

基于改进 LSTM 的电抗器故障预警方法

李冲冲, 史 操*

(青岛科技大学信息科学技术学院, 山东 青岛 266100)

摘 要: 针对传统的电抗器故障预警方法存在过拟合以及缺乏通用性的问题,提出了基于长短期记忆神经网络(LSTM) 和孪生神经网络的电抗器故障预警方法(以下简称 HDDse)。该方法结合长短期记忆神经网络和孪生神经网络的优势,LSTM 结构用于学习电抗器健康状态下的动态变化行为,孪生神经网络结构可以降低电抗器信息映射到高维空间的学习效率,该方法已成功地应用于电抗器的故障诊断中。实验结果表明,与已有电抗器故障预警方法相比,HDDse 可以极大提升电抗器故障预警准确性。

关键词: 电抗器故障预警; 预警模型; 长短期记忆神经网络; 孪生神经网络; 故障诊断

中图分类号: TP3-05 **文献标识码:** A **文章编号:** 1009-2552(2024)07-0076-08

DOI: 10.13274/j.cnki.hdzj.2024.07.013

Reactor fault early warning method based on improved LSTM

LI Chong-chong, SHI Cao

(Qingdao University of Science and Technology, Qingdao 266100, Shandong Province, China)

Abstract: To solve the shortcomings of over-fitting and lack of generality of traditional reactor fault early warning methods, a reactor fault early warning method based on short-term and short-term memory neural network and twin neural network (hereinafter referred to as HDDse) is studied. This method combines the advantages of long and short term memory neural network and twin neural network. LSTM structure is used to learn the dynamic change behavior of reactor under healthy state. Besides, twin network structure can reduce the learning efficiency of reactor information mapping to high-dimensional space. This method has been successfully applied in reactor fault diagnosis. The experiment results show that HDDse can greatly improve the accuracy of reactor fault early warning compared with the existing reactor fault early warning methods.

Key words: reactor fault prediction; early warning model; long and short term memory neural network; twin neural network; fault diagnosis

0 引 言

随着电力行业的发展,我国的电力系统正在向着超高压、特高压方向发展^[1],近几年,我国无

论是装机容量还是年发电量都已进入世界前列。电力是人类最主要的终端能源^[2],但从全球范围看,电力仍然存在着供应不足的情况。维护已有的电气设备,提高设备的运行可靠性以确保国民经济的发展以及人民生活水平的提高,十分必要。

电抗器是重要的电力设备,对整个系统的安全运行起着决定性的作用^[3]。随着电力工业的迅速发展,电抗器的电压等级和容量越来越高,电

基金项目: 国家自然科学基金(61806107)

作者简介: 李冲冲(1990 -),男,硕士研究生,研究方向为电网业务中台及设备管理系统设计研发。

* **通讯作者:** 史操(1981 -),男,博士,讲师,研究方向为深度学习、图像处理。

抗器故障带来的损失也会越来越大。因此,电抗器的故障预警具有重要的现实意义。

在电抗器的故障预警研究当中,大多是通过机器学习的方法^[4],利用电抗器的状态数据建立预警模型,在准确率达到一定稳定程度之后,维护人员可以根据模型的预警结果,将可能发生故障的电抗器进行替换,降低由于电抗器故障导致电力系统瘫痪的风险^[5]。目前的电抗器故障模型存在一定的误报率,会导致电力设备替换电抗器的代价提升。

人工智能技术的迅速发展为电抗器故障预警提供了技术支持,但目前已有的电抗器故障预警模型效果存在一定不足^[6],因此本文将以电抗器

故障的预警问题为切入点,通过人工智能技术解决目前电抗器故障预警存在的问题。

1 基于改进 LSTM 的故障预警模型(HDDse)

通用电抗器故障预警系统总体结构如图 1 所示,包括样本池模块、电抗器故障预警模块以及决策模块。样本池模块存储采集的电抗器运行参数数据,形成模型的样本数据集,训练集包括带有标签的电抗器样本和没带标签的样本。故障预警模块是 HDDse 的核心模块,用于电抗器故障预警,样本池模块数据作为电抗器故障预警模块的输入,输出二分类的预警结果。决策模块将若干连续样本的预警结果映射到最终的运行健康状况模块。

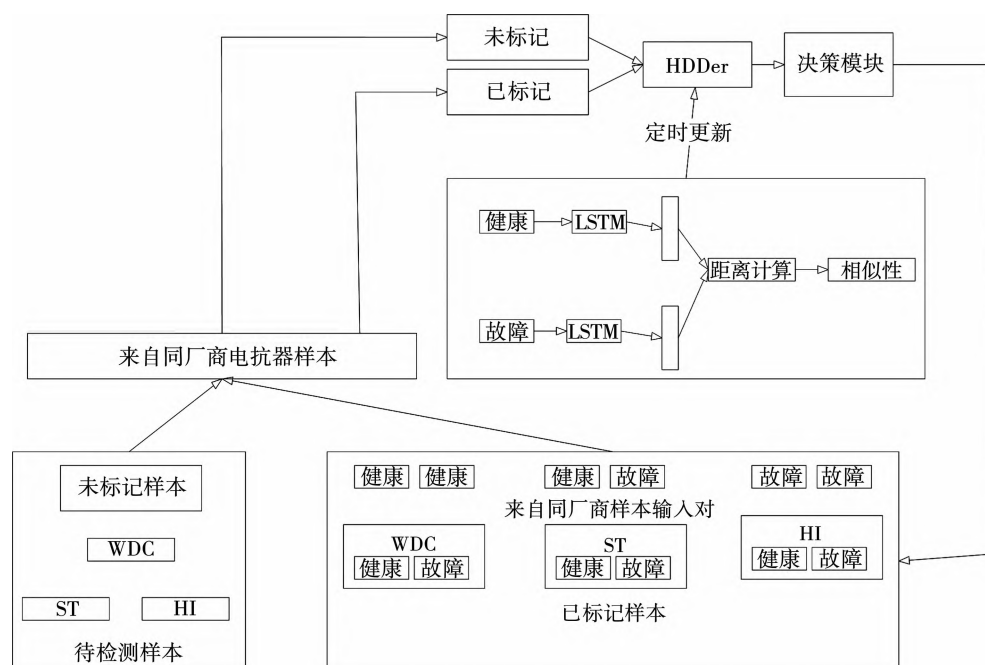


图 1 HDDse 的系统结构

1.1 孪生网络结构

目前常见的电抗器故障预警机器学习模型都依赖于假设,即电抗器将发生故障时,电抗器的运行状态值会发生突变。常见的电抗器故障预警方法将电抗器运行状态值的单个快照样本作为模型的训练集,对模型进行训练建模,并没有考虑到电抗器状态是随着时间变化而变化的,电抗器的状态变化顺序具有依赖性,因此,目前常见的方法缺乏使用时间序列数据进行建模的能力。电抗器运

行状态值在电抗器发生故障前会动态变化,因此电抗器的当前状态与长期的历史趋势信息有关。为此,本文将 LSTM 应用于顺序电抗器运行状态值的时序数据解析。

目前常用的分类方法,如支持向量机^[7]、深度神经网络^[8]等方法,存在以下问题:

①需要事先知道所有的待分类类别才能进行预警模型的训练,对于不同厂商不同型号的电抗器数据需要重新训练模型,以适应这些新的电抗

器型号引入的数据分布变化,获得更加可靠的预警效果。

②对于小样本量的电抗器预警易出现过拟合现象。

孪生神经网络是共享权重的神经网络,由两个相同的网络组成,可以用来训练输入数据之间

的相似程度,对于没有出现的样本数据也可以进行分类^[9]。另外,孪生神经网络可以生成样本数,提升样本量,避免过拟合现象。

综上所述,本文提出了基于 LSTM 的孪生神经网络对不同厂商不同型号的电抗器进行故障预警,网络结构如图 2 所示。

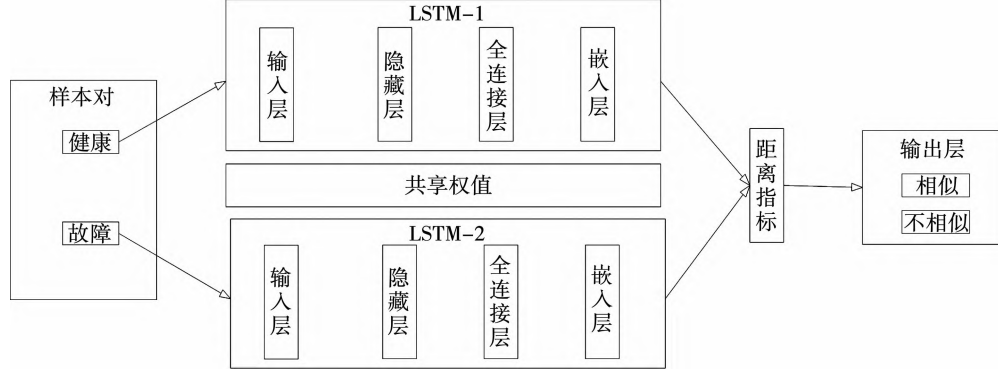


图 2 基于 LSTM 的孪生网络结构

由图 2 可知,从样本池中随机选择样本数据,可以将其表示为 $\langle I, I' \rangle$ 。两个 LSTM 网络接收两个电抗器运行状态值样本对,两个 LSTM 网络共享权重,对映射之后的样本进行相似度对比,每个 LSTM 包含一个输入层 U ,四个隐藏层 H ,一个具有 128 维的全连接层 D 和一个嵌入层 E 。LSTM 针对时间序列向量进行操作,获取电抗器状态的历史信息。在全连接层后连接一个 256 个神经元的嵌入层,并加一层来计算孪生网络的距离。最后利用 sigmoid 函数^[10]进行归一化,并对输入样本对的特征向量进行样本相似度计算。

模型学习过程:假设 $E_\omega(I)$ 和 $E_\omega(I')$ 是两个网络输入样本对 I 和 I' 由嵌入层网络函数 E_ω 计算后在嵌入空间中的投影,输出层用来计算孪生网络相似度值。计算公式如式(1)所示:

$$ED_\omega(I, I') = f_{sig} \left(\sum_{i=0}^{N-1} \beta_i |E_\omega^{(i)}(I) - E_\omega^{(i)}(I')| \right) \quad (1)$$

其中, f_{sig} 是 sigmoid 函数, β_i 是模型在训练期间学习的附加参数,假设 N 是数据集 $D = (I^{(i)}, I'^{(i)}, Y^{(i)})$ 电抗器运行状态值样本对的总数,其中 i 代表第 i 个输入训练对, $Y(I^{(i)}, I'^{(i)})$ 是对应的真实标签。当 $I^{(i)}, I'^{(i)}$ 是相同的电抗器状态标签时则

$Y(I^{(i)}, I'^{(i)}) = 1$, 否则记 $Y(I^{(i)}, I'^{(i)}) = 0$ 。预警模型总损失函数由式(2)给出:

$$L_\omega(D) = \lambda \|\omega\| + \frac{1}{2N} \sum_{i=0}^{N-1} \ell_\omega^{(i)}(I^{(i)}, I'^{(i)}, Y^{(i)}) \quad (2)$$

ω 是神经网络的权重, λ 是权重衰减,在训练模型时将 λ 设置为 0.001。样本损失函数 $\ell_\omega^{(i)}$ 包含以下两项:相似($Y=1$)的情况 L_s 和不相似($Y=0$)的情况 L_d ,如式(3)所示:

$$\ell_\omega^{(i)} = Y^{(i)} \ell_s(I^{(i)}, I'^{(i)}) + (1 - Y^{(i)}) \ell_d(I^{(i)}, I'^{(i)}) \quad (3)$$

相似和不相似情况下的损失函数由式(4)给出:

$$\begin{aligned} l_s(I, I') &= (ED_\omega)^2 \\ \ell_d(I, I') &= \begin{cases} (m - ED_\omega)^2, & m < ED_\omega \\ 0, & \text{otherwise} \end{cases} \end{aligned} \quad (4)$$

其中, m 是一个边距距离,以区分两个样本的相似程度。

1.2 样本池模块

样本池模块收集电抗器运行状态值实例,包括已标记和未标记的实例,从 T_0 天开始收集,当在 T_f 天发生故障,则采集 T_0 天到 T_f 天的电抗器运行

状态值数据,因此,在 t 时间点收集的向量可以表示为: $I^t = \{I_0^t, I_1^t, I_2^t, \dots, I_a^t\}$ 。

LSTM 的输入数据是固定连续 14 天的电抗器行为状态,将 14 天数据作为一个训练样本,该样本数据带有标签,标记是健康状态或故障状态。

当发现电抗器发生故障时,即停止收集数据,并收集故障发生前 7 天的数据作为样本,标记为故障状态,即 $y = 1$ 。对于健康的电抗器样本数据,标记为 $y = 0$ 。电抗器实例、样本和输入样本对之间的关系,如图 3 所示。



图3 电抗器实例、样本和输入样本对之间的关系

孪生网络结构接收的输入样本是电抗器实例样本对 (I, I') , 可以自由标记不同的样本对, 与目前常见的方法对比, 有以下两个优势:

①可以自由组合样本, 减少数据不平衡性。不平衡程度是多数类样本与少数类样本的比率, 值越大说明样本越不平衡, 对于大小为 A 的样本集, 多数样本大小为 αA , 则 $y = 1$ 和 $y = 0$ 的数目如式(5)所示:

$$\begin{cases} n_1 = C_A^2 + C_{\alpha A}^2 \\ n_0 = C_A^1 + C_{\alpha A}^1 \end{cases} \quad (5)$$

训练数据新的不平衡度 IDe' 可以由式(6)给出:

$$IDe' = \frac{n_1}{n_0} = \frac{\alpha}{2} - \frac{1 + \alpha - A}{2A\alpha} \quad (6)$$

变量 $A, \alpha > 1$, IDe' 的值将在 $\frac{\alpha}{2}$ 左右, 不平衡度减少 2 倍。与采样方法对比, 该方法将标记的样本输入作为样本对, 不会丢失采集的数据。

②通过本方法组合训练样本时, 样本对数量成倍增加, 新数据集的样本对数量如式(7)所示:

$$M = \frac{[A(1 + \alpha)]!}{2! [A(1 + \alpha) - 2]!} \quad (7)$$

本文方法可以极大提升训练样本数量, 利用深度学习算法预警电抗器故障, 避免因缺少训练样本而出现拟合现象。

1.3 决策模块

对于目标电抗器,用同一厂商所有带标签的训练样本形成相应的样本输入对,检测步骤为:

①将标记为 $y = 1$ 的故障电抗器与目标电抗器作为样本对输入,如果预警结果与样本相似,则标记为 $y = 1$,否则进入第二步。

②为了提升决策的效率,选择同一型号占总量 10% 的健康电抗器,以月为单位收集健康电抗器样本($y = 0$),与目标电抗器形成样本对输入,如果预警结果为相似,则标记为 $y = 0$ 。否则进行第三步。

③将同一厂商不同型号的故障电抗器与目标

预警电抗器作为样本对进行预警,如果预警结果为相似,则标记为故障电抗器 $y = 1$ 。

④与第二步相似,将来自同一厂商的不同型号样本与目标电抗器作为样本对输入进行预警。如果样本池中不存在目标样本的电抗器型号时,直接从第三步开始执行。

由于每个目标待预警电抗器由多个在连续时刻的待预警样本组成,这些样本的输出结果仅代表某一时刻的电抗器运行状态,而不是一段时间内的电抗器运行状态。为此,引入滑动窗口投票机制对电抗器的整体故障情况进行预警。基于投票滑动窗口方法^[11]的总体结构如图 4 所示。

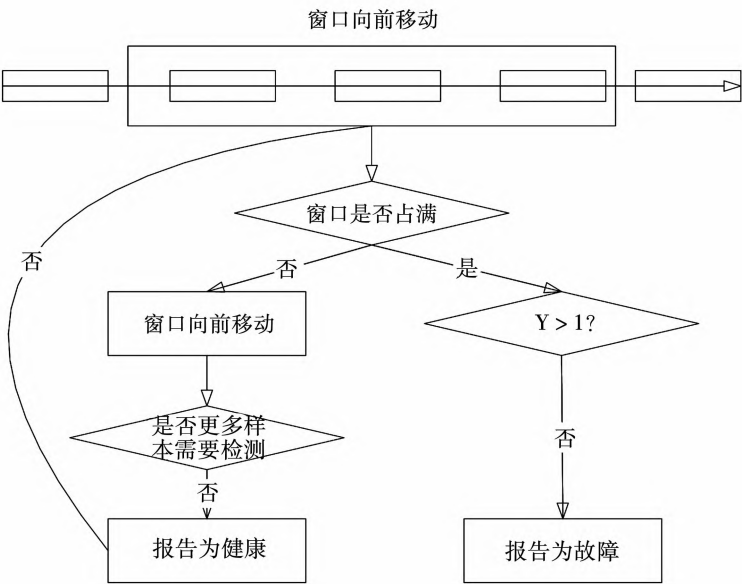


图 4 基于投票滑动窗口方法的总体结构

依据决策模块的预警结果,定义长度为 W 的滑动窗口,每天会向前移动一个单位,如果在滑动窗口中有决策模块步骤一和步骤三的预警结果则为故障,进行报警。否则不是故障,不报警。

2 实验与结果分析

2.1 实验设置

为验证 HDDse 的有效性,本实验选择南网某 500kV 变电站高压并联电抗器集。数据集包括连续 58 个月的电抗器数据,146203 个健康电抗器和 8256 个故障电抗器,使用其中的 80% 做训练数据集,剩下 20% 做测试样本集。对这些电抗器

进行标记,故障的电抗器标记为 1,健康的电抗器标记为 0。

将样本对按照 8:2 的比例分别作为训练集和测试集,确保训练集和测试集相互独立,同时为减少模型的不稳定性,HDDse 通过交叉验证的方法验证实验。实验的迭代次数最大设置为 1000 次,权重衰减系数为 0.001。

2.1.1 评价指标

FDR: (Failure Detection Rate,故障检测率,也被称为召回率)^[12]。即正确预警故障电抗器的比例。FDR 越高,模型预警能力越强^[13]。

FAR(False Alarm Rate,误报率)^[14]。是被错误预警为故障的健康电抗器的比例。FAR 越低,模型预警能力越强^[15]。

F-Measure 值是两个指标 FDR 和预警精度 (Prediction Precision) 之间的一个平衡^[16]。F-Measure 越高,模型的预警能力越好。

C-MTTDL 是一种基于成本的平均数据丢失时间指标 (Mean Time To Data Loss, MTTDL) ,用来评估模型的可靠性和可用性^[17]。

2.1.2 建模过程

①构建基于 LSTM 的孪生网络的结构,获取电抗器状态的历史信息,利用 sigmoid 函数进行归一化^[18],并对输入样本对的特征向量之间进行样本相似度计算。

②收集电抗器运行状态值实例,包括已经标记和未标记的实例,对样本数据进行标记,故障状态为 $y = 1$,健康状态为 $y = 0$ 。并利用孪生结构网络增加样本对数量。

③利用决策模块对电抗器的故障进行判定,引入滑动窗口的投票机制对目标电抗器作出最终的故障预警。

2.2 实验结果

对本文算法的电抗器故障预警准确率和召回率进行实验,实验结果分别如图 5 和图 6 所示。图 5 的横坐标表示的是实验的次数,纵坐标是真实分类,由图 5 可知,本文方法的电抗器故障预警准确率在 95% 以上,体现了较好的预警准确性。图 6 横坐标表示的是实验的次数,纵坐标是召回率结果,由图 6 可知,本文方法电抗器故障预警召回率都在 90% 以上,体现了较好的召回率。

此外,将 HDDse 方法和基于距离的异常检测方法 (Distance-based Anomaly Detection, DAD)^[18]、基于浅层的机器学习预警方法 (Shallow Machine Learning, SML)^[19]、基于单分类器预警方法 (One-Class Classification, OCC)^[20] 和基于深度神经网络预警方法 (Deep Neural Network, DNN)^[21] 等方法从小样本预警性能、模型的可靠性、模型的效率等方面进行评估,具体如下。

2.2.1 故障预警性能

不同方法之间小样本电抗器故障预警的性能

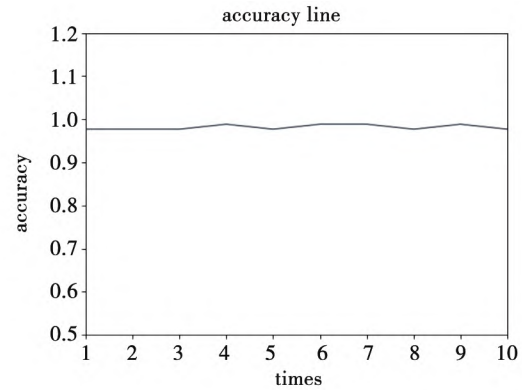


图 5 HDDse 方法电抗器故障预警准确率

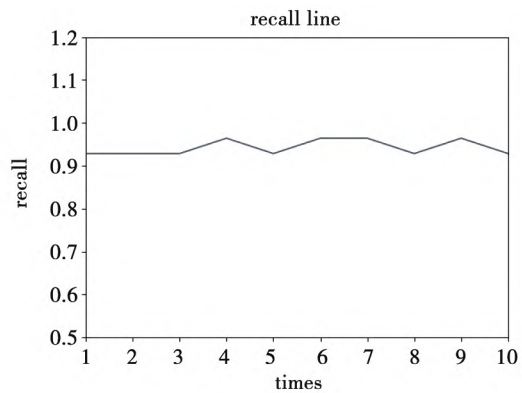


图 6 HDDse 算法的召回率

比较如图 7 所示。

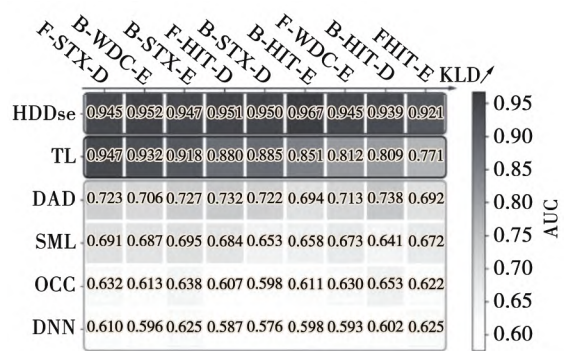


图 7 小样本电抗器故障预警的性能比较

将 HDDse 与 DAD、SML、OCC、DNN 等方法进行比较。由图 7 可知,横坐标是电抗器型号,已经按照 KLD 值 (KLD 值衡量的是数据集之间的分布差异) 进行排序。因 DAD、SML、OCC、DNN 等方法出现小样本过拟合现象,导致无法获得较高 AUC 值 (AUC 是一种针对各种不同阈值设置下的

分类问题的性能度量,AUC 越大就意味着该模型越能更好地区分故障电抗器和健康电抗器)。DNN 方法结果最差,因为神经网络需要大量的训练样本才能较好的拟合出神经网络中的参数权重,小样本数据集无法做到这点。HDDse 方法可以生成大量样本数据,避免过拟合现象而具有较好的预警性能。此外,基于 LSTM 的孪生网络学习方法对于异构的电抗器预警分类也有较好的效果。

2.2.2 模型的可靠性

模型可靠性结果如表 1 所示。

表 1 C-MMTDL 的改善结果

方法	FDR(%)	C-MTTDL	MTTDL
OCC	62.6	1.94	397.6
DAD	45.2	2.89	276.7
SML	72.6	3.37	504.10
DNN	85.3	7.13	814.13
本文方法	95.8	419.35	1656.3

如表 1 所示,HDDse 获得非常理想的 AUC 和 F-Measure。HDDse 与其他四个方法相比可以将 C-MTTDL 提高 2 个数量级,说明本文方法获取较好的预警性能,同时提升故障预警的可靠性。

2.2.3 模型效率

将本文方法与 DAD、SML、OCC 和 DNN 进行对比分析,不同方法模型训练和预警时间比较如图 8 所示。由图可知,DAD 方法训练方法比较简单,计算量较大,预警时间较高。HDDse 与 SML 方法相比花费了第二高的训练时间,与 DAD 方法相比则花费了第二高的在线预警时间。这是由于 HDDse 生成了更多的训练样本,因此需要更多的时间来确保模型收敛。考虑到电抗器故障预警频率为每周至每月进行一次,这样的时间成本是可以接受的。

2.3 实验结论

从实验结果可知,本文提出的基于 LSTM 的孪生网络学习方法对于电抗器故障预测的准确率以及召回率都有较好的表现;此外,基于 LSTM 的孪生网络学习方法对于异构的电抗器预警分类也有较好的效果;且在电抗器故障预测的可靠性以及效率方面也表现较好。

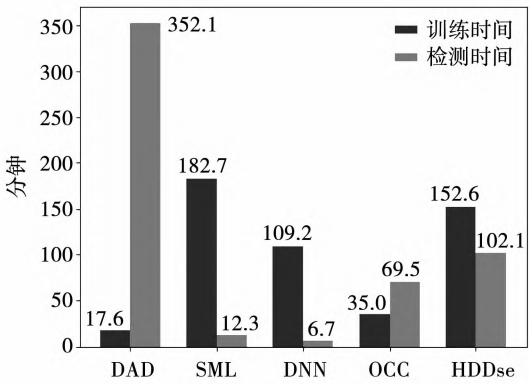


图 8 不同方法模型训练和预警时间比较

3 结束语

由于目前电抗器故障预警方法存在着样本较少导致的过拟合、性能不佳、模型可靠性差等问题,为此,提出基于长短期记忆神经网络的孪生神经网络方法 HDDse。HDDse 方法包括了样本池模块、电抗器故障预警模块以及决策模块。样本池模块存储动态收集的电抗器数据实例,形成 HDDse 训练数据集。模型训练时,带有标签的电抗器样本用于训练,没有标签的数据集作为预警数据集。故障预警模块是 HDDse 的核心模块,用于电抗器故障预警,样本池模块数据作为电抗器故障预警模块的输入,输出二分类的预警结果。决策模块将若干连续样本的预警结果映射到运行健康状况模块。通过实验表明本文方法相比目前已有研究方法在性能和可靠性方面表现更好。

参考文献:

[1] 朱庆龙,郑智勇,杨廷方. 基于 LSTSVR 模型的并联电抗器油中溶解气体浓度预警[J]. 机电信息,2020,13(30):2.

[2] Chen X,Liang Y,Wang G,et al. Parameter optimization design method for a fast-switch-based fault current limiter and circuit breaker[J]. International Journal of Electrical Power & Energy Systems,2020,41(27):24-25.

[3] 郭哲,郭志广,贾跟卯,等. 空心并联电抗器匝间故障三阶段特性分析及综合保护方法[J]. 电力系统自动化,2020,44(8):8.

[4] Hassan A,Mokhtar Z,Maskin M. Time dependent reliability analysis for a critical reactor safety system based on fault tree approach[J]. IOP Conference Series: Materials Science and Engineering,2022,1231(1):012015.

[5] 郭约法·马静勇. 500kV 纵宝线串联电抗器本体故障

- 类型及保护动作分析[J]. 电力系统装备, 2018(8): 146 – 147.
- [6] Luo X, Huang X, Zheng W, et al. Analysis of a dry-type reactor fault combined with COMSOL thermal field simulation and high frequency pulse oscillations [J]. Electric Power, 2018, 31(17): 108 – 109.
- [7] 高树国, 张明文, 吴书煜, 等. 基于振动感知的油浸式电抗器松动故障检测方法[J]. 振动与冲击, 2022, 23(41): 23 – 25.
- [8] 欧健良, 杨鹏, 曾智, 等. 500kV 高压并联电抗器的故障分析与防范措施[J]. 变压器, 2022, 59(2): 5.
- [9] Pang H, Xuan Q, Xie M, et al. Research on target tracking algorithm based on siamese neural network [J]. Mobile Information Systems, 2021, 23(4): 1 – 11.
- [10] Xiao Y, Yin H, Zhang Y, et al. A dual-stage attention-based Conv-LSTM network for spatio-temporal correlation and multivariate time series prediction [J]. International Journal of Intelligent Systems, 2021, 34(11): 45.
- [11] Tao C, Lu J, Lang J, et al. Short-term forecasting of photovoltaic power generation based on feature selection and bias compensation-LSTM network [J]. Energies, 2021, 14(12): 71.
- [12] Petmezas G, Haris K, Stefanopoulos L, et al. Automated atrial fibrillation detection using a hybrid CNN-LSTM network on imbalanced ECG datasets [J]. Biomedical Signal Processing and Control, 2021, 63(11): 67 – 68.
- [13] Gao X, Zha S, Li X, et al. Incremental prediction model of disk failures based on the density metric of edge samples [J]. IEEE Access, 2019(99): 1.
- [14] Shen J, Wan J, Lim S J, et al. Random-forest-based failure prediction for hard disk drives [J]. International Journal of Distributed Sensor Networks, 2018, 14(11): 63.
- [15] Rincon C C A, Pris J F, Vilalta R, et al. Disk failure prediction in heterogeneous environments [C]. 2017 International Symposium on Performance Evaluation of Computer and Telecommunication Systems. IEEE, 2017: 7.
- [16] Xie Y, Feng D, Wang F, et al. DFPE: explaining predictive models for disk failure prediction [C]. 2019 35th Symposium on Mass Storage Systems and Technologies (MSST). IEEE, 2019: 193 – 204.
- [17] Kara A. Multi-scale deep neural network approach with attention mechanism for remaining useful life estimation [J]. Computers & Industrial Engineering, 2022, 23(11): 251 – 254.
- [18] Zhang J, Zhou K, Huang P, et al. Minority disk failure prediction based on transfer learning in large data centers of heterogeneous disk systems [J]. IEEE Transactions on Parallel and Distributed Systems, 2020, 37(23): 1 – 10.
- [19] Wang H, Xu G Y, Zhou X, et al. Electrical parameters change of dry-type reactor under inter-turn fault [J]. Power & Energy, 2018, 41(12): 609 – 611.
- [20] Zhang C, Zhang K, Yuan Q, et al. GMove: group-level mobility modeling using geo-tagged social media [J]. ACM, 2016, 23(14): 39 – 48.
- [21] Chang, Xu, Gang, et al. Health status assessment and failure prediction for hard drives with recurrent neural networks [J]. IEEE Transactions on Computers, 2016, 65(11): 3502 – 3508.

(责任编辑: 丁玥)

(上接第 75 页)

- [7] Shi B, Bai B, Zhou Z, et al. 3DDeepPanoramic Representation for 3-D Shape recognition [J]. Signal Processing Letters. IEEE, 2015: 2339.
- [8] 金妍君, 万旺根. 基于注意力机制的三维模型多视图分类算法研究[J]. 工业控制计算机, 2020, 33(6): 54 – 56.
- [9] Su H, Maji S, Kalogerakis E, et al. Multi-view convolutional neural networks for 3D shape recognition [C]. In: Proceedings of IEEE International Conference on Computer Vision, Santiago, 2015: 945 – 953.
- [10] 丁博, 伊明. 基于卷积神经网络的三维 CAD 模型分类[J]. 哈尔滨理工大学学报, 2020, 25(1): 66 – 72.
- [11] 李根龙. 面向机器人装配的目标识别与位姿获取技术研究[D]. 西安: 西安理工大学, 2021.
- [12] 黄南天, 杨学航, 蔡国伟, 等. 采用非平衡小样本数据的风机主轴轴承故障深度对抗诊断[J]. 中国电机工程学报, 2020, 40(2): 563 – 573.

(责任编辑: 杨静)