IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received November 11, 2020, accepted December 1, 2020, date of publication December 4, 2020, date of current version December 22, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3042636

An Attention Guided Semi-Supervised Learning Mechanism to Detect Electricity Frauds in the Distribution Systems

ZEESHAN ASLAM[1], FAHAD AHMED[1], AHMAD ALMOGREN[2], (Senior Member, IEEE), MUHAMMAD SHAFIQ[3], MANSOUR ZUAIR[4], AND NADEEM JAVAID[1], (Senior Member, IEEE)

[1]Department of Computer Science, COMSATS University Islamabad, Islamabad Campus, Islamabad 44000, Pakistan
[2]Department of Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh 11633, Saudi Arabia
[3]Department of Information and Communication Engineering, Yeungnam University, Gyeongsan 38541, South Korea
[4]Computer Engineering Department, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia

Corresponding authors: Ahmad Almogren (ahalmogren@ksu.edu.sa) and Muhammad Shafiq (shafiq.pu@gmail.com)

This work was supported by the Deanship of Scientific Research at King Saud University under Grant RG-1437-035.

ABSTRACT Electricity theft is one of the main causes of non-technical losses and its detection is important for power distribution companies to avoid revenue loss. The advancement of traditional grids to smart grids allows a two-way flow of information and energy that enables real-time energy management, billing and load surveillance. This infrastructure enables power distribution companies to automate electricity theft detection (ETD) by constructing new innovative data-driven solutions. Whereas, the traditional ETD approaches do not provide acceptable theft detection performance due to high-dimensional imbalanced data, loss of data relationships during feature extraction and the requirement of experts' involvement. Hence, this paper presents a new semi-supervised solution for ETD, which consists of relational denoising autoencoder (RDAE) and attention guided (AG) TripleGAN, named as RDAE-AG-TripleGAN. In this system, RDAE is implemented to derive features and their associations while AG performs feature weighting and dynamically supervises the AG-TripleGAN. As a result, this procedure significantly boosts the ETD. Furthermore, to demonstrate the acceptability of the proposed methodology over conventional approaches, we conducted extensive simulations using the real power consumption data of smart meters. The proposed solution is validated over the most useful and suitable performance indicators: area under the curve, precision, recall, Matthews correlation coefficient, F1-score and precision-recall area under the curve. The simulation results prove that the proposed method efficiently improves the detection of electricity frauds against conventional ETD schemes such as extreme gradient boosting machine and transductive support vector machine. The proposed solution achieves the detection rate of 0.956, which makes it more acceptable for electric utilities than the existing approaches.

INDEX TERMS Electricity theft detection, smart grids, relational denoising autoencoder, electricity consumption, TripleGAN.

NOMENCLATURE

X	EC of consumer
$x_{i(t+1)}$	Next EC value
$x_{i(t-1)}$	Previous EC value
x_i	Current EC value
t	Time interval
c	Summation of previous and next EC values
Q_1	Lower quantile
Q_3	Upper quantile
H	Hidden representation of autoencoder
L	Reconstruction of autoencoder
$e(.)$	Encoding function of autoencoder
$se(.)$	Activation function of encoder
$sd(.)$	Activation function of decoder
W	Weights of neural networks
b	Biases of neural networks
X'	Reconstructed input by autoencoder
$d(.)$	Decoding function of autoencoder

The associate editor coordinating the review of this manuscript and approving it for publication was Ravindra Singh.

θ	Hyperparameters of autoencoder
$R(.)$	Associations' function in relational autoencoder
$\hat{X}$	Corrupted input by denoising autoencoder
X''	Reconstructed input by denoising autoencoder
α	Scale variable of relational autoencoder

I. INTRODUCTION

Currently, over 80% of the world uses electricity [1]. Therefore, secure, efficient and reliable distribution of electricity is an important concern of power utilities. Generally, non-technical losses (NTL) are the chief concerns for power utilities because they cover the largest proportion of the total electrical losses. Whereas, the technical losses (TL) include a minor and unavoidable portion of electricity distribution systems' losses, e.g., line losses. Specifically, NTL cover electricity frauds and meters' malfunctioning along with their installation problems and billing mistakes [2]. The act of electricity fraud is the prime cause of NTL, which further leads to power grid instability, inefficiency and poor reliability along with a significant proportion of economic losses. Therefore, the honest electricity consumers are penalized with heavy electricity bills because of the per unit increase in electricity price due to the energy scarcity caused by electricity theft.

Presently, electricity theft is a global issue, which is faced by both developed and underdeveloped nations. For instance, a report stated that a loss of \$89.3 billion is recorded globally on account of electricity theft [1]. Likewise, India and Brazil lose approximately \$4.5 billion annually [3], US has \$6 billion per annum [4], UK suffers around \$173 million/year [5], Canada losses \$100 million annually [6] and other emerging nations lose around 50% of the power utilities' revenue [1]. Hence, these statistics explain the necessity of electricity theft detection (ETD). The electric companies require to detect electricity frauds to ensure reliable and efficient energy distribution and to reduce potential revenue loss.

The common ways of ETD are to conduct manual inspections of selected consumers and audit their previous electricity bills. However, these approaches are inefficient because they are both time and labor-intensive. The advancement of traditional power grids with smart technologies and intelligence gives rise to the concept of smart grids that provide a two-way flow of energy and information [7]. Thus, enabling efficient power management [8], real-time billing and load surveillance [9]. Smart grids create large electricity usage profiles of consumers by recording their energy consumption at short intervals, typically, after every 30 minutes. Moreover, they help in designing and automating the data-driven solutions.

A substantial amount of research is proposed to develop hardware based, data-driven machine learning based and game-theory based solutions for ETD [10]. Hardware based solutions are also called state-driven solutions. These solutions are aimed to design and utilize sensors or physical components to identify the irregular behavior of electricity thefts [11]–[13]. However, these approaches are costly and vulnerable to seasonal effects. The game-theory based solutions rely on the formation of a contest between the power utility and the electricity fraudsters [14], [15]. The design of a utility function in game-theoretic schemes is a challenging and expensive task. Due to the limitations in state-driven and game-theoretic ETD methods, the solutions that rely on data-driven approaches are more practical and cost-efficient. These solutions analyze the consumers' large electricity consumption (EC) records for the identification of electricity fraudsters. Therefore, the prime intention of this work is to focus on data-driven solutions. We characterize these solutions into three groups: supervised solutions, unsupervised solutions and semi-supervised solutions.

A. SUPERVISED SOLUTIONS

Here we discuss the supervised learning mechanisms that require labeled EC data for ETD, which is obtained by on-site inspections. Zheng *et al.* [16] implement an approach that relies on the wide and deep convolutional neural network (WD-CNN) for ETD. Hasan *et al.* [17] present the combination of CNN and long short-term memory (LSTM) where CNN is practiced to derive abstract features while the LSTM is employed for theft detection. The authors in [18] design a hybrid of LSTM and random under-sampling boosting (RUSBoost) for the detection of dishonest consumers. LSTM is employed to capture long term dependencies while RUSBoost is employed for ETD. Likewise, Li *et al.* [19] propose a hybrid technique that utilizes CNN and random forest (RF) where RF acts as a classification layer of CNN. A similar case is presented in [20] where authors execute the extreme gradient boosting (XGBoost) tree method to rank the electricity consumers. It uses the smart meters data together with the auxiliary databases to improve ETD. In another study [21], the authors present firefly optimization based XGBoost for ETD in the smart grid environment. In the system, the meta-heuristic technique is employed to tune the hyperparameters of XGBoost. The authors in [22] apply three variations of the gradient boosting theft detector to identify the electricity misconducts by fraudulent consumers. Another study [23] offers the hand-crafted feature extraction mechanism along with the RUSBoost approach for the identification of abnormality in the EC profiles of consumers. Furthermore, Buzau *et al.* [24] introduce a methodology for sequential and nonsequential data. The former is served to LSTM for long-term dependencies and later is fed to the multi-layer perceptron (MLP) for auxiliary information. Another work [25] introduces a new data sampling mechanism to solve data imbalance concerns together with the usage of bi-directional gated recurrent unit for ETD. Likewise, in [26], an improved sampling technique is presented to handle the imbalanced data along with RF for ETD. Saeed *et al.* [27] present a new decision tree based supervised classification mechanism for the identification of NTL. However, supervised learning solutions solely require a large amount of labeled EC data, which is insufficient and sometimes infeasible in real life because it demands expensive on-field inspections. Moreover, these approaches confront the model's biasness issue due to the imbalanced data.

B. UNSUPERVISED SOLUTIONS

Here we describe some recent advances made for ETD by utilizing the unsupervised learning procedures that work with unlabeled EC histories of consumers. In [28], the authors employ the maximum information coefficient (MIC) together with the fast search and find of density peaks clustering for ETD. Similarly, the authors in [29] implement an unsupervised learning mechanism, namely, density-based spatial clustering of applications with noise to identify abnormalities in smart grids. Likewise, Joaquim *et al.* [30] use an unsupervised clustering technique, named as fuzzy Gustafson-Kessel, for the identification of NTL in smart grids. In [31], a mixture of clustering and deep learning based model is proposed for theft detection in smart grids. In particular, the authors employ K-means for grouping similar consumers while deep learning is used for ETD. The authors in [32] introduce the generative adversarial networks (GANs) based model: VAE-GAN, for dimensionality reduction. Furthermore, they use K-means for the grouping of consumers using extracted features by VAE-GAN. Cheng *et al.* [33] present an approach that utilizes ensemble of autoencoders and also provide a comprehensive study on autoencoders for abnormality detection. Moreover, in [34], the authors execute a clustering technique self-organizing maps to group similar EC instances. Then, the MLP based model is used for the identification of electricity fraudsters. The unsupervised solutions widely consist of clustering based methods, which have less capability to handle high-dimensional noisy data and also fail to extract refined meaningful features. Similarly, few of them are evaluated on synthetic data that do not provide a realistic assessment. Moreover, clustering methods search for centroids and group closer instances into a single cluster where the misjudgment uplifts the misclassification rate.

C. SEMI-SUPERVISED SOLUTIONS

This section investigates the ETD methods that efficiently utilize the supervised and unsupervised EC records. Tianyu *et al*. [35] present a multi-tasking based semi-supervised learning fraud detector for ETD in which authors gain the benefits of both labeled and unlabeled information. The authors in [36] introduce a semi-supervised mechanism, which is a combination of CNN, LSTM and stacked autoencoder (SAE). This mechanism uses the concept of transfer learning. Another work done in [37] implements a semi-supervised support vector machine (SVM), named as transductive SVM (TSVM), for ETD by utilizing the labeled and unlabeled information. Likewise, authors in [38] design a deep generative model, termed as semi-supervised autoencoder (SSAE), which makes use of semi-supervised information for ETD. A similar case is presented in [39] where a mean teacher based semi-supervised mechanism is used for the identification of NTL. In [40], the advantages of semi-supervised EC data is obtained by utilizing the semi-supervised SVM for the detection of abnormal consumers. From the aforementioned methods, it is learned that the semi-supervised solutions combine the benefits of both labeled and unlabeled information to improve the ETD. However, these methods require extensive involvement of human experts for acceptable ETD performance by a model. In summary, the following problems are the focus of this paper.

1) The limited essence of labeled EC data and auxiliary data makes it impossible to treat ETD as a fully supervised problem. Moreover, the performance of conventional methods is examined using synthetic EC profiles that do not provide a realistic assessment of a model.
2) The unbalanced EC data of consumers lead to the model's biasness problem. The majority class hides the classification capabilities of the minority class because the model learns more EC patterns of normal consumers.
3) There exists noisy high-dimensional imbalanced data with associations between its features. During the feature extraction phase, the conventional methods fail to derive correlations between these features.
4) Existing methods require considerable involvement of human experts during feature extraction and model training phases for superior performance regarding ETD. It shows the need for an automated mechanism.

We propose a new semi-supervised solution for ETD, which is a combination of relational denoising autoencoder (RDAE) and attention guided (AG) TripleGAN, termed as RDAE-AG-TripleGAN. In the proposed solution, RDAE reduces both the noise and dimensionality of data. It also derives the relationships between data features. Afterwards, the AG-TripleGAN is applied where AG dynamically weights the extracted features and also guides the TripleGAN to pay more attention to highly weighted features by adjusting the model's parameters. In this context, the proposed solution takes the advantages of EC data in terms of both labeled and unlabeled forms. The proposed solution is different from the traditional approaches as follows.

- A deep semi-supervised mechanism is proposed, named as RDAE-AG-TripleGAN, which utilizes the important information present in unlabeled cases and labeled representations. Therefore, it improves the model's generalization ability along with the improvement in ETD.
- A deep RDAE is presented that reduces the noise and dimensionality of features from the data. It also captures the relationships between data features during the feature extraction phase. Thus, it helps to improve the model's performance for ETD by maintaining the presence of features' associations.
- An AG-TripleGAN is proposed where the AG mechanism assigns weights to features and also acts as a supervisor to dynamically supervise the TripleGAN.
- Extensive simulations are conducted over the real EC data that examine the proposed methodology's significance against existing schemes. The simulation results demonstrate that the proposed solution gains

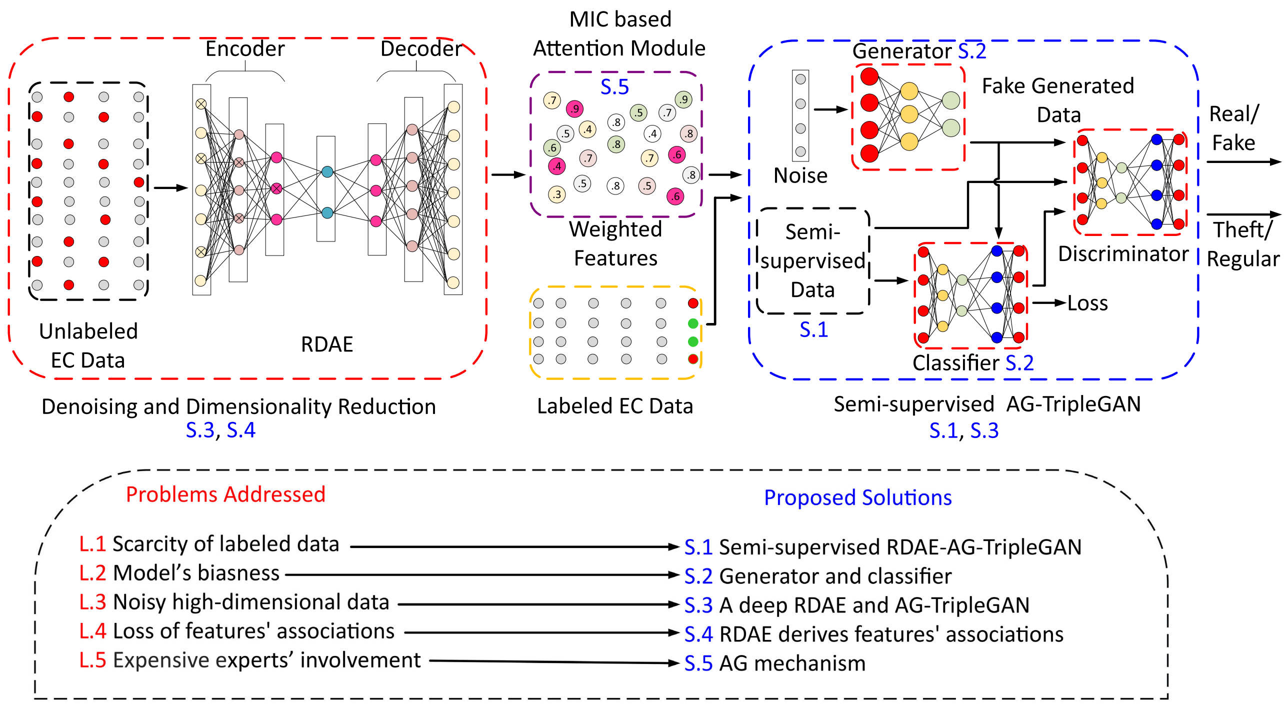


FIGURE 1. Overview of the proposed methodology.

superior results for ETD as compared to conventional approaches.

The remainder of this paper is as follows. Section II reports a detailed description of the proposed methodology. Whereas, the examination to show the significance of our proposed solution against other methods is presented by conducting simulations in Section III. Lastly, we conclude this paper in Section IV.

II. PROPOSED METHODOLOGY

This section describes the structure and components of the proposed methodology presented in this paper. The proposed methodology is constructed and evaluated using the real-time smart meters' data. It begins with the preprocessing of data. Afterwards, its subsequent modules are constructed to form a complete solution. The proposed methodology has three main stages: feature extraction, attention guidance and ETD, as shown in Fig. 1. It can be seen that the proposed semi-supervised RDAE-AG-TripleGAN solution addresses the limitations of conventional ETD schemes. The proposed solution uses semi-supervised data that solves the scarcity of labeled EC data, which is costly in real life due to expensive on-field inspections. Afterwards, the proposed solution employs deep RDAE for noise reduction and feature extraction. It also maintains the relationships between features during the feature extraction phase that fail to do the traditional feature extraction techniques. Furthermore, the proposed solution practices AG, which performs feature weighting and also acts as a guider to dynamically supervise the semi-supervised TripleGAN. The dynamic supervision of proposed TripleGAN via AG overcomes the expensive requirement of experts, which is required in conventional ETD schemes for better ETD results. The proposed RDAE-AG-TripleGAN for ETD is a deep learning based semi-supervised solution that efficiently deals with high-dimensional imbalanced data. Moreover, it also has the characteristic to generate fake samples and labels that solve the model's biasness problem due to imbalanced data. The following sub-sections describe the detailed description of each component.

A. DATA PREPROCESSING

The proposed solution is implemented on the smart meters data of daily electric power consumption of consumers, which is made available by the state grid corporation of China (SGCC) [41]. The dataset contains real on-field inspected honest and dishonest consumers. The dishonest consumers are considered as the ground truth in this work. The information about the dataset is provided in Table 1. The smart meters are present at consumers' end where they keep track of the power consumption of consumers and forward it to the electric utility over the communication network. This can lead to the presence of the missing and outliers/erroneous values in the dataset due to failure of meter components, memory issues or any other malfunctioning. Consequently, we need to apply the data preprocessing techniques to refine the dataset. The data preprocessing phase has three sub-tasks that handle the missing values, recover outliers and perform normalization. Therefore, the missing values are computed first using the concept of linear interpolation that uses the

TABLE 1. Information about the dataset.

Description	Numeric
Total duration	JAN 2014 - OCT 2016
Total consumers	42372
Electricity thieves	3615
Normal consumers	38757

following formula:

$$f(x_i) = \begin{cases} \frac{c}{2}, & x_i \in NaN,\ x_{i(t-1)}, x_{i(t+1)} \notin NaN, \\ 0, & x_i \in NaN,\ x_{i(t-1)} \text{ or } x_{i(t+1)} \in NaN, \\ x_i, & x_i \notin NaN, \end{cases} \quad (1)$$

where $c = x_{i(t-1)} + x_{i(t+1)}$. x_i is the current EC value at time t (i.e. a day). NaN indicates that x_i is not a numeric character. Besides, outliers are also present in data that negatively affect the model's performance. Therefore, we imply the notion of interquartile range [38], which is a statistical method to find and recover the erroneous instances, as given in Equation 2.

$$Z(x_i) = \begin{cases} m, & x_i > m, \\ 0, & x_i < 0, \end{cases} \quad (2)$$

where $m = Q_3(X) + [Q_3(X) - Q_1(X)] \times 3$. In addition, $Q_3(X)$ and $Q_1(X)$ denote the upper (75^{th}) and lower (25^{th}) quantiles for each consumer's profile (X), respectively. Afterwards, we normalize the data through min-max normalization because deep learning approaches are more susceptible to diverse data.

B. RELATIONAL DENOISING AUTOENCODER

This sub-section describes the noise and dimensionality reduction procedure by applying RDAE. In ETD, the first step is to obtain the most appropriate feature set due to the increasing size of EC data, i.e., high-dimensional data. Therefore, we utilize the idea of autoencoder due to its great success, which mainly consists of a neural network based unsupervised learning structure. A plain autoencoder's structure typically consists of an encoder and a decoder with the prime intention of dimensionality reduction. The reduced features cover important information of original high-dimensional data, which is obtained by minimizing the reconstruction loss [42]. In general, the encoder part maps the given input X over the hidden representation H, which is expressed as:

$$H = e(X) = se(WX + b_X), \quad (3)$$

where se is the encoder's activation function, W signifies the weights and b expresses the bias vector. Likewise, for l number of stacked layers, the encoder function is defined as:

$$H = e_l(\ldots e_2(e_1(X))). \quad (4)$$

Furthermore, the decoder part maps back the hidden representation H to obtain its reconstruction X', which is specified as:

$$X' = d(H) = sd(W'H + b_H), \quad (5)$$

where sd is the activation function of decoder. W' and b_H denote the weights and bias vector of decoder, respectively. For l number of stacked layers, the decoder function is stated as:

$$X' = d_l(\ldots d_2(d_1(X))). \quad (6)$$

whereas, a simple autoencoder achieves its objective by minimizing the reconstruction loss L, which is described as:

$$obj = \min_\theta L(X, X') = \min_\theta L(X, d(e(X))). \quad (7)$$

In this work, we exploit RDAE due to its great success, which is the type of relational autoencoder (RAE). It is an improved version of traditional deep autoencoders [42]. Therefore, we employ it to overcome the limitations of conventional autoencoders [32], [33], [36], [43], which do not maintain the relationships between features during the feature extraction mechanism. Consequently, this loss of features' associations affects the ETD results. A simple RAE achieves its purpose by minimizing the data reconstruction loss $L(X, X')$ and their associations' reconstruction loss $L(R(X), R(X'))$ [42], which is explained as:

$$obj = (1-\alpha)\min_\theta L(X, X') + \alpha \min_\theta L(R(X), R(X')), \quad (8)$$

where $R(X)$ and $R(X')$ capture associations among features in X and X', respectively. Particularly, θ indicates the three parameters of an autoencoder, including W, b_H and b_X, which have an influential impact on the minimization of L. Therefore, α is a scale variable that maintains the L of both features and their associations by controlling the W. Furthermore, RAE also provides an extension of denoising autoencoders (DAE), named as RDAE, which is employed in this paper for noise and dimensionality reduction, as shown in Fig. 1 (red dashed block). RDAE simply corrupts the given input X, called corrupted input $\hat{X}$ and also reconstructs X'' as the corrupted input. In this work, we corrupt X by adding the Gaussian noise, i.e., $\hat{X} \sim G(0, \sigma)$, where σ represents the standard deviation of X. Thus, the prime intention is to better learn and generate the feature representations along with the noise reduction to improve the ETD performance. The objective function of RDAE is computed by minimizing L between $\hat{X}$ and X'' together with their features' associations XX^T and corrupted features' associations $\hat{X}\hat{X}^T$ with rectifier function t [42], which is stated as:

$$obj = (1-\alpha)\min_\theta L(X, \hat{X}) + \alpha \min_\theta L({}_{Tt}XX^T, {}_{Tt}\hat{X}\hat{X}^T), \quad (9)$$

where the data associations are described according to their similarities like $R(X)$ demonstrates the multiplication of X and X^T. We employ RDAE for the first time in this paper for ETD. It extracts features and keeps their relationships. The simulation results prove the importance of features' associations, which are necessary to improve the ETD performance. RDAE significantly reduces noise, execution time, overfitting problem and also improves ETD results. The RDAE's structure used in this work is based on the stack of three hidden layers in each of its encoder and decoder sections. Moreover, it utilizes binary cross-entropy as a loss function along with Adam optimizer.

C. MIC BASED AG FOR DYNAMIC LEARNING

In this sub-section, we express the AG module, which applies MIC for features' weighting and model's guidance to focus on significant features, as shown in Fig. 1. More specifically, it takes the extracted features by RDAE and performs feature weighting according to the presence of associations and redundancy between them. Afterwards, the weighted features are served as unsupervised inputs to AG-TripleGAN. Where, AG acts as a supervisor to dynamically guide the TripleGAN, which keeps its focus on highly weighted features to better learn the complex representations. In particular, AG updates the TripleGAN's parameters according to the weights associated with features, which represent the feature importance. In this way, the proposed AG-TripleGAN efficiently learns the most difficult EC patterns of consumers, which results in accurate identification of electricity fraudsters. The proposed AG mechanism is different from other methods in the sense that it weights the extracted features in the range of 0 and 1 and also dynamically adapts TripleGAN's parameters. The weights near to 0 show no associations (independence) between features and near to 1 point the highest correlation between them.

This mechanism is the core part of the proposed solution that highly improves ETD performance, as validated in Section III-C. The proposed AG mechanism is based on the most recent and powerful approach, MIC, which uses the concept of information theory. It finds the degree of correlation and similarity between features and also weights them accordingly [28]. Traditionally, the Pearson correlation coefficient (PCC), mutual information (MI) and Spearman correlation coefficient (SCC) were employed to find the associations between features. However, PCC and SCC cannot find the complex nonlinear associations between data features. Whereas, MIC relies on the MI, which has the benefits of generality, equitability and better performance [29]. The mathematical illustration of MIC is expressed as:

$$MIC(v, z) = max(I(v, z)/log_2 min(a_v, a_z)), \tag{10}$$

where $I(v, z)$ is the MI between two random variables v and z. a_v and a_z indicate the number of bins where v and z are partitioned.

D. SEMI-SUPERVISED AG-TRIPLEGAN FOR CLASSIFICATION

Here we describe the last phase of the proposed methodology for ETD. A new semi-supervised learning procedure, AG-TripleGAN, is introduced to overcome the limitations of traditional techniques [16], [20], [36]–[38]. These techniques work on labeled EC data and auxiliary data, which is limited in the real world due to expensive on-field inspections. Therefore, we use a semi-supervised mechanism, TripleGAN, for ETD due to the rapid success of GAN [44]. The conventional GAN models are typically based on two neural networks, known as generator and discriminator [45]. Whereas, the TripleGAN comprises three neural networks with the incorporation of a classifier to enhance the performance of traditional GANs. Therefore, its architecture is mainly composed of generator G, discriminator D and classifier C, as shown in Fig. 1 (blue dashed block). It plays a three-player minimax game, in which the generator synthesizes fake instances that meet the statistics of the actual data to fool the discriminator. Likewise, the classifier takes the real EC data as input and generates fake labels. In this context, the fake generated samples and labels minimize the data imbalance problem that negatively affects the ETD. The authors in [43] solved this problem by applying a GAN based approach.

The discriminator appears as a classification technique that discriminates among the real and fake generated sample-label pairs as well as classifies the electricity fraudsters. TripleGAN plays a three-player minimax game in the sense that if the discriminator accurately identifies the fake generated samples and labels, then no parameters' refinement is required for the discriminator. Hence, the changes are made to the parameters of the generator and classifier to generate more tricky fake samples and labels. On the other hand, if the generator and classifier succeed to fool the discriminator, then no parameter updation is required for generator parameters. Thus, the parameter tuning is performed for the discriminator to enhance its discrimination power.

In this paper, the AG-TripleGAN for ETD is constructed using the structure formerly defined in [44]. In which, the dynamic adaptation of parameters with respect to the weighted features by AG is incorporated. Hence, we propose an AG-TripleGAN for ETD. It is served with the weighted unsupervised features along with the labeled EC profiles to take the advantages of both unlabeled and labeled information. It also learns the weights associated with each weighted feature and updates its network parameters dynamically according to these weights. Hence, the dynamic adaption of parameters enables the AG-TripleGAN to efficiently maximize the ETD's objective. Therefore, the proposed mechanism obtains more reliable results for theft detection as discussed in Section III-C.

III. SIMULATION RESULTS

In this section, the effectiveness of the proposed RDAE-AG-TripleGAN for detecting electricity frauds is analyzed. In addition, we demonstrate the efficiency and usefulness of the proposed solution as compared to the existing ETD techniques. The proposed solution is simulated over the real smart meters' data using the most well-known and powerful libraries of Python, identified as TensorFlow and Keras.

A. SIMULATIONS SETTING

The proposed methodology is trained and assessed using the smart meters dataset of SGCC on HP Intel core i5-2310M machine having 4GB RAM and 500GB hard-drive. The dataset contains the EC of consumers from JAN

2014 to OCT 2016. The information about the dataset is stated in Table 1. Moreover, SGCC has explicitly stated the ground truth of 3615 electricity thieves, validated through the on-field inspections. Furthermore, to get effective simulation results, this work begins with data preprocessing. In which, linear interpolation, interquartile range and min-max normalization are applied to preprocess the dataset, as mentioned in Section II-A. After that, the preprocessed dataset is distributed into three sets: training set, validation set and testing set. The training ratio is defined to be 80% and the remaining 20% is assigned to testing and validation sets, 10% each. Furthermore, each set contains the proportion of 90% unlabeled consumers while 10% of labeled consumers. Afterwards, RDAE is trained to perform denoising and feature extraction. Its structure consists of three fully connected layers in each of its encoder and decoder parts. In the encoder part of RDAE, the fully connected layers have neurons in an increasing order, i.e., 64, 128 and 256 while the layers of decoder have them in decreasing order, i.e., 256, 128 and 64. The structure of AG-TripleGAN is already defined in Section II-D. Furthermore, RDAE and AG-TripleGAN are trained by minimizing the binary cross-entropy as their cost functions using 60 training iterations (epochs). Specifically, we did not employ any special mechanism to tune the hyperparameters of the proposed solution. Rather, the proposed solution is trained and tested using the commonly used grid-search approach for finding the best hyperparameters' configuration where the proposed model efficiently minimizes the loss and improves the ETD results using real EC data.

B. PERFORMANCE INDEXES

Essentially, ETD is a highly complex anomaly detection mechanism due to imbalanced data of majority (regular) and minority (fraudsters) class. Therefore, there is a need to select the most suitable performance evaluation indexes to measure the performance of the proposed methodology. For classification, these performance indexes are obtained from the confusion matrix, which returns four possible outcomes, particularly, true positive (TP), true negative (TN), false positive (FP) and false negative (FN). We use six most useful and suitable performance metrics that are based on the outcomes of the confusion matrix. These metrics include area under the curve (AUC), precision, recall, Matthews correlation coefficient (MCC), F1-score and precision-recall area under the curve (PR-AUC) [23], [28], [46]. In particular, accuracy is the most common performance measure for classification problems. We do not consider accuracy as a performance measure in this work because the overall accuracy will always be high even if the proposed solution fails to identify electricity theft. The mathematical expressions of F1-score and MCC are described in Equations 11 and 12:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}, \tag{11}$$

$$MCC = \frac{C_1}{\sqrt{C_2 \times C_3 \times C_4 \times C_5}}, \tag{12}$$

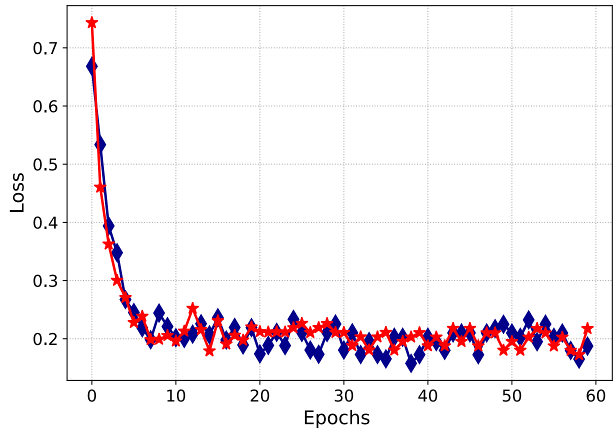


FIGURE 2. RDAE based loss analysis.

where $C_1 = (TN \times TP) - (FP \times FN)$, $C_2 = (FP + TP)$, $C_3 = (FN + TP)$, $C_4 = (FP + TN)$ and $C_5 = (FN + TN)$. Specifically, recall, also known as detection rate (DR) and true positive rate (TPR), is used to measure how perfectly the proposed solution detects the electricity fraudsters. Whereas, precision is applied to assess the performance of the proposed mechanism regarding FP. Moreover, F1-score derives the balance of precision and recall to obtain more realistic assessment of the proposed methodology. Likewise, AUC, also known as the area under the receiver operating characteristic curve (ROC-AUC), is a graphical description of TPR against the false positive rate (FPR) over the varying threshold. A higher ROC-AUC score, typically, more than 0.5, depicts better distinguishing capability of the model while less than 0.5 shows its inverse case. Furthermore, PR-AUC is obtained by representing the precision against recall score over the varying threshold. It shows that a perfect classifier for ETD is the one who which achieves accurate TP with a less number of FP and FN. It enables the model to appear at the top of other models. Lastly, MCC is the most suitable and appropriate performance measure. It finds the associations between four outcomes, i.e., TP, FP, TN and FN.

In summary, the prime intention of this work is to efficiently maximize DR and minimize FPR. The cost of FN is pretty high and significant because it shows the price of electricity used that is not paid by dishonest consumers. Whereas, the cost of FP is much lower than FN because it shows the cost of on-field inspections against the cost of energy stolen. Therefore, in ETD, more attention is given to recall than precision.

C. PROPOSED SOLUTION EVALUATION RESULTS

This sub-section interprets the performance evaluation results of the proposed RDAE-AG-TripleGAN methodology for ETD using the above-mentioned performance indicators. The proposed method begins with the process of denoising and extraction of features using RDAE, as shown in Figs. 2 and 3. The values of these plots depict the loss and accuracy of RDAE verses epochs during training. Fig. 2 shows the

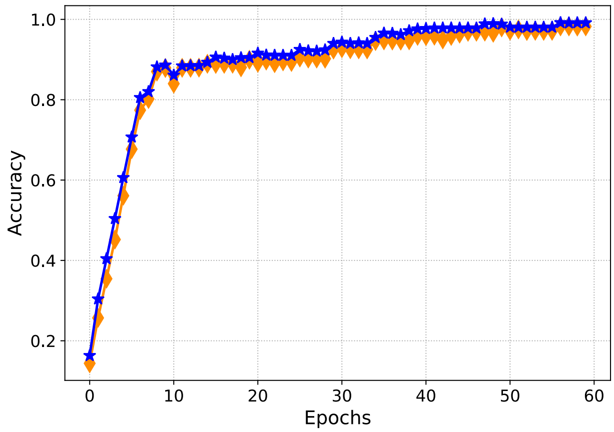


FIGURE 3. RDAE based accuracy analysis.

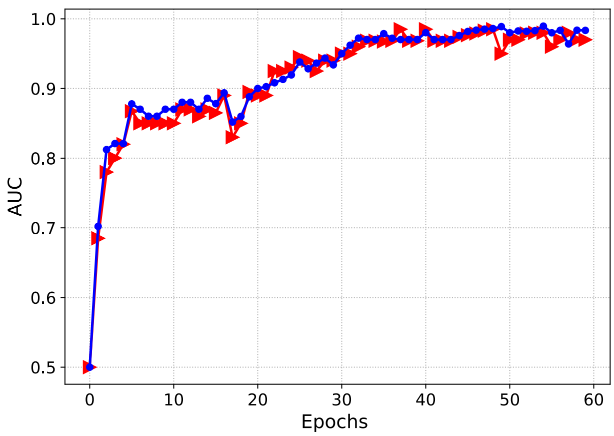


FIGURE 4. AUC based analysis of Proposed solution.

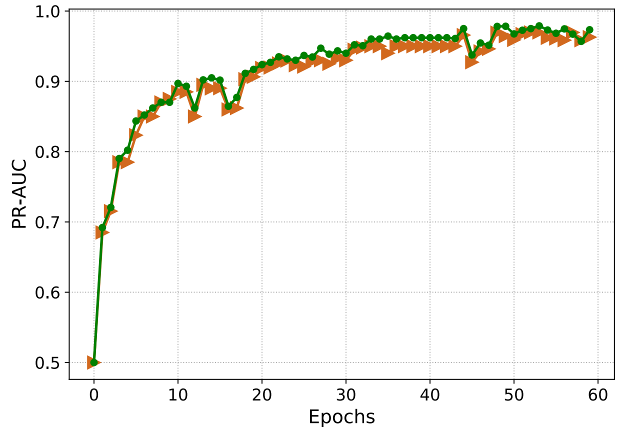


FIGURE 5. PR-AUC based analysis of Proposed solution.

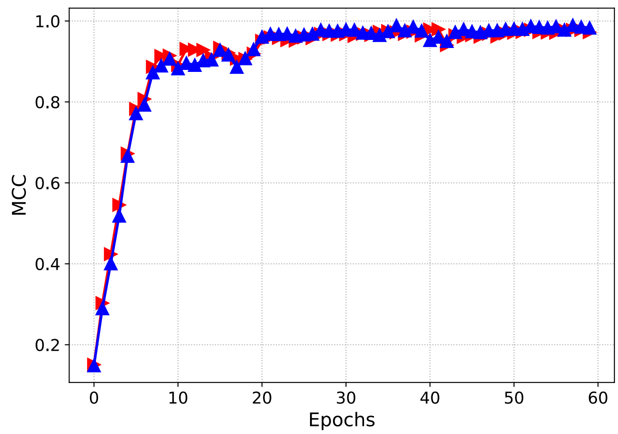


FIGURE 6. MCC based analysis of Proposed solution.

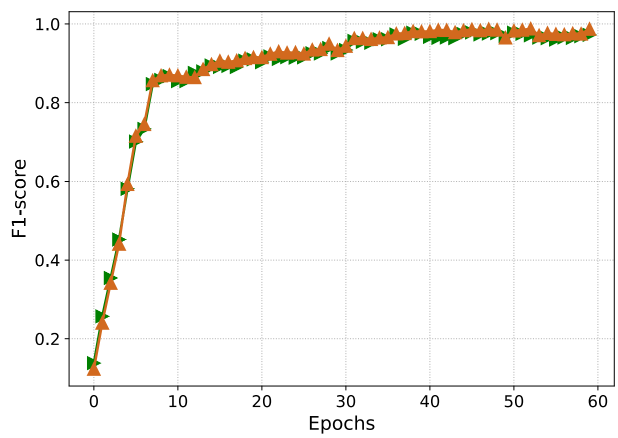


FIGURE 7. F1-score based analysis of Proposed solution.

convergence of loss during the training and validation phases of RDAE. It demonstrates that RDAE consistently learns the power consumption patterns throughout feature extraction and uniformly minimizes the loss. In Fig. 3, we demonstrate the analysis of the proposed RDAE in terms of accuracy. It efficiently performs feature extraction and also derives features' associations. Afterwards, the weights are assigned to the extracted features by the AG and served as the unlabeled feature representations to AG-TripleGAN. On the other hand, the limited amount of labeled EC cases are also supplied as input to the AG-TripleGAN to train in a semi-supervised fashion. Figs. 4 and 5 depict AUC and PR-AUC performance analysis of the proposed solution on training and validation sets verses training iterations, respectively. As it is seen that the proposed semi-supervised solution efficiently improves AUC and PR-AUC score throughout the training phase. This improvement depicts the dynamic detection capability of the proposed solution.

Likewise, Figs. 6 and 7 demonstrate the proposed solution's performance regarding MCC and F1-score, respectively. It is evident that the proposed model efficiently obtains excellent results for ROC-AUC, PR-AUC, MCC and F1-score, i.e., closer to 0.98, as the number of training iterations increase. These results clarify that the proposed solution efficiently captures irregularities from the dataset by taking the advantages of both labeled and unlabeled EC profiles. Furthermore, these results define the significance of the proposed RDAE and AG mechanism for high-dimensional imbalanced data.

RDAE-AG-TripleGAN creates a contest (minimax game) for its generator, classifier and discriminator parts,

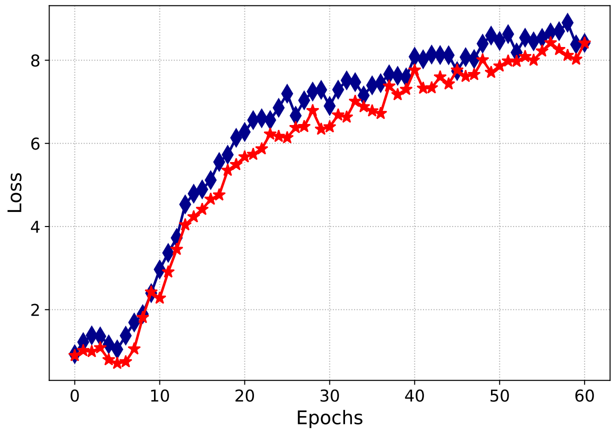


FIGURE 8. Generator based loss analysis.

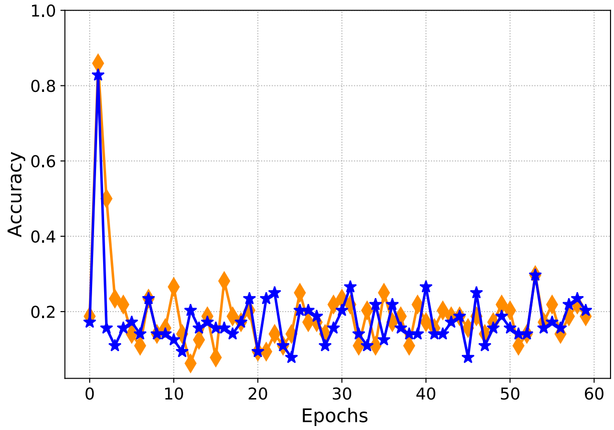


FIGURE 9. Generator based accuracy analysis.

as described in Section II-D. In this context, Figs. 8 and 9 show the loss and accuracy analysis of generator during the training phase, respectively, using training and validation sets. As it is evident in Fig. 8 that the loss of the generator increases with the increase in epochs. It synthesizes the fake samples to fool the discriminator. Therefore, Fig. 8 shows the remarkable discrimination capability of the discriminator to efficiently detect the fake samples, which increases the loss of generator sub-model. Similarly, Fig. 9 depicts the accuracy analysis of generator sub-model, which also demonstrates the excellent capability of the discriminator. In a similar manner, Figs. 10 and 11 analyze the performance of the classifier for generating fake labels to mislead the discriminator. It is evident that the discriminator accurately predicts the fake labels generated by the classifier. Afterwards, Figs. 12 and 13 show the loss and accuracy of the discriminator, respectively, while distinguishing the real and fake generated examples and labels. It is clear that the discriminator smoothly reduces the loss up to 0.08, together with the improvement in accuracy covering 0.99 over the training and validation sets. Therefore, the proposed semi-supervised solution gives remarkable performance for ETD during the training process.

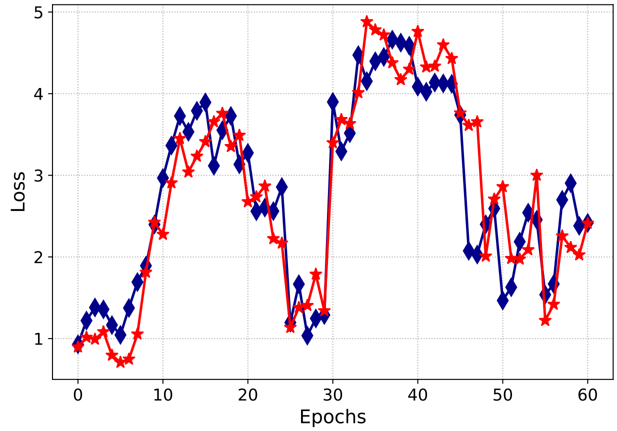


FIGURE 10. Classifier based loss analysis.

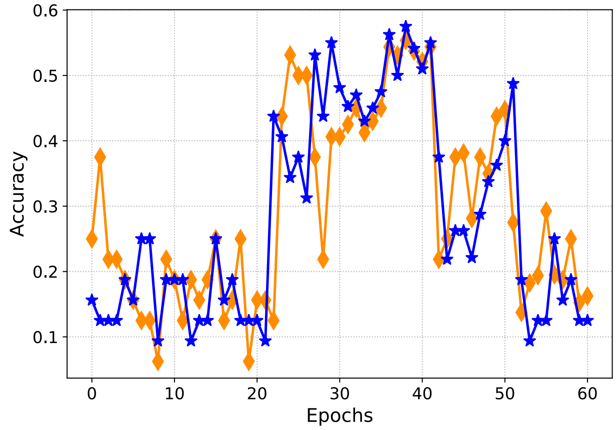


FIGURE 11. Classifier based accuracy analysis.

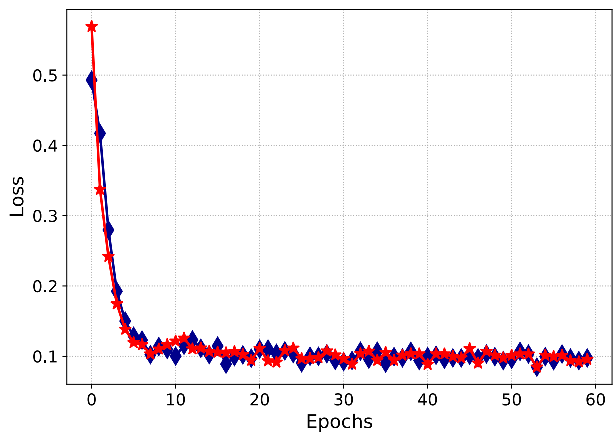


FIGURE 12. Discriminator based loss analysis.

In Figs. 14 and 15, we illustrate the graphical representation of ROC-AUC and PR-AUC of the proposed semi-supervised mechanism for ETD. In particular, Fig. 14 shows that the proposed solution achieves excellent scores in terms of ROC-AUC using training, validation and testing sets. Similarly, Fig. 15 depicts that it also obtains reliable results for PR-AUC performance measure using the same sets.

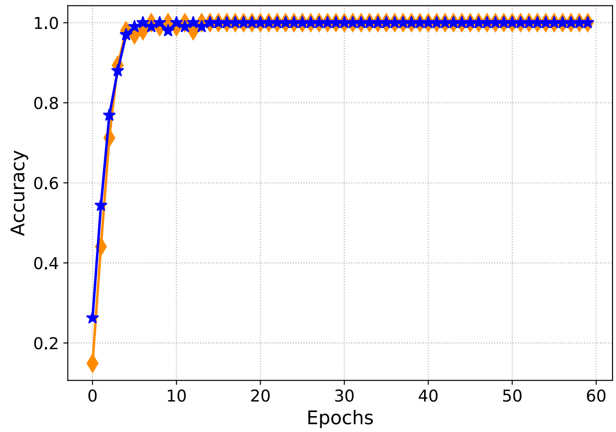


FIGURE 13. Discriminator based accuracy analysis.

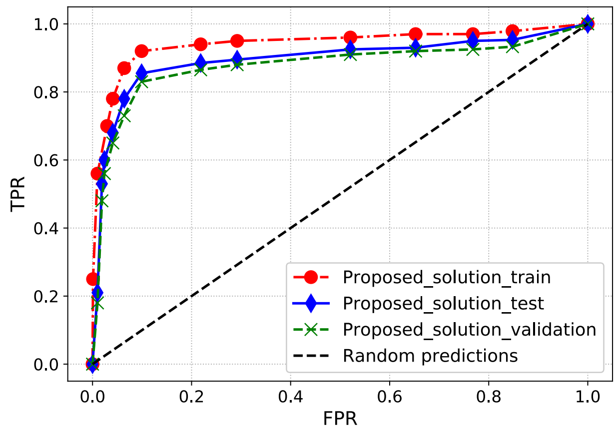


FIGURE 14. ROC-AUC based analysis of proposed solution.

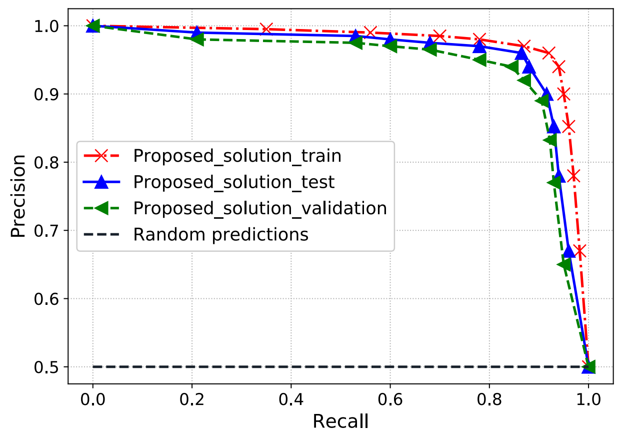


FIGURE 15. PR-AUC based analysis of proposed solution.

Furthermore, Fig. 16 justifies the competency of the proposed semi-supervised mechanism while demonstrating the effects of both labeled and unlabeled EC instances using F1-score, MCC and AUC. It depicts that with the increasing number of labeled cases, the proposed scheme uniformly uplifts its ETD results. Moreover, even with a few numbers of labeled instances, i.e., 500 and 1000, the proposed solution also proves to be perfection for ETD. Furthermore,

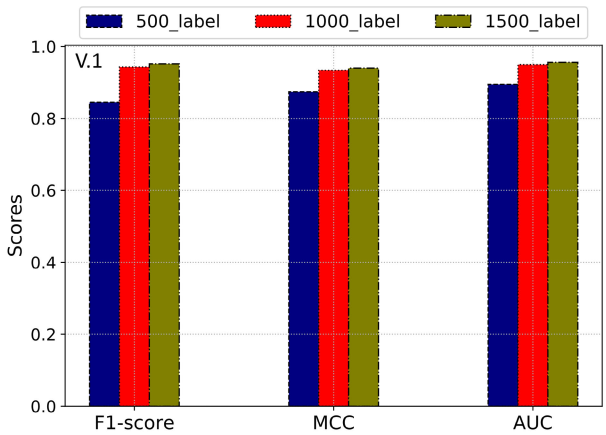


FIGURE 16. Effects of labeled and unlabeled data.

TABLE 2. Optimal hyperparameters' selection for TSVM and XGBoost using grid-search.

Hyperparameter	Range of values	Selected value
TSVM_C	0.1, 0.01, 0.001, 10	0.01
TSVM_γ	1, 1.5, 6, 10	10
XGBoost_Estimators	30, 50, 100, 200	100
XGBoost_Depth_size	3, 4, 5, 7	5
XGBoost_Learning_rate	0.1, 0.01, 0.001	0.001

using 1500 labeled EC histories make the proposed semi-supervised mechanism to touch its peak values.

D. BENCHMARK MODELS

This sub-section describes the performance comparison of the proposed solution with conventional approaches for ETD. It validates the uniqueness, effectiveness and appropriateness of the proposed RDAE-AG-TripleGAN methodology for real practices over traditional techniques. Following are the fundamental and conventional approaches for ETD.

1) TSVM

SVM is a well-known classifier used in many works [47], [48], for ETD. It creates a boundary between classes for class separation. The authors in [37] design a semi-supervised SVM, named as TSVM, to make use of both labeled and unlabeled information. Therefore, we consider it as a baseline. In [49], authors present the significance of hyperparameters' tuning in the smart grid environment. Therefore, the optimal hyperparameters' selection of TSVM is defined in Table 2.

2) XGBoost

It is a supervised learning approach that is widely used in recent works, [20], [22], for abnormality detection in smart grids. It employs the concept of gradient boosting decision trees in which a strong learner is produced by applying the ensemble of weak learners. In [20], the authors employ XGBoost over the real-time smart meters data together with the auxiliary databases to enhance the ETD performance.

Furthermore, the hyperparameters of XGBoost play an important role to optimize the detection performance. Therefore, their optimal values are selected using grid-search, as given in Table 2.

3) LSTM-MLP

Another approach, LSTM-MLP, is presented in [24] for NTL detection. The authors train LSTM over sequential EC histories to derive long-term associations between them while MLP is employed to capture non-sequential information from the auxiliary information. This work considers LSTM-MLP using the same hyperparameters' setting, which is given in [24].

4) CNN-LSTM-SAE

A new semi-supervised solution, CNN-LSTM-SAE, is proposed in [36] for NTL detection. It practices the concept of transfer learning to utilize the knowledge of labeled and unlabeled EC instances. Therefore, we use it as a baseline scheme.

5) SSAE

A semi-supervised learning based deep autoencoder, termed as SSAE, is proposed in [38] for the identification of NTL. This work considers it as a benchmark model for the performance comparison.

6) AG-TripleGAN

In order to show the importance of features' relationships during the feature extraction process, we remove RDAE module of the proposed methodology and consider it as a variant. Now, the AG-TripleGAN's training process begins with features' weighting mechanism where weights are assigned to the original unlabeled EC data and forwarded to the AG-TripleGAN for ETD.

7) RDAE-TripleGAN

It is considered as another variant of the proposed RDAE-AG-TripleGAN. In which, we simply remove the AG module to show its significance for better ETD results. RDAE-TripleGAN begins with the noise reduction and feature extraction phase. Afterwards, the extracted features and labeled samples are passed as inputs to TripleGAN for ETD.

Table 3 contains the summarized performance comparison results of the proposed solution and other conventional benchmark schemes using the SGCC dataset. It is clear that the proposed solution gains better results over existing schemes in terms of AUC, PR-AUC, precision, recall, MCC and F1-score. More specifically, the proposed model obtains 0.987, 0.956, 0.952, 0.958, 0.967 and 0.943 for precision, recall, AUC, PR-AUC, F1-score and MCC over the given dataset, respectively. Whereas, the DR for TSVM, SSAE, CNN-LSTM-SAE, LSTM-MLP and XGBoost is quite similar to each other, i.e., 0.657, 0.801, 0.824, 0.832 and 0.802, respectively. The proposed semi-supervised mechanism achieves the highest DR among all other conventional

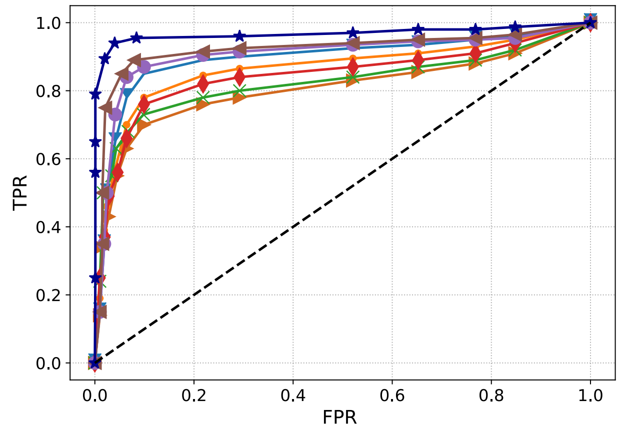


FIGURE 17. SGCC dataset based ROC-AUC comparison.

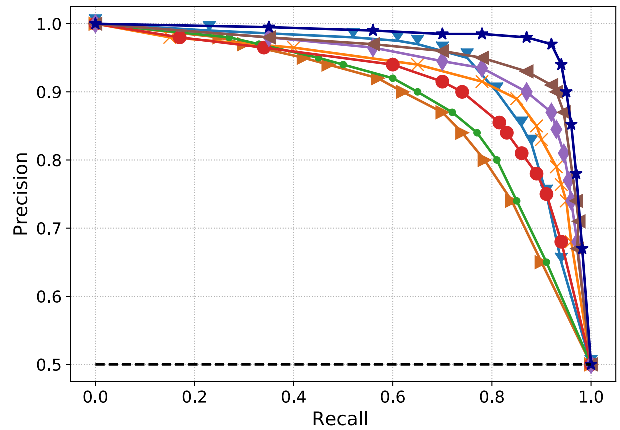


FIGURE 18. SGCC dataset based PR-AUC comparison.

techniques, i.e., 0.956. Consequently, it efficiently optimizes the ETD results. Moreover, the two variants, RDAE-TripleGAN and AG-TripleGAN, of the proposed methodology have also achieved better results than the traditional TSVM, SSAE, XGBoost, LSTM-MLP and CNN-LSTM-SAE models. Although, RDAE-TripleGAN and AG-TripleGAN have inferior performance than the proposed RDAE-AG-TripleGAN. It is seen from Table 3 that RDAE-TripleGAN gives values of 0.956, 0.914, 0.901, 0.896, 0.924 and 0.908 for precision, recall, AUC, PR-AUC, F1-score and MCC, respectively. These results are less than the proposed model's results, which prove that without using the AG module, the proposed solution requires attention for the complex representations. This also proves the significance of the proposed AG mechanism. Furthermore, AG-TripleGAN has values of 0.904, 0.895, 0.859, 0.886, 0.896 and 0.885 for precision, recall, AUC, PR-AUC, F1-score and MCC, respectively, without using the RDAE module. These results demonstrate that it is important to perform feature extraction and derive features' relationships.

Furthermore, Figs. 17 and 18 show the performance comparison of the proposed solution as compared to the conventional schemes using the SGCC dataset. It is clear

TABLE 3. Proposed model performance comparison.

Model	Precision	Recall	AUC	PR-AUC	F1-score	MCC
TSVM	0.715	0.657	0.706	0.763	0.687	0.646
SSAE	0.864	0.801	0.846	0.813	0.836	0.825
XGBoost	0.838	0.802	0.759	0.756	0.812	0.774
LSTM-MLP	0.846	0.832	0.822	0.808	0.829	0.791
CNN-LSTM-SAE	0.856	0.824	0.813	0.805	0.826	0.802
RDAE-TripleGAN	0.956	0.914	0.901	0.896	0.924	0.908
AG-TripleGAN	0.904	0.895	0.859	0.886	0.896	0.885
RDAE-AG-TripleGAN	**0.987**	**0.956**	**0.952**	**0.958**	**0.967**	**0.943**

that the proposed methodology achieves more superior results for ROC-AUC and PR-AUC performance measures against benchmark schemes using the real-time smart meters' dataset. These results depict the significance of semi-supervised data and the dynamic attention of the model.

E. PERFORMANCE ANALYSIS USING IRISH SMART METER DATASET

We have also evaluated the performance of the proposed semi-supervised mechanism using the Irish smart energy trails (ISET) dataset [50], which contains data of 5000 consumers during 2009 and 2010. It is a real-time smart meters' dataset that contains EC records with a frequency of 30 minute intervals. Moreover, it contains EC data of only regular consumers. Therefore, for ETD, we need to inject electricity frauds in the dataset where we first transform the half hourly data into hourly data, i.e., $x_t = \{x_1, \ldots, x_{24}\}$. Where, t denotes the time interval ranging from 1 to 24. Next, we employ the following six types of malicious attacks given in [4].

- $A_1(x_t) = rx_t,\ r = random(0.1, 0.8),$
- $A_2(x_t) = b_t x_t, \qquad b_t = \begin{cases} 0 & for\ t \\ 1 & else \end{cases}$
- $A_3(x_t) = yx_t,\ y = random(0.1, 0.8),$
- $A_4(x_t) = mean(x),$
- $A_5(x_t) = ymean(x),$
- $A_6(x_t) = x_{24-t}.$

In the first malicious attack, i.e., $A_1(x_t)$, the dishonest consumers return the malicious EC values by multiplying the original EC readings with some random number. In the second attack, i.e., $A_2(x_t)$, the electricity fraudsters either return or do not return the EC readings of a particular day. Likewise, in the third malicious attack, i.e., $A_3(x_t)$, the dishonest consumers multiply each EC reading with a unique random number. $A_4(x_t)$ and $A_5(x_t)$ attacks describe the actual mean of a particular day and a mean value multiplied with a random number, respectively. In the sixth attack, i.e., $A_6(x_t)$, the meter readings are reversed by the theft consumers. Therefore, by injecting these attacks in the ISET dataset, we generate 30% malicious data for the period of one year. Now, the ISET dataset contains EC records of both honest and fraudulent consumers. Afterwards, the dataset is partitioned into three sets: training, validation and testing, following the same proportion of the SGCC dataset. In the ISET dataset, we randomly designate 70% as unlabeled consumers and 30% as labeled consumers. In addition, similar data

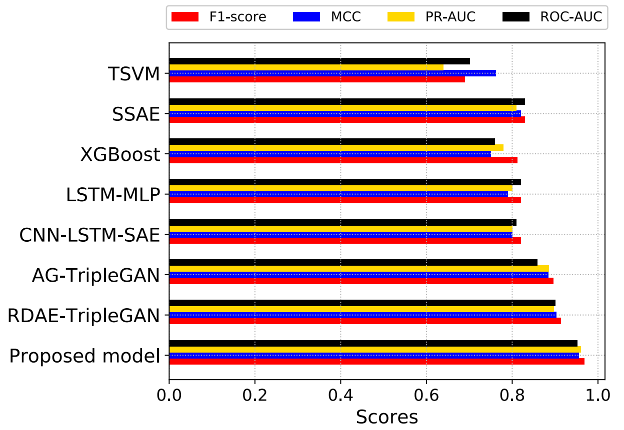


FIGURE 19. AUC, PR-AUC, F1-score and MCC based performance comparison on SGCC dataset.

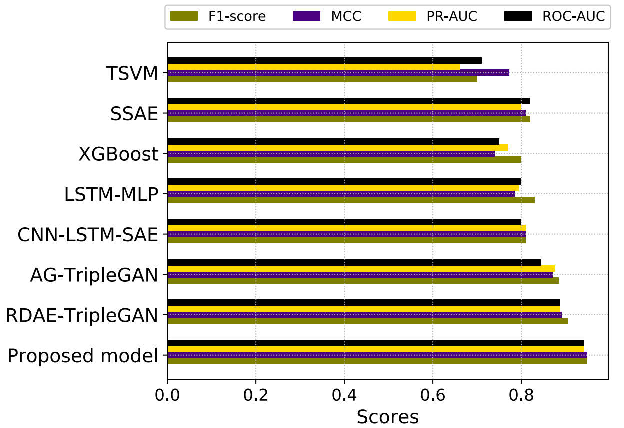


FIGURE 20. AUC, PR-AUC, F1-score and MCC based performance comparison on ISET dataset.

preprocessing steps are used for the ISET dataset as for the SGCC dataset (mentioned in Section II-A). Fig. 19 shows the performance comparison of the proposed semi-supervised solution with conventional approaches in terms of F1-score, MCC, PR-AUC and AUC over the SGCC dataset. Likewise, Fig. 20 displays the performance comparison in terms of F1-score, MCC, PR-AUC and AUC using the ISET dataset. It is clear in the figures that the proposed solution gives excellent results using both datasets. In the ISET dataset, we inject synthetic theft cases, which depicts the malfunctioning with meters' readings. Therefore, it is quite obvious that the proposed solution is also effective for cyber-attacks (data attacks) where electricity fraudsters change the actual energy consumption readings with fake ones. RDAE-AG-

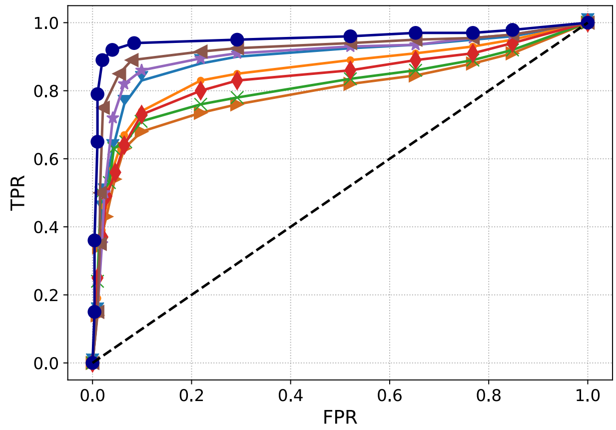


FIGURE 21. ISET dataset based ROC-AUC comparison.

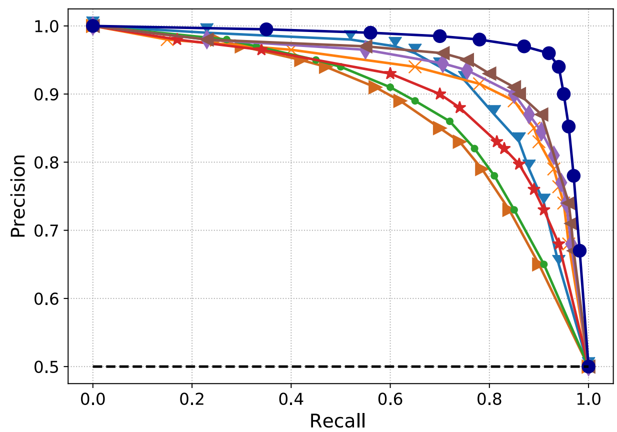


FIGURE 22. ISET dataset based PR-AUC comparison.

TripleGAN achieves excellent results using both real and synthetic data. This makes it acceptable for power distribution utilities to detect electricity fraudsters. Furthermore, Figs. 21 and 22 depict the performance comparison of the proposed solution with the standard benchmark schemes using the ISET dataset in terms of ROC-AUC and PR-AUC. In Fig. 21, it is seen that the proposed methodology achieves superior results than other techniques using the dataset with synthetic theft cases. Moreover, it is depicted that in terms of ROC-AUC, the proposed solution obtains excellent results for ETD where it demonstrates the highest TPR and less misclassification rate, i.e., FPR, against standard benchmark techniques. Besides, Fig. 22 shows the performance of the proposed methodology in terms of PR-AUC as compared to the conventional schemes for ETD. Where, the proposed solution obtained the highest values for precision and recall. These results demonstrate the excellence of the proposed solution for both honest and dishonest consumers using synthetic dataset. In particular, the proposed RDAE-AG-TripleGAN shows better performance for real on-field inspected SGCC dataset. The obtained results of the proposed solution are effective for ETD and reducing the cost of on-field inspections. Furthermore, in order to evaluate the scalability of the proposed semi-supervised RDAE-AG-TripleGAN solution, we conduct extensive simulations using two different datasets, i.e., SGCC and ISET. The former dataset consists of 42372 consumers' data whereas, the latter comprises data of 5000 consumers. Moreover, SGCC dataset covers data from Jan 2014 to Oct 2016 while ISET consists of data of only 2 years, i.e., 2009 and 2010. The simulation results depict that the proposed solution efficiently performs ETD using both datasets with a slight difference in the results obtained for different performance metrics. It implies that the proposed model is scalable and has the potential to handle both low-dimensional and high-dimensional data.

However, there is a trade-off between scalability and computational cost. To reduce the computational cost of the proposed solution, we employ RDAE that efficiently performs dimensionality reduction. The conventional approaches for ETD, such as TSVM and XGBoost, have high computational cost for high-dimensional data and poor ETD performance. Moreover, they require hand-crafted feature extraction to obtain the potential features and reduce computational time. Therefore, the proposed solution utilizes deep learning techniques, i.e., RDAE and TripleGAN, to automatically extract the potential features and reduce the computational time. Furthermore, the proposed solution gives more accurate and reliable results as compared to conventional machine learning techniques for ETD, such as TSVM and XGBoost. However, the proposed solution still inhibits high computation cost, given in terms of computational time complexity, i.e., 1.5h. It is due to the hardware constraints and not using the graphical processing unit. If the proposed solution would have been implemented using graphical processing unit, the computational time would have been reduced. However, in ETD, the computational time is not as crucial as FPR, i.e., misclassification rate. The primary objective of this work is to obtain more accurate ETD performance as compared to existing benchmark schemes. The simulation results proved that the proposed solution achieved the aforementioned objective efficiently.

In order to summarize this paper, we present the mapping between identified problems, proposed solutions and their validation in Table 4. The problems identified in traditional ETD approaches are mapped over the proposed solution of this paper. Afterwards, the validation is discussed to validate the problems against the proposed solutions. The proposed solution solves the scarcity of labeled EC data as this paper presents a semi-supervised solution that requires less number of labeled samples, as validated in Fig. 16. Furthermore, it is a deep learning based solution that effectively deals with the high-dimensional EC data, as shown in Figs. 2-7. Traditional feature extraction techniques fail to derive features' associations, as discussed in Section II-B. Therefore, we employ RDAE, which is a deep feature extraction technique to deal with the reduction of noise and dimensionality of features. Moreover, it also captures the features' associations during feature extraction and has the potential to efficiently deal with high-dimensional data, as illustrated in Figs. 2 and 3. The issue of model's biasness that occurs due to the imbalance between electricity honest and dishonest consumers is effi-

TABLE 4. Mapping of the problems identified and suggested solutions.

Problems Identified	Proposed Solution	Validation Results
L.1: Limited number of labeled EC cases	S.1: A semi-supervised RDAE-AG-TripleGAN	V.1: RDAE-AG-TripleGAN requires few labeled samples, as shown in Fig.16
L.2: Loss of features' associations during feature extraction	S.2: RDAE efficiently derives features' associations	V.2: Proposed RDAE efficiently captures features' correlations, as shown in Figs. 2 and 3
L.3: Noisy high-dimensional imbalanced EC data	S.3: A deep RDAE-AG-TripleGAN handles noisy and high-dimensional data	V.3: RDAE-AG-TripleGAN smoothly reduces dimensionality and noise in Figs. 2-7
L.4: Model's biasness due to imbalanced data	S.4: AG-TripleGAN's generator and classifier minimize data imbalance concerns	V.4: AG-TripleGAN's generator and classifier generate fake samples-labels to handle data imbalance concerns, as displayed in Figs. 8-13
L.5: Expensive experts' engagement needed for better ETD	S.5: An automated AG based supervisor	V.5: Proposed AG mechanism dynamically supervises the AG-TripleGAN, as shown in Figs. 12 and 13

ciently solved by the AG-TripleGAN's generator and classifier sub-models. These sub-models have the capability to generate more plausible fake samples-labels to fool the AG-TripleGAN's discriminator. Thus, the synthetic generation of sample-label pairs to avoid the model's biasness problem. It is seen in Figs. 8-13 that the generator and classifier generate fake sample-label pairs and the discriminator accurately identifies the real and fake instances. Furthermore, the traditional approaches for ETD require costly experts' engagement during the identification of electricity frauds, as discussed in Section I (3). To avoid this problem, AG is designed that dynamically supervises the proposed solution and improves ETD, as validated in Figs. 12 and 13. Consequently, this mapping of problem, solution and validation proves that the proposed semi-supervised RDAE-AG-TripleGAN solves the limitations of traditional ETD models in a systematic way. Moreover, the performance analysis of the proposed solution shows that it learns useful information from both labeled and unlabeled EC records. The above-mentioned simulation results prove that the proposed solution is efficient and has better ETD results than conventional approaches, which makes it more effective and appropriate for electric utilities.

IV. CONCLUSION

This paper proposes a semi-supervised mechanism, RDAE-AG-TripleGAN, for ETD in smart grids. It considers the most significant EC information in both labeled and unlabeled instances. RDAE is used for noise reduction and feature extraction together with maintaining features' relationship. The maintenance of the relationship plays a very important role in ETD. Afterwards, AG acts as a supervisor and assigns weights to the features extracted and forwarded by RDAE. It dynamically guides the AG-TripleGAN to pay more attention to the highly weighted features by adjusting its parameters. This mechanism significantly improves the generalization ability of the proposed solution along with its ETD capability. Furthermore, we conduct extensive simulations using the real EC dataset to prove the effectiveness of the proposed solution against conventional schemes. The simulation results validate that the proposed methodology achieves better results for ETD as compared to conventional schemes. The proposed solution obtains a DR of 0.956, which is a greater value than the conventional schemes. Whereas, the DR for TSVM, SSAE, XGBoost, LSTM-MLP, CNN-LSTM-SAE, RDAE-TripleGAN and AG-TripleGAN is 0.657, 0.801, 0.802, 0.832, 0.824, 0.914 and 0.895, respectively. Moreover, simulation results prove that RDAE-AG-TripleGAN achieves excellent performance gains for ETD over the real on-field inspected and synthetically injected theft cases. Therefore, simulations are performed for both the datasets, which ensure the proposed model's scalability. Hence, this makes RDAE-AG-TripleGAN a more practical and acceptable solution for power distribution companies. Furthermore, our prime intention in this paper is to achieve better and reliable ETD results than the conventional schemes using EC records. Therefore, in the future, we will take into account privacy preservation and auxiliary features, such as temperature and number of appliances. Moreover, we also intend to consider the heavy-tailed non-Gaussian noises like Laplace noise and Cauchy noise in future to evaluate the performance of our proposed model and further improve it.

REFERENCES

[1] P. Antmann, "Reducing technical and non-technical losses in the power sector," Background Paper WBG Energy Strategy, World Bank, Washington, DC, USA, Tech. Rep., 2009. [Online]. Available: https://openknowledge.worldbank.org/handle/10986/20786

[2] J. A. Meira, P. Glauner, R. State, P. Valtchev, L. Dolberg, F. Bettinger, and D. Duarte, "Distilling provider-independent data for general detection of non-technical losses," in *Proc. IEEE Power Energy Conf. Illinois (PECI)*, Champaign, IL, USA, Feb. 2017, pp. 1–5.

[3] K. M. Ghori, R. A. Abbasi, M. Awais, M. Imran, A. Ullah, and L. Szathmary, "Performance analysis of different types of machine learning classifiers for non-technical loss detection," *IEEE Access*, vol. 8, pp. 16033–16048, 2020.

[4] P. Jokar, N. Arianpoo, and V. C. M. Leung, "Electricity theft detection in AMI using Customers' consumption patterns," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 216–226, Jan. 2016.

[5] (2012). *Energy Theft: Incentives to Change*. Accessed: Apr. 2019. [Online]. Available: https://documents.worldbank.org/curated/en/829751468326689826/

[6] Complete Blood Count. *Electricity Theft by B.C. Grow-OPS Costs 100M a Year*. Accessed: Jan. 2019. [Online]. Available: http://www.cbc.ca/news/canada/british-columbia/story/2010/10/08/bc-hydro-grow-optheftw.html

[7] S. Mujeeb and N. Javaid, "ESAENARX and DE-RELM: Novel schemes for big data predictive analytics of electricity load and price," *Sustain. Cities Soc.*, vol. 51, Nov. 2019, Art. no. 101642, doi: 10.1016/j.scs.2019.101642.

[8] O. Samuel, N. Javaid, A. Khalid, W. Z. Khan, M. Y. Aalsalem, M. K. Afzal, and B.-S. Kim, "Towards real-time energy management of multi-microgrid using a deep convolution neural network and cooperative game approach," *IEEE Access*, vol. 8, pp. 161377–161395, 2020.

[9] M. Nabil, M. Ismail, M. M. E. A. Mahmoud, W. Alasmary, and E. Serpedin, "PPETD: Privacy-preserving electricity theft detection scheme with load monitoring and billing for AMI networks," *IEEE Access*, vol. 7, pp. 96334–96348, 2019.

[10] D. Yao, M. Wen, X. Liang, Z. Fu, K. Zhang, and B. Yang, "Energy theft detection with energy privacy preservation in the smart grid," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 7659–7669, Oct. 2019.

[11] J. B. Leite and J. R. S. Mantovani, "Detecting and locating non-technical losses in modern distribution networks," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 1023–1032, Mar. 2018.

[12] Q. Louw and P. Bokoro, "An alternative technique for the detection and mitigation of electricity theft in South Africa," *SAIEE Afr. Res. J.*, vol. 110, no. 4, pp. 209–216, Dec. 2019.

[13] S. McLaughlin, B. Holbert, A. Fawaz, R. Berthier, and S. Zonouz, "A multi-sensor energy theft detection framework for advanced metering infrastructures," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7, pp. 1319–1330, Jul. 2013.

[14] S. Amin, G. A. Schwartz, A. A. Cardenas, and S. S. Sastry, "Game-theoretic models of electricity theft detection in smart utility networks: Providing new capabilities with advanced metering infrastructure," *IEEE Control Syst.*, vol. 35, no. 1, pp. 66–81, Feb. 2015.

[15] C.-H. Lin, S.-J. Chen, C.-L. Kuo, and J.-L. Chen, "Non-cooperative game model applied to an advanced metering infrastructure for non-technical loss screening in micro-distribution systems," *IEEE Trans. Smart Grid*, vol. 5, no. 5, pp. 2468–2469, Sep. 2014.

[16] 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 Trans. Ind. Informat.*, vol. 14, no. 4, pp. 1606–1615, Apr. 2018.

[17] M. N. Hasan, R. N. Toma, A.-A. Nahid, M. M. M. Islam, and J.-M. Kim, "Electricity theft detection in smart grid systems: A CNN-LSTM based approach," *Energies*, vol. 12, no. 17, p. 3310, Aug. 2019, doi: 10.3390/en12173310.

[18] M. Adil, N. Javaid, U. Qasim, I. Ullah, M. Shafiq, and J.-G. Choi, "LSTM and bat-based RUSBoost approach for electricity theft detection," *Appl. Sci.*, vol. 10, no. 12, p. 4378, Jun. 2020, doi: 10.3390/app10124378.

[19] S. Li, Y. Han, X. Yao, S. Yingchen, J. Wang, and Q. Zhao, "Electricity theft detection in power grids with deep learning and random forests," *J. Electr. Comput. Eng.*, vol. 2019, pp. 1–12, Oct. 2019, doi: 10.1155/2019/4136874.

[20] 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 Trans. Smart Grid*, vol. 10, no. 3, pp. 2661–2670, May 2019.

[21] Z. A. Khan, M. Adil, N. Javaid, M. N. Saqib, M. Shafiq, and J.-G. Choi, "Electricity theft detection using supervised learning techniques on smart meter data," *Sustainability*, vol. 12, no. 19, p. 8023, Sep. 2020, doi: 10.3390/su12198023.

[22] R. Punmiya and S. Choe, "Energy theft detection using gradient boosting theft detector with feature engineering-based preprocessing," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 2326–2329, Mar. 2019.

[23] N. F. Avila, G. Figueroa, and C.-C. Chu, "NTL detection in electric distribution systems using the maximal overlap discrete wavelet-packet transform and random undersampling boosting," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 7171–7180, Nov. 2018.

[24] M.-M. Buzau, J. Tejedor-Aguilera, P. Cruz-Romero, and A. Gomez-Exposito, "Hybrid deep neural networks for detection of non-technical losses in electricity smart meters," *IEEE Trans. Power Syst.*, vol. 35, no. 2, pp. 1254–1263, Mar. 2020.

[25] H. Gul, N. Javaid, I. Ullah, A. M. Qamar, M. K. Afzal, and G. P. Joshi, "Detection of non-technical losses using SOSTLink and bidirectional gated recurrent unit to secure smart meters," *Appl. Sci.*, vol. 10, no. 9, p. 3151, Apr. 2020, doi: 10.3390/app10093151.

[26] Z. Qu, H. Li, Y. Wang, J. Zhang, A. Abu-Siada, and Y. Yao, "Detection of electricity theft behavior based on improved synthetic minority oversampling technique and random forest classifier," *Energies*, vol. 13, no. 8, p. 2039, Apr. 2020, doi: 10.3390/en13082039.

[27] M. Salman Saeed, M. W. Mustafa, U. U. Sheikh, T. A. Jumani, I. Khan, S. Atawneh, and N. N. Hamadneh, "An efficient boosted C5.0 Decision-Tree-Based classification approach for detecting non-technical losses in power utilities," *Energies*, vol. 13, no. 12, p. 3242, Jun. 2020, doi: 10.3390/en13123242.

[28] K. Zheng, Q. Chen, Y. Wang, C. Kang, and Q. Xia, "A novel combined data-driven approach for electricity theft detection," *IEEE Trans. Ind. Informat.*, vol. 15, no. 3, pp. 1809–1819, Mar. 2019.

[29] W. Zhang, X. Dong, H. Li, J. Xu, and D. Wang, "Unsupervised detection of abnormal electricity consumption behavior based on feature engineering," *IEEE Access*, vol. 8, pp. 55483–55500, 2020.

[30] J. L. Viegas, P. R. Esteves, and S. M. Vieira, "Clustering-based novelty detection for identification of non-technical losses," *Int. J. Electr. Power Energy Syst.*, vol. 101, pp. 301–310, Oct. 2018.

[31] A. Maamar and K. Benahmed, "A hybrid model for anomalies detection in AMI system combining K-means clustering and deep neural network," *Comput., Mater. Continua*, vol. 60, no. 1, pp. 15–39, 2019.

[32] Y. Zhang, Q. Ai, H. Wang, Z. Li, and X. Zhou, "Energy theft detection in an edge data center using threshold-based abnormality detector," *Int. J. Electr. Power Energy Syst.*, vol. 121, Oct. 2020, Art. no. 106162, doi: 10.1016/j.ijepes.2020.106162.

[33] C. Fan, F. Xiao, Y. Zhao, and J. Wang, "Analytical investigation of autoencoder-based methods for unsupervised anomaly detection in building energy data," *Appl. Energy*, vol. 211, pp. 1123–1135, Feb. 2018.

[34] M. A. de Souza, J. L. R. Pereira, G. D. O. Alves, B. C. de Oliveira, I. D. Melo, and P. A. N. Garcia, "Detection and identification of energy theft in advanced metering infrastructures," *Electr. Power Syst. Res.*, vol. 182, May 2020, Art. no. 106258, doi: 10.1016/j.epsr.2020.106258.

[35] T. Hu, Q. Guo, X. Shen, H. Sun, R. Wu, and H. Xi, "Utilizing unlabeled data to detect electricity fraud in AMI: A semisupervised deep learning approach," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 11, pp. 3287–3299, Nov. 2019.

[36] R. R. Bhat, R. D. Trevizan, R. Sengupta, X. Li, and A. Bretas, "Identifying nontechnical power loss via spatial and temporal deep learning," in *Proc. 15th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Anaheim, CA, USA, Dec. 2016, pp. 272–279.

[37] J. Tacon, D. Melgarejo, F. Rodriguez, F. Lecumberry, and A. Fernandez, "Semisupervised approach to non technical losses detection," in *Proc. Iberoamerican Congr. Pattern Recognit.*, 2014, pp. 698–705.

[38] X. Lu, Y. Zhou, Z. Wang, Y. Yi, L. Feng, and F. Wang, "Knowledge embedded semi-supervised deep learning for detecting non-technical losses in the smart grid," *Energies*, vol. 12, no. 18, p. 3452, Sep. 2019, doi: 10.3390/en12183452.

[39] J. Li and F. Wang, "Non-technical loss detection in power grids with statistical profile images based on semi-supervised learning," *Sensors*, vol. 20, no. 1, p. 236, Dec. 2019, doi: 10.3390/s20010236.

[40] X. Wang, I. Yang, and S.-H. Ahn, "Sample efficient home power anomaly detection in real time using semi-supervised learning," *IEEE Access*, vol. 7, pp. 139712–139725, 2019.

[41] *State Grid Corporation of China Electricity Theft Dataset*. Accessed: Jun. 2020. [Online]. Available: https://www.sgcc.com.cn/

[42] Q. Meng, D. Catchpoole, D. Skillicom, and P. J. Kennedy, "Relational autoencoder for feature extraction," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Juneau, AK, USA, May 2017, pp. 364–371.

[43] H. Liu, Z. Li, and Y. Li, "Noise reduction power stealing detection model based on self-balanced data set," *Energies*, vol. 13, no. 7, p. 1763, Apr. 2020, doi: 10.3390/en13071763.

[44] L. Chongxuan, T. Xu, J. Zhu, and B. Zhang, "Triple generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 4088–4098.

[45] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.

[46] A. Luque, A. Carrasco, A. Martín, and A. de las Heras, "The impact of class imbalance in classification performance metrics based on the binary confusion matrix," *Pattern Recognit.*, vol. 91, pp. 216–231, Jul. 2019.

[47] J. Nagi, K. S. Yap, S. K. Tiong, S. K. Ahmed, and M. Mohamad, "Nontechnical loss detection for metered customers in power utility using support vector machines," *IEEE Trans. Power Del.*, vol. 25, no. 2, pp. 1162–1171, Apr. 2010.

[48] J. Nagi, K. Siah Yap, S. Kiong Tiong, S. K. Ahmed, and F. Nagi, "Improving SVM-based nontechnical loss detection in power utility using the fuzzy inference system," *IEEE Trans. Power Del.*, vol. 26, no. 2, pp. 1284–1285, Apr. 2011.

[49] R. Khalid and N. Javaid, "A survey on hyperparameters optimization algorithms of forecasting models in smart grid," *Sustain. Cities Soc.*, vol. 61, Oct. 2020, Art. no. 102275, doi: 10.1016/j.scs.2020.102275.

[50] *Irish Smart Meter Dataset Irish Social Science Data Archive*. Accessed: Jun. 2020. [Online]. Available: https://www.ucd.ie/issda/data/commissionforenergyregulationcer/

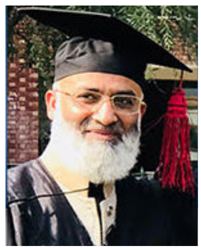
ZEESHAN ASLAM received the master's degree, under the supervision of Dr. N. Javaid, from the Communications Over Sensors (ComSens) Research Laboratory, Department of Computer Science, COMSATS University Islamabad, Islamabad Campus. His research interests include data science, optimization, security and privacy, energy trading, and smart grid.

FAHAD AHMED is currently pursuing the M.S. degree from COMSATS University Islamabad, Islamabad Campus. His research interests include data science, machine learning, energy trading, and smart grid.

AHMAD ALMOGREN (Senior Member, IEEE) received the Ph.D. degree in computer science from Southern Methodist University, Dallas, TX, USA, in 2002. He is currently a Professor with the Computer Science Department, College of Computer and Information Sciences (CCIS), King Saud University (KSU), Riyadh, Saudi Arabia, where he is currently the Director of the Cyber Security Chair, CCIS. Previously, he worked as the Vice Dean of the Development and Quality at CCIS. He also served as the Dean for the College of Computer and Information Sciences and the Head of the Academic Accreditation Council, Al Yamamah University. He served as the General Chair for the IEEE Smart World Symposium and a Technical Program Committee member of numerous international conferences/workshops, such as IEEE CCNC, ACM BodyNets, and IEEE HPCC. His research interests include mobile-pervasive computing and cyber security.

MUHAMMAD SHAFIQ received the master's degree in information technology (IT) from the University of the Punjab, Gujranwala, Pakistan, in 2006, the M.S. degree in computer science from the University Institute of Information Technology, Arid Agriculture University, Rawalpindi, Pakistan, in 2010, and the Ph.D. degree in information and communication engineering from Yeungnam University, South Korea, in February 2018. He was with the Faculty of Computing and IT, University of Gujrat, Gujrat City, Pakistan, as a Faculty Member, from 2010 to 2014, and formerly held the same position with the Department of Computer Science and IT, Federal Urdu University, Islamabad, Pakistan. His research interests include the Internet of Things (IoT), cognitive radio-based IoT networks-architecture and design, mobile ad hoc networks, wireless sensor networks, performance, management, and security, 5G cellular networks, admission control, and mobility management, device-to-device communications, medium access control protocols, Internet routing protocols, spectrum trading and auctions, information systems, design, and access control, and human–computer interaction.

MANSOUR ZUAIR received the B.S. degree in computer engineering from King Saud University and the M.S. and Ph.D. degrees in computer engineering from Syracuse University. He has served as the Chairman of CEN, from 2003 to 2006, the Vice Dean, from 2009 to 2015, and has been the Dean, since 2016. He is currently an Associate Professor with the Department of Computer Engineering, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia. His research interests include computer architecture, computer networks, and signal processing.



NADEEM JAVAID (Senior Member, IEEE) received the bachelor's degree in computer science from Gomal University, Dera Ismail Khan, Pakistan, in 1995, the master's degree in electronics from Quaid-i-Azam University, Islamabad, Pakistan, in 1999, and the Ph.D. degree in computer science from the University of Paris-Est, France, in 2010. He is currently an Associate Professor and the Founding Director of the Communications Over Sensors (ComSens) Research Laboratory, Department of Computer Science, COMSATS University Islamabad, Islamabad Campus. He has supervised 126 master's and 20 Ph.D. theses. He has authored over 900 articles in technical journals and international conferences. His research interests include energy optimization in smart grids and in wireless sensor networks using data analytics and blockchain. He was a recipient of the Best University Teacher Award from the Higher Education Commission of Pakistan, in 2016, and the Research Productivity Award from the Pakistan Council for Science and Technology, in 2017. He is also an Associate Editor of IEEE Access and an Editor of the *International Journal of Space-Based and Situated Computing* and *Sustainable Cities and Society*.