IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 15 September 2022, accepted 30 September 2022, date of publication 19 October 2022, date of current version 2 November 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3215532

RESEARCH ARTICLE

RFE Based Feature Selection and KNNOR Based Data Balancing for Electricity Theft Detection Using BiLSTM-LogitBoost Stacking Ensemble Model

PAMIR[1], NADEEM JAVAID[1,2], (Senior Member, IEEE), AHMAD ALMOGREN[3], (Senior Member, IEEE), MUHAMMAD ADIL[4], MUHAMMAD UMAR JAVED[1], (Graduate Student Member, IEEE), AND MANSOUR ZUAIR[5]

[1]Department of Computer Science, COMSATS University Islamabad, Islamabad 44000, Pakistan
[2]School of Computer Science, University of Technology Sydney, Ultimo, NSW 2007, Australia
[3]Department of Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh 11633, Saudi Arabia
[4]Department of Electrical and Computer Engineering, COMSATS University Islamabad, Islamabad 44000, Pakistan
[5]Department of Computer Engineering, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia

Corresponding authors: Ahmad Almogren (ahalmogren@ksu.edu.sa) and Nadeem Javaid (nadeemjavaidqau@gmail.com)

This work was supported by the King Saud University, Riyadh, Saudi Arabia, through Researchers Supporting under Project RSP-2022/184.

ABSTRACT Obtaining outstanding electricity theft detection (ETD) performance in the realm of advanced metering infrastructure (AMI) and smart grids (SGs) is quite difficult due to various issues. The issues include limited availability of theft data as compared to benign data, neglecting dimensionality reduction, usage of the standalone (single) electricity theft detectors, etc. These issues lead the classification techniques to low accuracy, minimum precision, low F1 score, and overfitting problems. For these reasons, it is extremely crucial to design such a novel strategy that is capable to tackle these issues and yield outstanding ETD performance. In this article, electricity theft happening in SGs is detected using a novel ETD approach. The proposed approach comprises recursive feature elimination (RFE), k nearest neighbor oversampling (KNNOR), bidirectional long short term memory (BiLSTM), and logit boosting (LogitBoost) techniques. Furthermore, three BiLSTM networks and a LogitBoost model are combined to make a BiLSTM-LogitBoost stacking ensemble model. Data preprocessing and feature selection followed by data balancing and electricity theft classification are the four major stages of the model proposed for ETD. It is obvious from the simulations performed using state grid corporation of China (SGCC)'s electricity consumption (EC) data that our proposed model achieves 96.32% precision, 94.33% F1 score, and 89.45% accuracy, which are higher than all the benchmarks employed in this study.

INDEX TERMS K nearest neighbor oversampling approach, bidirectional long short term memory, Logit-Boosting, deep learning, machine learning, stacking ensemble model, electricity theft detection, smart grid.

I. INTRODUCTION

Once electricity is produced by electricity generation plants, it is transmitted to consumers in a two hop process, i.e., transferred using high voltage transmission lines to the substations and then using low voltage distribution lines from substation to consumers. It is a reality that the amount of electric power generated at generation side and received at the consumer end are never the same. It means that the energy generated by the generators is always greater than the energy received by the consumers [1]. This mismatch in energy leads to imbalance between supply and demand. Different works have been

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

performed to tackle this imbalance like load scheduling, load forecasting, etc., [2], [3]. The deficiency in energy is unavoidable when energy flows using power transmission lines, conductors, transformers, and substation equipment [4]. The gap between the quantities of energy generated by generators and energy received by consumers is called distribution and transmission loss [5]. The categorization of the loss is further made into two types of losses: non technical and technical [6]. Technical losses (TLs) are unavoidable because they happen due to the electric energy dissipation in conductors, components employed for power transmission lines, transformers, etc., [4], [7]. Whereas, non technical losses (NTLs) can be avoided because they happen due to faulty meters, meter installation error, error in meter parameterization, and electricity theft [1], [8]. Electricity theft is a major reason of NTLs that dissipates a great amount of electricity. According to [9], 80% of the NTLs are caused due to electricity theft.

The term electricity theft is referred to as the illegal usage of electric energy in the unavailability of any contract or manipulation of the meter's data to either minimize or not pay the consumed electricity's bills [1]. It is commonly performed by electricity thieves using meter's bypassing, tampering, direct tapping, and putting of magnetic component inside the meter to slow down its process of measuring the energy consumption (EC). Electricity theft leads to a substantial amount of revenue losses in developed as well as developing nations of the world. On the whole, the world suffers a loss of more than $96 billion yearly due to the electricity theft [10]. According to [11], in 2015, Russia lost $5.1 billion, Brazil lost $10.5 billion, and India lost $16.2 billion due to electricity theft.

Currently, huge amount of studies are available in the literature for detecting electricity theft. Traditionally, electricity theft detection (ETD) was performed by physically checking the bypassed transmission cables, comparing the data of benign and theft consumers' meters, in person checking for problematic meter installation, etc., [12], which are highly inefficient, tedious, time-consuming, laborious, and costly tasks.

With the emergence of smart grids (SGs), the electric utilities find new weapons to fight electricity theft. An SG is the integration of information and communication technology (ICT) with conventional power grid [13]. With the combination of ICT, SGs find the quality to perform bidirectional communication between utility and consumer using advanced metering infrastructure (AMI). In this way, smart meters (SMs) are being integrated in AMI to collect EC data, electricity price information, grid's status information, etc., [12]. This data is useful for researchers to develop novel schemes for electricity price prediction [14], demand response management [15], energy scheduling for demand side management [16], [17], and demand side load forecasting [18]. In addition, recently, authors in [19], [20], [21], [22], [23], [24], and [25], prove that conducting data analysis task in SGs assist in detecting energy theft. However, following are some limitations in these ETD approaches.

1) Some of the approaches use a single and standalone machine learning (ML) or deep learning (DL) scheme for ETD, therefore, they obtain low ETD accuracy values.
2) Dimensionality reduction is ignored that results in low ETD performance.
3) Class imbalanced problem is not considered, which leads the model to generate false results for minority class and overfitting.
4) Some approaches employ random undersampling (RUS) to deal with data imbalanced problem. It randomly removes the majority class elements and leads the model to loss of important data, which further results in underfitting problem.

Therefore, in this research article, we focus on designing a novel ETD approach to efficiently address the above mentioned problems. We propose a BiLSTM-LogitBoost ensemble model, which is the combination of bidirectional long short term memory (BiLSTM) network and logit boosting (LogitBoost) models to pinpoint the electricity theft in SGs. This model combines the benefits of both ML and DL models that consequently gives better performance results in ETD. The novelty of the proposed work is as follows. BiLSTM-LogitBoost model is designed and applied for ETD in SGs. Moreover, k nearest neighbor oversampling (KNNOR) technique is applied in ETD domain to balance the EC data. Furthermore, the major contributions towards the research community made in the underlying work are enlisted below.

- The electricity theft happening in the SGs is detected using a novel BiLSTM-LogitBoost stacking ensemble model. The model consists of multiple BiLSTM networks and a LogitBoost model.
- The data imbalanced problem is tackled by KNNOR approach.
- Dimensionality reduction is performed using RFE; a feature selection (FS) method. It effectively selects those features that are much relevant in prediction of the output.
- Dropout layer is added to the proposed model. In our case, overfitting happens due to the imbalanced data based classification and model's complexity. Therefore, KNNOR and Dropout regularization methods are employed to balance the data by minority class' oversampling and probabilistically turning off some of the nodes (neurons) of the network, respectively. Turning off some neurons probabilistically results in model's complexity reduction, which consequently prevents the model from overfitting. Furthermore, balancing the data using KNNOR also avoids overfitting problem because when a model is trained using balanced data, it no more generates biased (one sided) results towards the majority class as models generate when training using imbalanced class data.
- Extensive simulations are performed on a huge real time EC dataset. Simulation results report that our

BiLSTM-LogitBoost stacking ensemble model followed by KNNOR and RFE gives better ETD performance than existing models.

The remaining sections of the article are presented in the following sequence. Section II provides the related work while Section III highlights the problem statement. Proposed system model for ETD, simulation results, and conclusion are presented in Section IV, Section V, and Section VI, respectively.

II. RELATED WORK

Research in the field of ETD in SGs has recently become a very hot topic for the researchers to leverage various techniques for performing ETD. The classification of these techniques can be done into three major classes: hardware based class, data mining based class, and hybrid (hardware and data mining) based class.

Hardware based class of ETD [26], [27], [28], [29] need some specialized hardware components, such as sensors, microcontrollers, and circuits to resolve ETD in SGs. These techniques are made to detect energy theft caused due to the SMs' and distribution power lines' physical interference. In this way, they are unable to detect cyber attacks on energy SMs. Cyber attack on energy SMs is a type of energy theft in which the EC readings is altered by SMs' hacking [11]. For example, in [26], hardware class based method is employed to detect electricity theft. The authors use temperature sensors for NTL estimation and detection. Moreover, in [27], energy theft is detected using transformer overloading. Once detection of overloading at transformer is done, a relay circuit is leveraged to turn off the transformer. In this way, electricity theft is discouraged. However, besides the incapability of these methods in cyber attacks' detection, they are also costly because of deployment and maintenance of the specialized hardware components in these methods.

Hybrid (hardware and data mining) based class techniques for ETD [19], [30], [31] use both data mining (ML and DL) and hardware related methods to detect NTLs. For instance, in [19], the authors employ support vector machine (SVM) for ETD in SGs. The distribution transformer meters' are also leveraged to identify the areas with high probability of electricity theft, and pinpointing the doubtful consumers by noticing anomalies in EC patterns. However, the authors use a standalone ML classifier for ETD that is why maximum FPR value is obtained. Moreover, the authors in [30], work on detection of new and complex type of electricity theft, called colluded NTL in which several electricity thieves collaborate to perform electricity fraud. An observer meter is deployed to maintain EC readings of multiple households located in a community. To ensure the safety of observer meter from physical and cyber attacks, the interference-resistant and surveillance camera devices are employed, respectively. In addition, a mathematical model is designed to use data from SMs and an observer meter to detect the SMs with tampered EC data. However, due to the requirement of special hardware components in the hybrid category, extra cost is incurred for the deployment and maintenance of hardware devices.

As previous both categories of ETD techniques are highly-expensive, therefore, many researchers have moved towards the third category that is data mining based ETD techniques [11], [20], [21], [22], [23], [32], [33], [34], [35], [36], [37], and [38], to handle the problem of electricity theft in SGs. Authors in [11] propose a deep model for electricity theft and non-theft consumers' classification in SGs. The problems of not a number (NaN) values and imbalanced data are tackled using interpolation and synthetic data points generation methods, respectively. Principal component analysis technique is used for dimensionality reduction. In addition, bayesian optimization is leveraged to enhance the ETD performance by tuning the model's hyperparameters. Moreover, adaptive moment estimation (Adam) optimizer is used to optimize model's parameters to obtain better results. The SGCC dataset is used in this study for analysis. Performance parameters used in the study are AUC, accuracy, Matthews correlation coefficient (MCC), F1 score, precision, and recall. Moreover, these parameters are calculated in time, frequency, and hybrid domains. However, bayesian optimizer is employed for hyperparameters' optimization, which requires more computational time to generate candidate solutions. Besides, in [20], the authors propose a deep stacked autoencoder with long short term memory (LSTM) based structure for detection of anomaly in EC readings. The deep autoencoder is considered to help in recognizing the complex patterns in data while LSTM is considered to capture long sequences of the long term time series EC data. Additionally, a sequential grid search optimizer is leveraged for hyperparameters' optimization. The proposed and benchmark models' training and testing are performed using Irish EC dataset. It is done to detect ETD in SGs. However, stochastic features' generation is neglected. Furthermore, an LSTM based DL model is designd in [21] to correctly classify electricity theft consumers. Consumption data from real SGCC dataset is used for analysis purpose. However, single and standalone techniques are used. Hence, low accuracy and high FPR are obtained.

In [22], authors develop a combined DL model to detect NTL in SGs. Hybrid DL model consists of multi layer perceptron (MLP) and LSTM models. LSTM is employed for the analysis of EC data while MLP is leveraged to use and analyse the non-sequential information. The proposed hybrid model is trained and tested using Spain's Endesa dataset to detect NTLs in SGs. However, the problem of data imbalance is ignored. In [23], a bagging ensemble random forest (RF) is employed for detecting electricity theft. Moreover, a stacked autoencoder is leveraged to extract important features in order to improve ETD performance of the RF classifier. Irish and Chinese EC datsets are used for theft analysis. Moreover, the data is balanced using RUS. However, RUS leads the proposed classifier to the problem of underfitting. Furthermore, the authors in [32] propose a deep model with low

FPR, known as (LFPR-DNN). Gradient descent optimizer is leveraged to update model's weights. Besides, focal loss function is used to reduce the influence of imbalanced data on ETD performance. Two-phase training of the proposed LFPR-DNN is considered. In the first phase, grid search is leveraged for hyperparameters' optimization. Whereas, in the second phase, the particle swarm optimization is used for tuning. I addition, the Irish EC dataset is used for analysis purpose to detect electricity theft. However, grid search is used that needs extremely high time complexity for hyperparameters' tuning.

Furthermore, the authors in [33]develop an efficient electricity theft detector wherein an autoencoder is employed for dimensionality reduction while bidirectional gated recurrent unit (BiGRU) is leveraged for final classification of electricity theft and normal consumers. Moreover, BiGRU is employed to recognize the drift patterns by its learning ability of long-time temporal dependency. Furthermore, the dropout regularization technique is leveraged to deal with the proposed model's overfitting issue. SGCC dataset is considered for the proposed model's training and testing. Moreover, SGCC data is balanced applying six synthetic theft attacks on honest data. Besides, in [34], an ensemble gradient boosting electricity theft detector is proposed. The proposed detector consists of categorical, extreme, and light gradient boosting techniques. Irish smart EC data is used for electricity theft analysis. Stochastic features are generated to enhance theft detection rate and FPR values of aforementioned three theft detectors. Moreover, feature extraction is also done using weighted feature importance function to decrease time and space complexities of the theft detectors' training. However, the fine tuning of hyperparameters relevant to the proposed boosting ensemble theft detectors is neglected.

A boosting ensemble (CatBoost) with feature engineering is employed in [35] for ETD in SGs. The SGCC data is used for the training and testing of the technique proposed for ETD. besides, k nearest neighbor (KNN) interpolation method is used for filling the missing data. Furthermore, the imbalance between data is tackled using a hybrid over and undersampling technique (SMOTE-Tomek). A feature extraction and scalable hypothesis method is leveraged to obtain reduced dimensions from the dataset. However, fine tuning of hyperparameters is ignored that leads a classifier to the local optimal stagnation and low ETD performance. Futhermore, in [36], an ensemble based DL electricity theft detector is developed. The proposed model is the combination of multiple DL models and the outputs of the multiple DL models are passed to the majority voting classifier to calculate the final classification result. The DL model used in the proposed model is gated recurrent unit (GRU). The EC samples from residential energy disaggregation dataset (REDD) are employed for theft analysis in SGs. Furthermore, in [37], fourteen classification techniques (including stacking and maximum voting ensembles) are employed with six data balancing (DB) techniques to finalize the best combination that detects NTLs in SGs. Finally, a combined synthetic minority oversampling technique and edited nearest neighbor (SMOTE-ENN) technique in combination with the stacking classifier is considered to be the best performer for ETD in SGs. Moreover, the EC data from SGCC dataset is leveraged for analysis purpose. Moving ahead, authors in [38] propose theft attacks, LSTM, and GRU (TLGRU) model for ETD. TLGRU consists of synthetic six theft attacks based DB, LSTM based feature extraction, and GRU based classification. Furthermore, proposed model's training and testing is done via the SGCC daily dataset. Proposed TLGRU is compared with the benchmark schemes in which the proposed model proves to be the best performer. However, fine tuning of the hyperparameters of the proposed TLGRU model is neglected. Besides, Table 1 presents an overview of the existing literature in terms of achievements, data analysed, performance measures, and limitations [38]. There are some other articles that use data mining techniques. However, they do not use them for ETD but for cyber attack and intrusion detection to secure internet of things networks, such as [50], [51], [52], and [53]. In [50], the authors use RF, extreme gradient boosting (XGBoost), and KNN techniques for cyber attacks' detection. Whereas, in [51], the authors employ a distributed stacking ensemble model using fog computing for intrusion detection that combines XGBoost, KNN, and gaussian naive bayes techniques at level-1. At level-2, the results of the first level classifiers are passed to random forest that acts as a second level learner to perform final classification. In [52], the authors employ convolutional neural network (CNN) for anomaly based intrusion detection to protect internet of things networks. A deep learning (DL) scheme, namely sparse evolutionary training-multi layer perceptron, is used for multiple cyber security attacks' detection in [53] for industrial internet of things.

III. PROBLEM STATEMENT

The authors in [19] and [21] employed SVM and LSTM classifiers for detecting NTL in SGs, respectively. However, these are standalone ML and DL classifiers; therefore, they achieved low ETD performance. Furthermore, dimensionality reduction is ignored in both [19] and [21] that results in curse of dimensionality issue. The issue consequently maximizes the model's complexity and minimizes theft detection accuracy. Furthermore, the authors in [22] proposed a hybrid DL model for ETD. It performs well in terms of ETD. However, the data imbalanced problem is ignored that leads the classifier to biasness (skewness) towards the majority class data (in our case, majority class is normal consumers' class that contains 91.47% records of the whole dataset). Besides, if class imbalanced issue is not properly tackled, it results in model overfitting. In addition, in [23], the data imbalanced problem is dealt using RUS method. However, in RUS technique, the data of the majority class is randomly deleted due to which important data is lost that leads the classifier to underfitting.

TABLE 1. Overview of the existing literature.

Models used	Achievements	Data analysed	Evaluation measures	Problems
Extreme gradient boosting [8]	To enhance ETD performance	Endesa	AUC, precision-recall and execution time	High computational time
Wide and deep CNN [12]	To secure SGs by detecting electricity theft	SGCC daily data	AUC and MAP	Data imbalance issue
Consumption pattern based ETD [19]	To improve ETD performance	SEAI	TPR, FPR and BDR	No feature extraction
Hybrid LSTM-MLP neural network [22]	To overcome NTLs	Endesa	AUC, precision, recall and precision-recall-AUC	Data imbalance issue
Stacked sparse denoising autoencoder [39]	To tackle NTLs	SGCC hourly data	FPR, TPR and AUC	Inadequate evaluation metrics
Bagging (RF, ET) and boosting (CatBoost, XGBoost, AdaBoost, LGB) ensemble schemes [40]	To detect energy theft in power grids	CER	Precision, AUC and accuracy	Ensemble techniques are computationally complex
DT-KSVM [41]	To decrease power losses	SEAI	AUC and accuracy	Inadequate performance metrics
RF [42]	To detect NTL behavior	Hebei province	AUC and accuracy	No feature extraction
Semi supervised autoencoder [43]	To reduce NTLs by employing semi-supervised data	SGCC daily data	Accuracy, TPR, FPR, precision, recall and F1 score	Inappropriate hyperparameter tuning
Random undersampling boosting [44]	To reduce NTLs	Honduras	F1 score, MCC, precision, recall, AUC and accuracy	Loss of important information due to RUS
Combined CNN and LSTM [45]	To detect abnormal EC profiles of consumers	SGCC daily data	F1 score, MCC, precision, recall and accuracy	Classes overlap due to SMOTE
Ensemble bagged tree [46]	To minimize NTLs	MEPCO	Accuracy, sensitivity, specificity, F1 score and FPR	Curse of dimensionality
Functional encryption based privacy-preserving ETD [47]	To detect ET by preserving consumers' privacy	CER	Highest difference (HD), FPR, DR and accuracy	High computational complexity due to improper hyperparameter optimization
Privacy-preserving ETD [48]	To perform ETD while maintaining consumers' privacy	CER	HD, DR and FPR	Improper hyperparameter tuning
Multiple linear regression model [49]	To overcome NTLs	Neighborhood area network dataset	Accuracy, sensitivity and specificity	Curse of dimensionality

IV. PROPOSED SYSTEM MODEL FOR ELECTRICITY THEFT DETECTION

This section comes up with a complete and detailed solution for ETD in SGs. The solution contains multiple sub-modules, such as data preprocessing, FS, DB, and classification. All these sub-modules are graphically abstracted in our proposed system model for ETD, shown in Fig. 1. All of the aforementioned sub-modules along with the used techniques are comprehensively discussed in the following subsections. Furthermore, Fig. 2 provides the proposed system model's flowchart wherein the working of the model is given.

The preprocessing of the data acquired from the state grid corporation of China (SGCC) dataset initiates the proposed mode's working. In this process, NaN values and outliers are checked, and normalization is performed. Afterwards, the FS is done using RFE technique. Then, the data splitting is performed using training and testing parts, in which 80% data is specified for training and 20% is specified for testing. Afterwards, the training data, i.e., (X_train, y_train) is balanced using KNNOR and the testing data, i.e., (X_test, y_test) remains the same. After balancing the training data using KNNOR, it is represented by (XK_train, yk_train). Afterwards, the balanced training data (X_train, y_train) is passed to BiLSTM-LogitBoost for fitting and then (X_test) is passed to the model for classification purpose. Finally, the classification labels generated by BiLSTM-LogitBoost are stored into the (BLB_pred) variable. At the end, the predicted labels (BLB_pred) and the actual labels (y_test) are passed to the performance evaluation process to measure the performance of our proposed system model.

A. DATA PREPROCESSING

Preprocessing the EC data is a necessary step to be taken for preparing the data by filling the NaN data, mitigating outliers, and scaling the data to a specific range so that an ML or DL classifier can easily learn from it. Therefore, our

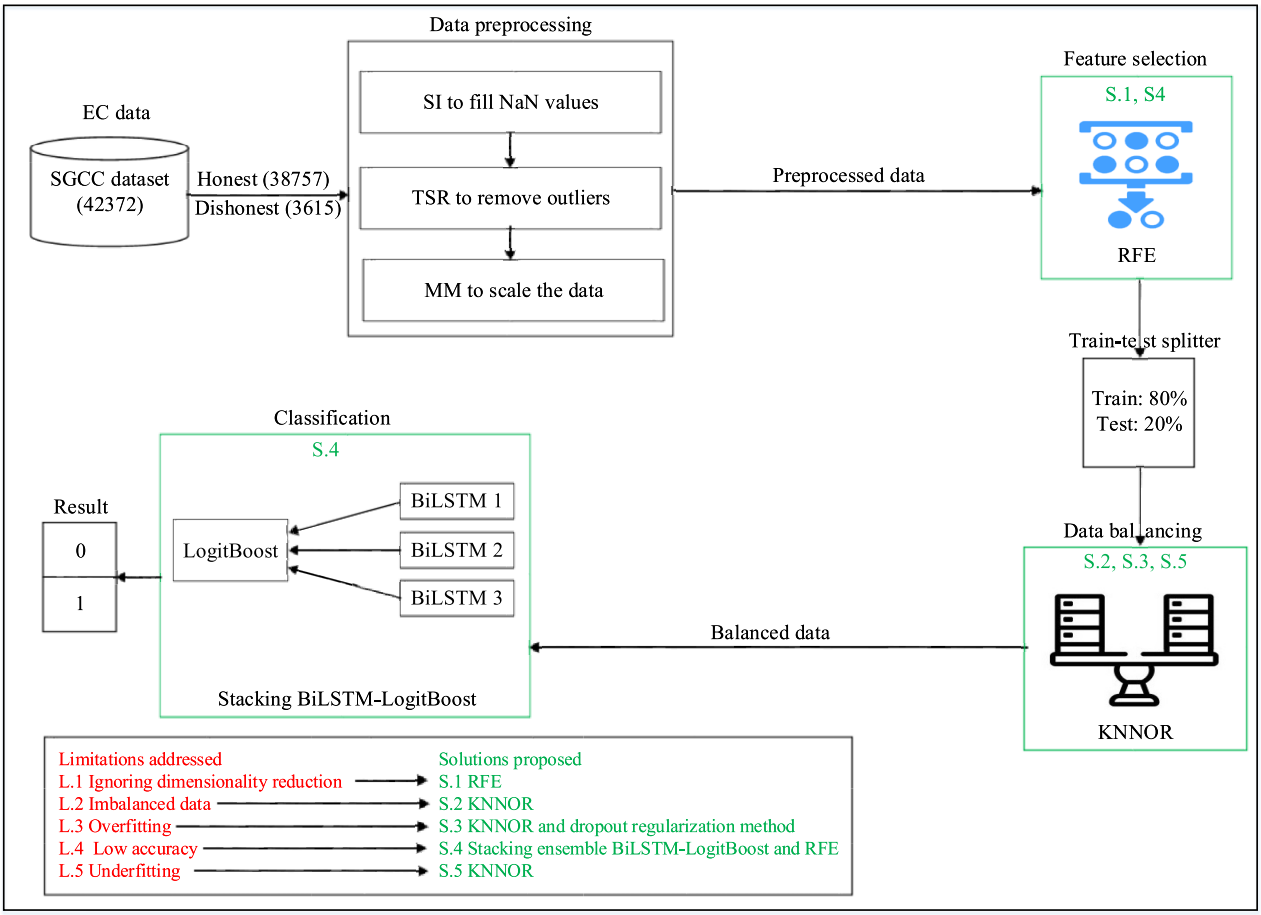


FIGURE 1. Proposed ETD model.

TABLE 2. Dataset information.

Description	Values
Duration	01-01-2014 to 31-10-2016
Benign records	38757
Theft records	3615
Total records	42372

proposed solution employs the simple imputer (SI) with mean method [54], [55], min-max (MM) scaler [12], and three sigma rule (TSR) [24] to preprocess the EC readings collected from residential consumers of the Fujian province in China. The EC data is collected by the SGCC [56], an electric utility company in China. The EC data is collected from those customers who have SMs installed in their residential places. Table 2 comprises the SGCC dataset's information.

The NaN values and outliers may exist in the dataset due to transmission error, broken component, storage loss, etc., [12], [57]. In our proposed system model, we leverage SI to deal with the NaN readings. There are three imputation strategies in SI: median, mean, and most_frequent (mode) [55]. We selected the default strategy of SI, i.e., mean, to impute the missing data in this study. Furthermore, TSR is leveraged to deal with the outlier readings by implementing Equation 1 [24].

$$f(x_{i,a}) = \begin{cases} M + 3.std(X), & x_{i,a} > M + 3.std(X), \\ x_{i,a}, & Otherwise. \end{cases} \quad (1)$$

where a shows a slot, i.e., day in our case. i indicates the consumer number. $x_{i,a}$ indicates the EC data of the ith number consumer at the ath day. Where $a = 1034$ and $i = 42372$. X is a dataframe that consists of multiple $x_{i,a}$ EC readings. M shows the average of X. $std(X)$ indicates the standard deviation for X. Furthermore, due to the sensitivity of DL classifiers' towards unscaled data, normalization is needed that is performed using the MM scaler by implementing Equation 2 [12].

$$f(x_{i,a}) = \frac{x_{i,a} - X_{min}}{X_{max} - X_{min}}, \quad (2)$$

where X_{min} represents the minimum value of X and X_{max} shows the maximum EC reading of X.

B. FEATURE SELECTION

In this subsection, FS performed in the proposed work is discussed. We employed a technique known as RFE to select the most effective features for ETD in SGs.

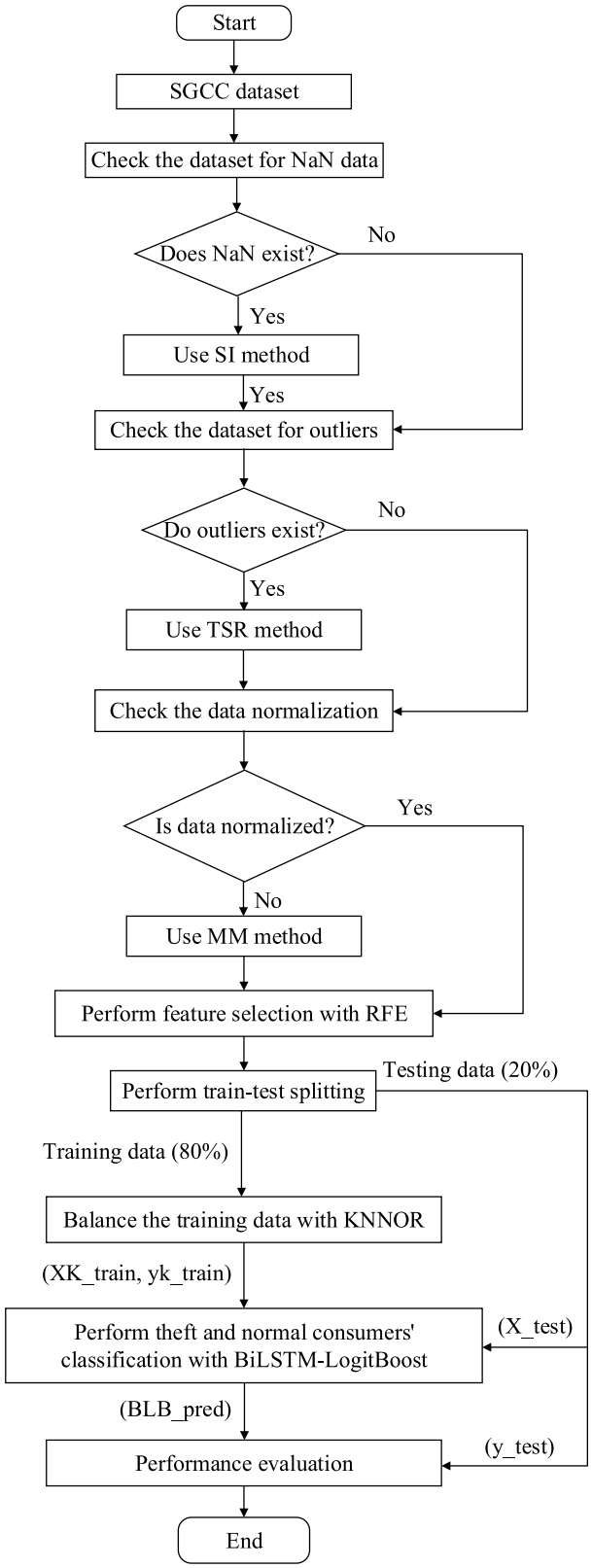


FIGURE 2. Flowchart of the proposed ETD system model.

FS is a task of feature engineering that is fulfilled by selecting only a small number of features (dimensions or columns) from a large number of features [58]. In other words, it is a process of automatically selecting those features that have a huge impact on the output [59]. Generally, there are three advantages of performing FS: achieve high accuracy, minimize overfitting, and minimize model's training time.

In this paper, we are using SGCC dataset that contains a broad range of EC data. As the dataset comprises a large number of features, so its execution needs huge amount of computational resources. We perform our simulations using Google Colaboratory, which provides 12 GB of RAM to its users. Besides, KNNOR method is used to balance the data present in the dataset. KNNOR is a KNN based oversampling approach, where KNN is a lazy learner. KNN performs predictions during testing phase by calculating the distance of each testing sample from all training samples and then assigning a specific label to the testing sample. It is a highly time-consuming and computationally complex process. Thus, KNN needs a huge amount of computational resources to perform prediction.

In addition, KNNOR needs large amount of computational resources to oversample and balance huge datasets like SGCC. Therefore, in order to perform analysis on such datasets without facing resource overloading issue, the number of rows and columns need to be decreased. In the proposed work, the number of rows are reduced from 42372 to 10000 by maintaining the same class distribution ratio of theft and normal classes as in the original dataset, i.e., 91.47% normal consumers and 8.53% theft consumers. To reduce the number of columns (features), we have many FS and extraction techniques, such as RFE, RF, linear discriminant analysis (LDA), principle component analysis, etc. In addition, some population based meta-heuristic techniques are also recently employed for FS, such as black hole algorithm [60] and binary jaya algorithm [61]. However, we choose RFE to perform FS for its popularity due to its enhanced effectiveness at picking those columns from the given dataset that largely affect the results [59]. There are two significant configuration choices to choose from when employing RFE method for FS. The first one is selecting the number of features (n_features_to_select) while the second one is selecting an optimal algorithm for FS.

Moreover, due to the availability of limited RAM, only 120 features are selected out of 1034 features using RFE to successfully balance the dataset. If the number of selected features is more than 120, the issue of session crash is faced. Due to this, the first configuration choice is used and the values of n_features_to_select is set to be 120. It ensures successful FS, DB, classification and generating ETD results. Initially, 120 features and 10000 records are passed to the train-test splitter for dividing the dataset into training split and testing split. Afterwards, the training data having 8000 records and 120 features is passed to KNNOR for DB.

C. DATA BALANCING

After performing FS using RFE, data splitting is performed. The ratio of 80:20% is selected for splitting the

dataset into training and testing sets. Now, in this subsection, the DB module of the proposed system model is discussed.

One of the challenging problems in detecting electricity theft in SGs is imbalanced distribution of the class observations (data imbalanced problem) where a class has more observations than another class. When DL and ML classifiers are trained using datasets with such behavior, classifiers become skewed towards the majority class (class owns more observations) and neglect the minority class (class owns less number of observations). Such situations negatively affect the classifiers and lead it to the overfitting issue. In the existing literature, the imbalance between data classes is tackled by some researchers using RUS [23], AdaSyn [37], and SMOTE [37], [62] techniques. However, RUS is prone to underfitting problem. Whereas, SMOTE and AdaSyn are prone to within class imbalance and small disjunct problems. Therefore, in this study, KNNOR [63] is used to tackle the class imbalanced, overfitting, small disjunct, within class imbalance, and underfitting problems. We replaced the RUS (an undersampler) with KNNOR (an oversampler) to resolve underfitting problem. RUS is a random undersampler that performs DB by randomly discarding the data points from the majority class, which causes the loss of important data that results in an underfitting problem. Moreover, KNNOR deals with the problems of with class imbalance and small disjunct by focusing on the compactness and location of the minority class as compared to the majority class. The EC data is balanced using KNNOR for the first time in ETD domain as per our knowledge.

KNNOR technique consists of the following three steps to find out safe and critical spaces, and augment the minority class' data by generating synthetic samples.

1) In this step, minority class samples are sorted according to distance to their k-th nearest neighbor. This empowers the technique to reach the place with huge amount of minority class samples. In this way, it ignores the noisy and outlier data points.
2) In the second step, the sorted data points and their relevant k-nearest neighbors are employed to artificially generate a sample. The action of new synthetic points initiates with an initial data point and another random point (discovered at random position between the initial data sample and its first closest neighbor). The newly discovered data point is now focused, and another artificial point is randomly found between initial's second nearest neighbor and the newly discovered data point. After n iterations, the data sample created is considered as synthetically generated nominee.
3) Finally, in this step, the synthetically created nominee that is the output of step 2 is tested using KNN classifier to explore if it is relevant to the minority class (the class which is being oversampled) or not. If it clears the test, it is maintained, otherwise, it is discarded.

The process flow of the KNNOR algorithm for data augmentation is provided in Algorithm 1 [63].

Algorithm 1 KNNOR Based Data Augmentation

Input: Training set (X_train) that consists of majority class data (X_train_{maj}) and minority class data (X_train_{min}), top distance threshold (d)% of the sorted list of minority sample ($Sorted_{min}$); count of neighbors (K); number of data samples to be generated (gen_count);
Output: Synthetic data point

```
Get the α value using Algorithm 2
while gen_count > 0 do
    for each data point in (Sorted_min) do
        source = data point
        for k closest neighbor n of data point do
            new point=data point created at random
            distance (0, α) on a line between source
            and n
            source = new point
        end for
        if new point is valid then
            gen_count = gen_count-1
            Add new data point to the augmented_data
            if gen_count == 0 then
                return augmented_data
            end if
        end if
    end for
end while
return augmented_data
```

Algorithm 2 α Value Calculation Algorithm

Input Training dataset (X_train) that contains the majority class data (X_train_{maj}) and minority class data (X_train_{min});
Output α's value

```
Overall_dist = distance between data points from (X_train_maj) to data points in (X_train_min)
Minimum_dist = Minimum (Overall_dist)
return Minimum_dist
```

After dataset splitting into 80% training and 20% testing parts, the the training samples (X_train) and their respective labels (y_train) are balanced using KNNOR. Fig. 3 illustrates the theft and normal classes' distribution before KNNOR implementation. Whereas, Fig. 4 shows theft and normal classes' frequency after KNNOR implementation. In Fig. 3, 7326 samples belong to the normal class and 674 samples belong to the theft class. Afterwards, KNNOR is used to oversample the theft class data until it becomes equal to the normal class data. So now the classes are balanced in a way that both the normal and theft classes contain the same number of records that is 7326, as shown in Fig. 4.

FIGURE 3. Normal and theft classes' distribution before KNNOR implementation.

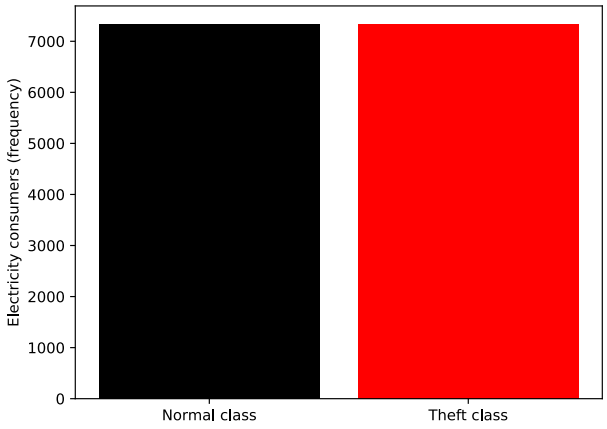


FIGURE 4. Normal and theft classes' distribution after KNNOR implementation.

D. ELECTRICITY THEFT AND NON-THEFT CLASSIFICATION

In this subsection, the proposed deep and machine learning (ML) stacking ensemble model (BiLSTM-LogitBoost) is discussed, which consists of two models BiLSTM [64] and LogitBoost [65]. The details of the proposed model are given in the below subsection.

1) BiLSTM-LogitBoost STACKING ENSEMBLE MODEL

To discuss BiLSTM, it is necessary to discuss LSTM first as it is the basic version. LSTM was developed by Hochreiter and Schmidhuber [66]. It is an impressive recurrent neural network (RNN) model particularly invented to address the gradient vanishing and gradient exploding problems that generally happen when learning the long duration dependencies [67]. This is avoided by constant error carousel (CEC), which keeps the error signal inside each cell. Actually, these cells are RNNs with an impressive and attractive architecture that CEC is expanded with added features, known as input and output gates. These gates build a memory cell. The self recurrent arrows present a feedback having lag of single time-step. A vanilla LSTM consists of input, output, and forget gates. These gates are employed to control the information flow in the cell. The cell maintains and remembers the data in arbitrary time-intervals. The BiLSTM [64] model is the enhanced version of the traditional LSTM [66].

LogitBoost [65] is a well-known boosting classifier that can be employed for both binary and multiclass classification tasks. The weak learners' construction is one of the key elements in boosting techniques that affect their performances. LogitBoost trains each weak learner, which we considered in this paper, i.e., stump weak classifier, separately for every single class.

In this article, we choose to develop stacking ensemble model for ETD in SGs. The reason is that it is the best strategy among different modern approaches based on the wining of numerous Kaggle and Netflix competitions for solving classification problems [68]. Stacking ensemble is a robust approach in which we can place multiple classifiers at level-0 (also called base classifiers) and a single classifier at level-1 (also called meta classifier) to obtain higher prediction performance. Particularly, in our stacking ensemble model, at level-0, three DL models are used, which are recent and very rarely employed in ETD domain. at level-1, an ML technique is employed, which is a novel technique and to our knowledge, used for the first time for ETD in SGs. Level-1 classifier is employed to combine the outputs of the level-0 classifiers and generate the final electricity theft detection results.

The base classifiers' selection for stacking ensemble model's construction can be performed into three different ways. Firstly, choose the heterogeneous (different) base models, secondly, use the same base models with different configurations, and thirdly, the same base models fitted on different datasets [59]. In this paper, we choose the second option to build BiLSTM-LogitBoost, which is also chosen by [69] and [70] to build their stacking ensemble models. Consequently, we develop a stacking ensemble model that consists of multiple BiLSTM (with different configurations) and a single LogitBoost. Three BiLSTM networks are considered as the base models and are used at level-0 while a single LogitBoost classifier is considered as the meta classifier and is used at level-1 of the proposed stacking ensemble model. The number of Dense layers, Dropout layers, and neurons are the internal configurations of the three BiLSTM models, which are differently chosen from each other to prepare level-0 classifiers of our stacking model. The more details about different configurations of the same BiLSTM model chosen in this study are given in Algorithm 3. As the proposed model is composed of three DL, i.e., three BiLSTM and an ML classifier, i.e., LogitBoost, thus, we can also call it as a stacking ensemble DL and ML model. The architecture and algorithm of the stacking BiLSTM-LogitBoost model are given in Fig. 5 and Algorithm 3, respectively.

V. DISCUSSION OF THE SIMULATION RESULTS

The proposed BiLSTM-LogitBoost stacking ensemble model, proposed for ETD in SGs, is evaluated and discussed in this section. Some recent benchmarks, such as SVM [19],

TABLE 3. Confusion matrix.

		Actual class	
		Theft consumer	**Benign consumer**
Predicted class	**Theft consumer**	TP	FP
	Benign consumer	FN	TN

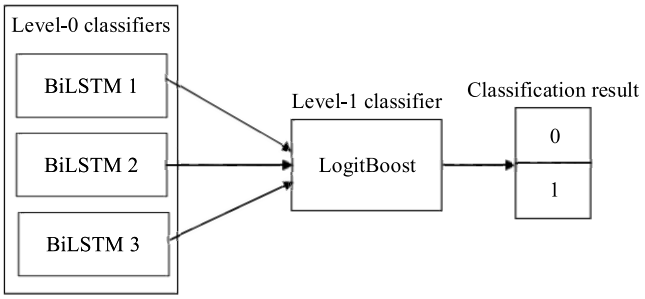


FIGURE 5. BiLSTM-LogitBoost stacking ensemble architecture.

[71], logistic regression (LR) [37], decision tree (DT) [37], LSTM [21], [71], adaptive boosting (AdaBoost) [37], BiLSTM [64], LogitBoost [65], and LSTM-AdaBoost [72] are also implemented for ETD and their results are compared with the proposed model. LogitBoost with *n_estimators* = 25 is employed as a benchmark technique to our proposed model. The proposed model is trained and tested using the real SGCC EC data. The implementation of the proposed model is done using Tensorflow, Keras, and LogitBoost libraries in Python programming language.

A. SIMULATIONS' SETUP

Due to the limited resources of our local system, Google Colaboratory is used to investigate the performance of our model for ETD in SGs. Our proposed BiLSTM-LogitBoost is implemented using the Tensorflow, Keras, and LogitBoost libraries using Python programming language. In addition, realistic EC data from SGCC dataset that belongs to the electricity users of Fujian province in China is leveraged for proposed model's training and validation. The data comprises EC history of 42372 users for 2 years and 10 months (01 Jan 2014 to 31 Oct 2016). In this data, 38757 consumers are normal and the rest 3615 are abnormal consumers, given in Table 2. However, for our model's training and testing, we computed 120 features out of 1034 using RFE method and at random selected only 10000 records out of 42372, the reasoning of which is mentioned in the subsection IV-B. Furthermore, to come up with outstanding simulation results, the proposed solution starts with data preprocessing. In this step, SI, TSR, and MM schemes are leveraged to perform data preprocessing, as already discussed in Subsection IV-A. Afterwards, FS is performed using RFE method, already given in Subsection IV-B. After that, the preprocessed data with selected features is splitted into training and testing sub datasets. The training part comprises 80% while the testing part comprises 20% of the total dataset. In the next step, DB is performed using a novel KNNOR scheme, as discussed in Subsection IV-C. Finally, the balanced data obtained from KNNOR is passed to our proposed BiLSTM-LogitBoost stacking ensemble model to perform classification task and separate electricity thieves from non-thieves, as discussed in detail in Subsection IV-D.

B. PERFORMANCE MEASURES

In this subsection, the parameters employed for evaluating the proposed model's performance are discussed. The SGCC dataset is leveraged for ETD in this paper. The dataset contains two classes, i.e., benign consumers and theft consumers. So ETD is considered as a two class (binary) classification problem. The initial results of classification techniques are generally shown using a confusion matrix (CM) and then further performance metrics are calculated using CM. The CM is given in Table 3.

Elements of the CM matrix are discussed as follows. True positive (TP) denotes that the theft consumers are accurately classified as theft consumers by the classifier, false negative (FN) means theft users are falsely predicted as benign users, false positive (FP) denotes that benign consumers are wrongly classified as theft consumers, and true negative (TN) means benign users are correctly predicted as benign users by the classifier. In supervised ML, the trained models are validated using their capability to successfully predict the correct labels for unseen and unlabeled data. To successfully accomplish this job, different performance evaluation parameters are available, as given in study [73]. On the other hand, it is not practical to employ all of the performance metrics mentioned in the study; therefore, in this paper, we employed few metrics from them that are most relevant and widely used in the recent literature of ETD [21], [71], [74]. Hence, the proposed model's performance evaluation is done via different performance metrics. The mathematical formulas of these metrics are provided in Equations 3-6 [37], [74].

$$Precision = \frac{TP}{FP + TP}, \tag{3}$$

$$Recall = \frac{TP}{FN + TP}, \tag{4}$$

$$F1\ score = 2 * \frac{Recall * Precision}{Recall + Precision}, \tag{5}$$

$$Accuracy = \frac{TN + TP}{FP + TP + FN + TN}, \tag{6}$$

All of the selected performance measures are computed using the elements available in the CM. Accuracy is the ratio between the accurately predicted samples and total records available in the dataset, F1 score is an important

Algorithm 3 BiLSTM-LogitBoost Stacking Ensemble Model

Input: Balanced dataset
Output: Electricity theft and non-theft classification

```
Creating and compiling BiLSTM 1:
model.add(Bidirectional(LSTM(100)))
model.add(Dense(100))
model.add(Flatten())
model.add(Dropout(0.2))
model.add(Dense(1, activation='sigmoid'))
model.compile('Adam', 'binary_crossentropy')
Creating and compiling BiLSTM 2:
model.add(Bidirectional(LSTM(100)))
model.add(Dropout(0.2))
model.add(Dense(100))
model.add(Dropout(0.2))
model.add(Dense(100)
model.add(Flatten())
model.add(Dropout(0.2))
model.add(Dense(1, activation='sigmoid'))
model.compile('Adam', 'binary_crossentropy')
Creating and compiling BiLSTM 3:
model.add(Bidirectional(LSTM(120)))
model.add(Dropout(0.2))
model.add(Dense(100))
model.add(Dropout(0.2))
model.add(Dense(100)
model.add(Dropout(0.2))
model.add(Dense(100)
model.add(Flatten())
model.add(Dropout(0.2))
model.add(Dense(1, activation='sigmoid'))
model.compile('Adam', 'binary_crossentropy')
Making keras classifiers for the above models:
BiLSTM_clf1=KerasClassifier(build_fn=BiLSTM1,
epochs=15, batch_size= 32)
BiLSTM_clf2=KerasClassifier(build_fn=BiLSTM2,
epochs=15, batch_size= 32)
BiLSTM_clf3=KerasClassifier(build_fn=BiLSTM3,
epochs=15, batch_size= 32)
Combining the BiLSTM_clf1, BiLSTM_clf2, and BiL-
STM_clf3 models:
intermediate=[BiLSTM_clf1,BiLSTM_clf2,
BiLSTM_clf3]
Stacking three BiLSTM and a LogitBoost models:
BiLSTM-LogitBoost=StackingClassifier(estimators=
intermediate, final_estimator=LogitBoost())
Fitting and prediction of BiLSTM-LogitBoost model:
BiLSTM-LogitBoost.fit(XK_train, yk_train)
BiLSTM-LogitBoost.predict(X_test)
```

metric that gives equal weight to both precision and recall [37]. Moreover, precision and recall are computed using the elements of the CM and their formulas are stated in Equations 3 and 4.

C. PROPOSED BiLSTM-LogitBoost MODEL PERFORMANCE RESULTS

In this subsection, we tackle and validate the problems mentioned in Section III, i.e., no dimensionality reduction is performed that leads the classifier to the problems of overfitting and low detection accuracy. Another issue is of imbalanced data that results in biased results and overfitting. Another limitation is that some authors employed RUS for balancing data that randomly removes data samples of the majority class and results in loss of important information and underfitting issue. The last limitation we tackled in this work is that some authors in the literature employed the standalone ML and DL schemes that leads to low detection accuracy. In order to conduct detailed evaluation of the proposed approach, its analysis is performed in multiple levels, i.e, BiLSTM-LogitBoost analysis with FS and BD steps versus BiLSTM-LogitBoost without FS and DB steps, BiLSTM-LogitBoost with KNNOR based DB versus BiLSTM-LogitBoost with other DB techniques, BiLSTM-LogitBoost with KNNOR based DB versus BiLSTM-LogitBoost with RUS based DB, and the final level analysis is done between proposed BiLSTM-LogitBoost model followed by KNNOR and RFE versus other benchmarks.

In first level of the proposed model analysis, we perform the comparison of results of the BiLSTM-LogitBoost stacking model (without FS and DB) and BiLSTM-LogitBoost with FS and DB, as depicted in Table 4 and Fig. 6. The results depict that BiLSTM-LogitBoost stacking model without FS and DB steps achieves better results as compared to the BiLSTM-LogitBoost stacking model with FS and DB steps. The reason is that without considering DB, BiLSTM-LogitBoost generates biased results and leads to overfitting problems. It means that the model learns and ultimately memorizes the normal class (represented by 0) data because model has a huge amount of normal class data, i.e., 91.47% samples for normal class. Whereas, it has very limited records (8.53%) for theft class (represented by 1), therefore, it generates the highest detection accuracy and falsely predicts the theft class data as normal as well.

In second level analysis of the proposed model, the BiLSTM-LogiBoost with KNNOR and other DB techniques comparison is performed. Table 5 and Fig. 7 show the analysis of our proposed model with KNNOR and other oversampling methods to examine the importance of the balanced and imbalanced class distributions. The simulation results show that BiLSTM-LogitBoost stacking model with KNNOR based DB yields the maximum results as compared to ADASYN and SMOTE based DB techniques. The reason is that SMOTE and ADASYN are prone to the within class imbalance and small disjunct problems. Moreover, SMOTE based DB leads the model to the overfitting problem because it overflows (overpopulates) the specific location by generating synthetic data instead of generating it throughout the data. Therefore, KNNOR based BiLSTM-LogitBoost outperforms the SMOTE and ADASYN based BiLSTM-LogitBoost models. However,

TABLE 4. Comparison of the proposed BiLSTM-LogitBoost with and without FS and DB steps.

Classifier	Precision	Recall or DR	F1 score	Accuracy
BiLSTM-LogitBoost (Without FS and DB)	0.9956	0.9143	0.9532	0.9110
BiLSTM-LogitBoost (With FS and DB)	0.9632	0.9241	0.9433	0.8945

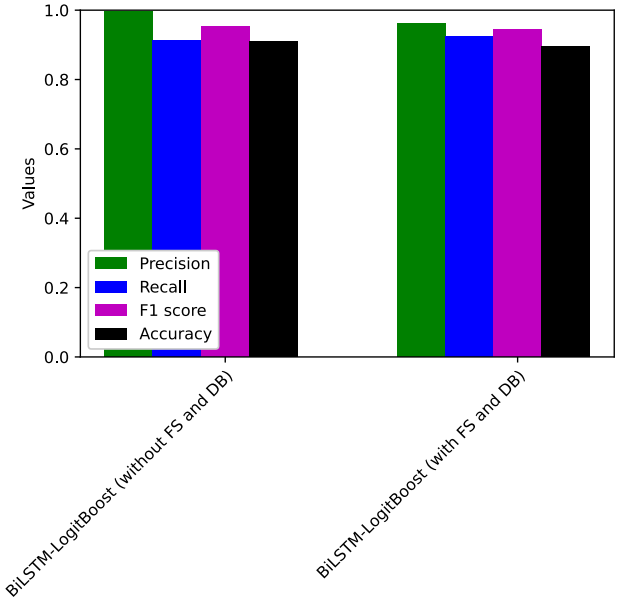


FIGURE 6. Comparison of the proposed BiLSTM-LogitBoost with and without FS and DB steps.

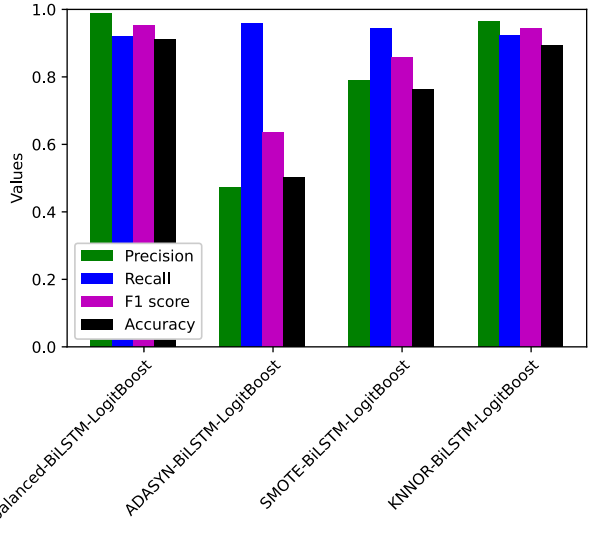


FIGURE 7. Comparison of our proposed BiLSTM-LogitBoost with KNNOR and other DB methods.

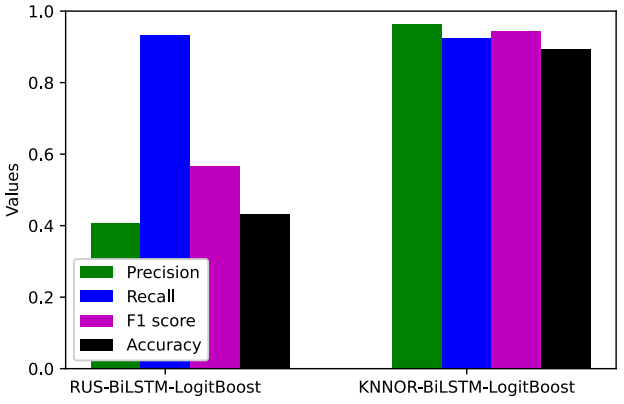


FIGURE 8. Comparison of the proposed BiLSTM-LogitBoost with KNNOR and RUS DB methods.

imbalanced data based BiLSTM-LogitBoost beats all the SMOTE, ADASYN, and KNNOR data balanincg based BiLSTM-LogitBoost models. The reason is that imbalanced data based BiLSTM-LogitBoost generates biased results. It means that the model memorizes the normal class data and gives the maximum detection accuracy. KNNOR-BiLSTM-LogitBoost obtains better results than others balancing techniques based BiLSTM-LogitBoost because KNNOR efficiently solves the issues raised in SMOTE and ADASYN schemes.

In third level analysis of the proposed model, the BiLSTM-LogitBoost model with KNNOR and BiLSTM-LogitBoost with RUS DB techniques' comparison is done. This comparison is actually done to validate that our proposed KNNOR based BiLSTM-LogitBoost successfully tackled the underfitting problem raised by the RUS DB based BiLSTM-LogitBoost. Table 6 and Fig. 8 present that proposed KNNOR DB based BiLSTM-LogitBoost technique yields the highest performance results as compared to the RUS DB based BiLSTM-LogitBoost. The RUS based balanced samples do not obtain good ETD performance results because it randomly discards the samples of the majority class, which results in loss of important data, and the loss of useful data leads to underfitting issue.

In the final analysis of the proposed methodology, the overall results considering all steps (FS, DB, and classification) for the proposed as well as benchmarks are provided. In this step, in order to perform realistic evaluation of our proposed stacking technique (followed by KNNOR AND RFE) and benchmarks, all the schemes (BiLSTM-LogitBoost and benchmarks) are fed with the same input. It means after performing FS using RFE and DB using KNNOR, the data is passed to the BiLSTM-LogitBoost as well as other benchmarks. Table 7 and Fig. 9 show that performance results of the proposed stacking model are improved than all the benchmarks in terms of precision, F1 score and accuracy. The reason that our proposed ensemble stacking model outperforms the standalone SVM, LR, DT, LSTM, AdaBoost, BiLSTM, and LogitBoost models is that our model is a stacking

TABLE 5. Comparison of our proposed BiLSTM-LogitBoost with KNNOR and other DB methods.

Classifier	Precision	Recall or DR	F1 score	Accuracy
Imbalanced-BiLSTM-LogitBoost	0.9879	0.9202	0.9529	0.9110
ADASYN-BiLSTM-LogitBoost	0.4739	0.9589	0.6343	0.5025
SMOTE-BiLSTM-LogitBoost	0.7886	0.9429	0.8588	0.7640
KNNOR-BiLSTM-LogitBoost	0.9632	0.9241	0.9433	0.8945

TABLE 6. Comparison of the proposed BiLSTM-LogitBoost with KNNOR and RUS DB methods.

Classifier	Precision	Recall or DR	F1 score	Accuracy
RUS-BiLSTM-LogitBoost	0.4069	0.9332	0.5667	0.4335
KNNOR-BiLSTM-LogitBoost	0.9632	0.9241	0.9433	0.8945

TABLE 7. Comparison of the BiLSTM-LogitBoost with other benchmarks.

Classifier	Precision	Recall or DR	F1 score	Accuracy
SVM	0.9462	0.9224	0.9341	0.8785
LR	0.8698	0.9395	0.9033	0.8305
DT	0.9138	0.9306	0.9221	0.8595
LSTM	0.9242	0.9268	0.9255	0.8645
AdaBoost	0.9165	0.9246	0.9206	0.8560
BiLSTM	0.9149	0.9271	0.9209	0.8570
LogitBoost	0.9346	0.9146	0.9245	0.8610
LSTM-AdaBoost	0.8962	0.9288	0.9122	0.8430
BiLSTM-LogitBoost	0.9632	0.9241	0.9433	0.8945

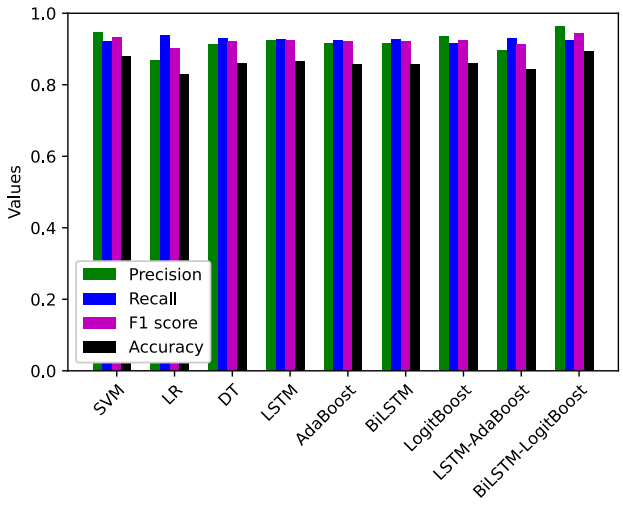


FIGURE 9. Comparison of the BiLSTM-LogitBoost with other benchmarks.

TABLE 8. Mapping table.

Limitations	Solutions	Validations
L.1 Ignoring dimensionality reduction [19], [21]	RFE technique is used for dimensionality reduction	Table 7 and Fig. 9.
L.2 Imbalanced data problem [22]	DB is performed using KNNOR	Table 5 and Fig. 7.
L.3 Overfitting issue [19], [21], and [22]	KNNOR and dropout regularization	Table 7 and Fig. 9
L.4 Low accuracy [19], [21]	RFE FS and BiLSTM-LogitBoost techniques	Table 7 and Fig. 9.
L.5 Underfitting due usage of RUS for DB [23]	KNNOR	Table 6 and Fig. 8.

ensemble model that combines three BiLSTM models with different configurations at level-0 and one LogitBoost model at level-1. That is why our ensemble model outperforms the other single models with respect to different performance metrics. This actually validates our claim. Secondly, our proposed stacking ensemble model also beats another stacking ensemble model, i.e., LSTM-AdaBoost in terms of various performance metrics. The reason is that we selected the most recent models in our stacking ensemble, i.e., BiLSTM as base learners and LogitBoost as a meta learner. Whereas, the benchmark ensemble contains LSTM and AdaBoost classifiers as its base and meta classifiers, respectively, which are old and conventional techniques. Therefore, our proposed stacking ensemble model beats the benchmark stacking ensemble model. Finally, after simulation results' comparison, it is proved that BiLSTM-LogitBoost stacking ensemble in combination with KNNOR and RFE is the outstanding performer among all the benchmarks. It ensures that KNNOR and RFE are well compatible with BiLSTM-LogitBoost as compared to other benchmarks. Hence, it is proved that our proposed BiLSTM-LogitBoost followed by KNNOR and RFE is the best performing model for ETD in SGs. Furthermore, our proposed BiLSTM-LogitBoost stacking ensemble model employs three BiLSTM models at level-0 and a LogitBoost model at level-1. Moreover, the deep models are scalable [75], [76], [77] and can perform well on large amount of data. In short, by increasing data samples, deep models' performance increases. Since we employed three deep models as level-0 learners in our proposed model, we can say that our proposed BiLSTM-LogitBoost is scalable and can be used for any amount of data.

The mapping table of the limitations, their respective solutions, and validations is given in Table 8, where the limitations tackled in the proposed work are denoted as L.1, L.2, L.3, L.4, and L.5.

VI. CONCLUSION

In this article, we come up with an efficient and effective stacking ensemble model ETD model, wherein BiLSTM-LogitBoost is followed by the usage of KNNOR and RFE trechniques. RFE is used for significant FS while KNNOR is used for DB. Further, BiLSTM-LogitBoost model is designed to generate the final ETD results. Extensive simulations are performed using SGCC data. The results yield 96.32% precision, 94.33% F1 score, and 89.45% accuracy values. The values are greater than all the compared benchmarks, which exhibits efficient performance of the proposed model. Consequently, it is concluded that the proposed BiLSTM-LogitBoost stacking ensemble model followed by KNNOR and RFE gives greater ETD performance. Furthermore, our proposed model is trained using daily EC data, which is considered as low frequency data as compared to the hourly or minute-wise EC data. In this way, it maintains consumers' privacy. Moreover, the performance of our proposed model is constrained when it comes to the detection of theft in high frequency EC patterns since it is trained using low frequency data. It happens to be a drawback of our proposed approach. In future, to avoid such issue, we will consider the high frequency EC data to train our model. In addition, by having access to huge amount of computational resources, we will consider more than one EC datasets with even more dimensions and samples than SGCC dataset in order to obtain a well generalized model. Moreover, in order to achieve much better ETD accuracy, we will consider the hyperparameters' tuning of BiLSTM-LogitBoost model using a novel meta-heuristic optimization algorithm.

REFERENCES

[1] Z. Yan and H. Wen, "Performance analysis of electricity theft detection for the smart grid: An overview," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–28, 2022.

[2] M. Ahmad, N. Javaid, I. A. Niaz, A. Almogren, and A. Radwan, "A cost-effective optimization for scheduling of household appliances and energy resources," *IEEE Access*, vol. 9, pp. 160145–160162, 2021.

[3] R. Khalid, N. Javaid, A. Almogren, M. U. Javed, S. Javaid, and M. Zuair, "A blockchain-based load balancing in decentralized hybrid P2P energy trading market in smart grid," *IEEE Access*, vol. 8, pp. 47047–47062, 2020.

[4] A. L. Shah, W. Mesbah, and A. T. Al-Awami, "An algorithm for accurate detection and correction of technical and nontechnical losses using smart metering," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 11, pp. 8809–8820, Nov. 2020.

[5] *EEP—Electrical Engineering Portal*. Accessed: Jun. 22, 2022. [Online]. Available: http://www.electrical-engineering-portal.com/

[6] J. L. Viegas, P. R. Esteves, R. Melício, V. M. F. Mendes, and S. M. Vieira, "Solutions for detection of non-technical losses in the electricity grid: A review," *Renew. Sustain. Energy Rev.*, vol. 80, pp. 1256–1268, Dec. 2017.

[7] *Electrical India Magazine*. Accessed: Jun. 22, 2022. [Online]. Available: http://www.electricalindia.in/

[8] 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.

[9] K. Fei, Q. Li, C. Zhu, M. Dong, and Y. Li, "Electricity frauds detection in low-voltage networks with contrastive predictive coding," *Int. J. Electr. Power Energy Syst.*, vol. 137, May 2022, Art. no. 107715.

[10] *PR Newswire: Press Release Distribution, Targeting, Monitoring and Marketing*. Accessed: Jun. 22, 2022. [Online]. Available: http://www.prnewswire.com/

[11] L. J. Lepolesa, S. Achari, and L. Cheng, "Electricity theft detection in smart grids based on deep neural network," *IEEE Access*, vol. 10, pp. 39638–39655, 2022.

[12] 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.

[13] S. Bera, S. Misra, and J. J. P. C. Rodrigues, "Cloud computing applications for smart grid: A survey," *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 5, pp. 1477–1494, May 2015.

[14] L. Tschora, E. Pierre, M. Plantevit, and C. Robardet, "Electricity price forecasting on the day-ahead market using machine learning," *Appl. Energy*, vol. 313, May 2022, Art. no. 118752.

[15] L. P. Raghav, R. S. Kumar, D. K. Raju, and A. R. Singh, "Analytic hierarchy process (AHP)—Swarm intelligence based flexible demand response management of grid-connected microgrid," *Appl. Energy*, vol. 306, Jan. 2022, Art. no. 118058.

[16] M. Enayati, G. Derakhshan, and S. M. Hakimi, "Optimal energy scheduling of storage-based residential energy hub considering smart participation of demand side," *J. Energy Storage*, vol. 49, May 2022, Art. no. 104062.

[17] N. Javaid, S. M. Mohsin, A. Iqbal, A. Yasmeen, and I. Ali, "A hybrid batcrow search algorithm based home energy management in smart grid," in *Proc. Conf. Complex, Intell., Softw. Intensive Syst.* Cham, Switzerland: Springer, 2018, pp. 75–88.

[18] Y. Li, R. Wang, and Z. Yang, "Optimal scheduling of isolated microgrids using automated reinforcement learning-based multi-period forecasting," *IEEE Trans. Sustain. Energy*, vol. 13, no. 1, pp. 159–169, Jan. 2022.

[19] 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.

[20] A. Takiddin, M. Ismail, U. Zafar, and E. Serpedin, "Deep autoencoder-based anomaly detection of electricity theft cyberattacks in smart grids," *IEEE Syst. J.*, vol. 16, no. 3, pp. 4106–4117, Sep. 2022.

[21] B. Kocaman and V. Tümen, "Detection of electricity theft using data processing and LSTM method in distribution systems," *Sādhanā*, vol. 45, no. 1, pp. 1–10, 2020.

[22] 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.

[23] G. Lin, X. Feng, W. Guo, X. Cui, S. Liu, W. Jin, Z. Lin, and Y. Ding, "Electricity theft detection based on stacked autoencoder and the undersampling and resampling based random forest algorithm," *IEEE Access*, vol. 9, pp. 124044–124058, 2021.

[24] F. Shehzad, N. Javaid, A. Almogren, A. Ahmed, S. M. Gulfam, and A. Radwan, "A robust hybrid deep learning model for detection of non-technical losses to secure smart grids," *IEEE Access*, vol. 9, pp. 128663–128678, 2021.

[25] Z. Aslam, F. Ahmed, A. Almogren, M. Shafiq, M. Zuair, and N. Javaid, "An attention guided semi-supervised learning mechanism to detect electricity frauds in the distribution systems," *IEEE Access*, vol. 8, pp. 221767–221782, 2020.

[26] H. O. Henriques, R. L. S. Corrêa, M. Z. Fortes, B. S. M. C. Borba, and V. H. Ferreira, "Monitoring technical losses to improve non-technical losses estimation and detection in LV distribution systems," *Measurement*, vol. 161, Sep. 2020, Art. no. 107840.

[27] U. Ramani, R. Ramya, R. Chandraprabha, and R. Nithya, "Development of raspberry pi based embedded scheme for power theft monitoring," in *Proc. 2nd Int. Conf. Smart Electron. Commun. (ICOSEC)*, Oct. 2021, pp. 1–4.

[28] J. Astronomo, M. D. Dayrit, C. Edjic, and E. R. T. Regidor, "Development of electricity theft detector with GSM module and alarm system," in *Proc. IEEE 12th Int. Conf. Humanoid, Nanotechnol., Inf. Technol., Commun. Control, Environ., Manage. (HNICEM)*, Dec. 2020, pp. 1–5.

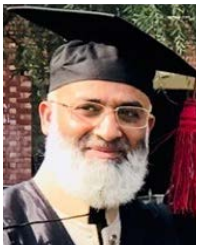
[29] P. Leninpugalhanthi, R. Janani, S. Nidheesh, R. V. Mamtha, I. Keerthana, and R. S. Kumar, "Power theft identification system using IoT," in *Proc. 5th Int. Conf. Adv. Comput. Commun. Syst. (ICACCS)*, Mar. 2019, pp. 825–830.

[30] W. Han and Y. Xiao, "A novel detector to detect colluded non-technical loss frauds in smart grid," *Comput. Netw.*, vol. 117, pp. 19–31, Apr. 2017.

[31] S. Sahoo, D. Nikovski, T. Muso, and K. Tsuru, "Electricity theft detection using smart meter data," in *Proc. IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf. (ISGT)*, Feb. 2015, pp. 1–5.
[32] D. Gu, Y. Gao, K. Chen, S. Junhao, Y. Li, and Y. Cao, "Electricity theft detection in AMI with low false positive rate based on deep learning and evolutionary algorithm," *IEEE Trans. Power Syst.*, early access, Feb. 10, 2022, doi: 10.1109/TPWRS.2022.3150050.
[33] Pamir, N. Javaid, U. Qasim, A. S. yahaya, E. H. Alkhammash, and M. Hadjouni, "Non-technical losses detection using autoencoder and bidirectional gated recurrent unit to secure smart grids," *IEEE Access*, vol. 10, pp. 56863–56875, Apr. 2022, doi: 10.1109/ACCESS.2022.3171229.
[34] 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.
[35] S. Hussain, M. W. Mustafa, T. A. Jumani, S. K. Baloch, H. Alotaibi, I. Khan, and A. Khan, "A novel feature engineered-CatBoost-based supervised machine learning framework for electricity theft detection," *Energy Rep.*, vol. 7, pp. 4425–4436, Nov. 2021.
[36] M. J. Abdulaal, M. I. Ibrahem, M. M. E. A. Mahmoud, J. Khalid, A. J. Aljohani, A. H. Milyani, and A. M. Abusorrah, "Real-time detection of false readings in smart grid AMI using deep and ensemble learning," *IEEE Access*, vol. 10, pp. 47541–47556, 2022.
[37] A. Banga, R. Ahuja, and S. C. Sharma, "Accurate detection of electricity theft using classification algorithms and Internet of Things in smart grid," *Arabian J. Sci. Eng.*, vol. 47, pp. 1–17, Aug. 2021.
[38] Pamir, N. Javaid, S. Javaid, M. Asif, M. U. Javed, A. S. Yahaya, and S. Aslam, "Synthetic theft attacks and long short term memory-based preprocessing for electricity theft detection using gated recurrent unit," *Energies*, vol. 15, no. 8, p. 2778, 2022.
[39] Y. Huang and Q. Xu, "Electricity theft detection based on stacked sparse denoising autoencoder," *Int. J. Elect. Power Energy Syst.* vol. 125, Feb. 2021, Art. no. 106448.
[40] S. R. Gunturi and D. Sarkar, "Ensemble machine learning models for the detection of energy theft," *Electr. Power Syst. Res.*, vol. 192, Mar. 2021, Art. no. 106904.
[41] X. Kong, X. Zhao, C. Liu, Q. Li, D. Dong, and Y. Li, "Electricity theft detection in low-voltage stations based on similarity measure and DT-KSVM," *Int. J. Elect. Power Energy Syst.*, vol. 125, Feb. 2021, Art. no. 106544.
[42] 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, 2020.
[43] 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, 2019.
[44] 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.
[45] M. Hasan, R. N. Toma, A. A. Nahid, 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, 2019.
[46] M. S. Saeed, M. W. Mustafa, U. U. Sheikh, T. A. Jumani, and N. H. Mirjat, "Ensemble bagged tree based classification for reducing non-technical losses in Multan electric power company of Pakistan," *Electronics*, vol. 8, no. 8, p. 860, 2019.
[47] M. I. Ibrahem, M. Nabil, M. M. Fouda, M. M. E. A. Mahmoud, W. Alasmary, and F. Alsolami, "Efficient privacy-preserving electricity theft detection with dynamic billing and load monitoring for AMI networks," *IEEE Internet Things J.*, vol. 8, no. 2, pp. 1243–1258, Jan. 2021.
[48] 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.
[49] G. Micheli, E. Soda, M. T. Vespucci, M. Gobbi, and A. Bertani, "Big data analytics: An aid to detection of non-technical losses in power utilities," *Comput. Manage. Sci.*, vol. 16, no. 1, pp. 329–343, 2019.
[50] P. Kumar, G. P. Gupta, and R. Tripathi, "Toward design of an intelligent cyber attack detection system using hybrid feature reduced approach for IoT networks," *Arabian J. Sci. Eng.*, vol. 46, no. 4, pp. 3749–3778, 2021.
[51] P. Kumar, G. P. Gupta, and R. Tripathi, "A distributed ensemble design based intrusion detection system using fog computing to protect the Internet of Things networks," *J. Ambient Intell. Humanized Comput.*, vol. 12, no. 10, pp. 9555–9572, 2021.
[52] T. Saba, A. Rehman, T. Sadad, H. Kolivand, and S. A. Bahaj, "Anomaly-based intrusion detection system for IoT networks through deep learning model," *Comput. Elect. Eng.*, vol. 99, Aug. 2022, Art. no. 107810.
[53] R. V. Mendonca, J. C. Silva, R. L. Rosa, M. Saadi, D. Z. Rodriguez, and A. Farouk, "A lightweight intelligent intrusion detection system for industrial Internet of Things using deep learning algorithms," *Expert Syst.*, vol. 39, no. 5, 2022, Art. no. e12917.
[54] Z. Zhang, "Missing data imputation: Focusing on single imputation," *Ann. Transl. Med.*, vol. 4, no. 1, pp. 1–5, Jan. 2016.
[55] Scikit. *Sklearn.impute.SimpleImputer*. Accessed: Sep. 12, 2022. [Online]. Available: http://www.scikit-learn.org/stable/modules/generated/sklearn.impute.SimpleImputer.html
[56] *State Grid Corporation of China*. Accessed: Jun. 9, 2022. [Online]. Available: http://www.sgcc.com.cn/html/sgcc_main/col2017011822/column_2017011822_1.shtml
[57] Z. Yan and H. Wen, "Electricity theft detection base on extreme gradient boosting in AMI," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–9, 2021.
[58] KDnuggets. *Feature Selection—All You Ever Wanted to Know*. Accessed: Jul. 18, 2022. [Online]. Available: https://www.kdnuggets.com/2021/06/feature-selection-overview.html
[59] Scientist, David Dalisay Junior Data, and Kevin Beaulieu Software Engineer. *Machine Learning Mastery*. Accessed: Jul. 18, 2022. [Online]. Available: http://www.machinelearningmastery.com/
[60] C. C. O. Ramos, D. Rodrigues, A. N. de Souza, and J. P. Papa, "On the study of commercial losses in Brazil: A binary black hole algorithm for theft characterization," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 676–683, Mar. 2018.
[61] Y. Li and Z. Yang, "Application of EOS-ELM with binary Jaya-based feature selection to real-time transient stability assessment using PMU data," *IEEE Access*, vol. 5, pp. 23092–23101, 2017.
[62] G. Lin, H. Feng, X. Feng, H. Wen, Y. Li, S. Hong, and Z. Ni, "Electricity theft detection in power consumption data based on adaptive tuning recurrent neural network," *Frontiers Energy Res.*, vol. 9, Nov. 2021, Art. no. 773805.
[63] A. Islam, S. B. Belhaouari, A. U. Rehman, and H. Bensmail, "KNNOR: An oversampling technique for imbalanced datasets," *Appl. Soft Comput.*, vol. 115, Jan. 2022, Art. no. 108288.
[64] M. Asif, O. Nazeer, N. Javaid, E. H. Alkhammash, and M. Hadjouni, "Data augmentation using BiWGAN, feature extraction and classification by hybrid 2DCNN and BiLSTM to detect non-technical losses in smart grids," *IEEE Access*, vol. 10, pp. 27467–27483, 2022.
[65] M.-T. Cao, N.-M. Nguyen, K.-T. Chang, X.-L. Tran, and N.-D. Hoang, "Automatic recognition of concrete spall using image processing and metaheuristic optimized LogitBoost classification tree," *Adv. Eng. Softw.* vol. 159, Sep. 2021, Art. no. 103031.
[66] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
[67] G. Van Houdt, C. Mosquera, and G. Nápoles, "A review on the long short-term memory model," *Artif. Intell. Rev.*, vol. 53, no. 8, pp. 5929–5955, 2020.
[68] I. U. Khan, N. Javeid, C. J. Taylor, K. A. A. Gamage, and X. Ma, "A stacked machine and deep learning-based approach for analysing electricity theft in smart grids," *IEEE Trans. Smart Grid*, vol. 13, no. 2, pp. 1633–1644, Mar. 2022.
[69] J. Moon, S. Jung, J. Rew, S. Rho, and E. Hwang, "Combination of short-term load forecasting models based on a stacking ensemble approach," *Energy Buildings*, vol. 216, Jun. 2020, Art. no. 109921.
[70] M. Mohammed, H. Mwambi, I. B. Mboya, M. K. Elbashir, and B. Omolo, "A stacking ensemble deep learning approach to cancer type classification based on TCGA data," *Sci. Rep.*, vol. 11, no. 1, pp. 1–22, 2021.
[71] H. Jain, M. Kumar, and A. M. Joshi, "Intelligent energy cyber physical systems (iECPS) for reliable smart grid against energy theft and false data injection," *Elect. Eng.*, vol. 104, no. 1, pp. 331–346, 2022.
[72] G. A. Busari and D. H. Lim, "Crude oil price prediction: A comparison between AdaBoost-LSTM and AdaBoost-GRU for improving forecasting performance," *Comput. Chem. Eng.*, vol. 155, Dec. 2021, Art. no. 107513.
[73] G. M. Messinis and N. D. Hatziargyriou, "Review of non-technical loss detection methods," *Electr. Power Syst. Res.*, vol. 158, pp. 250–266, May 2018.

[74] R. N. Santos, S. Yamouni, B. Albiero, R. Vicente, J. A. Silva, T. F. B. Souza, M. C. M. F. Souza, and Z. Lei, "Gradient boosting and Shapley additive explanations for fraud detection in electricity distribution grids," *Int. Trans. Electr. Energy Syst.*, vol. 31, no. 9, 2021, Art. no. e13046.
[75] C. Wang, L. Gong, Q. Yu, X. Li, Y. Xie, and X. Zhou, "DLAU: A scalable deep learning accelerator unit on FPGA," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 36, no. 3, pp. 513–517, Mar. 2017.
[76] G. Thamilarasu and S. Chawla, "Towards deep-learning-driven intrusion detection for the Internet of Things," *Sensors*, vol. 19, no. 9, p. 1977, 2019.
[77] T. Schlegl, S. Schlegl, N. West, and J. Deuse, "Scalable anomaly detection in manufacturing systems using an interpretable deep learning approach," *Proc. CIRP* vol. 104, pp. 1547–1552, Jan. 2021.

PAMIR received the B.S. degree in software engineering from the National University of Modern Languages (NUML), Islamabad, Pakistan, in 2016, and the M.S. degree in software engineering from the Communication Over Sensors (ComSens) Research Laboratory, Department of Computer Science, COMSATS University Islamabad (CUI), Islamabad, under the supervision of Dr. Nadeem Javaid, in 2018, where he is currently pursuing the Ph.D. degree in computer science. He has authored two journals and 12 conference proceedings in international journals and conferences. His research interests include data science, smart grids, and optimal power flow.



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 from the University of Paris-Est, France, in 2010. He is currently a Professor and the Founding Director of the Communications Over Sensors (ComSens) Research Laboratory, Department of Computer Science, COMSATS University Islamabad, Islamabad Campus. He is also working as a Visiting Professor with the School of Computer Science, University of Technology Sydney, Australia. He has supervised 158 master's and 30 Ph.D. theses. He has authored over 900 papers in technical journals and international conferences. His research interests include energy optimization in smart/microgrids and in wireless sensor networks using data analytics and blockchain. He was a recipient of the Best University Teacher Award (BUTA'16) from the Higher Education Commission (HEC) of Pakistan, in 2016, and the Research Productivity Award (RPA'17) from the Pakistan Council for Science and Technology (PCST), in 2017. He is an Associate Editor of IEEE Access and an Editor of Sustainable Cities and Society.

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 also the Director of the Cyber Security Chair, CCIS. Previously, he worked as the Vice Dean of the Development and Quality at CCIS. He also worked as the Dean of the College of Computer and Information Sciences and the Head of the Academic Accreditation Council, Al Yamamah University. His research interests include mobile-pervasive computing and cyber security. 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.

MUHAMMAD ADIL received the M.S. degree in electrical engineering from the Department of Electrical and Computer Engineering, COMSATS University Islamabad (CUI), Islamabad, Pakistan, under the co-supervision of Prof. Nadeem Javaid. His research interests include data science, optimization, security and privacy, energy trading, and smart grid.

MUHAMMAD UMAR JAVED (Graduate Student Member, IEEE) received the bachelor's and master's degrees in electrical engineering from the Government College University Lahore, Lahore, Pakistan, in 2014 and 2018, respectively. He is currently pursuing the Ph.D. degree with the Communications Over Sensors (ComSens) Research Laboratory, Department of Computer Science, COMSATS University Islamabad, Islamabad, under the supervision of Prof. Nadeem Javaid. He has authored more than 20 research publications in international journals and conferences. His research interests include smart grid, electric vehicles, and blockchain.

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.