

DOI:10.19651/j.cnki.emt.2212123

## 少标签下油浸式变压器双层故障诊断模型\*

汤茂祥<sup>1</sup> 王聪<sup>1</sup> 朱超平<sup>2</sup> 马萍<sup>1</sup> 王伟<sup>1</sup>

(1.新疆大学电气工程学院 乌鲁木齐 830000; 2.新疆送变电有限公司 乌鲁木齐 830000)

**摘要:** 针对油浸式变压器数据样本标记困难,有标签样本数据量较少,传统故障诊断方法精度低的问题,提出了基于GBDT与K-means增益聚类的少标签下油浸式变压器双层故障诊断模型。首先,采用SAE对表征变压器状态的高维特征气体进行降维,去除冗余信息,得到包含变压器运行状态的低维特征向量作为后续分类器的输入;其次,构建双层故障诊断模型;针对无标签样本,引入GBDT方法作为所提模型首层,获取无标签样本的伪标签;为进一步提高诊断精度,提出基于无标记样本伪标签的K-means聚类增益,作为新的特征向量,输入末层模型K-means用以实现二次诊断的目的。实验分析表明,在少标签状态下,所提的方法可有效提升变压器故障诊断精度,相较于传统方法,在诊断精度上至少提升了6%。为少标签下的油浸式变压器故障诊断提供了新的思路。

**关键词:** 油浸式变压器;故障诊断;少标签;K-means增益

**中图分类号:** TM411 **文献标识码:** A **国家标准学科分类代码:** 520.20

## Transformer fault diagnosis method based on GBDT and K-means gain clustering

Tang Maoxiang<sup>1</sup> Wang Cong<sup>1</sup> Zhu Chaoping<sup>2</sup> Ma Ping<sup>1</sup> Wang Wei<sup>1</sup>

(1. School of Electrical Engineering, Xinjiang University, Urumqi 830000, China;

2. Xinjiang Power Transmission and Transformation Co., Ltd., Urumqi 830000, China)

**Abstract:** In view of the difficulties in labeling the data samples of oil immersed transformers, the small amount of labeled samples and the low accuracy of traditional fault diagnosis methods, a two-layer fault diagnosis model for oil immersed transformers with few labels based on GBDT and K-means gain clustering is proposed. Firstly, a stacked autoencoder is used to reduce the dimension of the high-dimensional characteristic gas characterizing the transformer state, remove redundant information, and obtain the low-dimensional characteristic vector containing the transformer operating state as the input of the subsequent classifier. Secondly, a two-layer fault diagnosis model is constructed; For unlabeled samples, the GBDT method is introduced as the first layer of the proposed model to obtain the false labels of unlabeled samples. In order to further improve the diagnosis accuracy, the K-means clustering gain based on the false label of unlabeled samples is proposed as a new feature vector, which is input into the end layer model K-means to realize the secondary diagnosis. Experimental analysis shows that the proposed method can effectively improve the accuracy of transformer fault diagnosis under the condition of few tags, and the diagnosis accuracy is improved by at least 6% compared with other methods. It provides a new idea for fault diagnosis of oil immersed transformer with few labels.

**Keywords:** oil immersed transformer; fault diagnosis; less labels; K-means gain

## 0 引言

随着我国工业化建设的推进,对于电力资源的需求越来越大,油浸式变压器在电力传输和配送过程中发挥着重要的作用,其运行的可靠性关系着整体电力系统的安全与稳定<sup>[1]</sup>。

基于油中溶解气体分析(dissolved gas analysis, DGA)可以有效的判断变压器的运行状态,为工程维护人员进行设备检测和故障诊断提供有效的信息指导<sup>[2-3]</sup>。在DGA理论基础上,国内外学者提出了Doernenburg比值、大卫三角形、Rogers比值和IEC三比值等方法体系<sup>[4]</sup>。这些传统方

收稿日期:2022-11-17

\* 基金项目:新疆维吾尔自治区自然科学基金青年项目(2022D01C89)、国家自然科学基金(52065064,51967019)项目资助

法简单易用,操作难度低,但是基于阈值绝对和编码缺失等原因,诊断精度无法满足变压器故障检测的需要<sup>[5]</sup>。

随着人工智能的发展,机器学习被广泛应用到变压器故障诊断中,表现优异而受到业界的广泛关注<sup>[6-7]</sup>。文献[8]采用深度置信网络(deep belief networks, DBN),使用 ReLU 作为激活函数,避免了反向调参过程中的梯度消失与梯度爆炸,加快了模型的训练速度;文献[9]采用支持向量机(support vector machine, SVM)进行多级故障诊断,降低单层分类器的数据处理难度,提高模型的诊断精度;文献[10]采用麻雀搜索算法优化的自适应增强算法(adaptive boosting, Adaboost),不仅加快了超参数调优的速度,而且提高了模型的诊断稳定性;文献[11-12]采用优化算法优化的极限梯度提升算法(extreme gradient boosting, XGBoost),引入正则化,避免分类器的过拟合。这些算法需要使用大量的标签数据进行模型驱动,但是油浸式变压器数据样本标记困难,有标签样本数据量较少,无法满足模型的数据量需求<sup>[13-15]</sup>。

可学习理论提出将多个弱分类器进行有序组合,从而获得理想学习效果。在变压器故障诊断中,若有标签数据量较少,无法满足模型训练的需要,模型会被训练成弱分类器。受可学习思想启发,为提高少标签下变压器故障诊断精度,本文使用集成学习将多个同质弱分类器组合,提出基于梯度提升树(gradient boosting decision tree, GBDT)与 K-means 增益聚类的少标签油浸式变压器双层故障诊断模型。首先,采用堆栈自编码器(stacked autoencoder, SAE)对表征变压器状态的高维特征气体降维至 2 维;其次,构建双层故障诊断模型;针对无标签样本,引入 GBDT 方法作为所提模型首层,获取无标签样本的伪标签;为进一步提高诊断精度,提出基于无标记样本伪标签的 K-means 聚类增益,作为新的特征向量,输入末层模型 K-means 用以实现二次诊断的目的。

### 1 油中溶解气体

当油浸式变压器处于工作状态,其中的绝缘材料在高温高压的催化下发生分解反应,产生 H<sub>2</sub>、CH<sub>4</sub>、C<sub>2</sub>H<sub>6</sub>、C<sub>2</sub>H<sub>4</sub> 和 C<sub>2</sub>H<sub>2</sub> 等气体,逐步溶解在绝缘油中。若变压器运行出现故障,产生的气体含量相较于无故障状态下存在着较大差异。气体的含量、种类与故障的类型之间存在着紧密的联系,依据油中溶解气体可以快速诊断变压器的运行状态。

构建表征油浸式变压器运行状态的特征向量有特征气体法,IEC 三比值法和特征气体加无编码比值法,为更好的反应变压器的运行工况,以特征气体加无编码比值法构建 14 维特征向量,特征向量维度主要是 C(H<sub>2</sub>)、C(CH<sub>4</sub>)、C(C<sub>2</sub>H<sub>6</sub>)、C(C<sub>2</sub>H<sub>4</sub>)、C(C<sub>2</sub>H<sub>2</sub>)、C(CH<sub>4</sub>)/C(H<sub>2</sub>)、C(C<sub>2</sub>H<sub>4</sub>)/C(C<sub>2</sub>H<sub>2</sub>)、C(C<sub>2</sub>H<sub>4</sub>)/C(C<sub>2</sub>H<sub>6</sub>)、C(CH<sub>4</sub>)/C(C<sub>1</sub>+C<sub>2</sub>)、C(C<sub>2</sub>H<sub>6</sub>)/C(C<sub>1</sub>+C<sub>2</sub>)、C(C<sub>2</sub>H<sub>4</sub>)/C(C<sub>1</sub>+C<sub>2</sub>)、C(C<sub>2</sub>H<sub>2</sub>)/C(C<sub>1</sub>+C<sub>2</sub>)、C(H<sub>2</sub>)/C(H<sub>2</sub>+C<sub>1</sub>+C<sub>2</sub>)、C(CH<sub>4</sub>+C<sub>2</sub>H<sub>4</sub>)/

(C<sub>1</sub>+C<sub>2</sub>)。

其中,(C<sub>1</sub>+C<sub>2</sub>)=(CH<sub>4</sub>+C<sub>2</sub>H<sub>6</sub>+C<sub>2</sub>H<sub>4</sub>+C<sub>2</sub>H<sub>2</sub>),C 表示气体浓度。

## 2 GBDT 与 K-means 聚类模型

### 2.1 GBDT

集成学习是基于梯度提升(gradient boosting)思想,将多个弱分类器有序组合,形成强分类器的一种机器学习方法。后序列弱分类器不断拟合前序列弱分类器损失函数的负梯度,逐步降低分类器误差,最终形成高分类精度和强泛化能力的强分类器。通过一对多(one vs rest, OVR)策略,可将多个 GBDT 二分类模型组成一个 GBDT 多分类模型。在处理二分类问题时,设置迭代次数 M, y<sub>i</sub> 表示预测目标,第 1 个分类器为:

$$F_0(x) = \frac{1}{2} \log\left(\frac{\sum_{i=1}^N y_i}{\sum_{i=1}^N (1-y_i)}\right), k \in [1, N] \quad (1)$$

F<sub>0</sub>(x) 与 y<sub>i</sub> 之间的损失函数记作 L(y<sub>i</sub>, F(x<sub>i</sub>)), 其负梯度  $\tilde{y}_i$  为:

$$u_i = \frac{1}{1 + \exp(-F_{m-1}(x_i))}$$

$$L(y_i, F_{m-1}(x_i)) = -(y_i \log u_i + (1-y_i) \log(1-u_i)) \quad (2)$$

$$\tilde{y}_i = -\left[\frac{\partial L(y_i, F_{m-1}(x_i))}{\partial F_{m-1}(x_i)}\right]$$

m ∈ [1, M]

生成的决策树的叶子节点数值大小为:

$$\gamma_{jm} = \frac{\sum_{x_i \in R_{jkm}} \tilde{y}_{ik}}{\sum_{x_i \in R_{jkm}} (y_i - \tilde{y}_i)(1 - y_i + \tilde{y}_i)} \quad (3)$$

其中, R<sub>jkm</sub> 表示第 m 次迭代的第 k 棵决策树的第 j 个叶子节点的数值。根据  $\gamma_{jm}$  和学习率  $\alpha$  可以更新分类器。

$$F_m(x) = F_{m-1}(x) + \alpha \sum_{j=1}^J \gamma_{jm} I(x \in R_{jm}) \quad (4)$$

当 L(y<sub>i</sub>, F(x<sub>i</sub>)) 小于设定的阈值时,则表示 GBDT 模型训练完成。GBDT 多分类模型如图 1 所示。

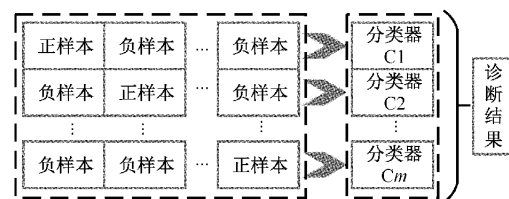


图 1 GBDT 多分类模型

### 2.2 K-means 聚类

K-means 是一种以特征向量的维度间相似度作为度量标准进行聚类的无监督学习算法。聚类中心随机初始化后,通过样本点欧氏距离不断更迭变化,最终趋于稳定而实

现全局聚类,超参数  $K$  表示聚类簇的数量<sup>[16]</sup>。聚类结束后,同簇数据间的相似度较高,异簇数据间的相似度较低。为降低初始聚类中心对聚类结果的影响,会多次重复聚类,选择全体欧式距离最低的聚类作为最终的聚类结果。

设存在  $m$  条  $n$  维数据  $D = \{D_1, D_2, \dots, D_m\}, D_i \in R^n$ , 将其分为  $K$  簇,记为  $C = \{c_k, k = 1, 2, \dots, K\}, c_i$  的聚类中心为  $\sigma_k$ , 数据点到聚类中心的欧氏距离为:

$$J(c_k) = \sum_{x_i \in c_k} \|D_i - \sigma_k\|^2 \quad (5)$$

全体数据的欧氏距离为:

$$J(C) = \sum_{k=1}^K J(c_k) = \sum_{k=1}^K \sum_{x_i \in c_k} \|D_i - \sigma_k\|^2 \quad (6)$$

$J(C)$  的数值越小,表示聚类的效果越好,初始聚类中心选择的越准确。

### 3 基于 GBDT 与 K-means 增益聚类的双层故障诊断模型

在构建的 GBDT 与 K-means 增益聚类的双层故障诊断模型中,GBDT 作为模型首层,以特征气体加无编码比值法构建的 14 维特征向量作为输入向量;K-means 增益聚类作为模型末层,引入 SAE 对特征气体加无编码比值法构建的 14 维特征向量进行降维,与 GBDT 对无标签数据构建的伪标签经过系数拉升后组成三维向量,作为末层分类器的输入特征。

#### 3.1 末层特征向量构建

为降低特征向量维度间的信息冗余,减少诊断模型训练时间,引用 SAE 对以特征气体加无编码比值法构建的 14 维特征向量进行降维处理,构建能够表征故障信息的低维特征向量。

SAE 采用“逐层贪婪算法”分层次对输入的特征向量进行降维,以首层自动编码器(auto-encoders, AE)为例,  $W_{11}, b_{11}$  为编码层结构权重、偏置,  $W_{12}, b_{12}$  为解码层结构权重、偏置,  $y(x)$  为降维后的数据。计算公式为:

$$y(x) = f(W_1 x + b_1) \quad (7)$$

$$z(x) = g(W_2 y(x) + b_2) \quad (8)$$

理想状态下,  $x = z(x)$ , 损失函数为:

$$Loss(x, z) = \frac{1}{2} \sum_{i=1}^N \|x^{(i)} - z(x^{(i)})\|^2 \quad (8)$$

在多次的迭代训练中,  $Loss(x, z)$  逐渐趋于设定的阈值。堆栈自编码器结构如图 2 所示。

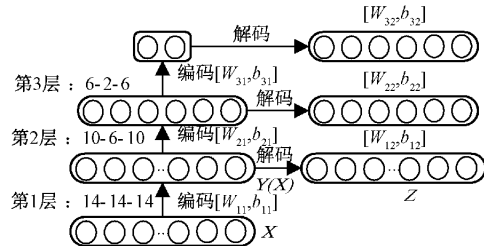


图 2 堆栈自编码器拓扑结构

如图 2 所示,SAE 的拓扑结构为 14-10-6-2。首先,使用首层 AE 将特征向量由 14 维降至 10 维;其次,使用中层 AE 将首层 AE 输出的 10 维特征向量降至 6 维;最后,使用末层 AE 将中层 AE 输出的 6 维特征向量降至 2 维。构成 SAE 的三层 AE 前后衔接,相较于直接使用拓扑结构为 14-2 的 AE,SAE 不仅具有拓扑结构的优化空间,且更能够保留数据的重要信息。

使用有限的有标签样本训练 GBDT, 然后进行无标签样本的伪标签预测, 预测标签记为  $\eta$ , 伪标签  $\eta$  乘以系数  $\alpha$  形成新维度标签  $\gamma$ 。降维特征向量与新维度标签  $\gamma$  组合成末层特征向量。

$$\gamma = \alpha \cdot \eta \quad (9)$$

#### 3.2 双层故障诊断模型

针对少标签样本下的油浸式变压器故障诊断, 本文提出一种 GBDT 与 K-means 增益聚类的双层诊断模型, 模型结构如图 3 所示。

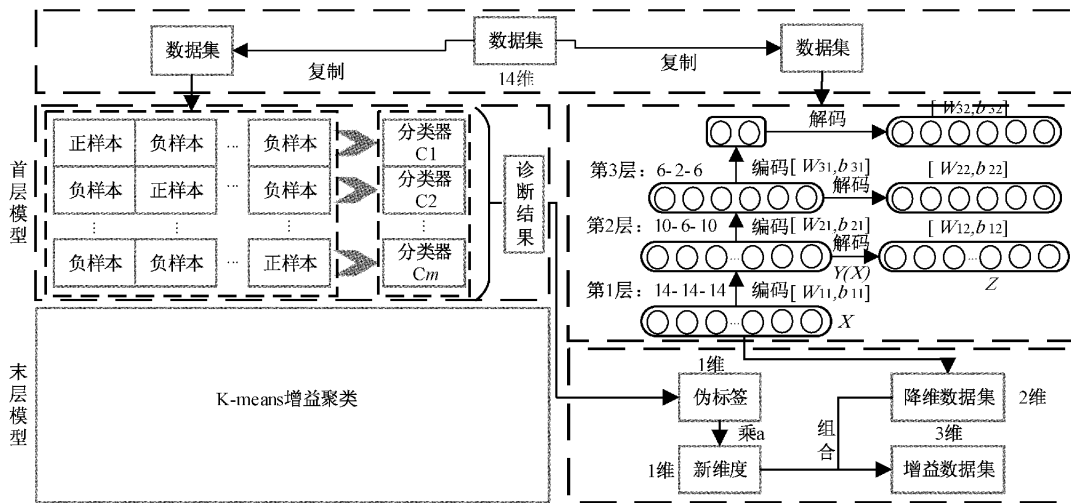


图 3 双层故障诊断模型

具体步骤如下:

1) 选择以特征气体加无编码比值法构建的 14 维向量作为降维模型 SAE 和首层模型 GBDT 的特征向量;

2) 数据集包含少量的有标签样本和大量的无标签样本,使用少量的有标签样本训练 GBDT,然后进行无标签样本的伪标签预测,预测标签记为  $\eta$ ;

3) 使用 SAE 对特征向量进行降维处理,构建能够表征故障信息的低维特征向量;

4) 伪标签  $\eta$  乘以系数  $\alpha$  形成新维度标签  $r$ ,低维特征向量与新维度标签  $r$  组合形成的 3 维向量作为末层模型特征向量;

5) 低维特征向量构成的降维数据集和新维度标签构成的新维度向量合成增益数据集,进行 K-means 增益聚类。

#### 4 算例分析

采用新疆送变电有限公司公布的案例数据和硕博文

献的附录数据作为数据集<sup>[17]</sup>,共计 2 141 组,其中无标数据 1 601 组,有标数据 540 组。样本分布如表 1 所示。采用 PyCharm 作为软件仿真平台,选取 Python 3. 9. 7 作为解释器,设定的各分类器超参数如表 2 所示。

表 1 样本分布

设备状态	状态编码	有标签样本	无标签样本	样本总数
中低温过热	1	90	315	405
高温过热	2	90	293	383
局部放电	3	90	93	183
低能放电	4	90	205	295
高能放电	5	90	280	370
正常	6	90	415	505

表 2 分类器超参数设置及分类结果

分类器种类	参数设置	训练时间/s	测试时间/s	准确率/%
SVM	C=11. 3137, gamma=0. 7071	21. 83	0. 075	66. 12
DBN	structure:14-9-18-6	247. 65	0. 0245	73. 23
DT	max_depth=7, criterion='gini', max_leaf_nodes=59	145. 62	0. 012	77. 15
RandomForest	n_estimators=18, max_leaf_nodes=29	345. 21	0. 015	77. 72
AdaBoost	n_estimators=12, max_depth=8	124. 26	0. 023	83. 91
GBDT	n_estimators=21, max_depth=11, max_leaf_nodes=29	86. 49	0. 036	83. 22
XGBoost	max_depth=5, n_estimators=15, alpha=1, reg_lambda=4	145. 65	0. 045	52. 13
GBDT+K-means	n_estimators=40, max_depth=11, max_leaf_nodes=29, K=6, $\alpha=4. 09$	86. 75	0. 047	90. 47

原始数据集存在同状态样本同维度数值差异较大的问题,会影响后期 GBDT 无标签样本伪标签预测的有效性和 K-means 增益聚类的准确率,因此需要对 14 维特征向量构成的数据集进行 Z-Score 归一化。 $N$  表示数据集样本数,公式如下:

$$Z_i = \frac{x_i - \frac{1}{N} \sum_{i=1}^N x_i}{\sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \frac{1}{N} \sum_{i=1}^N x_i)^2}} \quad (10)$$

#### 4.1 GDBT 首层模型有效性分析

针对特征向量选取及利用 GBDT 作为所提模型首层获取无标签样本的伪标签的有效性,设计以下实验验证特征向量选取及 GBDT 方法的有效性。有标签样本分布如表 3 所示。

采集有标签故障数据 540 组,每类 90 组,随机选取 80 组作为训练集,剩余 10 组作为测试集,分别训练随机森林(random forest)、GBDT、Adaboost、XGBoost 和决策树(decision tree,DT)5 种模型,并横向比较 3 种特征向量优势——特征气体法、IEC 三比值法和特征气体加无编码比

表 3 有标签样本分布

状态编码	训练集	测试集
1	80	10
2	80	10
3	80	10
4	80	10
5	80	10
6	80	10

值法。分类器使用网格搜索优化超参数,使用决策树作为基分类器,用五折交叉验证法评估模型诊断性能。

如表 4 所示,在诊断精度方面,相较于 IEC 三比值法或特征气体法构建的特征向量,选择特征气体加无编码比值法构建特征向量具有更明显的优势;相较于其余 4 种分类器,GBDT 模型表现出更强的诊断性能。

#### 4.2 K-means 增益有效性分析

使用 SAE 对特征气体加无编码比值法构建的 14 维特征向量进行降维处理,构建可以表征变压器运行状态的 2 维特征向量。如图 4 所示,高能放电数据分布密集,其他

表 4 不同特征向量、诊断模型的诊断精度

%

方法	IEC 三比值法		特征气体法		特征气体加无编码比值法	
	训练精度	测试精度	训练精度	测试精度	训练精度	测试精度
Random Forest	82.08	63.72	88.05	66.37	91.59	73.45
Adaboost	99.78	58.41	99.78	66.37	99.78	74.34
XGBoost	85.84	60.18	90.93	70.80	95.13	74.34
DT	85.40	59.29	85.18	65.49	90.71	71.68
GBDT	99.78	58.41	99.78	67.26	99.78	75.22

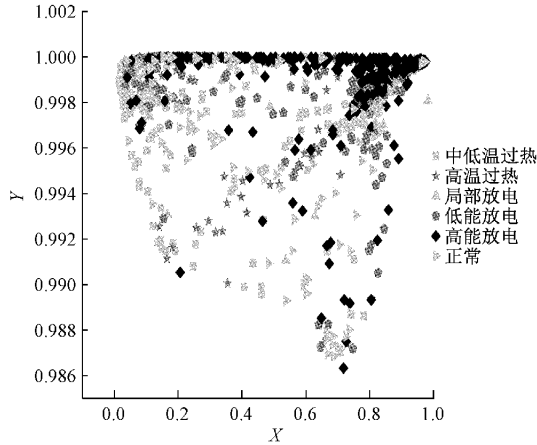


图 4 降维数据平面图

故障数据分布分散,因此,在进行 K-means 聚类之前,需要对数据进行进一步的处理。

本文所提出的双层故障诊断模型利用首层 GBDT 模型进行无标签样本的伪标签预测,伪标签  $\eta$  乘以系数  $\alpha$  形成新维度标签  $\gamma$ , 新维度标签  $\gamma$  构成 K-means 增益向量,降维数据集与 K-means 增益向量形成增益数据集,作为 K-means 增益聚类的输入数据集。

如图 5 所示,在 K-means 增益向量的引导下,降维数据集映射到新空间的数据分布更适合进行 K-means 聚类。

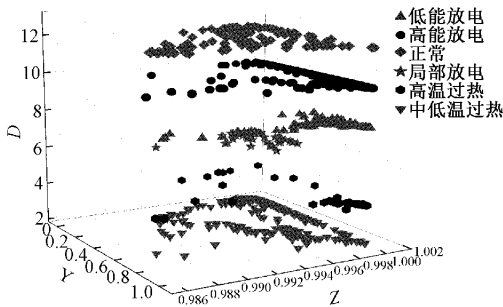


图 5 特征向量图 ( $\alpha=2$ )

系数  $\alpha$  的取值在一定程度上影响着增益数据集的空间分布,决定着最终 K-means 增益聚类的准确率。如图 6 所示,当  $\alpha = 4.09$  时,K-means 增益聚类诊断精度最高达到 90.47%。

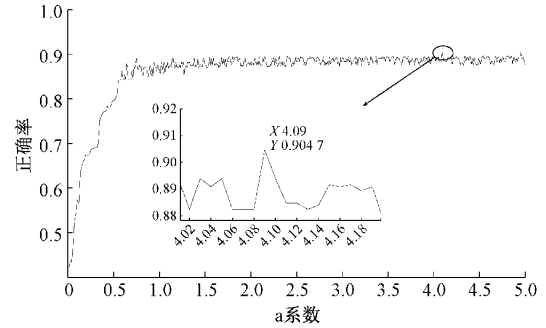


图 6 系数  $\alpha$  与聚类效果关系

### 4.3 模型诊断精度分析

选取 SVM、DBN、DT、Random Forest、Adaboost、GBDT、XGBoost 和 GBDT+K-means 8 种模型进行对比分析。

首先,验证各模型在油浸式变压器故障诊断中的精度和所耗时间,对于前 7 组分类器,将 540 组有标数据集作为训练集,1 301 组无标数据集作为测试集。超参数和诊断精度如表 2 所示。8 种模型故障诊断精度从高到低依次为 GBDT+K-means、AdaBoost、GBDT、Random Forest、DT、DBN、SVM 和 XGBoost;GBDT+K-means 的训练时间比 SVM 长,比其余 6 种分类器短;8 种分类器的测试时间表现出很小的差异。

其次,分析有标签数据量的占比对模型准确率的影响。分别选取有标签数据量为每类 20, 30, 40, 50, 60, 70, 80 和 90, 对 8 种诊断模型进行训练,得到诊断结果。诊断结果如图 7 所示,各分类器的诊断精度随着有标签数据量的增加呈现明显的上升趋势;在任何标签数据量下,GBDT 与 K-means 增益聚类的双层故障诊断模型均表现出最优的诊断性能。

最后,对比分析各模型查准率和查全率。查准率和查全率的公式如式(11)所示。

$$precision_i = \frac{num(predict(x \in S_i) \& real(x \in S_i))}{num(predict(x \in S_i))}$$

$$recall_i = \frac{num(predict(x \in S_i) \& real(x \in S_i))}{num(real(x \in S_i))}$$

$$i = 1, 2, 3, 4, 5, 6$$

(11)



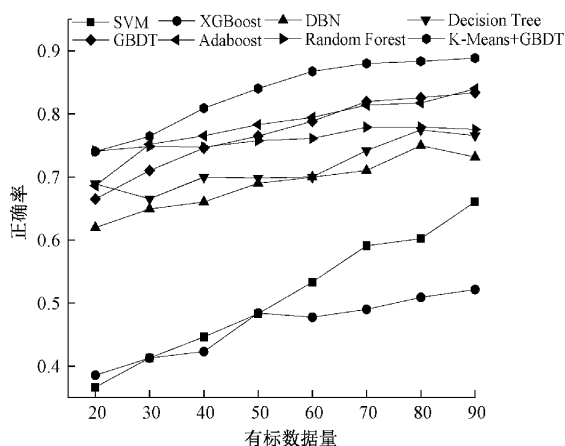


图 7 有标数据量对分类器预测准确率影响

如图 8,9 所示,本文所提模型在中低温过热故障、高温过热故障、局部放电故障、低能放电故障、高能放电故障和正常的查准率均高于对比模型;中低温过热故障的查全率低于 SVM 模型,其余状态查全率均高于对比模型。

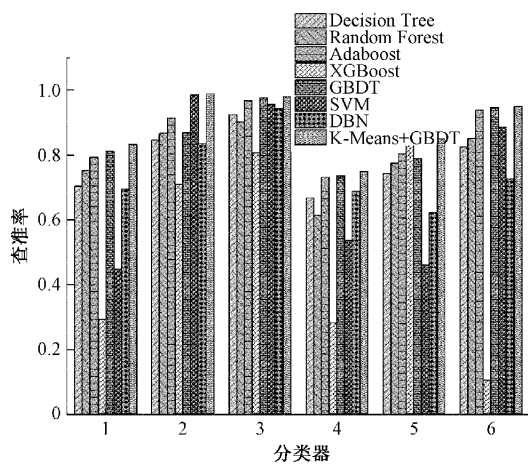


图 8 多模型查准率对比

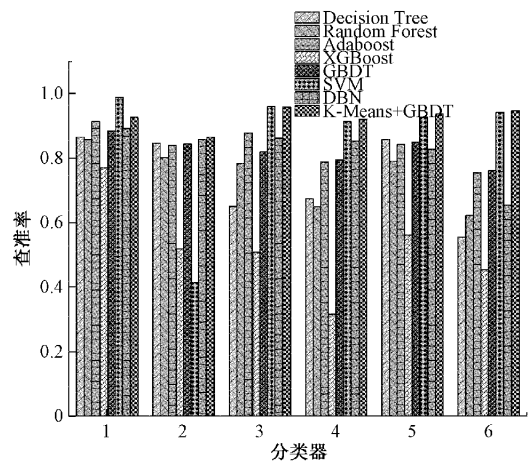


图 9 多模型查全率对比

## 5 结 论

针对油浸式变压器数据样本标记困难,有标签样本数据量较少,传统方法诊断精度低的问题,本文提出了一种 GBDT 与 K-means 增益聚类的双层故障诊断模型,与传统方法相比,具有如下优势:

引入 SAE 模型对以特征气体加无编码比值构建的特征向量进行降维,消除了特征维度间的信息冗余,实现了数据聚类的可视化;

使用 GBDT 进行无标签样本的伪标签预测,利用伪标签构建增益向量,降维数据集在增益向量的引导下,映射到新空间的数据分布更适合进行 K-means 聚类;

与传统方法相比,GBDT 与 K-means 增益聚类的双层故障诊断模型具有诊断精度高,诊断时间相对少,诊断效果稳定等优势。

本文为少标签下的油浸式变压器故障诊断提供了新的思路。下一步将基于油浸式变压器同种故障之间的内在联系,先后融合降维算法、核函数、数据增广和聚类算法,构建无标签数据下的油浸式变压器无监督故障诊断模型。

## 参考文献

- [1] 江秀臣,刘亚东,傅晓飞,等. 输配电设备泛在电力物联网建设思路与发展趋势[J]. 高电压技术,2019,45(5): 1345-1351.
- [2] SOUAHLIA S, BACHA K, CHAARI A. MLP neural network-based decision for power transformers fault diagnosis using an improved combination of Rogers and Doernenburg ratios DGA[J]. International Journal of Electrical Power and Energy Systems, 2012, 43(1): 1346-1353.
- [3] 汪可,李金忠,张书琦,等. 变压器故障诊断用油中溶解气体新特征参量[J]. 中国电机工程学报, 2016, 36(23): 6570-6578.
- [4] ABDELAZIZ L, FOUAD T. Bayesian duval triangle method for fault prediction and assessment of oil immersed transformers[J]. Measurement & Control, 2017, 50(4): 77-85.
- [5] 宋斌,刘志雄,李恩文,等. DGA 中三比值缺码研究初探[J]. 电力自动化设备, 2015, 35(12): 60-65.
- [6] 孙琛. 基于机器学习算法的变压器故障诊断[D]. 上海: 上海交通大学, 2019.
- [7] 王雨虹,王志中,付华,等. 多策略改进麻雀算法与 BiLSTM 的变压器故障诊断研究[J]. 仪器仪表学报, 2022, 43(3): 87-97.
- [8] 代杰杰,宋辉,杨祎,等. 基于油中气体分析的变压器故障诊断 ReLU-DBN 方法[J]. 电网技术, 2018, 42(2): 658-664.

- [9] 李春茂,周妹末,高波,等. 基于邻域粗糙集与多核支持向量机的变压器多级故障诊断[J]. 高电压技术, 2018, 44(11):3474-3482.
- [10] 单亚峰,段金凤,付华,等. 基于 SSA-AdaBoost-SVM 的变压器故障诊断[J]. 控制工程, 2022, 29(2):280-286.
- [11] 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.
- [12] 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.
- [13] 黄新波,王享,田毅,等. 基于 PSO-ELM 融合动态加权 AdaBoost 的变压器故障诊断方法[J]. 高压电器, 2020, 56(5):39-46.
- [14] 曹伟嘉,杨留方,徐天奇,等. 基于精英反策略麻雀搜索优化随机森林的变压器故障诊断[J]. 国外电子测量技术, 2022, 41(2):138-143.
- [15] 李黄曼,张勇,张瑶. 基于 ISSA 优化 SVM 的变压器故障诊断研究[J]. 电子测量与仪器学报, 2021, 35(3):123-129.
- [16] 周平,马景义. 基于路径跟随方法的光滑子区间 K 均值聚类算法[J]. 统计与决策, 2022, 38(12):17-22.
- [17] 段效琛. 基于 DGA 的电力变压器多分类模型与故障诊断研究[D]. 昆明:昆明理工大学, 2017.

#### 作者简介

汤茂祥, 硕士, 主要研究方向为大数据下智能故障诊断与状态监测的研究。

E-mail: 2539036355@qq.com