

广域级电网在运智能电能表的烧损故障关联分析和预测方法

谈丛^{1,2†}, 黄红桥^{1,2}, 陈石东^{1,2}, 李恺^{1,2}, 解玉满^{1,2}, 刘谋海^{1,2}

[1. 国网湖南省电力有限公司供电服务中心(计量中心), 湖南长沙 410004;
2. 智能电气量测与应用技术湖南省重点实验室, 湖南长沙 410004]

摘要:针对智能电能表在运行过程中出现烧损的现象,在对各类因素进行关联分析后,提出了一种基于XGBoost算法的智能电能表烧损预测方法,以某省份2019—2020年的数据为例进行了测试验证.采用该方法结合电能表基本信息数据、运行数据和环境数据进行烧损识别,并与K最邻近法(K-NearestNeighbor, KNN)、朴素贝叶斯和支持向量机等传统机器学习算法进行对比.结果表明,基于极限梯度提升算法(eXtreme Gradient Boosting, XGBoost)的算法精度达到91%,召回率达到66%,综合指标 F_1 达到76.51%,远高于传统算法.算法模型在进行系统部署的过程中,运用长短期记忆算法(Long Short Term Memory, LSTM)对部分缺失值进行了填充,经试点验证,该模型可较为准确地预测低压台区电能表烧损现象.

关键词:智能电表;电表故障;烧表预测;XGBoost算法;不平衡数据

中图分类号:TM930.9

文献标志码:A

Burning Fault Correlation Analysis and Prediction of Smart Meters in Operation in Wide-area Power Grid

TAN Cong^{1,2†}, HUANG Hongqiao^{1,2}, CHEN Shidong^{1,2}, LI Kai^{1,2}, XIE Yuman^{1,2}, LIU Mouhai^{1,2}

[1. State Grid Hunan Electric Power Limited Company Power Supply Service Center(Metrology Center), Changsha 410004, China;
2. Hunan Province Key Laboratory of Intelligent Electrical Measurement and Application Technology, Changsha 410004, China]

Abstract: Aiming at the burning loss of smart electricity meters, after correlation analysis of various factors, this paper proposes a burning fault prediction for smart meter based on XGBoost algorithm. Taking the data of a province from 2019 to 2020 as an example, the proposed method is tested and verified. Using the basic information data, operation data and environmental data, the proposed method is compared with the traditional algorithms such as KNN, naive Bayes and support vector machine. The results show that the burning fault prediction of the XGBoost algorithm is better than the traditional algorithms. The precision of the XGBoost is 91%, the recall is 66%, and F_1 -score is

* 收稿日期:2021-10-25

基金项目:国家电网公司总部科技资助项目(5700-202155204A-0-0-00), Science and Technology Project of State Grid Cooperation(5700-202155204A-0-0-00); 国网湖南省电力有限公司科技项目(5216AG21000K), Science and Technology Project of State Grid Hunan Electric Power Cooperation(5216AG21000K)

作者简介:谈丛(1992—),男,湖南长沙人,国网湖南省电力有限公司供电服务中心(计量中心)工程师,硕士

† 通信联系人, E-mail: 6101889016@qq.com

76.51%. In the process of system deployment, LSTM algorithm is used to fill some missing values. The experimental results show that the model can accurately predict the burning fault of smart meter in low-voltage platform area.

Key words: smart meters; meter fault; burning fault prediction; XGBoost algorithm; imbalanced data

智能电能表目前已经在我国实现了大面积覆盖,智能电能表的发展为人们生活带来了便利,但受智能电能表安装施工工艺、运行年限、运行环境以及用电负荷等因素的影响,智能电能表烧表的现象时有发生,由此造成的停电以及表箱或线路火灾给用电的安全以及便利带来了诸多风险,给供电企业造成了较大损失.以某省份数据为例,2019—2020年,每年共有数万件的烧表事故,由此带来的经济效益损失与社会效益损失巨大,建立一种能提前预测烧表的方法或者模型,成为行业内的一种迫切的需求.

由于电能表烧损是一种偶发现象,且引发烧表的因素多种多样,基层供电服务员工的运维水平、接线材料、电能表运行环境等多重因素都可能导致烧表,这些因素通常不可直接对其进行量化表示,基于以上困难,业界未曾对烧表建立起一种普遍可行的预测机制.

近年来,与智能电能表配套发展的用电信息采集系统不断完善,积累了大量智能电能表的运行数据,通过对这些数据的挖掘和分析,构造烧表特征模型,成为破解烧表预测模型的一种可能^[1-2].

目前,国内外学者针对电能表运行中的故障提出了多种预测和分析方法,对于解决烧表问题具有一定的借鉴意义.文献[3]提出的智能电能表故障类型分类和预测模型,主要基于 Bayes Network,模型主要考虑了电能表产品特征数据,采用了基于评分搜索的方法构建 Bayes Network 结构,构建了电能表故障预测模型,但该模型并未使用智能电能表的运行数据,评价的维度不够全面.文献[4]提出了智能表故障风险预估算法,针对智能电表的首检数据进行筛查,将筛查结果划分为四个风险等级,但评价模型只应用了智能电表的首检数据,没有全面利用智能电表的产品特征数据和电气运行数据,模型的评价也不够全面.文献[5]提出了一种多分类集成树模型,首先利用聚类方法实现了数据的预处理和维度的压缩,从数据预处理的角度克服了不平衡数据集在机器学习过程中的弊端,本模型在数据预处理方面取得了较大的进展,但由于没有引入电气运行数

据,预测精度较低.文献[6]设计了基于梯度提升决策树(Gradient Boosting Decision Tree, GBDT)的故障大类、故障小类以及设备寿命周期的预测,但使用的数据维度较少,仅限于部分到货批次信息以及气候数据信息.文献[7]提出了一种多分类融合模型的智能电表故障预测算法,通过传统的欠采样和过采样相结合的方式解决了数据集中类不平衡问题,但欠采样和过采样都引入了较大随机因素,传统的欠采样会导致样本信息的缺失,传统的过采样可能会导致过拟合.文献[8]提出了一种智能电表故障多分类方法,主要针对现有算法无法实现标签值与概率值的基类模型融合问题,但该方法选取的特征均为静态变量,对电能表使用过程中的变化无法进行监测,无法解释故障的偶发性.文献[9]设计了一种基于深度信念网络对计量装置故障溯源的方法,使用卷积神经网络对历史电流、功率曲线序列进行了特征提取,实现了异常电能表的故障溯源,该方法既使用了电能表的产品批次信息,也利用了运行电气数据,虽然未对烧表问题进行研究,但具有较好的借鉴作用.文献[10]提出了一种智能电能表异常数据检测方法,能够对窃电用户和异常设备进行在线监测,能够提升供电公司的线损指标,该方法侧重于研究电能表数据异常,但对于电能表故障的研究不够深入.文献[11]设计了基于智能电能表运行故障数据的纵向分析模型,可以对不同厂家、不同批次的智能电能表的故障率随时间的变化进行分析,该方法也是侧重于利用电能表产品信息进行建模,没能对电能表运行数据进行分析.文献[12]提出了一种基于条件变分自编码器-卷积神经网络模型的不平衡多分类方法,将类标签作为约束条件,搭建由全连接层构成的 CAVE 网络生成少数类样本,学习各类分布特点和数据集全局特征,提高生成数据质量,对于解决数据不平衡问题,可以借鉴该方法.文献[13]提出了一种基于 LSTM 的电力负荷预测方法,该方法能够有效预测电力负荷数据,对于电能表的运行缺失数据,可以借鉴该方法,对缺失值进行填充.

GBDT 的基本算法思路是将弱分类器进行线性组合后得到强分类器^[14],在训练的过程中,GBDT 能

够根据上一轮的结果,迭代更新梯度的方向,使得模型能够以最大的速率向误差不断减小的方向进化^[15-16].

XGBoost 算法在 GBDT 的基础上进行了优化,能够根据计算机 CPU 的运行情况进行多线程的并行计算;同时还能在一定程度上提高泛化能力,将单决策树模型的复杂度作为正则化的优化目标;在公式推导中引入了二阶导数,准确率更高,不易过拟合^[17].

为了构建电能表烧损故障关联分析和预测模型,本文研究思路如下:

1)结合某省供电公司部分电能表的产品信息等静态数据和电能表的历史运行数据,进行烧损关联性分析,使用 XGBoost 模型对包含“烧表”和“不烧表”的二分类数据集进行训练,形成算法模型.

2)使用 XGBoost 算法模型进行交叉验证测试,并与 KNN、SVM 和 Naive Bayes 等传统机器学习算法进行对比,以观察 XGBoost 算法对实验准确度的提升是否有影响.

3)将模型进行系统部署,以某省供电公司所有在运的电能表为数据源,针对缺失的运行数据,使用 LSTM 算法对其进行预测并填充,构建完整的测试数据集,验证模型运行的实际效果.

1 电能表多维数据统计与分析

1.1 数据源

某省公司计量中心拆回分拣实验室从 2019 年开始对退出运行的电能表进行集中回收处置,对因烧坏而退出运行的电能表进行了逐块记录,形成了烧表原始数据集.数据集包括已烧坏电能表所在市州、生产厂家、安装时间、运行时间、所在台区近年历史烧表数量以及退出运行前的每日电量、电压和电流数据,所在台区每日线损数据共形成了 22 567 条烧表信息记录.同时,在用电信息采集系统选取 185 710 块正常运行的电能表的同样维度的数据,形成了未烧表数据集.

1.2 数据统计与分析

1)根据历史统计结果,不同市州的表计烧表率有较大的区别,全省 14 个市州的年烧表率范围大致为 0.2%~0.5%,将各市州属性按风险程度等级进行分类,赋予不同的标签,如表 1 所示.

2)根据历史统计结果,不同生产厂家的表计烧表率有较大的区别,运行在全省的电能表所属厂家共有 78 家,不同厂家 2020—2021 年平均每年的烧表

率范围大致为 0~1%,将各生产厂家属性按风险程度等级进行分类,赋予不同的标签,如表 2 所示.

表 1 市州属性风险程度等级分类

Tab.1 Risk degree classification of different cities

年烧表率/%	风险级别	标签
[0.2, 0.3)	1 级	1
[0.3, 0.4)	2 级	2
[0.4, 0.5)	3 级	3
[0.5, +∞)	4 级	4

表 2 生产厂家属性风险程度等级分类

Tab.2 Risk degree classification of different manufacturers

年烧表率/%	风险级别	标签
[0.0, 0.2)	1 级	1
[0.2, 0.4)	2 级	2
[0.4, 0.6)	3 级	3
[0.6, 0.8)	4 级	4
[0.8, 1.0)	5 级	5
[1.0, +∞)	6 级	6

3)根据统计结果,不同安装年份的烧表率有着明显差异,如图 1 所示,安装时间在 2013 年到 2018 年电能表烧损居多,占总烧表率数量的 74.16%.

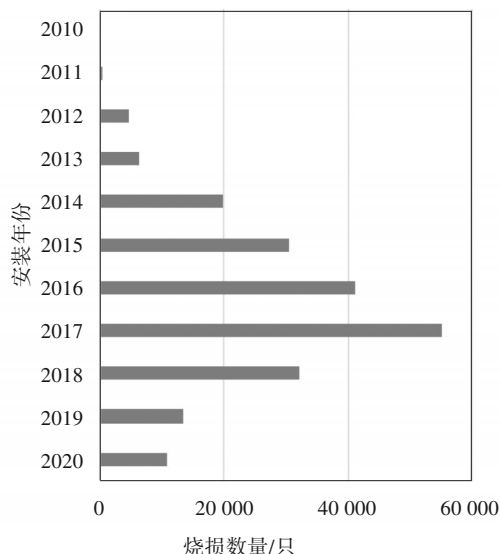


图 1 不同安装年份烧损电能表数量

Fig. 1 Number of burned smart meters in different installation years

4)根据统计结果,每年不同的月份烧表率也有着明显的差异,以某城市为例,2019—2021 年夏季月

平均气温每升高 1 °C,月烧表数量增加约 719 块;2019—2021 年冬季月相对湿度每下降 1%,烧表数量增加约 108 块。

5)根据统计结果,在烧表前一段时间内,每日电量、每日最大电压和最大电流会有不同程度的突变的现象。

6)根据统计结果,存在一部分烧表会导致电能表“表码”出现异常,会影响其所在台区的线损。

7)根据统计结果,电能表运行终码越大,其烧损的概率更高,终码在一定程度上能够表示电能表内元器件工作运行的实际时长。

8)根据统计结果,台区历史烧表数量,表示电能表所在台区上一自然年的总烧表数量,该数据能够在一定程度上反映出该台区的运维水平。

2 算法原理

XGBoost 算法是是基于 GBDT 改进而来的,它采用了 GBDT 的生成算法和梯度提升思想,并进行了相应的提升,使得算法效果进一步得到提高^[18-21]。

对于样本属性集合 $x = \{x_1, x_2, \dots, x_N\}$ 和样本标签组合 $y = \{y_1, y_2, \dots, y_N\}$,XGBoost 算法包括 K 棵树的加法模型为:

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), f_k \in F \quad (1)$$

式中: f_k 为基类模型的目标函数; F 为对应基类模型组成的函数空间。

目标函数计算公式为:

$$\text{Obj}(\theta) = \sum_{i=1}^N l(y_i, \hat{y}_i) + \sum_{j=1}^k \Omega(f_j), \quad f_j \in F \quad (2)$$

$$\text{where } \Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{t=1}^T \omega_t^2$$

式中: $l(y_i, \hat{y}_i)$ 表示真实值与预测值之间的误差; T 表示叶子节点总数; ω_t 表示第 t 个叶子节点的分数; γ 和 λ 表示 $L1$ 和 $L2$ 正则化系数。

$\hat{y}_i^{(t)}$ 可以写成:

$$\hat{y}_i^{(t)} = \sum_{j=1}^t f_j(x_i) = \hat{y}_i^{(t-1)} + f_t(x_i) \quad (3)$$

为了极小化目标函数,XGBoost 直接基于目标函数进行二阶泰勒展开。

$$\begin{aligned} \text{Obj}^{(t)} &= \sum_{i=1}^N l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) = \\ & \sum_{i=1}^N [l(y_i, \hat{y}_i^{(t-1)}) + g_i f_t(x_i) + \\ & \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t) \end{aligned} \quad (4)$$

$$\text{式中: } g_i = \frac{\partial l(y_i, \hat{y}_i^{(t-1)})}{\partial \hat{y}_i^{(t-1)}}, h_i = \frac{\partial^2 l(y_i, \hat{y}_i^{(t-1)})}{\partial^2 \hat{y}_i^{(t-1)}}.$$

继续简化,得目标函数为:

$$\text{Obj}^{(t)} = \sum_{i=1}^N (g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)) + \Omega(f_t) \quad (5)$$

找到最优分裂节点的步骤如下:

1) 每个叶子节点 j 上包含的样本集合 $I_j = \{i | q(x_i = j)\}$,为了进一步简化,目标函数可以改成如下形式:

$$\text{Obj}^{(t)} = \sum_{j=1}^r [G_j \omega_j + \frac{1}{2} (H_j + \lambda) \omega_j^2] + \gamma T \quad (6)$$

$$\text{式中: } G_j = \sum_{i \in I_j} g_i; H_j = \sum_{i \in I_j} h_i.$$

2)为使目标函数得到最优解,对 ω 进行求导,使导数为 0,得 $\omega_j = -\frac{G_j}{H_j + \lambda}$ 。

代入式(6)得:

$$\text{Obj}^{(t)} = -\frac{1}{2} \sum_{j=1}^r \left(\frac{G_j^2}{H_j + \lambda} \right) + \gamma T \quad (7)$$

3)最优的分裂节点能够使信息增益 G 最大。 G 的计算方式如下:

$$G = \frac{1}{2} \left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda} \right] - \gamma \quad (8)$$

式中: H_L 和 H_R 分别为目标函数在分裂后左、右节点的一阶导数之和; G_L 和 G_R 分别为目标函数在分裂后左、右节点目标函数的二阶导数之和。

3 实验流程

经过分析,电能表烧损可能与多个因素存在关联性,为了对电能表的烧损建立统一的分析模型,选取电能表所在市州、生产厂家、安装时间、运行时间、所在台区 2020—2021 年历史烧表数量以及退出运行前 30 d 的每日电量数据,每日最大电压和最大电流数据,所在台区每日线损数据作为特征,结合 XGBoost 模型的特点进行实验分析。

3.1 参数调优

XGBoost 模型包含多个参数,按照一般的调参顺序,先从框架参数开始进行调参:

1)“booster”:由于本课题涉及的特征数量较多,选择树模型;

2)“objective”:本课题研究属于分类问题,选择

softmax 分类器;

3)“n_estimators”:XGBoost 算法是由若干个弱分类器组合而成,在一定程度上,弱分类器的数量越多,分类的准确度越高.为了找到最优参数,保持其他参数为默认参数,进行实验,结果如图 2 所示,XGBoost 算法的识别正确率在迭代 700 次时达到最高点,之后逐渐降低.

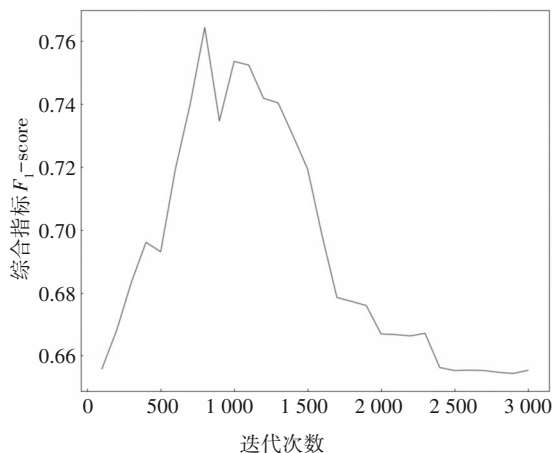


图 2 不同迭代次数时模型的 F_1 -score

Fig. 2 F_1 -score of the algorithm at different estimators

然后,开始调节弱学习器的参数:

4)“max_depth”:固定 XGBoost 算法的迭代次数为 700 次,尝试采用不同树的高度,观察实验结果与树的高度关系.实验结果如图 3 所示,当树的高度分别为 2、4、6、8、10、12 时,随着迭代次数的增加,算法的识别正确率逐渐提高,并且很快趋于稳定,当树的高度为 10 时,算法的性能达到最优.

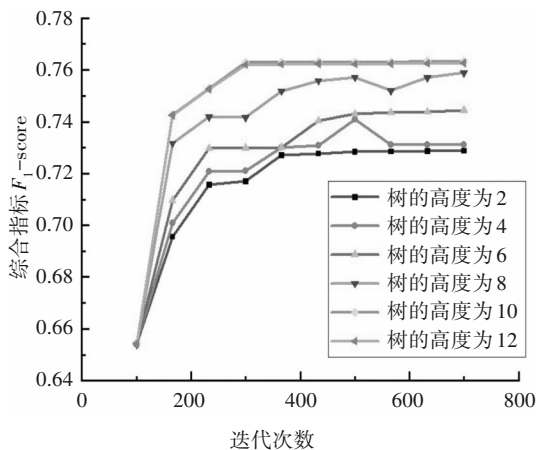


图 3 不同树的高度时模型的 F_1 -score

Fig. 3 F_1 -score of the algorithm at different tree height

5)其他参数:在确定了前面 4 个参数之后,采用

网格搜索,对学习率、gamma 参数、lambda 正则系数、alpha 正则系数等参数按给定步长在一定范围内依次进行调整,最终结果如表 3 所示.

表 3 XGBoost 参数调优

Tab.3 XGBoost parameters' optimization

参数	调整范围	步长	最优值
迭代次数	100~3 000	30	700
树的高度	2~12	2	10
学习率	[0.01, 0.05, 0.07, 0.1, 0.2]	—	0.1
gamma 参数	0.1~0.6	0.1	0.2
lambda 正则系数	[0.05, 0.1, 1, 2, 3]	—	0.1
alpha 正则系数	[0.05, 0.1, 1, 2, 3]	—	1
最小叶子节点样本权重和	1~6	1	2

3.2 特征的重要性排序

使用 XGBoost 建模过程中,可以根据模型自带的 feature_importances 排序工具,评价每个特征对模型的贡献程度,从而对特征变量进行进一步的筛选.图 4 为重要性排在前十的特征量,可以看出,运行时长、台区历史烧表数量、表计止码和烧表前 1 d 的电量为模型中比较重要的特征.

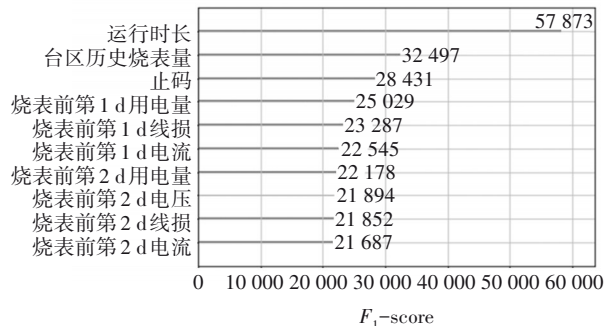


图 4 模型的特征重要性

Fig. 4 Feature importance of the algorithm

因为考虑到模型需要进行实际应用,任意一个网省公司的用户数量都是千万级别,模型的计算量相当大.因此,尝试进行降维,适当压缩电量、电压、电流、线损追溯的天数,在固定其他参数和特征不变的情况下,追溯天数与 F_1 指标的变化情况如图 5 所示.由图 5 可知,往前追溯 15 d 左右,即可得到较好的模型效果,所以为了模型计算更加快捷,将模型需要的追溯时间调整为烧表前的 15 d.

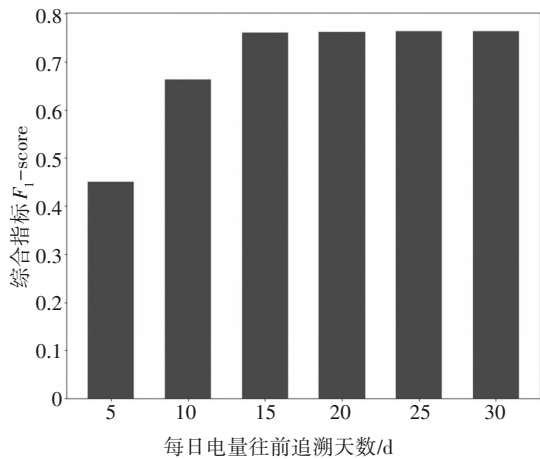


图5 不同追溯天数时模型的 F_1 -score

Fig. 5 F_1 -score of the algorithm at different days traced

3.3 其他机器学习模型结果对比

本文还采用了KNN、朴素贝叶斯、决策树、支持向量机、随机森林、GBDT等进行实验,与XGBoost算法的比对结果如表4所示.实验采取了五折交叉验证的方式,从表4可以看出,基于支持向量机、朴素贝叶斯和KNN的识别指标比较低,分别为27.08%、22.86%和40.92%;采用决策树、随机森林和GBDT的算法识别准确率分别为56.00%、56.00%和72.75%;而基于XGBoost算法的识别准确率为76.51%,远高于传统算法.

表4 与其他机器学习模型的结果对比

Tab.4 Comparison of results of different machine learning algorithm

模型名称	识别指标 F_1 /%
KNN	40.92
朴素贝叶斯	22.86
决策树	56.00
支持向量机	27.08
随机森林	56.00
GBDT	72.75
XGBoost	76.51

3.4 实验结果分析

根据电能表的实际运行数据,烧表情况占比很小,以某网省公司为例,每年烧表的比例不足1%,但由于基数比较大,所以烧表造成的经济损失相当大.

从理论上说,在实现二分类的过程中,如果类别分布不均衡,会在一定程度上影响模型分类的准确性.而在本实例中,数据集中包含的烧表样本仅占总

数的10%.

根据以上所有模型的表现,所有模型的精度与召回率如表5所示.基于朴素贝叶斯的算法精度只有40%;基于KNN、朴素贝叶斯、支持向量机的算法召回率分别为28%、16%和16%;而基于GBDT和XGBoost的算法精度和召回率都有明显提升,其中XGBoost的精度达到91%,召回率也达到66%,相比传统算法,已经有了显著的提升.

表5 精度和召回率对比

Tab.5 Comparison of precision and recall

模型名称	精度/%	召回率/%
KNN	76	28
朴素贝叶斯	40	16
决策树	84	42
支持向量机	88	16
随机森林	84	42
GBDT	88	62
XGBoost	91	66

以上结果表明,本文所提出的基于XGBoost的电能表烧表预测模型能够提高分类的准确性,在不平衡样本中也可以得到较为满意的效果.

4 应用

根据实验模型结果,尝试将该模型应用于实际生产数据进行验证,图6为某电力公司用电信息采集系统的烧表预测模块系统界面.

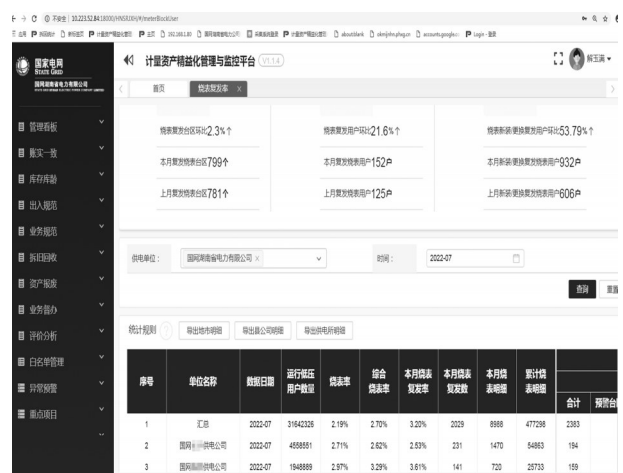


图6 烧表预测模块系统界面

Fig. 6 System interface of the module of burning fault prediction for smart meter

在试运行的时间段内,通过基层供电服务人员
对系统预测情况和电能表的实际运行情况进行现场
确认,测试该模型的综合指标能否达到预期。

在验证过程中,由于全省电能表数据基数大,电
量、电压、电流等电能表运行数据受到自然条件影
响,每天的数据采集成功率无法达到 100%,因此存
在很多的缺失值,需要进行缺失值的补充.这与训练
集数据有所区别,训练集数据都是历史数据,对于存
在缺失数据的样本,可以直接剔除,但验证数据集的
样本是准实时数据,不能直接剔除,如果直接剔除,
会导致样本越来越少.缺失值的填充有多种方法,常
见的填充方法有均值填充和插值填充等,这些方法
可以用于电压数据的填充,因为电压基本稳定在
220 V,但电流和电量数据通常情况下具有较大的波
动性,均值和插值算法都会带来一定程度的误差.由
于电流和电量的预测基本上类似于短期电力负荷的
预测,可以参考文献[13]使用 LSTM 算法对缺失电量
和电流数据进行预测和填充。

为了验证 LSTM 预测填充法和均值、线性插值法
的差异,在训练集中任意选取 1 000 组电能表烧损数
据和 1 000 组未烧损数据,在每组数据中,随机剔除
1 项电流数据或者电量数据,之后,用不同的方法对
剔除的数据进行填充.对比均值填充、线性插值填充
和 LSTM 算法预测填充的准确性,以平均绝对百分比
误差(MAPE)作为评价指标,结果如表 6 所示。

表 6 不同填充方法对比

Tab.6 Comparison of different data imputation methods

插值方法	MAPE/%
均值填充	22.21
线性插值	13.52
LSTM(学习率=0.005)	8.76

经过缺失值的填充与处理,测试集数据预处理
全部完成,用 XGBoost 模型对试点验证的某市公司
800 余个台区,共 10 万余块在运电能表进行了验证,
共验证了 263 块烧表,精确率达到 82.45%,召回率达
到 56.73%,基本上达到了预期水平.目前,该烧表预
测模型已在系统部署并全省运行。

5 结 语

本文针对供电公司低压台区烧表率高的现状,
提出了一种基于 XGBoost 算法的电能表烧表预测方

法,通过对电能表生产厂家的基本信息、每日电量等
运行信息和温度等环境信息,进行单变量和多变量
组合分析,实现了烧表现象的准确预测.模型应用结
果表明,所提算法有效实现了烧表现象的预测,并且
具有广泛推广应用的适用性。

参考文献

- [1] 薛禹胜,赖业宁.大能源思维与大数据思维的融合(一)大数据
与电力大数据[J].电力系统自动化,2016,40(1):1-8.
XUE Y S, LAI Y N. Integration of macro energy thinking and big
data thinking part one big data and power big data[J]. Automation
of Electric Power Systems, 2016, 40(1): 1-8. (In Chinese)
- [2] KABALCI Y. A survey on smart metering and smart grid communi-
cation[J]. Renewable and Sustainable Energy Reviews, 2016, 57:
302-318.
- [3] 郑安刚,张密,曲明钰,等.基于贝叶斯网络的智能电能表故障
类型预测[J].电测与仪表,2018,55(21):143-147.
ZHENG A G, ZHANG M, QU M Y, et al. Prediction of the fault
type of smart meters based on the Bayesian network[J]. Electrical
Measurement & Instrumentation, 2018, 55(21): 143-147. (In
Chinese)
- [4] 朱江森,张月倩,黄艳,等.智能电能表故障风险预估算法的研
究[J].电测与仪表,2018,55(8):97-100.
ZHU J M, ZHANG Y Q, HUANG Y, et al. Research on the algo-
rithm of fault risk prediction for the smart meter[J]. Electrical
Measurement & Instrumentation, 2018, 55(8): 97-100. (In Chi-
nese)
- [5] 张璨辉,赵丹,贺星,等.电能表故障预测的代价敏感模型研究
[J].电测与仪表,2021,58(3):180-184.
ZHANG C H, ZHAO D, HE X, et al. Research on cost-sensitive
fault prediction model for smart meters[J]. Electrical Measure-
ment & Instrumentation, 2021, 58(3): 180-184. (In Chinese)
- [6] 刘金硕,刘必为,张密,等.基于GBDT的电力计量设备故障预
测[J].计算机科学,2019,46(S1):392-396.
LIU J S, LIU B W, ZHANG M, et al. Fault prediction of power me-
tering equipment based on GBDT[J]. Computer Science, 2019, 46
(S1): 392-396. (In Chinese)
- [7] 陈叶,韩彤,魏龄,等.基于多分类融合模型的智能电表故障预
测[J/OL].电测与仪表,(2020-05-20)[2020-08-07]. https://
kns.cnki.net/kcms/detail/23.1202.TH.20200806.1627.012.html.
CHEN Y, HAN T, WEI L, YU X L, L X X. Prediction on fault clas-
sification of smart meters based on multi-classification integration
model[J/OL]. Electrical Measurement & Instrumentation, (2020-
05-20)[2020-08-07]. https://kns.cnki.net/kcms/detail/23.1202.
TH.20200806.1627.012.html. (In Chinese)
- [8] 高欣,刁新平,刘婧,等.基于模型自适应选择融合的智能电表
故障多分类方法[J].电网技术,2019,43(6):1955-1961.
GAO X, DIAO X P, LIU J, et al. A multi-classification method of
smart meter fault type based on model adaptive selection fusion

- [J]. *Power System Technology*, 2019, 43(6):1955-1961. (In Chinese)
- [9] 李宁,费守江,刘国亮,等. 基于深度信念网络的计量装置故障溯源研究[J]. *智慧电力*, 2020, 48(7):118-124.
LI N, FEI S J, LIU G L, *et al.* Fault traceability of metering device based on deep belief network[J]. *Smart Power*, 2020, 48(7):118-124. (In Chinese)
- [10] 肖勇,马喆非,罗鸿轩,等. 基于深度信念网络与数据聚合模型的智能电表数据异常检测方法[J]. *南方电网技术*, 2021, 15(1):99-106.
XIAO Y, MA Z F, LUO H X, *et al.* Anomaly detection method of smart meter based on deep belief network and data aggregation model[J]. *Southern Power System Technology*, 2021, 15(1):99-106. (In Chinese)
- [11] 刘紫熠,刘卿,王崇,等. 基于智能电表运行故障数据的纵向分析模型[J]. *计算机科学*, 2019, 46(S1):436-438.
LIU Z Y, LIU Q, WANG C, *et al.* Vertical analysis based on fault data of running smart meter[J]. *Computer Science*, 2019, 46(S1):436-438. (In Chinese)
- [12] 高欣,纪维佳,赵兵,等. 不平衡数据集下基于CVAE-CNN模型的智能电表故障多分类方法[J]. *电网技术*, 2021, 45(8):3052-3060.
GAO X, JI W J, ZHAO B, *et al.* Multi-classification method of smart meter fault types based on CVAE-CNN model under imbalanced dataset[J]. *Power System Technology*, 2021, 45(8):3052-3060. (In Chinese)
- [13] 张宇航,邱才明,贺兴,等. 一种基于LSTM神经网络的短期用电负荷预测方法[J]. *电力信息与通信技术*, 2017, 15(9):19-25.
ZHANG Y H, QIU C M, HE X, *et al.* A short-term load forecasting based on LSTM neural network [J]. *Electric Power Information and Communication Technology*, 2017, 15(9):19-25. (In Chinese)
- [14] 王恒,姜亚楠,张欣,等. 基于梯度提升算法的岩性识别方法[J]. *吉林大学学报(地球科学版)*, 2021, 51(3):940-950.
WANG H, JIANG Y N, ZHANG X, *et al.* Lithology identification method based on gradient boosting algorithm [J]. *Journal of Jilin University (Earth Science Edition)*, 2021, 51(3):940-950. (In Chinese)
- [15] 张津京,王兴贵,宋汶秦,等. 基于GBDT的串联结构微电网与电网交互稳定性评估[J]. *湖南大学学报(自然科学版)*, 2021, 48(2):87-95.
ZHANG J J, WANG X G, SONG W Q, *et al.* Assessment of interactive stability between grid and series structure micro-grid based on GBDT [J]. *Journal of Hunan University (Natural Sciences)*, 2021, 48(2):87-95. (In Chinese)
- [16] 毕云帆,撤奥洋,张智晟,等. 基于模糊Bagging-GBDT的短期负荷预测模型研究[J]. *电力系统及其自动化学报*, 2019, 31(7):51-56.
BI Y F, HAN A Y, ZHANG Z S, *et al.* Study on short-term load forecasting model based on fuzzy bagging-GBDT [J]. *Proceedings of the CSU-EPSA*, 2019, 31(7):51-56. (In Chinese)
- [17] 王新伟,钱虹,冷述文,等. 基于XGBoost算法的汽轮机转子故障原因定位方法[J]. *动力工程学报*, 2021, 41(6):460-467.
WANG X W, QIAN H, LENG S W, *et al.* Fault location of steam turbine rotor based on XGBoost algorithm [J]. *Journal of Chinese Society of Power Engineering*, 2021, 41(6):460-467. (In Chinese)
- [18] CHEN T, GUESTRIN C. Xgboost: a scalable tree boosting system [C]//*Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*. New York: ACM, 2016: 785-794.
- [19] 叶倩怡,饶泓,姬名书. 基于Xgboost的商业销售预测[J]. *南昌大学学报(理科版)*, 2017, 41(3):275-281.
YE Q Y, RAO H, JI M S. Sales prediction of stores based on xgboost algorithm [J]. *Journal of Nanchang University (Natural Science)*, 2017, 41(3):275-281. (In Chinese)
- [20] 李大中,王超,李颖宇. 基于XGBoost算法的风机叶片结冰状态评测[J]. *电力科学与工程*, 2019, 35(9):43-48.
LI D Z, WANG C, LI Y Y. Evaluation of fan blade icing based on XGBoost algorithm [J]. *Electric Power Science and Engineering*, 2019, 35(9):43-48. (In Chinese)
- [21] 蒋晋文,刘伟光. XGBoost算法在制造业质量预测中的应用[J]. *智能计算机与应用*, 2017, 7(6):58-60.
JIANG J W, LIU W G. Application of XGBoost algorithm in manufacturing quality prediction [J]. *Intelligent Computer and Applications*, 2017, 7(6):58-60. (In Chinese)