

基于机器学习的变压器预测性维护方案

王子祥, 王恩荣, 张海龙

南京师范大学电气与自动化工程学院, 江苏 南京

收稿日期: 2023年3月14日; 录用日期: 2023年6月23日; 发布日期: 2023年6月30日

摘要

为了改善现有的配电变压器维护方案, 更好地实现电力大数据的应用, 本文提出了一种基于机器学习的变压器预测性维护方案, 以变压器油中溶解气体数据特征数据, 首先对变压器原始采集数据进行处理, 然后使用隐半马尔可夫模型(HSMM)确定变压器的运行状态, 进一步使用改进卷积神经网络对异常数据进行分类预测, 最终实现辅助变电运维人员进行非周期定向维护, 从而有效降低运维成本、减少机组停机时间, 提高变压器的任务完成率的需要。

关键词

变压器, 预测性维护, 故障诊断, 隐半马尔可夫模型, 改进卷积神经网络

Machine Learning Based Predictive Maintenance Solution for Transformers

Zixiang Wang, Enrong Wang, Hailong Zhang

School of Electrical & Automation Engineering, College of Nanjing Normal University, Nanjing Jiangsu

Received: Mar. 14th, 2023; accepted: Jun. 23rd, 2023; published: Jun. 30th, 2023

Abstract

In order to improve the existing maintenance scheme for distribution transformers and better realise the application of big data in electricity, this paper proposes a machine learning-based predictive maintenance scheme for transformers, using data characterised by dissolved gas data in transformer oil, firstly processing the original transformer collection data, then using a hidden semi-Markov model (HSMM) to determine the operational status of the transformer, and further using an improved Convolutional neural networks are used to classify and predict abnormal data,

ultimately realising the need to assist substation operators and maintenance personnel to carry out off-cycle targeted maintenance, thereby effectively reducing operation and maintenance costs, reducing unit downtime and improving the task completion rate of transformers.

Keywords

Transformer, Predictive Maintenance, Fault Diagnosis, Hidden Semi-Markov Model, Improved Convolutional Neural Network

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

电力设备是保证电力系统正常运行的关键，而电力变压器作为变电站中用于电压转化的关键设备，是电力系统的核心，其健康水平和运行可靠性与电力系统稳定性息息相关。目前使用的配电变压器维护方案通常为周期性的预防维护，存在着维护成本浪费，信息更新缓慢的缺陷。因此，有必要研究一种更具维护效率的变压器运维方案作为替代。

本文利用变压器运行过程中运行参数逐步变化的特点，设计了一种基于机器学习的变压器预测性维护方案，以变压器油中溶解气体数据特征数据，首先对变压器原始采集数据进行处理，使用隐半马尔可夫模型(HSMM)确定变压器的运行状态，进一步使用改进卷积神经网络对异常数据进行分类预测，并通过测试集数据对模型预测准确率进行了验证，验证结果证明了该方案在变压器运维上的可行性。

2. 预测性维护研究背景

预测性维护是以状态为依据的维修，其目标是利用先进的科学技术，有计划、有组织、有针对性地对设备进行状态检测和故障诊断，在其性能退化的初期，尤其在还没有造成重大危害时，发现早期故障隐患，定量评价设备健康状态，预测设备的剩余寿命，并在此基础上确定对设备的最佳维护时机，是一个十分重要和实用的课题，这一技术的关键在于数据数据采集与处理。

根据现有研究的成果，机械设备预测性维护方法分为三种类型，分别是基于物理模型、知识和数据的预测方案，其中基于数据的健康预测方法可以通过数理统计和数据挖掘的方法分析设备劣化的内在规律。这种预测性维护方案基于大量的设备运行数据，建立设备未来的运行状态预测模型，以设备故障率或可靠度作为制定维护方案的依据。学者 Bunks 率先提出把 HMM 的概率用于机械设备的状态预测[1]，Baruah 和 Chinnam 使用基于 HMM 的方法来对传感器信号进行建模，不仅能确定系统的运行状态，还可以对设备的有效剩余寿命进行评估[2] [3]，刘波和宁芊等人使用连续型 HMM 对轴承的退化状态进行分析，在划分轴承为若干个退化状态后使用 PSO-SVM 进行预测模型的训练，能实现对轴承剩余寿命的有效预测[4]。对于一些需要考虑时间参数的问题，一些学者使用 HSMM 来代替 HMM，这种方法比原版添加了一个时间组件，可以描述出状态持续时间的概率问题。曾庆虎等人基于 HSMM 对机械设备的演化规律进行了分析与建模，并在滚动轴承的故障预测中验证了方案的可行性[5]。Wang 等人针对设备进行状态识别和故障预测的问题，提出了一种基于时变转移概率的 HSMM 模型，并通过实验证明该方法优于现存的传统方法[6]。刘勤明，李亚琴等人设计了针对设备健康诊断与寿命预测问题的自适应隐式半马尔可夫模型(AHSMM)，经验证表明相比传统的 HSMM 更有效[7]。

3. 变压器维护

预测性维护应用在变压器日常维护上的应用可以有效的降低运维成本，减少机组停机时间，是提高变压器的任务完成率的最佳选择。变压器的维护方案发展可以分为三个阶段，分别是被动维护，预防性维护和预测性维护。早期的被动维护不必关注设备的运行状态，仅在设备发生故障后，才进行事后的维修操作，既存在很高的安全隐患还会产生高昂的维护成本，目前已经被淘汰。现在常用的预防性维护是通过周期性的巡检，定时主动的对设备的状态进行分析，虽然提高了设备的安全性，但还存在一定的成本浪费与故障风险。基于设备状态的预测性维护方案是目前最新的研究成果，它基于数据与分析实现故障前的设备维护，在安全性和经济性两方面都实现了合理优化，是目前进行设备维护的最佳方案。

目前，油浸式变压器的正常寿命大约是 20 至 25 年，而变压器的设计寿命一般都长于其使用寿命[8]。为了节约成本，提高设备的利用效率，有必要对变压器运行可靠性和维护方案进行研究。目前常见的变压器故障分析方法有油中溶解气体分析法、振动分析法、红外图像分析法等等[9]，其中油中溶解气体的气相色谱分析(Dissolved Gas Analysis, DGA)已经成为电力变压器状态诊断与故障分类的标准手段，在发现变压器内部潜伏性故障，判断变压器内部材料老化程度的方面有着显著的优势。本文就以变压器 DGA 数据为依据，建立了针对配电变压器的预测性维护方案。

4. 基于 HSMM 的变压器状态诊断

HSMM 在语音识别、自然语言处理和模式识别等领域取得了广泛的应用，其用于解决序列化问题的数据处理技巧在设备健康状态预测方面有很好的应用效果。一般来说，设备的健康状态是随时间序列而逐步劣化的，当退化达到一定的程度就会引发故障。对油浸式变压器来说，在变压器正常运行中，内部绝缘油会逐步分解，异常变换还会导致反应加剧，累计引发变压器故障。使用 HSMM 可以对此类劣化过程进行建模，基于数据判断变压器后续的运行工况发展。

设备的状态识别是实施设备健康预测的基础，只有识别出设备的健康状态后，才能对该状态下设备的未来运行状态进行预测，继而实现对设备的维护安排。基于变压器的历史运行数据，可以基于实时数据分辨变压器的运行状态，对后续运行中的设备故障率进行预测，其具体流程如图 1 所示。



Figure 1. HSMM state classification flow chart

图 1. HSMM 状态分类流程图

HSMM 的模型参数为

$$\lambda = (\pi, A, B, D)$$

式中， A 为状态转移概率矩阵， B 是观测概率矩阵， π 是初始状态概率向量， D 为持续时间分布。

一般来说，变压器在使用过程中会历经很多不同的健康状态转换点，在 HSMM 中，针对任意状态序列都能够估算出其相应的健康状态转换点，根据状态转换点分布，可以得到变压器运行在何种状态下，并根据状态转移概率矩阵预测设备故障率。

基于 HSMM 的变压器运行状态分析的计算需要解决三个基本问题：评估(也是分类)问题、解码(也是

识别)问题和学习(训练)问题。虽然可以采用最直接的枚举法处理评估问题,但过于庞大的数据量往往导致无法寻找到全局最优解,因此需要使用不同的算法解决相对于的问题。

使用前向-后向算法可以解决分类问题,可以求出在给定的观测序列 $O = o_1 o_2 \cdots o_T$ 和参数值序列 $\lambda = (\pi, A, B, D)$ 的 HSMM 中,该观测序列出现的概率大小,可以根据给定的数据信息求解设备状态的转换概率与平均驻留时间。Viterbi 算法可以解决 HSMM 中实际的问题,在给定的观测序列中寻求最优的状态序列。Baum-Welch 算法可以对计算得出的模型参数进行重估,最终确定变压器的初始状态分布、状态转移概率和状态驻留时间分布,使之更符合历史数据呈现的运行规律。

以某型号变压器为例,其同型号变压器的油中溶解气体数据记录如图 2 所示。

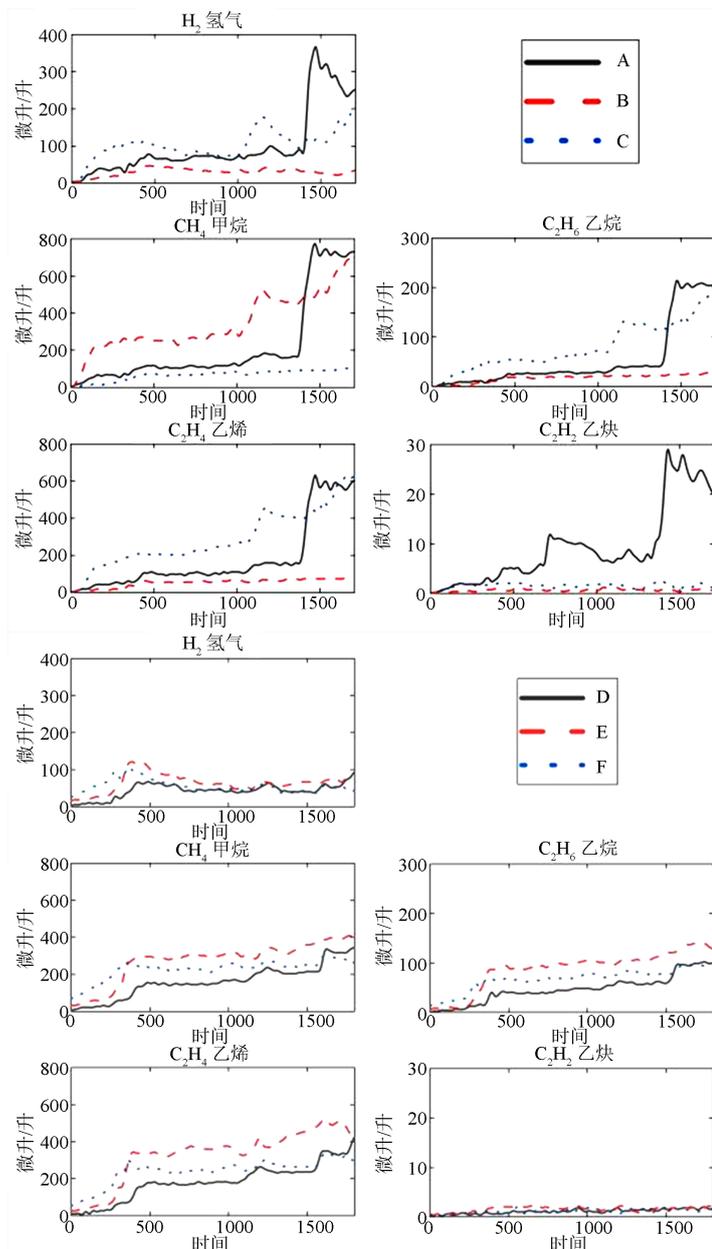


Figure 2. DGA data changes of six transformers of a certain model

图 2. 某型号六台变压器 DGA 数据变化情况

共得到六台同型号变压器运行中 1678 个阶段的 DGA 数据, 根据 IEEE 中有关于油浸式变压器中产生气体的解释指南可以根据 DGA 数据对油浸式变压器进行风险分类, 设变压器运行分为 ABCD 四种状态, 分别为正常运行状态, 不正常运行状态, 预警状态和故障状态。按照基于 HSMM 的设备运行状态分析步骤进行建模, 计算每个检测周期间转移概率分布, 得到变压器状态转移概率矩阵如表 1。

Table 1. Transformer state transition probability matrix
表 1. 变压器状态转移概率矩阵

变压器状态	状态转为 A	状态转为 B	状态转为 C	状态转为 D
A	0.9916	0.0084	0	0
B	0.2873	0.7035	0.0092	0
C	0.0819	0.3110	0.5868	0.0203
D	0	0	0	1

由以上状态转移概率估计得到变压器在各个状态的驻留时间如表 2:

Table 2. Average dwell time of each transformer state
表 2. 变压器各状态平均驻留时间

变压器状态	状态 A	状态 B	状态 C	状态 D
驻留均值	472.45	256.79	84.34	∞

根据表 1 中的变压器状态转移概率矩阵, 即可对变压器在检测时候后各时间段内的故障概率做出预估。由状态转移矩阵可知, 当变压器的运行状态转变为状态 D, 即故障状态时, 后续运行中不会自然恢复至前三类运行状态。根据建立的 HSMM 得到的概率矩阵创建马尔可夫链, 对变压器在后续 80、160 个检测周期内的运行状态进行模拟, 得到变压器在计划周期内转变为故障状态的概率分布, 即为变压器在计划周期内的故障率曲线, 如图 3 所示:

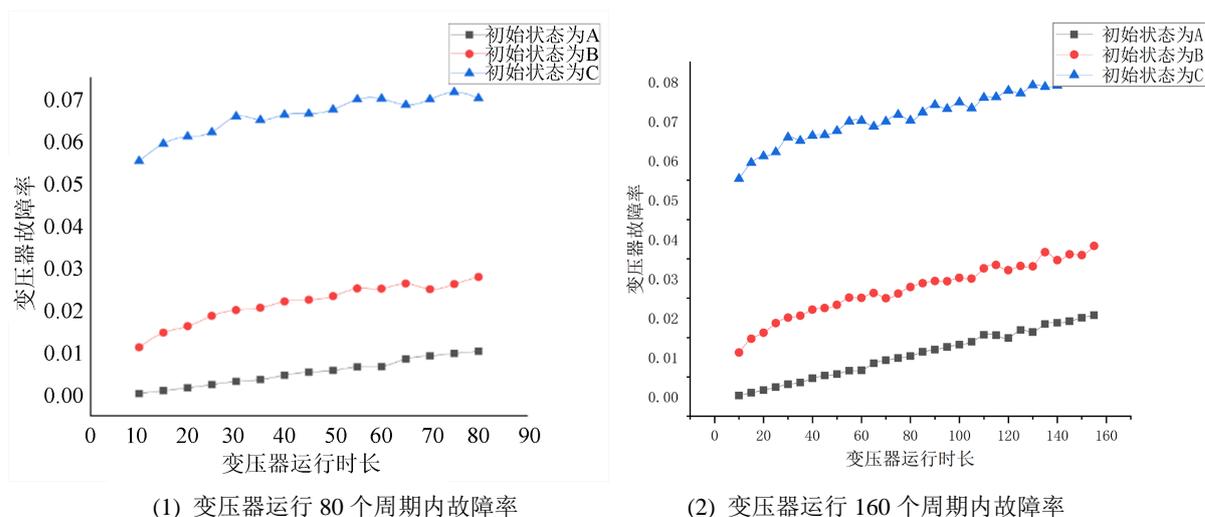


Figure 3. Transformer failure rate prediction based on HSMM
图 3. 基于 HSMM 的变压器故障率预测

当判断变压器运行状态为 A, 其故障率基本在 80 个检测周期内低于 1%, 在 160 个检测周期内低于

2.5%；当变压器运行状态为 B，故障率比于正常状态略高，随着运行时长增加更为明显；当变压器运行状态处于状态 C 时，其设备故障率始终在较高状态，在 10 个周期后就高于 5%，若不进行相关检修其故障风险还会不断提高，因此有必要在变压器运行状态进入状态 C 之后就对变压器进行相关维护。

考虑到变压器的运维安排和设备重要性，可以设置合理设置变压器的维护阈值，当变压器预期故障率低于维护阈值时，直至下一次检测周期都无需对变压器进行维护，避免过度维护；当模型预测的变压器故障率高于所设阈值时，则有必要对变压器进行针对性维护。在设计维护计划时，为了提高对变压器运行状态的了解，减小错误维护的发生几率，有必要对变压器的潜在故障类型进行判断，基于通道注意力机制的改进卷积神经网络(SE-Res-CNN)可以实现这一分类任务[10]。

5. 基于 SE-Res-CNN 的变压器故障分类模型

卷积神经网络(CNN)是一种前馈神经网络，其各层之间的连接线受动物视觉皮层的启发，主要可以分为三层：卷积层、池化层和完全连接层。这些模块可以用于对数据的特征学习，完成预测模型的建立。卷积层的基本单元为神经元，神经网络由大量的神经元相互连接而成。每个神经元接受线性组合的输入后，最开始只是简单的线性加权，后来给每个神经元加上了非线性的激活函数，从而进行非线性变换后输出。每两个神经元之间的连接代表加权值，称之为权重(weight)。不同的权重和激活函数，则会导致神经网络不同的输出。

神经网络的每个神经元如图 4 所示：

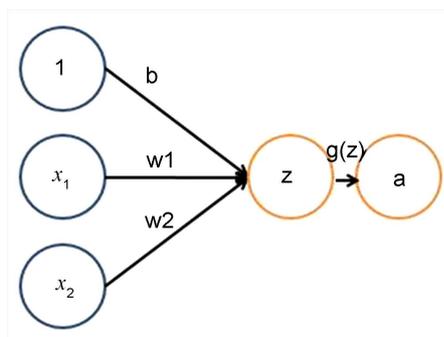


Figure 4. Schematic diagram of neurons
图 4. 神经元示意图

传统 CNN 为了提高其特征表达能力，设置的通道数较大，这样的深度 CNN 虽然可以提高性能，但并不是所有通道都富含故障信息，一些通道中学习的甚至是噪声的分布特征，对通道的无差别对待会导致 CNN 结构复杂并造成算力的浪费，影响 CNN 的学习效率。针对这样的不足，有学者提出了考虑通道注意力机制的卷积神经网络算法。注意力机制以权重的生成和分配重新调制网络参数，并通过训练最终使网络聚焦于关键特征，从而优化 CNN 的学习机制。在调制网络的同时不损失原有的故障特征、避免负优化的情况发生是使用注意力机制的前提。

在神经网络中，注意力机制能够硬性选择输入的某些部分，或者给输入的不同部分分配不同的权重。使用注意力机制能够从大量的信息中筛选出重要的信息，在神经网络中引入注意力机制有很多种方法，SE 注意力机制(Squeeze-and-Excitation Networks)是在通道维度增加注意力机制的一类方法，它可以通过自动学习的方式，使用另外一个新的神经网络，获取到特征图的每个通道的重要程度，然后用这个重要程度去给每个特征赋予一个权重值，从而让神经网络重点关注某些特征通道。提升对当前任务有用的特征图的通道，并抑制对当前任务用处不大的特征通道。

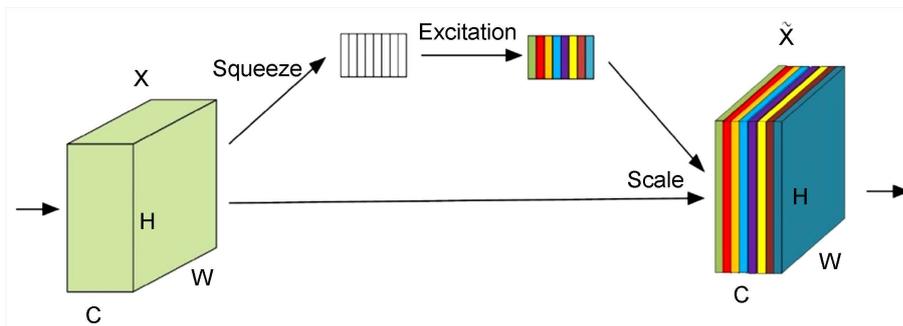


Figure 5. Structure of channel attention
图 5. 通道注意力结构图

SE 注意力机制的实现的实现的关键操作在于 Squeeze 和 excitation (图 5)。在在输入 SE 注意力机制之前(左侧 C)，特征图的每个通道的重要程度都是一样的，通过 SENet 之后(右侧彩色 C)，不同颜色代表不同的权重，使每个特征通道的重要性变得不一样了，使神经网络重点关注某些权重值大的通道。其核心思想在于通过全连接网络根据 loss 来自动分类特征权重，虽然在一定程度上增加了一些参数和计算量，但却可以获得更好的学习效果。

残差网络(Res)则可以用于缓解卷积神经网络中，随着网络深度增加产生的过拟合与泛化的问题，基本步骤是对两层卷积层的进行残差计算，并将差异信息作为学习的目标，使用的 Res 基本结构如图 6 所示。

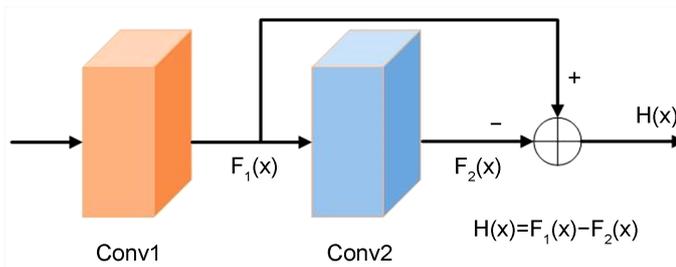


Figure 6. Res network structure diagram
图 6. Res 网络结构图

最终所用的 SE-Res-CNN 模型结构如图 7 所示。

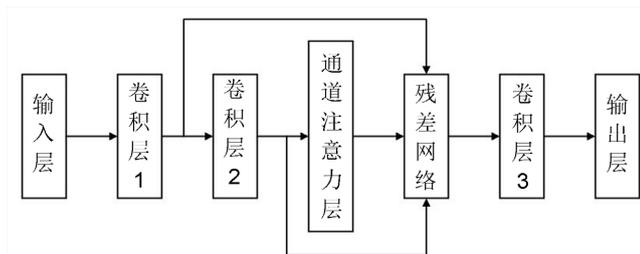


Figure 7. SE-Res-CNN model structure
图 7. SE-Res-CNN 模型结构

使用基于以上机制的学习模型，对同型号变压器的故障数据进行学习，采集到的数据同样为各状态下变压器油中溶解气体数据。所用的数据样本共 1258 条，分析的气体共有五类，包含 H_2 ， CH_4 ， C_2H_6 ，

C_2H_4 , C_2H_2 五类气体成分数据, 类型为中低温过热、高温过热、局部放电、低能放电和高能放电五类故障状态以及一类正常状态, 共六种类型, 其中正常状态数据 315 条, 中低温过热 216 条, 高温过热 229 条, 局部放电 96 条, 低能放电 157 条, 高能放电 245 条, 表 3 展示了部分原始数据内容。

Table 3. Early transformer failure data samples

表 3. 变压器早期故障数据样本

氢气 H_2	甲烷 CH_4	乙烷 C_2H_6	乙烯 C_2H_4	乙炔 C_2H_2	故障类型
1.38	6.15	9.21	76.63	6.63	高温过热
11.30	4.20	0.50	0.20	0.00	正常
5405.00	561.00	5.00	70.00	128.00	高能放电
12.15	40.50	18.90	28.35	0.09	中低温过热
1651.00	90.00	33.00	45.00	2.00	局部放电
...

使用归一化的五类气体数据作为特征数据, 将数据集分为训练样本和测试样本, 使用 SE-Res-CNN 进行模型训练。使用交叉验证的方式对模型进行验证, 得到混淆矩阵如图 8 所示, 主对角线为诊断结果与真实状态一致的数量, 非主对角线为误诊数量。

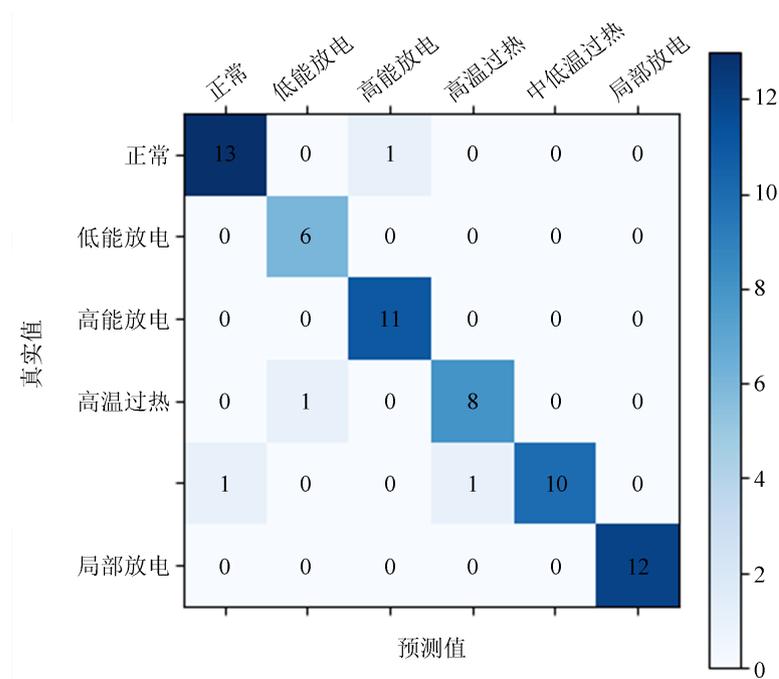


Figure 8. Results of SE-Res-CNN model test set

图 8. SE-Res-CNN 模型测试集结果

经测试, SE-Res-CNN 模型具有很高的诊断准确度, 在 64 组测试集数据上, 只有 3~4 次分类错误, 最终的准确率为 95% 左右。对相同的数据集以原始 CNN 网络、DBN 网络、仅考虑 SE 机制的 CNN 和仅考虑 Res 的残差网络进行训练, 得到各网络的预测准确率如表 4 所示。可以看出 SE-Res-CNN 模型诊断的准确率和稳定性均优于其他诊断模型。

Table 4. Accuracy of various fault classification methods
表 4. 各类故障分类方法准确率

故障分类方法	平均预测准确率
CNN	82%
DBN	90%
SE-CNN	90.24%
Res-CNN	91.17%
SE-Res-CNN	94.87%

6. 基于机器学习的变压器预测性维护方案

制定基于变压器健康状态评估的预测性维护计划是发配电企业的在新时期实现降低成本，减少运维过程碳排放，提高运维效率的有效方案，基于实时状态检测的运维方案不仅可以降低变压器定期维护的各项元件成本和人力成本，更可以有效的延长变压器设备的使用寿命，在时间维度上提高设备利用率。

根据上述两种数据分析模型，设计了如图 9 所示的基于机器学习的变压器预测性维护方案设计流程如下：

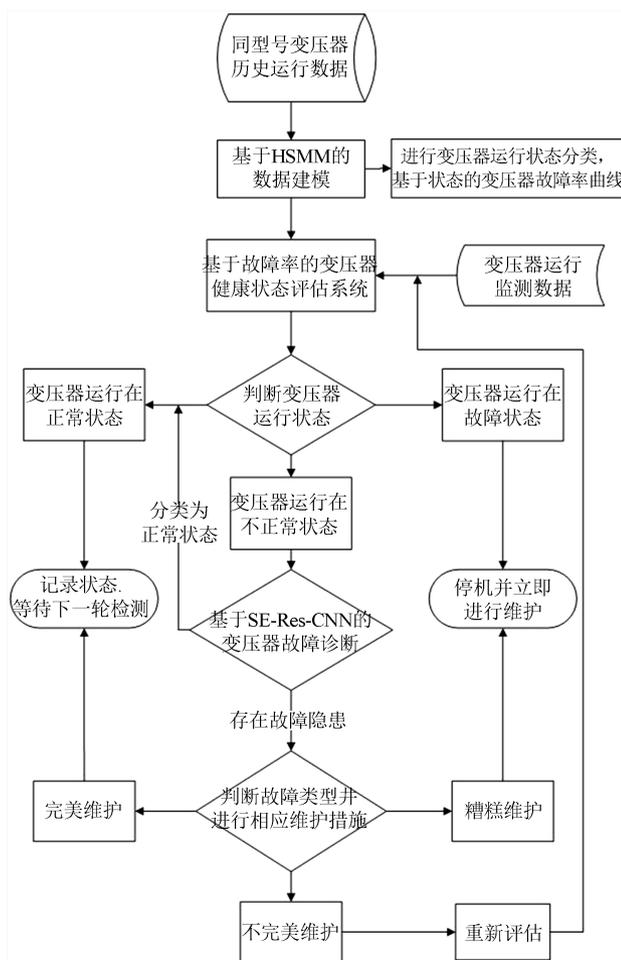


Figure 9. Overall flow chart of transformer health state assessment based on machine learning
图 9. 基于机器学习的变压器健康状态评估总体流程图

预测性维护方案的过程为:

- 1) 由同类型变压器机组的历史运行数据训练 HSMM 模型, 实现对变压器状态分类和参数计算。
- 2) 在变压器的运行过程中周期性测量相关的特征数据, 经由 HSMM 模型判断变压器机组的所属状态及基于时序的故障率分布。
- 3) 设立变压器故障率阈值, 预测故障率低于所设阈值时不进行维护, 预测故障率高于所设阈值时, 使用基于 SE-Res-CNN 的故障诊断模型判断故障类型, 运维人员基于诊断结果进行维护。
- 4) 运维人员在现场实际考察变压器机组运行状态, 根据数据预测结果和现场的实际情况对变压器进行维护, 在维护后对维修结果进行评定, 得出完美维护、不完美维护和糟糕维护的结论。对于判定为完美维护的维护结果, 认为维护操作已经完全解决了变压器的潜在风险, 系统中的状态可标记为正常状态, 记录归档后等待下一个检测周期。对判定为不完美维护的维护结果, 则需要在短期内重新获取检测数据, 进行变压器运行状态认定, 若仍存在故障风险需要重新进行维护。对判定为糟糕维护的维护结果, 则需立即停机检修, 彻底的消除故障隐患, 避免变压器故障的发生。

7. 结论与展望

1) 本文所提出的变压器预测性维护方案是基于数据驱动的预测性维护方案。在收集大量变压器历史运行数据基础上的, 通过 HSMM 确定变压器运行状态转移概率矩阵和运行状态平均驻留时间, 使用 SE-Res-CNN 模型对超过设定阈值的变压器 DGA 数据进行进一步分类, 模型总体分类效果较好, 预测准确率符合期望要求。

2) 使用预测性维护方案的变压器运维策略, 可以有效的避免滞后维护和过度维护, 对变压器的维护建立在两种模型的预测基础上。当变压器预测故障率低于所设阈值时, 无需对设备进行维护操作以减少过度维护, 当变压器故障率高于所设阈值时, 使用 SE-Res-CNN 分类模型对变压器数据再分析, 根据分类结果辅助变压器运维人员作出维护决策, 可以有效的避免滞后维护。

3) 在实际应用中, 可以对变压器的运行状态进行更加详细的划分, 新获取的数据也可以用于更新模型的相关参数, 达到更加理想的状态分析效果。

参考文献

- [1] Bunks, C., McCarthy, D. and Al-Ani, T. (2000) Condition-Based Maintenance of Machines Using Hidden Markov Models. *Mechanical Systems and Signal Processing*, **14**, 597-612. <https://doi.org/10.1006/mssp.2000.1309>
- [2] Baruah, P. and Chinnam, R.B. (2005) HMMs for Diagnostics and Prognostics in Machining Processes. *International Journal of Production Research*, **43**, 1275-1293. <https://doi.org/10.1080/00207540412331327727>
- [3] Chinnam, R.B. and Baruah, P. (2004) A Neuro-Fuzzy Approach for Estimating Mean Residual Life in Condition-Based Maintenance Systems. *International Journal of Materials and Product Technology*, **20**, 166-179. <https://doi.org/10.1504/IJMPT.2004.003920>
- [4] 刘波, 宁芊, 刘才学, 等. 基于连续型 HMM 和 PSO-SVM 的滚动轴承剩余寿命预测[J]. 计算机应用, 2019, 39(S1): 31-35.
- [5] 曾庆虎, 邱静, 刘冠军. 基于 HSMM 的机械故障演化规律分析建模与预测[J]. 机械强度, 2010, 32(5): 695-701.
- [6] 王宁, 孙树栋, 李淑敏, 等. 基于 DD-HSMM 的设备运行状态识别与故障预测方法[J]. 计算机集成制造系统, 2012, 18(8): 1861-1868.
- [7] 刘勤明, 李亚琴, 吕文元, 等. 基于自适应隐式半马尔可夫模型的设备健康诊断与寿命预测方法[J]. 计算机集成制造系统, 2016, 22(9): 2187-2194.
- [8] 熊婕. 浅析油浸式变压器的老化和寿命评估[J]. 江西化工, 2017(5): 213-214.
- [9] 张云朋, 胡海燕, 刘全楨, 等. 变压器典型故障分析及在线监测技术应用[J]. 安全、健康和环境, 2018, 18(9): 11-14.
- [10] 王陈恩, 蔡涌烽, 谢振华, 殷豪. 基于通道注意力与残差卷积神经网络的变压器故障诊断[J]. 黑龙江电力, 2022, 44(1): 68-74.