个新的数据集用来训练元学习器,最后由第2层元学习器使用该数据集得到最终的预测结果,从而获得更准确的预测结果<sup>[19]</sup>,并得到比单模型更优且鲁棒性更强的预测结果的方法,其框架如图2所示。

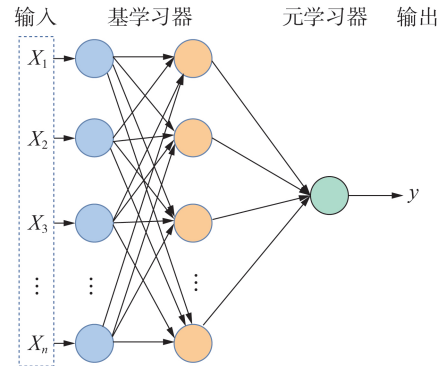


图2 Stacking 集成学习框架

Fig.2 Stacking ensemble learning framework

### 2.2 基学习器

基模型学习能力越强,差异性越大,则 Stacking 集成模型的效果越好<sup>[20]</sup>。本文将 KNN、SVM、RF、XGBoost 这4种在光伏故障诊断与检测中应用成熟且原理存在差异性的算法模型作为基学习器。下面对各学习器进行简要介绍。

1) KNN。在故障分类中, KNN 方法是一种通过检查新对象与特征空间中最近的邻近训练样本的距离来对新对象进行分类的方法。

2) SVM。SVM 的核心思想在于构建一个最优超平面将不同数据分类,并使分类间隔最大,对于非线性可分数据,可通过核函数将样本数据映射到更高维度的空间,将最优分类超平面由低维不可建变为高维可建,进而对数据进行分类<sup>[21]</sup>,本文研究中所用核函数为“linear”,最优超平面分类函数 $f(x)$ 的表达式为:

$$f(x) = \text{sgn} \left( \sum_{i=1}^n \alpha_i y_i K(x_i \cdot x_j) + b \right) \quad (6)$$

式中: $\text{sgn}(\cdot)$ 为符号函数,若自变量为正则取值为1,若为0则取值为0,若为负则取值为-1; $\alpha_i$ 为拉格朗日乘子; $n$ 为样本数量; $y_i$ 为样本 $i$ 的预测值; $K(x_i \cdot x_j)$ 为核函数, $x_i$ 、 $x_j$ 分别为样本集第 $i$ 、 $j$ 个样本; $b$ 为分类阈值。

3) RF。RF 使用重新采样的方法,从一个原始的训练集中反复随机地选择 $n$ 个样本,并进行替换,以生成一个新的训练集 $\{(x_i, y_i)\} (i \in [1, n])$ ,该训练集通过考虑这些采样点附近的加权函数,对未标记类

别的新样本  $x_{\text{new}}$  进行分类预测,新数据的分类由投票的分类树数量形成的分数确定<sup>[22]</sup>。具有自权重  $\omega$  的一组  $k$  棵树的森林平均预测如式(7)所示。

$$\hat{y} = \sum_{i=1}^n \left( \frac{1}{k} \sum_{j=1}^k \omega(x_j, x_{\text{new}}) y_i \right) \quad (7)$$

式中:  $\hat{y}$  为新样本  $x_{\text{new}}$  的预测值。

4) XGBoost。XGBoost 属于基于树集成的学习模型,其基本思想是通过整合多棵基本树计算最终的分类结果。同时在 GBDT 的基础上,引入损失函数对预测结果的二阶导数,并在目标函数中加入树模型的复杂度作为正则化项,防止过拟合,提高模型的泛化性能<sup>[23]</sup>。

### 2.3 元学习与故障诊断流程

轻梯度提升机(light gradient Boosting machine, LightGBM)算法是一种基于决策树算法的分布式梯度提升算法,它具有运行速度快、内存消耗低、精度高特点<sup>[24]</sup>,因此将其作为元学习器。

根据传感器在线测量的光伏阵列输出的时序电压与电流数据,截取固定时间窗内的故障信号序列数据通过所述统计分析实现故障特征的选取与计算,得到如表 B1 所示的特征  $F_{\text{feature1}}$ 、 $F_{\text{feature2}}$  等,并构建输入特征向量  $[X_1, X_2, \dots, X_n]$ ,将其输入训练好的 Stacking 模型中,输出得到不同类型的故障标签,并对此进行可解释性分析构建基于可解释 Stacking 集成的光伏阵列故障诊断方法,诊断流程如图 3 所示。

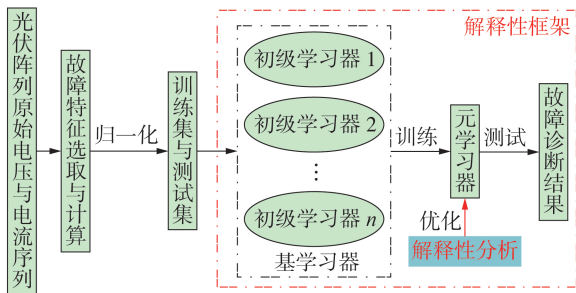


图3 故障诊断流程图

Fig.3 Flowchart of fault diagnosis

### 2.4 可解释性方法研究

虽然大量研究中使用不同研究方法根据不同特征得到了光伏故障的诊断结果,但在实验中不能获取除模型输出结果以外相应足够的可解释信息,不可解释性使得难以对导致故障的关键特征进行分析,故而难以了解模型如何进行决策,制约了其实际应用。为此,本文以一种附加特征归因法——沙普利可加性特征解释(Shapley additive explanation, SHAP)方法作为模型的解释框架,综合应用局部可解释性等方法对所提的 Stacking 集成模型进行解释。

SHAP 分析方法以 SHAP 值作为特征重要性的

统一度量,SHAP 值将对每个特征进行调节时预期模型预测的变化归因于该特征<sup>[25]</sup>。SHAP 方法依赖于加性特征属性的思想,使用模型  $g(x)$  拟合经过训练的模型的输出。模型  $g(x)$  的输出表示为预测基准值和每个特征的贡献(即 SHAP 值)之和<sup>[26]</sup>,如式(8)所示。

$$g(x) = \varphi_0 + \sum_{i=1}^{n_{\text{feature}}} \varphi_i \quad (8)$$

式中:  $\varphi_0$  为模型对样本预测的基准值;  $\varphi_i$  为特征  $x_i$  的特征归因 SHAP 值;  $n_{\text{feature}}$  为特征数量。

SHAP 值  $\varphi_i$  反映了特征  $x_i$  与不同特征子集联合作用下的模型输出值之间的差异以及  $x_i$  单独作用下的输出值。因此,SHAP 值旨在计算  $x_i$  边际贡献的平均值。基于合作博弈论中的价值分配问题,通过计算具有特征和不具有每个特征的模型预测,获得特征的边缘贡献,取所有特征贡献的平均值,以此来计算该特征 SHAP 值  $\varphi_i$ ,即在所有特征值组合上加权与求和,计算式如式(9)所示。

$$\left\{ \begin{aligned} \varphi_i &= \sum_S \frac{|S|!(n-|S|-1)!}{n!} (f(S \cup \{x_i\}) - f(S)) \\ S &\subseteq \{x_1, x_2, \dots, x_{i-1}, x_{i+1}, x_n\} \end{aligned} \right. \quad (9)$$

式中:  $S$  为不包含  $x_i$  的特征子集;  $f(S)$  为仅存在  $S$  时模型的输出值;  $|S|$  表示当前正在计算的子集  $S$  中包含的特征数量;  $f(S \cup \{x_i\})$  为同时存在  $S$  和  $x_i$  时模型的输出值。

SHAP 值的正负反映的是该故障特征对诊断结果的正向或反向影响关系,SHAP 值的大小反映的是该特征对结果的影响程度。因此,基于特征值的分布及其对应的 SHAP 值,可以挖掘特征值对诊断结果的影响关系和影响程度。同时,当使用所提出的 Stacking 集成模型诊断某故障样本时,每个特征的 SHAP 值可以根据式(8)和式(9)进行计算,获取不同特征所起的作用,并找到准确识别不同故障时所倚靠的关键特征,同时可与所做的故障特性分析进行对比,以此既可以提供故障诊断模型所做决策背后的推理,也可以对所做分析进行验证,提高模型结果的可靠性。

基于对模型的全局分析,本文还应用局部近似可解释性方法对故障诊断模型进行局部分析。局部近似可解释性方法的总体目标是用可解释的表示形式来拟合一个可解释的模型,这个模型只需要在局部上实现对一个分类器很好的近似,从而可以可信地解释任何分类器预测。在故障诊断模型中解释的实现方式为:  $x$  作为一个特征样本,故障模型  $f(x)$  代表其属于某故障类的概率,同时使用  $\Pi_x(y)$  作为样本  $y$  到样本  $x$  之间的一个接近性度量,以便定义样本  $x$  周围其他样本的位置。然后采用保真函数

$L(f, g, \Pi_x)$  定义局部区域内的解释模型  $g$  近似  $f$  的不可信程度<sup>[27]</sup>, 其中  $f$  为原模型,  $\Pi_x$  为样本之间的接近性度量。局部近似可解释性产生的局部解释如下:

$$\xi(x) = \operatorname{argmin}_{g \in G} L(f, g, \Pi_x) \varphi_0 + \Omega(g) \quad (10)$$

式中:  $\xi(x)$  为衡量 2 个模型之间差异的目标函数;  $\operatorname{argmin}$  表示函数取值最小时的自变量取值;  $G$  为可解释模型集合;  $\Omega(g)$  为复杂性度量函数。

### 3 实验验证与分析

#### 3.1 仿真和数据获取

通过 MATLAB / Simulink 构建光伏发电系统, 模型结构与图 A1 所示结构相同, 光伏阵列包括 4×7 (4 块并联、7 块串联) 块光伏组件。并采用仿真来构建光伏阵列故障集。其中, 故障场景中光伏阵列温度处于 20~35 °C, 光照强度在 200~1 000 W / m<sup>2</sup> 内。故障场景模拟总时间为 0.6 s, 设置线对线、开路、硬遮光故障发生在 0.2 s, 且故障发生后故障一直存在至仿真结束, 对于暂时性遮光故障, 则在仿真时间内设置辐射度的波动变化和阴影的组件数量来进行模拟, 对这段时间内的数据进行采集, 并计算 1.3 节中所述的故障特征。最终生成数据集包括多个特征与一个类标签。不同运行状态对应不同标签, 具体描述见附录 B 表 B2。

#### 3.2 故障诊断结果与分析

实验过程中分别以 30%、50%、70%、90% 比例的数据集作为训练集来验证 Stacking 集成学习模型的泛化性能, 通过与各初级学习器的诊断结果进行对比验证其准确性。各模型采用贝叶斯优化方法对参数进行寻优, 以准确率  $A_{cc}$  和 F1 评分  $F_1$  作为诊断结果优良的评价指标, 其表达式分别为:

$$A_{cc} = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{TN} + N_{FP} + N_{FN}} \times 100\% \quad (11)$$

$$\begin{cases} F_1 = 2 \frac{P_{re} R_c}{P_{re} + R_c} \\ P_{re} = \frac{N_{TP}}{N_{TP} + N_{FP}}, R_c = \frac{N_{TP}}{N_{TP} + N_{FN}} \end{cases} \quad (12)$$

式中:  $N_{TP}$  为预测本类正确样本数;  $N_{TN}$  为预测其他类准确样本数;  $N_{FP}$  为错将其他类预测为本类样本数;  $N_{FN}$  为错将本类预测为其他类样本数;  $P_{re}$  为查准率, 表示被分类器判定正例中实际为正样本的比重;  $R_c$  为召回率, 表示预测准确的样本中实际为正的比重。

对比各算法在不同比例数据集实验中的性能表现, 不同算法诊断的评价指标  $A_{cc}$  与  $F_1$  如附录 B 表 B3 所示。所提算法诊断各运行状态的  $F_1$  如图 4 所示。

由表 B3 可知, 将 30% 数据集作为训练集的实验

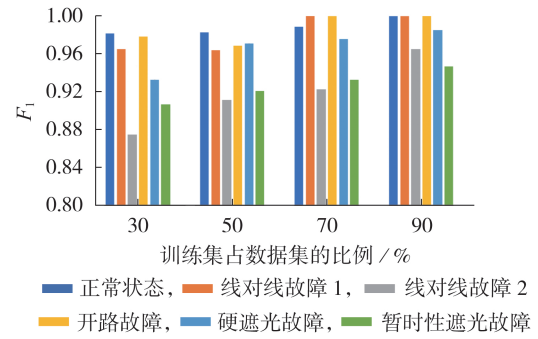


图4 不同运行状态下的诊断结果

Fig.4 Diagnostic results of different operating states

中, Stacking 集成算法的准确率仅略微低于 XGBoost 算法, 低比例数据集中 KNN、SVM 算法表现一般, 随着训练集占比的增加, 各算法准确率随之提升, 但是 XGBoost 出现过拟合现象, 准确率反而降低。而 Stacking 集成算法的准确率在不同比例数据集实验中与各算法相比都较高, 结果综合表明所提 Stacking 集成算法对不同规模的数据集诊断准确率较高, 具有很强的泛化性能。由图 4 可知, 在不同类型运行状态的诊断中, 所提 Stacking 集成算法随着数据集的增大, 对于各类故障都有较高的准确率, 因此证明其能有效诊断不同类型的故障。综上所述, 所提模型对于不同类型的故障和不同大小的数据样本都有较好的性能, 证明了其优越性与泛化性。

#### 3.3 诊断模型与结果的解释

首先从全局角度对所提光伏阵列故障诊断模型进行解释, 分析实际模型决策过程中故障特征对各类别故障的贡献程度。本文使用 SHAP 方法框架研究特征的重要性, 并了解 Stacking 集成故障诊断模型如何依赖它们来判断光伏阵列故障类型, 形成故障特征与故障类别之间的直接映射。

如 2.4 节所述, SHAP 值显示了给定特征对判别结果的影响程度。由于以 90% 数据集作为训练集的实验中模型具有很高的准确性, 以此作为实验样本解释所提模型在该实验中的测试集受特征影响的实际情况。图 5 从不同角度展示了各特征(特征含义见附录 B 表 B1)对模型输出的影响。

图 5(a) 为全局特征重要性图, 展示的是各种特征对故障诊断结果产生的全局性影响, 全局重要性被视为该特征在所有给定样本中的平均绝对值, 其清楚地表明特征中直流侧电压方差  $V_{var}$  的贡献最大, 其次是电流峰值系数  $I_{CF}$ 、电压峰值系数  $V_{CF}$  和输出电流方差  $I_{var}$ 。图 5(b) 为特征值及其对应的 SHAP 值, 图中每一个散点为一个样本, 散点颜色代表特征值大小(越红表示特征值越大、越蓝表示特征值越小)。以特征  $V_{var}$  和  $I_{CF}$  为例, 可以观察到对诊断结果有正面贡献时通常  $V_{var}$  特征值较小,  $I_{CF}$  特征值较大。



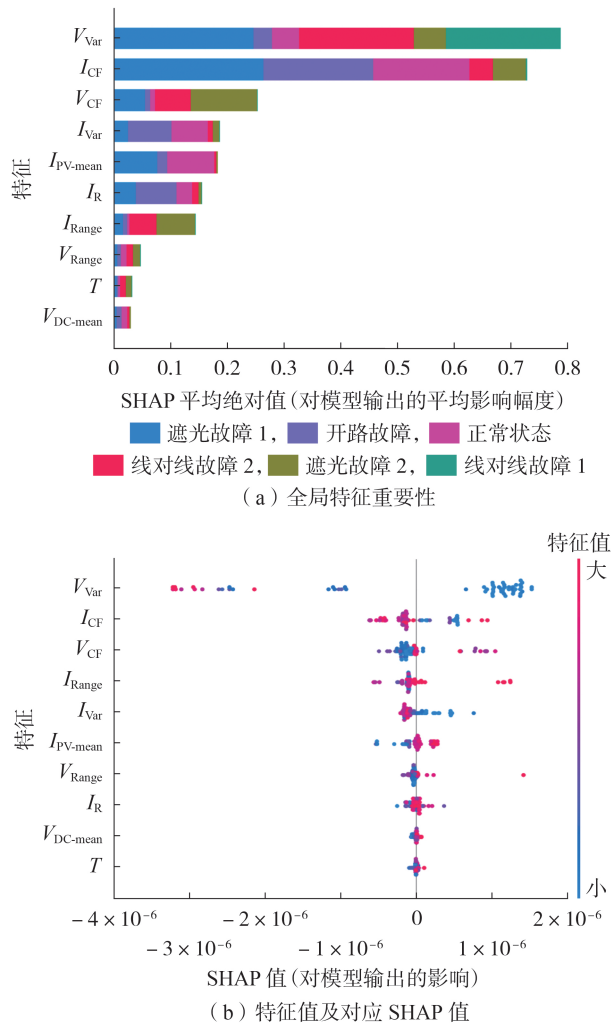


图 5 特征对模型输出的定量影响

Fig.5 Quantitative influence of features on model output

从实际角度来进行分析,当电压方差较小时,意味着时序电压波形较为平稳,电压方差较大时,波形动荡,说明  $I_{CF}$  越大代表电流波形振幅越大,电流波形存在突变,反映了不同运行状态下电流电压变化特点。

可以从上述分析中获得故障特征在模型决策过程中的全局贡献程度,但在模型决策过程中,了解各特征是如何影响模型的决策过程,并结合相应的电气量变化趋势和物理原理进行分析,利于提升模型结果的可信度和对结果的理解。结合 SHAP 与局部近似可解释性方法对模型的局部预测进行分析,并对具体的故障样本进行分析,当模型诊断某故障样本为运行状态  $F_{11}$  时,决策图如附录 C 图 C1 所示,局部近似可解释性方法对故障输出的局部解释图见附录 C 图 C2。结合图 C1 与图 C2 可知模型在得到故障诊断结果时各特征是如何影响其决策的,对模型做出正确预测(即判别类别为 1)中贡献最大的特征为  $V_{Var}$ ,其特征值较大为 1.99,意味着其电压值与其均

值之间的差异较大,出现了较大波动,与发生线对线故障时电压时序特征相吻合。

上文分析了不同故障特征对各类别故障的贡献程度和模型正确决策的过程,为了更加全面清晰地了解模型的决策,对实验中模型错判样本也进行了决策解释,以其中准确率最低的故障类型,将  $F_{32}$  误判为  $F_{31}$  的样本为例,其决策过程如附录 C 图 C3 所示。由图可知,模型将该样本判断为故障  $F_{31}$  和故障  $F_{32}$  的概率分别为 0.58 和 0.33,模型认为其相似性较高,但更偏向于为硬物遮挡造成的遮光故障。在特征贡献方面,特征  $I_{Range}$ 、 $T$ 、 $I_{CF}$  等为模型作出判断提供了正面贡献,从图 C3(b)模型决策解释中也可以发现是这些特征导致模型作出差异化判断,而特征  $V_{Var}$ 、 $I_{CF}$  则是使模型区分这两类和其他类的重要因素。从实际情况分析,该样本环境因素中辐射度  $I_R$  ( $-0.40$ ) 较低,温度  $T$  较高,光伏阵列输出电流和电压显著降低,会减小波形的变化幅度,也使得温度对模型的预测结果影响程度更大; $I_{CF}$  较大符合遮光故障的电流波形变化特性,而  $V_{Var}$ 、 $V_{Range}$  较小,表明电压波动性较低,更符合故障  $F_{31}$  的电压波形特征,因而模型易做出上述的误判结果。

以上对所提 Stacking 集成光伏故障诊断模型进行了可解释性分析,获得了在该智能模型中各故障特征对不同类别故障的贡献程度,澄清了模型的决策过程,并联系实际情况下的电气量的变化趋势对因果关系进行了分析,分析结果对增加“黑箱”模型的透明度以及提高结果的可信度发挥了重要作用,也为改善输入特征的选择和根据模型学习的内容对其进行优化提供了有力依据。

## 4 结论

本文提出了一种基于 Stacking 集成学习方法的光伏阵列故障诊断方法,并对此提出了面向该模型训练与决策的可解释性体系。通过实验验证了该模型的泛化性、准确性和有效性,并从全局与局部角度出发,建立了对完整系统的理解和对特定实例模型如何决策的认知,可得出以下结论。

1) 根据光伏阵列的输出电压、电流的时序波形,基于统计分析提取了有效故障特征,保证了时效性并极大地简化了数据的采集,节约了成本。

2) 所提 Stacking 集成模型相比于其他单模型在不同比例数据集的实验中有很好的泛化性能,也能准确诊断不同类型的故障。

3) 基于 SHAP 框架与局部近似可解释性方法对 Stacking 集成模型进行解释,通过解释方法阐明了故障特征对各类别故障的贡献程度,澄清了模型的决策过程及各特征对模型输出的定量影响,表明了基于数据驱动建立的特征与故障标签映射关系符合物

理规律。从全局和局部角度帮助管理人员理解完整的模型系统,提高了所提集成智能模型结果的可信度。

本文所提方法对于其他会导致光伏阵列时序电压、电流变化的故障,通过获取相应的数据与特征,即可实现诊断。可解释性模型为结果提供了合理的解释,后续将利用模型解释为故障特征选取、模型优化所提供的方向进行适当调整。

附录见本刊网络版(<http://www.epae.cn>)。

#### 参考文献:

- [1] 高伟,黄俊铭. 基于SSELM的光伏组件故障智能诊断方法[J]. 太阳能学报,2021,42(12):465-470.  
GAO Wei,HUANG Junming. Intelligent fault diagnosis method of photovoltaic module via SSELM[J]. Acta Energetica Sinica,2021,42(12):465-470.
- [2] PILLAI D S, RAJASEKAR N. A comprehensive review on protection challenges and fault diagnosis in PV systems[J]. Renewable and Sustainable Energy Reviews,2018,91:18-40.
- [3] 孙培胜,陈堂贤,程陈,等. 基于SOA-SVM模型的光伏阵列故障诊断研究[J/OL]. 电源学报. [2023-08-22]. <http://kns.cnki.net/kcms/detail/12.1420.TM.20220823.1109.002.html>.
- [4] HONG Y Y,PULA R A. Methods of photovoltaic fault detection and classification:a review[J]. Energy Reports,2022,8:5898-5929.
- [5] 乔苏朋,杨艳,陈世群,等. 光伏阵列故障检测方法综述[J]. 电气技术,2021,22(7):1-6.  
QIAO Supeng,YANG Yan,CHEN Shiqun,et al. Review on photovoltaic array diagnosis methods[J]. Electrical Engineering,2021,22(7):1-6.
- [6] SABBAGHPUR A M,HEJAZI A M. On-line faults detection and classification in PV array using Bayesian and K-nearest neighbor classifier[J]. Energy Engineering & Management,2018,8(2):14-25.
- [7] ESKANDARI A,MILIMONFARED J,AGHAEI M,et al. Autonomous monitoring of line-to-line faults in photovoltaic systems by feature selection and parameter optimization of support vector machine using genetic algorithms[J]. Applied Sciences,2020,10(16):5527.
- [8] LIU Y,YAN B F,QIAN D,et al. Research on fault diagnosis of photovoltaic array based on random forest algorithm [C]//2021 IEEE International Conference on Power Electronics,Computer Applications(ICPECA). Shenyang,China:IEEE,2021:194-198.
- [9] ADHYA D,CHATTERJEE S,CHAKRABORTY A K. Performance assessment of selective machine learning techniques for improved PV array fault diagnosis[J]. Sustainable Energy, Grids and Networks,2022,29:100582.
- [10] 顾崇寅,徐潇源,王梦圆,等. 基于CatBoost算法的光伏阵列故障诊断方法[J]. 电力系统自动化,2023,47(2):105-114.  
GU Chongyin,XU Xiaoyuan,WANG Mengyuan,et al. Cat-Boost algorithm based fault diagnosis method for photovoltaic arrays[J]. Automation of Electric Power Systems,2023,47(2):105-114.
- [11] 贾嵘,李云桥,张惠智,等. 基于改进BP神经网络的光伏阵列多传感器故障检测定位方法[J]. 太阳能学报,2018,39(1):110-116.  
JIA Rong,LI Yunqiao,ZHANG Huizhi,et al. Multi-sensor fault detection and positioning method of photovoltaic array based on improved BP neural network[J]. Acta Energetica Sinica,2018,39(1):110-116.
- [12] ZAKI S A,ZHU H L,AL FAKIH M,et al. Deep-learning-based method for faults classification of PV system[J]. IET Renewable Power Generation,2021,15(1):193-205.
- [13] SOFFIAH K,MANOHARAN P S,DEEPAMANGAI P. Fault detection in grid connected PV system using artificial neural network[C]//2021 7th International Conference on Electrical Energy Systems(ICEES). Chennai,India:IEEE,2021:420-424.
- [14] 蒲天骄,乔骥,赵紫璇,等. 面向电力系统智能分析的机器学习可解释性方法研究(一):基本概念与框架[J]. 中国电机工程学报,2023,43(18):7010-7030.  
PU Tianjiao,QIAO Ji,ZHAO Zixuan,et al. Research on interpretable methods of machine learning applied in intelligent analysis of power system(part I):basic concept and framework[J]. Proceedings of the CSEE,2023,43(18):7010-7030.
- [15] CHEN S Q,YANG G J,GAO W,et al. Photovoltaic fault diagnosis via semisupervised ladder network with string voltage and current measures[J]. IEEE Journal of Photovoltaics,2021,11(1):219-231.
- [16] MELLIT A,TINA G M,KALOGIROU S A. Fault detection and diagnosis methods for photovoltaic systems:a review[J]. Renewable and Sustainable Energy Reviews,2018,91:1-17.
- [17] GHONEIM S S M,RASHED A E,ELKALASHY N I. Fault detection algorithms for achieving service continuity in photovoltaic farms[J]. Intelligent Automation & Soft Computing,2021,29(3):467-479.
- [18] 袁铁江,郭泽林,方形. 基于运行数据时空特征和Stacking集成学习的质子交换膜燃料电池故障诊断[J]. 中国电机工程学报,2023,43(14):5461-5471.  
YUAN Tiejiang,GUO Zelin,FANG Tong. Proton exchange membrane fuel cell fault diagnosis based on operation data temporal and spatial characteristics and Stacking ensemble learning[J]. Proceedings of the CSEE,2023,43(14):5461-5471.
- [19] 崔树银,汪昕杰. 基于最大信息系数和多目标Stacking集成学习的综合能源系统多元负荷预测[J]. 电力自动化设备,2022,42(5):32-39,81.  
CUI Shuyin,WANG Xinjie. Multivariate load forecasting in integrated energy system based on maximal information coefficient and multi-objective Stacking ensemble learning[J]. Electric Power Automation Equipment,2022,42(5):32-39,81.
- [20] 陆万荣,许江淳,李玉惠. 面向Stacking集成的改进分类算法及其应用[J]. 计算机应用与软件,2022,39(2):281-286.  
LU Wanrong,XU Jiangchun,LI Yuhui. Improved classification algorithm for Stacking integration and its application[J]. Computer Applications and Software,2022,39(2):281-286.
- [21] 陈逸枫,张大海,李宇欣,等. 基于DWT-MOSMA-SVM的多目标优化短期母线负荷预测[J]. 电力建设,2023,44(3):49-55.  
CHEN Yicong,ZHANG Dahai,LI Yuxin,et al. Multi-objective optimization based on DWT-MOSMA-SVM for short-term bus load forecasting[J]. Electric Power Construction,2023,44(3):49-55.
- [22] CHEN Z C,HAN F C,WU L J,et al. Random forest based intelligent fault diagnosis for PV arrays using array voltage and string currents[J]. Energy Conversion and Management,2018,178:250-264.
- [23] 栗然,丁星,孙帆,等. 基于Wide&Deep-XGB2LSTM模型的超短期光伏功率预测[J]. 电力自动化设备,2021,41(7):31-37.  
LI Ran,DING Xing,SUN Fan,et al. Ultra-short-term photovoltaic power prediction based on Wide & Deep-XGB2LSTM

- model[J]. Electric Power Automation Equipment,2021,41(7):31-37.
- [24] MENG Y, YANG Q, CHEN S L, et al. Multi-branch AC arc fault detection based on ICEEMDAN and LightGBM algorithm [J]. Electric Power Systems Research, 2023, 220: 109286.
- [25] LUNDBERG S M, LEE S I. A unified approach to interpreting model predictions[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, California, USA: ACM, 2017: 4768-4777.
- [26] 朱春霖, 余成波. 基于LightGBM算法的光伏并网系统孤岛检测及其集成的可解释研究[J]. 电力自动化设备, 2023, 43(7): 80-86, 150.
- ZHU Chunlin, YU Chengbo. Islanding detection of grid-connected photovoltaic system based on LightGBM algorithm and its integrated interpretability analysis[J]. Electric Power Automation Equipment, 2023, 43(7): 80-86, 150.
- [27] 林志萍. 可解释的机器学习及应用研究[D]. 广州: 华南理工大学, 2021.
- LIN Zhiping. Interpretable machine learning and its application research[D]. Guangzhou: South China University of Technology, 2021.

#### 作者简介:

陈泽(1999—),男,硕士研究生,主要研究方向为新能源状态评估(**E-mail**: czhenshuai1113@163.com);  
 刘文泽(1973—),男,副教授,博士,主要研究方向为能源互联网、电力系统可靠性、配电网规划等(**E-mail**: liuwzh@scut.edu.cn)。

(编辑 李玮)

## Interpretable intelligent integration method for photovoltaic array fault diagnosis

CHEN Ze<sup>1</sup>, LIU Wenze<sup>1</sup>, WANG Kangde<sup>1</sup>, YU Tao<sup>1,2</sup>, HUANG Zhanhong<sup>1,2</sup>

(1. School of Electric Power, South China University of Technology, Guangzhou 510641, China;

2. Guangdong Provincial Key Laboratory of Intelligent Measurement and Advanced Metering of Power Grid, Guangzhou 510641, China)

**Abstract:** Aiming at the problems of weak generalization and poor interpretability of existing intelligent methods in photovoltaic array fault detection and diagnosis, an interpretable intelligent integration method is proposed. The feature mining is performed on the collected output time-series of voltage and current waveforms of the photovoltaic array, and multiple mature intelligent algorithms that have been applied to photovoltaic fault diagnosis are used as different base learners and meta learners to construct a Stacking ensemble learning model that combines the advantages of different intelligent algorithms and is more generalized. Then, taking the Shapley additive explanation method as the overall framework, combined with the local approximate interpretable method, the model training process and results are explained and analyzed. By obtaining the contributions of each feature, analyzing the decision-making mechanism of the integrated model, and understanding how to diagnose it, the reliability and credibility of the model are improved. The experimental results of case study show that the proposed interpretable intelligent integration method achieves high-precision fault diagnosis in testing on datasets of different sizes. The interpretability results of the model indicate that the mapping of fault features and diagnostic results established by the intelligent integration model follows physical insights, enhancing the credibility and transparency of the intelligent method.

**Key words:** photovoltaic array; fault diagnosis; Stacking ensemble; interpretable intelligent method; Shapley additive explanation method



## 附录 A

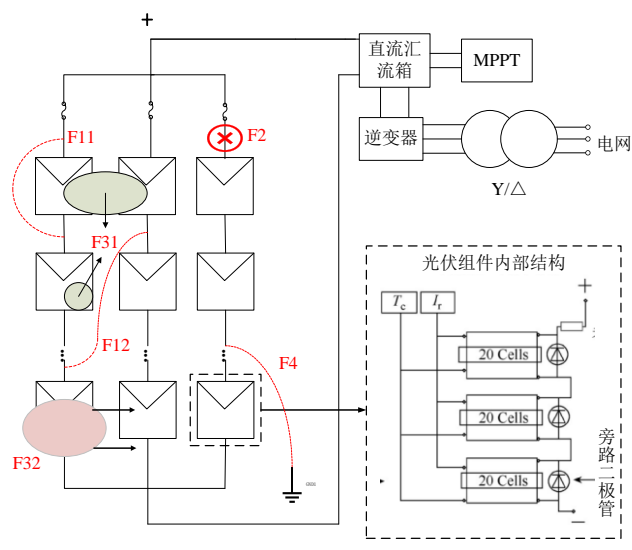


图 A1 光伏发电系统结构

Fig.A1 Structure of photovoltaic power generation system

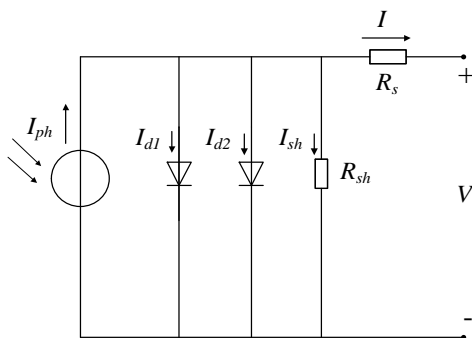


图 A2 光伏组件等效电路模型

Fig.A2 Equivalent circuit model of photovoltaic module

## 附录 B

表 B1 故障特征及描述

Table B1 Fault characteristics and description

特征编号	特征名称	特征含义描述	特征编号	特征名称	特征含义描述
$F_{feature1}$	$I_{PV-mean}$	阵列输出平均电流	$F_{feature6}$	$V_{Range}$	电压最大与最小值差
$F_{feature2}$	$V_{DC-mean}$	直流侧平均电压	$F_{feature7}$	$I_{CF}$	电流峰值系数
$F_{feature3}$	$I_{Var}$	输出电流方差	$F_{feature8}$	$V_{CF}$	电压峰值系数
$F_{feature4}$	$V_{Var}$	直流侧电压方差	$F_{feature9}$	$I_R$	辐射度
$F_{feature5}$	$I_{Range}$	电流最大与最小值差	$F_{feature10}$	$T$	环境温度

表 B2 故障数据集描述

Table B2 Description of fault data set

运行状态	标注	数据集数量	数据占比/%	标签
正常	Normal	120	20.0	0
线对线 1	F <sub>11</sub>	85	14.2	1
线对线 2	F <sub>12</sub>	60	10.0	2
开路	F <sub>2</sub>	100	16.7	3
遮光 1	F <sub>31</sub>	135	22.5	4
遮光 2	F <sub>32</sub>	100	16.7	5

表 B3 各算法性能表现

Table B3 Performance of each algorithm

训练集数据占数据集比例/%	F <sub>1</sub> /%					A <sub>cc</sub> /%				
	KNN	SVM	RF	XGBoost	Stacking	KNN	SVM	RF	XGBoost	Stacking
30	89.42	88.17	93.31	95.74	95.13	89.28	88.09	93.33	95.75	95.09
50	92.47	91.90	95.69	96.03	96.03	92.33	91.76	95.66	96.11	96.11
70	95.73	94.88	96.68	98.58	97.43	95.85	95.0	96.66	98.45	97.50
90	98.33	97.16	96.5	98.25	98.31	98.33	97.16	96.67	98.33	98.33

## 附录 C

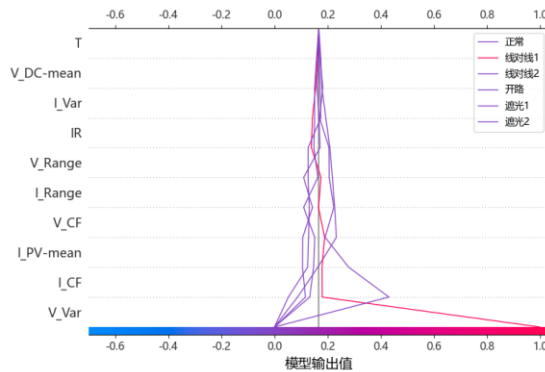


图 C1 模型决策图

Fig.C1 Model decision diagram

图 C1 中，中间灰色垂直直线标记了模型的基础值，红色线是预测结果类，线的轨迹表示左侧每个特征对模型决策的影响，模型输出值为各类型的预测概率。

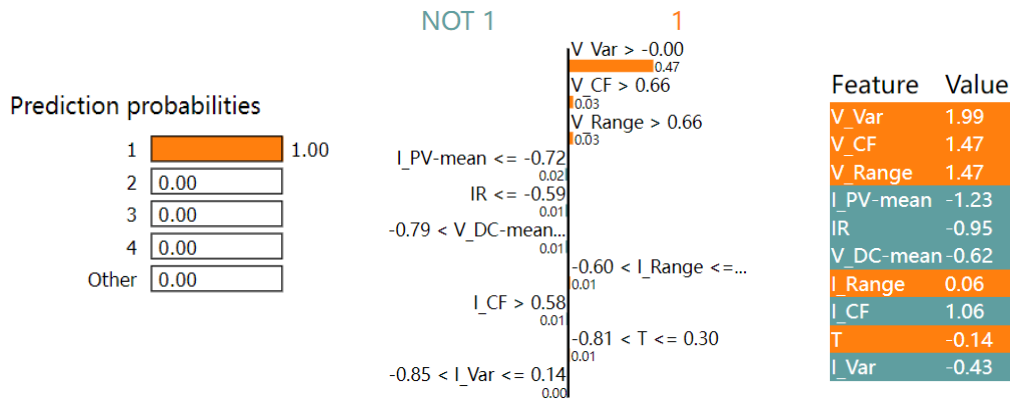


图 C2 局部决策解释

Fig.C2 Local decision interpretation

图 C2 中，左侧代表原诊断模型对解释点的输出值，(NOT 1/1) 为权重图，其左侧代表消极权重，右侧为积极权重，右侧为模型特征值，为标准化处理后的结果。

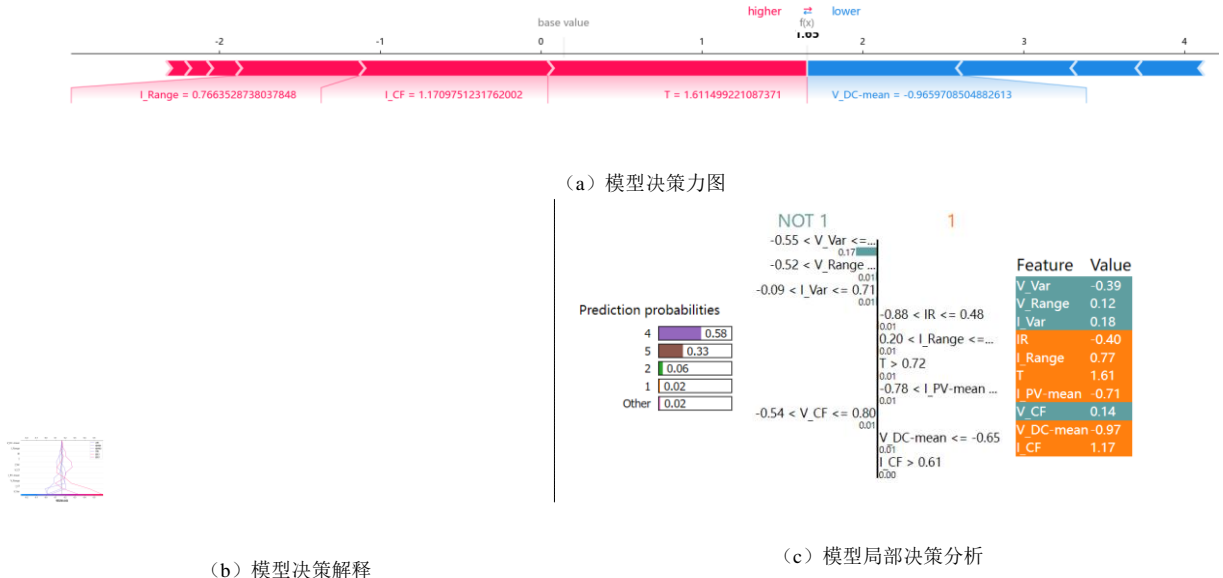


图 C3 错判样本的决策过程

Fig.C3 Decision-making process of misjudging sample

图 C3 (a) 模型决策力，图中  $f(x)$  为模型输出值，base value 为基值，将预测推高的特征用红色表示，将预测推低的特征用蓝色表示，箭头越长，特征对输出的影响越大；图 C3 (b) 中灰色线为模型输出，蓝色线表示特征值对模型输出的作用；图 C3 (c) 中特征值为标准化值。