

DOI:10.12158/j.2096-3203.2020.03.023

基于知识嵌入和 DNN 的工商业用户异常用电检测

李江腾, 王非

(华中科技大学电信学院, 湖北 武汉 430074)

摘要:目前的异常用电检测研究主要以居民用户为对象,并不适用于工商业用户。针对此问题,文中提出一种融合了电能计量原理的基于深度学习的异常用电检测方法。首先,分析了各类异常用电的数据现象,结果说明单纯采用智能电表数据不足以准确检测异常用电。文中遵循电能计量原理,将描述电气参量内部逻辑关系的指标作为知识嵌入智能电表数据,构建深度学习样本模型。然后提出一个改进的深度混合残差神经网络,从海量智能电表数据中学习用于识别异常用电的高级特征。实验结果表明,相比多个基准算法,文中方法在所有评估指标上均取得了明显的提升。

关键词:智能电网;人工智能;深度学习;神经网络;异常用电

中图分类号:TM933

文献标志码:A

文章编号:2096-3203(2020)03-0158-08

0 引言

长期以来,由于窃电、电表故障和安装错误等引起的损失统称为非技术损失(non-technical losses, NTL)。NTL不仅给全球电网运营商每年带来巨额经济损失^[1-2],还会影响电网的稳定性和可靠性^[3],是电网最关心的问题之一。而世界各国工商业用户在全社会用电量中占有非常大的比例,在我国该比例已接近85%^[4],因此在工商业用户中检测异常用电显得尤为重要^[3]。

为解决异常用电检测问题,国内外学者进行了广泛研究。早期研究主要围绕基于硬件的技术方案,但所需传感器数量较大,成本和效率难以满足现代电网运营的需求^[5]。随着智能电网的发展和用电信息采集系统的建设,电网采集并保存了海量智能电表数据,因此后续的研究主要以面向数据分析的方法为主^[5]。目前,这些方法大致分为三类:(1)基于统计学的方法,主要通过拟合数据的概率分布模型参数检测异常值^[6-7],这类方法误报率较高,不适合电网的大规模实际应用。(2)基于专家知识的方法,主要将专家的异常检测知识与经验转化为一组规则和阈值用于识别异常^[8]。此类方法受限于规则和阈值选取的主观性,通用性和适应性不理想。(3)基于机器学习的方法,利用各类机器学习算法,对海量用电数据进行分析,达到检测异常的目的。这类方法根据采用的机器学习方法不同,又可以分为基于监督学习的方法,主要包括决策树^[9]、支持向量机(support vector machine,

SVM)^[10-11]、K近邻^[12]、人工神经网络^[5]等;以及基于无监督学习的方法,主要包含聚类^[13]、离群点检测^[14]等。特征选择是这类方法的重要环节,直接影响到异常用电检测的最终效果。由于用电数据存在漂移、波动等特性,人工设计特征模型越来越困难。

近年来,深度学习因其强大的特征学习能力,在计算机视觉、语音识别等领域取得了令人瞩目的进展^[15-16]。文献[2,5,17]分别提出了基于深度学习的窃电检测方法,从海量用电数据中自动学习高级特征。实验结果表明,这些方法均提高了窃电检测的准确性,然而这些研究主要针对居民用户,并不适用于工商业用户。为了克服该问题,文中提出一种基于电能计量知识和深度混合残差网络的异常用电检测方法,适用于合同容量高于50 kV·A的10 kV工商业专变用户。首先,基于原始智能电表数据,设计了一种融合电能计量知识的样本模型;其次,提出了一种改进的深度混合残差网络,从海量数据中提取更有效的高级特征。最后,基于真实案例数据通过实验验证了文中方法用于异常用电检测时的准确性和稳定性。

1 电能计量原理与异常用电分析

1.1 电能计量原理

目前,几乎所有工商业用户都配备了三相智能电表。根据接线方式的不同,可以分为三相四线和三相三线。文中以三相四线为例进行分析。首先,对于单相智能电表,电能计算为:

$$E = \sum_{n=0}^{N-1} P_n \Delta t \quad (1)$$

式中: Δt 为计算周期; E 为一定时间段 $N\Delta t$ 内的有

收稿日期:2019-12-03;修回日期:2020-01-12

基金项目:国家自然科学基金资助项目(61501195)

功电能; P_n 为时刻 $n\Delta t$ 处的平均有功功率。通过电压和电流, 可以进一步计算 P_n :

$$P_n = U_n I_n \cos\varphi_n \quad (2)$$

式中: U_n, I_n 分别为时刻 $n\Delta t$ 的平均电压和电流; φ_n 为时刻 $n\Delta t$ 电压和电流之间的相位差; $\cos\varphi_n$ 为功率因数。因此, 有功电能方程可写为:

$$E = \sum_{n=0}^{N-1} U_n I_n \cos\varphi_n \Delta t \quad (3)$$

基于式(3), 三相四线智能电表的总有功电能为:

$$E_{\text{total}} = \sum_{n=0}^{N-1} P_{\text{total}} \Delta t = \sum_{n=0}^{N-1} [P_{n,A} + P_{n,B} + P_{n,C}] \Delta t \quad (4)$$

式中: $P_{n,A}, P_{n,B}, P_{n,C}$ 分别为时刻 $n\Delta t$ 处的三相平均有功功率; P_{total} 为时刻 $n\Delta t$ 处的总有功功率。

由此可见, 电能的计量完全依赖电压、电流的测量, 各电气参量的内部逻辑关系由智能电表物理模型决定。此外, 工商业用户相比居民用户, 其用电模式更多、更复杂, 即使同属一个城市或行业, 彼此间的用电模式也完全不同^[3]。因此工商业用户的正常用电与异常用电在用电模式上较难区分。单纯对电气参量进行统计学习, 并不能很好地学习到彼此之间的物理模型关系, 有必要将相关先验知识或内在规律引入深度学习。

1.2 异常用电特征分析

在正常情况下, 智能电表数据必须遵循上述电能计量原理。相反, 异常用电的智能电表数据必然会违背电能计量规律以减少电能的计量。因此, 异常用电可以大致分为:

(1) 分流。对于三相智能电表, 分流是指电压互感器 (potential transformer, PT) 或电流互感器 (current transformer, CT) 的输出被分流, 将使测量到的电压或电流低于实际值。图 1 对比了正常用户与异常用户的三相电压和三相电流曲线。一般情况下, 三相用户的电压和电流几乎是平衡的^[18], 如图 1(a) 和 (b) 所示。分流将破坏三相电压或电流的平衡, 如图 1(c) 和 (d) 所示。且三相电流的不平衡程度还与用户的负载水平有关, 如图 1(d) 所示。分流是最复杂的异常用电行为, 因为正常用户也可能出现类似现象。在图 1(b) 中, 当负载较低时, 正常用户的电流也会失衡。

(2) 相移, 即人为改变电压和电流之间的相位差。通过降低功率因数, 达到降低有功功率测量值的目的。通常这种情况下, 总功率因数会低于国家标准。

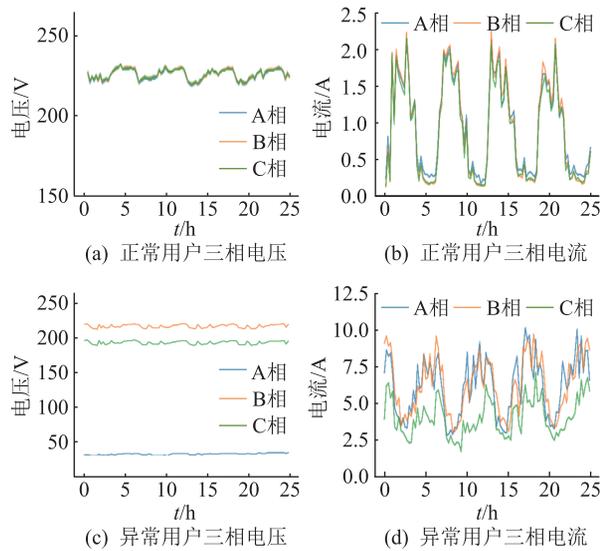


图 1 正常用电与异常用电的电压和电流曲线

Fig.1 Curves of voltage and current of normal and abnormal cases

(3) 相序混乱, 电流与错误的电压耦合。图 2 为这种情况的典型示例。A 相的 CT 和 PT 输出错误地与 B 相的 PT 和 CT 输出结合在一起。由图 2 可以发现, 电压、电流和功率因数都很正常。但在图 2(d) 中, 电表测得的归一化总有功功率与估算的归一化总有功功率之间存在巨大差距, 主要是由于相序混乱带来的电压、电流相位错误, 最终导致测量得到的总有功功率与估算值相比偏低。

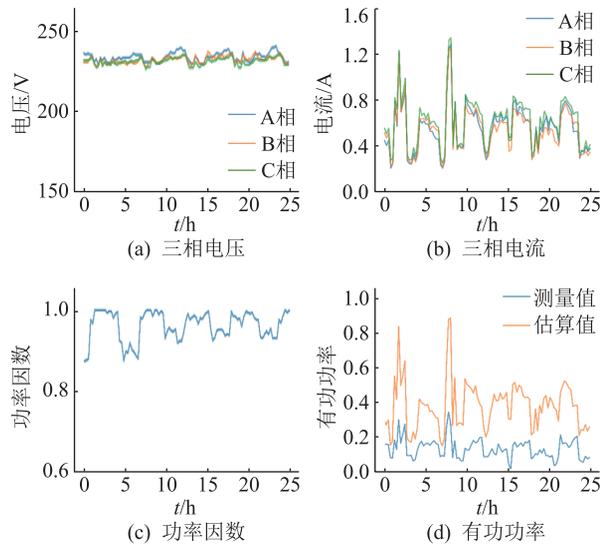


图 2 某相序混乱用户的数据曲线

Fig.2 Curves of a phase disordered case

(4) 相位反转, PT 或 CT 的输出正负极反转。电压和电流之间的相位差从 φ 变为 $\pi - \varphi$ 。智能电表将测量到负有功功率或较小的总有功功率。相位反转时归一化有功功率和功率因数曲线如图 3 所示, 这种异常情况存在有功功率和功率因数出现负

相关的现象,其他数据全部正常。

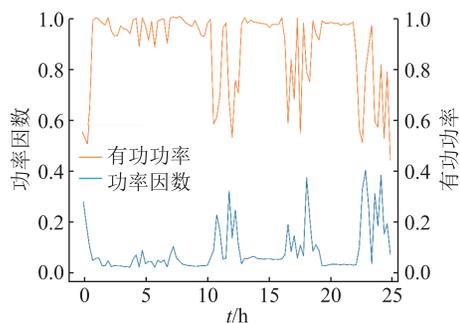


图3 相位反转

Fig.3 The of phase reverse

分析上述真实案例可知,工商业用户的用电模式会随时随机变化,且异常用电可能从一开始就存在,例如反相和相序混乱。因此,仅凭智能电表数据和深度神经网络(deep neural network, DNN)很难在工商业用户中准确查找异常用电,还需要在样本数据中嵌入电能计量领域的相关知识。

2 基于电能计量知识的样本模型

样本数据的组织结构对于 DNN 的重要性类似于特征模型对于分类算法的重要性。在文献[5]中,用电量数据被组织为一维向量和二维矩阵输入 DNN。不同于用电量数据,智能电表数据是一个多时间序列数据,包括三相电压、三相电流、总有功功率和总功率因数等电气参量。文中将其组织为具有多个通道的一维向量。通过比较不同的时间跨度后,发现太短的时间跨度蕴含的信息量太低,太长的时间跨度噪声大,且因为数据缺失的原因,可供训练和测试的样本数量将大幅下降。综合而言,以周为单位的智能电表数据可以更好地描述用电行为的周期性特征。因此,文中选择以周作为样本的时间跨度,并设计一个移动窗口,以天为滑动步数构造多个样本,如图4所示。

根据文中对异常用电数据现象的分析,并结合电能计量原理所揭示的各个电气参量之间的物理模型关系,文中补充以下参数作为附加通道:

$$\text{电压不平衡度} = \frac{U_{\max} - U_{\min}}{U_{\max}} \quad (5)$$

$$\text{电流不平衡度} = \frac{I_{\max} - I_{\min}}{I_{\max}} \quad (6)$$

$$\text{相位系数} = \frac{(U_A I_A + U_B I_B + U_C I_C) F_{\text{总}} - P_{\text{总}}}{P_{\text{总}}} \quad (7)$$

$$\text{负载率} = \frac{U_A I_A + U_B I_B + U_C I_C}{\text{合同容量}} \quad (8)$$

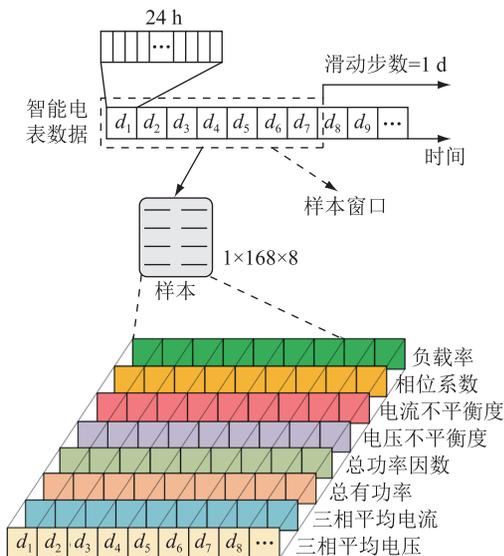


图4 嵌入电能计量知识的样本模型

Fig.4 Electricity measurement knowledge embedded sample model

式中: U_{\max} 为三相电压中的最大电压; U_{\min} 为三相电压中的最小电压; I_{\max} 为三相电压中的最大电流; I_{\min} 为三相电压中的最小电流; U_A, U_B, U_C 分别为三相电压; I_A, I_B, I_C 分别为三相电流。

此外,三相三线用户的 B 相电压和电流因为接线方式的原因,在智能电表数据中恒为 0。为统一分析三相四线和三相三线用户数据,文中采用三相平均电压和电流取代原始电压和电流数据,避免三相三线用户的 B 相电压和电流为 0 带来的影响。平均电压,平均电流以及 $U_{\max}, U_{\min}, I_{\max}, I_{\min}$ 均基于有效电压和电流。

3 深度混合残差网络

深度学习的优势在于更深的神经网络,但是实验发现,DNN 中存在梯度消失问题^[19],造成 DNN 的训练非常困难。针对这个问题,何凯明等人提出了残差网络(ResNet)^[19]。通过引入神经网络单元输出的特征 $F(x)$ 与输入 x 之间的残差 $H(x)$,使网络学习的对象不再是 $F(x)$,而是残差 $H(x)$,从而使梯度可以在更深的神经网络中传递,有效解决了网络深度与梯度消失之间的矛盾。

目前,残差网络的设计主要针对图像数据,不论是采用的卷积网络类型,还是对数据的感受范围,均不太适合智能电表数据。考虑到电气参量会在不同的时间尺度上产生关联,并且同一时刻不同电气参量之间也具有各种不同的相关性。因此,文中借鉴 Inception 网络^[20]的思想,设计了图5所示的混合残差网络结构。该网络结构不仅允许构建更深的神经网络,还可以更广泛地从不同角度、不同

5 实验与分析

5.1 数据集介绍

所有训练和测试数据都来自某地电力公司提供的真实案例数据,其中包含 346 户已通过现场检查的 10 kV 专变用户,涵盖三相四线和三相三线 2 种接线方式。每个用户的数据的时间跨度从 10 至 12 个月不等,文中以现场检查报告的时间为准,截取之前约 3 个月的数据。因此,同一用户的所有样本根据现场检查结果标注为相同的标签。表 1 提供了有关数据集的详细信息。

表 1 数据集概况
Table 1 Overview of dataset

异常用电类型	用户数	样本数
电压分流	11	998
电流分流	25	2 271
相移	13	1 182
相序混乱	20	1 816
反相	45	4 093
正常	232	21 089

5.2 异常用电检测评估指标

受试者工作特征曲线(receiver operating characteristic, ROC)是一种在不平衡数据集上验证分类器性能的流行方法,由于在实际场景中正常用户与异常用户极不平衡,因此得到广泛应用^[3,5]。通常,将 ROC 曲线下的面积(area under ROC, AUC)用作主要指标。此外,部分文献选择 F1 分数来评估算法的性能^[21]。为了更全面、准确地评估,文中选择所有上述指标评估算法的性能。具体指标定义如下:

$$\text{查准率} = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (12)$$

$$\text{查全率} = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (13)$$

$$F1 = \frac{2N_{TP}}{\text{测试样本总数} + N_{TP} - N_{TN}} \quad (14)$$

式中: N_{TP} (真阳性)为正确检测到的非技术性损失样本数; N_{FP} (假阳性)为被分类为正常的非技术性损失样本数; N_{TN} (真阴性)为正确分类的正常样本数; N_{FN} (假阴性)为正常样本被归类为非技术性损失样本的数量。

5.3 实验设置

5.3.1 测试基准

为了证明文中方法的有效性,选择了几种基准算法进行比较。相关算法参数的选择参考文献[3, 10—11],先确定范围,然后采用网格搜索的方法在

文中数据集上获得最优参数。具体如下:

(1) SVM。对核函数和惩罚参数分别在[线性函数,径向基函数]和[0.001,0.01,0.1,1,10,100]的范围内搜索,最终确定径向基核函数,惩罚参数为0.01。 γ 由下式计算而得:

$$\gamma = \frac{1}{\text{特征模型维度}} \quad (15)$$

此外,正常和异常类别根据样本的比例分别赋予 1 和 2 的权重。

(2) 极端梯度提升树(XGBoost)。对树的数量和树的最大深度分别在[100,200,300,400,500]和[3,4,5]的范围搜索,最终确定树的数量为 200,每棵树的最大深度为 3,学习率为 0.01。

(3) ResNet-18。网络模型基于文献[19],使用一维卷积替换所有二维卷积,其他参数保持不变。

参考文献[22],SVM 和 XGBoost 特征模型直接采用样本降维后的结果。通过对比主成分分析方法(principal component analysis,PCA)和截断奇异值分解(truncated SVD, TSVD) 2 种降维方法,可知 TSVD 的效果更好,最佳的特征维度设定为 50。

5.3.2 模型实现与参数设置

Keras 是一个支持高级编程接口的深度学习框架,简单且易于实现。因此,文中网络模型以及测试基准中的 ResNet-18 均基于 Keras 实现。

文中所有实验的混合残差网络模型均包含 6 个混合残差网络单元,滤波器的参数如图 5(a)所示。

5.4 结果与分析

5.4.1 样本模型性能分析

表 2 为原始数据样本和文中知识嵌入模型样本的详细性能。通过在 2 个样本模型中比较各个测试基准算法和混合残差网络的性能指标,可以发现所有算法的性能指标均得到了明显改善。即使对于 SVM 和 XGBoost 算法,使用降维数据作为特征,嵌入样本的电能计量知识依然可以提升算法识别异常用电的性能。

表 2 2 种样本数据模型的性能比较

Table 2 Performance comparison between two sample models

方法	模型	F1 分数	AUC 分数
SVM	原始数据	0.713	0.892
	文中模型	0.746	0.937
XGBoost	原始数据	0.649	0.912
	文中模型	0.765	0.934
ResNet-18	原始数据	0.720	0.921
	文中模型	0.786	0.937
混合残差网络	原始数据	0.743	0.929
	文中模型	0.846	0.970

图 7 为不同时间跨度样本模型下,深度混合残差网络对异常用电的检测效果。首先,从图 7 可以看出 7 d 的时间样本模型,不论 AUC 分数还是 F1 分数都是最好的,尤其是 F1 分数明显优于其他时间跨度的样本模型。不仅因为数据维度足够大,包含足够的信息量;而且较好地体现了工商业用户的用电行为的周期性特征。1 d 和 3 d 的样本模型,由于数据维度太低,包含的信息量少,深度混合残差网络很难学习异常用电特征,导致异常用电检测效果相对较差。5 d 的样本模型尽管在 AUC 分数上与 7 d 的样本模型非常接近,但 F1 分数有明显差距。主要因为 5 d 的时间跨度仍不能很好地体现用电行为的周期性特征,造成样本模型对判决阈值更加敏感,这将影响深度混合残差网络的泛化能力。

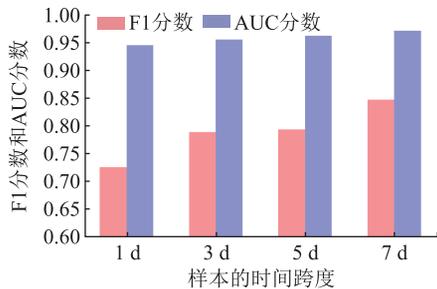


图 7 不同时间跨度的样本结构的性能
Fig.7 Performance of sample structures with different time spans

5.4.2 深度混合残差网络性能分析

从以下 2 个方面评估文中模型的性能:

(1) 泛化性能。表 3 为深度混合残差网络的 5 折交叉验证的结果。由表可知,F1 分数和 AUC 分数分别介于 $[0.822, 0.853]$ 和 $[0.967, 0.974]$,没有出现非常明显的性能差异,总体上说明文中方法在不同的训练集和测试集上具有较好的泛化性能。对比 F1 分数和 AUC 分数的分布情况可知,F1 分数的波动范围大于 AUC 分数,即使第 5 折的 F1 分数只有 0.822,AUC 分数依然达到了 0.967。主要是因为 F1 分数是基于 0.5 的判决阈值,而过大的阈值使神经网络模型对数据的变化更加敏感。

表 3 文中模型的泛化性能

Table 3 Generalization performance of proposed model

算法	第 1 折	第 2 折	第 3 折	第 4 折	第 5 折
F1 分数	0.835	0.852	0.853	0.846	0.822
AUC 分数	0.968	0.974	0.972	0.970	0.967

(2) 收敛性分析。图 8 为深度混合残差网络的收敛性。从图中可知,在预热训练阶段,训练损失大于测试损失,主要是因为文中应用了随机增强方

法对训练进行增强处理,使得网络在最开始几轮训练中训练损失较测试损失大,但随着训练继续,训练损失很快下降。由此可见,通过预热训练阶段,模型很快找到了合适的初始状态。后续的收敛训练阶段,AUC 分数和 F1 分数稳步提升,训练损失和测试损失也稳步下降,总体上损失曲线和性能曲线都比较平滑,证明了模型性能的稳定性。

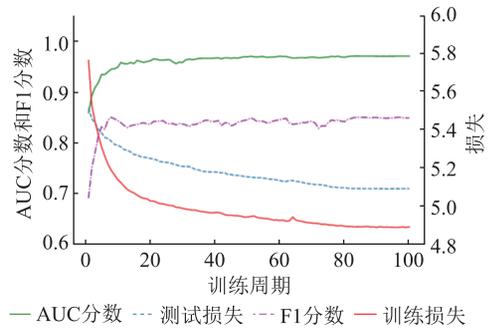
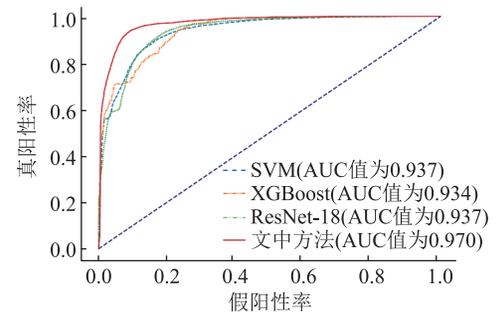


图 8 文中模型的收敛性能

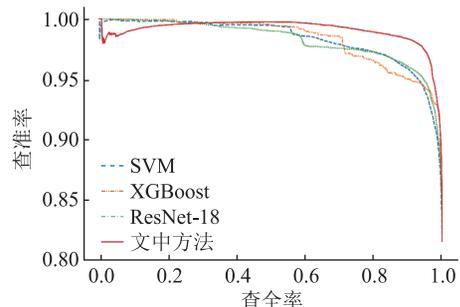
Fig.8 Convergence performance of proposed model

5.4.3 同基准算法的性能对比

图 9 为文中方法与测试基准算法的异常用电类 ROC 曲线和查全率-查准率 (precision-recall, PR) 曲线的比较。ROC 曲线和 PR 曲线均通过尝试不同的判决阈值,展示不同的综合分类性能。ROC 曲线和 PR 曲线均基于表 3 中第 4 折训练集和测试集获得。



(a) 异常用电的 ROC 曲线



(b) 异常用电的 PR 曲线

图 9 文中模型与测试基准的性能对比

Fig.9 Performance comparison between the proposed model and baselines

从图 9(a)可以看出,文中方法获得较大的优势,其他 3 个基准算法的性能处于同一水平。尤其

是 SVM 和 XGBoost 是基于降维而非人工优化的数据特征,而其自身并不具备特征学习能力,因此综合性能要差于 ResNet-18 和文中深度混合残差网络,主要因为后者具有强大的特征学习能力。同样,图 9(b)也得到了类似的结论。同时,SVM 和 XGBoost 的 PR 曲线也显示其更加依赖特征的选择。而文中方法的 PR 曲线则将其他 3 个基准算法的 PR 曲线完全“包住”,说明文中方法在异常用电检测性能更优。

进一步对比文中深度混合残差网络和原生的 ResNet-18,在使用完全相同的数据模型的前提下,文中方法优于原生的 ResNet-18,说明混合多个不同卷积尺度的 1 维卷积模块,更加适合在智能电表数据中学习更高级的特征,从而获得更好的性能。

6 结语

针对在工商业用户中准确检测异常用电的问题,文中提出了一种新颖的嵌入了电能计量知识的样本数据模型和深度混合残差神经网络。实验结果说明,在基于深度学习的异常用电检测中,通过融合电能计量知识构建样本数据模型,可以有效提升异常用电检测的准确率。其次,文中提出的深度混合残差网络可以更好地适配智能电表数据,学习其不同时间范围内的运行规律。5 折交叉验证结果显示文中方法所学习到的高级特征具有更好地泛化性能,异常用电检测性能更加稳定、可靠,可有效支撑在实际环境中识别异常用电。

参考文献:

- [1] ANTMANN P. Reducing technical and non-technical losses in the power sector, in: background paper for the WBG energy strategy [R]. Technical Report, Washington, DC, USA: The World Bank, 2009.
- [2] 赵文清,沈哲吉,李刚.基于深度学习的用户异常用电模式检测[J]. 电力自动化设备, 2018, 38(9): 34-38.
ZHAO Wenqing, SHEN Zheji, LI Gang. Anomaly detection for power consumption pattern based on deep learning[J]. Electric Power Automation Equipment, 2018, 38(9): 34-38.
- [3] BUZAU MM, TEJEDOR-AGUILERA J, CRUZ-ROMERO P, et al. Detection of non-technical losses using smart meter data and supervised learning [J]. IEEE Transactions on Smart Grid, 2019, 10(3): 2661-2670.
- [4] 2019 年 1—8 月全国电力工业统计数据一览表 [R/OL]. <http://www.cec.org.cn/guihuayutongji/tongjinxixiyuedushuju/2019-09-23/194417.html>, 2019.
Monthly statistics of China power industry (1—8, 2019) [R/OL]. http://www.cec.org.cn/guihuayu_tongji/tongjinxixiyuedushuju/2019-09-23/194417.html, 2019.
- [5] ZHENG Z, YANG Y, NIU X, et al. Wide and deep convolutional neural networks for electricity-theft detection to secure smart grids [J]. IEEE Transactions on Information Technology, 2018, 14(4): 1606-1615.
- [6] AGYEMANG M, BARKER K, ALHAJJ R. A comprehensive survey of numeric and symbolic outlier mining techniques [M]. Holland: IOS Press, 2006.
- [7] BORIAH S, CHANDOLA V, KUMAR V. Similarity measures for categorical data: a comparative evaluation [C]//Siam International Conference on Data Mining, SDM 2008. Atlanta, Georgia, USA, 2008: 243-254.
- [8] 林树青. 基于计量自动化系统的计量装置故障判断规则的研究与应用 [D]. 广州: 华南理工大学, 2017.
LIN Shuqing. Metering device fault diagnosis rules based on the measurement automation system the research and application [D]. Guangzhou: South China University of Technology, 2017.
- [9] COMA-PUIG B, CARMONA J, GAVALDA R, et al. Fraud detection in energy consumption: a supervised approach [C]// In IEEE International Conference on Data Science and Advanced Analytics (DSAA). 2016: 120-129.
- [10] NAGI J, YAP KS, TIONG SK, et al. Nontechnical loss detection for metered customers in power utility using support vector machines [J]. IEEE Transactions on Power Delivery, 2009, 25(2): 1162-1171.
- [11] JOKAR P, ARIANPOO N, LEUNG V C. Electricity theft detection in AMI using customers' consumption patterns [J]. IEEE Transactions on Smart Grid, 2015, 7(1): 216-226.
- [12] RAMOS CCO, DE SOUZA AN, FALCAO AX, et al. New insights on nontechnical losses characterization through evolutionary-based feature selection [J]. IEEE Transactions on Power Delivery, 2011, 27(1): 140-146.
- [13] KRISHNA V B, WEAVER G A, SANDERS W H. PCA-based method for detecting integrity attacks on advanced metering infrastructure [C]//In Proceedings 12th International Conference on Quantitative Evaluation of System (QEST). 2015: 70-85.
- [14] RUSSAKOVSKY O, DENG J, SU H, et al. Imagenet large scale visual recognition challenge [J]. International Journal of Computer Vision, 2015, 115(3): 211-252.
- [15] DENG L, HINTON G, KINGSBURY B. New types of deep neural network learning for speech recognition and related applications: an overview [C]//International Conference Acous. Speech Signal Proc., 2013: 8599-8603.
- [16] JÚNIOR LAP, RAMOS CCO, RODRIGUES D, et al. Unsupervised non-technical losses identification through optimum-path forest [J]. Electric Power Systems Research, 2016, 140: 413-423.
- [17] 胡天宇, 郭庆来, 孙宏斌. 基于堆叠去相关自编码器和支向量机的窃电检测 [J]. 电力系统自动化, 2019, 43(1): 119-127.
HU Tianyu, GUO Qinglai, SUN Hongbin. Nontechnical loss detection based on stacked uncorrelating autoencoder and support vector machine [J]. Automation of Electric Power Systems,

- 2019,43(1):119-127.
- [18] 王若丞. 配电网三相不平衡对线损增加率及电压偏移的影响[J]. 电力工程技术,2017,36(4):131-136.
WANG Ruochen. Influence of distribution network three-phase unbalance on line loss increase rate and voltage offset[J]. Electric Power Engineering Technology, 2017, 36(4): 131-136.
- [19] HE K M, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016:770-778.
- [20] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception-v4, inception-resnet and the impact of residual connections on learning[C]//Thirty-first AAAI Conference on Artificial Intelligence, 2017.
- [21] ZANETTI M, JAMHOUR E, PELLEZZI M, et al. A tunable fraud detection system for advanced metering infrastructure using short-lived patterns [J]. IEEE Transactions on Smart Grid, 2019, 10(1):830-840.
- [22] RAMOS C C O, ANDRÉ N, SOUZA C G, et al. A novel algorithm for feature selection using harmony search and its application for non-technical losses detection[J]. Computers & Electrical Engineering, 2011, 37(6):886-894.

作者简介:



李江腾

李江腾(1995),男,硕士在读,研究方向为深度学习(E-mail:lijiangteng@hust.edu.cn);

王非(1977),男,博士,副教授,研究方向为机器学习与电力数据分析。

Non-technical loss detection based on energy measurement knowledge and deep neural network among industrial and commercial customers

LI Jiangteng, WANG Fei

(School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China)

Abstract: The current research on non-technical loss (NTL) detection is mainly aim on residential customers, however, the related methods are not suit industrial and commercial customers. According to this problem, a deep learning based NTL detection method by embedding the principle of electricity measurement is proposed. Firstly, various NTL is analyzed and the phenomena show that only smart meter data is not enough for detecting NTL. Hence, smart meter data and some principles of electricity measurement are organized which describe the inherent relationship among electricity magnitudes as samples for deep learning. Secondly, an improved hybrid residual neural network is proposed to extract advanced features of NTL from massive smart meter data for detecting NTL. The experiment results show that the approach in this paper has achieved significant improvement on all metrics by comparing with the baselines.

Keywords: smart grid; artificial intelligence; deep learning; neural network; non-technical loss detection

(编辑 方晶)