

基于特征优选与 IPSO-LSTM 的变压器故障诊断

胡俊泽¹, 杨耿煌^{1,2}, 耿丽清^{1,2}, 刘新宇¹

(1. 天津职业技术师范大学 自动化与电气工程学院, 天津 300222;
2. 天津职业技术师范大学 天津市信息传感与智能控制重点实验室, 天津 300222)

摘要:针对变压器故障诊断精度差、准确率低的问题,提出一种基于数据特征优选与改进粒子群优化算法的长短期记忆网络(IPSO-LSTM)的变压器故障诊断方法。首先对原始数据集进行预处理,使用合成少数类样本过采样技术(SMOTE)扩充数据数量;其次利用特征比值法扩充特征维数至20维,使用随机森林(RF)算法判断特征重要程度进行特征优选,降低过拟合风险;然后引入自适应惯性权重对PSO算法进行改进,利用改进后的PSO算法来优化LSTM最优超参数;最后输入特征优选后的数据进行变压器故障诊断。结果表明所构建的故障诊断模型诊断精度为91.6%。该优化模型与LSTM, HBA-LSTM 和 PSO-LSTM 诊断模型相比,准确率分别提高了10.12%, 5.95%, 3.57%, 证明IPSO-LSTM 诊断模型有更高的诊断准确率,在变压器故障诊断领域有一定的实际意义。

关键词:变压器故障诊断;特征优选;随机森林;长短期记忆网络;粒子群优化算法

中图分类号:TM411 文献标识码:A DOI:10.19457/j.1001-2095.dqcd26465

Fault Diagnosis of Transformers Based on Feature Selection and IPSO-LSTM

HU Junze¹, YANG Genghuang^{1,2}, GENG Liqing^{1,2}, LIU Xinyu¹

(1. School of Automation and Electrical Engineering, Tianjin University of Technology and Education, Tianjin 300222, China; 2. The Tianjin Key Laboratory of Information Sensing and Intelligent Control, Tianjin University of Technology and Education, Tianjin 300222, China)

Abstract: To address the issues of low accuracy and precision in transformer fault diagnosis, a transformer fault diagnosis method based on an improved particle swarm optimization-long short-term memory (IPSO-LSTM) network was proposed, integrating data feature selection and an improved particle swarm optimization algorithm. Firstly, the raw dataset was preprocessed using the synthetic minority oversampling technique (SMOTE) to increase the data volume. Secondly, the feature dimensions were expanded to 20 using the feature ratio method, and the random forest (RF) algorithm was employed to evaluate feature importance and perform feature selection, reducing the risk of overfitting. Subsequently, adaptive inertia weights were introduced to improve the PSO algorithm, which was then utilized to optimize the hyperparameters of the LSTM network. Finally, the feature-selected data was input into the model for transformer fault diagnosis. Results demonstrate that the proposed diagnostic model achieves an accuracy of 91.6%. Compared with LSTM, HBA-LSTM, and PSO-LSTM diagnostic models, the accuracy improves by 10.12%, 5.95%, and 3.57%, respectively, validating that the IPSO-LSTM diagnostic model provides superior diagnostic accuracy and holds practical significance in the field of transformer fault diagnosis.

Key words: transformer fault diagnosis; feature selection; random forest (RF); long short-term memory (LSTM) network; particle swarm optimization (PSO) algorithm

随着我国电网容量的不断增长,接入电网的发电设备和用电设备数量呈现持续增长趋势。

变压器作为电力系统中的重要设备,在输配电网中起着至关重要的作用。由于变压器成本高、

基金项目:天津市科技计划项目(23YDTPJC00320)

作者简介:胡俊泽(2001—),男,硕士,主要研究方向为控制工程,Email:2057916167@qq.com

通讯作者:杨耿煌(1978—),男,博士,教授,主要研究方向为电气工程及其自动化,Email:yenghuang@126.com

维修难度大,其故障一旦发生将造成严重影响。因此,及时准确地发现变压器的故障状态对于保障电网的稳定运行具有重要意义^[1-2]。

油浸式变压器的故障早期检测主要采用油中溶解气体分析(dissolved gas analysis, DGA)技术^[3]。该技术涵盖了多种分析方法,如国际电工委员会(international electrotechnical commission, IEC)提出的比值法、改良三比值法以及 Duval 三角法等^[4-5]。然而这类传统的油中溶解气体分析法存在编码不全和依赖专家经验等缺陷^[6],因此将人工智能技术与 DGA 方法相结合已成为当今研究的热点。文献[7]利用数据增强技术增加训练数据量来提升长短期记忆网络(long short-term memory, LSTM)模型的性能,结果证明准确率、查准率、查全率及 F1 值均有提高;另外与支持向量机(support vector machine, SVM)模型相比有更好的预测效果。然而,LSTM 模型自身存在全局搜索能力不足且容易陷入局部最优的问题,准确率在很大程度上受到数据集质量的影响。文献[8]借助少数类样本过采样(synthetic minority oversampling technique, SMOTE)算法扩充样本,并利用贝叶斯算法优化 LSTM 网络,进而进行变压器故障诊断,结果表明该方法模型使过拟合度降低了 20%,准确率提升了 10%,但优化效果依赖于概率模型的选择和先验知识的准确性。同时贝叶斯优化需要构建和更新概率模型,使得面临训练时间长且计算量大的问题。文献[9]利用随机森林(random forest, RF)算法对小型固定翼无人机故障数据进行特征选择,并构建基于 RF-LSTM 的故障诊断模型,实验表明相较于 LSTM 准确度提高 5.6%,但算法结构单一,稳定性差。文献[10]提出了一种粒子群优化(particle swarm optimization, PSO)算法与长短期记忆网络相结合的水电机组故障诊断方法,相较于标准的 LSTM 和 PSO-SVM 的故障诊断方法,该模型虽然提高了全局搜索能力和收敛速度,但其本身具有一定的复杂性,在处理大规模数据集时,存在时间长、效率低的问题。

综上所述,本文提出了一种基于特征优选与改进粒子群优化(improved particle swarm optimization, IPSO)算法的优化长短期记忆网络的变压器故障诊断模型。首先,对变压器故障数据进行预处理,去除内部缺失值和明显异常值,并利用 SMOTE 算法扩充其数据量,以及使用特征比值法

将样本特征扩展至 20 维;其次,使用随机森林算法按照特征重要程度进行排序,去除冗余特征,降低数据的复杂程度,进而选择最优特征数;再者,引入非线性自适应惯性权重改进 PSO 算法,利用改进后的 PSO 算法来优化 LSTM 最优超参数;最后,输入特征优选后的数据进行变压器故障诊断。结果表明,利用改进粒子群算法优化长短期记忆网络(improved particle swarm optimization-long short-term memory, IPSO-LSTM)的诊断模型在对变压器故障诊断中具有较高的准确率,且与粒子群算法和蜜獾算法(honey badger algorithm, HBA)相比,效果更佳。

1 算法原理

1.1 SMOTE 算法

由于变压器故障诊断数据集中故障样本数量有限,致使模型泛化能力不足,较易造成过拟合风险。为解决这一问题,本文采用 SMOTE 算法来生成相似样本,从而扩充数据集并进行辅助训练。SMOTE 算法是通过在样本及其邻近同类样本的连线上随机创建虚拟样本,以此来实现不同特征样本比例的平衡与扩展^[11]。其具体流程如下:

1)选择少数类样本:在少数类样本中选取样本 x_i 作为生成新样本的基准,以欧式距离为标准,计算该样本与其余故障样本的欧式距离。

2)近邻样本选择:根据样本不平衡比例设定采样倍率 N ,对于每个基准样本 x_i ,从其 k 个近邻中随机选择 N 个样本 x_{ij} 。

3)新样本合成:按照下式生成新样本:

$$X_{\text{new}} = x_i + r \times (x_{ij} - x_i) \quad (1)$$

式中: x_i 为少数类中的第 i 个样本; x_{ij} 为样本 x_i 的第 j 个近邻样本, $j=1, 2, \dots, k$; X_{new} 为在样本 x_{ij} 和 x_i 之间插值合成的新样本; r 为 $[0, 1]$ 之间的随机数。

1.2 RF 算法

RF 是一种集成学习(ensemble learning)方法,它通过集成多个决策树,可有效解决分类和回归问题。在 RF 模型中,使用 Bagging 方法抽取样本,其中约 2/3 的样本被用于训练模型,而未被用来训练的数据则被称为袋外数据(out-of-bag, OOB)。利用 OOB 数据,RF 模型能够评估各特征对分类精度的影响。并且,通过使用该数据计算出平均精度下降(mean decrease accuracy, MDA)指标^[12],进而对特征重要程度进行排序。数学公式如下:

$$MDA(X_k) = \frac{1}{N} \sum_{j=1}^N M A_j(X_k) \quad (2)$$

式中: N 为决策树数量。

通常情况下 MDA 值越大,代表其重要程度越高。

1.3 LSTM 算法

LSTM 的记忆神经元由内部状态、输入节点、输入门、遗忘门和输出门组成,用来控制信息流动与更新^[13]。基本结构图如图 1 所示。

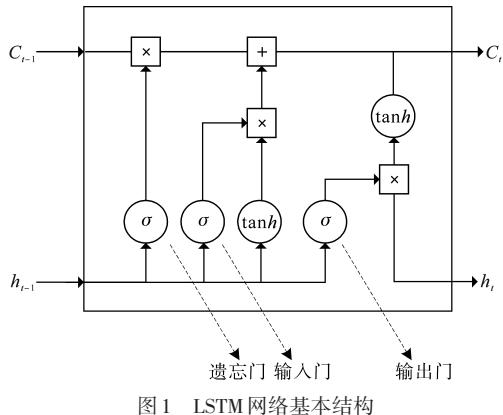


Fig.1 Basic structure of the LSTM network

步骤一:遗忘门通过 Sigmoid 函数对细胞状态中的信息进行删减和保留,将有用的信息长期记忆,公式如下:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (3)$$

式中: f_t 为 t 时刻遗忘门输出; σ 为 Sigmoid 激活函数; b_f 为偏置项。

步骤二:输入门分成 Sigmoid 层和 \tanh 层两部分,对将要写入细胞状态的信息进行预处理筛选,并生成新的候选值向量,以更新细胞状态。公式如下:

$$C'_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (4)$$

式中: C'_t 为 t 时刻的细胞状态。

步骤三:输出门负责从更新后的细胞状态中提取信息作为当前时间步的输出。公式如下:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (6)$$

式中: W 为权重矩阵; b 为偏置项; h_t 为 $t+1$ 时刻隐藏层的输出信号。

1.4 改进粒子群优化算法

粒子群优化算法是受到鸟群迁徙行为的启发而提出的一种群体智能算法。在粒子群算法中每个粒子的位置表示模型的超参数,速度表示控制超参数的更新幅度。粒子通过位置和速度的更新公式调整运动状态。更新速度和位置的公式如下:

$$v_i^{k+1} = \omega v_i^k + c_1 r_1 (pbest_i - x_i^k) + c_2 r_2 (gbest_i - x_i^k) \quad (7)$$

$$x_i^{k+1} = x_i^k + v_i^{k+1} \quad (8)$$

式中: $pbest_i$ 为个体最优; $gbest_i$ 为群体最优; ω 为惯性权重; c_1 为个体学习因子; c_2 为群体学习因子; r_1, r_2 为随机数。

传统粒子群算法处理高维数据时容易陷入局部收敛,出现误差大、精度低等问题。为解决上述问题,本文提出一种改进的动态自适应 PSO 算法。该算法对粒子群算法的惯性权重 ω 值进行改进,在迭代次数增加的过程中,通过非线性变化的形式予以调整。惯性权重 ω 是调整粒子搜索行为的关键参数,优化初期较大的 ω 值有利于全局探索,优化后期较小的 ω 值有利于局部搜索^[14-15]。

引入离散度 $k(n)$ 结合 Sigmoid 激活函数将惯性权重设为以下表达式:

$$\omega(t) = \omega_{\max} + (\omega_{\min} - \omega_{\max}) \frac{1}{1 + \exp[-10b(\frac{2t}{k(n) \cdot T} - 1)]} \quad (9)$$

式中: $\omega_{\max}, \omega_{\min}$ 为最大、最小惯性权重; T 为最大迭代次数; b 为阻尼因子,一般取 $[0, 1]$; $k(n)$ 为当前代与上一代适应度标准差的比值。

1.5 IPSO-LSTM 模型构建

LSTM 模型的参数包含常规参数和超参数两种类别。其中,常规参数由可训练的权重和偏置项组成,并通过 TensorFlow 内嵌的优化算法来实现参数调整^[16]。超参数是人工设置的参数,包括隐藏层大小,学习率和 L2 正则化参数,这些参数决定了模型的容量和训练时间,并对模型的性能有很大影响。为解决上述问题,本文采用 IPSO 算法对 LSTM 超参数进行寻优,并将优化后的超参数输入至 LSTM 中,构建 IPSO-LSTM 的故障诊断模型。主要步骤如下:

1) 设置 IPSO 算法参数,算法关键参数配置如下:粒子群规模设置为 20,迭代次数设为 50,加速因子为 2,惯性权重采用动态调整策略,取值范围设定为 $[0.4, 0.9]$,阻尼因子 b 取 0.5。

2) 确定算法的适应度函数,本文适应度函数设置为分类准确率。将预测结果与测试集的真实标签进行比较,得出分类准确率。

3) 计算每个粒子的适应度值,更新个体历史最优解 $pbest_i$ 和全局最优解 $gbest_i$,并同步更新对应的最优适应度值。

4) 进行寻优迭代,若达到最大迭代次数,结束循环并输出最优超参数。

5) 输入超参数,包括最优隐藏层单元数、最优学习率、最优 L2 正则化参数。

6) 设置 LSTM 模型,模型结构包括输入层、LSTM 层、激活层、全连接层、softmax 层和分类层,训练选项设置为使用 adam 优化器,并设置迭代次数。

IPSO-LSTM 变压器故障诊断模型流程图如图 2 所示。

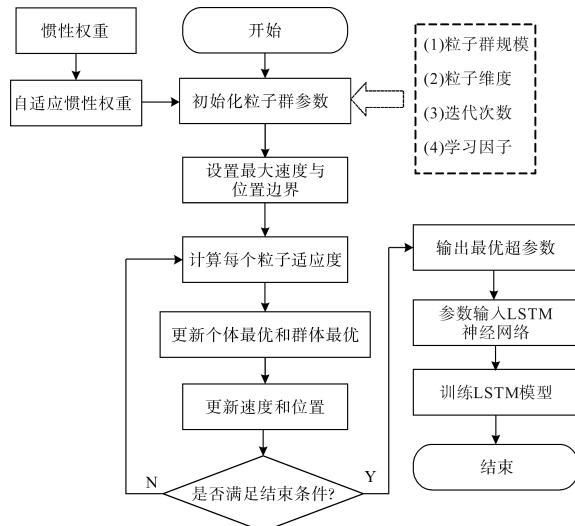


图 2 IPSO-LSTM 故障诊断流程图

Fig.2 IPSO flowchart of IPSO-LSTM fault diagnosis

2 变压器故障诊断具体实现过程

针对变压器故障诊断精度差、准确率低的问题,提出了基于 RF 特征优选的 IPSO-LSTM 变压器故障诊断方法如下:

1) 样本扩充:删除原始故障数据集中的缺失值,使用 SMOTE 算法根据式(1)合成新样本,实现对数据集样本数量的扩充。

2) 特征扩充:以 5 种基本的变压器油中溶解气体为基础,使用气体特征比值法将气体特征扩展,并生成 20 维待选特征集合。

3) RF 特征重要度排序:首先设置随机森林模型参数并训练,然后依据模型输出的特征重要性指标 MDA 值,对 20 维变压器特征按从高到低的顺序排列。

4) 特征优选:将 1 至 20 维变压器特征数据集输入到 LSTM 算法中,运行 10 次取平均诊断精度,选择平均诊断精度最高的特征维数作为故障诊断实验数据集。

5) IPSO 优化搜索:初始化粒子群,设置 IPSO 迭代次数、种群数量以及优化参数的取值范围,然后进行寻优迭代,寻找最优超参数。

6) 输入 IPSO-LSTM 模型:将 IPSO 模型所得的最优超参数输入至 LSTM 模型中,然后使用训练数据对 IPSO-LSTM 模型进行多轮迭代优化。最终,将测试集输入至 IPSO-LSTM 模型,对变压器故障类型进行识别。

流程图如图 3 所示。

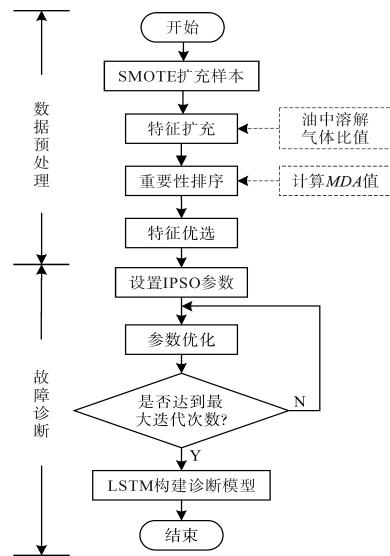


图 3 故障诊断方法流程图

Fig.3 Fault diagnosis flowchart

3 实验结果分析

3.1 故障类型和故障特征

研究数据来源于 500 kV 油浸式变压器故障记录,包含故障变压器绝缘油中溶解的各类气体(H_2 , CH_4 , C_2H_4 , C_2H_6 , C_2H_2);同时变压器的状态类别可分为 6 种,样本扩充表如表 1 所示,20 种气体特征量如表 2 所示。

表 1 样本扩充前后对照表

Tab.1 Comparison of samples before and after augmentation

状态类型	对应编号	扩充前	扩充后
中低温过热	1	113	140
高温过热	2	92	140
正常	3	140	140
局部放电	4	83	140
低能放电	5	45	140
高能放电	6	77	140

表 2 中,编号 6 至 20 是基于数据比值法扩充的特征。表中“总气”表示 5 种特征气体(H_2 , CH_4 , C_2H_4 , C_2H_6 , C_2H_2)的总和,而“总烃”则表示 4 种烃

表2 20种待选特征及其编号

Tab.2 20 Features to be selected with their numbers

编号	特征量	编号	特征量
1	H_2	11	$H_2/\text{总烃}$
2	CH_4	12	$CH_4/\text{总烃}$
3	C_2H_6	13	$C_2H_6/\text{总烃}$
4	C_2H_4	14	$C_2H_4/\text{总烃}$
5	C_2H_2	15	$C_2H_2/\text{总烃}$
6	CH_4/H_2	16	$H_2/\text{总气}$
7	C_2H_4/C_2H_6	17	$CH_4/\text{总气}$
8	C_2H_2/C_2H_4	18	$C_2H_2/\text{总气}$
9	C_2H_2/CH_4	19	$C_2H_4/\text{总气}$
10	C_2H_2/C_2H_6	20	$C_2H_4/\text{总气}$

类气体(CH_4 , C_2H_4 , C_2H_6 , C_2H_2)的总和。

3.2 特征优选结果分析

为避免预测受到不重要特征的干扰,使用RF算法对数据的20个特征进行优选,运行10次后取平均值,将重要度由高到低进行排序,如图4所示。

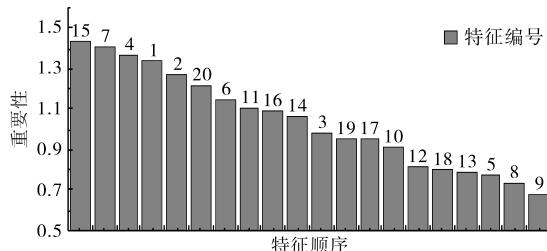


图4 重要度由高到低进行排序图

Fig.4 Feature importance sorted from high to low

由图4可知,各个特征的重要性存在差异,为消除冗余特征,需要对特征进行优选。按照重要性由高到低依次构建包含1至20个特征的数据集,分别使用LSTM算法、LSSVM算法、ELM算法和SVM算法进行预测。运行10次取平均诊断精度如图5所示。

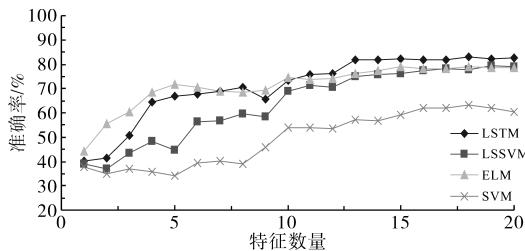


图5 基础模型平均精度图

Fig.5 Average accuracy of basic models

由图5可知,相比于SVM,LSSVM和ELM算法,基于LSTM算法的变压器故障诊断模型在平均精度上表现最佳。LSTM算法平均诊断精度和

最高诊断精度如图6所示。

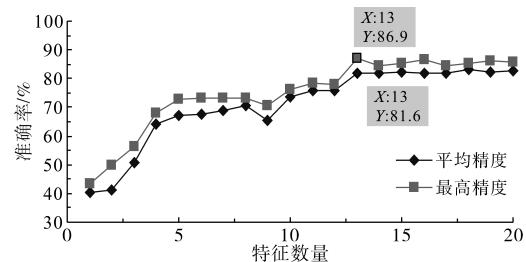


图6 LSTM模型平均精度图

Fig.6 Average accuracy of the LSTM model

由图6可知,特征个数从1到13增加的过程中,诊断精度逐步提高;当特征数增加到13个时,最高精度达到峰值为86.9%。特征个数从14增加到20的过程中,诊断精度逐渐趋于平稳,并且略低于使用13个特征时的准确率。分析上述实验结果:在特征数量有限的情况下,模型训练不充分,从而导致分类准确率偏低;而当特征数量过多时,模型的复杂度会增加,引起过拟合现象,进而使得分类准确率不增反降。基于图6的实验结果,选择排序后的前13维特征作为模型训练和诊断的数据集。

3.3 基于IPSO-LSTM的变压器故障诊断结果分析

将特征优选后的13维数据集带入到IPSO-LSTM模型中进行故障诊断。诊断结果如图7和图8所示。

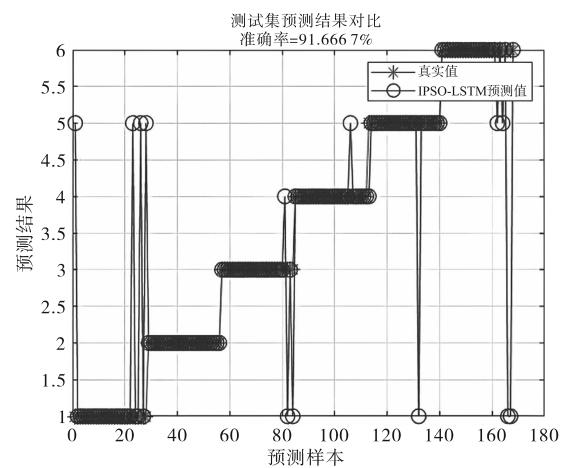


图7 故障诊断预测结果

Fig.7 Fault diagnosis prediction results

结合图7和图8可知,测试数据共168组,IPSO-LSTM正确分类154组,错误分类14组,总体故障诊断准确率为91.66%。其中高温过热和低能放电准确率为100%,高能放电的错误分类最多,准确率只有78.6%。分析其原因在于,故障特征存在区分度低且重叠的问题,低能放电与局部

Confusion Matrix for Test Data						
真实类 1 2 3 4 5 6	24			4		
	28					85.7% 14.3%
		25	1			100.0%
			27	1		89.3% 10.7%
				28		96.4% 3.6%
					22	100.0%
	3		1	2	22	78.6% 21.4%
82.8% 100.0% 100.0% 93.1% 80.0% 100.0%						
17.2% 2 3 4 5 6						
预测类						

图8 IPSO-LSTM 故障诊断分类结果

Fig.8 Fault diagnosis classification results based on IPSO-LSTM model

放电会产生与高能放电相同的气体特征,导致难以通过气体浓度准确区分高能放电故障。为了进一步评估 IPSO-LSTM 模型性能,对于故障诊断结果引入了准确率(accuracy)、精确率(precision)、召回率(recall)、F1_score 和 kappa 系数进行进一步的描述。根据表3所示的各项指标结果,可以得出该模型分类能力强,在故障诊断方面有较高的准确性与可靠性。

表3 IPSO-LSTM 算法性能指标

Tab.3 Performance metrics of the IPSO-LSTM algorithm

指标	精确率	准确率	召回率	F1_score	kappa 系数
实际值	0.9265	0.9167	0.9244	0.9181	0.9

单次变压器故障诊断模型性能的评估具有偶然性与随机性。为了进一步说明该诊断模型的优点,本文使用 IPSO 算法对不同方法处理的数据集进行训练对比,并与不同的优化算法对比分析。

3.4 不同处理方法对比分析

为验证基于随机森林特征优选方法的有效性,本文采用原始5维特征、20维全特征、RF特征优选的13维特征、SMOTE扩充后5维特征、SMOTE扩充后20维全特征和本文方法,这6种特征处理方式进行对比验证。诊断模型均选用 IPSO-LSTM 诊断模型,实验参数设置与上节保持一致,连续运行10次后求得平均值。实验结果如表4所示。

由实验结果得出,通过SMOTE数据扩充以及增加特征维度并采用特征优选的数据处理方法,可以有效提升变压器故障诊断模型的准确率。

表4 不同特征诊断准确率

Tab.4 Diagnostic accuracy for different features

标号	特征选取	准确率/%
1	原始5维特征	77.27
2	20维全特征	82.72
3	RF特征优选(13维)	85.45
4	SMOTE扩充后5维特征	83.92
5	SMOTE扩充后20维全特征	86.01
6	本文方法	90.11

3.5 不同优化算法的对比分析

为进一步验证 IPSO-LSTM 故障诊断模型的优秀性能,将经本文方法处理后的样本数据输入到 LSTM, HBA-LSTM 和 PSO-LSTM 模型中进行训练。为了确保实验中变量控制的一致性,实验参数设置与3.2节保持一致。分类结果如图9~图11所示,对比结果如表5所示。

Confusion Matrix for Test Data						
真实类 1 2 3 4 5 6	24	1	1		5	
	26			2		75.0% 25.0%
	2	1	24	1		92.9% 7.1%
	2			25		85.7% 14.3%
			1	3	21	89.3% 10.7%
		2		1	5	75.0% 25.0%
				20		71.4% 28.6%
84.0% 86.7% 92.3% 78.1% 67.7% 83.3%						
16.0% 13.3% 7.7% 21.9% 32.3% 16.7%						
1 2 3 4 5 6						
预测类						

图9 LSTM 分类结果

Fig.9 LSTM classification results

Confusion Matrix for Test Data						
真实类 1 2 3 4 5 6	21		1	6		
	1	27				75.0% 25.0%
	2	1	24	1		96.4% 3.6%
				26		85.7% 14.3%
			1	27		92.9% 7.1%
			1	8	19	96.4% 3.6%
						67.9% 32.1%
87.5% 96.4% 100.0% 86.7% 62.8% 100.0%						
12.5% 3.6% 0 13.3% 37.2% 0						
1 2 3 4 5 6						
预测类						

图10 HBA-LSTM 分类结果

Fig.10 HBA-LSTM classification results

在4种算法模型中,LSTM准确率最低,引入优化算法后HBA-LSTM和PSO-LSTM模型的准确率有所提升。这些模型中IPSO-LSTM模型表现

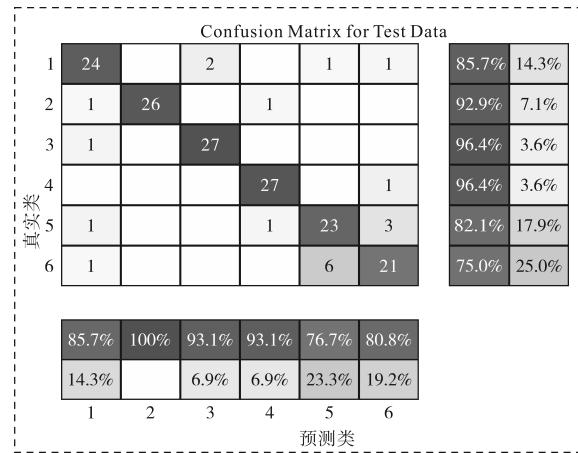


图 11 PSO-LSTM 分类结果

Fig.11 PSO-LSTM classification results

表 5 不同算法诊断准确率

Tab.5 Diagnostic accuracy of different algorithms

状态类型	故障诊断准确率/%			
	LSTM	HBA-LSTM	PSO-LSTM	IPSO-LSTM
中低温过热	75	75	85.71	85.71
高温过热	92.85	96.42	92.85	100
正常	85.71	85.71	96.42	89.28
局部放电	89.28	92.85	96.42	96.42
低能放电	75	96.42	82.14	100
高能放电	71.42	67.85	75	78.57
综合精度	81.54	85.71	88.09	91.66

最佳,相比于PSO-LSTM,HBA-LSTM,LSTM算法,IPSO-LSTM模型诊断精度分别提高了3.57%,5.95%,10.12%。因此,IPSO-LSTM模型有较强的分类能力。

3.6 不同数据集分析

单一数据集可能存在局限性,为进一步验证模型的有效性,本文采用110 kV油浸式变压器故障数据集进一步分析验证模型有效性。

110 kV变压器数据集状态类型分为5种,分别为中低温过热、高温过热、低能放电、高能放电和正常,原始数据集故障数据453组,使用本文方法,首先采用SMOTE算法扩充至585组,其次使用比值法将特征维数扩展至20维,再次利用RF算法进行特征优选,优选后的特征维数为12维,最后输入至IPSO-LSTM算法分析结果。诊断结果如图12、图13所示。

结合图12和图13可知,测试数据175组,正确分类156组,错误分类19组,总体故障诊断准确率为89.1%。由此可见,该故障诊断模型具有较强的泛化能力,适用于不同的变压器故障数据集。

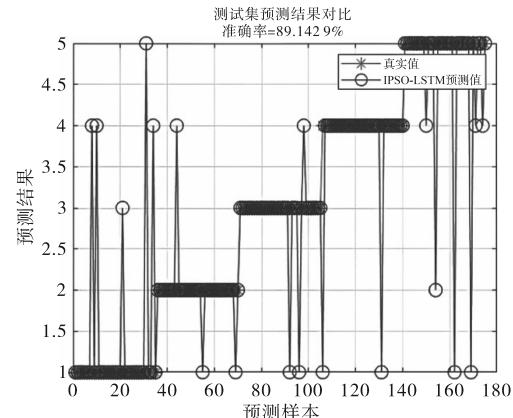


图 12 IPSO-LSTM 故障诊断结果

Fig.12 The IPSO-LSTM fault diagnosis results

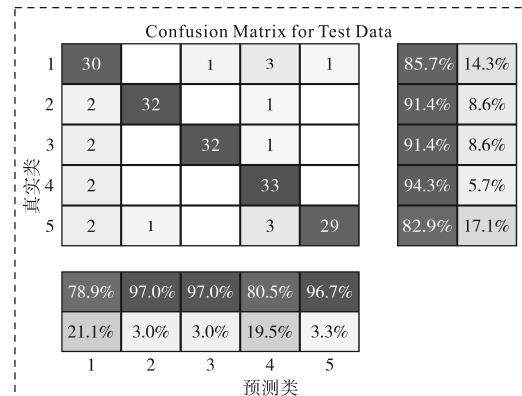


图 13 故障诊断分类结果

Fig.13 Fault diagnosis classification results

4 结论

本文提出一种基于特征优选的IPSO-LSTM变压器故障诊断方法,经过实验分析,得出结论如下:

1) 使用SMOTE算法扩充样本数量,通过特征比值法扩充至20维特征,利用RF算法优选前13维特征,去除了冗余特征。相较于原始数据集(5种DGA气体),本文所提方法准确率提高了14.4%。

2) 本文所提出的惯性权重优化方法有效平衡了算法全局探索能力与局部开发能力。经过IPSO优化后的LSTM模型故障诊断准确率可由寻优前的81.54%进一步提升至91.66%,相比于未改进的PSO-LSTM以及常用的HBA-LSTM准确率提升了3.57%和5.95%,同时,模型在不同数据集中的诊断精度达到89.1%,表现出了较强的泛化能力。

因此本文所提出的基于特征优选的IPSO-LSTM变压器故障诊断模型,分类精度高,实用性强,具有一定的理论意义与实际价值。

参考文献

- [1] 王巍津,任本东. 电力变压器故障检测技术的现状和发展趋势分析[J]. 现代工业经济和信息化,2021,11(8):176-178.
WANG Weijin, REN Bendong. Analysis of the current status and development trend of fault detection technology of power transformer[J]. Modern Industrial Economy and Informationization, 2021, 11(8): 176-178.
- [2] 林凡勤,李明,郭红. 变压器故障诊断技术综述[J]. 计算机与现代化,2022 (3):116-126.
LIN Fanqin, LI Mingming, GUO Hong. Review on fault diagnosis technology of transformer[J]. Computer and Modernization, 2022 (3): 116-126.
- [3] 张鹏,齐波,李成榕,等. 电力变压器油中溶解气体特性影响因素的量化分析[J]. 中国电机工程学报,2021,41(10):3620-3631,3686.
ZHANG Peng, QI Bo, LI Chengrong, et al. Quantitative analysis of influence factors of dissolved gas characteristics in power transformer oil[J]. Proceedings of the CSEE, 2021, 41(10): 3620-3631, 3686.
- [4] 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.
- [5] DUVA M. A review of faults detectable by gas-in-oil analysis in transformers[J]. IEEE Electrical Insulation Magazine, 2002, 18 (3):8-17.
- [6] 汪繁荣,李州. 基于SCSSA-BiLSTM的变压器故障诊断模型[J/OL]. 南方电网技术:(2024-07-31)[2025-02-20]. <http://kns.cnki.net/kcms/detail/44.1643.tk.20240730.1605.013.html>.
WANG Fanrong, LI Zhou. Transformer fault diagnosis model based on SCSSA-BiLSTM[J/OL]. Southern Power System Technology: (2024-07-31) [2025-02-20]. <http://kns.cnki.net/kcms/detail/44.1643.tk.20240730.1605.013.html>.
- [7] 蔡晨. 基于数据增强技术的LSTM模型变压器故障诊断研究[J]. 红水河,2024,43 (1):102-105,111.
CAI Chen. Transformer fault diagnosis of LSTM model based on data enhancement technology[J]. Hongshui River, 2024, 43 (1):102-105,111.
- [8] 张宏杰,陈贵凤,闫宏伟,等. 基于SMOTE与Bayes优化的LSTM网络变压器故障诊断[J]. 中国电力,2023,56 (10): 164-170.
ZHANG Hongjie, CHEN Guifeng, YAN Hongwei, et al. Fault diagnosis of LSTM network transformer based on SMOTE and Bayes optimization[J]. Electric Power, 2023, 56 (10) : 164-170.
- [9] 孙菱,张振宇,郭健,等. 基于RF-LSTM模型的无人机执行器故障诊断算法[J]. 南京理工大学学报,2024,48 (5):661-668.
SUN Ling, ZHANG Zhenyu, GUO Jian, et al. Fault diagnosis algorithm for UAV actuator based on RF-LSTM model[J]. Journal of Nanjing University of Science and Technology, 2024, 48 (5) : 661-668.
- [10] 罗玮,陈媛. 基于PSO-LSTM网络的水电机组振动故障诊断方法[J]. 机械设计与制造工程,2024,53 (9):94-98.
LUO Wei, CHEN Yuan. Vibration fault diagnosis method for hydroelectric units based on PSO-LSTM network[J]. Machine Design and Manufacturing Engineering, 2024, 53 (9): 94-98.
- [11] 邹德旭,徐赫,权浩,等. 基于DGA和改进SMOTE的小样本变压器故障诊断方法[J]. 广东电力,2024,37 (7):118-126.
ZOU Dexu, XU He, QUAN Hao, et al. Small sample transformer fault diagnosis method based on DGA and enhanced SMOTE [J]. Guangdong Electric Power, 2024, 37 (7):118-126.
- [12] 胡乐怡,付晓雷,蒋晓蕾,等. 基于LSTM、RF、SVR三种机器学习方法的径流预测研究[J]. 水文,2024,44 (5):17-24.
HU Leyi, FU Xiaolei, JIANG Xiaolei, et al. Research on runoff prediction based on machine learning: LSTM, RF and SVR[J]. Journal of China Hydrology, 2024, 44 (5):17-24.
- [13] 郝泽嘉,施玉群,成博超,等. 基于PSO-LSTM的大坝变形组合预测模型[J]. 长江科学院院报,2025,42 (5) : 208-214, 222.
HAO Zejia, SHI Yuqun, CHENG Bochao, et al. Combined prediction model of dam deformation based on PSO-LSTM[J]. Journal of Changjiang River Scientific Research Institute, 2025, 42 (5) : 208-214, 222.
- [14] 乔建刚,李硕,刘怡美,等. 改进PSO-LSTM算法预测高速公路交通量[J]. 科学技术与工程,2024,24(15):6466-6472.
QIAO Jiangang, LI Shuo, LIU Yimei, et al. Improved PSO-LSTM algorithm for forecasting expressway traffic volume[J]. Science Technology and Engineering, 2024, 24 (15) : 6466-6472.
- [15] 王生亮,刘根友. 一种非线性动态自适应惯性权重PSO算法[J]. 计算机仿真,2021,38 (4):249-253,451.
WANG Shengliang, LIU Genyou. A nonlinear dynamic adaptive inertial weight particle swarm optimization[J]. Computer Simulation, 2021, 38 (4):249-253,451.
- [16] 郭明辰,张润润,闻余华. 基于PSO-LSTM模型的平原河网汛期水位预测[J]. 水利水电科技进展,2024,44(6):64-70.
GUO Mingchen, ZHANG Runrun, WEN Yuhua. Prediction of water level of plain river network during flood season based on PSO-LSTM model[J]. Advances in Science and Technology of Water Resources, 2024, 44(6):64-70.

收稿日期:2025-02-20

修改稿日期:2025-04-15