

9-10-2024

Transformer fault diagnosis based on DGA and TPE-LightGBM

Jinxin YANG

Department of Energy and Electrical Engineering, Nanchang University, Nanchang 330031, China

Caibo LIAO

Department of Energy and Electrical Engineering, Nanchang University, Nanchang 330031, China,
Lcb1752@126.com

Xiong HU

Department of Energy and Electrical Engineering, Nanchang University, Nanchang 330031, China

Wenqing ZHU

Department of Energy and Electrical Engineering, Nanchang University, Nanchang 330031, China

Xu ZHANG

Department of Energy and Electrical Engineering, Nanchang University, Nanchang 330031, China

See next page for additional authors

Follow this and additional works at: <https://jepst.researchcommons.org/journal>

Recommended Citation

YANG, Jinxin; LIAO, Caibo; HU, Xiong; ZHU, Wenqing; ZHANG, Xu; and LIU, Bang (2024) "Transformer fault diagnosis based on DGA and TPE-LightGBM," *Journal of Electric Power Science and Technology*. Vol. 39: Iss. 4, Article 8.

DOI: 10.19781/j.issn.1673-9140.2024.04.008

Available at: <https://jepst.researchcommons.org/journal/vol39/iss4/8>

This Smart grid is brought to you for free and open access by Journal of Electric Power Science and Technology. It has been accepted for inclusion in Journal of Electric Power Science and Technology by an authorized editor of Journal of Electric Power Science and Technology. For more information, please contact dlxb04@163.com.

Transformer fault diagnosis based on DGA and TPE-LightGBM

Authors

Jinxin YANG, Caibo LIAO, Xiong HU, Wenqing ZHU, Xu ZHANG, and Bang LIU

引用格式:杨金鑫,廖才波,胡雄,等.基于 DGA 与 TPE-LightGBM 的变压器故障诊断[J].电力科学与技术学报,2024,39(4):70-77.

Citation: YANG Jinxin, LIAO Caibo, HU Xiong, et al. Transformer fault diagnosis based on DGA and TPE-LightGBM[J]. Journal of Electric Power Science and Technology, 2024, 39(4): 70-77.

基于 DGA 与 TPE-LightGBM 的变压器故障诊断

杨金鑫, 廖才波, 胡雄, 朱文清, 张旭, 刘邦

(南昌大学能源与电气工程系, 江西 南昌 330031)

摘要:油中溶解气体分析(dissolved gas analysis, DGA)对变压器故障的早期预警及诊断具有重要意义。为了提升变压器故障诊断的准确性及可靠性,提出一种基于树结构概率密度估计(tree-structured parzen estimator, TPE)算法优化轻量级梯度提升机(light gradient boosting machine, LightGBM)的变压器故障诊断方法。首先,建立包含油中气体比值、编码等 16 维 DGA 特征集合,采用最小绝对收缩和选择(least absolute shrinkage and selection operator, LASSO)算法选择用于变压器故障诊断的有效特征量;其次,构建基于 LightGBM 的变压器故障诊断方法,并引入 TPE 算法对 LightGBM 诊断模型参数进行优化,形成最优故障诊断模型;最后,选用精确度、召回率和 F_1 分数等评价指标对所提诊断模型性能进行评估。研究表明,TPE-LightGBM 的平均准确率为 90.23%,其诊断精度及鲁棒性均优于 RF 和 XGBoost 等算法。同时,与现场常用的三比值法进行对比,所提方法的准确性和可靠性均有显著提升。该方法可有效提升电力变压器的智能运维水平。

关键词:变压器;油中溶解气体;故障诊断;树结构概率密度估计;LASSO 算法;轻量级梯度提升机

DOI: 10.19781/j.issn.1673-9140.2024.04.008 **中图分类号:** TM721 **文章编号:** 1673-9140(2024)04-0070-08

Transformer fault diagnosis based on DGA and TPE-LightGBM

YANG Jinxin, LIAO Caibo, HU Xiong, ZHU Wenqing, ZHANG Xu, LIU Bang

(Department of Energy and Electrical Engineering, Nanchang University, Nanchang 330031, China)

Abstract: Dissolved gas analysis (DGA) is significant for early warning and diagnosis of transformer faults. To enhance the accuracy and reliability of transformer fault diagnosis, a transformer fault diagnosis method is proposed based on the tree-structured parzen estimator (TPE) algorithm to optimize the light gradient boosting machine (LightGBM). Firstly, a 16-dimensional DGA feature set including gas ratios and encodings in oil is established, and the least absolute shrinkage and selection operator (LASSO) algorithm is used to select effective feature quantities for transformer fault diagnosis. Secondly, a transformer fault diagnosis method based on LightGBM is constructed, and the TPE algorithm is introduced to optimize the parameters of the LightGBM diagnosis model, forming an optimal fault diagnosis model. Finally, evaluation metrics such as accuracy, recall, and F1 score are selected to assess the performance of the proposed diagnosis model. The research results indicate that the average accuracy of TPE-LightGBM is 90.23%, and its diagnostic accuracy and robustness are superior to algorithms such as RF and XGBoost. At the same time, compared with the commonly used three-ratio method in practice, the proposed method shows significantly improved accuracy and reliability. This method can effectively enhance the level of intelligent operation and maintenance of power transformers.

Key words: transformer; dissolved gas in oil; fault diagnosis; tree-structured parzen estimator; least absolute shrinkage and selection operator; light gradient boosting machine

油浸式变压器是输变电系统中关键的设备,其运行状态对国民经济和社会生产有重要影响^[1]。随

着设备运行年限增长,变压器将不可避免地发生绝缘损坏或老化等缺陷,导致出现大面积停电等重大

收稿日期:2023-08-25;修回日期:2024-01-15

基金项目:国家自然科学基金(62163025;52367001)

通信作者:廖才波(1989—),男,博士,副教授,主要从事输变电设备状态评估方面的研究;E-mail:Leb1752@126.com

事故^[2-3]。因此,及时准确地检测变压器早期内部故障,对维持电网的安全可靠运行具有重要意义^[4]。

油中溶解气体分析(dissolved gas analysis, DGA)技术在不影响设备正常运行的情况下,通过检测气体组分或体积分数辨识变压器内部故障,已成为电力行业公认的故障检测手段^[5]。由于运行设备会受到电场、温度及外部环境等因素干扰,变压器内部故障与特征气体间呈现出较为复杂的关系^[6]。对此,国内外学者提出了多种方法诊断变压器内部故障,如IEC三比值^[7]、Rogers比值^[8]和Duval三角法^[9]等方法,并在现场运用中发挥了重要作用。但上述传统方法存在编码缺失和过于依赖运维人员的现场经验,会出现较大的诊断偏差。

近年来,中国相继发布的《新一代人工智能发展规划》和《中国电力大数据发展白皮书》等重要文件,明确了电力设备的运行维护正朝着数字化和智能化方向发展^[10-12]。人工智能技术由于具有自适应性等特点,其在变压器状态检测及故障诊断方面得到了广泛应用,如卷积神经网络^[13]、支持向量机(support vector machine, SVM)^[14]、阶梯网络^[15]、随机森林(random forest, RF)^[16-18]、极端梯度提升树(extreme gradient boosting, XGBoost)^[19]和分类回归树(classification and regression tree, CART)^[20-22]等。文献[14]采用秃鹰搜索算法优化SVM的核函数参数,实现了变压器的故障诊断;文献[17]结合DGA气体含量,提出了贝叶斯优化随机森林算法的变压器故障诊断方法;文献[21]利用鲸鱼优化算法优化决策树集成模型,改善了变压器故障诊断效果。上述方法对变压器故障诊断的准确性有一定的提升,但以气体体积分数作为诊断模型输入,其分散性会降低模型的可靠性。同时,DGA特征信息的全面性和多样性方面可进一步考虑。以变压器故障诊断准确性为目标,将差异化特征量作为模型输入量,并选择合适的超参数,对提高诊断模型的准确性及可靠性具有重要意义。

为此,本文提出基于DGA和树结构概率密度估计(tree-structured parzen estimator, TPE)算法优化轻量级梯度提升机(light gradient boosting machine, LightGBM)的变压器故障诊断方法。为进一步提取油中溶解气体特征信息,该方法首先对气体进行特征衍生处理;然后,采用最小绝对收缩和选择(least absolute shrinkage and selection operator, LASSO)算法筛选出DGA有效特征量;最后,引入

TPE算法对LightGBM模型参数进行优化,形成基于TPE-LightGBM的变压器故障诊断模型,实现故障的准确诊断。算例表明,所提方法具有较好的诊断准确性和可靠性。

1 模型算法及原理

1.1 LASSO特征选择

Tibshiranni提出了LASSO算法,其主要通过加入L1正则化作为惩罚项,对模型系数加以约束^[23]。由于变压器色谱数据具有大量特征量,且多个特征存在高度相关性,LASSO算法倾向于选择与变压器故障相关性较强的特征,将不重要的特征量和其他高度相关特征量的系数稀疏为0,减少特征量的冗余信息,提高模型的稳定性。

假定变压器样本为 $X=[x_1, x_2, \dots, x_i, \dots, x_n]^T$, $x_i=[x_{i,1}, x_{i,2}, \dots, x_{i,m}]$,故障对应的目标编码为 $Y=[y_1, y_2, \dots, y_n]^T$,其中 n 为样本量, m 为特征维度。线性回归模型可以表示为

$$Y = X^T \beta + \epsilon \quad (1)$$

式中, $\beta=[\beta_1, \beta_2, \dots, \beta_m]^T$ 为特征量对应的回归系数; $\epsilon=[\epsilon_1, \epsilon_2, \dots, \epsilon_n]^T$ 为误差项。

基于线性回归模型的最小二乘法估计数学表达式为

$$\min_{\beta} \left[\frac{1}{2n} \sum_{i=1}^n \left(y_i - \sum_{j=1}^m x_{i,j} \beta_j \right)^2 \right] \quad (2)$$

当最小二乘法估计加入L1正则化约束时,得到LASSO算法的目标函数为

$$\min_{\beta} \left[\frac{1}{2n} \sum_{i=1}^n \left(y_i - \sum_{j=1}^m x_{i,j} \beta_j \right)^2 + \alpha \sum_{j=1}^m |\beta_j| \right] \quad (3)$$

式中, α 为正则化参数,其值越大,则更多 X 对应的 β 会变为0,从而稀疏相关性较弱的特征量。 \min_{β} 对回归系数 β 进行优化,使目标函数最小。

1.2 LightGBM算法

LightGBM模型是一种改进的梯度提升决策树(gradient boosting decision tree, GBDT)算法^[24]。相比梯度提升算法,LightGBM基于直方图和梯度提升算法能够有效地处理大规模的变压器数据集,同时不断迭代新的弱分类器并校正错误,进而提高整体模型的泛化性能。LightGBM对于异常值具有一定的鲁棒性,即使变压器色谱数据中存在一些异常值,仍然可以保持较好的模型诊断性能。

假设变压器训练样本集为 $D=\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$,LightGBM初始化的弱评估器为

$$f_0(x) = \arg \min_w \sum_{i=1}^n L(y_i, w) \quad (4)$$

式中, $L(y_i, w)$ 为损失函数; w 为拟合以决策树为弱评估器的输出值。

LightGBM 不断迭代新的弱评估器, 第 t 轮损失函数的负梯度估计值:

$$g_{it} = - \left[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)} \right]_{f(x)=f_{t-1}(x)} \quad (5)$$

利用 (x_i, g_{it}) 拟合第 t 个弱评估器, 每个叶节点对应的最佳拟合值为

$$w_{it} = \arg \min_w \sum_{x_i \in G_{ij}} L(y_i, f_{t-1}(x_i) + w) \quad (6)$$

其中, G_{ij} 为第 t 个决策树的叶节点区域。此时, 第 t 轮评估器对应的函数为

$$f_t(x) = f_{t-1}(x) + \sum_{j=1}^J w_{ij} I_j \quad (7)$$

式中, J 为第 t 个决策树的叶节点数; I_j 为全部样本在第 t 个决策树上输出的结果。训练结束后的模型表示为

$$f(x) = f_0(x) + \sum_{t=1}^T \sum_{j=1}^J w_{tj} I_j \quad (8)$$

LightGBM 在拟合弱评估器前主要采用梯度单边采样方法进行样本筛选, 并结合互斥特征捆绑方法减少特征维度。在拟合弱评估器过程中, LightGBM 采用直方图算法对样本数据进行离散化, 节省了存储空间。在决策树节点分裂过程中, LightGBM 通过按叶生长策略对具有最大分裂增益的叶子进行分裂, 从而减少叶节点的分裂次数, 提升了模型的运行效率。

1.3 TPE 算法

LightGBM 算法通过集成多个弱评估器大幅度提高模型的性能, 但存在模型参数众多和调参复杂等问题。为了找到合适的参数组合, 在给定的配置空间中本文对 LightGBM 模型参数进行寻优, 得到最优解:

$$u^* = \arg \min L(u) \quad (9)$$

式中, $L(u)$ 为损失函数; u 为配置空间中任意值。

贝叶斯优化算法是基于损失函数的前序采样结果建立的概率代理模型, 从而实现模型参数的自动寻优^[25]。Bergstra 等人^[26]提出一种基于贝叶斯优化变体的 TPE 算法, 采用如下公式:

$$p(u|v) = \begin{cases} l(u), & v < v' \\ h(u), & v \geq v' \end{cases} \quad (10)$$

式中, $p(u|v)$ 为参数 u 的先验概率分布; $l(u)$ 、 $h(u)$ 为参数 u 在损失函数中的 2 种概率密度估计; v 为采样值; v' 表示最优采样阈值。

TPE 优化算法采用了提升期望 (expected improvement, EI) 值来衡量新的超参数对应的损失值所能降低的幅度, 表达式为

$$E_I(u) = \int_{-\infty}^{v'} (v' - v) p(u|v) dv \quad (11)$$

令 $p(u) = \int_R p(u|v) p(v) dv$ 、 $\gamma = p(v < v')$, 其中 R 为实数, 式(11)可以简化为

$$E_I(u) \propto \left(\gamma + (1 - \gamma) \frac{h(u)}{l(u)} \right)^{-1} \quad (12)$$

为寻找最优解, 应逐步迭代 $h(u)/l(u)$ 逼近最小值, 使提升期望处于最大值, 此时返回模型对应的参数值。

1.4 模型应用流程

基于 DGA 与 TPE-LightGBM 的变压器故障诊断方法主要包括变压器样本数据预处理、诊断模型参数优化和故障诊断分析这三部分。数据预处理主要在高维特征集合中利用 LASSO 算法筛选差异化特征量; 模型优化主要采用 TPE 算法对 LightGBM 模型进行优化, 得出最优诊断模型; 故障诊断分析主要包括选用多个评价指标对模型诊断性能进行评估。本文所提诊断模型应用流程如图 1 所示, 具体步骤如下。

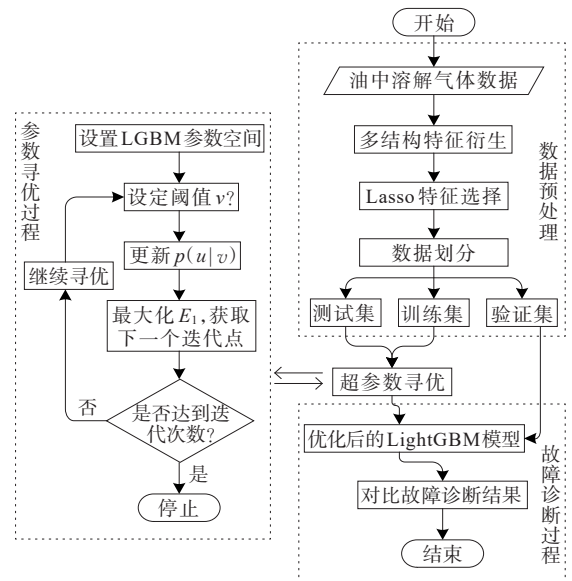


图 1 模型应用流程

Figure 1 Model application process

1) 收集已知变压器状态的在线或离线油中溶解气体数据, 提取原始特征量进行特征衍生处理,

形成高维的 DGA 特征集合;

2) 调整 LASSO 算法中正则化参数不断稀疏特征量系数,并对不同组合的特征量进行效果分析,选择出有效特征量;

3) 对 DGA 特征选择后的数据集进行样本划分,引入 TPE 算法对 LightGBM 诊断模型进行优化,并验证模型的泛化能力;

4) 对比多个评价指标下的诊断结果,对优化后的 LightGBM 模型性能进行评估,最后输出最优故障诊断模型。

2 DGA 特征量选择

2.1 故障类型和特征量

从文献[27]和 IEC TC 10 变压器故障数据库^[9]中选择 439 组已知变压器运行状态的样本作为 DGA 原始样本集。根据 IEC 60599^[9],将变压器故障诊断结果分为正常(N)、低温过热(T1, $T < 300\text{ }^\circ\text{C}$)、中温过热(T2, $300\text{ }^\circ\text{C} < T < 700\text{ }^\circ\text{C}$)、高温过热(T3, $T > 700\text{ }^\circ\text{C}$)、局部放电(PD)、低能放电(D1)和高能放电(D2),分别对其编码为 0~6。DGA 样本数据分布如表 1 所示。

表 1 DGA 样本数据分布
Table 1 Data distribution of DGA samples

运行状态	训练样本	测试样本	验证样本	样本总数
N	40	24	11	75
T1	33	7	3	43
T2	35	6	7	58
T3	45	27	15	87
PD	33	18	10	61
D1	30	15	7	52
D2	43	13	7	63
总计	259	120	60	439

在现场应用中,三比值方法主要采用 $\varphi(\text{C}_2\text{H}_2)/\varphi(\text{C}_2\text{H}_4)$ 、 $\varphi(\text{CH}_4)/\varphi(\text{H}_2)$ 和 $\varphi(\text{C}_2\text{H}_4)/\varphi(\text{C}_2\text{H}_6)$ 这 3 对比值特征量实现变压器故障诊断,其油色谱诊断准确率约为 82%^[28]。文献[29]通过总结变压器内部油中气体产生规律,提出了无编码比值诊断方法。对此,本文结合气体比值、三比值编码 T 和气体状态编码 C ,共扩展得到 16 维 DGA 特征量,如表 2 所示。

气体状态编码特征量的阈值判定规则主要参照文献[30]中的注意值,并结合现场应用需求对 $C(\varphi(\text{H}_2))$ 、 $C(\varphi(\text{C}_2\text{H}_2))$ 和 $C(\varphi(\text{TH}))$ 这 3 个特征量进行分箱操作,如表 3 所示。

表 2 DGA 特征量

Table 2 Characteristic quantity of DGA

序号	特征量	序号	特征量
S ₁	$T(\varphi(\text{C}_2\text{H}_2)/\varphi(\text{C}_2\text{H}_4))$	S ₉	$(\varphi(\text{H}_2)+\varphi(\text{C}_2\text{H}_2))/\varphi(\text{TG})$
S ₂	$T(\varphi(\text{CH}_4)/\varphi(\text{H}_2))$	S ₁₀	$(\varphi(\text{H}_2)+\varphi(\text{C}_2\text{H}_4))/\varphi(\text{TG})$
S ₃	$T(\varphi(\text{C}_2\text{H}_4)/\varphi(\text{C}_2\text{H}_6))$	S ₁₁	$(\varphi(\text{H}_2)+\varphi(\text{C}_2\text{H}_6))/\varphi(\text{TG})$
S ₄	$C(\varphi(\text{H}_2))$	S ₁₂	$(\varphi(\text{CH}_4)+\varphi(\text{C}_2\text{H}_2))/\varphi(\text{TH})$
S ₅	$C(\varphi(\text{C}_2\text{H}_2))$	S ₁₃	$(\varphi(\text{CH}_4)+\varphi(\text{C}_2\text{H}_6))/\varphi(\text{TH})$
S ₆	$C(\varphi(\text{TH}))$	S ₁₄	$(\varphi(\text{CH}_4)+\varphi(\text{C}_2\text{H}_4))/\varphi(\text{TH})$
S ₇	$\varphi(\text{H}_2)/\varphi(\text{TG})$	S ₁₅	$(\varphi(\text{C}_2\text{H}_4)+\varphi(\text{C}_2\text{H}_6))/\varphi(\text{TH})$
S ₈	$(\varphi(\text{H}_2)+\varphi(\text{CH}_4))/\varphi(\text{TG})$	S ₁₆	$(\varphi(\text{C}_2\text{H}_4)+\varphi(\text{C}_2\text{H}_2))/\varphi(\text{TH})$

注:TH 为油中气体中烃类气体;TG 为 H₂ 和烃类气体。

表 3 分箱操作规则

Table 3 Rules of binning operation

特征量	体积分数阈值/ 10^{-6}	编码
$C(\varphi(\text{H}_2))$	< 75	0
	$[75, 150)$	1
	> 150	2
$C(\varphi(\text{C}_2\text{H}_2))$	< 0.5	0
	$[0.5, 1)$	1
	> 1	2
$C(\varphi(\text{TH}))$	< 75	0
	$[75, 150)$	1
	> 150	2

2.2 LASSO 特征选择

基于 LASSO 算法的 DGA 特征量筛选标准主要通过调整正则化参数 α 确定与变压器故障相关性较强的特征量,并结合诊断效果对 DGA 特征量进行筛选。为得到最佳性能的正则化参数值,本文设定 α 按 $q=10$ 的比率从 10^{-5} 变化到 1,表 2 中 DGA 特征量对应特征系数的变化情况如图 2 所示。特征系数大于 0,表示特征量和故障编码呈正相关,小于 0 则呈负相关,等于 0 表示该特征量被稀疏。

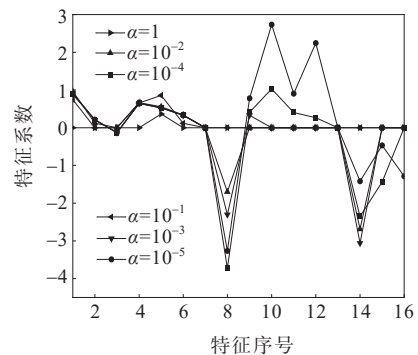


图 2 α 对特征系数的影响

Figure 2 Impact of α on characteristic coefficients

由图2结果表明, α 越大,LASSO算法特征量系数越稀疏。当 $\alpha=1$ 时,仅有特征量 S_5 对应系数不为0。将不同 α 对应的特征量组合作为LightGBM诊断模型的输入量,变压器色谱诊断结果如图3所示。当 $\alpha=10^{-3}$ 时,基于LightGBM模型的诊断准确率为85.2%,此时分类效果最佳。对此,本文选择 $\alpha=10^{-3}$ 对应的9个差异化特征量作为DGA有效特征量,如表4所示。

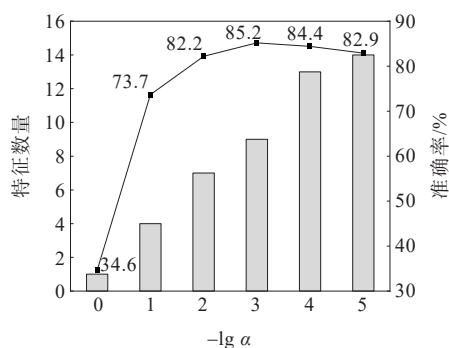


图3 不同 α 对应的诊断结果

Figure 3 Diagnostic results with different α

表4 DGA有效特征量

Table 4 Effective characteristic quantity of DGA

序号	特征量	序号	特征量
S_1	$T(\varphi(C_2H_2)/\varphi(C_2H_4))$	S_6	$C(\varphi(TH))$
S_2	$T(\varphi(CH_4)/\varphi(H_2))$	S_8	$(\varphi(H_2)+\varphi(CH_4))/\varphi(TG)$
S_3	$T(\varphi(C_2H_4)/\varphi(C_2H_6))$	S_9	$(\varphi(H_2)+\varphi(C_2H_2))/\varphi(TG)$
S_4	$C(\varphi(H_2))$	S_{14}	$(\varphi(CH_4)+\varphi(C_2H_4))/\varphi(TH)$
S_5	$C(\varphi(C_2H_2))$		

3 基于TPE-LightGBM的故障诊断

3.1 模型参数优化

在Inter(R) Core(TM) i7、1.8 GHz、内存16 GB和Python3.7的环境下,TPE算法LightGBM诊断模型的超参数进行寻优。LightGBM模型主要参数如表5所示;TPE算法优化过程如图4所示。

表5 LightGBM模型的主要参数

Table 5 Main parameters of LightGBM model

参数名称	参数意义	参数空间	步长
n_estimators	评估器数目	[10,300]	1.00
max_depth	树最大深度	[10,50]	1.00
learning_rate	学习率	[0.01,0.10]	0.01
max_leaves	最大叶节点数	[10,50]	1.00
reg_alpha	L1正则化参数	[0.01,0.50]	0.01
reg_lambda	L2正则化参数	[0.01,0.50]	0.01

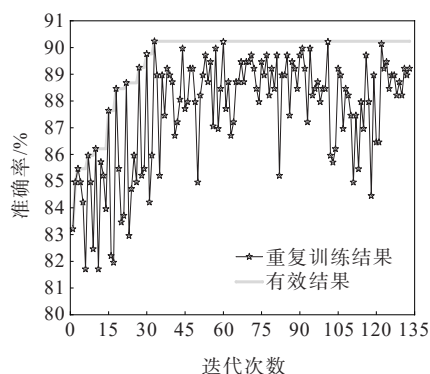


图4 TPE优化过程

Figure 4 Optimization process of TPE

由图4结果表明,在迭代至第33次时,TPE算法优化LightGBM诊断模型的平均准确率为90.23%,且随后连续100次未出现升高,模型优化过程结束。TPE算法优化后的LightGBM模型的评估器数量为240,树最大深度为8,模型学习率为0.06,最大叶节点数为25,L1和L2正则化参数值分别为0.36和0.20。

3.2 故障诊断

为分析TPE算法优化后的LightGBM模型的诊断效果,引入混淆矩阵对诊断结果进行可视化,如图5所示,结果表明,优化后的LightGBM诊断模型对测试集的准确率为90.8%,其中高温过热的准确率达到100%,局部放电和高能放电准确率均超过90%。同时,LightGBM模型在验证集的准确率达到93.3%,表明该诊断模型具有较好的泛化能力。

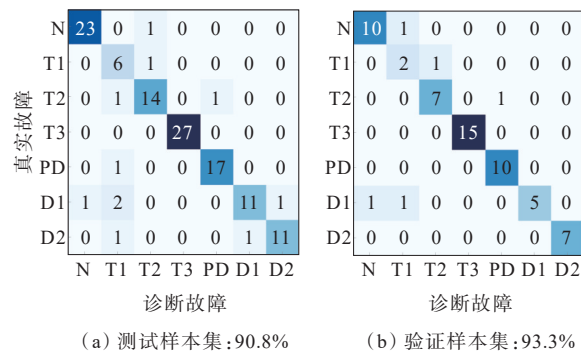


图5 TPE-LightGBM模型诊断结果

Figure 5 Diagnostic results of TPE-LightGBM model

为评估TPE-LightGBM模型对变压器故障的诊断效果,选取逻辑回归(logistic regression, LR)、CART、RF、GBDT、XGBoost这5种算法,采用 F_1 分数、召回率和精确度对模型性能进行评估,如图6所示,可知TPE-LightGBM模型的3种评价指标值均大于0.9,高于其他对比算法中的评价指标值,表明具有较好的故障识别能力。

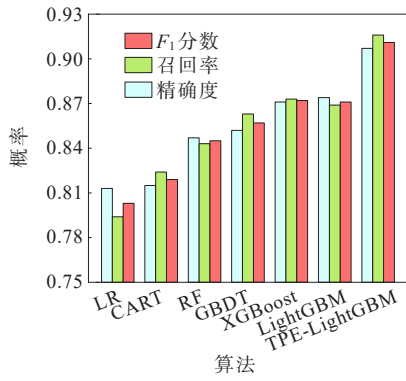


图 6 不同评价指标对比

Figure 6 Comparison of different evaluation indicators

为进一步确认 TPE-LightGBM 模型的对变压器故障诊断的准确性和稳定性,对表 1 中 439 组样本进行随机抽样,并进行多次重复训练,诊断结果如图 7 所示,小提琴是箱型图和密度图的结合体,中间的箱线盒表示下四分位到上四分位的数据,中心黑色条代表 1.5 倍四分位距的数据,中间白点表示中位数,外部轮廓表示核密度估计量。由图 7 可以看出,优化后的 LightGBM 诊断模型的平均诊断准确率高于其他机器学习对比模型,说明本文所提诊断方法具备较好的准确性和鲁棒性。

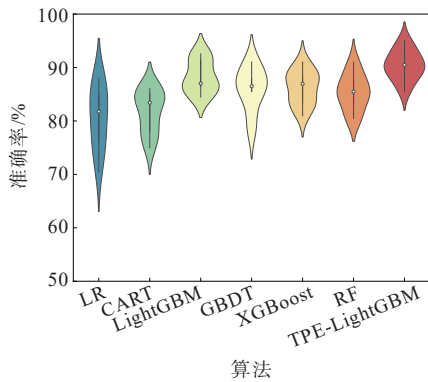


图 7 模型诊断效果对比

Figure 7 Comparison of model diagnosis effect

3.3 算例分析

为验证本文所述方法的应用效果,从 IEC TC 10 数据库和现场中共选取 10 组变压器油色谱数据进行故障诊断分析,如表 6 所示。

采用现场常用的三比值方法对表 6 中变压器样本进行故障诊断,并与优化后的 LightGBM 模型对比分析,结果如表 7 所示。结果表明,三比值法在诊断第 1、6 条色谱数据时,由于存在三比值编码缺失问题,难以判断具体的故障类型;在处理第 7 条色谱数据时,由于样本数据中出现了少量 C_2H_2 ,但未超

过注意值,不满足三比值的先验判定条件。因此,本文诊断方法能够有效弥补三比值法的准确性不高和编码缺失等问题,有效提升变压器故障诊断的准确性和可靠性。

表 6 变压器故障样本数据

Table 6 Sample data of transformer faults

序号	气体体积分数/ 10^{-6}				
	$\varphi(H_2)$	$\varphi(CH_4)$	$\varphi(C_2H_6)$	$\varphi(C_2H_4)$	$\varphi(C_2H_2)$
1	37.97	30.94	7.87	23.00	0.00
2	1450.00	940.00	211.00	332.00	61.00
3	360.00	610.00	259.00	260.00	9.00
4	107.00	143.00	34.00	222.00	2.00
5	46.00	180.70	49.00	325.60	0.70
6	157.80	3.20	0.60	0.70	0.00
7	29.53	15.00	0.08	0.53	0.44
8	60.00	10.00	4.00	4.00	4.00
9	60.00	5.00	2.00	21.00	21.00
10	13.00	3.00	1.00	3.00	6.00

表 7 故障样本诊断结果

Table 7 Diagnosis results of fault sample

序号	IEC 三比值法		LightGBM	真实故障
	编码	诊断结果		
1	001	无	T1	T1
2	101	D2	T2	T2
3	021	T2	T2	T2
4	022	T3	T3	T3
5	022	T3	T3	T3
6	011	无	PD	PD
7	—	正常	PD	PD
8	101	D2	D1	D1
9	112	D2	D2	D2
10	102	D2	D1	D2

4 结语

针对传统变压器故障诊断准确率低等问题,本文将 DGA 方法与人工智能技术相结合,建立了基于 TPE-LightGBM 的变压器故障诊断方法。

1) 构建了涵盖油中溶解气状态编码及比值等多维度的特征集合,采用 LASSO 算法选择得到表征变压器故障诊断的差异化特征量。

2) 采用 TPE 算法对 LightGBM 模型进行优化,可有效提升模型诊断效果,且优于其他机器学习算法模型,具有更好的稳定性和泛化性能;

3) 与 IEC 三比值诊断方法对比,本文方法能够实现变压器故障的准确诊断,弥补了传统诊断方法编码缺失和诊断精度不高等问题,具有一定理论研究和工程实际意义。

本文所提变压器故障诊断模型经算例验证,其诊断性能优异,但变压器故障特征信息可以进一步丰富,如可以考虑电气试验参量。此外,不同模型对不同故障类型的诊断精度不同,后续可以考虑多模型融合方法进一步提升模型诊断精度,并使用更多现场案例数据验证所提方法的适用性。

参考文献:

- [1] 周秀,怡恺,李刚,等.基于邻域粗糙集与 AMPOS-ELM 的变压器 DGA 故障诊断[J].电力科学与技术学报,2022,37(3):157-164.
ZHOU Xiu, YI Kai, LI Gang, et al. A transformer DGA fault diagnosis approach based on neighborhood rough set and AMPOS-ELM[J]. Journal of Electric Power Science and Technology, 2022, 37(3): 157-164.
- [2] 赵洪山,常杰英,曲岳哈,等.基于二元非线性 Wiener 随机过程的变压器油纸绝缘剩余寿命预测方法[J].电工技术学报,2023,38(15):4040-4049.
ZHAO Hongshan, CHANG Jieying, QU Yuhan, et al. Residual life prediction method of transformer oil-paper insulation based on binary nonlinear Wiener random process[J]. Transactions of China Electrotechnical Society, 2023, 38(15): 4040-4049.
- [3] 何宁辉,朱洪波,李秀广,等.基于贝叶斯网络和假设检验的变压器故障诊断[J].电力科学与技术学报,2021,36(6):20-27.
HE Ninghui, ZHU Hong-Bo, LI Xiuguang, et al. Transformer fault diagnosis based on Bayesian networks and hypothesis testing[J]. Journal of Electric Power Science and Technology, 2021, 36(6): 20-27.
- [4] 廖才波,杨金鑫,邱志斌,等.一种基于夏普利值及油中溶解气体分析的可解释变压器故障诊断方法[J].电网技术,2024,48(4):1752-1762.
LIAO Caibo, YANG Jinxin, QIU Zhibin, et al. An interpretable transformer fault diagnosis method based on SHAP value and dissolved gas analysis of transformer oil[J]. Power System Technology, 2024, 48(4): 1752-1762.
- [5] 朱庆东,朱文兵,王浩哲,等.基于油中溶解气监测的变压器在线半监督故障诊断方法研究[J].电网技术,2023,47(3):1031-1037.
ZHU Qingdong, ZHU Wenbing, WANG Haozhe, et al. Online semi-supervised fault diagnosis of transformer based on dissolved gas in oil[J]. Power System Technology, 2023, 47(3): 1031-1037.
- [6] 马鑫,尚毅梓,胡昊,等.基于数据特征增强和残差收缩网络的变压器故障识别方法[J].电力系统自动化,2022,46(3):175-183.
MA Xin, SHANG Yizi, HU Hao, et al. Identification method of transformer fault based on data feature enhancement and residual shrinkage network[J]. Automation of Electric Power Systems, 2022, 46(3): 175-183.
- [7] ROGRTS R R. IEEE and IEC codes to interpret incipient faults in transformers, using gas in oil analysis[J]. IEEE Transactions on Electrical Insulation, 1978, 13(5): 349-354.
- [8] DUVAL M. Dissolved gas analysis: it can save your transformer[J]. IEEE Electrical Insulation Magazine, 1989, 5(6): 22-27.
- [9] DUVAL M, DEPABLA A. Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases[J]. IEEE Electrical Insulation Magazine, 2001, 17(2): 31-41.
- [10] 盛戈峰,钱勇,罗林根,等.面向新型电力系统的数字化电力设备关键技术及其发展趋势[J].高电压技术,2023,49(5):1765-1778.
SHENG Geha, QIAN Yong, LUO Linggen, et al. Key technologies and development trends of digital power equipment for new type power system[J]. High Voltage Engineering, 2023, 49(5): 1765-1778.
- [11] 廖才波,杨金鑫,胡雄,等.数据与经验混合驱动下的变压器故障分层诊断方法[J].高电压技术,2023,49(5):1841-1850.
LIAO Caibo, YANG Jinxin, HU Xiong, et al. Hierarchical diagnosis method for transformer faults driven by mixed data and experience[J]. High Voltage Engineering, 2023, 49(5): 1841-1850.
- [12] 李佳迪,汤会增,相黎阳,等.基于 5G 与 AR 技术的高压 GIS 数字化运检平台设计与研究[J].高压电器,2023,59(12):83-93.
LI Jiadi, TANG Huizeng, XIANG Liyang, et al. Design and study on digital operation and inspection platform of UHV GIS based on 5G and AR technology[J]. High Voltage Apparatus, 2023, 59(12): 83-93.
- [13] 陆云才,廖才波,李群,等.基于声纹特征和集成学习的变压器缺陷诊断方法[J].电力工程技术,2023,42(5):46-55.
LU Yuncai, LIAO Caibo, LI Qun, et al. Transformer fault diagnosis method based on voiceprint feature and ensemble learning[J]. Electric Power Engineering Technology, 2023, 42(5): 46-55.
- [14] 周晓华,冯雨辰,陈磊,等.改进秃鹰搜索算法优化 SVM 的变压器故障诊断研究[J].电力系统保护与控制,2023,51(8):118-126.

- ZHOU Xiaohua, FENG Yuchen, Chen Lei, et al. Transformer fault diagnosis based on SVM optimized by the improved bald eagle search algorithm[J]. Power System Protection and Control,2023,51(8):118-126.
- [15] 郑尚直,仲林林,王同磊,等.基于改进半监督阶梯网络的有载分接开关故障诊断[J].电力工程技术,2023,42(2):197-205.
- ZHENG Shangzhi,ZHONG Linlin,WANG Tonglei,et al. Fault diagnosis method for OLTC based on improved semi-supervised ladder networks[J]. Electric Power Engineering Technology,2023,42(2):197-205.
- [16] HAQUE N, JAMSHED A, CHATTETJRR K, et al. Accurate sensing of power transformer faults from dissolved gas data using random forest classifier aided by data clustering method[J].IEEE Sensors Journal,2022, 22(6):5902-5910.
- [17] 王雪,韩韬.基于贝叶斯优化随机森林的变压器故障诊断[J].电测与仪表,2021,58(6):167-173.
- WANG Xue, HAN Tao. Transformer fault diagnosis based on Bayesian optimized random forest[J].Electrical Measurement & Instrumentation,2021,58(6):167-173.
- [18] 马红明,李倩,陶鹏,等.基于贝叶斯与数据驱动的智能电能表状态感知技术研究[J].电测与仪表,2022,59(2): 176-182.
- MA Hongming,LI Qian,TAO Peng,et al.Research on the state perception technology of smart meters based on Bayesianand data-driven[J]. Electrical Measurement & Instrumentation,2022,59(2):176-182.
- [19] 张又文,冯斌,陈页,等.基于遗传算法优化XGBoost的油浸式变压器故障诊断方法[J].电力自动化设备,2021, 41(2):200-206.
- ZHANG Youwen, FENG Bin, CHEN Ye, et al. Fault diagnosis method for oil-immersed transformer based on XGBoost optimized by genetic algorithm[J]. Electric Power Automation Equipment,2021,41(2):200-206.
- [20] 谢桦,陈俊星,赵宇明,等.基于SMOTE和决策树算法的电力变压器状态评估知识获取方法[J].电力自动化设备,2020,40(2): 137-142.
- XIE Hua, CHEN Junxing, ZHAO Yuming, et al. Knowledge acquisition method of power transformer condition assessment based on SMOTE and decision tree algorithm[J].Electric Power Automation Equipment, 2020,40(2):137-142.
- [21] 张国治,陈康,方荣行,等.基于DGA与鲸鱼算法优化 Logit Boost-决策树的变压器故障诊断方法[J].电力系 统保护与控制,2023,51(7):63-72.
- ZHANG Guozhi, CHEN Kang, FANG Rongxing, et al. Transformer fault diagnosis based on DGA and a whale algorithm optimizing a LogitBoost-decision tree[J]. Power System Protection and Control,2023,51(7):63-72.
- [22] 卢雪琴,李长安,吴忠强.基于鲸鱼算法优化极限学习机的微电网故障诊断[J].智慧电力,2022,50(2):15-21.
- LU Xueqin, LI Changan, WU Zhongqiang. Microgrid fault diagnosis based on extreme learning machine optimized by whale algorithm[J]. Smart Power, 2022, 50 (2):15-21.
- [23] TIBSHIRANI R. Regression shrinkage and selection via the lasso[J].Journal of the Royal Statistical Society,1996, 58(1):267-288.
- [24] MA X, SHA J, WANG D, et al. Study on a prediction of P2P network loan default based on the machine learning LightGBM and XGBoost algorithms according to different high dimensional data cleaning[J]. Electronic Commerce Research and Applications,2018,31:24-39.
- [25] 程路明,楼平,诸骏豪,等.基于APRIORI-贝叶斯优化 XGBoost的电力通信网告警预测[J].电力建设,2022, 43(1):113-121.
- CHENG Luming, LOU Ping, ZHU Junhao, et al. Root alarm prediction of power communication network applying APRIORI-Bayesian optimization XGBoost[J]. Electric Power Construction,2022,43(1):113-121.
- [26] BERGSTRA J, BARDENT R, BENGIO Y, et al. Algorithms for hyper-parameter optimization[C]//24th International Conference on Neural Information Processing Systems, Guangzhou, China,2017.
- [27] 尹金良.基于相关向量机的油浸式电力变压器故障诊断方法研究[D].北京:华北电力大学,2013.
- YIU Jinliang. Study on oil-immersed power transformer fault diagnosis based on relevance vector machine[D]. Beijing:North China Electric Power University,2013.
- [28] DUVAL M,ALZIEU E,BOCSI I,et al.Advances in DGA interpretation[R].Paris, France:CIGRE,2019.
- [29] 孙才新.电气设备油中气体在线监测与故障诊断技术 [M].北京:科学出版社,2003:78-82.
- SUN Caixin. On-line monitoring and fault diagnosis technology of gas in oil of electrical equipment[M]. Beijing: China Electric Power Press,2003:78-82.
- [30] DL/T 722—2014. 变压器油中溶解气体分析和判断导 则[S].
- DL/T 722—2014. Guide to the analysis and the diagnosis of gases dissolved in trans-former oil[S].