

数据驱动的建筑电气系统诊断方法研究

郑建波

(中铁十六局集团第一工程有限公司)

【摘要】随着机器学习的发展,建筑电气系统诊断逐渐走向自动化,然而传统的方法需要烦琐的特征工程,极大限制了诊断速度和准确率。针对上述问题,拟提出一种基于深度学习的建筑电气系统诊断方法,利用深度学习独有的特征自学习能力,实现复杂建筑电气系统诊断目的,并对此展开了深入研究和分析。

【关键词】特征工程;建筑电气系统;深度学习;故障诊断

中图分类号: TU855

DOI: 10.13655/j.cnki.ibci.2023.08.040

Research on Data-driven Diagnostic Methods for Building Electrical Systems

ZHENG Jian-bo

(China Railway 16th Bureau Group 1st Corporation Limited)

【Abstract】With the development of machine learning, building electrical system diagnosis is gradually becoming automated. However, traditional methods require tedious feature engineering, which greatly limits the speed and accuracy of diagnosis. In response to the above problems, a deep learning-based diagnostic method for building electrical systems is proposed, which uses the unique feature self-learning ability of deep learning to achieve the purpose of diagnosing complex building electrical systems, and in-depth research and analysis have been carried out.

【Keywords】feature engineering; building electrical system; deep learning; fault diagnosis

1 引言

近年来,智能建筑在我国取得了显著进展,建筑电气系统是整个建筑工程体系中的重要组成部分,主要包括建筑供配电、日常办公和机械设备用电等需求。然而,随着用电设施类型和数量的增加,电气系统规模日益增大,各结构之间联系更加复杂,导致故障率显著提高。因此,快速准确的故障诊断技术越发重要^[1]。

目前,我国电气的系统故障检测仍处于初步阶段,大部分依赖人工诊断故障原因,这不仅费时费力,还增加了检修成本^[2]。机器学习技术的飞速发展,为电气故障检测带来了重大突破,通过用电气系统的数据建模可以快速诊断电气故障,但是此类方法严重依赖特征工程,即需要研究人员根据经验提取故障特征以进行模型训练,因此无法应对当下复杂的电气系统^[3]。本课题拟建立一种基于深度学习的建筑电气系统检测方法,无须人工给出特征的前提,真正实现建筑电气故障检测的自动化、智能化。

2 信号处理在建筑电气故障诊断中的应用

在建筑电气故障诊断方法中,信号处理技术将

被测信号中重要部分凸显出来,然后根据信号与故障间存在的关联判断和分析,相较于传统的人工检测方法,其诊断效率高,不易受人为因素影响。本课题的研究方法是基于信号处理的,因此着重介绍信号处理方法,大致包括以下几种方法,时域处理方法、频域处理方法、统计处理方法、模式识别方法和深度学习等方法。

1) 时域处理方法

信号时域处理方法是针对原始被测信号的时间特性进行分析,进而识别故障的特征。常见的时域处理方法有卷积、相关性分析和信号积分等。信号积分是对原始信号的动态特性进行研究的,对于信号的动态特征变化有较强的敏感性。相关性分析方法是通过对信号的时间相关性进行故障诊断。卷积方法能够提取信号的时域特征。

2) 频域处理方法

频域处理方法是针对信号的频谱进行分析。常见的方法有小波变换和傅立叶变换。傅立叶变换的数学原理是将一个函数表示成多个函数的线性组合,在信号处理领域能够对信号的频谱进行分析,衍生出连续傅立叶变换和离散傅立叶变换等多种变体。20 世纪初期,阿尔弗雷德·哈尔首次提出小波概念,该变换常用于信号处理,由离散小波变换和连续小

波变换组成。实现原理为通过平移和缩放去匹配故障特征,从而达到去噪效果。小波变换继承和发展了短时傅立叶变换局部化的思想,同时又克服了窗口大小不随频率变化等缺点,能够提供一个随频率改变的“时间-频率”窗口。相较于傅立叶变换,小波变换能同时获得时域和频域特征,是信号处理的有效工具。

3) 统计学方法

统计学方法是对被测信号的均值、方差、能量、相关系数等特征进行分析,建立此类特征与故障分类的关系。其中均值是确定信号中心位置的值,方差能够反映信号的波动程度,衡量数据的分布情况。相关系数分析是对信号间的相关性进行分析的值,能够确定信号数据间的相关程度。

4) 模式识别方法

模式识别方法是近年来逐渐兴起的方法,其主要是利用电气系统的历史数据,通过建立一系列规则对电气数据进行分析,以找出常见的故障模式,并用这些模式来诊断当前故障,模式识别的方法主要分为两类:监督学习和非监督学习。监督学习是指训练数据包含标签,且标签在模型训练过程中指导模型训练。非监督学习是指训练数据不包含标签,模型在训练过程中不采用标签信息。显然二者的区别很明显,因此二者的算法也具有本质的区别。下面将分别介绍监督学习和非监督学习的代表算法。

监督学习的常见算法有:多层感知器、决策树、随机森林、支持向量机、朴素贝叶斯等。决策树是一种基于树形结构的算法,可同时用于分类和回归问题。具体分为决策树建立、分裂特征选择、生成和剪枝操作。决策树每个分裂特征是由信息熵和信息增益等原则决定的,且在后续递归的生成子树,直至满足停止条件为止,在这个过程中,为了使决策树具有较高的鲁棒性,剪枝操作旨在移除过深或简单的节点。随机森林则是建立在多棵决策树的基础上的,通过随机采样和随机特征选择两个策略,并且在输出最终结果时采用投票机制,即构造的多棵决策树输出结果的众数作为随机森林的输出结果。随机森林的优势在于不易发生过拟合情况,并且泛化能力强。此外,随机森林算法能够给出数据特征的重要性顺序,是一种应用广泛的机器学习算法。

非监督学习的常见算法有:K-means、DBSCAN、混合高斯模型等。K-means 算法的基本原理是初始

化K个质心,K值也是类别数,计算每个数据点到这K个质心的聚类,距离最近的点归属为该类别,至此每个数据点划分为不同类别,通过计算所有数据点的均值更新每个类别数据的质心位置,重复上述步骤直到各个数据点所属类别不在变化停止更新。DBSCAN 是基于密度的聚类方法,步骤为首先初始化圆半径和最小数据量,这决定了数据点如何被视为密集区域,然后从每个数据点开始,循环迭代找到该数据点周围的样本并视为一个簇,直到周围的数据点全部遍历结束,从其他未遍历的点开始重复上述任务并聚集为新簇。该方法对离群点比较敏感。

5) 深度学习方法

深度学习方法是通过对大量的数据中提取有价值的内容,需要注意的是该方法需要数据量特别大。不像机器学习算法,少量数据也可以做模式识别。20 世纪深度学习方法已经出现,受限于当时计算机算力以及数据量的缺乏,一度陷入低谷。随着计算机的飞速发展以及存储数据的增加,之前的限制因素已然消失,因此深度学习方法迎来了又一次发展高峰,针对不同形式的数据,深度学习均有模型与之对应。本文的信号数据,经过简单的处理后即可输入进深度学习模型进行训练。具体的深度学习方法将在第三章内容中详细介绍。

3 深度学习方法介绍

随着科学技术的发展,人工智能逐渐进入我们的视野。其中机器学习是人工智能的核心,通过设计算法让计算机实现自动学习的功能,机器学习算法的缺点是需要人工手动提取分类特征,为了克服这一缺点,深度学习逐渐成为研究热点,受启发于人类大脑的工作方式,模拟人脑进行分析,无须复杂的特征工程,网络模型在训练的过程实现特征自提取功能,已广泛应用于计算机视觉、语音识别、机器翻译等领域,且衍生出许多经典模型,例如深度神经网络、卷积神经网络、长短时记忆网络和生成对抗网络等。深度神经网络是深度学习的雏形,也是神经网络的简单形式。卷积神经网络是根据人眼观察事物的主要特征发展而来的,通过卷积计算,能够捕获图像数据中的重要成分,权值共享机制能够大大降低神经网络训练的参数量。长短时记忆网络用于处理数据具有前后时序关系的情况,例如机器翻译、文本对话,其特有的门控机制可以保留有用信息,同时丢

弃无用内容,从而将有用信息传输到后面再处理。

本文拟采用深度神经网络模型,故重点介绍该网络结构及其训练方法。该网络模型由输入层、中间层和输出层组成。数据通过输入层输入进网络,通过多个中间层节点,最后由输出层输出结果。为了解决非线性问题,在网络层中引入激活函数,常见的激活函数有 Sigmoid 函数、Softmax 函数和 Relu 函数等。Sigmoid 函数常用于二分类问题,例如建筑电气系统是否故障;Softmax 函数常用于多分类问题,例如建筑电气故障的类型,此时模型输出结果为多类别的概率,所有类别的概率之和为 1,其中概率值最高的类为预测类别。

4 基于深度学习的建筑电气系统故障诊断方法

4.1 输入数据预处理

本课题拟建立建筑电气系统故障检测模型,因此收集到的数据经过简单的预处理即可输入进模型。预处理共分为两步:数据归一化和数据切片。由于被测信号波动范围不同,采用归一化将其统一在 $[0,1]$ 的范围内,即信号的最小值变为 0,最大值变为 1。数据切片是将整个时间序列信号分解为更小的部分,以便神经网络能够有效地提取其固有特征。在输入数据中,每个切片信号的个数与网络模型输入层匹配。至此,数据预处理工作完成。相比于其他基于特征工程的诊断方法,大大缩减了数据预处理的时间。

4.2 模型构建

本项目的主要目标是构建一个神经网络模型,用于检测建筑电气系统是否存在故障。模型的输入是来自系统的信号数据。该神经网络由四层构成,节点数分别为 256、512、256 和 128。目标是解决二分类问题,即判断电气系统是否存在故障,因此在模型的输出层只需 1 个节点。该节点结合使用了 sigmoid 激活函数来输出表示故障概率的值,范围为 0 到 1。为了增强模型的非线性表达能力,并保持一定的鲁棒性,我们在各层之间使用了 LeakyRelu 作为激活函数。考虑到避免过拟合的问题,每一个激活函数后添加 Dropout 层。Dropout 层可以随机地在训练过程中忽略一部分神经元,帮助防止模型过度依赖训练数据的特定特征,增强了模型的泛化能力。

4.3 评价指标及案例分析

4.3.1 评价指标

模型训练结束后,需要对模型的效果进行评估,业界常用的评价指标为准确率。即模型预测正确的个数占数据总数的比例。通常会预设准确率阈值,如果准确率较低,则通过调整网络模型层数和节点数以达到较高准确率。

4.3.1 案例分析

为了检验本次设计的基于深度学习的建筑电气系统诊断方法的效果,选取 M 公司的冷水机系统近 5 年运行数据作为研究对象。根据前 4 年的数据作为训练集,最后 1 年的数据作为测试集,其中共有 7 种故障类别和 1 种正常模式,但本试验只区分正常模式和故障模式。

实验使用前 4 年的数据训练深度学习模型,并使用最后一年的数据进行测试。结果显示,该模型能够以 98% 的准确率区分冷水机系统的正常运行和故障状态。表明基于深度学习的故障诊断系统具有出色的故障检测能力,可以大幅度提高冷水系统的故障诊断速度。这进一步强调了深度学习在设备故障预警和诊断中的应用潜力。

5 结语

本文主要介绍了几种建筑电气系统故障诊断方法,并讨论了拟构建的诊断模型和评估指标。利用神经网络特征自学习能力建立分类器,经过简单预处理后的数据输入进模型进行预测,省去了复杂烦琐的特征工程过程,并预期达到较高的诊断准确率,实现建筑电气故障诊断智能化和自动化。但课题并未考虑数据类别的平衡性,正常数据比故障数据多,这将影响模型的诊断准确性,下一步将考虑引入生成对抗网络模型,对缺少的故障数据进行数据增强,以达到两类数据平衡的目的。

参考文献

- [1] 钱文博,熊建斌,岑健,王颀,余得正,吴润杰.建筑电气系统故障诊断综述[J].计算机工程与应用,2021,57(16):27-39.
- [2] 崔小飞,罗磊,白青松.建筑电气系统故障检测研究[J].智能建筑与智慧城市,2022(11):153-155.
- [3] 程思铭.故障诊断技术及监测方法在建筑电气系统上的应用[D].江苏:中国矿业大学,2021.