

基于深度学习的变压器故障诊断方法研究

杨涛, 黄军凯, 许逵, 吴建蓉, 陈仕军

(贵州电网有限责任公司电力科学研究院, 贵州 贵阳 550002)

摘要:油中溶解气体分析方法(DGA)是变压器内部故障诊断的重要方法,广泛应用于变压器在线监测和定期试验检测中,传统的特征气体法和三比值法等诊断方法在实际应用中普遍存在着一定的局限性,导致故障诊断精度偏低。针对这一问题,本文提出了一种基于深度学习技术中的多层感知机的变压器故障综合诊断方法,利用开源的Scikit-learn机器学习框架及TensorFlow深度学习框架构建了变压器故障诊断模型,并应用实际工程中的故障样本数据,对故障诊断模型进行了训练和测试。试验结果表明,基于多层感知机技术的变压器故障诊断模型能够对变压器故障进行正确诊断,与传统的三比值法及支持向量机技术相比,多层感知机的诊断准确率更高,具有更优的故障诊断性能,能够为变压器的检修提供更为准确的参考信息。

关键词:深度学习;油中气体分析;变压器;故障诊断;多层感知机

文章编号:2096-4633(2018)06-0023-08 **中图分类号:**TM42 **文献标志码:**B

变压器是电力系统中最重要设备之一,其运行状态直接影响到电力系统的安全可靠运行,变压器的任何故障都可能导致电力系统局部故障甚至整个系统瘫痪,影响用户正常生产生活供电,造成巨大经济损失。因此,对变压器故障的早期诊断是至关重要的^[1-3]。

运行中的变压器在电、热的作用下,其内部的油和纸等绝缘材料会逐步发生老化和分解,产生氢气、烃类、一氧化碳和二氧化碳等气体,这些溶解在油中的气体成份可以反映变压器内部绝缘老化、放电和发热等故障^[4]。油中溶解气体分析(dissolved gas-in-oil analysis, DGA),通过分析油浸式变压器绝缘油中甲烷(CH₄)、乙烷(C₂H₆)、乙烯(C₂H₄)、乙炔(C₂H₂)、氢气(H₂)、一氧化碳(CO)和二氧化碳(CO₂)等气体的含量来判断变压器的运行状态,基本不受外界环境噪声干扰影响,且可以在不停电的情况下开展,是发现油浸式变压器潜伏性故障最有效的措施之一,广泛应用于变压器在线监测和定期试验检测之中^[5]。

基于DGA的变压器故障诊断方法主要分为传统的故障诊断方法和智能诊断方法。传统诊断方法主要有特征气体法、气体含量比值法和气体比值图示法等^[6],这些方法操作简单易行,在实际工作中得到了广泛应用,但普遍存在故障诊断准确率偏低的缺点。智能诊断方法结合人工智能技术,主要包

括支持向量机^[7-8]、贝叶斯网络^[9-10]、专家系统^[11-12]、模糊理论^[13-14]、BP神经网络^[15-17]、极限学习机^[18]等。这些智能诊断技术不同程度上提高了变压器故障诊断的准确性,但也存在各自的问题,如专家系统需要在实际中比较难以获取丰富的专家知识和经验;支持向量机本质上是二分类算法,在多分类问题上存在分类重叠和不可分类的情况;BP神经网络存在着收敛速度慢、易陷入局部最优等问题。

本文在分析和研究现有变压器故障诊断方法的基础上,结合最新的人工智能算法,提出了一种基于深度学习方法的变压器故障诊断新方法。

1 深度学习

深度学习概念中的“深度”,是相对传统机器学习算法支持向量机(SVM)、BP神经网络、最大熵等“浅层学习”方法而言的。深度学习方法中,非线性算法的层级数更多。传统的浅层学习方法,需要人工提取数据样本特征,模型学习后所得到的是没有层次结构的单层特征,而深度学习通过对原始数据样本进行逐层特征变换,自动学习得出数据样本的层次化特征表示,更有利于数据分类的实现^[19]。深度学习技术是当前机器学习领域一个新的研究方向,在语音识别、图像识别、自然语言处理等领域取得了突破性进展,目前在电力领域的应用还处于刚

刚起步阶段。

深度学习的概念由 G. E. Hinton 等人于 2006 年提出,指基于样本数据通过一定的训练方法得到包含多个层级的深度网络结构的机器学习过程^[20]。深度学习技术所采用的深度网络包含大量的单一神经元,每个神经元与大量其他神经元相连接,神经元间的连接强度(权值)在训练过程中进行调整并决定了网络的功能。通过深度学习所得到的深度网络本质上就是深层次的人工神经网络,即深度神经网络(deep neural networks, DNN)^[21]。

多层感知机(multilayer perceptron, MLP),也叫深度前馈网络(deep feed forward network),是深度神经网络中一种典型的最基本的模型,其结构如图 1 所示。前馈网络的目标是近似某个函数 f^* 。例如,对于分类器, $y = f^*(x)$ 将输入 x 映射到一个类别 y 。前馈网络定义了一个映射 $y = f(x; \theta)$, 并且学习参数 θ 的值,使它能够得到最佳的函数近似。

多层感知机包含一个输入层、多个隐藏层和一个输出层,图中每个节点代表一个神经元,其输入输出关系可由式(1)表示。

$$y = h\left(\sum_j w_j x_j + b\right) \quad (1)$$

式中, y 为神经元的输出值, x_j 为神经元的输入值, $h(x)$ 是激活函数(常用 Sigmoid 函数, tanh 函数和 ReLU 函数), w_j 为权重值, b 为偏置值。单个神经元的输出值为输入值的加权和与偏置值之和,经过激活函数进行非线性变换后的值。

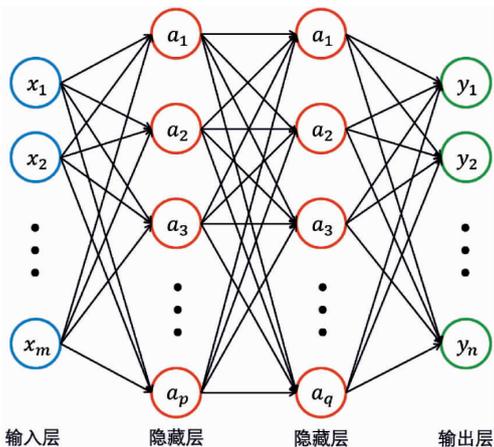


图 1 多层感知机结构示意图

Fig. 1 Schematic diagram of multilayer perceptron

多层感知机的流程分为前向过程和反向过程。在前向过程中,输入层作为第一层隐藏层的

输入,第一层隐藏层的输出作为第二层隐藏层的输入,依此类推,得到一个多层的神经网络结构,最后输出层得到的输出值就是所得到的预测结果。多层感知机的训练过程是一种监督学习过程,训练所用样本是已知输出值的有标签样本。首先对神经网络中的权重值、偏置值进行随机初始化,将样本的特征值作为神经网络的输入,经过神经网络的前向计算,得到输出值,然后根据此输出值与样本真实输出值的误差,逐步反向调整各层神经网络的权重值和偏置值,完成调整后重新计算网络输出值并得到输出误差。如此反复,直到输出误差满足一定要求,此时神经网络中各个神经元通过自我学习得到了合适的权重和偏置值,神经网络通过学习自动得出了输入到输出之间的关系,也即完成了神经网络的训练过程。以下给出前向传播和反向传播过程的详细数学表示。

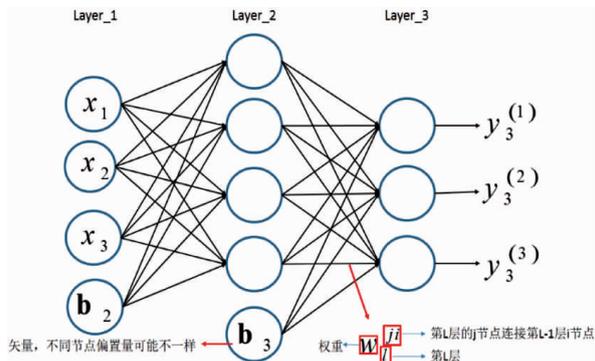


图 2 多层感知机前向传播过程

Fig. 2 Forward broadcasting process of multilayer perceptron

前向传播指的是信息从第一层逐渐地向高层进行传递的过程。以图 2 为例来进行前向传播的过程的分析。假设第一层为输入层,输入的信息为 $[x_1, x_2, x_3]$ 。对于层 l ,用 L_l 表示该层的所有神经元,其输出为 y_l ,其中第 j 个节点的输出为 $y_l^{(j)}$,该节点的输入为 $u_l^{(j)}$,连接第 l 层与第 $(l-1)$ 层的权重矩阵为 W_l ,上一层(第 $(l-1)$ 层)的第 i 个节点到第 l 层第 j 个节点的权重为 $w_l^{(ji)}$ 。可得到:

$$\begin{cases} y_2^{(1)} = f(u_2^{(1)}) = f\left(\sum_{i=1}^n w_2^{(1i)} x_i + b_2^{(1)}\right) \\ y_2^{(2)} = f(u_2^{(2)}) = f\left(\sum_{i=1}^n w_2^{(2i)} x_i + b_2^{(2)}\right) \\ y_2^{(3)} = f(u_2^{(3)}) = f\left(\sum_{i=1}^n w_2^{(3i)} x_i + b_2^{(3)}\right) \end{cases}$$

转化为矩阵表达式有:

$$y_2 = \begin{bmatrix} y_2^{(1)} \\ y_2^{(2)} \\ y_2^{(3)} \end{bmatrix} = f \left\{ \begin{bmatrix} w_2^{(11)} & w_2^{(12)} & w_2^{(13)} \\ w_2^{(21)} & w_2^{(22)} & w_2^{(23)} \\ w_2^{(31)} & w_2^{(32)} & w_2^{(33)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} b_2^{(1)} \\ b_2^{(2)} \\ b_2^{(3)} \end{bmatrix} \right\} = f(W_2 X + b_2)$$

推广到网络中的任一层,有:

$$\begin{cases} y_l^{(j)} = f(u_l^{(j)}) \\ u_l^{(j)} = \sum_{i=1}^n w_l^{(ji)} y_{l-1}^{(i)} + b_l^{(j)} \\ y_l = f(ul) = f(W_l y_{l-1} + b_l) \end{cases}$$

其中, f 为激活函数, $b_l^{(j)}$ 为第 l 层的第 j 个偏置值。这样,就完成了多层感知机的前向传播过程。

在模型搭建完成后,训练过程所做的就是完成模型参数的更新。由于多层感知机存在的多层结构,无法直接对中间的隐层利用损失值来进行参数

$$\begin{cases} \frac{\partial E}{\partial w_l^{(ji)}} = \frac{\partial E}{\partial y_l^{(j)}} \frac{\partial y_l^{(j)}}{\partial w_l^{(ji)}} = \frac{\partial E}{\partial y_l^{(j)}} \frac{\partial y_l^{(j)}}{\partial u_l^{(j)}} \frac{\partial u_l^{(j)}}{\partial w_l^{(ji)}} = \frac{\partial E}{\partial y_l^{(j)}} f'(u_l^{(j)}) y_{l-1}^{(i)} \\ \frac{\partial E}{\partial b_l^{(j)}} = \frac{\partial E}{\partial y_l^{(j)}} \frac{\partial y_l^{(j)}}{\partial b_l^{(j)}} = \frac{\partial E}{\partial y_l^{(j)}} \frac{\partial y_l^{(j)}}{\partial u_l^{(j)}} \frac{\partial u_l^{(j)}}{\partial b_l^{(j)}} = \frac{\partial E}{\partial y_l^{(j)}} f'(u_l^{(j)}) \end{cases}$$

由于下一层所有节点的输入都与前一层的每个节点输出有关,因此损失函数可以认为是下一层的每个神经元节点输入的函数,则有:

$$\begin{aligned} \frac{\partial E}{\partial y_l^{(j)}} &= \frac{\partial E(u_{l+1}^{(1)}, u_{l+1}^{(2)}, \dots, u_{l+1}^{(K)})}{\partial y_l^{(j)}} \\ &= \sum_{k \in L_{l+1}} \frac{\partial E}{\partial u_{l+1}^{(k)}} \frac{\partial u_{l+1}^{(k)}}{\partial y_l^{(j)}} \\ &= \sum_{k \in L_{l+1}} \frac{\partial E}{\partial y_{l+1}^{(k)}} \frac{\partial y_{l+1}^{(k)}}{\partial u_{l+1}^{(k)}} \frac{\partial u_{l+1}^{(k)}}{\partial y_l^{(j)}} \\ &= \sum_{k \in L_{l+1}} \frac{\partial E}{\partial y_{l+1}^{(k)}} \frac{\partial y_{l+1}^{(k)}}{\partial u_{l+1}^{(k)}} w(kj)_{l+1} \end{aligned}$$

定义节点的灵敏度为误差对输入的变化率,有:

$$\delta = \frac{\partial E}{\partial u}$$

那么第 l 层第 j 个节点的灵敏度为:

$$\begin{aligned} \delta(j)_i &= \frac{\partial E}{\partial u_l^{(j)}} \\ &= \frac{\partial E}{\partial y_l^{(j)}} \frac{\partial y_l^{(j)}}{\partial u_l^{(j)}} \\ &= \frac{\partial E}{\partial y_l^{(j)}} f'(u_l^{(j)}) \end{aligned}$$

更新,因此需要通过损失值从顶层到底层的反向传播来进行参数的更新。

假设多层感知机用于分类,在输出层有多个神经元,每个神经元对应一个标签。输入样本为 $x = [x_1, x_2, \dots, x_n]$, 其标签为 t ; 对于网络的最后一层输出层第 k 层,定义损失函数为:

$$E = \frac{1}{2} \sum_{j \in L_k} (t^{(j)} - y_k^{(j)})^2$$

为了极小化损失函数,通过梯度下降来进行推导,得到:

于是可得到:

$$\begin{aligned} \frac{\partial E}{\partial y_l^{(j)}} &= \sum_{k \in L_{l+1}} \frac{\partial E}{\partial y_{l+1}^{(k)}} \frac{\partial y_{l+1}^{(k)}}{\partial u_{l+1}^{(k)}} w_{l+1}^{(kj)} \\ &= \sum_{k \in L_{l+1}} \delta_{l+1}^{(k)} w_{l+1}^{(kj)} \end{aligned}$$

两边同时乘上 $f(u_l^{(j)})$, 则有:

$$\begin{aligned} \delta(j)_i &= \frac{\partial E}{\partial y_l^{(j)}} f'(u_l^{(j)}) \\ &= f'(u_l^{(j)}) \sum_{k \in L_{l+1}} \delta_{l+1}^{(k)} w_{l+1}^{(kj)} \end{aligned}$$

对于输出层,并不满足上式,但输出层的输出直接和误差相关,因此可以用损失函数来直接求取偏导数,那么:

$$\begin{aligned} \delta_l^{(j)} &= \frac{\partial E}{\partial y_l^{(j)}} f'(u_l^{(j)}) \\ &= \begin{cases} f'(u_l^{(j)}) \sum_{k \in L_{l+1}} \delta_{l+1}^{(k)} w_{l+1}^{(kj)} & l \text{ 层为隐层} \\ f'(u_l^{(j)}) (y_l^{(j)} - t^{(j)}) & l \text{ 层为输出层} \end{cases} \end{aligned}$$

可得到损失函数对各参数的梯度为:

$$\begin{cases} \frac{\partial E}{\partial w_l^{(ji)}} = \frac{\partial E}{\partial u_l^{(j)}} \frac{\partial u_l^{(j)}}{\partial w_l^{(ji)}} = \delta_l^{(j)} y_{l-1}^{(i)} \\ \frac{\partial E}{\partial b_l^{(j)}} = \frac{\partial E}{\partial u_l^{(j)}} \frac{\partial u_l^{(j)}}{\partial b_l^{(j)}} = \delta_l^{(j)} \end{cases}$$

最后可得到在每一次训练过程中,参数的更新公式为:

$$\begin{cases} w_l^{(j)} := w_l^{(j)} - \eta \delta_l^{(j)} y_{l-1}^{(i)} \\ b_l^{(j)} := b_l^{(j)} - \eta \delta_l^{(j)} \end{cases}$$

式中, η 为学习速率。通过这个式子不断迭代,直到梯度的幅值接近于 0,就完成了模型参数的学习。

传统的神经网络算法在训练层数较多多层神经网络时,会遇到网络训练中的梯度消失和收敛到局部最小等问题,导致训练结果不够理想。近年来,学者提出了一系列解决深度神经网络训练问题的措施,如优化的参数初始化方法、正则化、提前终止、Dropout 方法等,使得训练多层深度神经网络成为了可能。

2 机器学习及深度学习框架

2.1 机器学习框架

机器学习(machine learning, ML)是一门多领域交叉学科,涉及概率论、统计学、逼近论、凸分析、算法复杂度理论等多门学科。专门研究计算机怎样模拟或实现人类的学习行为,以获取新的知识或技能,重新组织已有的知识结构使之不断改善自身的性能。

从广义上讲,机器学习是一种能够赋予机器学习的能力,让它以此完成直接编程无法完成的功能的方法。但从实践的意义上讲,机器学习是一种通过利用数据,训练出模型,然后使用模型预测的一种方法。

Scikit-learn 是基于 Python 语言的一个开源机器学习框架,实现了包含支持向量机、随机森林、k-均值聚类在内的多种分类、回归、聚类、模型选择和数据预处理等算法,通过调用相关模块,可以快速实现机器学习算法。Scikit-learn 在机器学习算法和研究和应用方面得到了广泛应用。

2.2 深度学习框架

随着深度学习研究的不断深入,许多机构都开发出了各种开源的深度学习框架,其中包括 TensorFlow、Caffe、Keras、CNTK、Torch7、MXNet、Leaf、Theano、DeepLearning4、Lasagne、Neon,等等。其中 Google 公司开发的 TensorFlow 在这些框架中异军突起,在研究领域和生产领域都得到了广泛

的应用,是目前最流行的深度学习开发框架,有快速、灵活并适合产品级大规模应用等特点,极大简化了深度学习项目的开发过程,并具有极高的计算性能。

TensorFlow 是相对高阶的机器学习库,用户可以通过它方便地设计和实现神经网络结构,而不必亲自编写底层的 C++ 或 CUDA 代码。其底层采用 C++ 实现,保证了运行效率,而在上层提供了 Python 接口,方便用户迅速编写测试机器学习算法。

TensorFlow 支持常见的深度学习神经网络结构,如多层感知机(MLP)、卷积神经网络(CNN)和循环神经网络(RNN)。TensorFlow 支持跨操作系统平台,可运行在 Windows、Linux 和 MacOS 平台上,甚至移动设备的 iOS 和 Android 平台上,同时还支持多种硬件平台,支持 Intel 和 AMD 的 CPU,同时还可支持通过 Nvidia 的 GPU 进行加速。

3 基于深度学习的变压器故障诊断方法

3.1 选取样本数据

选取 IECTC 10 数据集及部分工程实际中的变压器油中气体数据共 333 条作为样本,包含有 6 种故障类型和正常状态,其中局部放电样本 28 条,低能放电样本 37 条,高能放电样本 88 条,低温过热样本 43 条,中温过热样本 33 条,高温过热样本 59 条,正常样本 45 条,如表 1 所示,其中部分样本值如表 2 所示。将样本拆分为 2 个部分,其中 266 条作为训练集,用来训练神经网络,67 条作为测试集,用来评估系统性能。

表 1 样本故障类型

Tab. 1 Fault type in the samples

故障类型	样本数量	故障代码
局部放电	28	PD
低能放电	37	D1
高能放电	88	D2
低温过热 $T < 300^{\circ}\text{C}$	43	T1
中温过热 $300 < T < 700^{\circ}\text{C}$	33	T2
高温过热 $T > 700^{\circ}\text{C}$	59	T3
正 常	45	No

表2 部分样本值
Tab. 2 Part of the samples

H ₂	CH ₄	C ₂ H ₆	C ₂ H ₄	C ₂ H ₂	CO	CO ₂	故障代码
93	19	14	2	0	109	0	No
137	18	16	2	0	119	0	PD
44	23	11	7	0	319	0	No
0	385	92	714	0	265	0	T3
35	551	219	344	0	393	0	T2
23	412	102	771	0	304	0	T3
22	443	110	745	0	312	0	T3
18	461	194	263	0	454	0	T2
19	428	176	235	0	333	0	T2
15	15	4	17	22	82	0	D1
22	443	110	745	0	312	0	T3
15	365	97	585	0	293	0	T3
12	338	95	550	0	261	0	T3
14	322	85	474	0	243	0	T3
7 855	7 450	899	8 564	54	205	0	D2
1 691	730	202	837	726	1 499	0	D2
54	15	0	0	0	61	0	No
3 635	1 490	146	1 784	4 127	898	0	D2
75	33	12	15	16	252	0	D1

3.2 特征值选取

考虑油中气体成分与变压器故障之间的关系特点,选取 H₂、CH₄、C₂H₆、C₂H₄、C₂H₂这5种气体组分作为特性变量。为了使特征变量值更符合神经网络的输入需要,按照式(2)对数据进行规范化处理。

$$x_{scaled} = \frac{x - x_{mean}}{x_{std}} \quad (2)$$

式中, x_{scaled} 为规范化处理后的气体含量数据, x 为气体含量原始值, x_{mean} 为气体含量平均值, x_{std} 为气体含量标准。

3.3 变压器故障诊断模型

```

from sklearn.model_selection import StratifiedShuffleSplit
from sklearn.preprocessing import LabelEncoder
from sklearn.preprocessing import StandardScaler
encoder = LabelEncoder()
split = StratifiedShuffleSplit(n_splits=1, test_size=0.2, random_state = 22)
for train_index, test_index in split.split(dga_data, dga_data['Fault Type']):
    dga_train_set = dga_data.loc[train_index]
    dga_test_set = dga_data.loc[test_index]
X_train = dga_train_set[['H2', 'CH4', 'C2H6', 'C2H4', 'C2H2']]
y_train = dga_train_set['Fault Type']
X_test = dga_test_set[['H2', 'CH4', 'C2H6', 'C2H4', 'C2H2']]
y_test = dga_test_set['Fault Type']
y_train = encoder.fit_transform(y_train)
y_test = encoder.fit_transform(y_test)
scaler = StandardScaler()
X_train_scaled = scaler.fit_transform(X_train)
X_test_scaled = scaler.transform(X_test)
    
```

图3 变压器故障诊断模型源码

采用前节所述的多层状态感知机作为变压器故障诊断模型,利用 Scikit-learn 机器学习框架和 TensorFlow 深度学习框架进行了实现,部分代码如图3所示。

多层状态感知机主要的参数包括输入层节点数量、输出层节点数量、隐藏层数和隐藏层节点数量。在本文中,输入参数为5种特征气体含量,输出参数为6种故障类型的概率,因此,输入层取5个节点,输出层取6个节点。所需调节的参数为隐藏层数和隐藏层节点数量(图4)。

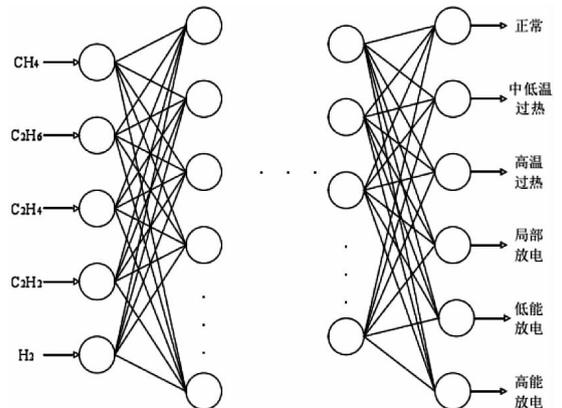


图4 基于多层状态感知机的变压器故障诊断模型

Fig. 4 Transformer fault diagnosis model based on multilayer perceptron

Fig. 3 Source code of the transformer diagnosis model

分别测试了每个隐藏层 50 节点,隐藏层数从 1 到 15 时系统在训练集和测试集上的正确率,如图 5 所示。

从图 5 曲线可看出,当隐藏层数小于 6 层时,训练正确率随层数增加而增大,隐藏层数大于 6 层时,训练正确率随层数增加缓慢下降。测试正确率在隐藏层为 10 层时达到最大,为 85%。

表 3 给出了多层状态感知机取不同隐藏层数及不同节点数时在训练集和测试集上的准确率。综合系统的训练正确率和测试正确率,并考虑系统的计算复杂度,在系统取 3 层隐藏层,每层 50 个节点时的模型是此数据集上的最优模型。

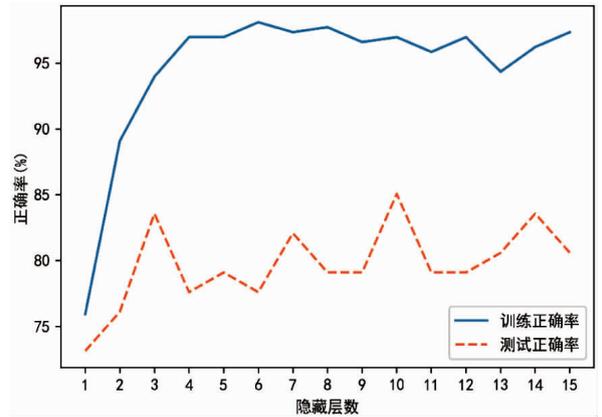


图 5 不同网络层数时系统正确率

Fig. 5 System accuracy at different network layers

表 3 不同层数和节点数下模型准确率

Tab. 3 Model accuracy at different layers and nodes

隐藏层数	各隐藏层节点数	训练集准确率/%	测试集准确率/%
1	50	75.9	73.1
1	100	78.2	76.1
2	50 - 50	89.1	76.1
2	100 - 100	91.0	77.6
3	30 - 30 - 30	92.9	82.1
3	50 - 50 - 50	92.9	83.6
3	100 - 100 - 100	95.4	79.1
4	50 - 50 - 50 - 50	97.0	79.1
6	50 - 50 - 50 - 50 - 50 - 50	98.1	77.6
10	50 - 50 - 50 - 50 - 50 - 50 - 50 - 50 - 50 - 50	97.0	85.1

作为对比,在相同的数据测试集上,还应用改良三比值法和支持向量机(SVM)两种故障诊断方法进行了试验,试验结果如表 4 所示。

表 4 基于三比值法、支持向量机及多层状态感知机的变压器故障诊断情况

Tab. 4 Transformer fault diagnosis based on three ratios method, support vector machine and multi-layer perceptron

方法	测试集正确率/%
三比值法	65.0
SVM	70.1
MLP	83.6

通过表 4 的结果对比可知,本文所采用的多层感知机深度学习方法,与三比值法和 SVM 方法相比,具有更高的故障诊断准确率,可更准确地为运行检修人员提供诊断参考信息。

测试结果还表明,当增加训练数据集中样本数量时,模型的故障诊断准确率会有明显提高。因此,可通过在模型训练数据集中增加实际变压器故障样本数据来提高模型的故障诊断性能^[22-24]。

4 结论

本文构建了基于多层感知机的油中气体分析故障诊断模型,并用工程实际样本数据对模型进行了训练和测试。

(1)测试结果表明,多层感知机模型能够通过

油中气体数据,对变压器故障进行正确诊断。

(2) 多层感知机比传统三比值法及支持向量机法具有更高的诊断准确率,故障诊断性能更优。

参考文献:

- [1] 张霖,杨超,李诗勇,等. 一起500千伏变压器多维状态监测与故障诊断[J]. 电力大数据,2018,21(02):67-73.
ZHANG Lin, YANG Chao, LI Shiyong, et al. Multi-dimensional state monitoring and fault diagnosis of 500 kV Transformer [J]. Power Systems and Big Data, 2018, 21(02): 67-73.
- [2] BACHA K, SOUAHLIA S, GOSSA M. Power transformer fault diagnosis based on dissolved gas analysis by support vector machine [J]. Electric Power Systems Research, 2012, 83(01):73-79.
- [3] SHINTEMIROV A, TANG W, WU Q H. Power transformer fault classification based on dissolved gas analysis by implementing bootstrap and genetic programming [J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 2009, 39(01):69-79.
- [4] 姜有泉,黄良,王波,等. 基于DGA和深度置信网络的变压器内部故障诊断[J]. 武汉大学学报(工学版), 2017, 50(05):749-753.
JIANG Youquan, HUANG Liang, WANG Bo, et al. Transformer internal fault diagnosis based on DGA and deep belief network [J]. Engineering Journal of Wuhan University, 2017, 50(05):749-753.
- [5] 石鑫,朱永利. 深度学习神经网络在电力变压器故障诊断中的应用[J]. 电力建设, 2015, 36(12):117-122.
SHI Xin, ZHU Yongli. Application of Deep Learning Neural Network in Fault Diagnosis of Power Transformer [J]. Electric Power Construction, 2015, 36(12):117-122.
- [6] 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, 2014.
- [7] BENMAHAMED Y, TEGUAR M, BOUBAKEUR A. Application of SVM and KNN to Duval Pentagon I for transformer oil diagnosis [J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2017, 24(06):3443-3451.
- [8] ZHENG H B, LIAO R J, GRZYBOWSKI S, et al. Fault diagnosis of power transformers using multi-class least square support vector machines classifiers with particle swarm optimisation [J]. IET Electric Power Applications, 2011, 5(09):691-696.
- [9] 白翠粉,高文胜,金雷,等. 基于3层贝叶斯网络的变压器综合故障诊断[J]. 高电压技术, 2013, 29(02):330-335.
Bai Cuifen, GAO Wensheng, JIN lei, et al. Integrated Diagnosis of Transformer Faults Based on Three-layer Bayesian Network [J]. High Voltage Engineering, 2013, 29(02):330-335.
- [10] ROLIM J G, MAIOLA P C, BAGGENSTOSS H R, et al. Bayesian Networks Application to Power Transformer Diagnosis [C]. 2007 IEEE Lausanne Power Tech, 2007:999-1004.
- [11] LIN C E, LING J M, HUANG C L. An expert system for transformer fault diagnosis using dissolved gas analysis [J]. IEEE Transactions on Power Delivery, 1993, 8(01):231-238.
- [12] 师瑞峰,史永锋,牟军,等. 油中溶解气体电力变压器故障诊断专家系统[J]. 电力系统及其自动化学报, 2014, 26(12):49-54.
SHI Ruifeng, SHI Yongfeng, MOU Jun, et al. Power transformer fault diagnosis expert system with dissolved gas analysis in oil [J]. Proceedings of the CSU-EPSA, 2014, 26(12):49-54.
- [13] ATALAR F, KUNTMAN A. Diagnosis of power transformers faults using fuzzy logic method [C]. 2016 National Conference on Electrical, Electronics and Biomedical Engineering (ELECO), 2016:28-32.
- [14] NIŦU M C, ACIU A M, NICOLA C I, et al. Power transformer fault diagnosis using fuzzy logic technique based on dissolved gas analysis and furan analysis [C]. 2017 International Conference on Optimization of Electrical and Electronic Equipment (OPTIM) & 2017 Intl Aegean Conference on Electrical Machines and Power Electronics (ACEMP), 2017:184-189.
- [15] SUN Y, ZHANG S, MIAO C, et al. Improved BP neural network for transformer fault diagnosis [J]. Journal of China University of Mining & Technology, 2007, 17(01):138-142.
- [16] ZHANG Y, CHEN E, GUO P J, et al. Application of improved particle swarm optimization BP neural network in transformer fault diagnosis [C]. 2017 Chinese Automation Congress (CAC), 2017:6971-6975.
- [17] ZHANG Y, DING X, LIU Y, et al. An artificial neural network approach to transformer fault diagnosis [J]. IEEE Transactions on Power Delivery, 1996, 11(04):1836-1841.
- [18] 苑津莎,张利伟,王瑜,等. 基于极限学习机的变压器故障诊断方法研究[J]. 电测与仪表, 2013, (12):21-26.
YUAN Jinsha, ZHANG Liwei, WANG Yu, et al. Study of transformers fault diagnosis based on extreme learning machine [J]. Electrical Measurement & Instrumentation, 2013, (12):21-26.
- [19] GOODFELLOW I, BENGIO Y, COURVILLE A. Deep learning [M]. Massachusetts: MIT Press, 2016.
- [20] HINTON G E, OSINDERO S, TEH Y - W. A Fast Learning Algorithm for Deep Belief Nets [J]. Neural Computation, 2006, 18(07):1527-1554.
- [21] 尹宝才,王文通,王立春. 深度学习研究综述[J]. 北京工业大学学报, 2015, (01):48-59.
YIN Baocai, WANG Wentong, WANG Lichun. Review of deep learning [J]. Journal of Beijing University of Technology, 2015, (01):48-59.
- [22] 李中胜,刘林. 一种改进高斯核度量的HEC算法在变压器故障诊断中的应用[J]. 广东电力, 2016, 29(12):104-109.
LI Zhongsheng, LIU Lin. Application of hyper-ellipsoidal clustering algorithm based on improved Gaussian kernel metric in

transformer fault diagnosis[J]. Guangdong Electric Power, 2016, 29(12):104-109.

- [23] 刘世欣, 韩玮琦, 杨志刚, 等. 500 kV 电力变压器内部过热故障分析及处理[J], 内蒙古电力技术, 2017, 35(03):90-92.
LIU Shixin, HAN Weiqi, YANG Zhigang, et al. Analysis and processing of internal overheat fault of 500 kV transformer[J]. Inner Mongolia Electric Power, 2017, 35(03):90-92.
- [24] 黄皓炜. 变压器油中溶解气体在线监测系统的原理及应用[J]. 浙江电力, 2016, 35(02):31-35.
HUANG Haowei. Principle and application of online monitoring

system of dissolved gases in transformer oil[J]. Zhejiang Electric Power, 2016, 35(02):31-35.

收稿日期: 2018-05-18

作者简介:



杨涛(1987), 男, 硕士, 高级工程师, 主要从事输变电设备状态监测评价方面的研究工作。

(本文责任编辑: 范斌)

Diagnosis method of power transformer fault based on deep learning

YANG Tao, HUANG Junkai, XU Kui, WU Jianrong, CHEN Shijun

(Electric Science Research Institute of Guizhou Power Grid Co., Ltd., Guiyang 550002 Guizhou, China)

Abstract: The dissolved gas analysis method (DGA) in oil is an important method for internal fault diagnosis of transformers. It is widely used in transformer on-line monitoring and periodical test detection. Traditional diagnostic methods such as characteristic gas method and three-ratio method are commonly used in practical applications. These methods have some limitations which caused the fault diagnosis accuracy to be low. In order to solve this problem, this paper proposes a transformer fault diagnosis method based on multi-layer perceptron in deep learning technology. It uses the open source Scikit-learn machine learning framework and TensorFlow deep learning framework to construct a transformer fault diagnosis model. The fault diagnosis model was trained and tested by using the fault sample data in practical engineering. The test results show that the transformer fault diagnosis model based on multi-layer perceptron technology can correctly diagnose the transformer fault. Compared with the traditional three-ratio method and the support vector machine techniques, the multi-layer perceptron has a higher diagnostic accuracy and better fault diagnosis performance, which can provide more accurate reference information for transformer overhaul.

Key words: deep learning; dissolved gas analysis in oil; transformer; fault diagnosis; multilayer perceptron