Wiley
Journal of Electrical and Computer Engineering
Volume 2024, Article ID 6225510, 16 pages
https://doi.org/10.1155/2024/6225510

WILEY

Research Article

Electricity Theft Detection in a Smart Grid Using Hybrid Deep Learning-Based Data Analysis Technique

Camille Franklin Mbey,[1] **Jacques Bikai**,[2] **Felix Ghislain Yem Souhe**,[1,3,4] **Vinny Junior Foba Kakeu**,[1,4] **and Alexandre Teplaira Boum**[1]

[1]*Department of Electrical Engineering, ENSET, University of Douala, Douala, Cameroon*
[2]*Department of Energy Engineering, University of Ngaoundere, Douala, Cameroon*
[3]*Laboratory of Technology and Applied Sciences, IUT, University of Douala, Douala, Cameroon*
[4]*Department of Electrical Engineering and Industrial Computer Science, IUT, University of Douala, Douala, Cameroon*

Correspondence should be addressed to Felix Ghislain Yem Souhe; felixsouhe@gmail.com

Received 20 March 2024; Revised 22 June 2024; Accepted 28 June 2024

Academic Editor: Ping-Feng Pai

With the popularization of smart meters around the world and the appearance of a large amount of electricity consumption data, the analysis of smart meter data is of major interest to electricity distributors around the world. Therefore, we proposed a hybrid artificial intelligence (AI) technique considering sudden changes of consumption in order to accurately predict fraudulent consumers in the smart network. Thus, the proposed hybrid model is based on the support vector machine (SVM) and a particle swarm optimization (PSO) algorithm to detect energy fraudsters in the network. In addition, a real smart grid dataset is used to verify the effectiveness of the proposed algorithm. Moreover, a smart calendar context is modeled showing the scheduling of energy consumption. The effectiveness of the proposed technique is evaluated using performance coefficients such as precision, recall, F1-score, and area under ROC curve (AUC). We also perform sensitivity analysis through regression, variance, and variogram analysis. The results obtained give a performance of 98.9% in the detection of irregular consumers in the smart power grid. These results demonstrate the effectiveness of the proposed method compared to that in the literature.

1. Introduction

Smart meter data analysis comprises all the strategies consisting of information processing collected using smart meters [1]. The advantages that the integration of smart meters brings into a network are diverse, notably, the acquisition of energy data, management of electricity consumption, safer billing, more reliable service, a reduction of technical losses, a reduction of energy theft, an improvement of security profile for consumers, and the optimization of data management [2]. Smart meters will therefore allow daily reading and collection of consumption data in real time [3].

The analysis of these data could be useful for electricity companies in understanding customer consumption behaviors [4]. In addition, these data can also be transmitted to the data center for possible storage and automated processing. It also allows unprecedented possibilities of analyzing, storing, and combining energy system data for both entities which are consumers and operators of the electrical power network [5]. Energy distribution systems are victims of significant energy losses; this represents a shortfall and a challenge for managers of the traditional electrical network and the smart grid [6]. These power losses can be technical losses or nontechnical losses. Technical losses arise in power distribution lines, transformers, and other network

components [7]. Nontechnical losses occur during the breakdown of a transformer or a defective smart meter. The capacity given to the smart grid ensures an intelligent and automated anticipation functionality to face the occurrence of both technical and nontechnical losses in the network [8].

In addition, some events such as fraud, undetected and unbilled consumption, and illegal connection to the network, measurement, and installation errors can be the cause of irregular and fraudulent electricity consumption [9]. In this case, these frauds can lead to an overload of the electrical system, a significant deterioration of energy quality and the increasing of billing costs, which would be unfair for honest consumers. In order to reduce this energy deficit, energy companies constantly carry out field inspections. Unfortunately, these inspections are expensive, which urge these companies to use automated calculation tools in order to make a consumption forecast to deal with these fraudulent problems [10]. The use of smart meters allows the reduction of technical losses; however, it makes fraudulent activities more difficult to detect, and in addition, their costs are very high and they require new infrastructure for data collection.

Electricity theft generally consists of any attempt to subtract energy from the electricity network without consumption being officially recorded. In fact, many electrical energy consumers make installations to avoid the recording of their energy consumption from electrical meters [11]. Theft detection in this context is defined as monitoring customer behavior (load curve) in order to estimate, detect, or avoid unwanted behavior. Works in [12] use data in the smart grid for anomaly detection. Using smart meters, the operator can know customer consumption data and thus identify any anomaly that may occur.

Many research studies have been carried out in smart meter data analysis for energy theft detection. In [13], a research work on electricity theft detection in smart grid systems has been carried out. It proposes a combined approach made of robust CNN-LSTM hybrid model, which is the integration of a convolutional neural network (CNN) and a long short-term memory (LSTM), with a preprocessing algorithm on the electricity consumption signature dataset to solve the binary classification problem. The performance comparison results of the proposed model based on the raw dataset and the transformed dataset showed that the number of fraudulent users was relatively small compared to the real users. Therefore, the overall accuracy is almost the same in both cases. However, due to class imbalance, the model could not classify fraudulent users very successfully. In [14], a study on energy theft and faulty smart meter detection in the smart grid architecture in a neighborhood area network (NAN) was carried out. For this purpose, two algorithms based on multiple linear regression models were designed successively. The first algorithm is a scheme based on linear regression for the detection of energy theft and defective smart meters, called LR-ETDM. The second algorithm, called CVLRETDM, is a scheme based on linear regression with improved categorical variables for the detection of energy theft and defective smart meters by dummy coding. Simulation results in MATLAB R2014b showed that fraudulent consumers can be detected regardless of whether they steal energy at a constant or variable rate.

2. Related Work

Energy theft represents the most significant problem related to the implementation of smart grids. It is estimated that power companies lose millions of dollars each year due to energy theft around the world. Considering that smart grids are proposed to modernize the current electricity grid, energy theft can become a very serious problem since smart meters used in smart grids are vulnerable to several attacks, especially in advanced metering infrastructure (AMI). The smart meter being the main constituent element of the smart grid having replaced analog electricity meters with the association of a computer system which transmits information through digital communication interfaces. Through strict rules and effective planning, energy theft can be effectively reduced.

To this end, stopping these energy theft practices is as important as energy production. Thus, it is important to develop effective and reliable methods based on machine learning and deep learning to accurately identify illegal consumers who commit energy theft. Several works have been carried out with the aim of reducing this energy theft. Table 1 gives the overview on smart meter data analysis and electrical theft detection.

Based on previous works, we propose a novel theft detection method which considers the irregularities of sudden changes in user consumption to identify fraudulent consumers with great precision and thus limit nontechnical losses for the distribution of electricity. We precisely address to the accuracy challenges of the model to have the best performance for the electricity theft detection.

The major contribution of our work is presented as follows:

(1) First, we present a general review on smart meter data analysis techniques and artificial intelligence algorithms for the detection of consumption fraud using data from smart meters.

(2) Second, we propose a hybrid SVM-PSO method for data analysis composed of the SVM and a metaheuristic algorithm based on PSO. The SVM can detect inappropriate correlations of abrupt changes in consumption, and the PSO can optimize the appropriate SVM parameters.

(3) Third, we use a real consumption dataset in a smart grid to implement the proposed hybrid algorithm with the aim to verify the effectiveness of the proposed method.

(4) Fourth, we develop a calendar context model for scheduling energy consumption in the electricity theft detection.

(5) Finally, we evaluate the effectiveness of the proposed hybrid model using several performance coefficients such as precision, recall, F1-score, and AUC in order

Table 1: Overview on smart meter data analysis and electrical theft detection.

Authors	Problem	Methodology	Outcomes	Future possibilities
[15]	Analyze smart meter data collected from 1000 households in Poland	Machine learning	Data are more random and volatile	Analysis of daily load profiles
[16]	Energy theft detection in advanced metering infrastructure	SVM	Detection rate is around 60 to 70%	Implement a state estimation technique
[17]	Prediction of nontechnical losses from smart meter data	Neural networks	Detection rate is from 3 to 60%	Modify the data granularity
[18]	Identification of energy theft and faulty smart meters	Deep bidirectional recurrent neural network	Methods can effectively locate fraudulent consumers and broken meters	Improve the detection framework
[19]	Data analysis of energy efficiency in the residential sector	K-nearest neighbor	Good prediction quality and accuracy of almost 70%	Use consumption characteristics of residences
[20]	Analysis and characterization of data in a smart grid	Stochastic gradient descent	Fast variations and random power distributions in smart meter data	Detecting the presence of specific load patterns in real-time power data
[21]	Analyzing and visualizing smart meter data and estimating the typical load profile	Artificial intelligence platform	Good precision coefficient	Estimating the temporality and consumer segmentation
[22]	Anomaly detection in energy consumption data from the University of California	Encoder and decoder frameworks with neural networks	High accuracies in anomaly detection	Improve the detection platform
[23]	Anomaly detection in smart grids	Supervised machine learning technique	Fast anomaly detection	Implement this algorithm for real-world data
[24]	Suspicious energy consumption detection	Multiprofiling-based machine learning	Classification and selection of the most appropriate profiles	Consider climate and periodic context
[25]	Developing a consumption program using data collected from smart meters	Demand participation program with two-way communication	Accuracy around 90%	Improve production and demand capacity
[26]	Analysis of residential smart meter data of 50 houses every 15 minutes	Abnormal behavior detection framework	Demand response identification, abnormal behavior detection, and fault diagnosis	Build a dynamic signature database and optimize the distribution of data processing capabilities between devices and control centers
[27]	Data analysis of advanced metering infrastructure of South Asian companies	Smart metering infrastructure network	Data storage and data processing	Conduct experiments through analysis operations of advanced metering systems
[28]	Smart meter data analytics	Data analytic framework	Data collection and application development	Improve the performance of the machine
[29]	Visualizing smart meter data	Virtual reality platform	Data cleaning and classification	Ensure a balance in energy demand
[30]	Integration and identification smart meter data	Supervision and intelligent techniques	Offering specific services to consumers	Improve the data supervision
[31]	Smart meter data management for energy theft detection	Smart threat model using the new ARMA detector-GLR	Evaluating and comparing detectors' anomalies	Improve the detection of an invariant attack for the composite hypothesis testing format
[32]	Automatic electricity theft detection on electricity consumption data of the China Network Corporation	CNN model with the blue monkey algorithm	Precision around 0.92 for reduction of database features	Find the best configuration of the sequential model for the classification and identification of electricity theft
[33]	Cyber-attacks and energy theft detection	Deep recurrent neural networks and nondominated sorting genetic algorithm associated with the Pareto frontier	Good reduction of their electricity bill	Improve the performance of the detection mechanism

Table 1: Continued.

Authors	Problem	Methodology	Outcomes	Future possibilities
[34]	Detection of energy theft in a public Chinese network consumption dataset	Decision tree (DT), artificial neural network (ANN), DNN, and AdaBoost	Better precision, recall, and F1-score	Improve the performance of the detector using other supervised algorithms on different types of dataset
[35]	Energy theft detection on a Chinese network corporation database	SVM and ANN	High accuracies	Implement a data imbalance technique
[36]	Energy theft detection for reduction of nontechnical losses in a smart grid	SVM	Accuracy of almost 90%	Improve the accuracy
[37]	Electricity theft detection on an IEEE 57 bus, IEEE 30 bus, and IEEE 14 bus network	Compressed sensing and sparse representation techniques	Good probability of detecting energy theft requires a large number of sensors	Optimally place the detectors in the distribution network in order to limit the number of sensors used
[38]	Detection of energy theft in a smart grid	Machine learning techniques	Good precision	Optimize the security of smart networks
[39]	Energy theft detection on a real dataset of the China Network Corporation	Optimized differential evaluation random under sampling boosting classifier	High precision of 96%	Improve the classifier using a metaheuristic optimization algorithm
[40]	Energy theft detection and faulty meter detection in AMI	Linear regression model	Successful detection of all fraudulent consumers and effective discovery of faulty smart meters	Improve the performance and reliability of the proposed detection scheme
[41]	Identification of power theft and data center distribution lines in a smart grid	Dynamic programming algorithm	Inserting a minimum number of relay protection devices in data centers	Better study the impact of different cases of energy theft by inserting protection relay algorithms; consider malicious attacks on both owners' and neighbors' smart meters
[42]	Detection of energy theft in microgrids	Synthetic adaptive minority oversampling technique	Detection rate of 0.981 and 0.976 and declassification rate of 0.015 and 0.033	Improve the performance of the classifiers
[43]	Energy theft detection on the China Smart Grid Corporation dataset	Gated recurrent unit and SVM	Accuracy of 82.65% and 64.33%	Consider imbalance problems
[44]	Real-time anomaly detection	Universal Lempel–Ziv algorithm	Normal behaviors reading of smart meter data and alert provider	Implementation and modification using real-time data to effectively adjust the detection of the dynamic nature of electricity theft
[45]	Electricity thieves' detection and powerful diagnostics on smart meter data from the Florida company	Distributed intelligent platform and Stackelberg game theoretical model	Efficient detection of fraudulent meters	Improve the analysis of the strategic interactions
[46]	Detection of electricity theft	LSTM and RUSBoost	Better precision, recall, F1-score, and receiver operating characteristics	Ensure normalization and interpolation of process data
[47]	Detection of electricity theft based on the dynamic price	Modified machine learning-based XGBoost technique	Effective detection of electricity theft	Implement a new type of power theft case
[48]	Detection of electricity theft in the smart grid	ANN, MLP, LGBM, NMB-SMOTE, and TBSSVM	Database contains 9% fraudulent consumers	Consider the unbalanced nature of the data
[49]	Electricity theft detection on consumers' actual electricity consumption data from the China Grid Corporation	CNN, GRU, and PSO	Good performance using AUC, precision, recall, and F1-score	Reduction of dimensionality and redundancy in the database
[50]	Electricity theft detection on the database of the Chinese network corporation	KNN	Accuracy of 91.0%	Improve the accuracy

Table 1: Continued.

Authors	Problem	Methodology	Outcomes	Future possibilities
[51]	Electricity theft detection	CNN-RF	Good precision in the detection	Improve the granularity and duration of smart meter data
[52]	Electricity theft detection	Maximum information coefficient technique	Good correlations between nontechnical losses and consumer behavior	Find abnormal consumers among load profiles
[53]	Electricity theft detection	XGBoost, CatBoost, and LightBoost	Good accuracy	Implement a hybrid model
[54]	Detection of electricity theft	GRU and DRNN	Effective detection	Improve the performance of the model
[55]	Detection of electricity theft in smart grids	Privacy-preserving approach	Reduction of fraudulent consumers	Build a more secure smart system
[56]	Electricity theft detection with smart meter data	Physically inspired data pipe model	Rapid detection of thieves	Optimize the billing method
[57]	Electricity theft detection in smart utility networks	Theoretical game model	Better classification of smart meter data	Implement another optimize game model
[58]	Electricity theft detection in smart grids	Random matrix theory	Good reduction of computing errors	Improve the matrix theory of a metaheuristic algorithm
[59]	Electricity theft detection in smart grids	Dynamic time warping and KNN	F1-score of 94%, TPR of 93%, and FPR of 1.1%	Plan on-site inspections to catch fraudulent customers
[60]	Electrical fault detection using smart meter data	Neuro fuzzy algorithm	Accuracy of 97%	Improve the hybrid model
[61]	State estimation in smart grids using smart meter data	Kalman filter and WLS	Precision of 95%	Build a neuro fuzzy filter

to demonstrate the superiority of the proposed method compared to those in the literature. Moreover, a sensitivity analysis has been performed to prove the accuracy of the results.

The rest of our work is presented as follows: Section 3 gives the materials and methods, where we present the household consumption dataset as well as the software used to implement the artificial intelligence algorithm based on the SVM and the PSO algorithm. Section 4 presents simulation results and discussion. To this end, we present fraudulent consumers over a period of one year in order to reduce nontechnical losses in the network. Section 5 presents the conclusion and perspectives of the future.

3. Materials and Methods

3.1. Materials

3.1.1. Description of the Dataset. The dataset was obtained from the Cameroon Electricity Company. This dataset contains data from nearly 1000 consumers in which 85.2% are normal consumers and 14.8% have abnormal consumption which could be considered as electricity theft. The dataset was collected over an interval from January 1, 2020, to December 31, 2020. Table 2 presents a description of this dataset.

3.1.2. OpenDSS-G. Some functionalities such as line geolocalization can only be implemented in OpenDSS-G which is the evolution of simulation tools based on OpenDSS. Its interface has adopted OpenDSS functionalities to make advanced components of the platform easier for the user. This version includes parallel processing of OpenDSS elements. OpenDSS-G also allows modification of network elements during the simulation process.

3.1.3. OMNet++. Objective modular network in C++ (OMNet++) is a tool for simulating components intended for communication networks. OMNet++ is also an Eclipse platform that expands on other features such as editor and design wizard. OMNet++ also has properties for creating and configuring models (NED and ini file) ensuring batch execution and analysis of simulation results, while Eclipse provides the C++ editor and other optional elements (UML modeling, access to the database). It allows you to simulate a real existing communication network without any risk. Simulation results can be analyzed to determine how the real network may be affected.

In our cosimulation, OMNet++ makes it possible to create a NAN network to recover data at the transformer stations on the electricity network. It is used to build and simulate the communication network and enable cosimulation with OpenDSS through the COM interface. OMNet++ is capable of creating our own communication network, cosimulating communication in the smart grid with OpenDSS, and ensuring better retransmission of information in the various local networks.

3.1.4. MATLAB. MATLAB is software that was initially developed by Cleve Moler in the 1970s. In addition, the MathWorks Company ensures its continuous development to this day. It is used for matrix calculations in order to analyze data and classify these. MATLAB also allows programming for the intelligent resolution of data mining problems.

In our work, MATLAB allows us to implement all AI algorithms such as SVM and PSO. All simulations of this work were carried out using the MATLAB R2020b 64 bit version.

3.1.5. Computer. All simulations were carried out on a computer with the following characteristics: icore5, 3.5 GHz, 8 GB RAM, 500 GB hard disk, and Windows 7/64 bit system.

3.2. Methods

3.2.1. Cosimulation Platform. The smart grid simulation (SG-SIM) platform is developed using OpenDSS software for the electrical power network and the OMNet++ software for the communication network. OMNet++ contains the INET component which has all the elements of the communication network such as the PMU, Solver, and PDC. SG-SIM also has an optimization and communication interface that can operate according to standards such as IEEE c37.188 for the PMU. This cosimulation platform makes it possible to implement the behavior of the intelligent network through real-time operation of the two software programs to collect consumption data in real time from smart meters. Figure 1 illustrates the structure of this platform.

3.2.2. SVM. The SVM was developed by Vladimir Vapnik in the 1980s using statistical, optimization, and neural network tools. The architecture of the SVM depends only on the parameter C and the kernel function. The SVM is based on hyperplanes that separate the training data between the two subgroups. Among all the separation planes between the two classes, there is an optimum space plane whose distance between the closest optimal hyperplane and the training element is maximum, thus allowing a clearer distinction of the regions having drop points across each group. Figure 2 shows the hyperplane that correctly separates the data.

The expression for the optimal hyperplane and the hyperplane separators are given by equations (1)–(3), respectively.

$$g(x) = w * x + b = 0, \tag{1}$$

$$p_{+1} = w * x + b = +1, \tag{2}$$

$$p_{-1} = w * x + b = -1. \tag{3}$$

Equation (4) gives the objective function under the constraint of equation (5).

Table 2: Description of the dataset.

Explanation	Values
Electricity consumption period	January 1, 2020, to December 31, 2020
Total number of consumers	1000
Number of normal consumers	852
Number of fraudulent consumers	148

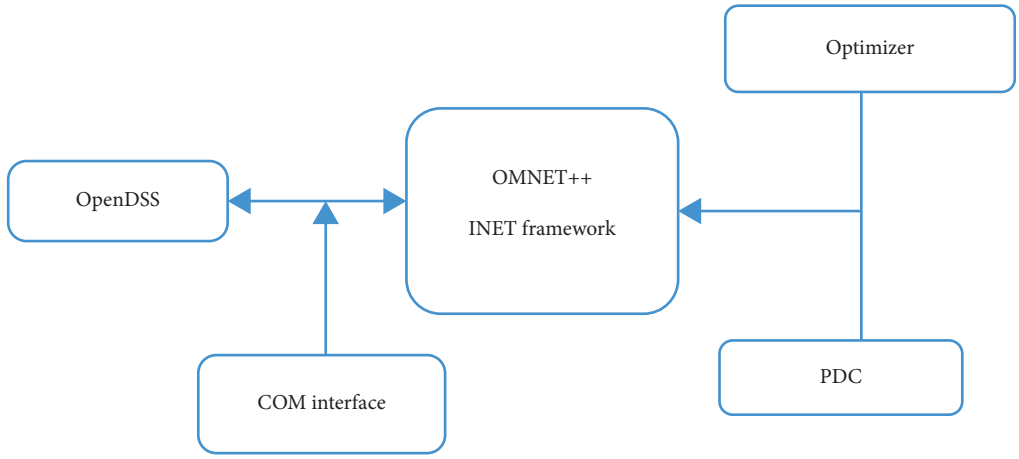

Figure 1: Smart grid cosimulation platform.

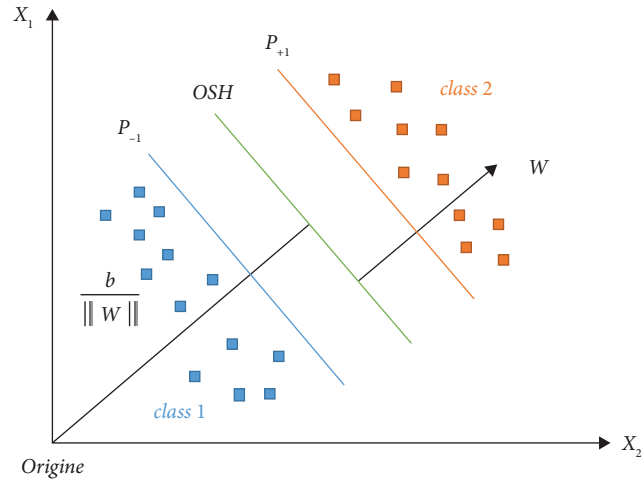

Figure 2: Data separation hyperplane [3].

$$\min_{w} \frac{1}{2}(\vec{w}.\vec{w}), \tag{4}$$

$$y_i(\vec{w}.\vec{x}_i + b) \geq 1. \tag{5}$$

The objective function and its constraint represent the quadratic optimization problem. It can be solved using the Lagrange multipliers given by the following equation:

$$L(\vec{w}, b, \vec{\alpha}) = \frac{1}{2}\|\vec{w}\|^2 - \sum_{i=1}^{n} \alpha_i \left[y_i(\vec{w}.\vec{x} + b) - 1 \right]. \tag{6}$$

The vast separator is a quadratic minimization problem which is valid for the case of a separable problem. In the case of a nonseparable classification problem, the machine will assign a false output to a vector x_i if the corresponding ξ_i is greater than 1. The sum of all ξ_i therefore represents a bound of the number of errors.

Thus, instead of looking for the weight vector w_0 which minimizes the square of the norm (w, w), we then seek to minimize, under the constraints expressed in the following equation:

$$Q(w,\xi) = \frac{1}{2}(w,w) + C\sum_{i=1}^{m}\xi_i, \tag{7}$$

where C is the parameter that can be chosen by the user (the larger this parameter, the more it amounts to assign a strong penalty to errors).

We thus obtain the equation of a hyperplane which is now optimal. The support vectors are still the example vectors closest to the hyperplane, but this time there are example vectors which are located in the wrong half-space and they will therefore not be considered to construct the hyperplane.

Carrying out a learning program using SVM essentially boils down to solve an optimization problem involving a quadratic programming resolution system in a space of significant dimension by considering the objective function of the following equation:

$$\underset{\alpha_i}{\text{Max}}\, W(\alpha) = \sum_{i=1}\alpha_i - \frac{1}{2}\sum_{i,j} y_i y_j \alpha_i \alpha_j k\{x_i, x_j\}. \tag{8}$$

Under the constraints of the following equation:

$$\begin{cases} \sum_i y_i \alpha_i = 0, \\ \forall i \in \{1,\ldots,n\}, \end{cases} \tag{9}$$

where $0 \le \alpha_i \le C$.

We can express this optimization problem in its dual formulation in the matrix form presented in the following equation:

$$\max_{\alpha_i} \frac{1}{2}\alpha^t H\alpha + C^t\alpha, \tag{10}$$

where $H = ZZ^t$ and $C^t = (-1,\ldots,-1)$.

Under constraints, $\alpha^t Y = 0$, $\alpha_i \ge 0$, and $i = 1,\ldots,n$.

With

$$Z = \begin{pmatrix} y_i x_i \\ . \\ . \\ . \\ y_n x_n \end{pmatrix}, \qquad T = \begin{pmatrix} y_i \\ . \\ . \\ . \\ y_n \end{pmatrix}, \tag{11}$$

where α is the Lagrange multiplier and H is the Hessian matrix, such as $H = y_i y_j k\{x_i, x_j\}$.

The SVM optimization problem can be solved analytically only when the number of examples is reduced or when, in the separable case, it is known which examples are the support vectors.

3.2.3. Particle Swarm Optimization. PSO is a metaheuristic technique introduced by Kennedy and Eberhart in 1995. It is based on the behavior of birds and insects. In the PSO, the population is called a swarm and is made up of individuals called particles. The initialization of the PSO algorithm is done with a number of iterations and particles. The movement of each particle consists of convergence towards the region having the best potential and having the optimal solutions. Each particle has a memory function and gradually adjusts its path based on its own experiences and the experience of other particles. The path traveled by each particle is based on the best particular position *p*best and the best global position *g*best of the swarm. Each particle has two vectors, notably the position vector X_i and the speed vector V_i. Each iteration of the PSO consists of moving the particle in space in order to find the optimal solution. The velocity vector and the position vector can be expressed by the following equations:

$$V_i^{t+1} = wV_i^t + c_1 r_1\left(p\text{best} - X_1^t\right) + c_2 r_2\left(g\text{best} - X_i^t\right) \tag{12}$$

$$X_i^{t+1} = X_i^t + V_i^{t+1}, \tag{13}$$

where V_i^{t+1} is the speed of the particle i on the next position, V_i^t is the speed of the particle i on the previous position, *p*best is the best particular position, *g*best is the best overall position, X_i^t is the previous position of the particle i, X_i^{t+1} is the next position of the particle i, c_1 and c_2 are the positive acceleration constants making it possible to maintain the balance between individual and social behavior, r_1 and r_2 are the numbers generated in a looped manner in the intervals 0 and 1, and w is the weight of inertia making it possible to maintain the balance between exploration and exploitation. Figure 3 shows the flow chart of the PSO algorithm.

3.2.4. Methodology of Electricity Theft Detection Using the SVM-PSO Technique. The analysis performed in this study focuses on abrupt changes in customers' consumption habits across their load profiles. Indeed, customers are represented by their load profiles or by using the temporal evolution of electricity consumption during a given period. These profiles are characterized by different shapes, which significantly represent their general behavior and it is possible to evaluate the measure of similarity between each customer and their consumption habits. This creates a measure of overall similarity between normal customers and potential fraudsters as a whole. Identification and detection are developed by the support vector classification (SVC), which is the intelligent "classification engine." For this purpose, we use SVM which is a recent artificial intelligence method based on a class of learning algorithms for the detection and identification of fraudulent activities as well as a PSO algorithm for the optimization of initial parameters of the SVM such as the velocity and position of the particle. To develop the detection system using the hybrid SVM-PSO method, we followed three main steps: preprocessing the data, developing a model for classification, and final processing of

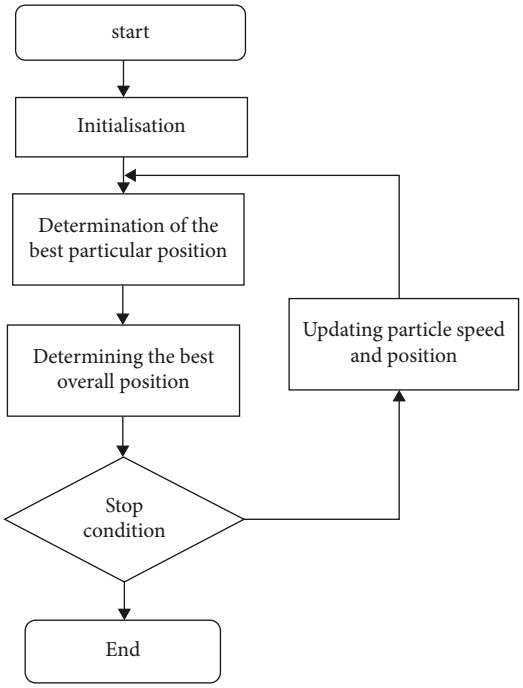

Figure 3: Flowchart of the PSO algorithm.

the data. Figure 4 presents the electricity theft detection methodology using the hybrid SVM-PSO algorithm.

(1) First, the data search consists of retrieving all the data in the field or at the energy distribution agency. This concerns all customers with a normal or fraudulent profile.

(2) Second, data preprocessing presents data mining techniques used for preprocessing customer information and billing data for feature selection and extraction. In this section, the operator must group all the data into an Excel or CSV file that can easily be used by MATLAB in the editor.

(3) Third, feature extraction involves extracting essential parameters such as customer consumption indexes for learning the SVM.

(4) Fourth, initialization of SVM parameters involves initializing the basic parameters of the SVM to ensure its operation.

(5) Fifth, SVM parameter optimization involves estimating the accuracy of the SVC learning model by adjusting the SVC kernel parameters and the penalty error parameter.

(6) Sixth, training the SVC makes it possible to build the SVC model thanks to the optimization of the parameters using the PSO according to the number of samples. Model training allows the model to train the inputs and adjust these data so that these can converge towards the desired output.

(7) Seventh, the development of the SVC classification model consists of defining training values to predict future variables. It illustrates learning and training of SVC for parameter optimization, development of SVC classifier (model), and testing and validation of the SVC model.

(8) Finally, the final data processing describes the development of an algorithm which makes the choice and selection of suspicious customers based on the forecast made by the SVM classification and the actual consumption of customers. It allows you to generate potentially fraudulent customers based on the model

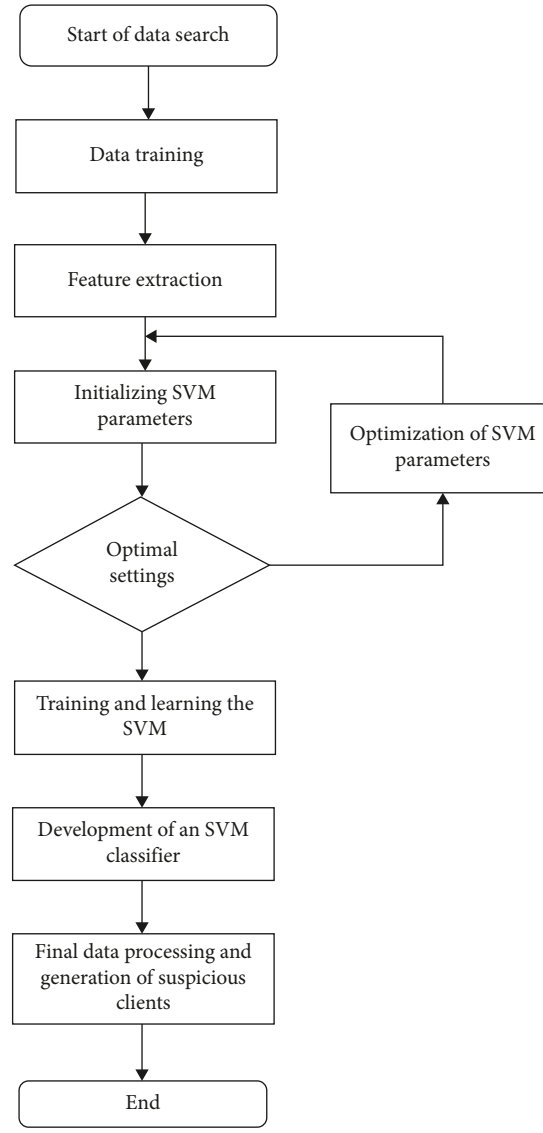

Figure 4: Electricity theft detection flowchart.

3.2.5. Calendar Context of Electricity Theft Detection. In this work, we propose a calendar context as shown in Figure 5 for electricity theft detection which learns user preferences, relationships, dependencies, and interactions between the smart system and its external environment comprising consumer, producer, and distributed generation, for fully automated and highly effective event scheduling during the

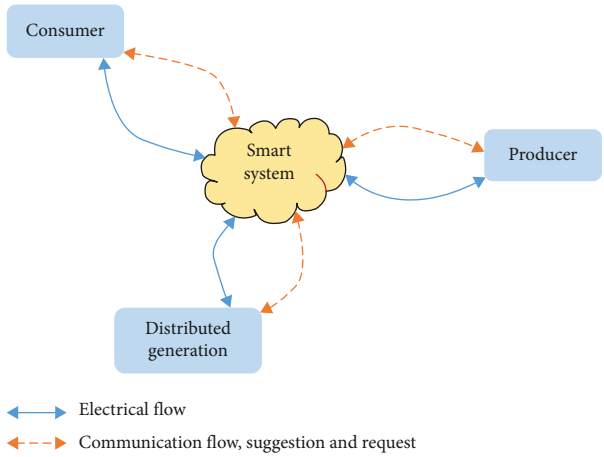

Figure 5: Calendar context for electricity theft detection.

theft detection operations. We identify much more calendar events for the smart electricity theft detection system to learn scheduling personal events about electricity consumption, and we further utilize the system for multiattendee event scheduling in the identification of electricity thieves. The proposed calendar smart context successfully incorporates deep learning techniques such as support vector machine, artificial intelligence, and particle swarm optimization for learning the preferences of each consumer and understanding the consumption behavior for the detection of electricity theft.

In the calendar context, the consumer requests the smart system to schedule electrical energy either from producer or distributed generation. At the same time, the producer provides sufficient energy to the system through communication flow. The distributed generation can give energy to consumers using the smart system to allow permanent availability of energy. The smart system considers each user's preference and calendar context and the purpose of the event to operate objectively and satisfy each element of the framework. In case of electricity fraud, the smart system can switch off the flow of electricity and send the penalty to the fraudster.

We precisely verify that the detected sudden changes are due to theft activities because after collecting the consumption data, we run the data analysis using the deep learning technique; once the electricity fraud is detected, we send a team to the house of consumers and verify if they are at home. Once the verification is done, we can send a penalty bill to consumers.

3.2.6. Evaluation of Performances and Sensitivity Analysis. The dataset used in this work is an unbalanced dataset in which the number of normal consumers varies continuously compared to fraudulent consumers. The dataset is implemented in a biased classifier in order to distinguish fraudulent consumers from regular consumers. To this end, the results were obtained using coefficients such as precision, recall, F1-score, and AUC.

(1) Precision refers to the quotient of correctly categorized positive classes over the total number of positive classes.

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} * 100\%. \tag{14}$$

(2) Recall refers to the rate of correctly classified positive classes.

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} * 100\%. \tag{15}$$

(3) F1-score is more suitable for the unbalanced class distribution including the weighted average of recall and precision.

$$\text{F1} - \text{score} = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}}, \tag{16}$$

where TP, FP, TN, and FN, respectively, are true positive, false positive, true negative, and false negative consumptions. These coefficients represent the possibilities of the consumers being either a true consumer or a fake consumer.

Performance metrics are also detailed based on the electricity theft detection using smart meter data on different datasets. Therefore, fraudulent customers are detected as fraudulent or normal with the highest precision. Moreover, normal customers are detected as fraudulent/normal when the recall and F1-score are highest as possible.

In this study, we performed sensitivity analysis on the process by calculating outcomes of electricity theft detection under alternative electrical assumptions to evaluate the impact of fraudulent consumption. Therefore, the sensitivity analysis allows us to test the robustness of the data analysis results and increase the understanding of the relationships between input and output variables on electrical consumption data. It also allows us to reduce the uncertainty through the identification of the hybrid deep learning model inputs, research the errors in the model, simplify the complexity of the model in the space of input factors, and identify the correlations between observations. These processes aim to study how the uncertainty in the output of the model can be divided and allocated to different sources of uncertainty in its inputs.

The sensitivity analysis is performed using the following operations:

(1) Regression analysis which involves fitting linear regression to the model response

(2) Analysis of variance which quantify the input and output uncertainties as probability distribution as follows:

$$\text{Var}\left(E_{X_{\sim i}}\left(Y \otimes X_i\right)\right). \tag{17}$$

where Var and E, respectively, are the variance and expected value operators and $X_{\sim i}$ is the set of all input variables except for X_i

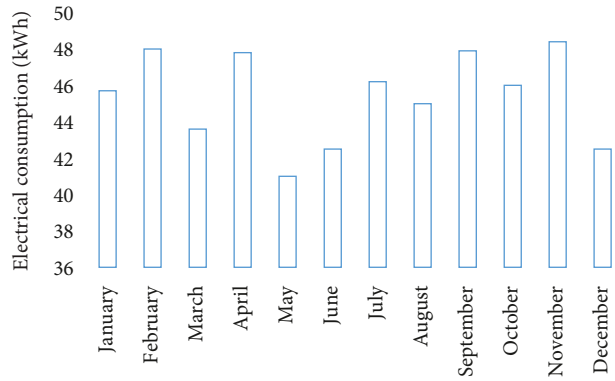

Figure 6: Normal power consumption.

The variance formulation measures the contribution X_i alone to the uncertainty in Y and is known as the first-order sensitivity index

(3) Variogram analysis of response (VARS) which addresses the weakness of directional variograms and covariograms through recognizing a spatially continuous correlation structure to the value Y and hence to the values $(\partial Y/\partial x_i)$

4. Results and Discussion

Data collection allows analysis of customer electricity consumption behavior across demographic trends and social and financial affiliations. The data are collected over a period of one year and contains information from 1000 consumers. Consumers who participated in this research study have installed smart meters in their homes. The normal consumption corresponding to the regularity of customers is presented in Figure 6. In Figure 6, we observe high consumption during the months of February, April, and November. Low consumption is mainly observed in May. For the database analysis, we used a hybrid artificial intelligence model based on the SVM and PSO for the detection of a few potential malicious consumers. Figure 7 gives a comparison between normal consumption and potential electrical energy fraudsters during the one-year period. Sudden variations in the consumption of suspicious customers allow the SVM to detect the fraud that occurs. The SVM therefore made it possible to detect 4 potential fraudsters among all the consumers in the network sample. The performance of the detector of potential electrical energy thieves is evaluated at 98.9%.

In addition, we present simulation results for different data classification methods. So, we divided our dataset for training data and test data. Each algorithm implemented in this work made it possible to obtain performance coefficients such as precision, recall, F1-score, and AUC.

Figures 8, 9, 10, and 11 show the performance coefficient results for all implemented methods. Theoretically, the training database was used for tuning the models, while the testing database was used for aligning the machine learning methods. In addition, the proposed methods were implemented on 3 databases of 400, 600, and 800 consumers, respectively.

Figure 8 shows the accuracy rating for each AI algorithm. This result demonstrates the accuracy in the training and classification of data for the hybrid SVM + PSO model. From this model, we obtain an accuracy of 98.9%, 98.7%, and 98.5%, respectively, for 400, 600, and 800 consumers in the classification and detection of electricity thieves. Most classifiers significantly maintain good accuracy in increasing the number of consumers to analyze except for the ANN model. In addition, we correctly observe that the LSTM and GRU classifiers present an interesting performance when the number of consumers changes from 400 to 800.

Figure 9 gives the recall evaluation for each AI model. It can be observed that the hybrid SVM + PSO model has a better recall percentage compared to the other implemented models. We also observe that the hybrid model gives a recall of 49.2%, 34.1%, and 25.4%, respectively, for 400, 600, and 800 consumers, thus demonstrating its outperformance in detecting fraudulent consumers. The behavior of the LSTM model is similar to the GRU when it gradually increases the recall percentage for each database category. We can clearly observe that the recall value of the SVM + PSO model is significantly better than other comparison models.

Figure 10 shows an evaluation of the F1-score for each proposed algorithm. We observe that the hybrid SVM + PSO algorithm presents a high F1-score value of 28.5, 21.9, and 19.4, respectively, for 400, 600, and 800 consumers, thus showing capabilities in the classification of malicious consumers. Although LSTM has a low F1-score compared to GRU, its value is necessarily higher than that of the KNN, RNN, and CNN. In addition, the ANN gives a relatively low score compared to other implemented methods.

Figure 11 gives the evaluation of the AUC for each model implemented in this work. It is obvious that the hybrid SVM + PSO model gives a better value compared to other methods. The proposed hybrid model gives an AUC of 0.98, 0.97, and 0.95, respectively, for 400, 600, and 800 consumers demonstrating its outperformance in detecting fraudulent

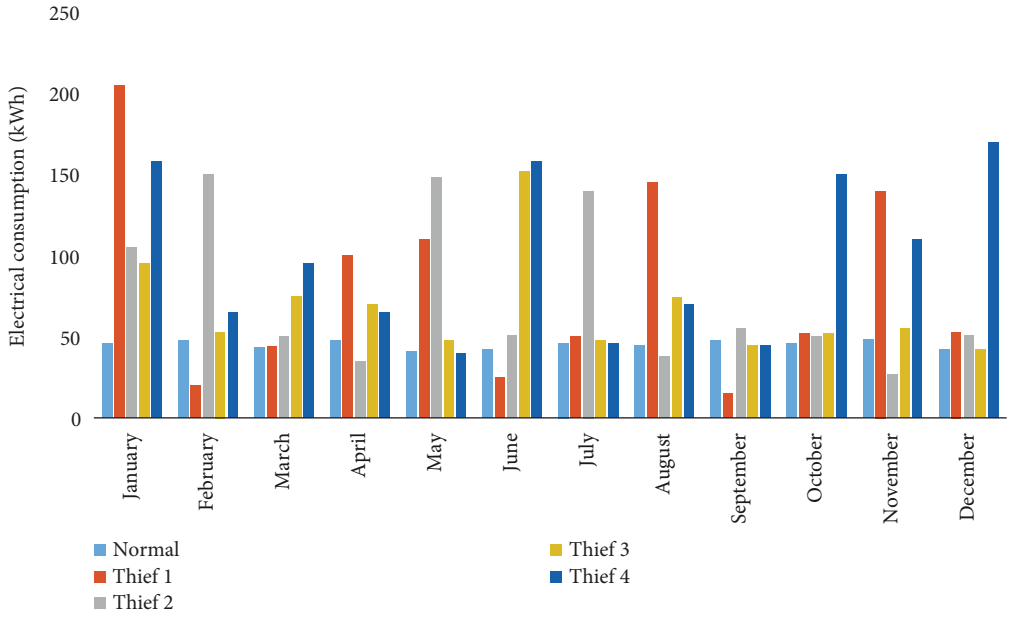

Figure 7: Normal consumers and potential thieves.

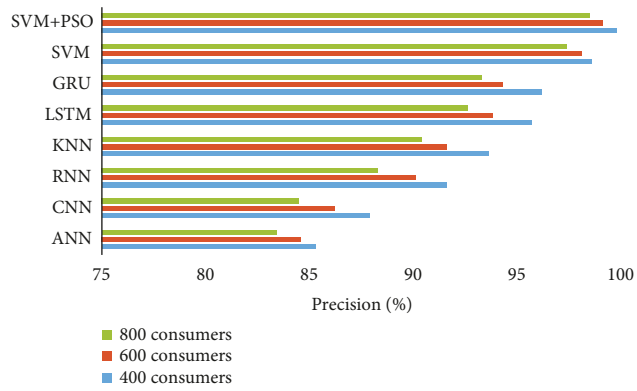

Figure 8: Accuracy assessment for each algorithm.

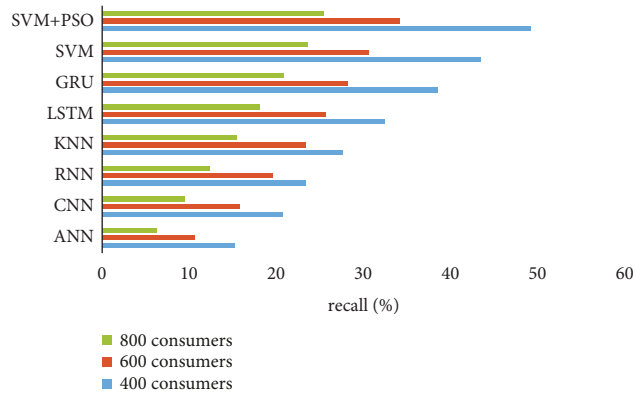

Figure 9: Evaluation of recall for each algorithm.

consumers. In addition, we observe that as the database increases, the AUC value decreases. The ANN model exhibits similar behavior to the CNN despite its reduced data classification capabilities.

Overall, we observe that all the implemented classifiers give different performance values, notably for precision, recall, F1-score, and AUC. Moreover, these performance values are distinct for each database studied.

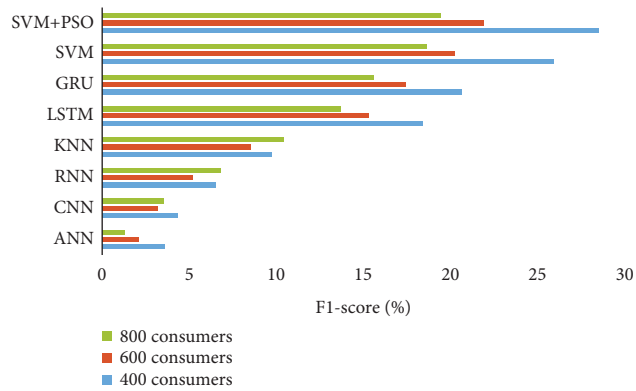

Figure 10: Evaluation of the F1-score for each algorithm.

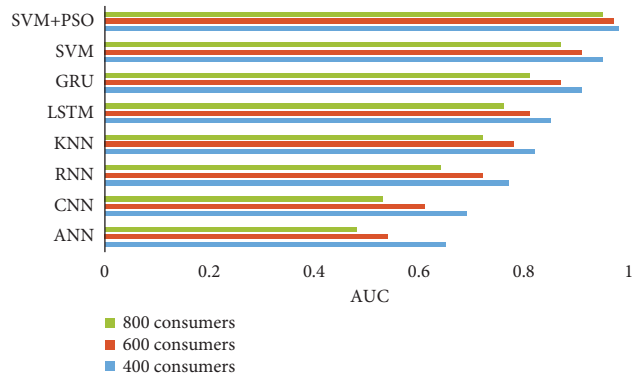

Figure 11: Evaluation of the AUC for each algorithm.

Table 3: Comparison with the literature.

Works	Methods	Precision (%)	Recall (%)	F1-score (%)	AUC
[33]	Machine learning	92.5	32.5	11.5	0.751
[34]	Supervised learning	93.2	37.4	12.6	0.781
[37]	Sparse representation techniques	93.8	38.5	15.4	0.824
[40]	Linear regression model	94.1	38.9	19.7	0.865
[41]	Dynamic programming algorithm	95.5	40.3	21.3	0.877
[43]	Gated recurrent unit	95.6	41.4	22.3	0.924
[47]	XGBoost	96.1	42.3	22.9	0.946
[48]	MLP and LGBM	97.4	43.7	23.5	0.962
[49]	GRU-PSO	97.8	44.6	25.1	0.976
[51]	CNN-RF	98.2	45.7	26.3	0.978
[54]	Deep recurrent neural network	98.5	46.2	27.3	0.982
[57]	Game theoretical models	98.7	48.7	28.1	0.984
Writers	Proposed model	98.9	49.2	28.5	0.989

Table 3 presents a comparison of the results obtained in the literature. We can observe the superiority of our model compared to models in the literature. This is explained because we consider various linguistic variables implemented in our model.

5. Conclusion

This work presented a hybrid AI model for electricity theft detection in a smart power grid. Therefore, we used an OpenDSS-OMNet++ cosimulation platform to simulate the behavior of a real smart grid in order to collect electricity consumption data in real time from smart meters installed for subscribers. Furthermore, we proposed a model using the SVM to determine the correlations between the actual values and the predicted values as well as the PSO algorithm for the optimization of the index parameters of the SVM. Thus, the proposed intelligent model allows us to detect fraudulent values of electrical energy consumption. A consumption dataset of 1000 households was used to verify the effectiveness of the proposed method over one-year period. The simulation results give a performance of 98.9% for the detection of electricity fraud in a smart grid based on data obtained from smart meters. In addition, the detection time

is relatively reduced and the AUC is increased demonstrating the effectiveness of the proposed method compared to that in the literature. This method can be effective for implementation in a larger power network with thousands of consumers, thereby enriching the learning database in the long term. The proposed model can also be optimized by integrating metaheuristic prediction algorithms and wireless connection devices. The limitations of this work concern the nonlinearity of the smart meter data which can affect the performance of the deep learning model. Moreover, further research can be conducted on the implementation of supervised deep learning techniques with different datasets in order to obtain better performance.

Data Availability

The data used to support the findings are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors thank the Electrical Engineering Department of ENSET of the University of Douala.

References

[1] Z. Jiang, R. Lin, and F. Yang, "A hybrid machine learning model for electricity consumer categorization using smart meter data," *Energies*, vol. 11, no. 9, pp. 2235–2319, 2018.

[2] B. P. Bhattarai, S. Paudyal, Y. Luo et al., "Big data analytics in smart grids: state-of-theart, challenges, opportunities, and future directions," *IET Smart grid*, vol. 2, no. 2, pp. 141–154, 2019.

[3] F. G. Yem Souhe, C. F. Mbey, V. J. Foba Kakeu, A. E. Meyo, and A. T. Boum, "Optimized forecasting of photovoltaic power generation using hybrid deep learning model based on GRU and SVM," *Electrical Engineering*, pp. 1–20, 2024.

[4] V. J. Foba Kakeu, A. T. Boum, and C. F. Mbey, "Optimal reliability of smart grid," *International journal of smart grid*, vol. 5, no. 2, pp. 74–82, 2021.

[5] B. Volker, A. Reinhardt, A. Faustine, and L. Pereira, "Watts up at home smart meter data analytics from a consumer-centric perspective," *Energies*, vol. 14, no. 3, p. 719, 2021.

[6] C. F. Mbey, V. J. Foba Kakeu, A. T. Boum, and F. G. Yem Souhe, "Solar photovoltaic generation and electrical demand forecasting using multi-objective deep learning model for smart grid systems," *Cogent engineering*, vol. 11, pp. 1–17, Article ID 2340302, 2024.

[7] F. N. Melzi, A. Same, M. H. Zayani, and L. Oukhellou, "A dedicated mixture model for clustering smart meter data: identification and analysis of electricity consumption behaviors," *Energies*, vol. 10, no. 10, p. 1446, 2017.

[8] D. Alahakoon and X. Yu, "Smart electricity meter data intelligence for future energy systems: a survey," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 1, pp. 425–436, 2016.

[9] C. F. Mbey, F. G. Yem Souhe, V. J. Foba Kakeu, and A. T. Boum, "A novel deep learning based data analysis model for solar photovoltaic power generation and electrical consumption forecasting in smart power grid," *Applied Computational Intelligence and Soft Computing*, pp. 1–22, 2024.

[10] E. W. S. dos Angelos, O. R. Saavedra, O. A. Carmona Cortes, and A. Nunes de Souza, "Detection and identification of abnormalities in customer consumptions in power distribution systems," *IEEE Transactions on Power Delivery*, vol. 26, no. 4, pp. 2436–2442, 2011.

[11] B. Rossi, S. Chren, B. Buhnova, and T. Pitner, "Anomaly detection in smart grid data: an experience report," in *Proceedings of the IEEE international conference on systems, man, and cybernetics (smc)*, Budapest, Hungary, October 2016.

[12] P. Lipcak, M. Macak, and B. Rossi, "Big data platform for smart grids power consumption anomaly detection," *Proceedings of the Federated Conference on Computer Science and Information Systems*, vol. 18, pp. 771–780, 2019.

[13] M. N. Hasan, R. N. Toma, A. A. Nahid, M. M. Manjurul Islam, and J. M. Kim, "Electricity theft detection in smart grid systems: a cnn-lstm based approach," *Energies*, vol. 12, no. 17, p. 3310, 2019.

[14] S. C. Yip, K. Wong, W. P. Hew, M. T. Gan, C. W. Phan, and Su W. Tan, "Detection of energy theft and defective smart meters in smart grids using linear regression," *International Journal of Electrical Power and Energy Systems*, vol. 91, pp. 230–240, 2017.

[15] G. Dudek, A. Gawlak, M. Kornatka, and J. Szkutnik, "Analysis of smart meter data for electricity consumers," in *Proceedings of the 15th International Conference on the European Energy Market (EEM)*, pp. 1–5, Lodz, Poland, June 2018.

[16] R. Jiang, R. Lu, Y. Wang, J. Luo, C. Shen, and X. Shen, "Energy-theft detection issues for advanced metering infrastructure in smart grid," *Tsinghua Science and Technology*, vol. 19, no. 2, pp. 105–120, 2014.

[17] S. S. S. R. Depuru, L. Wang, V. Devabhaktuni, and P. A. Nelapati, "Hybrid neural network model and encoding technique for enhanced classification of energy consumption data," *IEEE Power and Energy Society General Meeting*, vol. 25, pp. 1–8, 2011.

[18] S. C. Yip, W. N. Tan, C. W. Tan, M. T. Gan, and K. Wong, "An anomaly detection framework for identifying energy theft and defective meters in smart grids," *International Journal of Electrical Power and Energy Systems*, vol. 101, pp. 189–203, 2018.

[19] M. Sodenkamp, I. Kozlovskiy, K. Hopf, and T. Staake, "Smart meter data analytics for enhanced energy efficiency in the residential sector," in *Proceedings of the 13th International Conference on Wirtschaftsinformatik*, pp. 1235–1249, St.Gallen, Switzerland, February 2017.

[20] S. Barker, S. Kalra, D. Irwin, and P. Shenoy, "Empirical characterization, modeling, and analysis of smart meter data," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 7, pp. 1312–1327, 2014.

[21] A. J. Nezhad, T. K. Wijaya, M. Vasirani, and K. Aberer, "Smartd: smart meter data analytics dashboard," in *Proceedings of the Engineering, Computer Science Proceedings of the 5th international conference on Future energy systems*, pp. 1-2, Chengdu China, April 2013.

[22] Z. Fengming, L. Shufang, G. Zhimin, W. Bo, T. Shiming, and P. Mingming, "Anomaly detection in smart grid based on encoder-decoder framework with recurrent neural network," *The Journal of China Universities of Posts and Telecommunications*, vol. 24, no. 6, pp. 67–73, 2017.

[23] X. Liu and P. S. Nielsen, "Regression-based online anomaly detection for smart grid data," 2016, https://arxiv.org/pdf/1606.05781.

[24] T. Hartmann, A. Moawad, F. Fouquet et al., "Suspicious electric consumption detection based on multiprofiling using live machine learning," in *Proceedings of the IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pp. 891–896, Glasgow, Scotland, November 2015.

[25] M. Martinez Pabon, T. Eveleigh, and B. Tanju, "Smart meter data analytics for optimal customer selection in demand response programs," *Energy Procedia*, vol. 107, pp. 49–59, 2017.

[26] N. Lu, P. Du, X. Guo, and F. L. Greitzer, "Smart meter data analysis," in *Proceedings of the Conference: Transmission and Distribution Conference and Exposition, 2012 IEEE PES*, Orlando, FL, USA, May 2012.

[27] World Bank Group, *Data Analytics for Advanced Metering Infrastructure: A Guidance Note for South Asian Power Utilities*, International Bank for Reconstruction and Development, Washington, DC, USA, 2018.

[28] X. Liu, L. Golab, W. Golab, I. F. Ilyas, and S. Jin, "Smart meter data analytics: systems, algorithms and benchmarking," *ACM Transactions on Database Systems*, vol. 42, no. 1, pp. 1–39, 2016.

[29] E. McKenna, I. Richardson, and M. Thomson, "Smart meter data: balancing consumer privacy concerns with legitimate applications," *Energy Policy*, vol. 41, pp. 807–814, 2012.

[30] C. Flath, D. Nicolay, T. Conte, C. van Dinther, and L. Filipova-Neumann, *Cluster Analysis of Smartmetering Data*, Research Center for Information Technology, Karlsruhe Germany, 2011.

[31] D. Mashima and A. A. Cardenas, "Evaluating electricity theft detectors in smart grid networks," in *Proceedings of the 15th international conference on Research in Attacks, Intrusions, and Defenses*, Amsterdam The Netherlands, September 2012.

[32] N. M. Ibrahim, S. T. Faraj Al Janabi, and B. Al-Khateeb, "Electricity-theft detection in smart grids based on deep learning," *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 4, pp. 2285–2292, 2021.

[33] M. Nabil, M. Ismail, M. Mahmoud, and E. Serpedin, "Recurrent electricity theft detection in ami networks with evolutionary hyper-parameter tuning," in *Proceedings of the 2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, pp. 1002–1008, Rhodes, Greece, July 2019.

[34] F. A. Bohani, A. Suliman, M. Saripuddin, S. S. Sameon, N. S. Md Salleh, and S. Nazeri, "A comprehensive analysis of supervised learning techniques for electricity theft detection," *Journal of Electrical and Computer Engineering*, vol. 2021, pp. 1–10, Article ID 9136206, 2021.

[35] J. Pereira and F. Saraiva, "A comparative analysis of unbalanced data handling techniques for machine learning algorithms to electricity theft detection," in *Proceedings of the 2020 IEEE Congress on Evolutionary Computation (CEC)*, Glasgow, UK, July 2020.

[36] R. N. Toma, M. Nazmul Hasan, A. A. Nahid, and B. Li, "Electricity theft detection to reduce non- technical loss using support vector machine in smart grid," in *Proceedings of the 1st International Conference on Advances in Science, Engineering and Robotics Technology 2019 (ICASERT 2019)*, Dhaka, Bangladesh, May 2019.

[37] M. Lydia, G. Edwin Prem Kumar, and Y. Levron, "Detection of electricity theft based on compressed sensing," in *Proceedings of the 5th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India, March 2019.

[38] M. Adil, N. Javaid, Z. Ullah, M. Maqsood, S. Ali, and M. A. Daud, "Electricity theft detection using machine learning techniques to secure smart grid," in *Proceedings of the 14th International Conference on Complex, Intelligent and Software Intensive System (CISIS 20)*, Lodz, Poland, July 2020.

[39] S. Mujeeb, N. Javaid, R. Khalid, M. Imran, and N. Naseer, "De-rusboost: an efficient electricity theft detection scheme with additive communication layer," in *Proceedings of the IEEE International Conference on Communications (ICC)*, pp. 1–6, Dublin, Ireland, June 2020.

[40] S. C. Yip, C. K. Tan, W. N. Tan, M. T. Gan, and A. H. Abu Bakar, "Energy theft and defective meters detection in ami using linear regression," in *Proceedings of the IEEE International Conference on Environment and Electrical Engineering and IEEE Industrial and Commercial Power Systems Europe (EEEIC ICPS Europe)*, pp. 1–6, Milan, Italy, June 2017.

[41] M. Sivarathinabala and T. Niruban Projoth, "Energy theft detection in multi-tenant data centers and distribution line using smart grids," *ARPN Journal of Engineering and Applied Sciences*, vol. 14, no. 19, 2019.

[42] F. Shehzad, M. Asif, Z. Aslam et al., "Comparative study of data driven approaches towards efficient electricity theft detection in micro grids," in *Proceedings of the IEEE International conference*, Asan, Korea, July 2021.

[43] A. Pamir, A. Ullah, S. Munawar, M. Asif, B. Kabir, and N. Javaid, "Synthetic theft attacks implementation for data balancing and a gated recurrent unit based electricity theft detection in smart grids," in *Proceedings of the IEEE International conference*, Asan, Korea, July 2021.

[44] A. O. Otuoze, M. W. Mustafa, I. E. Sofimieari et al., "Electricity theft detection framework based on universal prediction algorithm," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 15, no. 2, pp. 758–768, 2019.

[45] L. Wei, A. Sundararajan, A. I. Sarwat, S. Biswas, and E. Ibrahim, "A distributed intelligent framework for electricity theft detection using benford law and stackelberg game," in *Proceedings of the resilience Week (RWS)*, pp. 5–11, Wilmington, DE, USA, September 2017.

[46] M. Adil, N. Javaid, U. Qasim, I. Ullah, M. Shafiq, and J. G. Choi, "Lstm and batbased rusboost approach for electricity theft detection," *Applied Sciences*, vol. 10, no. 12, pp. 4378–4421, 2020.

[47] R. Punmiya and S. Choe, "Tou pricing-based dynamic electricity theft detection insmart grid using gradient boosting classifier," *Applied Sciences*, vol. 11, pp. 401–415, 2021.

[48] A. Arif, "Employing machine learning and deep learning models for electricity theft detection in smart grids," COMSATS University, Islamabad, Pakistan, 2020, Ph.D.thesis.

[49] A. Ullah, N. Javaid, O. Samuel, M. Imran, and M. Shoaib, "Cnn and gru based deep neural network for electricity theft detection to secure smart grid," in *Proceedings of the International Wireless Communications and Mobile Computing (IWCMC)*, pp. 1598–1602, Limassol, Cyprus, June 2020.

[50] S. Aziz, S. Z. Hassan Naqvi, M. U. Khan, and T. Aslam, "Electricity theft detection using empirical mode decomposition and k-nearest neighbors," in *Proceedings of the International Conference on Emerging Trends in Smart*

Technologies (ICETST), pp. 1–5, Karachi, Pakistan, March 2020.

[51] S. Li, Y. Han, X. Yao, S. Yingchen, J. Wang, and Q. Zhao, "Electricity theft detection in power grids with deep learning and random forests," *Journal of Electrical and Computer Engineering*, vol. 2019, pp. 1–12, 2019.

[52] K. Zheng, Q. Chen, Y. Wang, C. Kang, and Q. Xia, "A novel combined datadriven approach for electricity theft detection," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 3, pp. 1809–1819, 2019.

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

[54] M. Nabil, M. Ismail, M. Mahmoud, M. Shahin, K. Qaraqe, and E. Serpedin, "Deep recurrent electricity theft detection in ami networks with random tuning of hyperparameters," in *Proceedings of the 24th International Conference on Pattern Recognition (ICPR)*, Beijing, China, August 2018.

[55] C. Richardson, N. Race, and P. Smith, "A privacy preserving approach to energy theft detection in smart grids," *Austrian institute of technology*, 2016.

[56] Y. Gao, B. Foggo, and N. Yu, "A physically inspired data-driven model for electricity theft detection with smart meter data," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 9, pp. 5076–5088, 2019.

[57] S. Amin, G. A. Schwartz, A. A. Cardenas, and S. S. Sastry, "Game theoretic models of electricity theft detection in smart utility networks," *IEEE Control Systems*, vol. 35, no. 1, pp. 66–81, 2015.

[58] F. Xiao and Q. Ai, "Electricity theft detection in smart grid using random matrix theory," *IET Generation, Transmission and Distribution*, vol. 12, no. 2, pp. 371–378, 2018.

[59] R. K. Ahir and B. Chakraborty, "Pattern-based and context-aware electricity theft detection in smart grid," *Sustainable Energy, Grids and Networks*, vol. 32, 2022.

[60] C. F. Mbey, V. J. Foba Kakeu, A. T. Boum, and F. G. Yem Souhe, "Fault detection d classification using deep learning method and neuro-fuzzy algorithm in a smart distribution grid," *Journal of Engineering*, vol. 1, pp. 1–9, 2023.

[61] F. G. Yem Souhe, A. T. Boum, P. Ele, C. F. Mbey, and V. J. Foba Kakeu, "A novel smart method for state estimation in a smart grid using smart meter data," *Applied Computational Intelligence and Soft Computing*, vol. 2022, pp. 1–14, 2022.