Energy-aware Theft Detection based on IoT Energy Consumption Data

Zunaira Nadeem*, Zeeshan Aslam†, Mona Jaber*, Adnan Qayyum‡, and Junaid Qadir§
* School of Electronic Engineering and Computer Science, Queen Mary University of London, London, UK
{z.nadeem, m.jaber}@qmul.ac.uk
† Department of Computer Science, Bahria University, Islamabad, Pakistan
zeeshanxh@gmail.com
‡ Department of Computer Science, Information Technology University, Lahore, Pakistan
adnan.qayyum@itu.edu.pk
§ Computer Science and Engineering Department, College of Engineering, Qatar University, Doha, Qatar
jqadir@qu.edu.qa

***Abstract*—With the advent of modern smart grid networks, advanced metering infrastructure provides real-time information from smart meters (SM) and sensors to energy companies and consumers. The smart grid is indeed a paradigm that is enabled by the Internet of Things (IoT) and in which the SM acts as an IoT device that collects and transmits data over the Internet to enable intelligent applications. However, IoT data communicated over the smart grid could however be maliciously altered, resulting in energy theft due to unbilled energy consumption. Machine learning (ML) techniques for energy theft detection (ETD) based on IoT data are promising but are nonetheless constrained by the poor quality of data and particularly its imbalanced nature (which emerges from the dominant representation of honest users and poor representation of the rare theft cases). Leading ML-based ETD methods employ synthetic data generation to balance the training the dataset. However, these are trained to maximise average correct detection instead of ETD. In this work, we formulate an energy-aware evaluation framework that guides the model training to maximise ETD and minimise the revenue loss due to mis-classification. We propose a convolution neural network with positive bias (CNN-B) and another with focal loss CNN (CNN-FL) to mitigate the data imbalance impact. These outperform the state of the art and the CNN-B achieves the highest ETD and the minimum revenue loss with a loss reduction of 30.4% compared to the highest loss incurred by these methods.**

***Index Terms*—Electricity theft detection, Convolutional neural network, Internet of Things (IoT), Smart meters, Data imbalance**

I. Introduction

The world is facing the worst energy crisis in decades affecting all countries and threatening lives.[1] Energy theft (ET) exacerbates the energy crisis harming power utilities and consumers in terms of increased market volatility, revenue loss, surging costs, and risks to public safety (such as fires and electric shocks). ET is primarily caused by bypassing or tampering with the electricity meter, direct tapping from feeders, injecting magnetic material to slow down metering and high line losses, and cyber-attacks for altering smart meters (SM) [1]. A study conducted in 2017, estimated the effect of ET on revenue loss at around \$96 billion worldwide[2]; it is expected that this will increase in coming years. Conventional methods for energy theft detection (ETD) often require conducting physical onsite inspections which are time-consuming, costly, labor-intensive, and prone to human or equipment error. The fast deployment of smart meters (SM) and the Internet of Things (IoT) have enabled the smart grid paradigm, whereby energy consumption data is available to energy providers in near-real time and with very high granularity [2]. However, IoT is often composed of constrained devices that are prone to cyber-attacks [3].

ET is a general problem afflicting numerous countries around the world [4]. For instance, in the United Kingdom (UK), an estimated 86% of the energy suppliers fail to meet the residential requirements of Office of Gas and Electricity Markets (Ofgem)[3], despite the directive to identify, investigate, and prevent energy theft. Illegal tampering with electricity results in revenue loss due to theft of more than £440 million annually, thereby seriously threatening the UK energy system.

With the advent of the IoT and the spread of SM usage, energy consumption (EC) is available in near-real time. This has resulted in a surge in ETD research using machine learning (ML) [5] [6] [7] [8] [9] [10]. However, despite the high generic accuracy and F1-score values reported in existing works, these do not reflect the actual theft detection rate instead, merely a successful classification rather than an accurate picture of theft. In contrast, this paper proposes an energy-aware deep learning (DL) based ETD method that employs a convolution neural network with positive bias (CNN-B) and focal loss biased CNN (CNN-FL) to counter data bias and improve theft representation while outperforming in revenue loss, thus benefiting energy companies in depicting an accurate picture of theft. Our results reveal that prior ETD art using DL at best identifies 55.2% of actual theft whilst mis-classifying 27.4% of honest users. In contrast, our proposed CNN-B approach has the best detection rate of theft 65.6% and a lower mis-

[1] https://www.spglobal.com/en/research-insights/featured/an-expanding-global-energy-crisis

[2] https://www.prnewswire.com/news-releases/96-billion-is-lost-every-year-to-electricity-theft-300453411.html

[3] https://www.power-technology.com

classification of honest users, 17.7% respectively . Given the inherent trade-off between improving ETD and reducing misclassification of honest users, we formulate a multi-objective metric based on the revenue loss due to incorrect ETD. Our proposed method achieves the highest ETD and the lowest revenue loss, 30.4% lower than the cost incurred by our implementation of [9].

Our contributions are summarized as follows:

1) We propose a novel set of energy-aware detection metrics true positive rate (TPR) and false positive rate (FPR) that focus on theft detection and derive the estimated revenue loss accordingly.
2) We propose CNN-B with a loss function that prioritise TPR and conduct a sensitivity analysis to tune the loss parameters for the best results.
3) A repository of the full code used in this work will be made available to promote research in this crucial area.[4]

The remaining paper is organized as follows. Section II discusses the background of this research and surveys the state of the art related work. Section III presents the problem formulation and dataset description. We illustrate our proposed CNN-B methodology in Section IV. Lastly, the experimental results are presented and discussed in Section V before concluding the paper in Section VI.

II. Background

Yan and Wen [11] identified three ETD directions that use EC data: game theory or game-based, state-based, and ML-based ETD. Game-based ETD formulates interactive decision-making between players: utility, thieves, and consumers. For example, authors in [12] formulate the ETD model as a game for utility and electricity thieves and find the Nash equilibrium of the game. Such methods are promising, but defining utility functions for each player is challenging and non-trivial. State-based ETD uses SM data to estimate bus voltage magnitudes and angles. Maximising or minimising selection criteria estimates state variables. To estimate feeder bus voltage and angle, Huang et al. [13] proposed a three-phase state. It is unlikely to know detailed topology and values to estimate their states.

ML-based ETD, on the other hand, aims to analyze the historical customer energy consumption data. The challenge for ML analysis is that real-time datasets have fewer malicious data samples, which renders the model biased to represent honest users. To this end, ML-based ETD often employs synthetic data to accommodate the imbalance ratio of datasets. Punmiya and Choe [5] proposed three gradient boosting techniques: extreme, categorical, and light gradient boosting machine (LGBM). Buzau et al. [6] proposed SM data and geographical information to analyze consumers' abnormal EC behavior. They employed supervised learning based on a real EC dataset collected in Spain and investigated the performance of conventional ML algorithms such as support vector machine (SVM), linear regression (LR), K-nearest neighbors (kNN), and XGBoost. SVM is similarly used in [7] with a different dataset provided by Tenaga National Berhad, the Malaysian multinational electricity company. The outcomes reported in both works are promising for EC-data-driven ETD but conventional ML suffers from two main shortcomings. First, they require manual feature engineering which is challenging and often relies on in-depth domain knowledge. Second, existing methods fail to handle high-dimensional data.

Deep learning (DL) methods have successfully automated feature engineering and perform well in multi-dimensional data. It follows that authors in [8] adopted DL techniques for ETD including long short-term memory (LSTM), UNet, and the assembling technique Adaptive Boosting (Adaboost). In [9], a simple variational-attention autoencoder with LSTM is proposed to predict energy theft in the State Grid Corporation of China (SGCC) dataset [1]. These papers present results that outperform traditional ML methods (e.g., SVM).

III. Problem Formulation and dataset

The smart electric model was adopted, in which a central server controls each home. Consider $N > 1$ residential locations where each household has an SM connected to the IoT. For each time instance t, each SM s^n gathers a single value representing the household's aggregate EC. Let $x^n(t) \in \mathbb{R}$ be a data instance that is collected from s^n SM at an hour t. Let $X \in \mathbb{R}^n$ be a sequence of n consecutive data instances that are collected by s^n SM over a duration T. Let y^n be the associated label with every sequence X which indicates whether ET was detected ($y^n = 1$) or not ($y^n = 0$) for every household being observed such that $y^n = \{0, 1\}$, where $n = \{1, 2 \cdots N\}$. Let $y^n \in \mathbb{R}^N$ be discrete labels whose values are to be modeled and predicted by the input data X. The minority class refers to the class with few samples in the data X, while the majority class refers to the class with many samples in the same data. The ratio between these two types of samples is referred to as *Imbalance ratio* IR, which is defined as a proportion of a number of samples in the minority class ($y^n = 1$) to the majority class ($y^n = 0$).

Our dataset contains data samples X and the corresponding discrete labels Y. In this dataset, $\mathcal{D} = \{(x_1, y_1), ..., (x_i, y_i)), , (x_N, y_N)\}$, where the each example pair in $\mathcal{D}$ denotes a data sample (i.e., feature vector) taken by SM s^n during a duration T and is labelled (through y^n) as either an honest sample (when $y^n = 0$) or theft (when $y^n = 1$). To this end, in our binary classification case, we denote data samples containing all theft samples (i.e., the minority class) as $\mathcal{D}_t \subset \mathcal{D} = \{(x_i, y_i) | y_i = 1\}$ and all honest samples (i.e., the majority class) as $\mathcal{D}_h \subset \mathcal{D} = \{(x_i, y^n) | y^n = 0\}$.

A. Dataset and Pre-Processing

The dataset $\mathcal{D}$ comprises real-time EC records of $42,372$ residential consumers (SGCC; Jan 2014-Oct 2016) with 10 time more honest users then dishonest. The daily EC records are available for 1034 days. From revenue loss perspective, it is fortunate that the number of dishonest users is low, however, such skewed dataset brings additional challenges to data-driven modelling [14].

[4] https://github.com/zunairanadeem/EnergyAwareTheftDetection

Attributes	Raw data	Clean data
Total Customers	$42,372$	$\lvert\mathcal{D}\rvert = 41897$
Honest Customers	38757	$\lvert D_h\rvert = 38,321$
Dishonest Consumers	3615	$\lvert D_t\rvert = 3,576$
Outliers	475 (39 theft)	

TABLE I: Overview of the SGCC Dataset

In addition, the dataset includes noisy samples (these are those with missing values or interrupted EC collection). It follows that, before being used for model training, the dataset needs to be pre-processed in two steps: (1) data imputation to replace missing values, and outlier detection to remove noisy samples. We adopt linear interpolation as an imputation method and the three-sigma rule method for outlier detection, similar to [8]. The cleaned SGCC dataset is described in Table I. Before using the clean data for model training, it is first scaled through the normalization technique ranges between 0 and 1. The remaining challenge in the data in Table I is the dominance of honest users over theft cases ($IR = 9.3\%$); such imbalance is bound to create a bias in the data modelling.To this end, we use combined under and over-sampling techniques, known as synthetic minority over-sampling (SMOTE) [15] to re-balance the training dataset and a customised loss to increase the model's sensitivity to maximising TPR [14].

IV. Methodology

In this section, we first introduce an energy-aware evaluation framework that reflects the actual theft detection rate. Next, two modified versions of the CNN are presented. The first, CNN-B, embeds positive bias in the loss function to combat the data imbalance. The second, CNN-FL, employs focal loss instead of cross entropy to account for the data imbalance.

A. Energy-aware Evaluation Framework

Due to its highly imbalanced nature, ETD demands a more nuanced energy-aware evaluation as in the presence of highly imbalanced data, the model accuracy Eq. (2) is not representative of the performance of any ETD model. Indeed, given the 91.5% share of honest users in the given dataset, if all users are classified as honest then the calculated accuracy would be 91.5% but the actual theft detection rate will be 0%.

Energy-aware performance evaluation needs to be measured from two perspectives. *Firstly*, TPR (see Eq.(1)), where the aim is to maximize the number of actual thefts that are correctly classified as such. *Secondly*, the FPR- (see Eq. (3)), where the aim is to reduce the number of mis-classified honest users. Both a low TPR and a high FPR incur revenue loss; the former causes direct loss and the latter requires a costly onsite inspection to remove the doubt of theft.

In an attempt to circumvent the pitfall of relying solely on accuracy Eq. (2) for ETD, Massaferro et al. [16] use a Weighted F1 Eq. (6) score instead. On the other hand, authors in [8], use the Weighted Recall (see Eq. (7)) for measuring the detection rate instead of the more ETD-specific formulation (see . (1)). We show in Section V how these metrics can be misleading and not fully representative of actual ETD.

In this work, we formulate the first enery-aware ETD evaluation framework in terms of revenue loss incurred by misclassification of ETD. We first define TPR Eq. (1) and FPR Eq. (3) where, P and N refer to the number of positive (theft) and negative (honest) data points, respectively. TP, TN, FP, and FN refer to the numbers of true positive, true negative, false positive, and false negative samples, respectively. TPR is defined as TP divided by all theft samples ($\lvert\mathcal{D}_t\rvert = TP{+}FN$). FPR is defined as the ratio between FP and the number of all honest samples ($\lvert\mathcal{D}_h\rvert = TN + FP$).

$$TPR = \frac{TP}{TP + FN} \quad (1) \qquad Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$FPR = \frac{FP}{TN + FP} \quad (3) \qquad Precision = \frac{TP}{TP + FP} \quad (4)$$

$$F1 = \frac{2{\times}TP}{2{\times}TP + FP + FN} \tag{5}$$

$$WF1 = \frac{(\lvert D_h\rvert{\times}F1\_0) + (\lvert D_t\rvert{\times}F1\_1)}{\lvert D\rvert} \tag{6}$$

$$Wrecall = \frac{(\lvert D_h\rvert{\times}recall\_0) + (\lvert D_t\rvert{\times}recall\_1)}{\lvert D\rvert} \tag{7}$$

Next, we formulate a multi-objective metric that optimises both TPR and FRP. A good model would result in high TPR and low FPR but there is a trade-off between these two objectives. From an energy provider perspective, both TPR and FPR are costly and an ETD method needs to find the right balance that will minimise the incurred cost. It follows that an energy-aware multi-objective metric should be formulated based on the incurred revenue loss from erroneous ETD (FPR and FNR). Based on the latest published report by Ofgem [17], the number of energy thefts in 2012 was $16,714$ and the retail value of the volume of electricity illegally taken was £$19,116,506$. Thus, the average annual revenue loss due to unaccounted electricity consumption is C_1=£1143.7 per case. The average duration of unidentified electricity theft was $M = 1.4$ years. The overhead annual cost for tackling electricity theft reported is £$6,395,000$ covering $46,447$ investigated cases; thus the cost for investigating a falsely identified theft can be estimated at C_2=£137.7 per case per year. It follows that the energy-aware cost of failed ETD can be formulated as follows ($\lvert D_t\rvert$ and $\lvert D_h\rvert$ from Table I):

$$C_{Total} = M \cdot (C_1 \cdot (1 - TPR) \cdot \lvert D_t\rvert) + C_2 \cdot FPR \cdot \lvert D_h\rvert \tag{8}$$

B. Positive Biased 1D-CNN (CNN-B)

We first build a one dimensional CNN (1D-CNN) to extract the sample features, followed by a sigmoid function that completes the role of classification. The adopted 1D-CNN signifies that for each input signal x_i with index $1 \leq i \leq N$, the sigmoid classifier produces an output $\mathbf{p}_i = \{p_{i,0}, p_{i,1}\}$ such that $0 \leq p_{i,\text{class}} \leq 1$ is the confidence level of input sample x_i matching signals of type 0 (honest) or 1 (theft). The proposed CNN architecture employs four convolutional layers with 128, 64, 32 and 32 filters, respectively (see Table II). The input layer of two 1D convolutional layers follows a max-pooling layer followed by a dropout layer. A flattened layer is

Layers	**No of Kernels**	**Size of Kernel**	**Activation**
Convo1D-1	128	3	ReLU
Convo1D-2	64	3	ReLU
MaxPooling1D	pool_size=2		-
Dropout	0.2		-
Convo1D-3	32	3	linear
Convo1D-4	32	3	linear
MaxPooling1D	pool_size=2		-
Dropout	0.2		-
Flatten			-
Dense	neuron=20		linear
Dense	neuron=1		sigmoid

TABLE II: Architecture of Proposed CNN

β	**0.5**	**0.6**	**0.7**	**0.9**
TPR	38.44	39.98	48.89	**65.58**
FPR	4.9	5.42	8.88	**17.73**
Cost (K£)	3753	3693	3370	**2890**

TABLE III: Sensitivity analysis of CNN-B for different β values.

	Binary loss (γ)				**Sigmoid loss(α=0.25,γ)**			
	0	**1**	**2**	**5**	**0**	**1**	**2**	**5**
TPR	40.46	**50.43**	45.92	39.11	26.46	29.14	**31.92**	29.24
FPR	5.5	**8.98**	8.56	6.02	1.96	2.47	**2.77**	2.75
Cost (K£)	3670	**3283**	3522	3773	4278	4153	4011	4162

TABLE IV: Sensitivity analysis for CNN-FL with different γ and α values.

then used to convert the last max pooling layer's output into a one-dimensional vector, followed by two fully connected layers with 20 and 1 neuron, respectively—the sigmoid activation function to calculate the probability for each class. The simulation hyperparameters are batch size 32, epochs 200, learning rate 0.001 and communication round one. The proposed model uses the objective function of binary cross-entropy (CE) since ETD is a binary classification problem.

During the model training, kernels are optimised through backpropagation which calculates a partial derivative of a loss function. In a binary classification using the sigmoid function, the loss function is based on the binary CE, as shown below:

$$\mathrm{L} = -y_i \cdot \log(p_{i,1}) - (1 - y_i) \cdot \log(p_{i,0}) \tag{9}$$

where, $y_i = [0, 1]$ is the ground truth label of sample x_i and $[p_{i,0}, p_{i,1}]$ are the confidence level calculated by the sigmoid function. To counter the bias in the dataset, we propose a biased version of binary CE using the factor β, which is set equal to the penetration of the majority class, i.e., honest samples, as shown below:

$$\mathrm{L}_{\mathrm{bias}} = -\beta y_i \cdot \log(p_{i,1}) - (1 - \beta)(1 - y_i) \log(1 - p_{i,0}) \tag{10}$$

Table III shows the performance of CNN-B with different values of β. As can be seen, increasing β results in better TPR with the side effect of increased FPR. From an energy-aware perspective, the choice would be geared by the associated cost of each of the false classifications. Practically, an unnoticed theft result is a direct and possibly long-term loss ($C_1 \times M$). On the other hand, a falsely classified theft would incur a site visit to verify the claim; this would be a one-time limited cost (C_2) in Eq (8).

In the presence of a skewed dataset, the value of β is often determined based on IR such that $\beta \sim (1-IR)$. In this case, IR is 0.092 and $\beta \sim 0.9$ would generate a suitable positive bias. In Table III, we implement the CNN-B with different β values (0.5,0.6,0.7,0.9) and confirm the expected impact on TPR and FPR; indeed, $\beta = 0.9$ give the best results as indicated by the incurred cost which is the lowest in Table III.

C. Focal Loss Biased CNN (CNN-FL)

CNN-FL is another known approach for addressing data imbalance and consists of applying weights to minority and majority classes. This is a simple extension of CE loss in Eq. (9). We implemented two variants of focal loss using TensorFlow addons: binary focal loss and sigmoid focal loss. A formulation of the alpha-balanced CE focal loss is $FL(p_i) = -\alpha_i \log(p_i)$. More precisely, we propose to add a modulating factor (1 - p_i)$^\gamma$ to the binary CE loss, with tunable focusing parameter gamma $\gamma \geq 0$. We define the binary focal loss as $FL(p_i) = -(1 - p_i)^\gamma \log(p_i)$.

The focal loss is visualized for several values of $\gamma \in$ [0, 5] in Table IV. We found that γ=1 performed best in our experiments. We use another variation sigmoid loss function where α-balanced variant of the focal loss can be seen in Eq. (11). In Table IV, the value of α is constant at 0.25 with variable values of γ. We found α=0.25, γ=2 performed well among all in terms of TPR.

$$\mathrm{FL}(p_i) = -\alpha_i (1 - p_i)^\gamma \log(p_i) \tag{11}$$

Comparing the results in Tables III and IV, CNN-B with the positive bias loss function seems to perform better at detecting the actual theft. This is an expected outcome as, for highly skewed datasets, a positive bias is often the recommended remedy to counter the data bias [14]. In the next section, both the tuned CNN-B and CNN-FL are benchmarked against leading state-of-the-art ETD methods.

V. Experimental Results

In this section, we present the results obtained using CNN-B with $\beta = 0.9$ and CNN-FL with $\gamma = 1$ and we compare these to state-of-the-art ETD models based on the energy-aware framework depicted in Section IV-A.

A. Energy-Aware Comparative Analysis

We first examine the performance of four prominent ETD ML algorithms, listed below, using the energy-aware framework. In addition, we evaluate the two modified 1D-CNN models, CNN-B and CNN-FL, described in Sections IV-B and IV-C, respectively:

- **LGBM** [5] has been selected for fast training speed and accuracy. In [5], the authors present 97% TPR and 7% FPR. In comparison to this work, our proposed solution gives 6.81% FPR.
- **LSTM** model is used to transform features into a sequence of time-based features [9]. It has been selected for

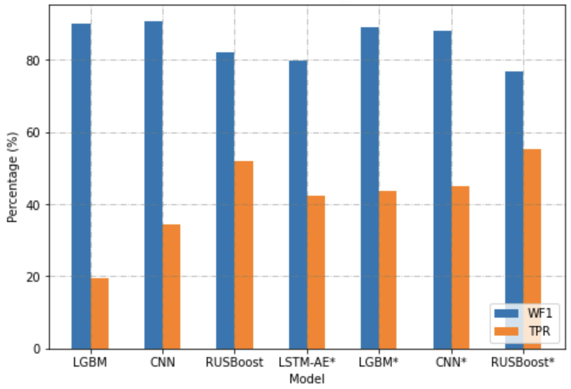

Fig. 1: The trade-off between WF1 and TPR. Note that the results indicated with * refer to those obtained after applying SMOTE.

its ability to work well with time series data. We implemented the encoder-decoder using a sequence of LSTM unit cells that continue to decrease the high-dimensional input vector to a low-dimensional input vector until latent space is reached. The results are based on one hidden LSTM layer, a dropout layer, a repeat vector layer where a timestamp of 1035 (full user record for one day) is selected, and two fully connected layers.

- **CNN** [1] has been selected for its known ability to extract temporal features. The results are based on area under the curve (AUC) and mean average precision. The authors used 100 epochs in their study, but also acknowledged that too few epochs may not allow the model to learn the underlying patterns in the data, while too many epochs may cause overfitting.
- **CNN-FL** is a CNN model with optimised focal loss to counter the effect of data bias, see Section IV-C.
- **CNN-B** is our proposed algorithm that leverages positive bias to counter the effect of data bias, see Section IV-B.

For each implementation, generic performance metrics accuracy Eq. (2), precision Eq. (4) and F1 Eq. (5) are calculated in addition to energy-aware metrics TPR Eq. (1) and FPR Eq. (3). The total cost of failed ETD is calculated based on Eq. (8) and for $|D_h|$ and $|D_t|$ values in Table I. The results prior to SMOTE implementation are first shown in Table V and those that are obtained after balancing the training dataset using SMOTE are shown in Table VI.

B. Discussion and Analysis

In this section, we investigate some insights gained from reproducing state-of-the-art ETD methods and using our proposed metrics to evaluate: the role of SMOTE, positive bias, and ML method in the effective ETD.

Interpretation of metrics: Conventional metrics that are often used in the presence of a biased dataset are deceptive, as shown in Figure 1. While LGBM achieves the highest Weighted F1-score (WF1) in both Tables V and VI, from an energy-aware evaluation framework, this method fails to detect theft 80% and 56% of the time with and without SMOTE, respectively. Furthermore, the ETD cost incurred using this detection method is the highest in Table V and ranks high in Table VI. This validates that WF1 is not an indicative performance metric to evaluate ETD. CNN is shown to yield the highest Weighted Recall (Wrecall) in Table VI, however it leaves 65% of theft undetected and a high revenue loss.

RUSBoost provides the best TPR results with and without SMOTE as seen in Tables V and VI. It is worth noting that this method uses a combination of RUS (random under-sampling) and the standard boosting procedure with ensemble learning. Thus, it is designed to better model the minority class by removing majority class samples. Therefore, the revenue cost without SMOTE is lower than that in Table VI, but remains 29% higher that CNN-B.

SMOTE: The effectiveness of tackling data bias with rebalancing the training dataset with SMOTE is further studied by comparing the results in Tables V and VI. SMOTE is seen to improve the TPR performance in LGBM and CNN by 125% and 31%, respectively. However, the impact on cost is much more limited with 23% and 10% reduction for LGBM and CNN, respectively.

Loss function: Next, we examine the impact of modified loss on ETD by comparing the gain achieved with SMOTE to the gain achieved with positive bias. As highlighted above, the TPR gain when comparing CNN to CNN-with-SMOTE is 31% and the cost reduction is 10%. However, the gain of the modified loss function with positive bias (CNN-B in Table V) over the CNN is 56% in TPR and 18% in cost reduction. These results indicate that in the case of highly skewed data, a loss function with positive bias is significantly more effective than synthetic data generation. When combining SMOTE with the positive bias loss function (CNN-B with SMOTE) the revenues loss is reduced by 25.2% compared to CNN without SMOTE.

Trade-off between TPR and FPR: We examine the trade-off between TPR and FPR in Figure 2, where the TPR is plotted against FPR using the AUC for each method listed in Table VI. We introduce a threshold that distinguishes between honest and dishonest users. With an AUC of 0.80, our proposed model effectively separates malicious users, as shown in Figure 2. It can be seen that CNN-B manages to find the best balance between high TPR and low FPR owing to the optimised positive bias in the loss function. This finding is confirmed when examining the incurred cost of energy theft. Whereas CNN achieves the highest weighted recall, it incurs an increase in the cost of 19.5% compared to CNN-B. The WF1 gain achieved by CNN-B in comparison with the next best method, LGBM, in Table VI is less than 1%. However, the cost reduction is 18.7% and the improvement of TPR is 50%. This analysis further reinforces the importance of energy-aware performance metrics in any ETD evaluation.

VI. Conclusions

This paper presented an energy-aware performance evaluation of machine learning (ML) based energy theft detection (ETD) methods using smart meter energy consumption (EC) data available obtained over the IoT. ML-based ETD is challenging due to the EC data being dominated by honest users which creates conflicting objectives of maximising actual theft

	Conventional			Energy-aware			
Model	**Acc**	**Wrecall**	**WF1**	**F1**	**TPR**	**FPR**	**Cost (K£)**
LGBM	**92.25**	**92.25**	**90.11**	30	19.46	**0.64**	4,605.2
RUSBoost	78.44	77.82	82.17	30.36	51.94	18.97	3,724.3
CNN	91.86	89.21	90.93	34.30	34.36	2.52	3,858.2
CNN-B	86.92	85.27	88.32	32.79	**53.59**	10.06	**3, 162.7**

TABLE V: Conventional and energy-aware metrics (%) for three existing models and the proposed CNN-B before applying SMOTE

	Conventional			Energy-aware			
Model	**Acc**	**Wrecall**	**WF1**	**F1**	**TPR**	**FPR**	**Cost (K£)**
LGBM	88.79	88.95	89.11	41.17	43.75	**6.81**	3,550.5
RUSBoost	71.05	71.36	77.00	25.57	55.16	27.39	3,983.8
LSTM-AE	79.84	84.40	79.84	23.79	42.47	16.77	4,146.3
CNN	87.70	**90.65**	88.04	41.34	45.06	6.36	3,452.5
CNN-FL	87.65	90.10	88.74	37.61	50.43	8.98	3,285.3
CNN-B	**89.60**	72.19	**89.93**	30.20	**65.58**	17.73	**2, 890.0**

TABLE VI: Conventional and energy-aware metrics (%) for four existing models and the proposed CNN-FL and CNN-B after applying SMOTE

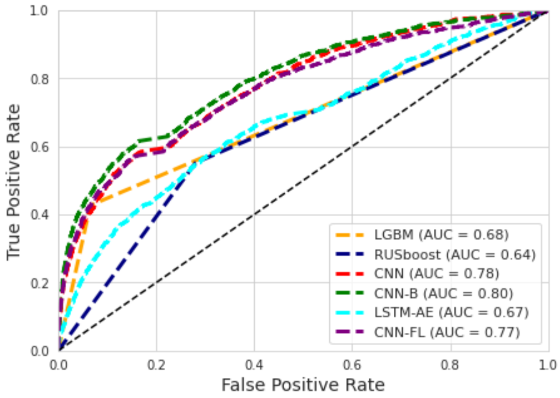

Fig. 2: AUC plots for all models with SMOTE in Table VI

detection (TPR) and minimising mis-classification of honest instances (FPR). To this end, we propose the first energy-aware multi-objective metric which is formulated based on the incurred cost of undetected theft and FPR. Next, we present a modified convolution neural network (CNN-B) and (CNN-FL) that mitigates the trade-off by adjusting the loss function to the imbalance ratio. Our model CNN-B achieves the highest theft detection rate of 65.58% and the least incurred revenue loss that is up to 30.4% less than the state-of-the-art.

References

[1] Z. Zheng, Y. Yang, X. Niu, H. N. Dai, and Y. Zhou, "Wide and deep convolutional neural networks for electricity-theft detection to secure smart grids," *IEEE Transactions on Industrial Informatics*, vol. 14, pp. 1606–1615, 4 2018.

[2] Y. Fathy, M. Jaber, and Z. Nadeem, "Digital twin-driven decision making and planning for energy consumption," *Journal of Sensor and Actuator Networks*, vol. 10, p. 37, Jun 2021.

[3] H. Alrubayyi, G. Goteng, M. Jaber, and J. Kelly, "Challenges of malware detection in the IoT and a review of artificial immune system approaches," *Journal of Sensor and Actuator Networks*, vol. 10, p. 37, Jun 2021.

[4] S. McLaughlin, D. Podkuiko, and P. McDaniel, "Energy theft in the advanced metering infrastructure," in *International Workshop on Critical Information Infrastructures Security*, pp. 176–187, Springer, 2009.

[5] R. Punmiya and S. Choe, "Energy theft detection using gradient boosting theft detector with feature engineering-based preprocessing," *IEEE Transactions on Smart Grid*, vol. 10, pp. 2326–2329, 3 2019.

[6] M. M. Buzau, J. Tejedor-Aguilera, P. Cruz-Romero, and A. Gomez-Exposito, "Detection of non-technical losses using smart meter data and supervised learning," *IEEE Transactions on Smart Grid*, vol. 10, pp. 2661–2670, 5 2019.

[7] J. Nagi, A. M. Mohammad, K. S. Yap, S. K. Tiong, and S. K. Ahmed, "Non-technical loss analysis for detection of electricity theft using support vector machines," in *2008 IEEE 2nd International Power and Energy Conference*, pp. 907–912, 2008.

[8] Z. Aslam, N. Javaid, A. Ahmad, A. Ahmed, and S. M. Gulfam, "A Combined Deep Learning and Ensemble Learning Methodology to Avoid Electricity Theft in Smart Grids," *Energies*, vol. 13, p. 5599, Oct. 2020.

[9] A. Takiddin, M. Ismail, U. Zafar, and E. Serpedin, "Deep autoencoder-based anomaly detection of electricity theft cyberattacks in smart grids," *IEEE Systems Journal*, 2022.

[10] S. Zhang, W. Li, Y. Wu, P. Watson, and A. Y. Zomaya, "Enabling edge intelligence for activity recognition in smart homes," pp. 228–236, Institute of Electrical and Electronics Engineers Inc., 12 2018.

[11] Z. Yan and H. Wen, "Performance analysis of electricity theft detection for the smart grid: An overview," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, 2022.

[12] A. A. Cardenas, S. Amin, G. Schwartz, R. Dong, and S. Sastry, "A game theory model for electricity theft detection and privacy-aware control in AMI systems," pp. 1830–1837, 2012.

[13] S. C. Huang, Y. L. Lo, and C. N. Lu, "Non-technical loss detection using state estimation and analysis of variance," *IEEE Transactions on Power Systems*, vol. 28, pp. 2959–2966, 2013.

[14] Y. Fathy, M. Jaber, and A. Brintrup, "Learning with imbalanced data in smart manufacturing: A comparative analysis," *IEEE Access*, vol. 9, pp. 2734–2757, 2021.

[15] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: Synthetic minority over-sampling technique," *J. Artif. Int. Res.*, vol. 16, p. 321–357, jun 2002.

[16] P. Massaferro, J. M. D. Martino, and A. Fernandez, "Fraud detection in electric power distribution: An approach that maximizes the economic return," *IEEE Transactions on Power Systems*, vol. 35, pp. 703–710, 1 2020.

[17] Smarter Markets, "http://thznetwork.net/index.php/thz-images," 2013. (Date last accessed 15-July-2014).