

基于改进型主动深度学习框架的 油浸式变压器故障辨识算法

黄灿, 仝杰, 唐鹏飞, 齐子豪, 张中浩, 龙天航
(中国电力科学研究院有限公司, 北京市 海淀区 100192)

An Algorithm of Oil-immersed Transformer Fault Classification Based on Improved Active Deep Learning Framework

HUANG Can, TONG Jie, TANG Pengfei, QI Zihao, ZHANG Zhonghao, LONG Tianhang
(China Electric Power Research Institute, Haidian District, Beijing 100192, China)

ABSTRACT: As the central node of the power system, the transformer has a critical impact on the safe operation of the power system. However, in the field of transformer condition assessment and fault investigation and identification, there are problems such as the difficulty of fault sampling and unbalanced data sets, which make the existing diagnosis methods very prone to false positives and false negatives, and their application in the practical operation and maintenance scenarios of transformers is limited. Therefore, this paper proposes a transformer fault diagnosis algorithm that combines improved active learning and deep neural networks. Four active learning strategies adapted with deep neural networks are proposed to enable deep neural networks to learn from oil chromatography data samples with directional selection to improve the learning performance of fault detection algorithms in small and unbalanced samples and to reduce data annotation and model training costs; Secondly, the label smoothing algorithm is introduced to improve the cross-loss function of deep neural networks to solve the overfitting problem caused by the inconsistency between the distribution of the training sets and the prediction sets after active learning sampling labels. The simulation experiment results show that the active learning algorithm proposed in this paper can effectively reduce the model training cost and improve the fault identification accuracy by more than 10% compared with traditional passive sampling learning methods such as random sampling, and improve the performance of the algorithm by more than 15% compared with common methods such as support vector machine.

KEY WORDS: transformer fault identification; active learning; deep learning; small sample learning; unbalanced dataset

摘要: 变压器作为电网的核心枢纽节点,对电网安全运行有着至关重要的影响作用。然而,在变压器状态评估及故障研判中,存在故障样本收集困难,数据集不平衡等问题。因此,现有诊断方法在实际应用时极易出现误报、漏报等情况。为此,该文提出了一种改进型主动学习-深度神经网络融合的变压器故障诊断算法。首先提出了4种与深度神经网络适配的主动学习策略,对油色谱数据样本进行定向选择学习,提高故障辨识算法在小样本及不均衡样本下的学习表现,降低数据标注及模型训练成本;其次,引入标签平滑算法对深度神经网络交叉损失函数进行改良,以改善训练集与预测集分布不一致导致的过拟合问题。仿真实验结果表明,所提出的主动学习算法与随机采样等传统的被动采样学习方法相比,能够有效降低模型训练成本,并提升故障辨识精度10%以上;与支持向量机等常见方法相比,算法性能提高15%以上。

关键词: 变压器故障辨识; 主动学习; 深度学习; 小样本学习; 非均衡数据集

DOI: 10.13335/j.1000-3673.pst.2024.0562

0 引言

变压器作为电压变换和电能传输的关键电力设备,其可靠稳定运行是电网安全稳定的基础^[1-2]。然而近年来,我国油浸式变压器相继发生严重的短路故障,多次引发火灾事故,给社会造成了严重经济损失和不良影响^[3],设备安全运行面临严峻挑战。因此,如何准确判断变压器工作状态,提高变压器设备供电可靠性,是电力领域中一个重要研究方向。

在不同的运行状态下,油浸式变压器中的绝缘油因电、热作用会产生不同含量和组分的气体。因此,油中溶解气体分析(dissolved gas analysis, DGA)法^[4]对于判别变压器运行状态具有重要参考价值。针对DGA故障判断原则,CIGRE^[5]、IEC

基金项目: 国家自然科学基金项目(U23B20135)。

Project Supported by the National Natural Science Foundation of China (U23B20135).

60599^[6]、DL/T 722—2014^[7]等标准中都进行了相应的归纳总结，但其判断标准大多较为简单，在实际复杂的变压器运行情况下，其诊断结果往往与实际偏差较大。近年来，以机器学习和深度学习为代表的人工智能模型大量涌现，因其强大的非线性拟合能力，得到多方专家学者的应用。文献[8]将相空间重构运用到油色谱数据预处理中，并选择其最优长度进行模型训练，减小了分类器算法计算成本。文献[9]提出了一种通过自决策主动纠偏方法，完成了在电力变压器油色谱数值误差修正过程和故障类别差异修正过程间进行自动决策与切换。文献[10]提出了基于方差分析和黑猩猩算法优化的支持向量机(bonobo optimizer-support vector machine, BO-SVM)模型，通过对模型输入进行筛选降维与支持向量机参数寻优提高了变压器故障模型辨识速度与精度。然而，多数人工智能算法本质上是数据驱动技术，其准确性依赖于训练数据集的全面性与均衡性。但变压器缺陷和故障属于小概率事件，及时收集缺陷及故障样本难度很大。同时，不同类型故障样本发生的频率不同，样本数量差异较大，使人工智能模型的应用效果受到严重制约和影响^[11]。

针对上述问题，国内外学者已进行了研究，可分为数据增强方法和分类器改造方法。数据增强方法即通过数据生成技术扩充少数类样本数据从而达到数据集平衡。刘云鹏等^[12]建立了梯度惩罚优化的条件式 Wasserstein 生成对抗网络模型，以指导油色谱多类别故障样本的生成过程，克服了原始生成对抗网络模型的训练不稳定问题。李雅欣等^[13]将策略梯度与 GAN 相结合扩充油色谱数据，提高了分类准确率。此外，运用过采样算法(synthetic minority oversampling technique, SMOTE)^[14-15]扩充的油色谱数据也为变压器故障诊断取得了良好的效果。分类器改造方法则通过改进与调整分类器结构，使其在数据集不平衡情况下仍能保持较高精度，常用算法有：基于基分类器优化、集成学习、多类分解技术^[16-18]等。然而，上述方法皆存在较为明显缺陷：数据增强方法生成的数据质量难有保障，易导致分类器混乱，从而降低准确率；分类器改造方法易造成模型复杂度上升，导致运算成本增加。

为提高人工智能算法在变压器故障诊断问题中数据量少且数据不平衡情况下的学习能力，本文基于主动学习(active learning, AL)思想，提出一种融合改进型主动学习-深度神经网络的变压器故障辨识模型(improved active learning-deep neural network transformer fault diagnosis algorithm,

IAL-DNN-TFDA)。首先，利用主动学习查询策略对油色谱数据进行目的性采样学习，从而提高神经网络在数据不平衡情况下的学习效率；其次，针对神经网络主动学习指向性采样标记后所产生的训练过拟合问题，对神经网络损失函数使用标签平滑算法(label smoothing, LS)进行改良，构建 LS 交叉熵函数，实现过拟合现象的缓解；最后，本文使用所提出算法与被动采样策略的深度神经网络算法以及传统变压器故障辨识模型和算法进行对比，结果表明，在面对小样本、不均衡的油色谱数据样本时，本文所提出的方法仍能在有限数据的条件下提高神经网络的准确性与泛化性。

1 算法框架

针对油浸式变压器故障辨识问题，本文提出了 4 种主动学习查询策略以实现主动学习算法与深度学习的适配，并对深度学习分类模型损失函数进行改进，构建基于改进型主动深度学习架构的油浸式变压器故障辨识算法框架，其结构如图 1 所示。该框架可总结为以下流程：

1) 标注初始集训练。从油色谱数据集内随机选择少量数据样本进行标记，构成标注初始集，并将初始集合输入神经网络进行训练。为防止算法因针对性采样训练产生的过拟合现象，使用 LS 交叉熵损失函数完成神经网络反向传播过程。

2) 查询策略运用。将未标注数据池中数据输入训练好的神经网络，获得其预测为各类别概率 P ，之后利用 P 配合 4 类查询策略函数对未标记数据进行评分，分数表示为 r ， r 越高则优先级越高，随后将数据 x 按照优先级进行排序。

3) 训练集更新。确定每轮选取样本数 f ，将排序后优先级高的前 f 组样本数据移交至专家/人工进行再次确认审核并标记。标记完成后，将数据从数据池中移除，并加入标注数据集中。

4) 神经网络再训练。将更新后的数据集输入变压器故障辨识模型，使其重新进行训练。后重复第 2、3、4 步，直至达到所设定的目标数据标记量。

5) 故障辨识。在标记训练完成后，将测试集输入训练完成的深度神经网络，计算故障判别结果。

2 改进型主动学习深度神经网络融合框架

2.1 基于不确定性分析的主动学习方法

2.1.1 主动学习

主动学习是一种数据采样学习技术手段，其关

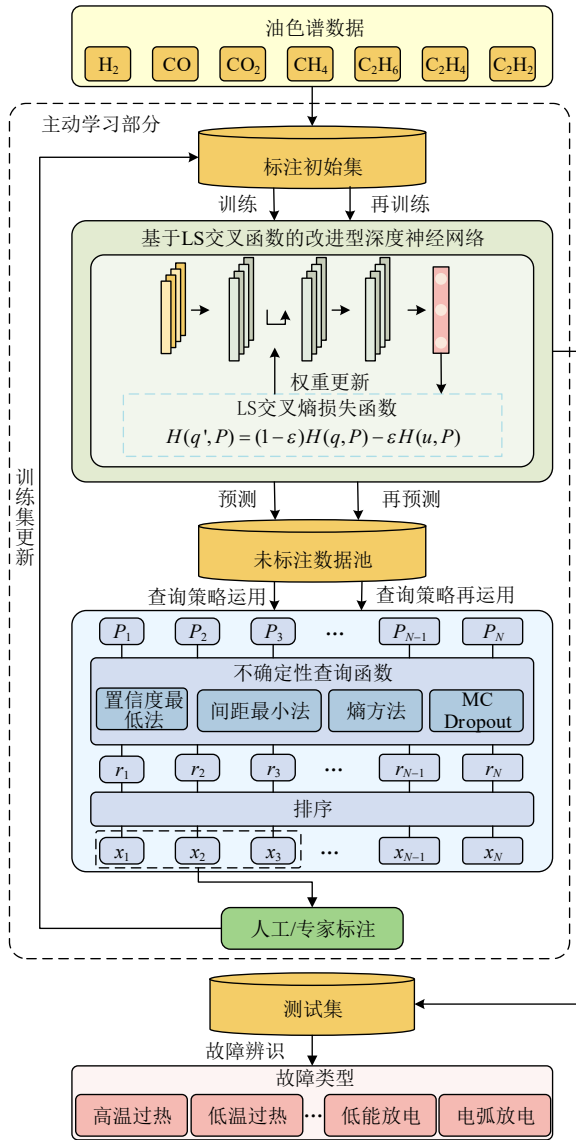


图1 基于改进型主动深度学习框架的油浸式变压器故障辨识算法框架

Fig. 1 Structure of transformer fault diagnosis algorithm combined improved active learning and deep neural network

键思想在于令分类模型主动地对更有学习价值的数据标记并进行重点学习^[19], 其训练流程可见附录表 A1。通过这种方法, 算法可以仅学习少量数据就能完成较高准确率辨识, 避免了对整个数据集进行遍历的冗余操作, 减少了人工标记的工作量。同时, 有方向性的学习也为不平衡数据的分类问题提供了新的解决方案。与一般的被动学习方法相比, 主动学习会通过查询策略从未标记的数据中选择具有丰富信息量的样本, 并交由专家进行人工标注, 从而扩充数据集, 达到将专家经验融于人工智能算法或机器学习模型当中的目的, 以解决欠采样导致的关键数据丢失与过采样生成错误数据造成的数据价值损失问题。

2.1.2 基于不确定性分析的查询策略

查询策略, 即主动选取样本并进行学习的策略依据, 是主动学习算法的核心, 其主要考虑3个标准进行数据采样: 信息性、代表性和多样性^[20]。信息性可通过数据结果的不确定性进行评价, 分类器所预测出的结果不确定性越强, 则说明数据样本的信息性越强; 代表性以优先选择位于簇中心或分布稠密处的样本为标准, 防止选择离群点; 多样性选择原则为所选数据应尽可能分散于样本特征空间中, 以便训练出一个良好的全局模型。

油色谱数据的不平衡性使代表性与多样性采样标准适用性不强, 而故障样本因在相同训练环境下, 与多数类样本即正常状态样本数据相比所获训练量更少, 训练更不充分, 预测不确定性增加, 更易造成变压器故障辨识模型在对其进行判别时出现误分、错分情况, 具有更强的学习价值, 即更强的信息性。针对油色谱数据的变压器故障辨识模型训练, 应着重于对油色谱数据信息性的参考。然而传统算法的采样策略往往对数据进行随机取样或只根据数据分布中心点运用欧氏距离最小方法对数据样本进行筛选^[21], 忽视了数据样本的信息性选取。因此本文提出由4种不确定性函数构成的主动学习查询策略框架, 对样本进行选择。

不确定性采样对数据在模型中输出的不确定性进行评估, 不确定性高的样本优先让模型进行学习。在分类问题中, 分类器往往会计算输入数据预测为各类别的概率, 类别间预测概率越相似, 说明该数据越难以区分, 数据的不确定性越大学习价值更高。本文采用的部分不确定性采样策略流程如图2所示。

1) 置信度最低法(least confident, LC): 置信度最低方法旨在选择预测后拥有最小的最大概率的样本进行采样标注, 其公式为

$$x_{LC} = \arg \max_{x \in D} (1 - P(y|x)) = \arg \min_{x \in D} P(y|x) \quad (1)$$

式中: x_{LC} 为经过置信度最低查询策略选出的采样标记变压器油色谱数据; $P(y|x)$ 为在数据池 $D(D=(x_1, x_2, x_N))$ 中的输入 x 经过模型预测得到变压器故障类别 y 的概率; y 为最大预测概率故障类别。

2) 间距最小法(margin sampling, MS): 间距最小方法选择预测结果最大和第二大类别概率差值最小的样本进行采样标注, 其公式为

$$x_{MS} = \arg \min_{x \in D} (P(y_1|x) - P(y_2|x)) \quad (2)$$

式中: x_{MS} 为经过间距最小查询策略选出的采样标记油色谱数据; y_1 为使用 x 输入模型后变压器故障

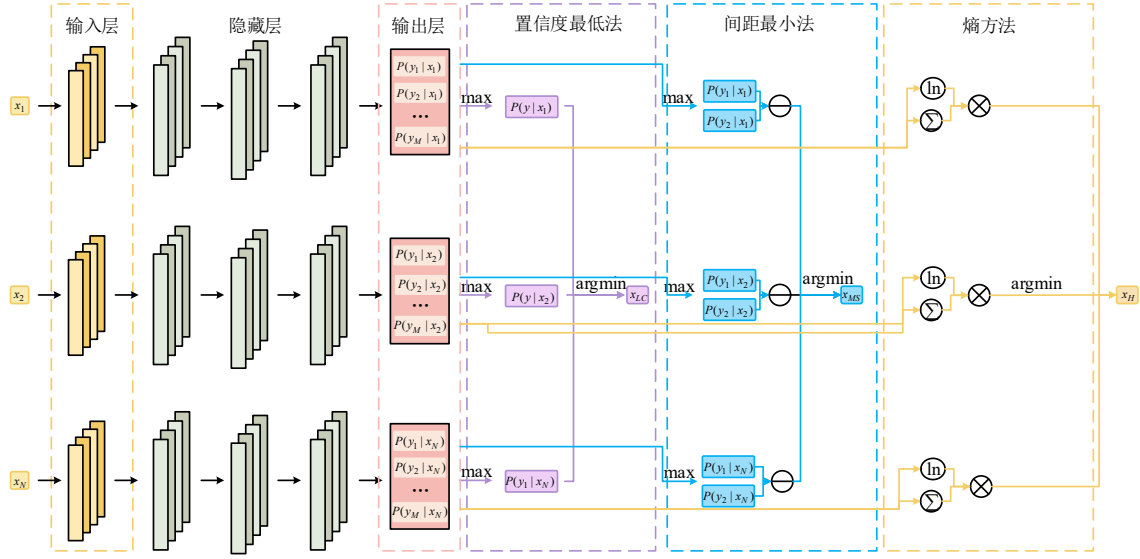


图2 置信度最低法、间距最小法、熵方法查询策略流程图

Fig. 2 Least confident, margin sampling, entropy strategy frameworks

最大预测可能类别； y_2 为变压器故障第二大可能类别。

3) 熵方法(Entropy): 熵方法选择熵更大的数据进行采样标记, 熵越大不确定性则越大:

$$x_H = \arg \max_{x \in D} (-\sum_i P(y_i | x) \ln P(y_i | x)) \quad (3)$$

式中: x_H 为经过熵方法选出的采样标记变压器油色谱数据; i 表示变压器故障第 i 类样本类别。

以上 3 类查询函数为最经典的不确定性函数, 可与不同的分类器针对不同分类问题进行配合使用, 运用范围广泛^[22-24]。

4) 基于 Monte Carlo-Dropout(MC Dropout)过程的查询策略^[25]: 该查询函数为样本数据的信息熵和总体平均信息熵之间的差值, 选择差值最大的样本进行标注, 公式为

$$x_{MC} = \arg \max_{x \in D} \alpha(x, M) = \arg \max_{x \in D} H[Y | X, D] - E_{\theta, P(\theta|D)}[H[Y | x, \theta]] \quad (4)$$

式中: x_{MC} 为基于 MC Dropout 过程查询策略选出的采样标记变压器油色谱数据; $H[Y | X, D]$ 为计算输入 x 预测值 $P(\theta|x, D)$ 的信息熵; $E_{\theta, P(\theta|D)}[H[Y | x, \theta]]$ 为计算分布 $P(\theta|x, D)$ 的信息熵; θ 为神经网络 M 的参数。

而上述公式无法直接进行计算, 故通过 Monte Carlo Dropout 过程(MC Dropout)对其近似函数进行模拟从而得到近似值, 近似函数为

$$x_{MC} \approx -\sum_i (\frac{1}{T} \sum_t \hat{P}_i^t) \lg(\frac{1}{T} \sum_t \hat{P}_i^t) + \frac{1}{T} \sum_{i,t} \hat{P}_i^t \lg \hat{P}_i^t \quad (5)$$

式中: t 表示 dropout 的次数, $t \in \{1, \dots, T\}$; i 为类别, $i \in \{1, \dots, C\}$; $\hat{P}_i^t$ 为输入值后的第 t 次 dropout 经过

网络预测为各类别 i 的概率。具体流程如图 3 所示。

MC Dropout 过程在将未标记数据输入训练好的分类模型进行预测时, 不会关闭向前传播过程中的 Dropout 层功能, 以此令同一输入可获得多个网络结构的输出, 从而达到使用多个分类模型预测评估的效果, 最后将得到的多个输出进行去平均值, 即可得到模型对预测值的评估结果。其多次向前传播过程为并行计算过程, 故所用时间与一次向前传播所用时间相同。MC Dropout 函数为针对深度神经网络搭配使用的新型查询函数, 在保证采集的油色谱数据具有较高信息性的同时, 利用开放 Dropout 层实现了对多个变压器故障辨识模型场景下的油色谱数据评估情况。

本文采用的基于不确定性函数的 4 类主动学习查询策略, 解决了无目的性遍历数据样本采样造成的分类模型学习能力低的问题, 使变压器故障辨识模型可在有限的样本数据中获取更高价值的样本信息, 从而使变压器故障辨识能力得到提高。

2.2 LS 交叉熵损失函数构建

近年来, 基于深度学习的故障诊断方法层出不穷且都有很好的效果, 但其始终需要大量数据做支撑, 且变压器油色谱数据特征值较少, 过于复杂的神经网络模型在提高计算复杂度的同时, 降低分类准确性^[26]。为缓解在低维数据上使用复杂神经网络算法易造成故障判别性能减退的问题, 本文选择使用深度神经网络(deep neural network, DNN)^[27]为基础分类模型对变压器故障进行判别, 在此损失函数部分提出 LS 交叉熵方法, 以预防主动学习过程中油色谱训练集与测试集分布不同所可能导致的过拟合现象。具体结构如图 4 所示, 其中 P_k 、 q_k 、 q_k'

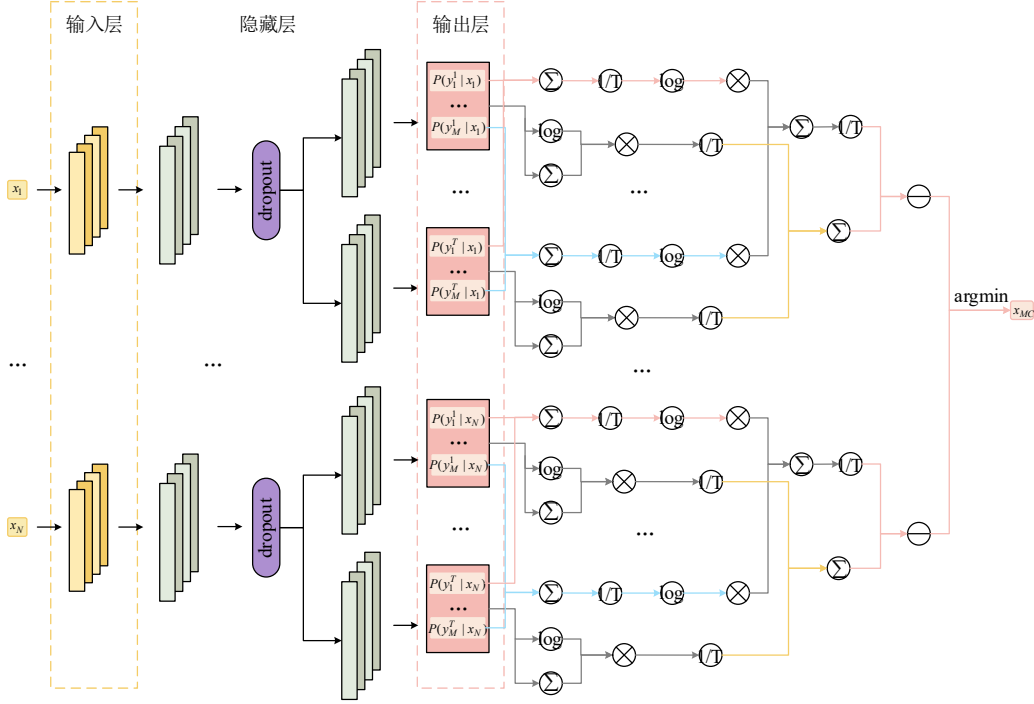


图3 MC Dropout 策略流程图

Fig. 3 MC Dropout strategy frameworks

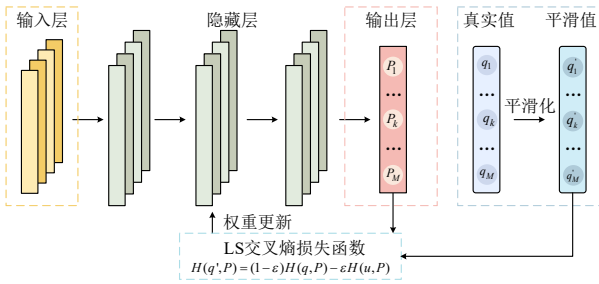


图4 基于LS交叉熵损失函数的DNN结构

Fig. 4 DNN structure using LS cross entropy loss function

与公式(6)~(15)中 $P(k|x)$ 、 $q(k|x)$ 、 $q'(k|x)$ 含义相同。

2.2.1 交叉熵损失函数

分类问题中, DNN 通常使用 Softmax 方法^[28] 计算输入样本 x 归类为不同标签的概率, Softmax 运算公式为

$$P(k|x) = \frac{\exp(o_k)}{\sum_{i=1}^K \exp(o_i)} \quad (6)$$

式中: o_k 为神经网络未规范化的预测值(logit); k 为标签, $k \in \{1, \dots, K\}$ 。

设经编码后的标签为 $q(k|x)$, 则有:

$$\sum_k q(k|x) = 1 \quad (7)$$

为简洁表达, 本文忽略预测变压器故障概率 P 与标签 q 对变压器油色谱样本 x 的依赖关系, 则传统交叉熵函数(cross-entropy loss)定义为

$$H(q, P) = -\sum_{k=1}^K q(k) \lg P(k) \quad (8)$$

将式(6)代入式(8)中, 得到:

$$H(q, P) = \lg \sum_{k=1}^K \exp(o_k) - \sum_{k=1}^K q(k) o_k \quad (9)$$

对其进行关于 o_k 求导:

$$\frac{dL}{do_k} = P(k) - q(k) \quad (10)$$

在传统的 one-hot 编码中, 通常将分类问题的目标类别标为 1 而非目标类别标为 0, 即:

$$q(k) = \begin{cases} 1, & k = y \\ 0, & k \neq y \end{cases} \quad (11)$$

式中: y 为目标类别, 由此 $q(k)$ 可等效为一个类迪拉克函数 $\delta_{k,y}$ 。

由式(9)(10)可得, 交叉熵损失函数表示为

$$H(q, P) = -\lg P(y) = -\lg \left(\frac{\exp(o_k)}{\sum_{i=1}^K \exp(o_i)} \right) \quad (12)$$

此时, 最小化交叉熵函数等同于最大化目标类别标签的对数似然函数, 为此, 目标类别对应的 logit 需远大于其他所有非目标类别的 logit 才可逼近对数似然函数最大值。由此, 暴露出 one-hot 编码方法存在的两个问题:

1) 易出现过拟合问题。由式(10)可知, 令交叉熵函数最小化, 需使 $P(k)$ 逼近 $q(k)$ 数值, 但数值过小的 $P(k)$ 不利于损失函数的减小, 易导致变压器故障辨识模型泛化性能降低。

2) 适应性降低。 o_y 与 o_k 之间差值过大易导致模型对自身辨识结果盲目自信, 输出结果偏激。

2.2.2 LS 交叉熵损失函数

在主动学习过程中, 变压器故障辨识模型会因针对性的油色谱数据采集学习, 使选用的训练集与

测试集分布不一致,易导致测试过程出现过拟合现象,搭配 one-hot 编码方法的交叉熵损失函数不再适用。为此,本文提出 Label Smoothing 算法对交叉熵损失函数进行改进,以配合主动学习稳定运算。

Label Smoothing 是一种正则化算法^[29],其通过对 one-hot 编码进行平滑化处理来减小非目标类别与目标类别概率差距,预防辨识模型过分自信地对样本进行预测标记,该算法如下所示:

$$q'(k|x) = (1-\varepsilon)\delta_{k,y} + \varepsilon u(k) \quad (13)$$

式中: ε 为大小位于 0 到 1 之间的可调参量; $u(k)$ 为表示独立于样本 x 的标签分布,本文使用均匀分布,则新标签表示为

$$q'(k|x) = (1-\varepsilon)\delta_{k,y} + \frac{\varepsilon}{K} \quad (14)$$

将 $q'(k|x)$ 替换入式(8),可得改进后的交叉熵函数,即 LS 交叉熵函数:

$$\begin{aligned} H(q',P) &= -\sum_{k=1}^K q'(k) \lg P(k) = \\ &= -\sum_{k=1}^K [(1-\varepsilon)\delta_{k,y} + \frac{\varepsilon}{K}] \lg P(k) = \\ &= (1-\varepsilon)H(q,P) + \varepsilon H(u,P) \end{aligned} \quad (15)$$

由式(15)分析,可将 LS 交叉熵损失函数分为原交叉熵函数和与 $u(k)$ 相关的交叉熵函数,当 $P(k)$ 偏离 $u(k)$ 的分布时, $H(u,P)$ 将对模型进行惩罚,达到正则化的目的,从而减轻过拟合现象。

3 实验验证

在实际工况下,变压器故障辨识可分为故障事后分析和故障在线检测两类业务场景。在故障事后分析中,业务人员已知变压器发生故障,需对其故障类型进行判断;在故障在线检测中,业务人员需依据数据判断变压器是否存在故障。因此,针对上述两种业务场景,本文设计了故障样本学习实验与正常样本混合实验,利用不平衡油色谱数据对本文提出算法与一系列传统算法进行性能对比。

3.1 被动采样策略

为证明本文所提出的算法能促使变压器故障辨识模型更高效地对少量且不平衡的数据样本进行学习,本文选用随机采样方法、贪婪采样方法两种常见的被动数据采样策略与其进行对比,并对其进行简要介绍。

3.1.1 随机采样策略

随机采样^[30]策略,即对未标记数据池内的数据随机选取后送于人工进行标记。其优点在于,可对不同类别数据进行均匀选取,且计算提取样本所需时间成本低,但当面对分类复杂且数据不平衡的情

况时,随机的数据选取方式容易导致所选数据缺乏代表性与信息性,算法无法进行高效学习,学习成本提高。

3.1.2 贪婪采样策略

贪婪采样(greedy sampling, GS)策略则是确定数据集分布中心点后,选取距离数据集中心最近的点进行采样标记,即采样函数为

$$x_{GS} = \arg \max_{x \in D} (\min \|x - x_C\|) \quad (16)$$

式中: x 为数据池中除去中心点的样本点; x_C 为数据集中心点,本文将该中心点定为数据池中心点。

该算法在油色谱分析变压器故障领域得到应用^[29],其围绕数据集分布中心点采样,可在一定程度上剔除噪点信息,提高算法泛化性,但其采样方法将导致采样区域过于集中,样本点缺乏信息性,易造成辨识算法泛化性能下降。

3.2 数据集构建与数据预处理

本文油色谱数据由国家电网公司所属的部分省级变电站提供,设备电压等级覆盖 110-500kV,共收集 868 组变压器故障油色谱数据与 1952 组状态正常数据进行算法研发及验证。由《DL/T 722—2014 变压器油中溶解气体分析和判断导则》可知,不同故障类型所产生的主要气体组成有所不同,具有判断变压器故障的价值气体分别为 H_2 , CO , CO_2 , CH_4 , C_2H_6 , C_2H_4 , C_2H_2 , 因此,为更充分搭建变压器内部气体组分与故障类别之间的关系,本文选择该 7 种气体含量作为变压器故障辨识依据。采集所得样本数据分为:正常、电弧放电、低能放电、局部放电、高温过热、中温过热、低温过热、低能放电兼过热和电弧放电兼过热 9 类,其类别、标签与数据量分配如表 1 所示。

表 1 变压器数据集构成			
Table 1 Composition of transformer dataset			
故障类型	样本数量	标签	占比/%
电弧放电	283	0	10
低能放电	57	1	6
局部放电	151	2	5
高温过热	164	3	3
中温过热	35	4	2
低温过热	69	5	2
低能放电兼过热	55	6	2
电弧放电兼过热	54	7	1
正常	1952	8	69

为减小各特征之间的数值差异,使用 L2 范数对数据进行归一化处理,具体公式可见附录式(B1)。

3.3 训练环境与训练流程

3.3.1 训练环境

本文实验使用 MacOS Big Sur 操作系统,内核

版本 11.4, CPU 为双核英特尔酷睿 i5 2.3GHz, 采用单显卡 Intel Iris Plus Graphics 640 1536 MB, 8GB RAM 内存。配合 Python3.9 使用 Pytorch 2.0 框架实现主动学习算法构建。

3.3.2 训练流程

利用 DNN 搭建的分类器模型架构如表 2 所示。实验采用 Adam 优化器, 使用 LS 交叉熵损失函数。

表 2 DNN 模型架构参数 Table 2 Parameters of DNN		
参数类型	参数含义	参数值
网络结构	—	7-50-50-50-6
激活函数	—	ReLU
lr	学习率	1e-4
dropout	丢弃率	0.5
batch size	批尺寸	128
epoch	训练轮数	1500

3.4 故障样本学习实验

为评价变压器故障辨识模型性能, 本文选用准确率、召回率、精确率与 F1 指数作为评价指标, 具体公式可见附录表 A2。

本文使用所提出的基于不确定性查询函数采样策略与传统采样策略首先针对故障样本集合进行故障学习实验, 并将数据打乱且按照 6:2:2 的比例将其分为待选数据池、初始数据集与测试集三部分, 具体分配样本量可见表 3。从未标注的数据池中选取 8×50 组数据(剩余 55 组未标记数据以保留不同算法的采样方法特性)加入初始数据集以分析比较结合各采样机制的神经网络面对不平衡样本的学习性能。其中 8 为采样轮数, 50 为每轮采样数。

训练情况如图 5 与附录图 A1 所示, 可见运用本文提出的算法损失值收敛更快, 准确率提升更高。

由表 4 可知采用了本文提出的算法性能皆高于采用传统采样策略的神经网络, 以置信度最低查询策略为例, 与随机采样相比, 其各指标分别超出 13.79%、20.00%、22.19%、21.43%; 与贪婪采样相比, 其各指标分别超出 11.49%、19.01%、21.82%、20.40%。附录图 A2 作为混淆矩阵进一步说明了主动学习策略的神经网络具有更高的识别精度, 并体现在了少数类样本中。

3.5 正常样本混合实验

3.5.1 正常样本混合后模型性能对比

为验证辨识模型在正常数据远超于各类异常数

表 3 故障样本集合及其样本量 Table 3 Sample sets and their sample size	
样本集合	样本量
数据池	520
初始数据集	174
测试集	174

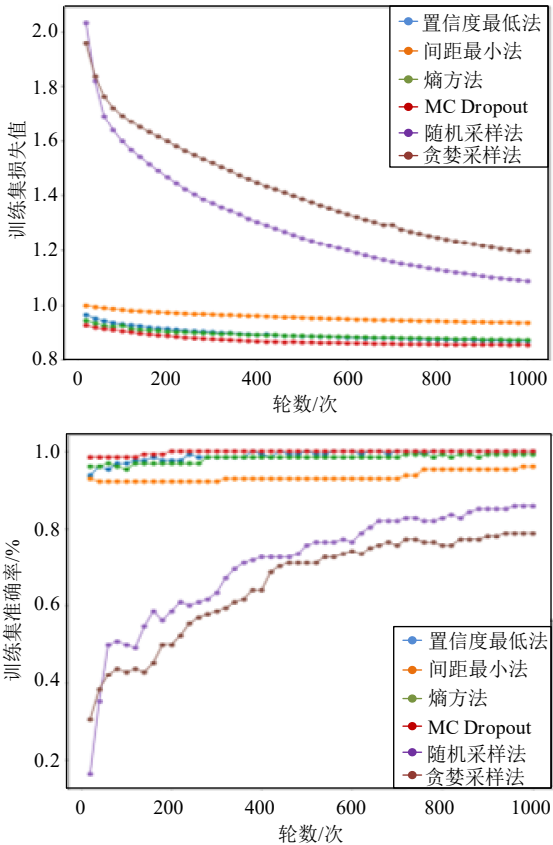


图 5 不同采样策略在训练集准确率、损失值的变化趋势

Fig. 5 Variation trend of accuracy and loss value in training set with different sampling strategies

表 4 主动学习与传统采样策略的性能对比
Table 4 Performance comparison between active learning and traditional sampling strategies

策略名称	准确率/%	召回率/%	精确率/%	F1 指数/%
置信度最低法	89.08	87.42	88.59	87.21
间距最小法	86.78	83.33	86.78	84.07
熵方法	89.08	87.58	88.22	87.43
MC Dropout	85.63	79.58	84.32	79.68
随机采样法	75.29	67.43	66.40	65.78
贪婪采样法	77.59	68.41	66.77	66.81

据样本数量的情况下对故障样本类别判断的性能, 本节分别加入 1388 组与 278 组正常数据到数据池与初始数据集中, 测试集加入 390 组。采样数据样本量与神经网络参数设置同 3.4 节。在实际工况中, 运维人员首先检查变压器是否存在故障, 其次再检测变压器出现了何种类型故障, 因此, 在进行故障辨识工作前, 本文先使用 IAL-DNN-TFDA 与正常采样方法对正常样本与故障样本进行二分类实验, 性能对比如附录表 A3; 随后再进行变压器故障诊断的多分类实验, 结果如表 5 所示。由上述两表可知, 主动学习方法可进一步提高变压器故障识别性能, 降低变压器故障辨识模型的误报、错报情况。

3.5.2 采样可视化分析

为更直观了解本文所提出的算法与使用传统采

表 5 加入正常样本主动学习与传统采样策略性能对比				
Table 5 Performance comparison between active learning and traditional sampling strategies adding normal samples				
策略名称	准确率/%	召回率/%	精确率/%	F1 指数/%
置信度最低法	96.45	89.32	91.77	89.62
间距最小法	96.63	89.88	90.56	89.41
熵方法	94.32	82.85	82.53	81.02
MC Dropout	96.28	88.62	91.49	88.97
随机采样法	89.01	62.27	65.32	62.42
贪婪采样法	92.02	70.94	73.84	71.34

样方法算法的不同，本节将数据池样本对应其故障类别及利用各类采样方法所选中的 10 个变压器油色谱采集样本使用 t-SNE^[31]进行二维可视化处理。

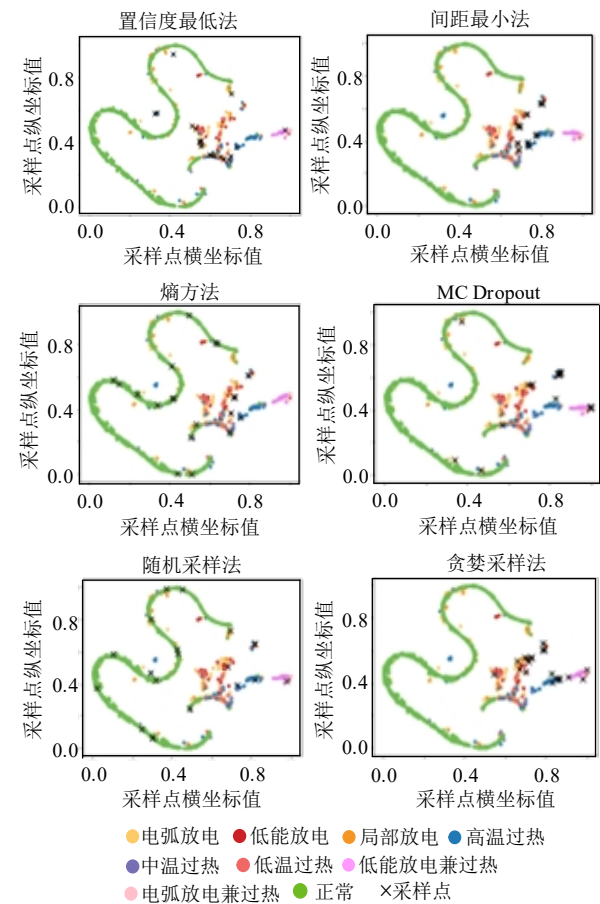


图 6 不同采样策略在数据池中所选样本 t-SNE 可视化
Fig. 6 Visualization of t-SNE for selected samples in the data pool using different sampling strategies

如图 6 所示，随机采样选择均匀随机取点，贪婪采样从样本分布的中心点出发进行扩散型的采样，两种方法皆未考虑数据样本的信息性选取，而本文使用的基于不确定性采样的查询策略都首先选择对较难以区分的样本点进行采样，即从故障样本开始进行采样学习，使算法得以更高效地学习数据样本中的信息，从而提高算法能力。

3.5.3 传统算法性能对比

目前较为常见的基于油色谱数据实现的变压器故障诊断模型有支持向量机(support vector

classifier, SVM)与卷积神经网络^[32-33]。因此，本文选取 SVM 与 CNN 以及传统诊断方法三比值法作为分类模型与本文所提出的 IAL-DNN-TFDA 模型进行性能对比。其中 IAL-DNN-TFDA 模型选用置信度最低函数为代表作查询策略；SVM 算法核函数设置为高斯核函数，惩罚项参数设为 1000，停止训练误差设为 0.001；CNN 模型激活函数、学习率、丢弃率、批尺寸、训练轮数设置同表 3，网络深度为 3。同时，3 种对比算法变压器油色谱训练集数据量和测试集数量与实验 3.5.1 节一致。

由表 6 可知，本文提出的 IAL-DNN-TFDA 模型对变压器故障准确率均超过传统算法，以召回率为例，IAL-DNN-TFDA 分别超过 SVM(17.20%)，CNN(16.58%)，三比值法(34.57%)，其余指标皆有不同程度的提升。取各指标平均值可得，本文算法相较于传统算法，性能提高 15%以上。

表 6 主动学习与传统算法性能对比				
Table 6 Comparison of sampling numbers between active learning and traditional sampling strategies				
模型名称	准确率/%	召回率/%	精确率/%	F1 指数/%
IAL-DNN-TFDA	96.45	89.32	91.77	89.62
SVM	88.82	72.12	75.98	70.82
CNN	89.36	72.74	77.83	73.10
三比值法	22.34	54.78	40.00	41.92

3.6 训练成本实验

为进一步验证本文所提出的 IAL-DNN-TFDA 模型具有降低数据标注成本能力，本文选择减少 IAL-DNN-TFDA 训练集的采样数量，达到降低模型准确率至与采用上述传统采样方法准确率相似水平的目的，采样数对比如表 7 所示。

表 7 主动学习与传统采样策略采样数对比				
Table 7 Comparison of sampling numbers between active learning and traditional sampling strategies				
策略名称	故障样本		正常样本混合	
	准确率/%	采集数	准确率/%	采集数
置信度最低法	77.59	150	93.09	150
间距最小法	77.01	200	89.72	150
熵方法	78.16	150	92.91	150
MC Dropout	77.01	150	92.38	150
随机采样法	75.29	150	89.01	400
贪婪采样法	77.59	150	92.02	400

由此说明了 IAL-DNN-TFDA 模型在运用少于传统采样标记方法模型约 50%的标记数据集即可获得与之相当的准确率，证明了本文所提出模型可有效降低人工标注数据集成本。

3.7 损失函数对比实验

本文提出了使用 LS 交叉熵损失函数对主动学习算法中的 DNN 分类模型进行改良，以防止出现

过拟合现象。为体现 LS 交叉熵损失函数运用效果, 本文将使用交叉熵损失函数的主动学习方法与使用 LS 交叉熵损失函数的主动学习方法进行训练损失值对比, 如图 7 与附录图 A3 所示。

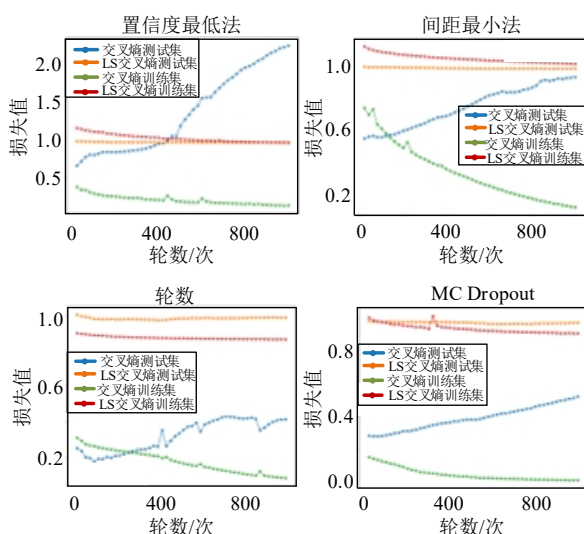


图7 未加入正常样本的各类主动学习运用不同损失函数损失值变化趋势对比

Fig. 7 Trend of loss values of different Active Learning using different loss functions without adding normal samples with training epochs

由上述图可知, 主动学习方法在使用传统交叉熵损失函数时, 训练集中损失值虽可取得更小的最小化值, 但在测试集上都出现了不同程度的过拟合现象; 使用 LS 交叉熵损失函数后, 训练集中损失值因 LS 交叉熵损失函数运算中的惩罚机制不进行过分最小化, 使测试集训练过程呈现稳定下降。

4 结论

本文针对人工智能算法训练变压器故障诊断模型过程中因故障数据样本量少、数据不平衡产生的故障诊断准确率低的问题, 提出了基于改进型主动深度学习框架的油浸式变压器故障辨识算法。本文的主要结论及创新点如下:

1) 基于主动学习算法改进了变压器油色谱故障辨识模型, 通过主动学习中的查询策略, 对变压器油色谱数据价值进行评估并进行针对性的学习, 相较于随机采样、贪婪采样等传统采样方法样本需求得以减少。

2) 针对主动学习选择数据的规律引入了 Label Smoothing 算法改进 DNN 的交叉熵损失函数, 缓解主动学习因针对性选取油色谱样本训练所产生的过拟合问题。

3) 实现了针对变压器油色谱数据的主动学习与 DNN 的融合适配, 为除变压器故障辨识以外的

其他分类及回归问题模型优化提供了新思路。

综上, 本文提出的算法模型使神经网络在变压器油色谱数据不平衡的情况下的故障辨识仍能维持较高准确性与泛化性, 但目前仍然存在一些问题有待进一步研究: 其一, 目前使用的网络为深度神经网络, 其数据集在到达一定规模后, 会出现训练周期较长的问题, 未来需寻找效率更高的神经网络分类模型; 其二, 本文从油色谱数据故障判别结果不确定性大的角度出发, 选择了 4 种主动学习查询策略, 尚未综合考虑多样性与代表性采样原则; 此外, 本文所用油色谱数据不包含时变信息, 因此暂未考虑产气速率等高阶数据特征, 下一步将开展考虑时变关系的连续型及高阶数值输入的主动学习策略的研究。

附录见本刊网络版 (<http://www.dwjs.com.cn/CN/1000-3673/current.shtml>)。

参考文献

- [1] 齐波, 王一鸣, 张鹏, 等. 面向变压器油色谱趋势预测的深度递归信念网络[J]. 电网技术, 2019, 43(6): 1892-1899.
QI Bo, WANG Yiming, ZHANG Peng, et al. Deep recurrent belief network model for trend prediction of transformer oil chromatography data[J]. Power System Technology, 2019, 43(6): 1892-1899(in Chinese).
- [2] 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(in Chinese).
- [3] 邓祥力, 朱慧, 杨梅, 等. 基于多状态模型的变压器绕组早期故障保护方法研究[J]. 中国电机工程学报, 2022, 42(18): 6704-6714.
DENG Xiangli, ZHU Hui, YANG Mei, et al. Research on early fault protection method of transformer winding based on multi-state model[J]. Proceedings of the CSEE, 2022, 42(18): 6704-6714(in Chinese).
- [4] 张鹏, 齐波, 张若愚, 等. 基于经验小波变换和梯度提升径向基的变压器油中溶解气体预测方法[J]. 电网技术, 2021, 45(9): 3745-3754.
ZHANG Peng, QI Bo, ZHANG Ruoyu, et al. Dissolved gas prediction in transformer oil based on empirical wavelet transform and gradient boosting radial basis[J]. Power System Technology, 2021, 45(9): 3745-3754(in Chinese).
- [5] MOLLMANN A, PAHLAVANPOUR B. New guidelines for interpretation of dissolved gas analysis in oil-filled transformers[M]. CIGRE, 1999: 31-51.
- [6] 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.
- [7] 国家能源局. 变压器油中溶解气体分析和判断导则: DL/T 722—2014[S]. 北京: 中国电力出版社, 2015.
- [8] 齐波, 张鹏, 荣智海, 等. 基于相空间重构的变压器油色谱数据最优长度选择方法[J]. 中国电机工程学报, 2018, 38(8): 2504-2512.
QI Bo, ZHANG Peng, RONG Zhihai, et al. Optimal length selection method of DGA data based on phase space reconstruction[J]. Proceedings of the CSEE, 2018, 38(8): 2504-2512(in Chinese).
- [9] 齐波, 王一鸣, 张鹏, 等. 基于自决策主动纠偏的电力变压器油色谱诊断模型[J]. 高电压技术, 2020, 46(1): 23-32.
QI Bo, WANG Yiming, ZHANG Peng, et al. Oil chromatography

- diagnosis model of power transformer based on self-decision active rectification[J]. High Voltage Engineering, 2020, 46(1): 23-32(in Chinese).
- [10] 康佳宇, 张沈习, 张庆平, 等. 基于 ANOVA 和 BO-SVM 的变压器故障诊断方法[J]. 高电压技术, 2023, 49(5): 1882-1891. KANG Jiayu, ZHANG Shenxi, ZHANG Qingping, et al. Fault diagnosis method of transformer based on ANOVA and BO-SVM[J]. High Voltage Engineering, 2023, 49(5): 1882-1891(in Chinese).
- [11] 李昂, 韩萌, 穆栋梁, 等. 多类不平衡数据分类方法综述[J]. 计算机应用研究, 2022, 39(12): 3534-3545. LI Ang, HAN Meng, MU Dongliang, et al. Survey of multi-class imbalanced data classification methods[J]. Application Research of Computers, 2022, 39(12): 3534-3545(in Chinese).
- [12] 刘云鹏, 许自强, 和家慧, 等. 基于条件式 Wasserstein 生成对抗网络的电力变压器故障样本增强技术[J]. 电网技术, 2020, 44(4): 1505-1513. LIU Yunpeng, XU Ziqiang, HE Jiahui, et al. Data augmentation method for power transformer fault diagnosis based on conditional Wasserstein generative adversarial network[J]. Power System Technology, 2020, 44(4): 1505-1513(in Chinese).
- [13] 李雅欣, 侯慧娟, 胥明凯, 等. 基于策略梯度和生成式对抗网络的变压器油色谱案例扩充方法[J]. 电力自动化设备, 2020, 40(12): 211-217. LI Yaxin, HOU Huijuan, XU Mingkai, et al. Oil chromatogram case generation method of transformer based on policy gradient and generative adversarial networks[J]. Electric Power Automation Equipment, 2020, 40(12): 211-217(in Chinese).
- [14] 韩笑, 王新迎, 韩帅, 等. 基于不平衡数据集集成学习的大型电力变压器状态评价方法[J]. 电网技术, 2021, 45(1): 107-114. HAN Xiao, WANG Xinying, HAN Shuai, et al. Ensemble learning method for large-scale power transformer status evaluation based on imbalanced data[J]. Power System Technology, 2021, 45(1): 107-114(in Chinese).
- [15] 王艳, 李伟, 赵洪山, 等. 基于融合少数类过采样均衡多分类数据的改进极限学习机的变压器故障诊断方法[J]. 电网技术, 2023, 47(9): 3799-3807. WANG Yan, LI Wei, ZHAO Hongshan, et al. Transformer fault diagnosis fused with synthetic minority over-sampling balanced multi-classification data based on improved extreme learning machine[J]. Power System Technology, 2023, 47(9): 3799-3807(in Chinese).
- [16] YAP B W, RANI K A, RAHMAN H A A, et al. An application of oversampling, undersampling, bagging and boosting in handling imbalanced datasets[C]//Proceedings of the First International Conference on Advanced Data and Information Engineering. Singapore: Springer, 2014: 13-22.
- [17] ABDI L, HASHEMI S. To combat multi-class imbalanced problems by means of over-sampling techniques[J]. IEEE Transactions on Knowledge and Data Engineering, 2016, 28(1): 238-251.
- [18] AGRAWAL A, VIKTOR H L, PAQUET E. SCUT: multi-class imbalanced data classification using SMOTE and cluster-based undersampling[C]//2015 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K). Lisbon, Portugal: IEEE, 2015: 226-234.
- [19] SETTLES B. Active learning literature survey[R]. Madison: University of Wisconsin, 2009: 762-769.
- [20] WU Dongrui. Pool-based sequential active learning for regression[J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(5): 1348-1359.
- [21] WU Dongrui, LIN C T, HUANG Jian. Active learning for regression using greedy sampling[J]. Information Sciences, 2019, 474: 90-105.
- [22] NÚÑEZ J A, CINCOTTA P M, WACHLIN F C. Information entropy[J]. Celestial Mechanics and Dynamical Astronomy, 1996, 64(1): 43-53.
- [23] JAIN P, KAPOOR A. Active learning for large multi-class problems [C]//2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, FL, USA: IEEE, 2009: 762-769.
- [24] ZHOU Jin, SUN Shiliang. Improved margin sampling for active learning[C]//6th Chinese Conference on Pattern Recognition. Changsha: Springer, 2014: 120-129.
- [25] TSYMBALOV E, PANOVA M, SHAPEEV A. Dropout-based active learning for regression[C]//7th International Conference on Analysis of Images, Social Networks and Texts. Moscow: Springer, 2018: 247-258.
- [26] 李平, 胡根铭. 基于数据增强型一维改进卷积神经网络的变压器故障诊断方法[J]. 电网技术, 2023, 47(7): 2957-2966. LI Ping, HU Genming. Transformer fault diagnosis based on data enhanced one-dimensional improved convolutional neural network[J]. Power System Technology, 2023, 47(7): 2957-2966(in Chinese).
- [27] 陈云霁, 李玲, 李威, 等. 智能计算系统[M]. 北京: 机械工业出版社, 2020: 23-27. CHEN Yunji, LI Ling, LI Wei, et al. AI computing systems[M]. Beijing: China Machine Press, 2020: 23-27(in Chinese).
- [28] WEI Yuxuan, LIU Qixuan, ZHANG Guixu, et al. Label smoothing technique for ordinal classification in cloud assessment[C]//IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium. Waikoloa: IEEE, 2020: 2264-2267.
- [29] 刘颖, 庞羽良, 张伟东, 等. 基于主动学习的图像分类技术: 现状与未来[J]. 电子学报, 2023, 51(10): 2960-2984. LIU Ying, PANG Yuliang, ZHANG Weidong, et al. Active learning-based image classification technology: status and future[J]. Acta Electronica Sinica, 2023, 51(10): 2960-2984(in Chinese).
- [30] TANG Pengfei, ZHANG Zhonghao, TONG Jie, et al. Research on transformer fault diagnosis based on active learning with imbalanced data of dissolved gas in oil[J]. Review of Scientific Instruments, 2024, 95(5): 055101.
- [31] VAN DER MAATEN L, HINTON G. Visualizing data using t-SNE[J]. Journal of Machine Learning Research, 2008, 9(86): 2579-2605.
- [32] 张育杰, 冯健, 李典阳, 等. 基于油色谱数据的变压器故障征兆新优选策略[J]. 电网技术, 2021, 45(8): 3324-3331. ZHANG Yujie, FENG Jian, LI Dianyong, et al. New feature selection method for transformer fault diagnosis based on DGA data[J]. Power System Technology, 2021, 45(8): 3324-3331(in Chinese).
- [33] KARI T, GAO Wensheng, ZHAO Dongbo, et al. Hybrid feature selection approach for power transformer fault diagnosis based on support vector machine and genetic algorithm[J]. IET Generation, Transmission & Distribution, 2018, 12(21): 5672-5680.



黄灿

在线出版日期: 2024-11-20。

收稿日期: 2024-04-08。

作者简介:

黄灿(2000), 女, 硕士, 主要研究方向为电力人工智能, E-mail: m19968097942@163.com;

全杰(1983), 通信作者, 男, 教授级高级工程师, 主要研究方向为电力人工智能、电力物联网等, E-mail: tongjie1@epri.sgcc.com.cn;

唐鹏飞(1997), 男, 工程师, 主要研究方向为电力人工智能、科学机器学习等, E-mail: tangpengfei@epri.sgcc.com.cn;

齐子豪(1999), 男, 硕士, 主要研究方向为电力人工智能、边缘计算等, E-mail: qizihao2021@163.com;

张中浩(1991), 男, 工程师, 博士, 主要从事电力人工智能等方面的研究工作, E-mail: zhangzhonghao@epri.sgcc.com.cn。

(编辑 李健一)

附录 A

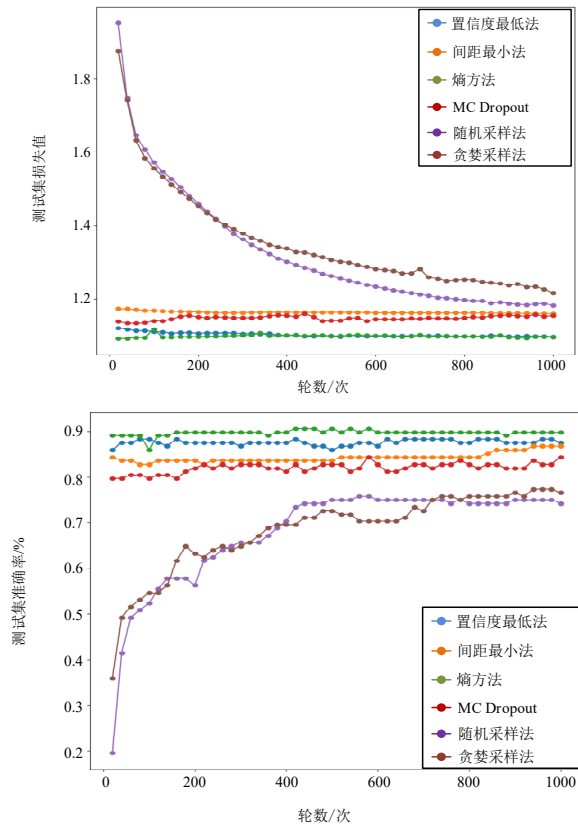


图 A1 不同采样策略测试集准确率、损失值的变化趋势
Fig. A1 The changing trend of accuracy and loss value of test set with different sampling strategies

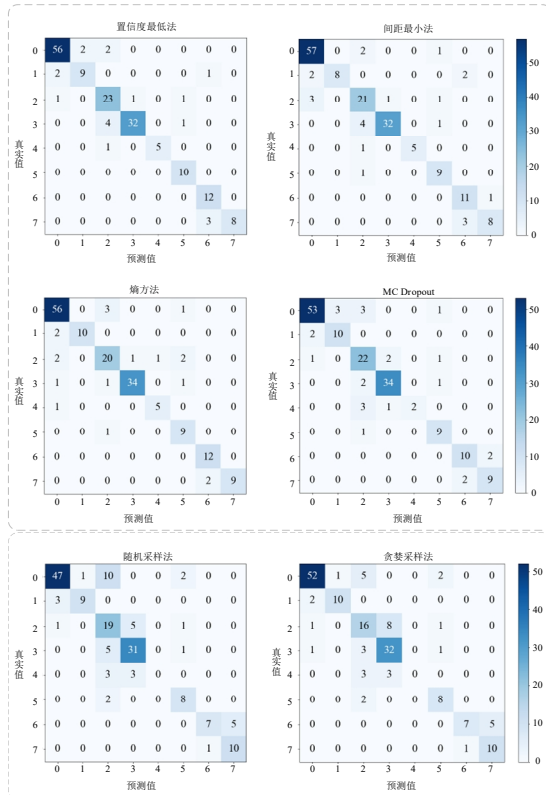


图 A2 不同采样策略在测试集的混淆矩阵
Fig. A2 The Confusion Matrix of Different Sampling Strategies in the Test Set

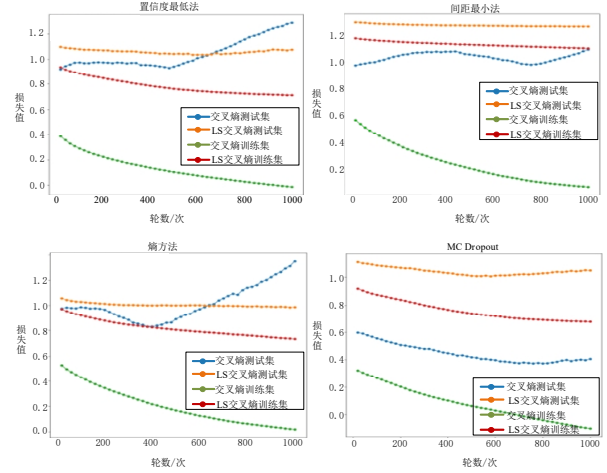


图 A3 各类主动学习运用不同损失函数损失值对比
Fig. A3 Trend of loss values of different Active Learning using different loss functions with training epochs

表 A1 基于主动学习的变压器诊断算法训练流程
Table A1 Training process of transformer diagnosis algorithm based on active learning

输入：油色谱数据(H_2 , CO , CO_2 , CH_4 , C_2H_6 , C_2H_4 , C_2H_2)
输出：故障辨识结果(电弧放电、低能放电、局部放电、高温过热、中温过热、低温过热)
初始化：数据池 P ，初始数据集 I ，添加样本轮数 m ，每轮选取样本数 f ，循环次数 i ，网络结构，网络参数。

- 1: for $i=1,2,\dots,m$:
- 2: 利用 I 训练神经网络
- 3: 将 P 投入神经网络进行预测
- 4: 利用查询函数计算预测结果分数 r
- 5: 选取 r 最高的前 f 个样本，构成数据集 S
- 6: 添加 I : $I := I \cup S$
- 7: 剔除 S : $P := P - S$
- 8: 训练整体已标注的数据样本

表 A2 评价指标
Table A2 Evaluation index

指标	计算公式
准确率	$r_{acc} = \frac{y_{TP} + y_{TN}}{y_{TP} + y_{TN} + y_{FP} + y_{FN}}$
召回率	$r_{recall} = \frac{y_{TP}}{y_{TP} + y_{FN}}$
精确率	$r_{pre} = \frac{y_{TP}}{y_{TP} + y_{FP}}$
F1 指数	$F_1 = \frac{2 \times r_{pre} \times r_{recall}}{r_{pre} + r_{recall}}$

其中 y_{TP} 表示真正例， y_{TN} 表示真负例， y_{FP} 表示假正例， y_{FN} 表示假负例。

表 A3 主动学习与传统采样策略故障判别指标对比
Tab.A3 Comparison of fault identification indicators between active learning and traditional sampling strategies

策略名称	准确率/%	召回率/%	精确率/%	F1 指数/%
置信度最低法	99.47	98.28	99.89	99.13
间距最小法	99.47	98.85	99.42	99.13
熵方法	99.64	98.85	99.89	99.42
MC Dropout	99.64	99.43	99.43	99.43
随机采样法	98.58	95.98	99.40	97.66

贪婪采样法	98.05	94.25	99.39	96.76
-------	-------	-------	-------	-------

附录 B

L2 范数公式为:

$$\|x\|_2 = (|x_1|^2 + |x_2|^2 + \dots + |x_n|^2)^{1/2} \quad (B1)$$

其中， $\|x\|_2$ 表示样本 x 每行 L2 范数值， $|x_n|$ 表示样本 x 每行的第 n 个元素。将各行的每个元素同除以各行的 $\|x\|_2$ ，即可得到 L2 范数归一化结果。