

考虑复杂时序关联特性的Bi-LSTM变压器 DGA故障诊断方法

吴晓欣,何怡刚,段嘉璐,张 慧,曾昭璐

(武汉大学 电气与自动化学院,湖北 武汉 430072)

摘要:当前基于油中溶解气体分析(DGA)的变压器故障诊断方法往往仅考虑单一时刻数据点,容错性差,难以充分挖掘在线监测数据的时序信息。提出一种考虑变压器油特征参量序列间复杂关系的基于双向长短时记忆(Bi-LSTM)网络的变压器故障诊断方法。首先构建了变压器油特征参量序列,基于序列数据构建了Bi-LSTM变压器故障诊断模型。工程实际中不同变压器油特征参量序列长短不一,需通过排序、分组填充对模型输入进行重构改进,然后对超参数进行优化。基于同一自建数据库对比所提方法与其他方法,结果表明:经过数据重构后所提方法的准确率可达91.9%;当特征指标数量减少约2/3时,所提方法的准确率仅下降约1%,而其他方法的准确率平均下降约6%;当采样数据存在10%的随机错误时,所提方法诊断准确率仅下降2%~6%,且通过改变隐藏层的数量可得到改善。

关键词:变压器;故障诊断;双向长短时记忆网络;时间序列;复杂关联关系;油中溶解气体分析

中图分类号:TM 41

文献标志码:A

DOI:10.16081/j.epae.202006004

0 引言

随着计算机存储及传感器技术的不断发展,电力变压器的在线监测数据将呈现出爆炸式增长态势^[1-2]。作为电力系统的枢纽性装备,电力变压器的健康状态直接影响着整个电网的安全稳定。油中溶解气体分析DGA(Dissolved Gas Analysis)是一项对变压器绝缘油中溶解气体成分及含量进行分析的在线监测技术^[3]。通过对油中溶解气体与变压器状态之间相关性的研究可以对变压器的健康状态进行有效的诊断,及时排除潜伏性故障。

目前学者们基于DGA数据和传统诊断方法已进行了大量的研究^[4-5]。文献[6-7]考虑数据预处理或清洗以获取更可靠的诊断数据,进而提高诊断准确率。文献[8-9]则考虑多种算法的有机融合,利用优势互补进一步完善诊断模型。近年来,由于深度学习对复杂非线性特征具有强大的提取能力,部分文献尝试将其引入变压器故障诊断领域,其故障诊断准

确率相比传统机器学习诊断模型有了显著提升^[10-12]。

随着深度学习的发展,循环神经网络RNN(Recurrent Neural Network)因其对数据的“记忆能力”,开始被广泛应用于序列建模^[13-14]。而长短时记忆LSTM(Long Short-Term Memory)网络作为RNN的变体,因其有效地解决了RNN在反馈序列信息的过程中存在的梯度消失等问题,逐渐取代RNN成为序列建模的主流方法^[15]。目前LSTM在变压器故障诊断领域的应用仍然处于起步阶段,仅有少量文献针对LSTM的预测进行了研究。将LSTM引入基于DGA的变压器诊断中,可有效提高诊断方法的准确率、稳定性和适用性。

基于DGA的故障诊断模型大多仅考虑某故障状态下单一时刻的DGA数据点,使得数据采样出错的偶然性偏差直接影响相关诊断模型的训练及准确率。此外,现有方法未能充分提取不同时序特征参量之间的复杂关联关系,为故障诊断模型提供反映变压器状态的有效信息^[16-17];另一方面,海量的监测数据因无法被充分利用反而成为冗余或无效数据。

本文考虑时序数据复杂关联关系,提出一种基于双向长短时记忆Bi-LSTM(Bi-directional Long Short-Term Memory)网络的变压器故障诊断方法,该方法充分提取了变压器油特征参量序列间的复杂关联信息,且有效降低了监测过程中可能存在的噪声干扰与数据错误对故障诊断准确率的影响。此外,本文还考虑到实际应用中序列数据长度难以统一标准的问题,对模型输入矩阵进行重构,采取排序、分组填充、分批训练的策略对模型进行改进。最后,本文通过理论分析及反复验证给出了适合的超

收稿日期:2019-08-14;修回日期:2020-04-05

基金项目:国家自然科学基金资助项目(51977153,51977161,51577046);国家自然科学基金重点资助项目(51637004);国家重点研发计划“重大科学仪器设备开发”资助项目(2016YFF0102200);装备预先研究重点项目(41402040301)
Project supported by the National Natural Science Foundation of China(51977153,51977161,51577046),the State Key Program of National Natural Science Foundation of China(51637004),the National Key R&D Program of China “Important Scientific Instruments and Equipment Development”(2016YFF0102200) and the Equipment Research Project in Advance(41402040301)

参数 M (minibatch), 并对比给出了量化不同诊断模型效果的指标值。

1 基于油特征参量序列的变压器诊断框架

1.1 变压器油特征参量序列的构建

与一般的变压器故障诊断模型仅输入单一时刻的数据进行诊断不同, Bi-LSTM 诊断框架的输入需为序列数据, 以便网络可以深入挖掘不同序列间的关联关系, 因此需要构建变压器油特征参量序列作为模型的输入。目前基于深度学习模型的变压器 DGA 故障诊断研究较少, 而模型对数据量的要求较高, 考虑到同时包含 5 种主要油中溶解气体 (H_2 、 CH_4 、 C_2H_6 、 C_2H_4 、 C_2H_2) 含量的数据样本相对较多, 为保持数据统一性、扩大样本数据量, 本文选择这 5 种气体的监测序列作为变压器油特征参量序列的组成部分。

针对单一时刻的 DGA 数据及其比值与变压器故障模式的映射关系已经有了较为系统的研究基础, 常见的几种传统方法有特征气体法、Dornenburg 比率法、罗杰斯比值法等^[3], 尽管这些方法仅仅通过几种油中溶解气体及其简单比值提取变压器故障的内在规律, 难以完全表征变压器内部复杂的故障特征, 但是以此为依据构建输入深度学习模型的特征参量, 却可以有效地改善深度学习模型的诊断效果^[18]。因此, 本文借鉴文献[19]中的无编码比值构造方法, 基于 5 种主要变压器油中溶解气体的监测序列, 构造了 8 种新的变压器油特征参量序列。构造方法对这 5 种气体的监测序列进行对应时刻求比值计算, 每种比值计算对应一串序列, 总共有 8 种比值, 分别为 $C(H_2)/C(CH_4)$ 、 $C(H_2)/C(C_2H_6)$ 、 $C(H_2)/C(TH)$ 、 $C(CH_4)/C(C_2H_2)$ 、 $C(CH_4)/C(C_2H_6)$ 、 $C(C_2H_2)/C(C_2H_4)$ 、 $C(C_2H_4)/C(TH)$ 、 $C(C_2H_6)/C(TH)$ 。其中, TH 为总烃, 包含 CH_4 、 C_2H_2 、 C_2H_4 、 C_2H_6 ; $C(H_2)/C(CH_4)$ 表示 H_2 、 CH_4 的监测序列在对应时刻的比值, 其他依此类推。求取比值时, 若分母为 0, 则将分母修改为极小值以避免出现无效值, 本文取该极小值为 10^{-8} 。分别对每种比值计算对应的序列, 则有 8 种新的变压器油特征参量序列。因此共有 13 种变压器油特征参量序列作为 Bi-LSTM 诊断框架的输入, 特征参量序列的构建方法如图 1 所示。通过比值法和气体原始数据的结合, 能够充分利用 LSTM 网络对复杂非线性特征强大的提取能力, 且有效缩短了运算时间。

1.2 LSTM 网络原理

LSTM 网络引入了门控单元以解决长序列数据长期依赖问题, 避免网络训练过程中出现梯度消失或爆炸, 以便保留每一循环单元的不同时间尺度信息特征^[20]。LSTM 单元结构图如附录中的图 A1 所

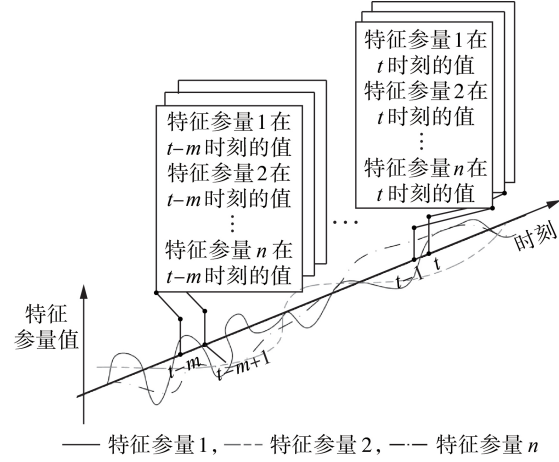


图1 变压器油特征参量序列的构建

Fig.1 Construction of characteristic parameter sequences of transformer oil

示。设隐藏层中, 某一 LSTM 单元的输入为 t 时刻的输入序列值 x_t 、 $t-1$ 时刻对应的 LSTM 单元的输出值 h_{t-1} 及该时刻的记忆单元状态 c_{t-1} , 该 LSTM 单元的输出为 t 时刻对应的 LSTM 单元的输出值 h_t 及该时刻记忆单元状态 c_t 。每个时刻对应的 LSTM 单元的输出不仅受制于该时刻的输入序列值, 还取决于历史数据对该时刻数据的影响, 而将不同数据对最终输出的影响进行结构化表示便得到 3 个门, 即输入门、输出门、遗忘门。其中, 输入门操控着 x_t 对 c_t 的影响, 遗忘门操控着 c_{t-1} 对 c_t 的影响, 输出门操控着 c_t 对 h_t 的影响。输入门、输出门、遗忘门的计算公式分别如式(1)~(3)所示。

$$i_t = \varphi(W_i[h_{t-1}, x_t] + b_i) \quad (1)$$

$$o_t = \varphi(W_o[h_{t-1}, x_t] + b_o) \quad (2)$$

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

其中, i_t 、 o_t 、 f_t 分别为输入门、输出门、遗忘门在 t 时刻的状态; W_i 、 W_o 、 W_f 分别为输入门、输出门、遗忘门的权值矩阵; b_i 、 b_o 、 b_f 分别为输入门、输出门、遗忘门的偏置项; φ 为 sigmoid 激活函数。故 t 时刻对应的 LSTM 单元的输出如式(4)、(5)所示。

$$h_t = o_t \odot \psi(c_t) \quad (4)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \psi(W_{is}[h_{t-1}, x_t] + b_{is}) = f_t \odot c_{t-1} + i_t \odot \psi(i_s) \quad (5)$$

其中, ψ 为激活函数, 通常取 tanh 函数; $\odot$ 表示按元素相乘; W_{is} 为输入状态权重矩阵; b_{is} 为输入状态偏置项。

1.3 基于 Bi-LSTM 网络的 DGA 诊断框架

Bi-LSTM 网络是由 2 层 LSTM 网络通过反向连接构成的, 可以充分提取各个数据序列间的关联关系及序列自身的前向反向关联关系, 最终得到序列与变压器状态之间的映射关系, 其结构图如图 2 所示^[20]。考虑到在变压器故障诊断领域, 已有文献大

多是对每个试验指标与变压器故障状态构成单独的映射关系,并未充分考虑变压器油特征参量序列间的横向关联联系及序列自身随时间变化而波动的纵向关系,结合 Bi-LSTM 网络的原理,本文构建了基于 Bi-LSTM 网络的 DGA 诊断框架,如图 2 所示。

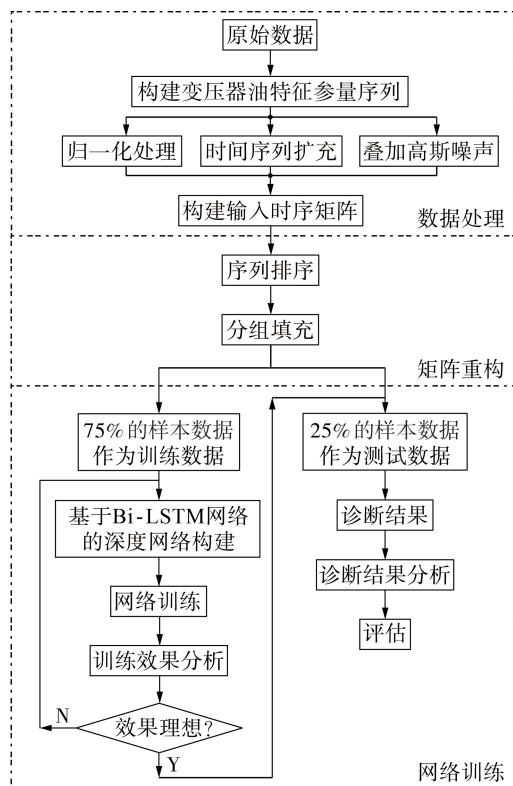


图2 基于 Bi-LSTM 网络的 DGA 诊断框架

Fig.2 DGA diagnosis framework based on Bi-LSTM network

最终利用训练样本对构建好的深度学习框架进行训练,过程如下。

(1)利用前向传播运算得到每个单元的输出,则代价函数如式(6)所示。

$$L(f(X_i; \theta), y_i) = -\frac{1}{N} \sum_{i=1}^N y_i \lg f(X_i; \theta; b) + (1 - y_i) \lg (1 - f(X_i; \theta; b)) \quad (6)$$

其中, $L(f(X_i; \theta), y_i)$ 为交叉熵函数, X_i 为样本值, θ 为网络中的权值矩阵, y_i 为对应的样本期望输出值,即对应的故障标签; N 为样本数; $f(X_i; \theta; b)$ 为第 i 个样本输入网络后的输出结果, b 为网络中的偏置矩阵。

(2)利用反向传播计算得出每层每个单元输出的偏差项,其中 LSTM 层的误差有向上一层传播和沿时间反向传播 2 个反向传播方向。

(3)由偏差项计算每个参数的梯度并更新参数值。

(4)反复迭代并修正参数值,使代价函数越来越

小直至达到预先设定的迭代次数。

2 基于 Bi-LSTM 网络的变压器故障诊断方法

2.1 样本数据获取及预处理

本文通过采集电网公司历年故障变压器的 5 种主要油中溶解气体实测数据及大量文献中已知故障类型的 DGA 数据,组成 401 组数据。限于篇幅,仅列出其中的 3 组数据及其对应的变压器状态如附录中的表 A1 所示^[21-22]。

利用 1.1 节中所述的变压器油特征参量序列构建方法进行数据初步处理,得到 401 组包含 13 种变压器油特征参量序列的数据集。采用式(7)对数据进行正则化,得到均值为 0、方差为 1 的新数据集。由于深度学习模型往往需要较大的数据量,通过叠加高斯噪声的方式将新数据集扩充至 1 203 例。叠加噪声的概率密度函数服从高斯分布(即正态分布),如式(8)所示。

$$g_i = \frac{X_i - E(X)}{D(X)} \quad (7)$$

其中, g_i 为归一化值; $E(X)$ 为均值; $D(X)$ 为方差。

$$p_G(g) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(g-\mu)^2}{2\sigma^2}} \quad (8)$$

其中, μ 为均值; σ 为标准差; $p_G(g)$ 为高斯噪声的概率密度函数。对应各个状态标签的数据量如表 1 所示。

表1 故障状态编码及样本数据分布

Table 1 Fault state coding and sample data distribution

故障序号	故障标签	样本总数	扩充后样本数	训练样本数	测试样本数
I	高温过热	83	249	187	62
II	中温过热	48	144	108	36
III	低温过热	45	135	101	34
IV	局部放电	54	162	122	40
V	低能放电	98	294	211	73
VI	高能放电	73	219	164	55

2.2 故障诊断框架应用流程

(1)输入矩阵重构。

工程实际中不同变压器油特征参量序列往往长短不一,难以统一标准,因此需考虑变压器油特征参量序列长度的不一致性,使模型在实际应用中更具可移植性。

首先利用三次样条插值对样本序列进行等间隔扩充,得到采样频率为 f_{si} 的 i 组数据。然后进行序列排序及分组填充,模型的输入数据在横向和纵向上分别按照时间顺序和不同数据组的各类特征指标进行排列,以矩阵形式形成网络批训练的输入数据集,如式(9)所示。

$$\mathbf{X}_{\text{input}} = \begin{bmatrix} \mathbf{X}_1^{t_1} \\ \mathbf{X}_2^{t_2} \\ \vdots \\ \mathbf{X}_k^{t_m} \end{bmatrix} = \begin{bmatrix} x_1^1 & x_1^2 & \cdots & x_1^{t_1} \\ x_2^1 & x_2^2 & \cdots & x_2^{t_2} \\ \vdots & \vdots & & \vdots \\ x_k^1 & x_k^2 & \cdots & x_k^{t_m} \end{bmatrix} \quad (9)$$

其中, $\mathbf{X}_k^{t_m}$ 为第 k 种指标在 t_m 时刻的时间序列串; $x_k^{t_m}$ 为第 k 种特征指标在 t_m 时刻的值。

将样本数据集中的 75% 作为训练集、25% 作为测试集。

(2) 基于 Bi-LSTM 网络的深度网络构建。

首先, 构建一个如图 3 所示的 5 层网络, 该网络包括输入层、Bi-LSTM 层 (其隐藏单元数为 H)、全连接层、softmax 层、分类输出层。其中, Bi-LSTM 网络单元采用 tanh 函数作为状态激励函数, 如式 (10) 所示, 采用 sigmoid 函数作为门激励函数, 如式 (11) 所示; 全连接层将 Bi-LSTM 层的输出作为输入向量, 而其输出的特征向量通过 softmax 层进行特征分类, 最终利用 softmax 函数输出故障类别的概率分布, 具体如式 (12) 所示。

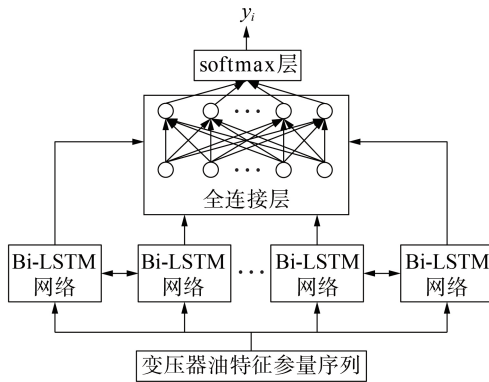


图3 构建的5层网络

Fig.3 Schematic diagram of five-layer network

$$\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \quad (10)$$

$$\text{sigmoid}(z) = \frac{1}{1 + e^{-z}} \quad (11)$$

$$\text{softmax}(\mathbf{x}) = \begin{bmatrix} p(y=1, \mathbf{x} | \boldsymbol{\theta} | \mathbf{b}) \\ p(y=2, \mathbf{x} | \boldsymbol{\theta} | \mathbf{b}) \\ \vdots \\ p(y=K, \mathbf{x} | \boldsymbol{\theta} | \mathbf{b}) \end{bmatrix} = \frac{1}{\sum_{i=1}^K e^{z_i}} \begin{bmatrix} e^{z_1} \\ e^{z_2} \\ \vdots \\ e^{z_K} \end{bmatrix} \quad (12)$$

其中, z 为激活函数的输入; $\mathbf{x}$ 为全连接层的输出, 即 softmax 层的输入; y 为诊断结果; $z_i = \boldsymbol{\theta}_i^T \mathbf{x} + b_i$, $\boldsymbol{\theta}_i$ 、 b_i 分别为 $\boldsymbol{\theta}$ 、 $\mathbf{b}$ 中的元素; K 为故障标签数量。

(3) 超参数设置及网络训练。

本文采用自适应矩估计 (Adam) 算法作为优化器进行网络训练, 网络超参数设计值及所采用的计算机性能参数如下。

a. 超参数设计值: 初始学习率为 0.001, 梯度阈

值为 0.1, 最小训练批次为 250, 总迭代次数为 300, 隐藏单元数为 1500。

b. 计算机性能参数: 内存为 16 GB, 处理器型号为 Intel(R) Core(TM) i7-7700@3.60 GHz, 显卡型号为 NVIDIA GeForce GT 730, 操作系统为 Windows 10 (64 bit)。

(4) 诊断准确率 γ_{accuracy} 的计算公式为:

$$\gamma_{\text{accuracy}} = \frac{\text{sum}(y_{\text{test}} = y_{\text{pred}})}{\text{num}(y_{\text{test}})} \times 100\% \quad (13)$$

其中, y_{test} 为数据实际对应的故障标签; y_{pred} 为模型诊断结果; sum 为求和函数; num 为计数函数。

2.3 矩阵重构效果检验

直接将该长度不一的数据序列作为模型的输入量, 网络训练过程中过多的序列填充步骤将对网络分类性能造成较大的负面影响, 因此需要对输入矩阵进行序列排序、分组填充的重构改进。

按照 2.2 节所述的故障诊断框架应用流程进行网络构建及模型训练, 并将矩阵重构前、后的网络训练结果进行对比, 结果如图 4 所示。由图可知, 序列长度随机分布时, 网络训练过程发生振荡, 稳定性很差。这是由于序列长度的随机分布, 使得对应于同一状态标签的数据长度差异甚大, 最终影响网络分类性能。

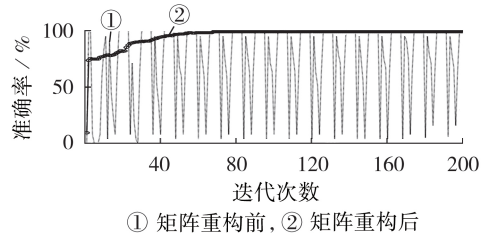


图4 矩阵重构前、后模型效果对比

Fig.4 Comparison of model effects between before and after matrix reconstruction

取最后 50% 的迭代结果进行计算, 量化结果分析如表 2 所示, 由表可知矩阵重构后的模型训练准确率提高了 29.8%。

表2 矩阵重构前、后模型效果参数对比

Table 2 Comparison of model effect parameters between before and after matrix reconstruction

情况	准确率 平均值 / %	准确率 标准差	训练损失 平均值	训练损失 标准差
矩阵重构前	77.040	33.94	1.2299	1.7286
矩阵重构后	99.996	0.04	0.0029	0.0042

最终, 进行矩阵重构后的模型诊断准确率可达 91.9%, 对于表 A1 中的示例均能诊断正确。例如对于表 A1 中某型号为 SFSZ9-90000/220 的 220 kV 主变压器的 DGA 监测数据, Bi-LSTM 变压器故障诊断模型的诊断结果为高能放电故障, 经检查发现该变

压器潜油泵存在接地故障,发生了电弧放电,出现严重烧损变黑,与模型诊断结果一致。

2.4 超参数的优化选择

随着分批训练所选择的批次大小的不同,故障诊断精度差异较大。原因在于当批次过大时,过多的序列填充将会对网络的分类性能造成较大的负面影响,而批次过小时,同一批次中序列长度接近的变压器油特征参量序列对应的状态标签占比较不均匀,也将直接破坏网络的分类性能,因此需要合理选定批次,即minibatch的值(本文用 M 表示)。

将 M 分别选取为50、100、200、250、300,网络训练过程中的训练准确率和训练损失分别如附录中的图A2、A3所示,表3给出了训练验证结果的参数指标。从表3可以看出,当 $M=250$ 时,在迭代次数及训练时间较少的情况下,诊断准确率即可达91.9%。故针对本文设定的应用场景,选取 $M=250$ 较为适宜。

表3 不同批次选择对训练及诊断结果的影响

Table 3 Influence of different batch selection on training and diagnosis results

M	准确率 / %	损失率	训练时间 / s	迭代次数
50	88.8	0.1105	246	1800
100	89.4	0.0719	195	1800
200	90.3	0.0015	140	400
250	91.9	0.0009	162	300
300	90.3	0.0716	166	300

3 对比分析

3.1 不同算法的对比

本文选取IEC三比值法、线性判别分析(LDA)法、支持向量机(SVM)法、K-近邻(KNN)法、梯度下降树(GBDT)法、随机森林(RF)法与本文方法进行对比。由于除本文方法外,其余方法均对单一时刻点进行分析,因此对自建数据库中的序列数据随机采样,取1个时刻的样本获得单一时刻点数据。

为进一步表现Bi-LSTM网络对复杂特征的提取能力,去掉额外构建的8个特征指标,仅将5种主要油中溶解气体的归一化值作为模型的输入参量,对模型进行训练,不同方法的诊断结果如表4所示。

表4 不同特征指标下各方法诊断准确率

Table 4 Diagnosis accuracy of each method upon different characteristic indicators

方法	模型诊断准确率 / %		下降百分比 / %
	13种特征指标	5种特征指标	
IEC三比值法	64.4	64.4	—
LDA法	68.2	49.8	27
SVM法	82.8	78.5	5
KNN法	79.2	85.1	-7
GBDT法	84.1	81.5	3
RF法	87.3	85.5	2
Bi-LSTM网络	91.9	90.8	1

由表可知,基于Bi-LSTM网络的深度学习模型故障诊断准确率最高,且当特征指标数量急剧减少到原来数量的36%时,大多数传统人工智能方法诊断准确率都有不同程度的下降,只有KNN法的诊断准确率反而上升,原因在于其数学本质是利用欧氏距离进行计算,而向量的维度越高,欧氏距离的区分能力反而越弱,因此KNN法适用于较简单的分类。而Bi-LSTM网络对复杂非线性特征具有强大的提取能力,因此在特征指标急剧减少时,其网络的诊断准确率仅下降约1%,可见Bi-LSTM网络的适应性较强。

3.2 序列数据诊断的容错性分析

在测试样本集中设置采样出错的数据,本文采取将出错时刻对应的数据置零的方法。在此情况下,基于单一时刻点数据构建的传统方法在处理采样出错的数据时无校正能力。为验证Bi-LSTM网络在数据采样偶然性出错时仍然具有较高的诊断准确率,对原测试集中所有序列数据随机设置10%的时刻为采样出错点,将对应的数据置0。利用这些数据对训练好的Bi-LSTM网络(训练时设置 $M=250$)进行测试,计算其故障诊断准确率,结果如表5所示。

表5 数据出错对故障诊断准确率的影响

Table 5 Influence of data error on fault diagnosis accuracy

H	准确率 / %		下降百分比 / %
	测试数据未出错	测试数据出错	
150	91.9	85.7	6.75
200	88.8	84.7	4.62
250	90.8	88.8	2.20
300	89.8	86.7	2.36
350	89.8	87.8	2.23
400	87.8	85.7	2.39

由表5可看出,在考虑所有测试数据均出现采样出错的极端情形下,基于Bi-LSTM网络的变压器故障诊断模型准确率略有下降,当 $H=250$ 时的诊断准确率最高,此时的下降幅度仅为2.20%,说明本文模型具有较强的鲁棒性;在数据未出错的情形下,Bi-LSTM层的隐藏单元数在150和250间波动,对诊断准确率影响不大,而考虑数据出错情形时,隐藏单元数越多,网络诊断准确率下降幅度越小。因此,为提高模型的鲁棒性,应选择隐藏单元数为250。

4 结论

本文将序列建模引入变压器DGA故障诊断领域中,提出了一种基于无编码比值和Bi-LSTM网络的变压器故障诊断方法,不同于传统仅考虑单一时刻点数据的诊断方法,本文方法充分考虑变压器油特征参量序列之间的复杂关联关系,同时避免了数据采样出错对诊断结果的严重影响,有效提高了故障诊断的准确率。本文所得的结论如下。

(1)本文方法将LSTM网络引入变压器故障诊断领域,构建Bi-LSTM网络诊断框架,通过深度学习能够有效提取多维数据时间序列间的复杂非线性特征;本文方法充分考虑不同序列间的复杂关联关系,诊断准确率达到91.9%,而基于同一自建数据库进行计算的传统诊断方法中,准确率最高也具有87.3%。

(2)相比传统利用单一时刻数据进行变压器故障诊断的方法,本文采用序列建模,既可充分利用实时监测信息,又可有效避免因数据的偶然性出错对故障诊断准确率产生的影响;在考虑测试数据均出现采样出错的极端情形时,基于Bi-LSTM网络的变压器故障诊断模型的准确率仅下降2.20%,说明其具有较强的鲁棒性。

(3)本文方法考虑了工程实际中不同变压器时间序列长短不一的特点,通过长度排序、分组填充、批训练的方式对模型输入矩阵进行重构,有效地将基于Bi-LSTM网络的变压器故障诊断模型的训练准确率提高了29.8%。

附录见本刊网络版(<http://www.epae.cn>)。

参考文献:

- [1] 刘云鹏,许自强,李刚,等. 人工智能驱动的数据分析技术在电力变压器状态检修中的应用综述[J]. 高电压技术, 2019, 45(2): 337-348.
- [2] 李刚,于长海,刘云鹏,等. 电力变压器故障预测与健康管理的挑战与展望[J]. 电力系统自动化, 2017, 41(23): 156-167.
- [3] BAKAR N, ABU-SIADA A, ISLAM S. A review of dissolved gas analysis measurement and interpretation techniques[J]. IEEE Electrical Insulation Magazine, 2014, 30(3): 39-49.
- [4] 杨廷方,张航,黄立滨,等. 基于改进型主成分分析的电力变压器潜伏性故障诊断[J]. 电力自动化设备, 2015, 35(6): 149-153, 165.
- [5] CRUZ V, COSTA A, PAREDES M. Development and evaluation of a new DGA diagnostic method based on thermodynamics fundamentals[J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2015, 22(2): 888-894.
- [6] 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.
- [7] 代杰杰,宋辉,杨伟,等. 基于栈式降噪自编码器的输变电设备状态数据清洗方法[J]. 电力系统自动化, 2017, 41(12): 224-230.
- [8] 范永敏,汪泓,孙秋芹,等. RVM和ANFIS用于变压器故障诊断及状态评估[J]. 电力系统及其自动化学报, 2018, 30(3): 35-41.
- [9] 王永强,律方成,李和明. 基于粗糙集理论和贝叶斯网络的电力变压器故障诊断方法[J]. 中国电机工程学报, 2006, 26(8): 137-141.
- [10] 周念成,廖建权,王强钢,等. 深度学习在智能电网中的应用现状分析与展望[J]. 电力系统自动化, 2019, 43(4): 180-197.
- [11] 代杰杰,宋辉,杨伟,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.
- [12] ZHAO Rui, YAN Ruqiang, CHEN Zhenghua, et al. Deep learning and its applications to machine health monitoring[J]. Mechanical Systems and Signal Processing, 2019, 115: 213-237.
- [13] LIPTON Z C, BERKOWITZ J, ELKAN C. A critical review of recurrent neural networks for sequence learning[J]. Computer Science. (2015-10-17)[2019-04-01]. <https://arxiv.org/abs/1506.00019v2>.
- [14] 王志远,王守相,陈海文,等. 考虑空间相关性采用LSTM神经网络的光伏出力短期预测方法[J/OL]. 电力系统及其自动化学报. (2019-07-22)[2020-04-07]. <https://doi.org/10.19635/j.cnki.csu-epsa.000298>.
- [15] 张宇帆,艾芊,林琳,等. 基于深度长短时记忆网络的区域级超短期负荷预测方法[J]. 电网技术, 2019, 43(6): 1884-1892.
- [16] 梁永亮,郭汉琮,薛永端. 基于特征气体关联特征的变压器故障诊断方法[J]. 高电压技术, 2019, 45(2): 386-392.
- [17] 赵文清,严海,周震东,等. 基于残差BP神经网络的变压器故障诊断[J]. 电力自动化设备, 2020, 40(2): 143-148.
- [18] SOUHLIA S, BACHA K, CHAARI A. SVM-based decision for power transformers fault diagnosis using Rogers and Doernenburg ratios DGA[C]//2013 10th International Multi-Conference on Systems, Signals & Devices. (2013-03-01)[2019-04-01]. <https://ieeexplore.ieee.org/document/6564073/>.

- [19] 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):6570-6578,6625.
WANG Ke,LI Jinzhong,ZHANG Shuqi,et al. New features derived from dissolved gas analysis for fault diagnosis of power transformers[J]. Proceedings of the CSEE,2016,36(23):6570-6578,6625.
- [20] 刘飞龙,郝文宁,陈刚,等. 基于双线性函数注意力Bi-LSTM模型的机器阅读理解[J]. 计算机科学,2017,44(增刊1):92-96,122.
LIU Feilong,HAO Wenning,CHEN Gang,et al. Attention of bilinear function based Bi-LSTM model for machine reading comprehension[J]. Computer Science,2017,44(Supplement 1):92-96,122.
- [21] 牛林. 变压器状态监测诊断技术[M]. 北京:中国电力出版社,

2013:85-97.

- [22] 徐康健,孟玉婵. 变压器油中溶解气体的色谱分析实用技术[M]. 北京:中国质检出版社,中国标准出版社,2011:148-170.

作者简介:



吴晓欣

吴晓欣(1997—),女,福建泉州人,硕士研究生,主要研究方向为深度学习、变压器故障诊断及预测(E-mail:739889197@qq.com);

何怡刚(1966—),男,湖南新邵人,教授,博士,主要研究方向为智能电网技术、射频识别技术等。

(编辑 任思思)

Bi-LSTM-based transformer fault diagnosis method based on DGA considering complex correlation characteristics of time sequence

WU Xiaoxin,HE Yigang,DUAN Jiajun,ZHANG Hui,ZENG Zhaorong

(School of Electrical Engineering and Automation,Wuhan University,Wuhan 430072,China)

Abstract: Transformer fault diagnosis methods based on DGA(Dissolved Gas Analysis) are commonly just consider data obtained at a single moment,which is susceptible to data errors and cannot fully mine the online monitoring data sequence information. Aiming at this problem,a Bi-LSTM(Bi-directional Long Short-Term Memory) network based transformer fault diagnosis method considering the complex correlation relationship among the characteristic parameter sequences of transformer oil is proposed. Firstly,the characteristic parameter sequences of transformer oil are constructed,and then the Bi-LSTM-based transformer fault diagnosis model is constructed based on the sequence data. Considering the different length of the sequences in engineering practice,the model inputs are reconstructed by sorting and group filling. The hyper-parameters are then optimized. Based on the same built database,the proposed method is compared with other methods. The results show that the diagnosis accuracy of the proposed method after data reconstruction can reach 91.9%,and only drops by about 1% when the number of characteristic parameter decreased by about two-thirds,while other methods decrease by about 6% on average. When the sampling data exists a 10% random errors,the diagnosis accuracy of the proposed method decreases by 2% to 6%,and can be improved by changing the number of hidden layers.

Key words:electric transformers;fault diagnosis;bi-directional long short-term memory network;time sequence;complex correlation relationship;dissolved gas analysis

附录

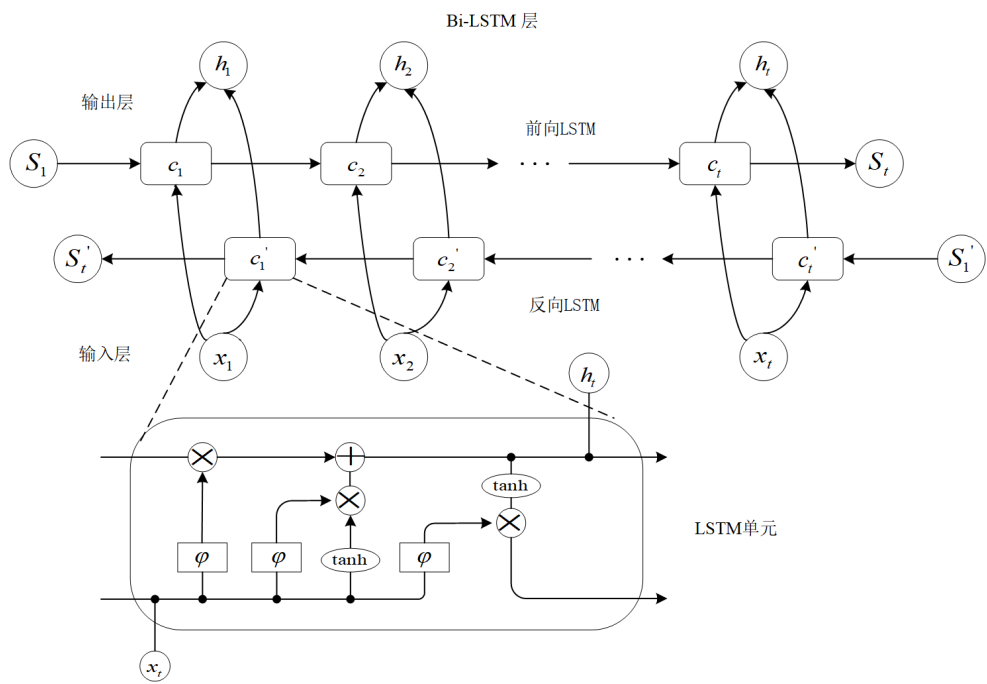


图 A1 LSTM 单元结构图和 Bi-LSTM 层图

Fig.A1 Structure diagram of LSTM unit and Bi-LSTM layer diagram

表 A1 变压器油中溶解气体序列示例

TableA1 Examples of characteristic parameter sequence of transformer oil constructed based on monitoring data

(a) 某发电公司 220kV 变压器								
变压器油特征参量	某发电公司：220kV 变压器油中溶解气体的在线监测数据						变压器状态	
	2007-4-30	2007-5-1	2007-5-2	2007-5-3	2007-5-8	2007-5-9		
	17:12	19:31	18:05	18:06	23:07	23:08		
H ₂	40.42	42.82	42.57	40.23	38.39	44.53	中温过热	
CH ₄	184.16	186.62	184.01	185.21	188.93	198.92		
C ₂ H ₆	92.72	91.44	92.8	99.15	103.43	104.66		
C ₂ H ₄	173.27	206.74	192.16	208.84	215.29	216.16		
C ₂ H ₂	0	0	0	0	0	0		
变压器油特征参量	上述同一变压器油中溶解气体的离线分析数据						变压器状态	
	2007-4-30	2007-5-1	2007-5-2	2007-5-3	2007-5-8	2007-5-9		
	H2	46.59	44.73	47.59	42.13	41.12		45.08
CH4	214.29	214.19	212.96	209.42	218.05	229.92		
C2H6	97.95	99.05	100.11	102.32	106.09	104.23		
C2H4	186.99	187.65	189.56	192.32	198.36	199.17		
C2H2	0	0	0	0	0	0		
(b) 陕西某 750kV 变电站 5 号主变压器								
变压器油特征参量	陕西某 750kV 变电站 5 号主变压器：变压器油中溶解气体数据						变压器状态	
	2010-5-29	2010-5-29	2010-5-29	2010-5-29	2010-5-30	2010-5-30		
	01:37	07:37	13:37	19:37	01:37	07:37		
H ₂	149.13	149.59	149.21	153.2	151.11	151.49	151.71	低能放电
CH ₄	22.63	22.43	21.67	22.19	22.77	22.67	22.39	
C ₂ H ₆	4.45	4.4	4.46	4.37	4.46	4.45	4.42	
C ₂ H ₄	10.27	10.14	.97	10.25	10.32	10.27	10.19	
C ₂ H ₂	2.28	2.34	2.26	2.24	2.36	2.3	2.27	
变压器油特征参量	2010-5-30	2010-5-3	2010-5-3	2010-5-3	2010-5-3	2010-6-1	...	
	19:37	1	1	1	1	1	...	
		01:37	07:37	13:37	19:37	01:37		
H ₂	150.38	147.93	150.63	153.88	153.77	153.3		
CH ₄	21.35	22.2	22.75	22.81	22.6	23.13		
C ₂ H ₆	4.37	4.47	4.53	4.45	4.47	4.47		
C ₂ H ₄	9.98	10.28	10.4	10.46	10.3	10.42		
C ₂ H ₂	2.23	2.33	2.38	2.31	2.31	2.38		
(c) 某 220kV 主变压器(型号 SFSZ9-90000/220)								
变压器油特征参量	某 220kV 主变压器(型号 SFSZ9-90000/220)：变压器油中溶解气体数据记录时间						变压器状态	
	2004-2-15	2004-4-1	2004-6-10	2004-6-17	2004-6-25	2004-7-8		
	H ₂	4.3	4.4	710	735	393		642
CH ₄	1.2	6.5	238	242	278	349		
C ₂ H ₆	0	3.3	38.4	36.5	40.0	49.5		
C ₂ H ₄	7.3	45.8	389	453	515	681		
C2H2	0	36.4	299	319	330	423		

注：针对较长序列或故障状态发生变化的序列，如表 A1(b)中数据，将其分段截取为多组样本数据。

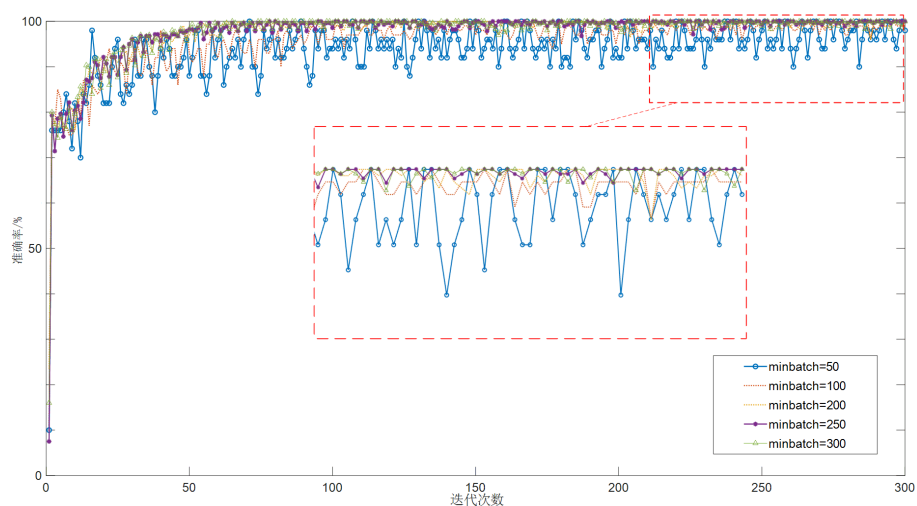


图 A2 不同训练批次对应的网络训练准确率

Fig.A2 Network training accuracy rate under different training batches

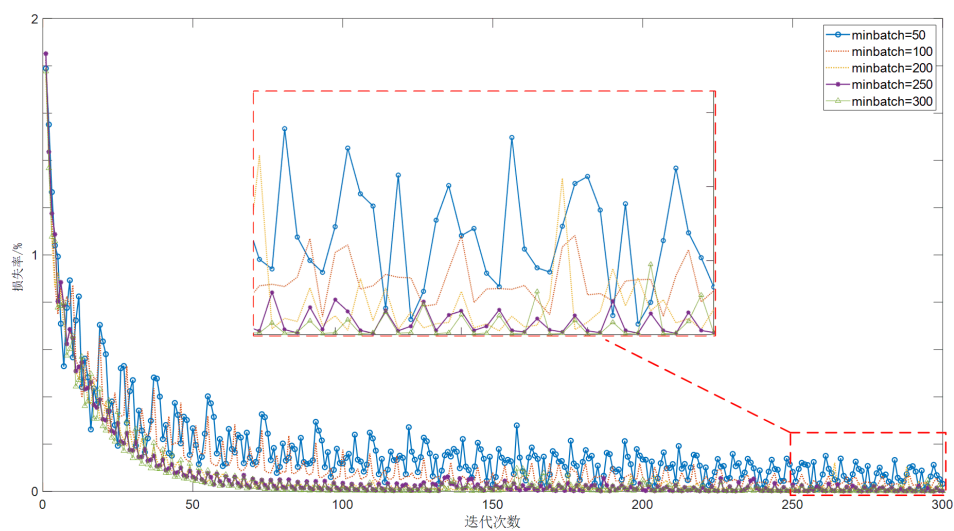


图 A3 不同训练批次对应的网络训练损失

Fig.A3 Network training loss under different training batches