Research Article

ENERGY EXPLORATION & EXPLOITATION

Energy Exploration & Exploitation
2024, Vol. 42(5) 1684–1714
© The Author(s) 2024
DOI: 10.1177/01445987241255394
journals.sagepub.com/home/eea
Sage

Improving Electricity Theft Detection Using Electricity Information Collection System and Customers' Consumption Patterns

Asif Iqbal Kawoosa[1], Deepak Prashar[2], G R Anantha Raman[3], Anchit Bijalwan[4], Mohd Anul Haq[5], Mohammed Aleisa[5] and Abdullah Alenizi[6]

Abstract

Electricity theft detection (ETD) techniques employed to identify fraudulent consumers often fail to accurately pinpoint electricity thieves in real time. The patterns associated with electricity use are leveraged to identify anomalies indicative of electricity theft. However, challenges in the benchmark ETD include overfitting and a high incidence of false positives (FPs) resulting from incorrect usage patterns formed by considering only electricity consumption patterns without accounting for external factors that contribute to variations in normal consumption patterns. Further investigation is required to precisely detect instances of electricity theft, thereby minimizing non-technical losses and forecasting future electricity demand to maintain a stable supply. This study employs a master energy meter located on the transformer side to monitor the amount of energy distributed to the region. Zones with a high likelihood of energy theft are detected by comparing the sum of readings from all the smart meters with the readings from the master energy meter installed on the HV of the substation transformer within the same electric feeder. Ensemble

[1]School of Computer Applications, Lovely Professional University Punjab, Phagwara, Punjab, India
[2]School of Computer Sciences and Engineering, Lovely Professional University Punjab, Phagwara, Punjab, India
[3]Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India
[4]School of Computing and Innovative Technologies, British University Vietnam, Húng Yên, Vietnam
[5]Department of Computer Science, College of Computer and Information Sciences, Majmaah University, Al-Majmaah, Saudi Arabia
[6]Department of Information Technology, College of Computer and Information Sciences, Majmaah University, Al-Majmaah, Saudi Arabia

Corresponding author:
Mohd Anul Haq, Department of Computer Science, College of Computer and Information Sciences, Majmaah University, Al-Majmaah 11952, Saudi Arabia.
Email: m.anul@mu.edu.sa

XGBoost machine-learning algorithm and K-Means algorithm are used for the classification of malicious and nonmalicious samples and grouping similar types of consumers together, respectively. This makes the proposed model resistant to false-positive rates caused by changes in usage patterns that aren't done on purpose. Furthermore, energy thieves are identified by detecting anomalies in consumption behavior while maintaining constant internal and external environmental variables. This novel model proposed here mitigates the FP rate found in the present research of electricity usage data. Our approach outperforms support vector machines, convolution neural network, and logistic regression in simulations. Precision, F1-score, recall, Matthews Correlation Coefficient, Receiver Operating Characteristics (ROC)-Area Under The Curve (AUC), and Precision Recal (PR)-Area Under The Curve (AUC) validate our model. The simulation results show that the proposed K-Means-LSTM-XGBoost model improved the classifier's F1-score, precision, and recall to 93.75%, 95.16%, and 92.38%, respectively. Our model classifies huge time series data with high precision and can be utilized by the utility for real time electricity theft detection.

Keywords
Electricity theft detection, false positives, nontechnical losses, nonmalicious, ensemble algorithm

Introduction

Many electricity theft detection (ETD) methods have been tried by electricity utilities to avoid losing huge money. Many authors had proposed different methods for theft detection in their scientific articles (Messinis and Hatziargyriou, 2018). The ETD models are broadly classified as hardware-based and data-driven-based methods. State estimation, game theory, machine-learning-based methods use data driven approach to tackle electricity theft.

The aim of this study is to explore the existing research in the field of ETD and explore the landscape and factors in detecting electricity theft in challenging environments characterized by inconsistent power supplies and seasonal weather changes. The primary objective of this innovative approach is to reduce false-positive rates (FPR) in areas where traditional machine-learning models struggle. By focusing on challenging environments with varying power supplies and seasonal changes, the integrated model aims to improve the accuracy and efficiency of ETD. The proposed integrated model combines smart meter data, uses K-Means clustering, and employs an integrated ensemble techniques to assist electric utilities with limited field staff in identifying theft clusters. This approach aims to reduce FPR in areas where seasonal changes, voltage fluctuations, and scheduled and unscheduled electricity shutdowns affect regular electricity supply.

The machine-learning methods are most widely used methods because of efficient, accurate detection and are economically cost-effective. Hardware-based solutions include use of sensors, Radio Frequency Identification (RFID) tags, etc. have huge deployment costs and require frequent maintenance due to system failure caused due to moisture or heat in extreme cold and warm climatic conditions. The power system is defined by its states, that is, state of voltage magnitudes and phase angles in its bus system. Estimation method is a tool for checking the abnormalities in the system operators. The state estimator calculates the system state by taking sets of random and redundant readings and measurements of the system variables. In the paper (Wen et al., 2018) authors propose a scheme for ETD. The bias estimation is compared with the true bias of all the users. If the predefined values are satisfied, we consider the user has no anomaly in usage pattern. A recursive filter estimates the consumer's electricity theft based on the state estimation method for detecting abnormal data. Game theory approach is represented as a play-off between electricity thieves and electricity suppliers or distribution utility amidst the different adversaries. The approach uses electricity thief as a player who tries to steal an amount of electricity, and the distribution utility that tries to maximize the probability of detecting and catching the electricity thief and minimize its operational cost for regulating anomaly detection. The name game theory is because consumers and supplier both participate as players in the game, electricity thieves (fraudulent consumer) minimize the possibility of being caught and electricity suppliers (distribution utility) try to augment the chance of detection, and both are using different strategies to complete their tasks (Cárdenas et al., 2012). This approach fails to construct a robust plan to control the electricity thieves and regulate them. The authors in the paper (Cárdenas et al., 2012) formulate a game between the supplier and the malicious user, which uses a probability density function, and tries to achieve Nash equilibrium using energy meter measurements. Data-driven approach based on artificial Intelligence in Nagi et al. (2008) proposes the use of the machine-learning algorithms like support vector machines (SVM), recurrent neural networks (RNN), extreme learning machine (ELM), extra trees (ET), convolution neural network (CNN), and online sequential extreme learning machine (OSELM), and so on for ETD. The approach mentioned in Nizar et al. (2008) identifies the abnormal load consumption of a consumer using load-profile evaluation technique. The machine-learning-based ETD model uses real-world electricity consumption dataset to find the abnormal usage behavior of the consumer for detection of the electricity theft. Classification-based approaches often face a challenge with nonmalicious factors that change electricity consumption patterns. Change of people living in house, appliances, seasons, etc. If overlooked, these factors could raise FPR. False positives in ETDs can be very expensive because for levying final charges against theft, On-site inspection is needed as proof for energy theft. The amount of data about electricity usage has multiplied exponentially in recent years due to the expanding use of smart meters (SMs), and so on. A popular area of research at the moment is the data mining, analysis, and efficient use of big data (Chen, 2014; Zhou et al., 2017). Additionally, use of ensemble machine-learning algorithms where more than one machine-learning method is used in combination to give more accurate and precise result than a single machine learning algorithm. Ensemble models are essentially used to improve the efficiency and performance of classification problem.

The following is a summary of the key contributions of our work

The proposed model leverages smart meter data to utilize K-Means, long short-term memory (LSTM), recurrent neural network architecture, and the XGBoost ensemble method. Its purpose is to aid electric utilities with limited inspection staff in identifying clusters of electricity theft and, consequently, pinpointing the fraudulent consumers engaged in theft. This approach is designed to diminish FPR, particularly in areas where climate fluctuations significantly affect regular electricity supplies.

The model operates in two stages. Initially, a master energy meter is installed on the transformer side to aggregate smart meter readings for the total electricity usage in the area, revealing nontechnical losses (NTLs). Anomalies are detected using the LSTM-XGBoost ensemble machine-learning model to identify deviations from normal electricity consumption behavior. This model relies on consumption patterns and short-term usage forecasting, showcasing enhanced detection rates (DR) and low FPR compared to other classification-based ETD models.

The study utilizes electricity consumption dataset from state grid corporation of China (SGCC), Kashmir Power Distribution Corporation Limited (KPDCL) for experimentation and validation. To address unbalanced datasets, theft data are incorporated into the dataset after anomalies are detected over time, and records are saved as sensitive records. If a physical inspection confirms the anomaly as theft, these sensitive records are added as theft records; otherwise, they are added as normal data in the dataset until a sufficiently large dataset is prepared. The dataset is then divided into "benign" and "malicious fake attack" sets, and additional data attacks are synthetically added to balance the dataset. This approach effectively mitigates the FPR, enhances the DR, and identifies various types of stealing attacks.

K-Means clustering is employed to group consumers with similar electricity state and availability. Forming clusters based on similar parameters and environments significantly reduces the false-positive rate. Clustering aids in training the classifier to detect anomalies in cities or provinces experiencing erratic supply due to feeder overloading, high demand, low energy availability, and frequent faults during adverse weather conditions.

The other parts of the paper are section II which presents the related work about ETD. Section III that describes the methodology approach used for identification of theft to address the false positive rates (FPR) in machine learning-based ETD. Section IV discusses the methodology application phase of proposed K-Means-LSTM-XGBoost (K-MLX) model in detail. Comparison, Training, and selection of model is described in Section V. Evaluation of model is described in Section VI. Finally, conclusion is presented in Section VII.

Related work

Several machine-learning and statistical models have been proposed for ETD in the literature; however, they require manual feature engineering and specific domain knowledge to be effective. Since it is difficult to extract underlying characteristics of electricity consumption from one-dimensional data as most of the current models are employing only one-dimensional data (Smith, 2004). The authors of Wen et al. (2018) make an identical argument, that a lack of adequate features engineering in many existing machine-learning models causes the models to produce poor generalization outcomes. As an added challenge, the current model use dataset which is multidimensional. Therefore, it's a very challenging task to extract the most abstraction features from the high-dimensional data. Weak feature engineering generally leads to high FPR, which lowers the system's performance.

The researchers acknowledge the fact that there isn't any adequate feature engineering processes described in previous methodologies in Salinas et al. (2012), Cárdenas et al. (2012), Nagi et al. (2008), and Nizar et al. (2008). The feature engineering methodology demands extended time and subject understanding. In contrast, the auto encoder is employed in Cárdenas et al. (2012) to extract semantic information from high-dimensional electricity user behavior. It still needs to be enhanced in order to detect sophisticated attacks, such as zero-day attacks, with accuracy. According to the researchers of Muniz et al. (2009a) and Cody et al. (2015), in the literature, the characteristics related to electricity consumption are essentially made directly employing subject expertise. But, arbitrarily changing patterns in usage of electricity by consumers are insufficient for efficient pattern recognition and NTL detection for general users and particularly for industrial users.

The authors of Muniz et al. (2009b) and Krishna et al. (2015) emphasize that several earlier research use various machine and deep learning models for efficient ETD and feature building. But, none was able sustain long-term temporal connection of a consumer electricity consumption pattern for reliable ETD. It is also challenging to identify underlying patterns with one-dimensional power consumption data. In comparison, common machine-learning models in Lo and Ansari (2013) and Khoo and Cheng (2011) have poor detection capabilities and performance due to a variety of nonmalicious reasons. A semisupervised learning-based method for ETD is proposed in Amin et al. (2012). However, this still demands enhancement to increase DR and decrease FPR.

Table 1. Comparison of various ETD techniques.

Study	Proposed	Novelty	Results
Cody et al. (2015)	Decision tree for classifying the fraud cases of energy consumption	Existing models rely on manual feature engineering and domain knowledge, posing challenges in extracting characteristics from one-dimensional and high-dimensional data. Weak feature engineering leads to false positive rates (FPR).	Low accuracy in detection and high values of false postives
Wen et al. (2018); Cárdenas et al. (2012); Nagi et al. (2008); and Amin et al. (2012)	State Estimation Based Energy Theft Detection	Lack of sufficient feature engineering processes, which are time-consuming and require domain expertise.	High cost involved in installing and maintenance of equipment
Chen (2014)	Utilizes an autoencoder to extract semantic information from high-dimensional electricity user behavior.	Autoencoder for feature extraction.	Improvements needed for accurate detection of sophisticated attacks, such as zero-day attacks.
Lo and Ansari (2013); Khoo and Cheng (2011)	Detection of nontechnical losses using state estimation and analysis of variance.	Direct use of subject expertise in creating features related to electricity consumption.	Arbitrary changes in usage patterns deemed insufficient for efficient pattern recognition and nontechnical loss (NTL) detection, especially for industrial users.
Cárdenas et al. (2012)	A game theory model for ETD including privacy-aware control.	Use of various machine and deep learning models for efficient energy theft detection (ETD) and feature building.	Long-term temporal connections of consumer electricity consumption patterns not sustained for reliable theft detection. One-dimensional power consumption data complicates the identification of underlying patterns.
Asif et al. (2021)	A Hybrid Deep Learning Approach for Detecting Non-Technical Losses.	Common machine-learning models.	Poor detection capabilities and performance due to nonmalicious reasons.
Chen (2014)	Semisupervised	Semisupervised learning approach.	Requires enhancement to

(continued)

Table 1. Continued.

Study	Proposed	Novelty	Results
	learning-based method for ETD.		improve detection rate (DR) and reduce FPR.
Jokar et al. (2016); Angelos et al. (2011)	Consumption pattern-based LSTM-XGBoost energy theft detection model.	LSTM-XGBoost model for ETD. Calculates NTLs by comparing smart meter readings with transformer master meter readings. Trains model for each individual customer using historical data and synthetic attack data.	Addresses high false rate in classification-based ETDs.

Based on the table summarizing the research studies in Table 1, several technical gaps are identified in existing methods that the proposed K-MLX model aims to address.

Existing ETD models heavily rely on manual feature engineering, requiring-specific domain knowledge, leading to time-consuming and challenging processes. Current models struggle to extract meaningful characteristics from one-dimensional data, hindering the identification of sophisticated attack patterns. Inadequate feature engineering often results in poor generalization outcomes and impacts model effectiveness in real-world scenarios. Complexity arises from the use of multidimensional datasets, making it challenging to extract abstraction features and detect theft patterns. Limited efficiency in pattern recognition for NTL detection is due to insufficient feature engineering, particularly affecting industrial users. Furthermore, existing models fail to sustain long-term temporal connections in consumer electricity consumption patterns, crucial for reliable theft detection. Common machine-learning models exhibit poor detection capabilities, especially when anomalies are nonmalicious. Semisupervised learning-based methods in ETD require enhancement to improve DR and reduce FPR. Many existing machine-learning and statistical models require manual feature engineering and specific domain knowledge to achieve effectiveness, as highlighted in Cody et al. (2015) and Muniz et al. (2009b).

The proposed K-MLX model aims to address this gap by automatically extracting relevant features from the data using the LSTM component, reducing the need for manual feature engineering.

Challenges in extracting features from high-dimensional data: Studies (Cody et al., 2015; Muniz et al., 2009b) mention the difficulty in extracting underlying characteristics from high-dimensional data using one-dimensional approaches. The K-MLX model aims to tackle this challenge by leveraging the capability of LSTM networks to handle sequential, high-dimensional data and capture temporal patterns. Lack of long-term temporal pattern recognition: As mentioned in Amin et al. (2012) and Enguo and Xuan (2011), existing methods struggle to sustain long-term temporal connections of consumer electricity consumption patterns, which are crucial for reliable theft detection. The LSTM component in the proposed model is designed to capture and learn these long-term dependencies, enabling better recognition of temporal patterns in electricity consumption data. Inefficient handling of nonmalicious anomalies: Common machine-learning models mentioned in Jokar et al. (2016), Enguo and Xuan (2011), Yang et al. (2011), and Anas et al. (2012) exhibit poor performance due to their inability to distinguish between malicious and nonmalicious anomalies in electricity consumption data. The K-Means component in the proposed model aims to cluster customers based on their consumption patterns, enabling the LSTM-XGBoost model to learn customer-specific patterns and better differentiate between malicious and nonmalicious anomalies. High FPR and low DR: Studies like Chen (2014) highlight the need for improving DR and reducing FPR in energy theft detection models. The combination of K-Means clustering, LSTM feature extraction and XGBoost classification in the proposed model aims to address this gap by providing a more accurate and robust energy theft detection system. By addressing these technical gaps, the proposed K-MLX model aims to provide an improved and more effective energy theft detection solution compared to existing methods.

The current study proposes a consumption pattern-based LSTM-XGBoost energy theft detection model to address the high false rate in classification-based ETDs. The novelty lies in the LSTM-XGBoost model and the approach of calculating NTLs by comparing smart meter readings with transformer master meter readings, training the model for each individual customer using historical data and synthetic attack data. This research utilizes a pattern-based approach combining K-Means, LSTM, and XGBoost for detection of theft in electricity usage, employing innovative techniques to address the high FPR in detection rate. The model begins by calculating NTLs by comparing combined reading of usage reported by the SMs with the total amount of consumption measured by transformer master meter in each neighborhood. If a NTL is recognized at this level, consumers in the area who exhibit aberrant patterns will be picked as potential malicious users. Model is selected after training it for each individual customer by employing both the user's historical data and a synthetic attack data. After that, the classifier is selected and put to use to determine whether or not a new sample is malicious or benign.

This study employs the K-MLX to solve the classification problem and use the combined approach on the SGCC and Power Distribution Corporation Limited (PDCL) electricity consumption dataset and simulate the various attacks as per the scenarios of real-world to constitute supervised learning problem. The model is used for classification of data samples to detect anomaly in the consumer's electricity consumption behavior based on abnormal pattern. The primary objective of this innovative approach is to reduce FPR in areas where traditional machine-learning models struggle. By focusing on challenging environments with varying power supplies and seasonal changes, the integrated model aims to improve the accuracy and efficiency of ETD. The proposed integrated model combines smart meter data, uses K-Means clustering, and employs an integrated ensemble of LSTM-XGBoost algorithms to assist electric utilities with limited field staff in identifying theft clusters. This approach aims to reduce FPR in areas where seasonal changes, voltage fluctuations, and scheduled and unscheduled electricity shutdowns affect regular electricity supply.

Problem statement

Let the electricity usage dataset be represented by users X (X $= X_1, X_2, \ldots, X_i, \ldots .X_N) \in R^{N*T}$ for N users , where $X_i = (x_1, x_2, x_3, \ldots, x_T) \in R^T$ denotes the data for the ith user over a period of T. The aim of ETD is to utilize the proposed model B to determine whether a consumer is indulged in electricity theft or exhibiting abnormal usage behavior and this process is defined in Equation (1). The detection model B is a function that takes in data regarding a consumer's electricity usage and produces the detection result (Liu et al., 2023). In this context, $\hat{D}_i$ represents the result of model: 0 indicates normal users, and 1 indicates abnormal users. The optimization goal of the model is to minimize the difference between the actual result D_i and the detection model result $\hat{D}_i$ which is expressed in Equation (2) (Liu et al., 2023)

$$\hat{D}_i = B(X_i) \tag{1}$$

$$\min_{B} \sum_{i=1}^{N} |D - \hat{D}_i| \tag{2}$$

Algorithm of K-MLX

The proposed K-MLX model addresses gaps in existing energy theft detection methods. It combines K-Means clustering to group customers by consumption patterns, LSTM networks to extract temporal features from sequential data, and XGBoost for robust classification. This approach tackles challenges like manual feature engineering, handling high-dimensional data, recognizing long-term patterns, and reducing false positives (FPs). By clustering customers, the model captures customer-specific behaviors and anomalies. The LSTM extracts relevant features, overcoming manual engineering. XGBoost handles nonlinear relationships, improving accuracy. Results demonstrate higher DR, better precision-recall balance, lower FPs, and robustness to non-malicious anomalies, improved temporal pattern recognition, and scalability over traditional methods.

Mathematically the model can be shown as:

Step I: K-Means Clustering Objective Function:

$$\min_{c} \sum (i = 1 \ to \ k\Big) \sum (x \in C_i)\mathrm{x} \; - \; \mu_{\mathrm{i}}^{2} \tag{3}$$

Where k is the number of clusters, C_i is the set of data points in cluster i, and μ_i is the centroid of cluster i.

Step II: LSTM network equations defined below governs the flow of information within an LSTM unit, allowing it to learn and retain long-term dependencies in sequential data. This is useful for detecting patterns indicative of electricity theft in consumption data. We have:

$$f_t = \sigma(W_f.[h_{(t-1)}, \; x_t] + b_f \tag{4}$$

Equation (4) known as forget gate describes what information should be discarded or kept in the cell state. Here f_t is the forget gate at time step t, W_f is the weight matrix for the forget gate, σ is the sigmoid activation function and $h_{(t-1)}, \; x_{(t)}$ is the previous hidden state and input at time step t respectively whereas $b_{(f)}$ is bias for the forget gate.

$$i_t = \sigma[W_i.(h_{(t-1)}, \; x_t) + b_i] \tag{5}$$

$$\tilde{C}_t = \tan h[W_c.(h_{(t-1)}, \; x_t) + b_c] \tag{6}$$

Equation (5) tells us which new information to store in the cell state. Equation (6) is described as candidate cell state, creates a candidate vector that could be added to the state. In Equation (5), i_t is the input gate at time step t, W_i is the weight matrix for the input gate. Whereas as in Equation (6), $\tilde{C}_t$ is the candidate cell state at time t, tan$h$ is hyperbolic tangent activation function, W_c is the weight matrix for the candidate cell state.

$$C_t = f_t \odot C_{(t-1)} + i_t \ \odot \ \tilde{C}_t \tag{7}$$

Equation (7) combines the previous cell state with the candidate cell state to update the current cell state. Where $\odot$ is the element-wise cell multiplication.

$$O_t = \sigma[W_O.(h_{(t-1)}, \ x_t) + b_O] \tag{8}$$

In Equation (8), the LSTM decides what the next hidden state should be based on the cell state. Where O_t the output is gate and W_O is the weight matrix of output gate.

$$h_t = O_t \odot \tanh(\ C_{(t)}) \tag{9}$$

Equation (9) defines the output based on the current cell state and the output gate. Where h_t is the hidden state. In the above equations, W and b are the weight matrices and bias vectors respectively.

Step III: XGBoost Objective Function:

$$\mathrm{L}(\varnothing) \sum (i = 1 \ to \ n\Big) \ 1(\ y_i \ , \tilde{y}_i) \sum (k = 1 \ to \ k) \ \Omega(f_k) \tag{10}$$

In Equation (10), L is the loss function, y_i is the observed values, $\tilde{y}_i$ is the predicted values, Ω is the regularization term, and f_k are the individual decision trees in the ensemble.

This section discusses in detail how each segment of proposed model works. Figure 1 shows how the proposed method works from start to end. We argue that the permissible FPR depends on regionally varying parameters such as the theft rate. The steps executed in model to accomplish

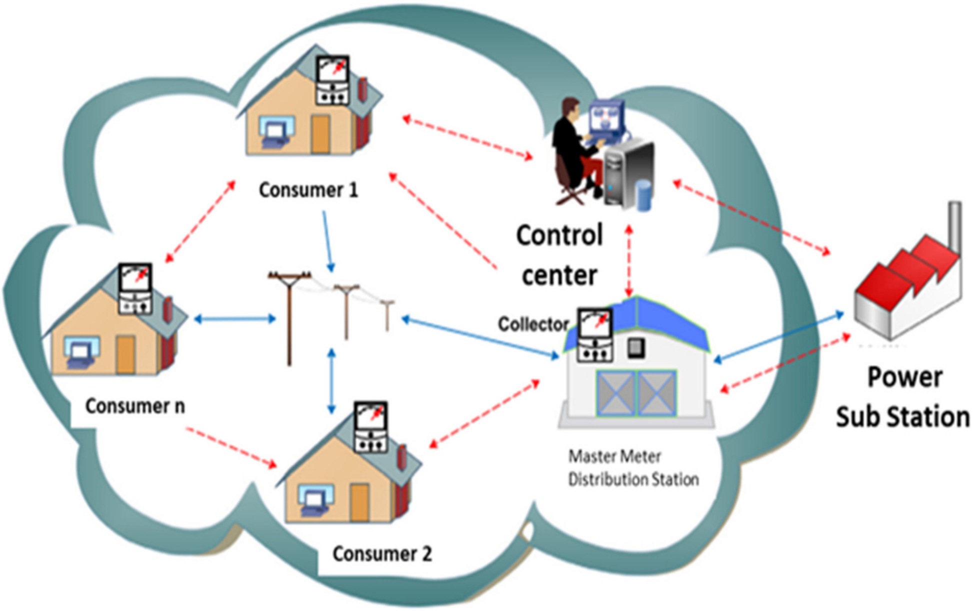


Figure 1. An overview of consumers smart meters connected to control center for precise anomaly detection.

our goal of ETD are based on ensemble ML method. The proposed model K-MLX is generally employed for detection of electricity theft on SMs data but more specific can be used to reduce FPs in the areas that experience harsher climatic conditions, erratic power supply and varying peak time demand during different seasons.

For data Preparation, interpolated median method is used for imputation to handle the missing values Three-sigma rule for removal of outliers. Min-Max method for quantifying categorical data in the dataset.

K-Means clustering is used for clustering of consumers where electricity is being provided in a similar way, considering those cities or provinces where supply is erratic due to feeder overloading/ high demand, and low availability of energy, or where electricity remains off due to frequent faults occurring due to wind, snow, rain, etc. A cluster is formed based on similarities in the availability and consumption of electricity.

Dataset is divided into two sub-sets one having malicious samples and another with genuine data. Synthetic data attacks are generated with six different types of attacks to remove the class imbalance, and over-fitting issues. For extracting important features and reducing the dimensionality of a high-dimensional electricity dataset, our model uses the LSTM technique.

Finally, the XGBoost ensemble method is used to determine the genuine and malicious samples.

To assess the model's effectiveness, interpretation and validation are carried out using the symmetric Mean Absolute Percentage Error (sMAPE) algorithm and standard metrics, and then compared to existing ML techniques such as k-NN, Extra Trees, MLP, Random Forest, SVM, logistic regression, AdaBoost, and so on. The results show that the proposed KSLX model outperforms existing techniques for detecting and removing FPs. After looking at the data of customers who stole energy and were penalized for it, the proposed model was 83% accurate.

Methodology proposal

The study's key contributions are as follows:

- Development of a novel hybrid model that integrates K-Means clustering, LSTM networks, and XGBoost for classification of theft and nontheft cases.
- Utilization of K-Means clustering to effectively segment customers based on their consumption patterns with similar features.
- Implementation of LSTM networks for extraction of temporal features from sequential data.
- Employment of XGBoost for robust classification, enabling the model to handle complex patterns and reduce FPs.
- Performing a thorough analysis to showcase the better performance of the proposed model over existing techniques, with potential for generalizing the model to other domains involving time-series data analysis.

The proposed model could be employed to identify potential energy theft instances by leveraging K-Means clustering for capturing consumption patterns, LSTM networks for temporal feature extraction, and XGBoost for robust classification. The model's ability to handle high-dimensional data and capture long-term dependencies could prove advantageous in detecting sophisticated theft techniques. By deploying the model, the utility company could reduce revenue losses, enhance operational efficiency, and promote responsible energy consumption practices. The insights gained from the model could inform targeted inspection and enforcement strategies, further deterring energy theft.

A significant amount of metering data collected from SMs is actually used in order to analyze the consumption behavior of a consumer. The data utilized for analysis includes electrical power parameters i.e., voltage, current, power factor and nontechnical parameters daily/weekly/monthly usage details of electricity, connected sub-station information, connected feeder information, GIS information, and so on are all included in the electricity data for analysis (Enguo and Xuan, 2011) as shown in Figure 1. This study takes the transmission losses constant in all the clusters under study.

The abnormality detection based on analysis of electricity consumption behavior of consumer consists of the three primary components: the collection of user consumption data, the processing of user electricity data, and finally the analysis and verification of the user detection results (Messinis and Hatziargyriou, 2018; Yang et al., 2011; Anas et al., 2012).

Dataset details

(i) Training Phase: This study puts to use the SGCC dataset, which is publicly available at (http://www.sgcc.com.cn/) and is a real electricity consumption dataset for China. Another dataset PDCL was used and recreated by injecting the theft data. The SGCC dataset comprises 38,752 honest and 3615 fraudulent consumers for a total of 42,372 users shown in Table 2.

Apart from dataset of electricity consumption from SMs, including both normal and anomalous (energy theft) consumption patterns. Auxiliary dataset like weather data and GIS data is also used. The dataset cover a diverse range of residential and industrial customers, spanning different time periods (e.g., daily, weekly, monthly). The dataset is properly labeled, with instances marked as either normal or energy theft cases.

(ii) Data preprocessing: The dataset of electricity consumption often has erroneous and missing values because of problems with the energy meters, cyber-attacks, servicing, data transfer,

Table 2. Dataset details.

Number of users	42,372
Number of fraudulent users	3615
Number of users who use electricity genuinely	38,752

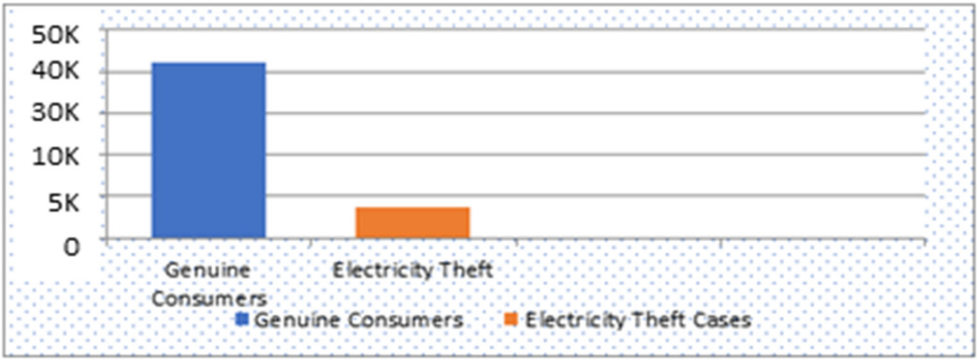


Figure 2. Class imbalance between theft and nontheft data in the real-world dataset.

and storage issues. Erroneous and noisy data has a poor impact on the model's performance. In this paper, the appropriate preprocessing techniques for specific problems are used (Asif et al., 2021). Linear interpolation method is used for filling in the missing values. It takes the average values of electricity consumption on the day before the current day and on the next day of the current day. Noise and outliers are also needed to deal so as to work the model well. In this paper, the "three-sigma rule of thumb" (TSR) is used to deal with the outliers. Finally, the data needs to be normalized to avoid poor performance of classification ensemble/ deep learning models on different forms of data (Asif et al., 2021). So Min-Max method is used for normalization. Set the range of values to similar metrics. Data from the user's electricity usage is normalized using a Min-Max technique. The split in the dataset into training, validation, and testing sets, ensure a representative distribution of normal and energy theft instances in each set.

(iii) Clustering the consumers: We determine the optimal number of clusters (k) for grouping customers based on their consumption patterns, using techniques called K-Means. K-Means clustering algorithm is used in training data to obtain customer clusters. The purpose of the K-Means algorithm is to group people with similar consumption patterns (Jokar et al., 2015). The proposed model can reduce FPs more accurately where consumers have similar electricity consumption and are supplied power for the same amount of time from the common grid station. This clustering can be formed among the electricity consumers that are fed by the same electric feeder, substation, area, or city because electricity consumption behavior is affected by various reasons, like erratic power supply due to feeder overloading/ high demand but lesser availability, or where electricity remains off due to faults arising from harsher climatic conditions like wind, snow, rain, etc. It can be used to cluster consumers based on the similarities in the availability and consumption of electricity. This type of clustering will reduce the frequency of FPs in ETD where the feeder or sub-station is shut off due to reasons like balancing the peak load demand and feeder tripping due to line faults or overloading, or erratic power supplied to the neighborhood area network (NAN) due to various climatic conditions. The electricity may also remain cut off due to the overloading of grid stations/substations/feeders. As after the restoration of electricity, the consumer's usage pattern increases or changes in a way that isn't normal and that may lead to false anomaly detection

Table 3. Shows the features extracted from the auxiliary databases.

Type of data	Type of Features
Tariff summary	Monthly EC "consumption drop" during last 1-year moving windowThe ratio of monthly ECs to contractual power.Minimum, maximum, standard deviation, and slope of a monthly ECs linear model.
Geographic information system	Latitude, longitude, town, village, smart metered/digital metered/EM metered/ nonmetered, 100% metered feeder/partially metered feeder.Percentage of theft detected on a radius from 1 to 10 km.Number of inspections, last inspection, inspections with/without anomalies and electricity theft detected before last inspection.
Technological characteristics	Type and model of smart meter (SM)Installation location of Smart Meter (inside/ outside house).Closed case combined SM's or separate, date of fabrication.
Contract database	Commercial/noncommercial, contracted power, tariff, number of complaints 2–24 months

by the machine-learning classifier. The machine-learning model needs to be trained so that it sees this pattern as normal in this kind of external environment. As the pattern changes almost for the whole cluster, a machine-learning model can learn by analyzing the overall changes in a cluster. K-Means clustering helps to group users who are connected to the same feeder and station. This way, the relative amount of power used by each cluster stays the same compared to the other clusters. K-Means algorithm maximizes the similarity in the cluster to reduce the FP rate.

Removal of Class Imbalance: The next step is to prepare a dataset for training on various classifiers. Using historical data, it's easy to obtain benign samples for each customer. However, malicious or anomalous samples may not be available easily because the stealing of electricity doesn't happen more often as can been seen in Figure 2.

One way to deal with the problem of imbalanced data is to use single-class classification techniques, in which the classifier is trained using only normal samples. But, the one-class classifier doesn't work well for this type of problems. Density function approximation methods, like the ones in (Jokar et al., 2016), are another way to solve the problem. But these methods only work if the approximation function can model the data accurately, which might be hard to do in real life. Instead, we want to use the benign samples to make a set of malicious samples. The goal of the stealing is to report less energy use than the real amount or to move high usage to times when prices are low. So, in this study, the malicious samples are artificially created taking the reference of benign ones. We assume $\boldsymbol{x} = (x_{1,}x_{2}, x_{3,.....,}x_{n})$ a vector of genuine consumption values for a 24-h period with n samples, and $x \boldsymbol{\epsilon} X$, in which X is a random vector and having P0 distribution. The utility will compute the electricity consumption on the values say $\boldsymbol{y} = (y_{1,}y_{2}, y_{3,.....,}y_{n})$ from the meter readings as considered in (Jokar et al., 2015). In case of honest customers, we have $y = x$, but for fraudsters, $y = h(x)$, where $y \boldsymbol{\epsilon} Y$ again here Y is a random vector and is having a P1 distribution, so that $E[Y] \leq E[X]$. It is possible to figure out h (.) by studying various energy theft scenarios and their effect on values that have been measured. For example, if $h(x) = \boldsymbol{\alpha}$x, then 0 <= $\boldsymbol{\alpha}$ <= 1 is a possibility. So, using the benign dataset to make malicious samples is a smart option. Even though it might not be possible to define

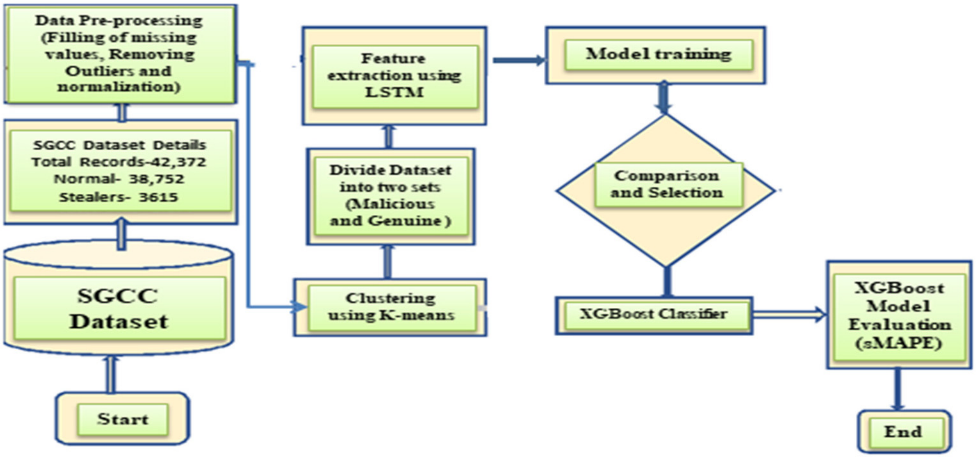

Data Pre-processing (Filling of missing values, Removing Outliers and normalization)
SGCC Dataset Details Total Records-42,372 Normal- 38,752 Stealers- 3615
SGCC Dataset
Start
Feature extraction using LSTM
Divide Dataset into two sets (Malicious and Genuine)
Clustering using K-means
Model training
Comparison and Selection
XGBoost Classifier
XGBoost Model Evaluation (sMAPE)
End

Figure 3. Flowchart of the proposed model for electricity theft detection (ETD).

all functions that lead to $E[Y] \leq E[X]$, a complete set of attack samples can be made by looking at different situations and using the generalization property of classifier. The various types of attacks defined as Type 1 to Type 6 in (Jokar et al., 2016) are used to generate theft data for balancing are:

A Type 1 is that attack where a smart meter's reading is multiplied by the same parameter (say $\boldsymbol{\alpha} t$) all the time. This attack is known as a "scaling attack" (Jokar et al., 2015) and it manipulates the readings of a smart meter by multiplying the readings by a constant factor ($\boldsymbol{\alpha} t$).

A Type 2 is that type of attack where smart meter's reading is multiplied by a different random number at different times (say $\boldsymbol{\alpha} t$). This attack is known as a "random scaling attack" (Jokar et al., 2015) and the attacker multiplies the readings of a smart meter by a different random factor ($\boldsymbol{\alpha} t$) at different times. This type of attack is harder-to-detect as compared to a simple scaling attack. Its impact is similar to a scaling attack, but is more complex to be detected.

A Type 3 attack is that type of attack, where a smart meter sends only half of the actual readings, i.e., $0.5 \times x$ (half of the actual load x) during peak load time, i.e., night hours during the winter, and actual load during off-peak hours. This type of attack is known as a "load shifting attack" (Jokar et al., 2015). In a load-shifting attack, the attacker changes the readings during peak load times and sends the real readings when the system isn't overloaded. This makes it hard to detect if someone is stealing electricity because the average load during peak hours looks normal. This kind of theft attack can be detected by using advanced machine-learning algorithms that analyze the use of electricity on a seasonal and hourly basis.

A Type 4 attack is a form of electricity theft in which the average value of the consumption reading is multiplied by a random factor ($\boldsymbol{\alpha} t$). It is known as a "random offset attack" (Jokar et al., 2015) where the attacker manipulates the smart meter readings by sending a reading that is the average consumption multiplied by a random factor ($\boldsymbol{\alpha} t$). This type of attack is more complex to detect than a simple scaling attack because the readings appear to be plausible but are manipulated or incorrect.

A Type 5 energy theft attack is an attack where the energy meters send the average value of energy consumption to the control center of the utility. This type of attack is known as a baseline attack (Jokar et al., 2015).

A Type 6 electricity theft attack, also known as a reverse order attack is an attack, where fraudulent consumers send the readings of their consumption in reverse order to the utility during the day (Jokar et al., 2015). During this kind of attack, the highest readings taken during peak hours are sent out during off-peak hours, and the lowest readings are sent out during peak hours. This type of attack is typically done to avoid high dynamic pricing during peak hours. The impact of this type of attack can be difficult to detect because the readings appear to be plausible.

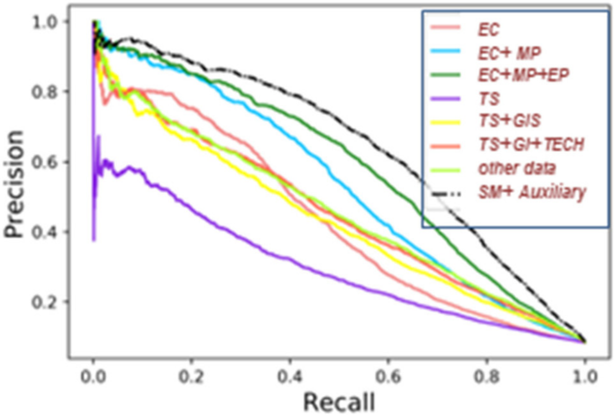
Table 4. Details of dataset before and after data preprocessing.

Number of anomalous users before applying data balancing.	3615
Number of benign users before applying data balancing.	38752
Number of abnormal users after applying data balancing.	18,279
Number of benign users after applying data balancing.	18,281
Number of consumers before initial preprocessing of data.	42,372
Number of consumers after initial preprocessing of data.	42,367

(iv) Feature extraction and dimensionality reduction: For each customer cluster, training is done using LSTM network on the time-series consumption data of customers in that cluster. The trained LSTM networks is used to extract relevant features from the sequential data, capturing temporal patterns and dependencies. Each data vector in the dataset contains a customer's electricity readings over a 24-h period; for example, for n measurements per hour, the data vector has 24 × n components. There are several methods for dimension reduction that minimizes performance degradation by extracting only useful data features. The unusual data information on energy consumption caused by electricity theft is not always a single parameter, but a situation of electricity theft would cause several abnormal phenomena at the same time. In order to carry out theft detection, extensive features are required to be extracted from a variety of abnormal electricity consumption phenomena and quantified characteristics induced by a variety of abnormal energy consumption behavior (Angelos et al., 2011; Messinis and Hatziargyriou, 2018).

Features extracted from auxiliary databases: The features extracted from the auxiliary databases like tariff summary (TS) database, which contains the monthly Energy Consumption (EC), maximum power usage in six different tariffs. The geographic information system (GIS) provides information about customers' location and the rate of NTL in the neighborhood. Also, the information about the technological characteristics like the smart meter details, its brand, its type and its location of installation is extracted from TECH database, the contract event information is also found from auxiliary databases like contract database. The contracts event in database also offers information about the activity type of the customer. The auxiliary dataset like GIS and weather data can be used to strengthen feature set for training the classifier (Table 3).

Long short term memory (LSTM), a RNN-based DL is used to strengthen the feature extraction. LSTM (Adil et al., 2020) extracts key features as the electricity consumption adds records every fifteen minutes and makes it huge. So, dimensionality reduction is much needed in these huge datasets like the dataset. If a conventional recurrent neural network (RNN) (Rumelhart et al., 1986) is used for dimensionality reduction, it encounters issues of vanishing gradient and exploding gradient. For exploding and vanishing gradients, LSTM, an advanced RNN variant, is widely used (Pamir et al., 2022). After the RNN finds the time correlation between both the input data and the prior data during training, the output is



*Smart meter data
(EC- Electricity Consumption;
MP- meter parameters;
EP- electricity parameters)
Auxiliary data
(TS-Tariff summary;
GIS- geographic info sys;
TECH- technical characteristics;
CONT- Contract summary).

Figure 4. Precision Recall (PR)-Recall curves on various feature-set of dataset.

completed. As its memory is temporary and short-lived, it can't figure out how the current and previous states are related in time. However, LSTM captures temporal correlation smoothly (Pamir et al., 2022). It reduces the dimensionality of vast time series data. Its unique memory cells use historical data. LSTM memorizes significant features from huge time series data. This is saved and kept by the long-term memory (cell state) of LSTM. These important parts of the dataset hold the most important information.

Training the model: In this study, electricity theft is detected in two stages. One is the installation of a master energy meter at the substation transformer to measure the difference in energy supplied and revenue generated by the total number of SMs installed in the neighborhood, and the other is the application of a machine-learning algorithm to the cluster to detect anomalies in the consumer's electricity usage.

Step I: The master energy meter installed on sub-station transformer measure the total amount of electricity supplied to customers in the neighborhood. Electricity supplied through the master EM is compared to sum of electricity consumed by consumers and recorded by their SMs installed and fed by the sub-station transformer equal to $\sum_i E_{SM_i}(t)_{MM}$. Anomaly is identified if, at any stage of a day $E_{MM}(t) > \sum_i E_{SM_i}(t) + E_{TL}(t) + \varepsilon$, where ε is the error in calculating TL. Every time a new sample is taken, this test is conducted afresh.

Step II: If Step 1 fails to detect anomaly in new sample and treat the sample as benign, the benign dataset is updated by adding the sample into it and the corresponding attack sample is made to add in the attack dataset.

Step III: If the classifier detected the attack in Step 1, the classifier detects consumer's SM anomaly in usage. When the anomaly is detected m number of times in a certain amount of time, this is treated as an indication of energy theft. During this study time of SM, all the new samples are temporarily stored in a temporary database. The onsite inspection could also be conducted to verify the energy theft. The onsite inspected is scheduled on priority if, the NTL is high in the smart meter and if many SM's in an area are having higher NTL. Further, if a theft is positive, the samples stored temporarily in another buffer database is transferred into the malicious database. If not, these attack patterns are not transferred into the malicious dataset but are added to the benign database.

Step IV: Another possibility is that the machine-learning classifier finds an anomaly, but Step I didn't find the NTL. Such a concern might happen by any one of three factors. It could be

Table 5. Hyper-parameters for K-NN model.

Hyperparameter	Range of values
K	2,4,8,16
P	2,3

Table 6. Hyper-parameters for binary LR model.

Hyperparameter	Range of values
C	0.001, 0.01,10,100
R	L1 norm, L2 norm

because the classifier has misclassified the anomaly or the NTL calculation isn't correct, but this cannot happen for many days in a row. Thirdly, it could happen when consumers have alterations in consumption habits because of changes in the use of electrical appliances that can cause changes in electricity consumption. So, when the classifier finds an anomaly or outlier but master energy meter in Step I shows no signs of NTL, this sample is stored temporarily in a temporary database. If this happens for long time over the next few days, the old dataset is replaced by this temporary dataset. The classifier is re-trained on the updated dataset if it is large enough. Every smart meter has a unique binary variable called a credibility factor (CF), which is initially set to 1. The CFI is reset to one when the problem has been rectified and zeroed out when a non-malicious anomaly is discovered, as explained above. Smart meters with CFI = 1 are more likely to take action when the algorithm determines that energy is being stolen. Through this algorithmic process, K-MLXS becomes robust against accidental shifts in consumption habits.

Step V: Another possibility is if the anomaly is identified by master meter, (which calculates the difference in the readings supplied at transformer levels and total sum of the reading of all the individual SMs connected through that transformer to compute the amount of NTLs considering the amount of the TL losses as constant) but machine-learning algorithm does not detect the theft attack and it continues to be like that for many samples over time. In this scenario, when the benign customer dataset is examined for evidence of data contamination attack, here, the adversary gradually modifies data and contaminates the dataset to deceive the machine-learning model by treating the theft data as normal data. In this case, consumer is studied for consumption over time. A declining long-term electricity usage curve indicates contamination attack and the model raises alarm, even if the machine-learning algorithm has not detected the contaminated attack on the historic data. Also, the model continues its normal activities for new sample in this scenario. This type of attack happens rarely if a new heavy consumption load is directly linked to a line.

Methodology application phase

The suggested K-MLX model is divided into five stages: (1) data preparation, (2) data balancing, (3) dimensionality reduction and feature extraction, (4) classification, and finally (5) validation as shown in Figure 3. This section describes the basic structure of the model architecture as well as the execution steps. This model elucidates the use of the supervised ML method for the detection of anomaly in the

Table 7. Hyper-parameters for SVM model.

Hyperparameter	Range of values
C	0.001, 0.01,10,100
Kernel	Linear, radial basis function

Table 8. Hyper-parameters for XGBoost model.

Hyperparameter	Range of values
Number of trees	1000, 2000
Learning rate	0.01, 0.1
Maximum depth	7, 15
Minimum child weight	1, 10

electricity consumption of a consumer. Dataset used “SGCC” and PDCL electricity consumption dataset publicly available and for experimentation and validation of results respectively.

1. Data Preparation: Data preprocessing is necessary for the identification of anomalies in the electricity dataset because raw data often contains errors, outliers, missing values, and irrelevant information. Preprocessing helps to clean, transform, and standardize the data, making it more suitable for analysis and modeling. This makes the results more accurate and reliable, making it easier to identify people who steal electricity. Also, preprocessing can help reduce the number of dimensions in the data and solve the problem of class imbalance, both of which come up often when analyzing data on electricity usage. The removal of imbalance in the cases of theft and nontheft after applying the data balancing technique is shown in Table 4.
2. Filling of missing values: Linear interpolation is a method that can be used to fill missing values in an electricity consumption dataset (Jokar et al., 2015). It estimates the missing readings based on the values of the other points in the dataset. Linear interpolation method uses a simple linear equation to estimate missing values based on the values of the other points in the dataset.

   Given two data points ($x1$, $y1$) and ($x2$, $y2$), the linear equation can be represented as:

   $y = mx + b$, where $m = (y2–\mathrm{y}1)/(x2–x1)$ is the slope of the line and $b = y1–mx1$ is the y intercept.

   To fill a missing value, we can use the linear equation to estimate the value based on the values of the two surrounding data consumption points (Jokar et al., 2016). For example, given a missing value at x, we can use the linear equation to estimate y as:

$$y = m(x - x1) + y1.$$

Table 9. Proposed and other benchmark models tuned parameters and execution time.

Model	Hyperparameter	Range of values
K-NN	K	2,4,8,16
	P	2,3
Logistic regression	C	0.001, 0.01,10,100
	R	L1 norm, L2 norm
SVM	C	0.001, 0.01,10,100
	Kernel	Linear, Radial basis function
XGBoost	Number of trees	1000, 2000
	Learning rate	0.01, 0.1
	Maximum depth	7, 15
	Minimum child weight	1, 10

Table 10. Showing the execution time of using LSTM with existing model and proposed model.

Model name	Input batch size	Execution time existing model	Execution time proposed model
K-NN	50–300	218	90
Logistic Regression	50–300	165	88
SVM	50–300	159	87
XGBoost	5-300	152	82

The value of m can be determined from the values of the two surrounding points, and the value of y can be estimated based on the value of x. This process is repeated for each missing value in the dataset, using the values of the surrounding points to estimate the missing values.

3. Outlier detection: The TSR is a statistical method used to identify outliers in a dataset (Jokar et al., 2016). The basic idea behind the TSR is that a data point can be considered an outlier if it lies more than three standard deviations from the mean of the data.

   Mathematically, the TSR can be expressed as follows:

   - Let x be the data point in question
   - Let μ be the mean of the data
   - Let σ be the standard deviation of the data

   If $|x - \mu| > 3\sigma$, then the data point x is considered an outlier.

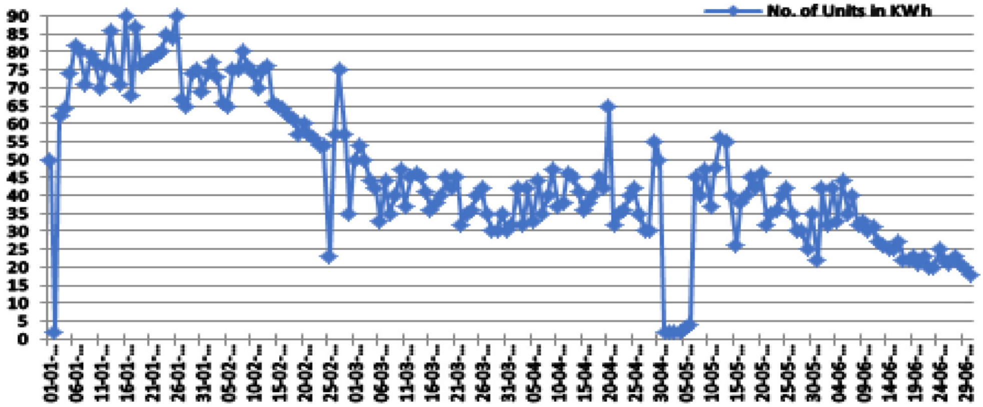


Figure 5. Typical energy usage pattern of a genuine consumer over six months, with occasional spikes in both low and high consumption for valid reasons.

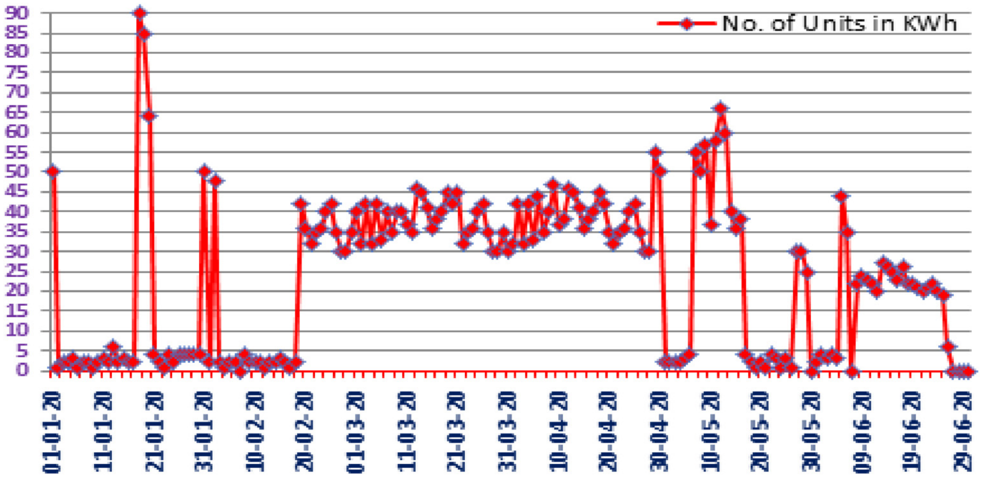


Figure 6. Energy usage pattern of a malicious consumer, featuring sporadic low consumption periods and intermittent spikes.

The TSR is a simple way to identify outliers in a dataset and is a widely used rule of thumb for identifying outliers; It can be a useful tool for quickly identifying outliers in large or complex datasets.

4. Clustering using K-Means algorithm**:** The K-Means algorithm is used in this study for grouping electricity consumers based on their consumption patterns within same city. The first step in the K-Means is to initialize the k cluster centroids randomly in the areas feed from the same feeder or a transformer. After preprocessing of the data, the number of clusters is determined using the silhouette score based on their electricity consumption KWh values under similar conditions. Clustering is accomplished by assigning each consumer to the nearest centroid iteratively based on their consumption values; Centroids are re-calculated iteratively based on the means of the assigned data values assigned to each cluster. This is done by taking the mean of all the data values in each cluster, and then updating the cluster centroid to the mean.

   This process is repeated until convergence or a maximum number of iterations are reached. The resulting clusters are analyzed to examine the patterns in the electricity consumption data that led to the formation of each cluster. The results of the K-Means algorithm are verified through a visual inspection before conclusions are drawn about the patterns in the electricity consumption data

   The K-Means algorithm can be represented by the following formula:

   $C(i) = 1/n \times \Sigma(\text{j} = 1$ to $n)$ $x(j)$, where $C(i)$ is the cluster centroid for the ith cluster, n is the number of data points in the cluster, and $x(j)$ is the jth data point in the cluster

   The root of squared distances between the data points and centroids are added and calculated as mentioned in Equation (11):

$$D(i,j) = (x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \ldots\ldots\ldots + (x_{in} - x_{jn})^2 \tag{11}$$

5. Feature extraction and dimensionality reduction using LSTM: LSTMs module consists of three gates, these are input (*xt*), output (*ot*), and forget (*ft*) gates. The forget (*ft*) gate selects whether the cell should keep or discard the current input (*xt*) and the hidden values (*ht*1) of previous state. The sigmoid (*s*) activation function uses *ht*1 and *xt* to decide whether to keep or discard the cell state's previous output by selecting 1 or 0. The input gate selects data to change memory or cell states (*Ct*). The second sigmoid repeats the details from *ht*1 and *xt* and decides whether to keep or discard them. The cell state (*Ct*) uses the tanh activation function on *ht*1 and *xt*, for storing the result. Adding the multiplication of the "cell state, input gate, forget gate, and cell state outcomes" gives an updated cell state (*Ct*0). Finally, *Ct*0 gets updated. The final output gate (*ot*) conducts the sigmoid function on *xt* and *ht*1 and stores the result. Multiplying tan*h*(*Ct*0) and ot yields the next hidden state (*ht*). The product of the multiplication is applied to the sigmoid function and stored in the variable *ht*. Equations (12)–(17) elucidate the mathematical expressions for the forget, input, and output gates (Javaid, 2021; Wu et al., 2022).

$$ft = s[Wf(xt, ht - 1)\ + bf] \tag{12}$$

$$it = s[Wi(xt, ht - 1)\ + bi] \tag{13}$$

$$Ct = tanh[Wc(xt, ht - 1)], \tag{14}$$

$$Ct0 =\ [ft * (Ct)]\ +\ [it * (Ct)], \tag{15}$$

$$ot = s[Wo(xt, ht - 1)\ + bo], \tag{16}$$

$$ht = s[ot * tanh(Ct0)]. \tag{17}$$

where *bo*, *bi*, and *bf* are the biases of output, input, and forget gate, *Wi*, *Wo*, and *Wf* are input, output and forget gate weights respectively. *Ct* and *Ct*0 represent past concealed state information and updated cell state information, respectively. LSTM feature extraction requires optimum hyperparameter settings. Two LSTM, two LeakyReLU, one batch normalization, and a single dropout layer comprise our LSTM feature extractor. The dropout probability

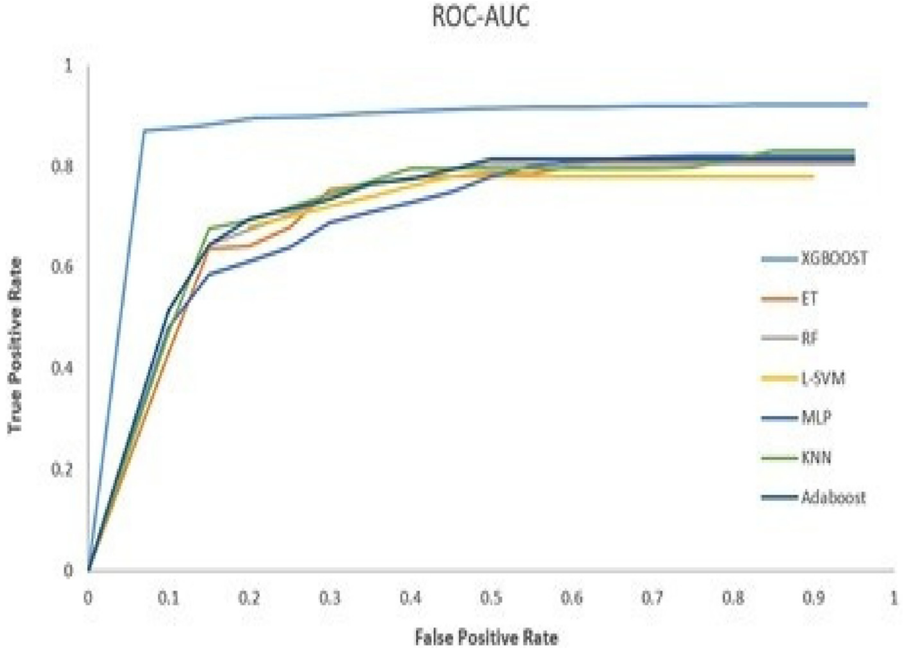


Figure 7. ROC of proposed XGBoost and other various models in electricity theft detection (ETD).

Table 11. Performance evaluation of the proposed KM-LX model using 10-fold cross-validation (CV).

Number of folds	Accuracy	Recall	Precision	F1 score	Kappa	MCC
1	0.9312	0.9226	0.9479	0.9345	0.8619	0.8622
2	0.9343	0.9278	0.9500	0.9388	0.8705	0.8708
3	0.9334	0.9229	0.9536	0.9385	0.8706	0.8711
4	0.9321	0.9196	0.9524	0.9357	0.865	0.8656
5	0.9370	0.9263	0.9542	0.94	0.8736	0.8741
6	0.9390	0.9292	0.9552	0.942	0.8777	0.8781
7	0.9285	0.9191	0.9454	0.9321	0.8567	0.8571
8	0.9331	0.9258	0.9476	0.9366	0.8659	0.8661
9	0.9313	0.923	0.947	0.9348	0.8623	0.8626
10	0.9344	0.9206	0.9548	0.9374	0.8686	0.8692
Mean	0.9338	0.9237	0.9508	0.9371	0.8673	0.8677
Std. Dev	0.0092	0.0033	0.0035	0.0028	0.0059	0.0059

for the dropout layer is 0.2, the learning rate (LR $= \alpha$) for both LeakyReLU layers is 0.001, and the first LSTM layer has 200 neurons (Figure 4).

Comparisons, training and selection of the model

A 10-fold nested cross-validation was applied for model selection and evaluation. The model selection of hyper-parameters was done with the help of SGCC and PDCL customers. The model was fitted using benchmark ETD models: random forest, SVM, logistic regression (LR), k-nearest neighbors (K-NN), and LSTM using the Scikit-Learn toolkit. The fitting of the model with XGBoost (Liu et al., 2023) was done by using grid-search method of Python API Scikit Learn.

Model comparisons

1. K-N: K-NN is the easiest algorithm to group things. At test time, K-NN uses training data to find the nearest neighbors. In ETD problems, the algorithm begins by using the previous theft data confirmed by field inspections or other means to help in finding anomaly in the usage behavior of the new consumer. So, to check out the likelihood of electricity theft in the new customer, the results of the previously detected fraudulent consumer's record of the nearest neighbor are averaged. Using $k = 16$ (neighbors) and $p = 2$ (power parameter set to 2) equal to the Euclidean distance, Table 5 shows the grid-search hyper parameters for the best results.
2. Binary LR algorithm: The sigmoid function used in LR classification takes input features X in the form of the matrix and multiplies each value with a weight matrix theta (Θ) as $g(z) = \Theta$, where $z = \Theta$ The LIBLINEAR solver (Yan and Wen, 2021) was used to train the classifier on a logarithmic loss function. Table 6 shows the hyper parameters that LR uses when it does a grid search. The C-value of hyper parameter represents inverse of the regularization strength, and avoids overfitting of model during its training. The regularization type is shown by the R hyper-parameter. The validation proves that a C of 0.01 and an L2 regularization produced the best results.
3. Support vector machines: SVM predicts decision values Instead of estimating probabilities,. To maximize the margin plane separating the two vector classes, the input characteristics are placed in a high-dimensional space and an SVM approach is applied (Yan and Wen, 2021). This gap is established by comparing the two classes' support vectors. The training dataset customer examples most similar to the decision function are the ones used to create the support vectors. The

Table 12. Displays the classification performance metrics, including f1-score, precision, recall, AUC-ROC, and PR-AUC, of the proposed model with the existing top models.

Classifiers	Precision	Recall	F1-Score	AUC-ROC	PR-AUC
K-NN	0.65	0.52	0.53	0.64	0.16
ExtraTrees	0.56	0.56	0.56	0.58	0.11
MLP	0.73	0.56	0.58	0.77	0.32
Random forest	0.90	0.93	0.84	0.80	0.84
SVM	0.88	0.84	0.87	0.87	0.72
Logistic Regression	0.85	0.80	0.79	0.78	0.81
AdaBoost	0.88	0.84	0.85	0.88	0.83
Proposed k-MLX	**0.98**	**0.93**	**0.96**	**0.96**	**0.94**

They are bold as they represents the parameters for the evaluation.

regularization strength is represented by the hyperparameter C, which is analogous to the LR parameter. The kernel parameter is used to create a hyperplane separation between nonlinear customer groups in high dimensional space. We found that a $C = 0.001$ and a linear kernel produced the best validation fold results. Grid-search SVM hyperparameters are displayed in Table 7.

4. Extreme Gradient Boosted (XGBoost) Trees: XGBoost algorithm employs gradient boosting (Zhou et al., 2015) in conjunction with a regularized cost-function. GB creates a model by additive integration of the predictions of a large number of “weak” classifiers. The model uses a regression tree which starts the training on a single regression tree. This single regression tree is in search for some rules that can separate consumers by some best possible ways with and without abnormalities. Following the construction of the initial tree, it builds a new regression tree on new round of training. During a new training round, it searches for those areas where the prior tree’s results were not correct or had an error in prediction and then it develops a new tree that corrects the previous tree’s mistakes making new set of rules.

Selection of a classifier

For each customer cluster, XGBoost classifier is trained to use the LSTM-extracted features as input and the labeled instances (normal or energy theft) as targets. Also XGBoost hyperparameters are optimized (e.g., learning rate, maximum depth, regularization parameters) using grid search technique.

XGBoost, a tree-boosting system is used on a huge datasets like electricity consumption dataset (Zhou et al., 2015) in which XGBoost constructs a model based on the training data, and determines the input data as malicious or benign. It labels the input samples as “1” and “0” means a theft case or a genuine case respectively. In the first step of the process, a generalization function is used to translate the one-dimensional data vector to a higher-dimensional, two-dimensional space where the data classes can be distinguished more easily. We use XGBoost classifier in KSLX model to construct a model based on the training data, and to determine or detect whether or not the data samples is malicious. In the first step of the process, a generalization function is used to translate the one-dimensional data vector to a higher-dimensional, two-dimensional space where the data classes can be distinguished more easily. XGBoost by way of using the algorithm called the “weighted quantile sketch algorithm” helps the classifier to get incorrectly classified data corrected during each new iteration. The loss function performs the regularization and avoids the possibility of overfitting (Zhou et al., 2015).

Table 13. Displays the DR, FPR, and HD as percentages for the benchmark ETD models compared to our proposed model.

Classifiers	DR%	FPR%	HD%
k-NN	75.40	1.60	64.22
MLP	74.11	4.12	69.80
Random forest	81.33	1.89	79.47
L-SVM	72.16	4.98	75.19
Logistic Regression	77.18	3.98	78.19
Extra Trees	79.11	3.56	79.70
Adaboost	88.56	3.10	80.16
Proposed K-MLX model	**93.15**	**3.58**	**82.64**

They are bold as they represents the parameters for the evaluation.

The hyperparameters utilized during grid-search for XGBoost are shown in Table 8. The following hyperparameters produced the best results for XGBoost. The chosen parameters range where the no. of trees are 1000 to 2000, learning rate is 0.01, 0.1; Maximum depth = 7,15; and Minimum child weight = 1–10. The best performance is obtained at LR = 0.01, max. depth =15 min. child wt. equal to 1, RT equal to 5000, and evaluation metrics is AUC. The optimal

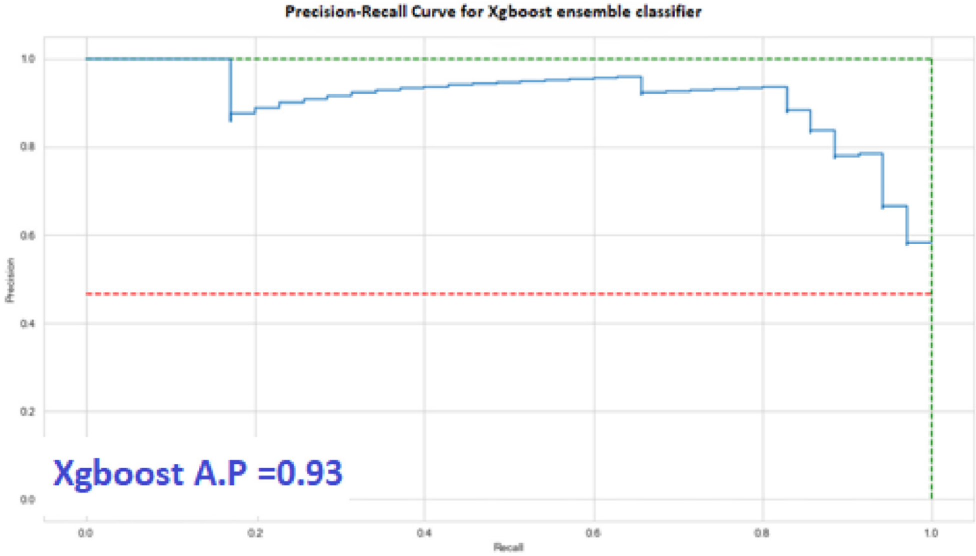


Figure 8. Precision recall PR curve for proposed XGBoost ensemble models for electricity theft detection (ETD).

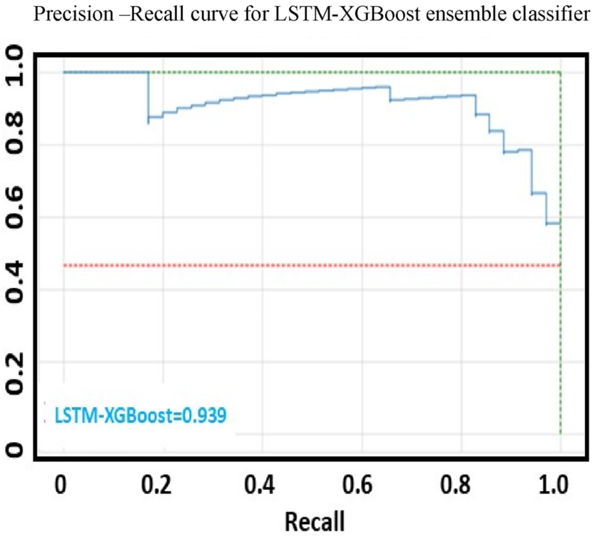


Figure 9. Illustrates the PR-recall curve of the proposed ensemble technique.

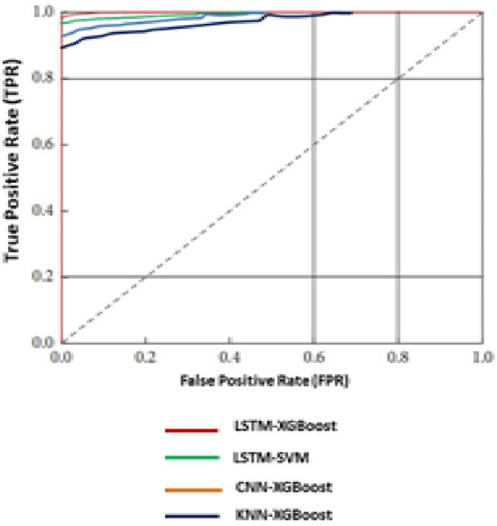


Figure 10. FPR-TPR comparison of proposed and existing models.

hyperparameters are obtained using 10-fold CV in GridSearchCV function of Scikit learn. The GridSearchCV chooses the best parameters in our model for XGBoost estimator. The GridSearchCV provided increased accuracy by choosing the best parameters as shown in Table 9.

The accuracy of our proposed model improved from 0.85 to 0.93 when parameters were set to default. Table 10 showing the execution time of using LSTM as feature extraction algorithm with XGBoost and existing classification model.

XGBoost predicts the input samples as “1” and “0” as theft case or genuine case respectively. XGBoost by way of using the algorithm called the weighted quantile sketch algorithm helps the classifier to get incorrectly classified data corrected during each new iteration. The objective function having loss function (L), a regularization term (Ω) and the parameter (θ) is mentioned below in Equation (18). It is mostly used in supervised learning problems where the target variable $\hat{y}_i$ is predicted by using the training data x_i with multiple features.

$$obj\ (\theta)\ =\ L\ (\theta)\ +\ \omega\ (\theta) \tag{18}$$

Dataset D is represented by Equation (19) as:

$$D = \{(x_i, y_i)\}(|D| = n,\ \ x \in R^m, y_i \in R) \tag{19}$$

where x_i, is a vector with n samples and m features and y_i the label. In this paper, x_i is the true meter reading for a day with n samples, and $y_i = \{0, 1\}$, where n is the number of samples (Zhou et al., 2015).

where 0 indicates normal power consumption and 1 indicates abnormal power consumption. Equation (20) is the result of a tree ensemble model function with inputs as K additive functions f_k .

$$\hat{y}_i = \sum_{k=1}^{K} f_k\ (x_i), f_k\ = F \tag{20}$$

Where F is the function space where all trees of classification are located.

It is pertinently mentioned that XGBoost optimizes functions rather tree model weights as mentioned in Equation (21) also used by authors in (Zhou et al., 2015).

$$\text{Obj}(\emptyset) = \sum_{i} l(\hat{y}_i, y_i) + \sum_{k} \Omega\big(f_k\big) \text{ with } \Omega\big(f\big) = \Upsilon T + \frac{1}{2}||w^2|| \tag{21}$$

The loss function measures how well a model fits during training by calculating the difference between y_i targetted value and $\hat{y}_i$ predicted value. T is total number of leaf nodes, and w represents the sum of all of their scores. Ω is the regularization measure of the XGBoost model's complexity. XGBoost builds a decision tree using gradient boosting and then adds regularization terms L1 and L2 to it. The regularization term helps to prevent over fitting. The training of the XGBoost model is performed in an additive manner mentioned in Equations (22)–(28). It is convenient to refer to the prediction term $\hat{y}_i^{(t)}$ for the ith instance at tth iteration. Iteratively optimizing the *i*th instance's goal at tth iteration saves time and effort (Zhou et al., 2015).

$$^{(t)} = \sum_{i=1}^{n} l(y_i \, , \hat{y}_i^{(t-1)} + f_t \, (x_i \,) + \, \Omega(f_t \,) \tag{22}$$

After the taking Taylor expansion of an objective

$$^{(t)} \triangleq \sum_{i=1}^{n} l(y_i \, , \hat{y}_i^{(t-1)} + g_i \; f_t \, (x_i \,) + \frac{1}{2} h_i \; f_t^2(x_t \,) + \, \Omega(f_t \,) \tag{23}$$

Where in the loss function, we show the g_i and h_i as first and second order gradient stochastics.

$$g_i \; = \; \delta_{\hat{y}_i^{(t-1)}} \; l(\, y_i \, , \hat{y}_i^{(t-1)}) \tag{24}$$

$$h_i \; = \; \delta^2_{\hat{y}_i^{(t-1)}} \; l(\, y_i \, , \hat{y}_i^{(t-1)}) \tag{25}$$

Let's call the set of all the instances of j as leaf $I_j = \{ \; i|q(x_i) = \text{j} \; \}$. Once the constants are eliminated and Ω is expanded, the objective function at time t is

$$^{(t)} = \sum_{i=1}^{n} \left[\; g_i f_t \, (x_i \,) \; + \frac{1}{2} h_i \; f_t^2(x_i \,) \right] + \Omega(f_t \,) \tag{26}$$

$$= \sum_{i=1}^{n} \left[\; g_i f_t \, (x_i \,) \; + \frac{1}{2} h_i \; f_t^2(x_i \,) \right] + \; \Upsilon T \; + \; \frac{1}{2} \sum_{j=i}^{T} w_j^2 \tag{27}$$

$$= \; \sum_{j=i}^{T} [\left(\sum_{i \in I_j} g_i \,) \, w_j \; + \; \frac{1}{2} \left(\sum_{i \in I_j} h_i \; + \; \tau w_j^2 \right) \right] + \Upsilon T \tag{28}$$

Since XGBoost allows user-defined loss functions, the above equations can be used as an optimization target for the new tree (Zhou et al., 2015). Regularizations are added to the loss function to simplify the new tree and prevent from over fitting.

Evaluation of model

Customers whose EC was used in the study were installed with a smart meter placed in their houses. This dataset is a valuable source in the study area and with a analyzing data from SMs since it contains a huge number and types of consumers, enough period of data collection for study. Additionally, it is available to the general public. Also a data file, that contains data in an interval of 30-min of each individual customer. We reduced the frequency of sampling to every hour, and then, for each user, we split the file into a dataset consisting of benign and malicious users, each of which had 24 components. The total data records in SGCC dataset have 42,372, customers with 38,752 normal users and 3615 fraudulent user records. The dataset have a sample frequency of 24 records per day. Since the malicious users are much fewer in number, six distinct types of malicious attacks are generated for each benign sample that had the values $x = $ "$x1$, $x2 \ldots x24$." If the NTL is detected falsely by the ML classifier when there is erratic power being supplied to the NAN cluster due to the reasons when one or many feeders are put off to accommodate the peak load demand in the area or because of the overloading of grids/substations/feeders or because of the feeder fault during heavy rains and snowfall. In that case, once the electricity is restored after some gap, the consumption increases, and the usage pattern changes leading to an increase in FPs as this change is treated as an anomaly by the machine-learning classifier. Our model forms a cluster using K-Means and this anomaly is neglected as the cluster is changed as a whole provided the number of SMs is fed from the same feeder or sub-station houses.

Classifying of data samples: It can be seen below the number of genuine and dishonest consumers registered in our data set. The genuine users have different usage patterns than the fraudulent users Electricity consumption pattern of fraudulent consumers is aberrant, i.e., it often shows periods of very low spikes and with periods of few highs as can be seen in Figure 6, also it has irregular patterns, contain low periodicity of electricity usage and their amount of usage is low for longer periods than usual which is possible in case of meter tempering or other methods. In contrast, the normal consumers follow a consistent pattern or periodicity and have usage of identical periods for the more consecutive days in a month or a week as shown in Figure 5. The machine-learning algorithms are trained on data to track the anomalous electricity usage behavior for identifying the electricity stealers. The electricity consumption as expected, after sunset reaches its peak points while after midnight to sunrise it go down to its minimum values most of the time. During standard working hours usually from 09:00 a.m. to 10:00 p.m. we have both high and average amounts of electricity consumption throughout the year. According to the graph between January & August electricity consumption amounts increases during colder season months than the other months (Figures 5 and 6).

We tried to figure out the AUC score by training the XGBoost classifier on a subset of features from SMs and other databases. A smart meter has data on how much electricity is used, the parameters of the smart meter, and the parameters of the electrical power. The features of the auxiliary database include tariff summaries, geographic information, technological characteristics, and contract data**.** The change in AUC with and without using the auxiliary database isn't seen at a very considerable level. However, by including the auxiliary feature subset, the FPR was reduced considerably.

The AUC of 0.88 was reached by using only the features the SM offered. Even with the extra databases, the AUC was found to be 0.87, which is less than the SM subset alone as can been seen in Figure 7. Auxiliary database features and smart meter features were used to get the FPRs of 5.92 and 2.01, respectively.

Proposed K-MLX model outcome interpretability using the symmetric Mean Absolute Percentage Error (sMAPE) algorithm:

The interpretability of the proposed K-MLX model outcomes is assessed using the sMAPE algorithm. We utilize the trained XGBoost model for prediction, validate its generalization, and employ the sMAPE to assess model performance as shown in Table 11. Specifically tailored for models like gradient boosted and other ensemble ML techniques, sMAPE quantifies the disparity between predicted and actual values. It calculates the sum of absolute differences between predicted and actual values, normalized by the average of the predicted and actual values. The formula for sMAPE calculation is as follows:

$$\text{SMAPE} = (100/n) \times \text{sum}\{\text{abs}(\text{actual\_}i - \text{predicted\_}i)/[\text{abs}(\text{actual\_}i) + \text{abs}(\text{predicted\_}i)]\} \tag{29}$$

where n represents data point count, actual_i is the actual value for data point i, and predicted_i is the corresponding predicted value. SMAPE has several advantages over other evaluation metrics such as mean absolute error (MAE) or mean squared error (MSE). Unlike MAE, which can be influenced by large errors, SMAPE gives equal weight to both over-prediction and under-prediction errors. Also, SMAPE is symmetric and more intuitive to interpret, as it provides a percentage error between actual and predicted values. A confusion matrix table shows the performance of a classification model, including the detection of electricity theft. It provides a summary of the model's predictions compared to the actual class labels of the data. The matrix typically contains four elements: true positive (TP), FP, true negative (TN), and false negative (FN).

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{30}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{31}$$

$$F_{\text{measure}} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \tag{32}$$

The confusion matrix can be utilized to compute various evaluation metrics such as accuracy, precision (PR), recall, and F1 score. These metrics provide a deeper insight into the model's performance and can serve as a roadmap for model enhancement. For detecting electricity theft, the positive class can be labeled as theft, while the negative class is labeled as no theft. A TP signifies a correct identification of theft, a FP indicates a wrongful identification of theft when none occurred, a false negative represents a failure to detect theft when it actually occurred, and a TN signifies a correct identification of no theft. In Equation (30), True Positive (TP) denotes the count of positive samples correctly classified as positive as used by researchers [32]. We utilize the F-measure, as depicted in Equation (32), to concurrently assess Recall and Precision. Additionally, the ROC Curve plots the number of FPs on the x-axis against the number of TP s on the y-axis. Precision is equivalent to TP. The ROC curve illustrates the detection capability in ETD of a binary classifier system as the threshold varies. The discrimination threshold measures the probability of theft occurrence. True negatives (TNs) represent negative samples correctly classified as negatives, while false negatives (FNs) denote negative samples erroneously classified as positives. In evaluating the proposed model, the "positive label" refers to electricity theft, and the "negative label" denotes no theft [32] as per the standard. We have also considered three evaluation indicators: detection rate (DR), false positive rate (FPR), and Bayesian detection rate A higher detection rate indicates better performance of a model [32] as a standard.

The AUC Score is a reliable metric to measure the performance of imbalanced datasets like ETD [40] as per the standard calculation. This metric measures the rate of change of TPs with the change in FP. If decision threshold is changed, ROC curve is an appropriate indicator to measure the change between true and false- positive rates. It is observed that KSPXS's XGBoost outperforms the widely used classifiers shown in Table 12 whilst L-SVM obtains the lowest performance. The DR, FPR and highest difference HD of our proposed model for ETD as shown in Table 13 is having 93.15%, 3.58% and 82.64% compared to the traditional models.

Our model achieved 80% precision when on-field inspections were done for the predictions. Figure 7 shows the AUC score of our model as compared to the other popular ETD models. Nevertheless, our approach doesn't (Figure 8).

go to find the high granularity as can been seen adopted by the authors in [33] and [34] as a part of their work to detect the intermittent frauds, but the high granularity of EC data will lead to privacy intrusion of customers. So an intermediate approach is adopted by our proposed model and the detection time is also better in the case of XGBoost method. The above figures depict the graphical representations of the Precision-Recall curve as shown in Figure 9 and the ROC curve in Figure 10. They emphasize the superior performance of the proposed K-Means-LSTM-XGBoost technique, showcasing higher True Positive Rates (DR) for a given False Positive Rate compared to existing benchmark models. These figures demonstrate that the proposed technique surpasses existing methods in achieving a higher True Positive Rate (DR) for a given False Positive Rate, indicating superior overall performance.

Future scope the proposed K-MLX model can be further improved by incorporating additional data sources, adapting to concept drift, developing explainable AI techniques, addressing deployment and integration challenges, and exploring generalization to other domains involving time-series data and anomaly detection

Conclusion

ETD presents a significant challenge, as it requires identifying anomalies in usage patterns and correlating them with routine anomalous measurements and other normal variations. Most of the studies primarily utilize electricity measurements (EC) from SMs and include few or no features from auxiliary databases, such as weather databases, curtailment schedules and tariff documents for theft detection. In contrast, our proposed approach incorporates additional features including feeder status, fault frequency, erratic power supply and local weather conditions in a cluster. These features play a crucial role in mitigating FP detections by machine-learning models. Focusing solely on usage behavior in this context often leads to a high number of FPs. The model's performance is assessed using standard and reliable metrics suitable for imbalanced datasets, including Precision, Accuracy, F1-Score, AUC-ROC, PR-AUC, and Matthews Correlation Coefficient (MCC). This approach outperforms existing conventional machine-learning models for ETD. The performance metrics, including Accuracy, Recall, Precision, F1-Score, AUC-ROC, PR-AUC, and MCC, further emphasize its effectiveness. The proposed model utilizes K-Means for clustering and combines LSTM-XGBoost for feature extraction and classification respectively. The proposed model achieved outstanding performance metrics, including precision, recall, F1-score, AUC-ROC, and PR-AUC values of 0.98, 0.93, 0.96, 0.96, and 0.94, respectively.

Acknowledgments

Mohammed Aleisa would like to thank the Deanship of Scientific Research at Majmaah University for supporting this work under Project No: R-2024-1083.

Data availability statement

The dataset employed in this research is available online at https://github.com/henryRDlab/ElectricityTheftDetection (accessed on 8 February 2023).

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Deanship of Scientific Research, Deanship of Scientific Research, Majmaah University (grant number Project No: R-2024-1083).

Ethical approval

This article does not contain any studies with human or animal participants.

ORCID iD

Deepak Prashar https://orcid.org/0009-0004-9926-1020

References

Adil M, Javaid N, Qasim U, et al. (2020) LSTM and bat-based RUSBoost approach for electricity theft detection. *Appl. Sci* 10(12): 1–21.

Amin S, Schwartz GA and Tembine H. (2012) Incentives and security in electricity distribution networks. In *International Conference on Decision and Game Theory for Security.* November 5–6, 2012, Budapest, Hungary: Third International Conference, GameSec 2012, pp. 264–280.

Anas M, Javaid N, Mahmood A, et al. (2012) Minimizing electricity theft using smart meters in AMI. *Proceedings of 2012 7th International Conference P2P, Parallel, Grid, Cloud Internet Computing (3PGCIC)*, 12–14 November 2012, Victoria, British Columbia, Canada: Institute of Electrical and Electronics Engineers (IEEE), pp. 176–182. doi: 10.1109/3PGCIC.2012.42

Angelos EWS, Saavedra OR, Cortés OAC, et al. (2011) Detection and identification of abnormalities in customer consumptions in power distribution systems. *IEEE Trans Power Deliv* 26(4): 2436–2442.

Asif M, Kabir B, Ullah A, et al. (2021) A hybrid deep learning approach for detecting non technical losses in smart grids. *Researchgate Net.* https://www.researchgate.net/profile/Nadeem-Javaid/publication/351133891_A_Hybrid_Deep_Learning_Approach_for_Detecting_Non_Technical_Losses_in_Smart_Grids/links/608a1f34a6fdccaebdf4e34e/A-Hybrid-Deep-Learning-Approach-for-Detecting-Non-Technical-Losses-in

Cárdenas AA, Amin S, Schwartz G, et al. (2012) A game theory model for electricity theft detection and privacy-aware control in AMI systems. In *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Oct 1 - Oct 5 2012, Monticello, IL: IEEE, pp. 1830–1837.

Chen C (2014) Intelligent analysis on the malfunction meter based on the electric energy data acquisition system. *Electr Meas Instrum* 51(15): 18–22.

Cody C, Ford V and Siraj A. (2015) Decision tree learning for fraud detection in consumer energy consumption. In *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*, 9–11 December 2015, Miami, FL: IEEE, pp. 1175–1179.

Enguo Z and Xuan L. (2011) Construction and application of electric energy information acquisition system. In *2011 IEEE 3rd International Conference on Communication Software and Networks*, 4–7 June 2020, Chengdu, China: IEEE, pp. 114–116.

Javaid N (2021) A PLSTM, AlexNet and ESNN based ensemble learning model for detecting electricity theft in smart grids. *IEEE Access* 9: 162935–162950.

Jokar P, Arianpoo N and Leung VCM (2015) Electricity theft detection in AMI using customers' consumption patterns. *IEEE Transactions on Smart Grid* 7(1): 216–226.

Jokar P, Arianpoo N and Leung VCM (2016) Electricity theft detection in AMI using customers' consumption patterns. *IEEE Transactions on Smart Grid* 7(1): 216–226.

Khoo B and Cheng Y. (2011) Using RFID for anti-theft in a Chinese electrical supply company: A cost-benefit analysis. In *2011 Wireless Telecommunications Symposium (WTS)*, 13–15 April 2011, New York, NY: IEEE, pp. 1–6.

Krishna VB, Weaver GA and Sanders WH. (2015) PCA-based method for detecting integrity attacks on advanced metering infrastructure. In *International Conference on Quantitative Evaluation of Systems*, pp. 70–85.

Liu Z, Ding W, Chen T, et al. (2023) A electricity theft detection method through contrastive learning in smart grid. *EURASIP J Wirel Commun Netw* 2023(1): 54.

Lo C-H and Ansari N (2013) CONSUMER: A novel hybrid intrusion detection system for distribution networks in smart grid. *IEEE Trans Emerg Top Comput* 1(1): 33–44.

Messinis GM and Hatziargyriou ND (2018) Review of non-technical loss detection methods. *Electr Power Syst Res* 158: 250–266.

Muniz C, Figueiredo K, Vellasco M, et al. (2009a) Irregularity detection on low tension electric installations by neural network ensembles. In *2009 International Joint Conference on Neural Networks*, 14-19 June 2009, Atlanta, GA: IEEE, pp. 2176–2182.

Muniz C, Vellasco MMBR, Tanscheit R, et al. (2009b) A neuro-fuzzy system for fraud detection in electricity distribution. In *IFSA/EUSFLAT Conference*, pp. 1096–1101.

Nagi J, Mohammad AM, Yap KS, et al. (2008) Non-technical loss analysis for detection of electricity theft using support vector machines. In *2008 IEEE 2nd International Power and Energy Conference*, pp. 907–912.

Nizar AH, Dong ZY and Wang Y (2008) Power utility nontechnical loss analysis with extreme learning machine method. *IEEE Trans Power Syst* 23(3): 946–955.

Pamir N, Javaid S, Javaid M, et al. (2022) Synthetic theft attacks and long short term memory-based preprocessing for electricity theft detection using gated recurrent unit. *Energies* 15(8). doi: 10.3390/en15082778.

Rumelhart DE, Hinton GE and Williams RJ (1986) Learning representations by back-propagating errors. *Nature* 323(6088): 533–536.

Salinas S, Li M and Li P. (2012) Privacy-preserving energy theft detection in smart grids. In *2012 9th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, 18–21 June 2012, Seoul, Korea (South), IEEE, pp. 605–613.

Smith TB (2004) Electricity theft: A comparative analysis. *Energy Policy* 32(18): 2067–2076.

Wen M, Yao D, Li B, et al. (2018) State estimation based energy theft detection scheme with privacy preservation in smart grid. *IEEE Int Conf Commun* 2018: 1–6.

Wu Q, Zhang M and Liao L (2022) Analysis of electricity stealing based on user electricity characteristics of electricity information collection system. *Energy Reports* 8: 488–494.

Yan Z and Wen H (2021) Electricity Theft Detection Base on Extreme Gradient Boosting in AMI. *IEEE Trans Instrum Meas* 70: 1–6.

Yang Y, Littler T, Sezer S, et al. (2011) Impact of cyber-security issues on Smart Grid. *IEEE PES Innovation Smart Grid Technology Conference Europe*, 5–7 December 2011, Manchester, UK, IEEE, pp. 1–7. doi: 10.1109/ISGTEurope.2011.6162722.

Zhou K, Yang C and Shen J (2017) Discovering residential electricity consumption patterns through smart-meter data mining: A case study from China. *Utilities Policy* 44: 73–84.

Zhou Y, Chen X, Zomaya AY, et al. (2015) A dynamic programming algorithm for leveraging probabilistic detection of energy theft in smart home. *IEEE Trans Emerg Top Comput* 3(4): 502–513.