

DOI:10.19651/j.cnki.emt.2211433

基于 XGBoost 的光伏阵列故障诊断方法研究*

刘行行¹ 帕孜来·马合木提¹ 程志江^{1,2} 李高原¹ 周昂¹

(1. 新疆大学电气工程学院 乌鲁木齐 830017; 2. 可再生能源发电与并网控制教育部工程研究中心 乌鲁木齐 830017)

摘要: 针对光伏阵列输出具有非线性并受最大功率点跟踪影响,从而导致传统的诊断方法精度低、模型性能差等问题,提出一种基于极端梯度提升的光伏阵列故障诊断方法。首先,在光伏电池单二极管模型的基础上,建立光伏阵列仿真模型,利用 PVsyst 软件对光伏阵列的输出特性和故障成因进行系统的模拟分析,得到了故障特征参数,并通过特征重要度排序验证了所选择故障特征参数的有效性;其次,提取光伏阵列不同故障状态下的故障特征,构建基于 XGBoost 的故障诊断模型;最后,利用网格搜索和交叉验证对诊断模型的超参数进行寻优,通过混淆矩阵计算评价指标来评估诊断模型的性能。并将该方法与决策树、随机森林以及梯度提升树相比,结果表明,该方法不仅能准确检测所有的故障种类,并且模型的泛化能力更好,诊断准确率更高。

关键词: 光伏阵列;故障诊断;极端梯度提升;网格搜索;混淆矩阵

中图分类号: TM914.4 **文献标识码:** A **国家标准学科分类代码:** 510.40

Research on fault diagnosis method of photovoltaic array based on extreme gradient boosting

Liu Hanghang¹ Pazilai·Mahemuti¹ Cheng Zhijiang^{1,2} Li Gaoyuan¹ Zhou Ang¹

(1. School of Electrical Engineering, Xinjiang University, Urumqi 830017, China; 2. Renewable Energy Power Generation and Grid Connection Control Engineering Research Center of the Ministry of Education, Urumqi 830017, China)

Abstract: An extreme gradient boosting (XGBoost) method for photovoltaic array fault diagnosis is proposed to solve the problems of low precision and poor model performance due to the nonlinear output of photovoltaic array and the influence of maximum power point tracking. Firstly, based on the single diode model of photovoltaic cell, a simulation model of photovoltaic array was established, and the output characteristics and fault causes of photovoltaic array were systematically simulated and analyzed by using PVsyst software, and the fault characteristic parameters were obtained and the validity of the selected fault feature parameters is verified by the feature importance ranking. Secondly, the fault characteristics of photovoltaic array under different fault states are extracted, and the fault diagnosis model based on XGBoost is constructed. Finally, grid search and cross validation were used to optimize the hyperparameters of the diagnostic model, and the performance of the diagnostic model was evaluated by confounding matrix calculation. Compared with decision tree, random forest and gradient lifting tree, the results show that the proposed method not only can accurately detect all kinds of faults, but also has better generalization ability and higher diagnosis accuracy.

Keywords: photovoltaic array; fault diagnosis; extreme gradient boosting; grid search; confusion matrix

0 引言

目前,日益增长的能源需求和传统化石能源带来的环境污染已成为当下急需解决的问题。光伏作为双碳战略中最重要的可持续能源之一,装机容量不断增加,截止 2021 年底,累计并网容量 3.06 亿 kW。光伏阵列作为光伏发电系统的关键部件,在恶劣条件下连续运行容易发生故障^[1]。

目前,光伏发电领域所面临的问题是降本增益和提高光伏发电效率,而高效的智能运维是光伏发电降本增益的核心之一^[2]。因此,对光伏系统进行快速准确的故障诊断,不仅能提高系统发电的可靠性,而且能大大的降低运维成本。

光伏发电系统包括光伏阵列、DC-DC 和 DC-AC 变换器,其输出特性中连续部分与离散部分相互耦合,是一个典型的混杂系统。光伏阵列作为光伏电站的核心设备之一,

收稿日期:2022-09-18

* 基金项目:国家自然科学基金(61963034)、自治区重点实验室开放课题(2021D04011)项目资助

对其故障进行实时的检测,有利于运维人员通过干预措施来保证系统稳定运行。传统的光伏阵列检测常采用过流保护和接地保护装置^[3],由于光伏阵列输出非线性,以及低辐照度和最大功率点跟踪的影响,保护装置容易漏报和误动作。目前,国内外学者对光伏阵列的诊断研究主要集中在以下方面:文献[4-6]利用红外热成像仪来对光伏阵列进行故障监测,这些方法不需要测量电气量就能对阵列进行有效检测,但难以区分多种故障,并且对红外设备的精度要求较高,容易受外界环境的干扰,工程应用较为局限。基于传感器检测的诊断方法^[7-10],贾嵘等^[8]通过布置电压和电流传感器来获取输出的电流和电压,虽然检测方法较为简单,但布置大量的传感器本身会造成系统的故障率升高,并且造成光伏装机成本的增加,不适用于大型的光伏电站。基于数学模型的诊断方法^[11-13],对单个故障检测准确度高,但十分依赖模型的精度,所建模型具有单一性,应用较为狭窄,同时难以定位故障位置。基于人工智能的诊断方法^[14-18],文献[15]利用输出电压和功率曲线来对阵列故障进行分类,文献[16]提出了一种 DA-SOM 算法来对光伏阵列进行诊断,文献[17]能准确检测出光伏发电系统直流串联微弱故障电弧故障,这些方法对故障的辨识精度较高,但所需的数据样本较大,文献[18]虽使用小样本,但特征选取困难,模型参数确定较为复杂。

以上研究对光伏阵列的故障诊断提供了各种思路,但仍存在诊断模型性能低、参数选取困难、检测设备成本昂贵、工程应用局限等问题。同时没有考虑组串之间短路和阵列老化故障的情况。因此,结合光伏组件的输出特性和外部环境因素,提出一种基于极端梯度提升(extreme gradient boosting, XGBoost)光伏阵列故障诊断方法。首先,建立了光伏电池的单二极管模型,利用 PV_{system} 软件对光伏组件的输出状态进行了研究,分析正常与故障条件下的输出特性及故障成因,确定了不同故障类型下的特征参数,并通过特征重要度排序验证了所选择故障特征的有效性。其次,利用网格搜索(GridSearchCV, GC)和交叉验证对 XGBoost 的参数进行寻优,并通过混淆矩阵进行多指标模型评估。最后通过实验验证提出的网格搜索寻优后的 XGBoost (gridsearchcv extreme gradient boost, GC-XGBoost)诊断方法光伏阵列故障诊断准确率更高、诊断模型性能更好,对光伏阵列的 6 种故障诊断准确率达到 94.56%,尤其对阵列老化和串间短路的准确率达到 99%。

1 光伏电池原理及特性分析

光伏电池作为光伏组件的基本单元,有阳光照射时,光伏电池内部将吸收光子产生电子-空穴对,由于电场的存在,电子开始移动,经过一定时间的聚集,从而形成端电压。当两端接上负载时,负载会有电流流过,从而产生功率。光伏电池在使用过程中会受到各种不确定性因素的影响,如环境因素(温度、光照强度、雨、雪、大风、沙尘等),导致光伏

电池输出不稳定,图 1 为光伏电池单二极管 5 参数等效电路。

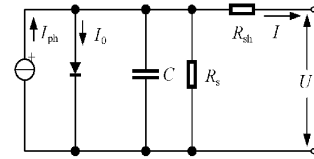


图 1 光伏电池等效电路

光伏组件由光伏电池串联而成的,而光伏阵列由光伏组件串并联形成。在理想状态下,光伏电池可被等同为一个理想的直流电源,再加并联一个二极管组成,图 1 中 R_s 为等效串联电阻, R_{sh} 为分流电阻,由基尔霍夫电流定律可得:

$$I = I_{ph} - I_0 - I_{RS} \quad (1)$$

式中: I_{ph} 为光生电流, I_0 为二极管反向饱和电流, I_{RS} 为流经 R_s 的电流。

光伏电池等效电路的输出特性方程为:

$$I = I_{ph} - I_0 \left\{ \exp \left[\frac{q(U + IR_s)}{AkT} \right] - 1 \right\} - \left(\frac{U + IR_s}{R_{sh}} \right) \quad (2)$$

式中: A 为 P-N 结理想因子, T 为光伏电池实际温度, k 为玻尔兹曼常量, $k = 1.381 \times 10^{-23}$ J/K, q 为电荷常数, $q = 1.602 \times 10^{-19}$ C。

以某光伏公司型号为 LG400N2W-A5 的组件为研究对象,在标准工况下(25 °C, 1 000 W/m²)具体的规格参数如表 1 所示。

表 1 LG400N2W-A5 规格参数表

组件参数	标称值	组件参数	标称值
电池片数量	72	短路电流(I_{sc})	10.47 A
最大功率(P_{max})	400.5 W	电流温度系数	0.030%/°C
开路电压(U_{oc})	49.3 V	最大功率 电流(I_m)	9.96 A
电压温度系数	-0.36%/°C	分流电阻(R_{sh})	1 500 Ω
最大功率电压 (U_m)	40.2 V	串联电阻(R_s)	0.33 Ω

PV_{system} 软件是一款工程应用级的光伏专业仿真软件,可以在其数据库查询不同组件的各项参数,并对光伏组件的各种状态进行仿真。光照强度保持不变,不同温度区间的光伏组件输出特性曲线如图 2 所示。

从图 2 中可以得出不同温度对输出的 I_{sc} 影响较小,而 U_{oc} 、 U_m 随着温度的升高而减小,最大功率 P_{max} 减小,同时对 I_m 影响较小。组件温度为 25 °C 时,不同光照强度区间的光伏组件输出特性曲线如图 3 所示。从图 3 中可得到组件的 I_{sc} 则随着光照强度的降低而减小, U_{oc} 有轻微的减小,最大功率点向正下方移动, P_{max} 减小, U_m 不受影响, I_m 随之减小。

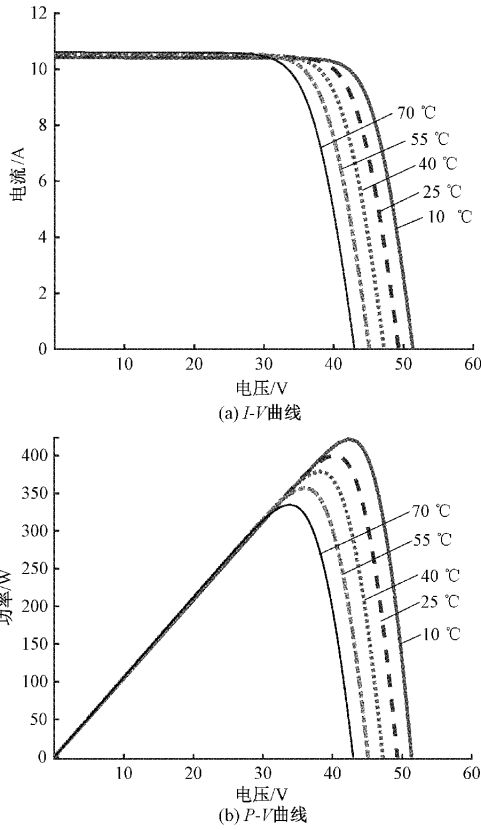


图 2 光照强度一定,不同温度时输出特性

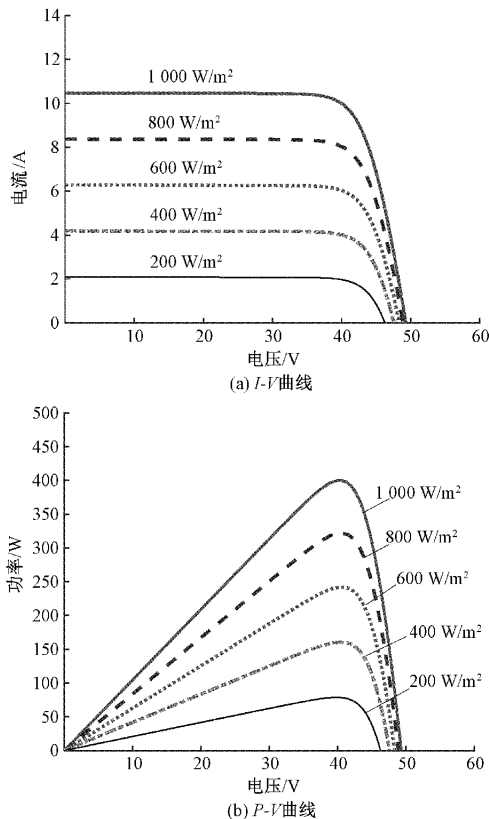


图 3 温度一定,不同光照强度时输出特性

2 光伏阵列故障类型及特征

光伏阵列的故障分为暂时和永久的故障,常见的故障有线对地、线对线,开路,电弧故障、阴影故障、光伏组件异常退化、模块不匹配、接线盒故障、电池裂纹、分层、热点、防反射涂层缺陷、积灰、接口老化腐蚀,最大功率点跟踪器(maximum power point tracking, MPPT)故障等。为了准确获取阵列故障时的状态和输出特性,搭建如图 4 所示功率为 6.4 kW 的 4×4 光伏阵列模型,用来分析组串开路故障串内串间短路、阴影故障、老化故障共 6 种类别的故障特性及成因。

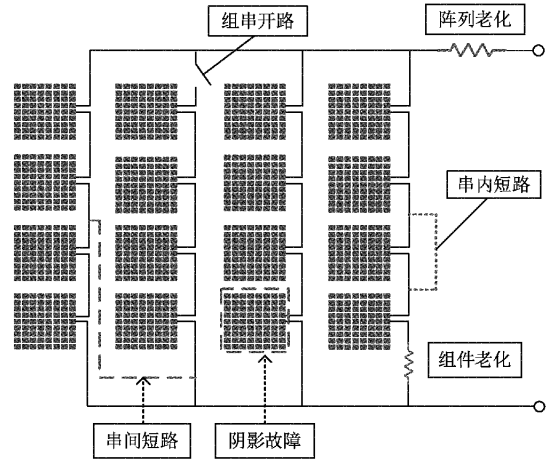


图 4 光伏阵列故障机理模型

模拟图 4 中故障类型,不同故障状态下的输出特性曲线如图 5 所示,可以得到各故障阵列的输出特性,从而推断出可能故障成因。

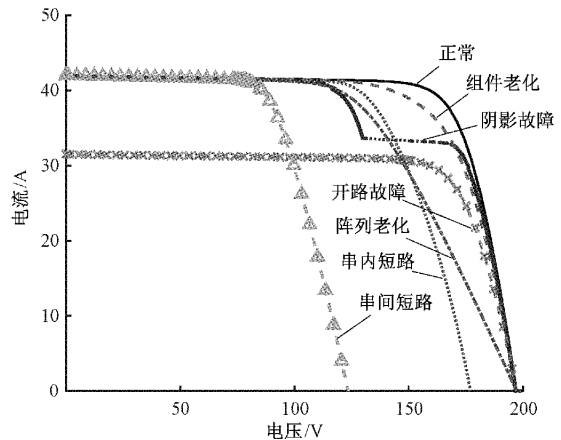


图 5 不同工况下阵列的输出特性 $I-V$ 曲线

阵列发生一个组串或多个组串开路时,短路电流随着组串开路数量的增加而递减,开路电压几乎不变,可能的故障原因是组件连接线断开,或是长期工作在恶劣环境或维护期间导致的电池板内部损伤。组串内部短路时,开路电压减小的幅度与组件短路的数量有关,组串之间短路时,开

路电压骤减,而短路电流波动较小,可能的故障原因是连接电缆绝缘性差,MC-4 连接器绝缘性低或机械损伤导致不同电位点的导通。发生阴影故障时,系统的输出功率随着阴影遮挡的阵列面积增加而减小,同时最大功率点电流减小较为明显,可能是由于云层,树木带来的阴影,这些是暂时性的,还有可能是长期积灰或鸟粪造成的遮挡。组件退化时系统的输出功率减小的幅度与组件老化的个数有关,同时最大功率点电压下降的较为明显,可能的原因是组件表面脏污导致热斑,长期暴露在恶劣环境导致的老化,从而导致串联电阻增加。经过以上深入分析,本文选取光伏阵列的电流、电压和最大功率作为主要的故障特征,光照强度和温度作为一般特征进行评判。

3 XGBoost 的故障检测模型

XGBoost 是一种基于提升树的模型^[19],由多个决策树共同进行决策,其主要思想是不断的添加新的树去拟合上次得到的残差获得新的目标函数,并进行迭代计算。假设光伏阵列的故障数据集为 $D = \{(x_i, y_i)\} (|D| = n, x_i \in R^m, y_i \in R)$, n 为故障样本集数目, m 为故障特征种类。使用 k 个附加函数构建预测模型,即:

$$\hat{y}_i = \varphi(x_i) = \sum_{k=1}^K f_k(x_i), f_k \in F \quad (3)$$

式中: $F = \{f(x_i) = w_q(x_i)\} (q: R^m \rightarrow T, w \in R^T)$ 由多个分类 CART 树组成, q 代表树结构, T 是树中的叶子结点的数目, f_k 对应于一个独立的树结构 q 和叶权重为 w , 与决策树的区别在于 XGBoost 将目标函数简化为:

$$L^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_k \Omega(f_k) \quad (4)$$

$$\Omega(f_k) = \gamma T + \frac{1}{2} \gamma \|w\|^2 = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (5)$$

式中: $\sum_i l(y_i, \hat{y}_i)$ 是目标损失函数, y_i 是真实的样本值, \hat{y}_i 是模型的预测结果, $\Omega(f_k)$ 是正则化项, $\sum_k \Omega(f_k)$ 表示树的复杂度函数,加入正则项可以调整模型的复杂度,预防模型出现过拟合,其值越小越简单,泛化能力越强, γ 和 λ 为惩罚项,抑制随着叶子节点增加导致模型更加复杂。经过 t 次迭代后,设前 t 迭代优化得到的模型为 $\hat{y}_i^{(t-1)}$, 则在第 t 代得到的模型为 $\hat{y}_i^{(t-1)} + f_i(x_i)$, $f_i(x_i)$ 是第 t 代新加入的模型,即:

$$L^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_i(x_i)) + \Omega(f_i) \quad (6)$$

利用二阶泰勒展开法,将常数项去掉并加入正则化项,可得:

$$\tilde{L}^{(t)} \cong \sum_{i=1}^n \left[g_i f_i(x_i) + \frac{1}{2} h_i f_i^2(x_i) \right] + \Omega(f_i) \quad (7)$$

式中: $g_i = \partial_{y^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)})$ 和 $h_i = \partial_{y^{(t-1)}}^2 l(y_i, \hat{y}_i^{(t-1)})$ 分别为损失函数的一阶偏导和二阶偏导。将式(5)代入可得:

$$\tilde{L}^{(t)} = \sum_{j=1}^T \left[\left(\sum_{i \in I_j} g_i \right) w_j + \frac{1}{2} \left(\sum_{i \in I_j} h_i + \lambda \right) w_j^2 \right] + \gamma T \quad (8)$$

则最优 w_j 的表达式如下:

$$w_j^* = - \frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda} \quad (9)$$

将最优 w_j 代入式(8)中可得到最优的目标值:

$$L(\tilde{t}) = - \frac{1}{2} \sum_{j=1}^T \frac{\left(\sum_{i \in I_j} g_i \right)^2}{\sum_{i \in I_j} h_i + \lambda} + \gamma T \quad (10)$$

$L(\tilde{t})$ 代表树的结构分数,其值越小精度越高。通过上述分析,将选择的特征作为模型的输入,根据特征来不断添加决策树,最终求取最优的目标值。所求的目标值越低,说明建立的模型越稳定,效果越好。同时利用网格搜索和 10 折交叉验证来获取模型的全局最优参数。网格搜索是通过遍历给定的参数区间不断的来寻找每一项参数的最优值,而采用 K 折交叉验证是将训练集分成 S 组,每次抽取 $S-1$ 组进行训练,得到 S 个模型,最后利用测试集进行模型测试。两者结合可以有效的减小训练过程中随机取样偏差,并获得最优诊断模型。GC-XGBoost 诊断流程如图 6 所示。

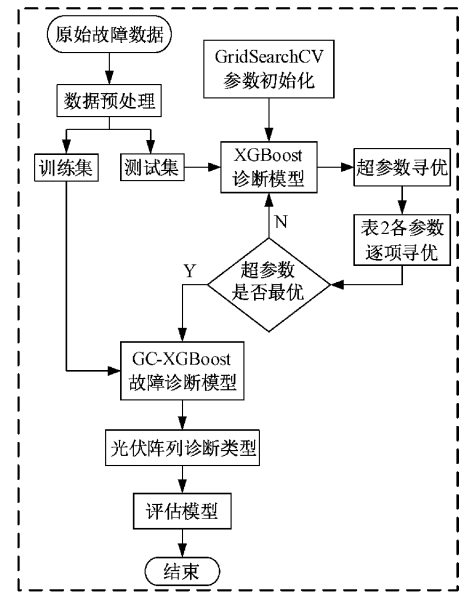


图 6 基于 GC-XGBoost 的故障诊断流程

诊断模型的具体步骤如下:

- 1) 数据预处理,对光伏阵列的不同故障状态进行模拟分析,获取原始故障数据样本集,并划分训练集和测试集。
- 2) 利用网格搜索法的搜索能力,对诊断模型中的关键参数进行寻优,并进行 10 倍交叉验证,得到诊断模型的最优参数。
- 3) 利用寻优后的模型参数建立 XGBoost 光伏阵列诊

断模型,使用训练集对模型进行训练。

4)将测试集送入优化后的模型进行测试,获得诊断结果,并利用混淆矩阵计算相关评估参数,对诊断模型的性能进行评估。

为了防止建立的诊断模型过拟合,并提高 XGBoost 模型分类预测的性能。因此,利用网格搜索来寻找模型的关键参数^[20]。由于 XGBoost 的参数较多,因此选取主要参数进行优化。通用参数有树模型和线性模型,在这里选用 gmtree 树模型,使用线程参数 nthread,它可以用于并行处理,本实验平台 CPU 是 6 核 12 线程的 i5-11400H,采用默认使用全部线程 nthread=1,可以根据处理器性能选择设置。学习目标参数,objective:设置为多分类 softmax,类别 num_class 为 7;而需要的主要优化参数的 boost 超参数如表 2 所示。

表 2 XGBoost 参数优化表

优化参数	参数范围	默认值	GridSearchCV 寻优后
迭代次数	$(0, +\infty)$	500	130
树的深度	$(1, +\infty)$	6	6
惩罚项系数	$[0, +\infty)$	0	0.07
最小权重总和	$(0, \infty)$	1	1
随机样本	$(0, 1)$	1	0.95
学习率	$[0, 1]$	0.3	0.13

为了验证诊断模型的性能,采用微观计算评估指标(micro-average),这样可以有效降低小样本量带来的误差影响。多分类评价指标包括召回率(recall)、召回率和精确率协调平均值(F1-score),受试者工作特征曲线(receiver operating characteristic curve, ROC),ROC 曲线下面积(area under the curve, AUC)。这些指标的可以利用混淆矩阵来计算,如表 3 所示。

表 3 混淆矩阵

混淆矩阵	预测样本		
		正	负
实际样本	正	真阳性(TP)	假阴性(FN)
	负	假阳性(FP)	真阴性(TN)

各评估指标计算公式如下:

$$accuracy = \frac{TP + TN}{TP + TN + FN + FP} \quad (11)$$

$$precision = \frac{TP}{TP + FP} \quad (12)$$

$$recall = \frac{TP}{TP + FN} \quad (13)$$

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \quad (14)$$

其中,accuracy 表示准确率;precision 是真阳性占正

确总阳性预测样本的比率;recall 表示正确分类的阳性样本与总阳性样本的衡量;F1 表示 precision 和 recall 的调和平均值,其值在 0~1 范围内,用于区分模型分类性能。获取这些评估指标可以更全面的评价诊断模型的性能。

4 实验结果及分析

在 MATLAB/Simulink 中搭建 4×4 功率为 6.4 kW 的光伏阵列,阵列的辐照度范围为 300~1 000 W/m²,温度范围为 25 °C~55 °C,模拟不同故障的状态。获得了 735 组样本数据,将 588 组数据作为训练模型样本,147 组为测试样本。不同故障的数据样本分布可视化如图 7 所示。

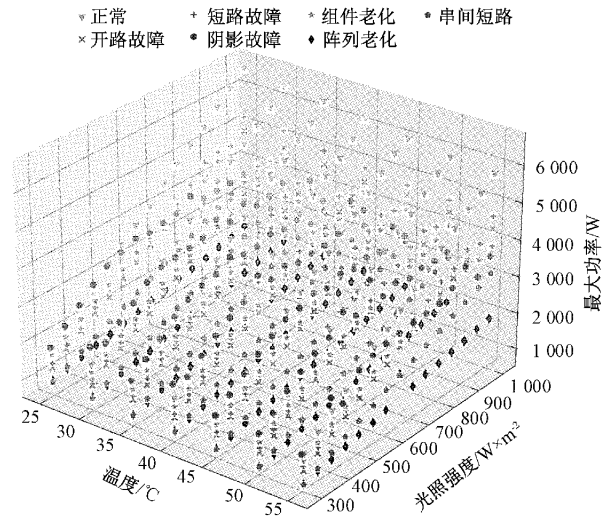


图 7 不同故障下的数据分布

由图 7 可得,不同条件下正常状态的数据分布可作为参考,将 6 种类别的故障数据分别与正常状态时相比较,可以剔除该类故障误差较大的异常点。同时利用 F-score 的数据分析方法,来评价数据故障特征的重要程度,如图 8 所示。

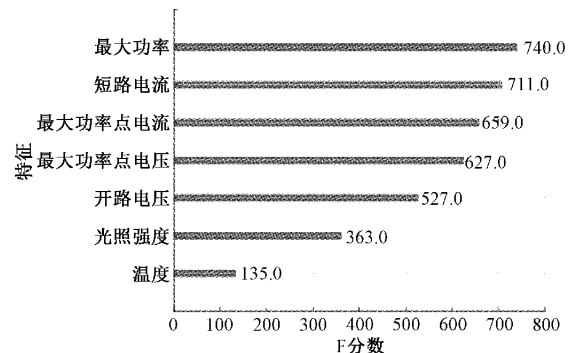


图 8 F-score

最大功率 P_{max} 和电流 I 的特征重要评分较高,表明发生故障时,功率和电流的波动较大;而光照强度 S 和温度 T 的较低,是由于光照强度和温度受外部环境的影响较大,说明选择的特征可以有效区分故障。

利用 GC 寻优后的参数建立 GC-XGBoost 光伏阵列诊断模型,设置对照实验组并进行 10 倍交叉验证,与决策树,随机森林(random forest, RF)和梯度提升树(gradient boosting decision tree, GDBT)进行对比,实验结果如表 4 所示。

表 4 不同模型诊断精度

故障类型	诊断准确率				
	决策树	GDBT	RF	XGBoost	GC-XGBoost
正常	0.82	0.83	0.86	0.88	0.91
开路故障	0.88	0.91	0.90	0.94	0.97
短路故障	0.92	0.88	0.91	0.95	0.96
阴影故障	0.80	0.83	0.82	0.83	0.84
组件老化	0.86	0.87	0.94	0.96	0.97
阵列老化	0.93	0.98	0.98	0.97	0.99
串间短路	0.92	0.96	0.98	0.98	0.99

由表 4 得,不同诊断模型对各种故障状态的精度不同,GC-XGBoost 的整体精度要优于其他诊断模型。为了验证不同诊断模型的性能,依据混淆矩阵及式(11)~(14)计算模型的评价指标如表 5 所示。

表 5 不同诊断模型评价

诊断模型	准确率/%	Recall	F1-score	AUC
决策树	89.12	0.89	0.90	0.81
GDBT	91.16	0.91	0.91	0.85
RF	92.52	0.92	0.93	0.84
XGBoost	93.20	0.93	0.93	0.87
GC-XGBoost	94.56	0.95	0.95	0.95

由表 5 结果可得,决策树、梯度提升树和随机森林诊断模型的准确率分别是 89.12%,91.16%,92.52%,未经优化的 XGBoost 的诊断准确率为 93.20%,而同等条件下,与其他模型相比,GC-XGBoost 模型准确率达到 94.56%,均高于其他模型。同时该模型的召回率和 F1-score 分别为 0.95 和 0.95,也优于其他模型。GC-XGBoost 诊断模型的 ROC 曲线如图 9 所示。

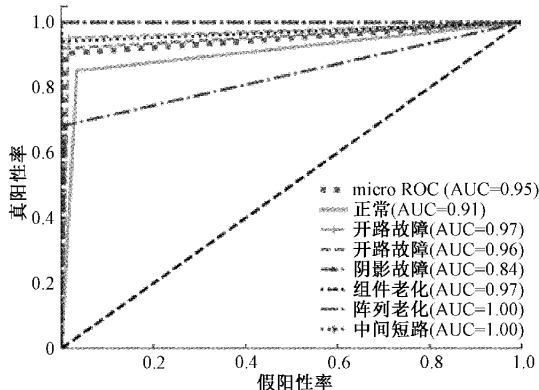


图 9 诊断模型 ROC 曲线

从图 9 中可以得出 GC-XGBoost 模型对阵列诊断的性能好,AUC 达到了 0.95,尤其对阵列老化和串间短路的效果较好,AUC 为 1。而图中阴影故障 AUC 较低的原因是,光伏阵列运行过程中具有环境不确定性和复杂性,特别是低照度下故障特征变化小,易被分类为正常状态,导致整体的准确率微弱下降,但从评价指标来分析,GC-XGBoost 诊断模型的总体性能优于其他诊断模型。

5 结 论

本文首先对光伏阵列发生故障时的输出状态进行了深入的研究,获得了不同故障状态下的输出特性,分析了光伏阵列故障的成因,同时利用 GC 对 XGBoost 诊断模型的超参数进行寻优,并与决策树、RF 和 GDBT 相比,结果表明,GC-XGBoost 的诊断效果更好,性能更优。该方法尤其对阵列老化故障和串间短路故障识别率更高。在工程应用方面,XGBoost 可以并行计算处理数据,能极大的提高计算速度,其分布式版本有广泛的移植性,在大型的光伏电站和分布式屋顶电站可以建立端到端系统,不仅能提高光伏发电的安全性和可靠性,而且能降低光伏装机成本。

参考文献

- [1] 郑伟烁,郑文悦,李志伟,等. 光伏电站在线监测技术现状与进步趋势展望[J]. 电测与仪表, 2021,58(9): 1-7.
- [2] 马铭遥,张志祥,刘恒,等. 基于 I-V 特性分析的晶硅光伏组件故障诊断[J]. 太阳能学报, 2021,42(6): 130-137.
- [3] 李光辉,段晨东,武珊. 基于半监督机器学习法的光伏阵列故障诊断[J]. 电网技术, 2020, 44(5): 1908-1913.
- [4] AHMED W, ALI M U, HUSSAIN S J, et al. Visual vocabulary based photovoltaic health monitoring system using infrared thermography [J]. IEEE Access, 2022,10:14409-14417.
- [5] 毛峡,李亚豪. 光伏阵列故障检测中的无人机红外图像拼接[J]. 太阳能学报, 2020,41(3): 262-269.
- [6] 蒋琳,苏建徽,李欣,等. 基于可见光和红外热图像融合的光伏阵列热斑检测方法[J]. 太阳能学报, 2022, 43(1):393-397.
- [7] AI-DAHOUD A, FEZARI M, BELHOUCHE F Z, et al. Remote monitoring system for solar power panels using intelligent sensors network[J]. Telecommunications Forum(TELFOR). IEEE, 2016:1-4.
- [8] 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.
- [9] 苏懿生,郑子君,张经炜,等. 光伏阵列无线传感节点设计及故障诊断应用[J]. 电子测量技术, 2022,45(7):

- 8-13.
- [10] HU Y, ZHANG J, CAO W, et al. Online two-section PV array fault diagnosis with optimized voltage sensor locations [J]. IEEE Transactions on Industrial Electronics, 2015, 62(11): 7237-7246.
- [11] 余玲珍, 杨靖, 龙道银, 等. 典型故障条件下光伏阵列建模与仿真[J]. 电网与清洁能源, 2020, 36(10): 103-111, 118.
- [12] 刘兴杰, 郭栋, 王凯龙. 基于电气外特性的光伏发电系统模型等效方法[J]. 电工技术学报, 2014, 29(10): 231-238, 265.
- [13] BACHA M, TERKI A. Diagnosis algorithm and detection faults based on fuzzy logic for PV panel[J]. Materials Today: Proceedings, 2022, 51: 2131-2138.
- [14] 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.
- [15] NIETO A E, RUIZ F, PATINO D, et al. Classification of electric faults in photovoltaic systems based on voltage-power curves [J]. IEEE Latin America Transactions, 2021, 19(12): 2071-2078.
- [16] 张晓阳, 李田泽, 张涵瑞, 等. 应用于光伏阵列故障诊断的 DA-SOM 算法研究[J]. 电源学报, 2022, 20(2): 122-128.
- [17] 唐圣学, 刁旭东, 陈丽, 等. 光伏发电系统直流串联微弱故障电弧检测方法研究[J]. 仪器仪表学报, 2021, 42(3): 150-160.
- [18] 孙海蓉, 李帅. 改进胶囊网络的小样本光伏热斑识别方法[J]. 中国测试, 2023, 49(2): 106-112.
- [19] CHEN T, GUESTRIN C. XGBoost: A scalable tree boosting system [C]. Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, 2016: 785-794.
- [20] 王雨虹, 王志中. 基于 RFRFE 与 ISSA-XGBoost 的变压器故障辨识方法[J]. 电子测量与仪器学报, 2021, 35(12): 142-150.

作者简介

刘行行, 硕士研究生, 主要研究方向为新能源发电、光伏系统故障诊断。

E-mail: 332062204@qq.com

帕孜来·马合木提, 教授, 主要研究方向为过程控制和智能故障诊断。

E-mail: lhh07052022@163.com

程志江, 博士, 副教授, 主要研究方向为可再生能源发电系统、微电网及储能控制。

E-mail: 67078267@qq.com

李高原, 硕士研究生, 主要研究方向为大功率并网逆变器故障检测及寿命预测。

E-mail: 1525533258@qq.com

周昂, 硕士研究生, 主要研究方向为大功率并网逆变器故障检测寿命预测。

E-mail: 710231465@qq.com