Journal of Science Research and Reviews
PRINT ISSN: 1595-9074
E-ISSN: 1595-8329
DOI: https://doi.org/10.70882/josrar.2026.v3i5.264
Homepage: https://josrar.esrgngr.org
Review Article

A Review of Machine Learning-Based Theft Detection for Smart Energy Meters

***Oladipupo Olatunji, Victor Edibo and Olumide Alamu**

Department of Electrical and Electronics Engineering, University of Lagos, Akoka, Lagos, Nigeria
*Corresponding Author's email: oladipupoolatunji272@gmail.com

KEYWORDS
Smart Grid,
Advanced Metering Infrastructure,
Machine Learning,
Power System,
Internet of Things,
Supervised learning.

ABSTRACT
The implementation of Advanced Metering Infrastructure (AMI) in smart grid environments has led to a paradigm shift in energy consumption management by means of high-throughput bi-directional communication. Unfortunately, the adoption of AMI is associated with growing Non-Technical Losses (NTL) such as electricity theft, estimated at global losses exceeding $96 billion a year, threatening the stability and resilience of electricity grids. As a response to the emerging challenge, traditional NTL detection, based on periodic field audits and threshold-based inspections, has been proven to be ineffective in addressing stochastic and increasingly intelligent cyber-physical manipulations. Hence, this paper provides a literature review of Machine Learning applications in the domain of Electricity Theft Detection (ETD) that is grounded in 71 scholarly studies produced between 2021 and 2026 and selected according to criteria of machine learning application in the process of theft/NTL detection and accompanying empirical validation. Out of those, 44 primary research works have been organized into tabulation and contrasted with regard to eight distinct categories of methods including classical ML, deep learning, hybrid and reinforcement and federated learning techniques. The comparative evidence reveals that the hybrid and stacked architectures provide the best detection performance whereas federated designs typically sacrifice their maximum accuracy in favor of better privacy and deployability. Challenges in assessing the comparative gains are discussed, and it is revealed that benchmarked datasets are scarce and inconsistent as well as only 9 of 44 studies included in the table reported their false positive rates that determine audit expenses.

CITATION
Olatunji, O., Edibo, V., & Alamu, O. (2026). A Review of Machine Learning-Based Theft Detection for Smart Energy Meters. *Journal of Science Research and Reviews*, *3*(5), 73-103. https://doi.org/10.70882/josrar.2026.v3i5.264

INTRODUCTION

The transition from the legacy electrical grid to the more advanced smart energy grid marks a radical revolution in energy generation, delivery, and consumption management. At the core of this revolution lies the Advanced Metering Infrastructure (AMI) technology, which includes IoT-enabled smart energy meters as an integral component (Qi et al., 2022; Tanwar et al., 2022). The interconnected nature of smart energy meters enables bidirectional communication between the utility provider and the end consumer, creating an extensive data environment that supports efficient energy consumption monitoring and demand-side energy management (El-Toukhy et al., 2023). The high-frequency, high-dimensional telemetry from connected IoT devices forms the

foundation of the data substrate that enables energy grid optimisation and the detection of anomalous behaviour.

Despite the advantages of implementing an AMI system, the technological leap in energy grids has also introduced new cyber-physical threats in the form of Non-Technical Losses (NTL). This refers to all forms of energy losses, excluding any physical attributes of the power grid, among which electricity theft represents the largest economic burden (Ibrahim et al., 2021). Not only does electricity theft lead to losses exceeding $96 billion USD per annum for global power utilities (Qi et al., 2022; Hussain et al., 2021), but it also causes systematic instabilities in the electric grid, resulting in frequent outages and potentially deadly consequences, such as fires and electrocutions (Cheng et al., 2021).

Traditionally, approaches to combating such problems include classical methods such as hardware state analysis, statistical auditing, and manual inspection (Lepolesa et al., 2022; Lin et al., 2021). However, these methods implicitly assume a constant and predictable consumer consumption profile (Cheng et al., 2021; Gu et al., 2022). In modern smart grids, however, energy thieves increasingly use complex, stochastic attack methods, such as intermittent energy theft, False Data Injection (FDI) attacks, and other stealth mechanisms that mimic the natural volatility of consumer behaviour. The dynamic and evolutionary nature of modern threat signatures makes classical detection methods unsuitable and ineffective, leading to excessively high false-positive rates and making the process prohibitively expensive.

Due to the fundamental inefficiencies of the classical detection methodology, considerable research interest has recently emerged in applying machine learning (ML) algorithms to process the vast, high-dimensional data streams arising from AMI deployments (Fang et al., 2023; Yan & Wen, 2021). A broad range of artificial intelligence tools, including both classical ML approaches and modern deep learning architectures, have been studied for their ability to automatically detect complex spatiotemporal features and discrimination factors. Moreover, reinforcement learning systems have also been explored for their ability to address the evolutionary nature of contemporary energy theft attack vectors. Despite the considerable progress made in that direction, significant challenges remain in using such methods to solve Electricity Theft Detection (ETD) problems. Some of the major hurdles are class imbalance among legitimate and illegitimate users, absence of labeled fraud cases, poor generalizability to other networks, and computation limitations in edge devices.

Given these important directions and challenges, this paper reviews recent advances in ML-based approaches to tackling the ETD problem. By analysing the recent literature on classical ML techniques, deep learning models, ensemble architectures, and federated ML frameworks with respect to their ability to address a variety of modern attack scenarios, this review offers a valuable perspective on future research directions. The main contribution of this work is the critical evaluation of algorithmic approaches to solving ETD problems, as well as the identification of important technical challenges to address in the future to develop a highly effective solution.

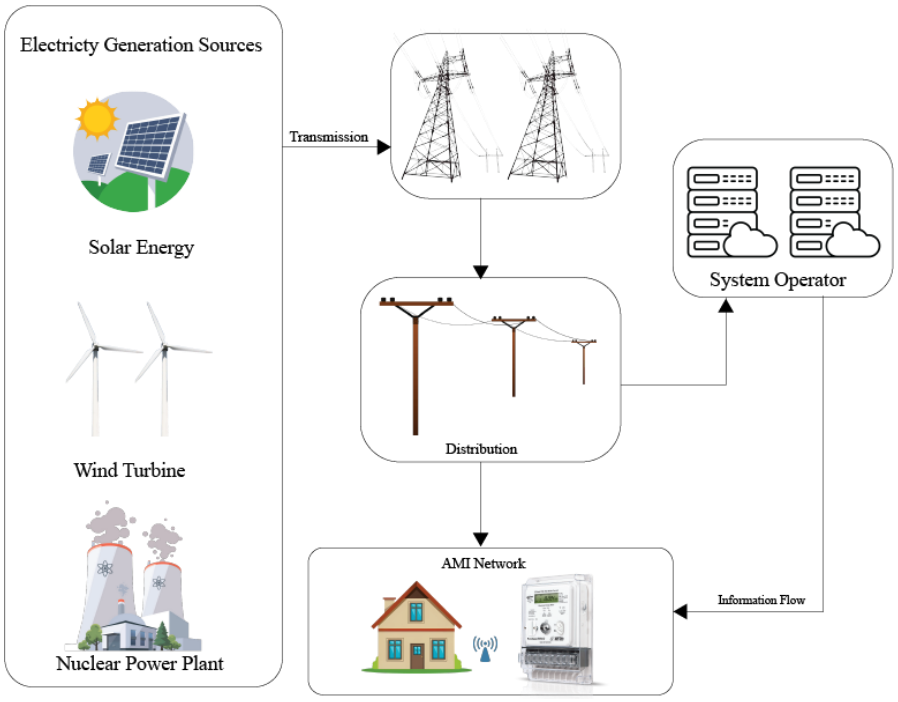

Figure 1: Smart Grid Architecture

Review Methodology

The present work is described as a structured review and not a systematic review. The corpus of studies has been developed and analyzed consistent with the scope, inclusion criteria, and synthesis framework described below. These aspects have been provided in order to provide an independent evaluation of the limits of the review, the criteria used for inclusion, and the comparative framework used. No attempt has been made to prospectively document the yield from searches using a PRISMA style approach.

Review Questions

The review is organised around five questions, each answered in an identified section of the paper:

RQ1. Which machine learning methods have been used for detecting electricity theft in smart metering systems, and what changes in their usage have been observed between 2021 and 2026? (Sections 5 and 6)

RQ2. What data sets and attacks scenarios have been used to devise and test these methods, and how well do they reflect the actual deployed AMI environment? (Sections 6 and 6.4)

RQ3. What measures have been used for evaluation, and does modern practice support cross-comparison between studies? (Section 6.4)

RQ4. What failure modes and practical limitations such as class imbalance, computational complexity, adversarial evasion, and distributional shifts hamper real-world usage? (Sections 7 and 8)

RQ5. What research areas remain to be explored, especially in regions where smart grids and AMIs have not yet been widely deployed? (Sections 9 and 10)

Scope and Search Strategy

Literature identification was accomplished by linking three conceptual domains – the loss phenomenon, the analytical approach, and the metering environment – through the use of Boolean operators. The search string used was:

("electricity theft" OR "energy theft" OR "non-technical loss" OR NTL) AND ("machine learning" OR "deep learning" OR "neural network" OR "anomaly detection") AND ("smart meter" OR "smart grid" OR "advanced metering infrastructure" OR AMI)

Search results were limited to peer-reviewed journal articles and conference papers published from January 2021 through February 2026 and written in English. This ensures that the evidentiary base captures the current status of research on the topic. The resulting set of literature includes 71 publications, of which 21 have been published in IEEE journals, 18 in Elsevier journals, 16 in MDPI, 6 in Hindawi and Wiley, 3 in IET, 3 in Frontiers, and 4 elsewhere.

Inclusion and Exclusion Criteria

A study was considered relevant if it met all three criteria: (1) it suggested, utilized, or assessed an ML or DL approach to electricity theft and non-technical loss detection; (2) at least one performance metric was provided; and (3) the problem dealt with a smart metering, AMI, or distribution network setting.

Studies that did not meet any one of the following criteria were not included: (1) if the paper focused solely on technical losses without the element of theft or fraud; (2) if the detection process involved rules or thresholds or any other hardware solution but no learning was involved; (3) if the paper did not contain an empirical assessment; (4) if the full text of the paper was not accessible in English; and (5) if the paper was preceded by a more comprehensive version published later by the same authors.

Quality Appraisal

The following selection criteria were applied to all included studies: (a) the clarity of the problem statement and the underlying attack model; (b) the clarity of the dataset and pre-processing chain; (c) the appropriateness of the evaluation methodology, in terms of the handling of class imbalance and use of appropriate evaluation metrics; and (d) the ease of reproduction, including access to the data or the code. Studies that did not fulfill criteria (a) or (c) were not included in any of the comparison tables; studies that did fulfill the criteria but had limitations were included in the summary tables with their limitations noted.

Data Extraction and Synthesis Framework

The same data extraction procedure was used in all the considered papers, which included the following elements: method class, model structure, the type of data used, types of attacks that were simulated and the strategy used to solve the problem of class imbalance, as well as the evaluation procedure, performance and the limitations of the proposed model mentioned by the authors of the research. The aforementioned extraction procedure was the analytical framework used in the current review.

Based on the application of the analytical framework, 44 articles could be classified into eight classes. The list of articles used is shown in Table 1 below and sorted by the following characteristics: method class, the dataset used, the evaluation measure and the position of the article in the paper. The results of per-class data extraction are presented in Tables 4 to 11 below, while the cross-class comparison is provided in Section 6.4.

Table 1: Mapping of the reviewed corpus by method class, dataset and evaluation metric

Method class	Studies (n)	Datasets named in the reviewed studies	Evaluation metrics reported
Tree-based and classical ensembles	13	SGCC; PLN (Indonesia) postpaid data; other utility records; synthetic injection	AUC, Accuracy, Precision, Recall, F1, DR, FPR
Support vector machines and kernel methods	4	Not consistently named in the reviewed reports	AUC, Accuracy, Precision, Recall, F1, DR, FPR, MAP
Convolutional neural networks	5	SGCC; synthetic theft injection	AUC, Accuracy, Precision, Recall, F1, DR, FPR
Recurrent networks (RNN, LSTM)	4	Synthetic and simulated consumption series	Accuracy, Precision, Recall, F1, FPR, RMSE, MAP
Hybrid deep and evolutionary architectures	6	SGCC; Irish CER; synthetic and simulated data	AUC, Accuracy, Precision, Recall, F1, MAP
Meta-learning and stacked ensembles	5	Utility records; synthetic theft injection	AUC, Accuracy, Precision, Recall, F1, MAE, FPR
Adaptive and reinforcement learning	3	Irish CER; synthetic attack scenarios	AUC, Accuracy, Precision, Recall, F1, DR, FPR
Federated learning and forecasting	4	Simulated distributed partitions; synthetic data	AUC, Accuracy, DR, FPR, MAE, RMSE
Total tabulated primary studies	44	—	—

Related Works

This section provides an overview of pertinent studies on Electricity Theft Detection (ETD) and security issues in the context of Smart Grids (SGs) and AMIs based on selected works from the relevant literature. A summary of these related papers is presented in Table 2. The use of Anomaly Detection Techniques (ADTs) for detecting data manipulation attacks in AMIs has been covered in (Al-Ghaili et al., 2021). Criteria for evaluating the performance of Anomaly Detection Techniques include detection accuracy, computational complexity, and node trust. To improve detection accuracy, a feature engineering technique is introduced that utilises the Finite Mixture Model (FMM) clustering algorithm and Genetic Programming (GP) techniques. Behavioural metrics for measuring trust in AMI nodes have also been utilised, and data manipulation attacks are categorised into four classes: additive, deductive, camouflage, and conflict. The review of recent research efforts on deep learning approaches for electricity fraud detection presented in (Badr et al., 2023) differs from the anomaly-centric framework proposed in (Al-Ghaili et al., 2021). Specifically, the authors have studied privacy-preserving countermeasures against adversarial attacks on data integrity and authenticity in smart meters, including encryption, federated learning (FL), secure multi-party computation (SMC), and functional encryption. In addition, the authors have identified several important gaps in the literature on net-metering fraud detection. In addition, the review highlights the vulnerability of machine learning detection models to adversarial attacks, namely poisoning and evasion, and suggests ensemble learning and distillation as defence mechanisms. The paper (Kgaphola et al., 2024) discusses various IoT-based, big data, and ML-driven approaches for securing SGs and AMIs. The empirical analysis shows that CNNs are the most common classification technique, ensemble DL networks are used for real-time detection, and Deep Siamese Networks mitigate class imbalance. An important gap in the literature concerns the scarcity of solutions for ETD in jurisdictions without SG/AMI infrastructure. Moreover, the researchers note that most existing ETD approaches are reactive rather than preventive. In (Yan & Wen, 2022), the authors provide a taxonomy of existing ETD approaches, including game-theoretic, power grid analysis, hardware anti-theft, and optimisation-driven methods for analysing consumption patterns using ML. Some of the most important problems identified by the authors include the lack of labelled data, imbalanced class distributions, privacy concerns, and poor generalizability across regions. Several publicly available datasets include the State Grid Corporation of China (SGCC) dataset, Irish Smart Energy Trial (ISET), and UMass Smart dataset. According to the authors, deep learning models, specifically convolutional neural networks (CNN) and recurrent neural networks (RNN), as well as ensemble models such as XGBoost, have achieved the best results in terms of detecting electricity thefts. At the same time, the standardisation of ETD systems and the limitations in protecting consumers' personal information have been recognised as important research problems. In (Stracqualursi et al., 2023), the evolution of anomaly detection methodologies from harmonic. The paper (Kim et al., 2024) analyses current developments in the field of semisupervised approaches, i.e., combining supervised and unsupervised methods such as pseudolabeling and

model ensembles, while discussing the emergence of generative AI technology as a promising tool for generating new labelled examples, detecting anomalies, and addressing imbalance issues. Additionally, problems related to high dimensionality, meter-reading errors, and label noise in publicly available datasets, such as SGCC and CER, have been addressed.

Motivation And Contributions

The rapid spread of AMI and the challenges associated with its implementation require the development of new control and maintenance methods to ensure maximum efficiency. The increased use of digital technologies in AMI (Al-Ghaili et al., 2021) makes the infrastructure vulnerable to a range of cyber-attacks characteristic of digitalised systems. The critical analysis of surveys (Al-Ghaili et al., 2021; Badr et al., 2023), and (Kgaphola et al., 2024) highlights significant coverage gaps that warrant a review. Firstly, (Al-Ghaili et al., 2021) discusses only Deep Neural Networks (DNN), Recurrent Neural Networks (RNN), and Artificial Neural Networks (ANN), whereas several other popular machine learning architectures, such as Feed-forward Neural Networks (FFNN) and Convolutional Neural Networks (CNN), can be used for anomaly detection. The review of the current state of ETD technology in (Badr et al., 2023) does not cover all the parameters of an optimal ETD system design. Moreover, (Kgaphola et al., 2024) directly mentions that the lack of SG and AMI is a significant gap in the current research and that there are no appropriate technologies for detecting meter bypass attacks in such territories.

It should be emphasised that machine learning algorithms, artificial intelligence, and big data analysis have become the most prevalent tools for meter bypass detection in SG and AMI. However, it is impossible to detect anomalies with a single classification model, which is why combining several techniques is common in state-of-the-art devices. Therefore, in what follows, we highlight our contributions to the existing research aimed at addressing some of the aforementioned gaps:

1. To have proper background knowledge from a technical point of view, we offer an overview of the different types of ML algorithms used in smart energy networks, namely supervised, unsupervised, reinforcement, and deep learning algorithms, presenting the most popular approaches for ETD as well as key steps required for their effective use in practice.
2. Given the above-mentioned types of ML algorithms, we critically assess the state-of-the-art solutions proposed in the literature on ETD by analysing the type of attack scenario, algorithm applied, findings made by the authors, and possible limitations of their studies, especially related to the lack of research in ETD for regions with no SG and AMI technology infrastructure, with multiple attack vectors taken into consideration.
3. Finally, we identify several challenges for developing effective ETD systems using ML algorithms, namely severe class imbalance, limited data samples, limited generality and scalability, and computational challenges in IoT, and outline further directions for addressing these limitations and creating efficient ML-based ETD solutions.

The rest of the paper is organized as follows. Section 2 discusses the methodology of the review, which includes review questions, scope, inclusion criteria and the model used to analyze the selected articles. Section 3 discusses previous review papers and points out the gaps covered in this paper. Section 4 explains the motivation and contributions. Section 5 gives a brief introduction to machine learning algorithms relevant to the smart energy grid. Section 6 gives an in-depth review on machine learning algorithms used for electricity theft detection in classes and concludes with the comparison of the studies discussed. Sections 7, 8, and 9 give the lessons learned, challenges in the design and future research areas, respectively. Section 10 describes recent research trends, while section 11 gives the conclusion of the study. In the whole paper, the acronyms used are defined in table 3.

The issue of electricity theft has become a burning problem in the modern smart grid environment because it causes revenue losses, grid instability, increased technical losses, and the inability to plan operations (Yan & Wen, 2022). Although sources have proposed methods for identifying theft, the literature has shown deficiencies in identification accuracy, data accessibility, standard procedures, and the usability of intelligent systems. In particular, the authors consider the following key challenges: limited generalisation across other datasets and consumer types, insufficient data, scalability issues related to distribution networks, and the absence of an all-round anti-theft protection system that combines algorithms, software, and implementation approaches from the literature. In addition, according to (Stracqualursi et al., 2023), the implementation of ETD increasingly relies upon artificial intelligence methods rather than static, rule-based auditing.

Table 2: Summary of review papers related on utilizing ML for Electricity Theft Detection

Ref	Year	Main Contribution	Machine Learning Techniques							
			DNN	FFNN	RNN	CNN	ANN	DT	RF	SVM
(Al-Ghaili et al., 2021)	2021	Anomaly detection techniques in smart grid systems; evaluation based on Accuracy, Trust, and Computation Time.	✓	✗	✓	✗	✓	✓	✗	✓
(Badr et al., 2023)	2022	Review of data-driven methods for electricity fraud detection; adversarial attack defenses and consumer privacy.	✓	✓	✓	✓	✗	✓	✓	✓
(Kgaphola et al., 2024)	2024	Systematic review of technology-based ETD solutions and their operational effectiveness.	✓	✗	✗	✓	✗	✓	✓	✓
(Yan & Wen, 2022)	2022	Comprehensive classification and comparative analysis of ETD methods; ML, grid analysis, hardware, and game theory approaches.	✗	✗	✓	✓	✓	✓	✓	✓
(Stracqualursi et al., 2023)	2023	Systematic review integrating energy theft practices with AI detection; generalized AI solution guidelines.	✗	✗	✓	✓	✓	✗	✓	✓
(Kim et al., 2024)	2024	Review of data-driven ETD approaches; challenges of limited and imbalanced datasets; generative AI for data augmentation.	✗	✗	✓	✓	✓	✓	✓	✓
This Paper	2026	Comprehensive review of ML techniques applicable to AMI-based smart energy meter theft detection across diverse attack scenarios.	✓	✓	✓	✓	✓	✓	✓	✓

DNN → Deep Neural Networks, RNN → Recurrent Neural Networks, ANN → Artificial Neural Network, DT → Decision Tree, RF → Random Forest, SVM → Support Vector Machine, FFNN → Feed-forward Neural Networks, CNN → Convolutional Neural Networks.

A Brief Overview of ML Techniques

In this section, we introduce the reader to the different ML methods employed in smart energy networks by simply identifying their key features. We do not attempt to provide a detailed, comprehensive analysis of all ML methods, which have been exhaustively covered in other reports (Yan & Wen, 2022), but we aim to provide the reader with some groundwork on their use in ETD. In traditional ML classification, algorithms can be divided into supervised, unsupervised, and reinforcement learning (Bohani et al., 2021).

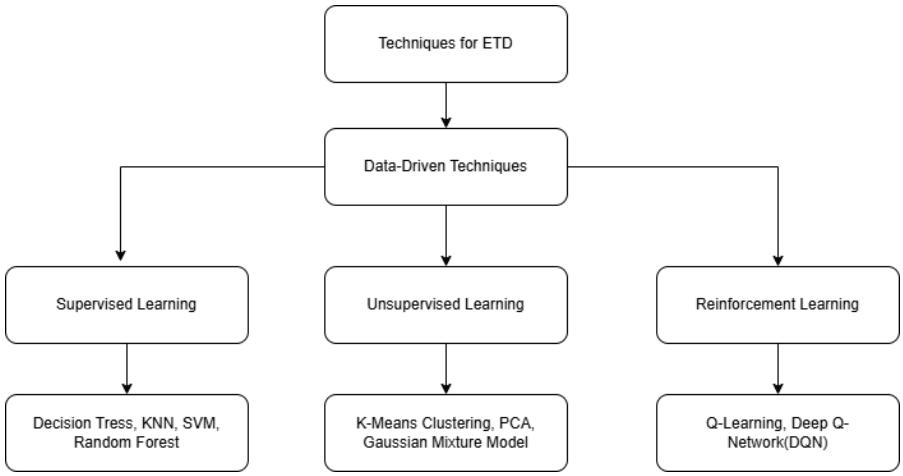

Figure 2: Taxonomy of ML Techniques in Electricity Theft Detection

Supervised learning

The objective of supervised learning is to learn predictive models from labelled training data, where a sample's target variable (or corresponding output label) is known (Alamu et al., 2025; Bohani et al., 2021). The approach is vital in ETD due to supervised algorithms' ability to classify consumption patterns as fraudulent or legitimate, as well as to predict future short- and long-term loads using regression models (Dash et al., 2021). In general, supervised algorithms exhibit excellent convergence properties and achieve high detection accuracy against recognised attack signatures. The main drawback of the technique is that it depends on extensive, accurately labelled datasets that are expensive to obtain, subject to labelling errors, and typically prone to severe class imbalance (Qi et al., 2022; Gu et al., 2022). Examples of supervised learning algorithms include Decision Trees (DT), K-Nearest Neighbours (KNN), Support Vector Machines (SVM), and Random Forests (RF).

Unsupervised Learning

The core difference between supervised and unsupervised paradigms lies in the absence of a priori knowledge of class labels in the latter (Qi et al., 2022). Unsupervised algorithms are especially relevant for detecting outliers and anomalies in ETD because it is impractical to collect extensive labelled datasets in operational conditions, as well as for detecting new zero-day attacks that do not have a representation in the training set (El-Toukhy et al., 2023). The downside of the technique is its tendency to yield higher False Positive Rates (FPRs), due to an inability to distinguish between legitimate changes in consumer behaviour patterns and fraud attack signatures (Gu et al., 2022). Representative examples of unsupervised learning algorithms are K-Means Clustering, Principal Component Analysis (PCA), and Gaussian Mixture Models (GMM).

Reinforcement Learning (RL)

As in RL, intelligent agents learn optimal decision-making policies through interactions with the environment, guided by scalar-valued reward signals obtained through trial-and-error exploration (El-Toukhy et al., 2023). In turn, the intrinsic dynamism of RL agents makes the approach highly suitable for solving autonomous decision-making, planning, and scheduling problems in a dynamic smart energy grid setting. The main benefit of the approach is the ability to optimise policies using feedback from the environment, without the need to provide the system with labelled data. The primary disadvantage of the technique is the need for extensive exploration and frequent retraining to adapt to changing consumption trends and attack methods, resulting in substantial computational costs for maintaining detectors across a broad consumer population (El-Toukhy et al., 2023). Representative examples of RL implementations include Markov Decision Processes (MDP), Q-Learning, and Deep Q-Networks (DQN).

Deep Learning (DL) is another highly important technique that continues to attract significant attention in smart grid analytics. In general, DL techniques are among the most relevant ML methods for smart grid applications. In contrast to classical ML algorithms, DL approaches are less sensitive to the limitations of manually engineered feature representations for high-dimensional input data

(Gu et al., 2022). DL algorithms use multi-layered neural networks that emulate brain structure and can automatically extract features from the data they are fed (Lepolesa et al., 2022). In turn, DL is particularly relevant for analysing large volumes of time-series data generated by Internet-of-Things devices embedded in smart meters. Just as in ML classification, DL models are classified by the type of learning involved: supervised, unsupervised, or reinforcement learning. These have been used in ETD for pattern identification, anomaly detection, and intelligent decision-making. Two of the most popular DL architectures used in ETD systems are CNN and RNN or LSTM.

Table 3: List and Definitions of Acronyms

Acronym	Full Meaning	Acronym	Full Meaning
1D-CONV-AE	1-Dimensional Convolutional Autoencoder	GRU	Gated Recurrent Units
ABC	Artificial Bee Colony	KNN / K-NN	K-Nearest Neighbors
ADASYN	Adaptive Synthetic Sampling	LDP	Local Differential Privacy
AMI	Advanced Metering Infrastructure	LR	Logistic Regression
ANN	Artificial Neural Network	LSTM	Long Short-Term Memory
BM	Blue Monkey (optimization algorithm)	ML	Machine Learning
CNN	Convolutional Neural Networks	MLP-AE	Multi-Layer Perceptron Autoencoder
DANN	Deep Artificial Neural Networks	NTL	Non-Technical Losses
DDQN	Double Deep Q-Networks	OCSVM	One-Class Support Vector Machine
DL	Deep Learning	PCA	Principal Component Analysis
DNN	Deep Neural Networks	PSO	Particle Swarm Optimization
DQN	Deep Q-Networks	RF	Random Forest
DT	Decision Tree	RNN	Recurrent Neural Networks
EDT	Electricity Theft Detection	SGs	Smart Grids
ETPS	Energy Theft Prevention System	SMC	Secure Multi-Party Computation
FDI / FDIA	False Data Injection (Attacks)	SMOTE	Synthetic Minority Over-sampling Technique
FFNN	Feed-forward Neural Networks	SVM	Support Vector Machine
FL	Federated Learning	TCN	Temporal Convolutional Network
FMM	Finite Mixture Model	TPR	True Positive Rate
GBDT	Gradient Boosting Decision Trees	TSRNN	Time-Series Recurrent Neural Network
GDPR	General Data Protection Regulation	VBGMM	Variational Bayesian Gaussian Mixture Models
GMM	Gaussian Mixture Models	WGAN	Wasserstein Generative Adversarial Networks
GP	Genetic Programming	XAI	Explainable Artificial Intelligence

A Comprehensive Review of ML Techniques in Machine Learning-Based Theft Detection for Smart Energy Meters

The following section presents a detailed analysis of the applications of ML algorithms for ETD, classified into three major research fields: classical machine learning, deep learning, and advanced hybrid reinforcement learning architectures. It should be noted that this classification is necessary to account for the multidimensional nature of real-world consumer consumption patterns and the wide spectrum of ETD attacks identified in the literature.

Classical Machine Learning in ETD

Classical ML algorithms have been widely researched as leading techniques for NTL detection in smart grids due to their efficiency, simplicity, and lower computational requirements. A summary of these contributions is provided in

.

Tree-Based Classifiers

The literature reveals a clear architectural evolution in ETD systems from stand-alone Decision Tree (DT) classifiers (Oprea & Bâra, 2021) to complex ensemble architectures that implement complementary learning approaches. Bagging ensemble classifiers (Random Forest (RF) and Extremely Randomised Trees (ExtraTrees)) were proven to be superior to individual classifiers in reducing variance errors with the same AUC of 0.90 for both architectures. Ensemble learning using the gradient boosting approach, namely frameworks of XGBoost, LightGBM, and CatBoost, has been studied. An accuracy of 93% and precision of 95% were reported for an architecture combining

CatBoost features with FRESH extraction (Hussain et al., 2021). Sequential XGBoost-based ETD systems were shown to be more successful in distinguishing malicious consumers from other customers than SVM and KNN classifiers. Studies like (Bohani et al., 2021) show that although shallow classifiers, such as ANNs, may provide basic ETD functionality, deep classifiers, such as DANNs, have superior discriminative ability for analysing large-scale SGCC data.

The class imbalance problem inherent in energy consumption data was addressed using advanced data preprocessing techniques (Bohani et al., 2021; Arif et al., 2021). To address skewness issues associated with TDD2022 (Zidi et al., 2023) and SGCC (Bohani et al., 2021), methods such as SMOTE, SMOTE-NM (Arif et al., 2021), and SMOTETomek (Appiah et al., 2023) were used. The SMOTETomek technique, combined with ExtraTrees classifiers, achieved 98% accuracy and an AUC of 0.9965 (Appiah et al., 2023). Meta-heuristic optimisation algorithms have increasingly been incorporated into classification processes; the Jaya algorithm applied to.

Ensemble tree-based architectures are increasingly demonstrating their potential for adaptation across a variety of smart meter and conventional environments. Although most studies focus on smart meter data at high sampling rates, some innovative frameworks designed for conventional low-sampling-rate data have achieved 99.7% accuracy using XGBoost with SMOTE (Norouzi et al., 2025). Hybrid RF-KNN models for postpaid residential consumers in conventional settings with conventional meters have been proposed, reaching AUCs of 0.89 for optimising field inspection activities for utility companies such as State Electricity Company (PLN) Indonesia (Taruna et al., 2025). Such operational versatility is enabled by the natural flexibility of the tree-based classifier model with respect to the dimensional and temporal characteristics of high-sampled smart grid data streams and low-dimensional conventional meter data (Oprea & Bâra, 2021), (Norouzi et al., 2025).

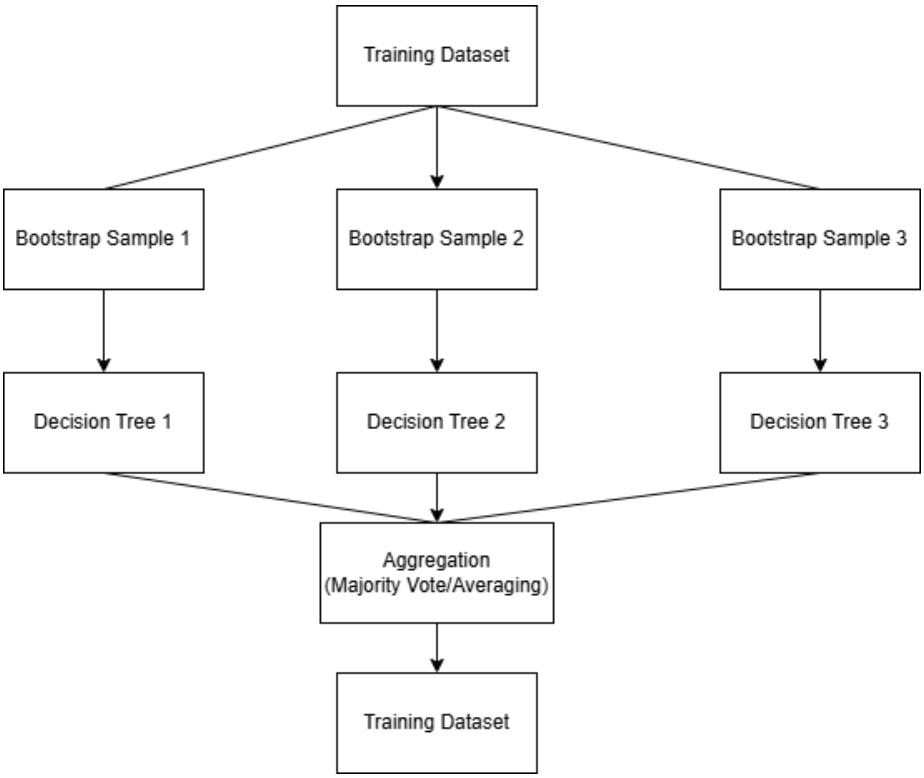

Figure 3: Schematic Diagram of Random Forest Algorithm

Support Vector Machines (SVM)

SVMs are a class of large-capacity discriminative classifiers based on the structural risk minimisation paradigm, which seeks the optimal separating hyperplane that maximises the distance between classes in a high-dimensional feature space. By applying kernel functions, most notably the Radial Basis Function (RBF) kernel, SVMs can map nonlinear consumption data into higher-dimensional feature spaces, making them linearly separable. SVMs' capability to generalise well from small training samples of support vectors makes them effective. However, scaling SVM models is a significant challenge when applying these algorithms in AMI. As the volume of smart meter data grows, the SVM training process becomes computationally intensive, as it amounts to solving a quadratic programming problem and requires storing all support vectors. Additionally, SVMs are highly sensitive to noise in IoT-based meter readings.

Applications in ETD (See summary in Table 5): Frameworks described in (Qi et al., 2022; Oprea et al., 2021), (Kong et al., 2021), and (Žarković & Dobrić, 2024) illustrate a range of SVM-related and alternative discriminative approaches for electricity theft detection. Specifically, (Qi et al., 2022) presents an unsupervised observer-meter-based scheme that provides ground-truth behaviour patterns for each consumer, reducing false alarms caused by natural low electricity consumption, without requiring any labelled theft data. The "Ratio Profile" of user vs observer data, processed with FCM clustering, yielded a 21% increase in AUC compared to regular detection patterns. In a big data analytics setting, (Oprea et al., 2021) employs a two-step approach to detecting electricity theft using the Spectral Residual-CNN for unsupervised anomalous user labelling combined with a supervised Two-Class Boosted Decision Tree classifier yielding an accuracy of 90%, precision of 0.875, and F1 score of 0.894. For NTL in LV substations, (Kong et al., 2021) proposes an extended DT-KSVM classification scheme, combined with a Wasserstein Generative Adversarial Network (WGAN) to balance class distribution, achieving a detection accuracy of 95.6% with an omission (false-negative) rate of 4.3%. Comparisons conducted in (Žarković & Dobrić, 2024) between statistical methods and various neural networks, ANNs and ANFIS, demonstrate that while autoencoders provide zero detection error, they do not classify types of theft. Conversely, ANN and ANFIS models achieve higher classification precision despite their higher error rates.

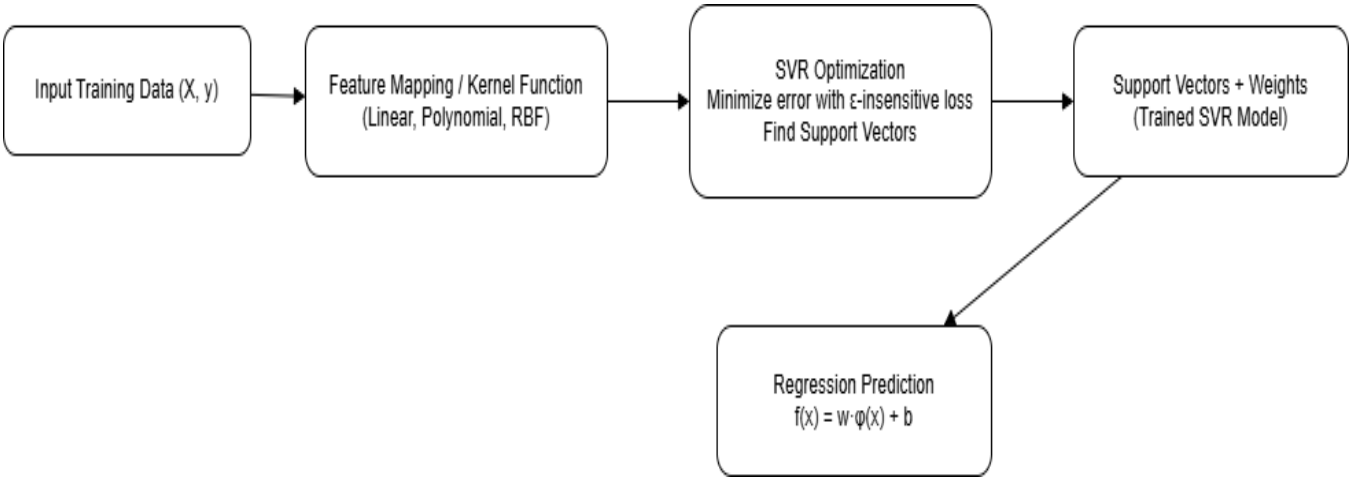

Figure 4: Architecture of Support Vector Regression

Deep Learning (DL) Architectures

As AMI moves away from traditional meters and adopts new metering systems that generate higher-frequency time-series data, deep learning becomes the required tool for processing such complex datasets.

Convolutional Neural Networks (CNN)

The relevant scientific literature indicates that CNN models are increasingly prevalent in smart grid studies due to their ability to automate feature extraction (see summary in Table 6). It was found that both 1D and 2D CNNs are suitable for smart grid analysis, where the former can detect patterns in sequential consumer behaviour data and the latter can detect periodic and seasonal structures in load profiles by converting the data into a matrix form (Ibrahim et al., 2021; Lepolesa et al., 2022). Moreover, recent research has suggested using ResNet models, which include residual connections to enable training of much deeper feature extraction pipelines, as CNNs suffer from the vanishing gradient problem in deeper networks (Arif et al., 2021). One approach that combines deep learning with ensemble classification techniques is the CNN-RF hybrid, which applies CNNs for deep feature extraction followed by classification via random forests (Cheng et al., 2021).

To overcome computational cost constraints associated with smart grid implementation, a Day-Week-Month CNN (DWMCNN) framework that uses CNNs for decentralised inference with multitemporal data processing was developed in (Cheng et al., 2021). This approach achieves an AUC score of 0.99 and an F1-score of 0.97, outperforming linear regression and gradient boosting baselines in terms of computational cost. In addition to the above approaches, meta-heuristic approaches can further improve model robustness by tuning its hyperparameters. As an example, using Particle Swarm

Nevertheless, the technical analysis presented in (Yan & Wen, 2021) shows that neural network–based detection models remain susceptible to overfitting when used alone, underscoring the need for approaches such as tree-based ensemble methods, such as XGBoost (Yan & Wen, 2021).

Recurrent Neural Networks (RNN) and LSTM

There is substantial literature on the application of RNNs to model the dynamic temporal dependence of electricity usage sequences typical of AMI systems (see Table 7 for a

summary of the presented papers). An Adaptive Time-Series Recurrent Neural Network (TSRNN), presented in (Lin et al., 2021), utilises multi-modal feature extraction of sharp, peak, and shoulder consumption characteristics to detect anomalous behaviour in consumption sequences by explicitly accounting for their stochastic dynamics. Although RNNs can be successfully used for modelling temporal dependence, the standard architecture suffers from the vanishing gradient problem when analysing long historical sequence data. To overcome such limitations, (Sebastian et al., 2024) proposes architectural enhancements to LSTM and GRU networks that leverage sophisticated gating mechanisms to retain information about fraud signatures throughout the sequence. Such an advancement is also justified in the GrAb scheme presented in (Tanwar et al., 2022), where the predictive power of LSTM-based architectures is leveraged to build a consumption baseline sensitive to behavioural anomalies, which are signs of advanced energy theft.

Moreover, the performance of recurrent models was further improved by incorporating parameter adaptation and distributed computation. The use of an adaptive weight adjustment technique in the TSRNN architecture enables automatic fine-tuning of the model using either real or synthetic data, achieving 95.1% discriminant accuracy (Lin et al., 2021). At the same time, a tendency to combine DL architectures with IoT-enabled smart systems and cloud computing environments can be observed (Sebastian et al., 2024; Si et al., 2025). Specifically, a DL model in the cloud has been created to handle massive amounts of time-series data collected by smart grids, where both computing and storage requirements for DL models become excessive, and latency-sensitive local preprocessing is assigned to the edge layer (Si et al., 2025). The empirical results indicate that although an LSTM-based architecture provides high performance for long-range dependency analysis, the trade-off between representation ability and storage/memory limitations of edge IoT devices must be considered during architecture design

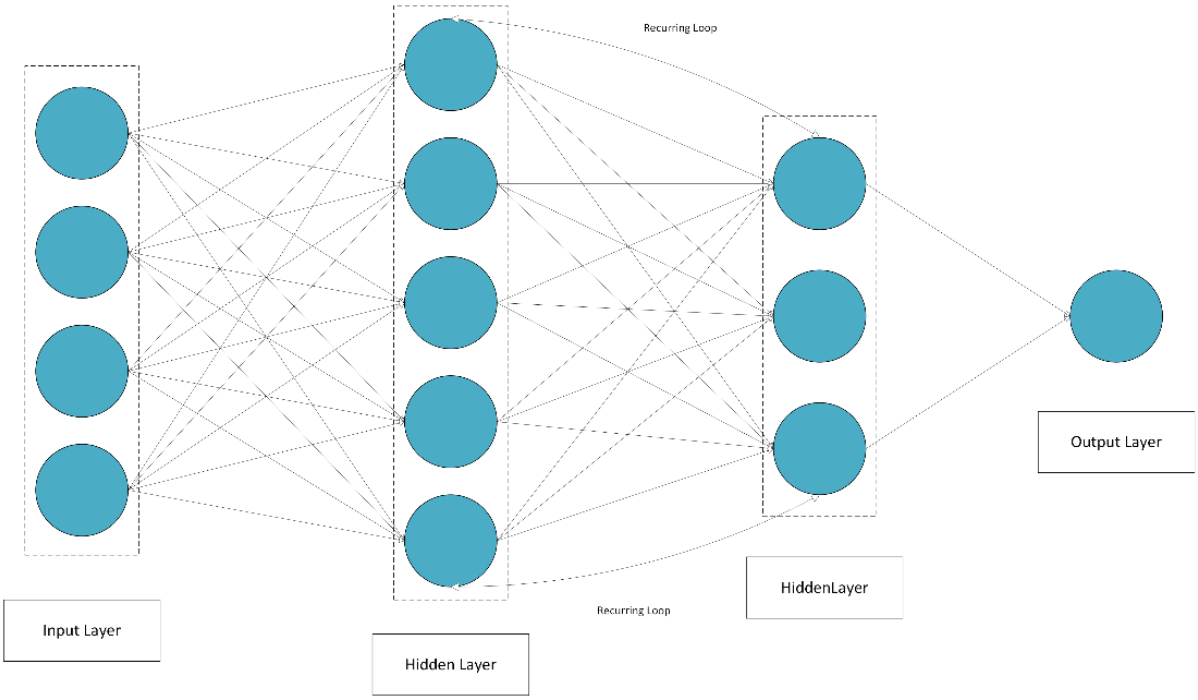

Figure 5: Recurrent Neural Network

Advanced Hybrid and Reinforcement Learning

Emerging research, on the other hand, seeks to overcome the weaknesses of single-algorithm frameworks by adopting hybridised architectures.

Hybrid Deep Learning & Evolutionary Algorithms

One trend in current literature concerns the implementation of deep learning techniques alongside evolutionary optimisation algorithms to account for the stochastic and complex nature of smart grid consumption data (see Table 8 for a summary of contributions in this domain). To mitigate the issue of local optima convergence, often resulting from manual hyperparameter tuning, the study in (Gu et al., 2022) develops a hybrid architecture combining a 1D CNN and ResNet with Particle Swarm Optimisation (PSO), achieving FPR and AUC values of 0.29% and 99.42%, respectively. Another innovative approach is directed towards the multi-objectivity aspect of detection optimisation, with opposing metrics being precision and recall.

The MOE-HDL architecture, presented in (Tursunboev et al., 2024), uses the ranking-based multi-objective evolutionary algorithm (RMOE) to simultaneously optimise model parameters and the detection window size, thereby enabling the extraction of consumption pattern cycles using a dynamically adjusted temporal analysis window.

In addition to algorithmic improvements, recent advances are also evident in the utilisation of new hardware platforms and signal processing techniques. The model in (Ullah et al., 2021) combines a CNN-GRU-PSO deep learning architecture with consumption data collected via intelligent-antenna smart meters, achieving 89% testing accuracy on the SGCC dataset and 93% accuracy, along with an AUC score of 95%, on the PRECON dataset. Deep reinforcement learning techniques add another layer of sophistication to a hybridised architecture, enabling the system to adaptively detect anomalies. A DDQN-based (El-Toukhy et al., 2023) framework coupled with the ADASYN algorithm for data balancing enables an optimal decision-making strategy to be learned autonomously through interactions with the environment. Lastly, the integration of Quantum Machine Learning (QML) into the FH-QVC-DRC framework shows considerable promise for high-dimensional false reading attack detection through quantum data re-uploading and entanglement layers in both consumption and net metering regimes (Blazakis et al., 2025).

Meta-Learning & Advanced Stacked Ensembles

Research literature indicates that more and more researchers are using stacking generalisation techniques to address basic challenges associated with individual classifiers, such as overfitting and the inability to map features appropriately (Shehzad et al., 2023). The meta-learning architecture presented in (Shehzad et al., 2023) uses a two-layer approach: a Fold-0 base layer, where different ML classifiers (Decision Trees, Random Forest, and gradient boosting) independently create their own representations of data; and a Fold-1 meta-layer, where a Multi-Layer Perceptron (MLP) learns and exploits non-linear relations between classifiers' representations to improve classification accuracy (Shehzad et al., 2023). The stacking generalisation framework is also used in (Khan et al., 2022), where the authors propose a combination of ML and DL models to enable the reliable analysis of complex theft patterns in heterogeneous smart grids. Research results in (Saqib et al., 2025) and (Shehzad et al., 2023) reveal that the stacking generalisation technique performs better than both classical voting ensembles and base models in terms of ability to exploit the non-linear structure of data in detecting theft attacks, with (Shehzad et al., 2023) reporting a ROC-AUC of 0.910 and (Saqib et al., 2025) achieving a ROC-AUC of approximately 0.94 via LR+RF stacking on ADASYN-balanced data.

Apart from conventional classification problems, recent research focuses on discovering new types of theft attacks and reducing the computational resources required to solve this problem in limited environments (Massarani et al., 2025; Chen et al., 2023). Prototype learning, implemented in (Massarani et al., 2025) using PCA and k-means clustering, has been shown to be highly effective for dimensionality reduction before ensemble-based anomaly detection. The Meta-OCSVM algorithm proposed by the authors of (Massarani et al., 2025) reduces data size by 92% and training time by 76.1%, while also improving classification accuracy for detecting novel theft types (Massarani et al., 2025). Moreover, stacked ensembles have been shown to be effective for data scarcity and class imbalance in smart grids (Saqib et al., 2025). LR+RF with ADASYN achieves 0.94 accuracy in classifying theft patterns, with stable performance across underrepresented regions. The summary articles discussed above is presented in

Adaptive & Reinforcement Learning (RL)

Significant emphasis is placed on the need for adaptive detection algorithms to combat the non-stationarity characteristic of the energy network and the evasion techniques employed by attackers. To overcome intermittent theft where perpetrators switch between fraudulent and normal consumption behaviours cyclically in order to avoid fixed threshold-based detection methods, (Fang et al., 2023) proposes a hybrid adaptive framework combining a supervised LightGBM algorithm and unsupervised Variational Bayesian Gaussian Mixture Models (VBGMM) for identifying theft in unlabeled consumption data with performance gains of 5.13% to 9.54%. In non-stationary grid systems, DRL-based agents, including Deep Q-Network (DQN) and Double Deep Q-Network (DDQN), are used to optimise adaptive rewards for avoiding costly offline model training. This adaptive framework, combined with the hybrid CNN-GRU-DDQN system, demonstrates high robustness against non-deterministic and stochastic adversarial behaviour in smart grids, achieving a benchmark accuracy of 97.33% (El-Toukhy et al., 2023).

In the empirical study of smart homes, the multidimensional and device-specific nature of smart grid environment power usage patterns indicates the requirement for high-dimensional, adaptive ETD detection frameworks. In (Abraham et al., 2024), the evaluation compares supervised ML models (e.g., XGBoost) with unsupervised DL algorithms (e.g., MLP-AE, 1D-CONV-AE). While unsupervised DL methods enable high-dimensional adaptivity against unexpected fraud, supervised models such as XGBoost demonstrate accuracy rates up to 95.83% in smart home deployment contexts.

Federated Learning & Forecasting

However, the literature highlights the need for a shift in paradigm toward privacy-aware, predictive architectures to protect smart grids from cyber-physical attacks that manipulate data. To counter the security challenges posed by the centralised smart grid, such as data exfiltration risks, and to reduce communication overhead, (Wen et al., 2022) proposes a new architecture – FedDetect – based on federated learning that leverages the Federated Temporal Convolutional Network (TCN). In this case, the consumers' data will be decentralised using the Local Differential Privacy and Lifted EC-ElGamal encrypted-aggregation approach, providing 92.5% classification accuracy while minimising computational overhead to 142ms. In addition to federated learning, (Dash et al., 2021) demonstrates that a forecasting-based detection model is also effective for predicting cyber-attacks, including T&D theft, utilising established household demand baselines. Using a demand forecasting algorithm based on RNN-GBRT, which outperforms other models such as ARIMA and SVR, T&D theft is statistically predicted; any deviation above the set threshold of 7.5kW is considered suspicious.

The implementation of ETD in IoT-based smart city infrastructure has led to the development of multidimensional ETPS to ensure smart grid stability in urban energy networks. The proposed multi-layer approach described in (Quasim et al., 2023), which includes multi-model prediction algorithms based on GRU and LSTM neural networks, operational monitoring of smart cities in real-time, as well as meta-heuristic optimisation using Grey Wolf Optimisation (GWO), with statistical validation performed via the Simple Moving Average (SMA) method, allows delivery ratio maximisation, as well as energy consumption reduction compared to other non-integrated approaches to ETD. Moreover, the reliability of multidimensional theft pattern detection is increased by using a multi-layer deep learning architecture. The two-stage model, incorporating the K-means algorithm for customer segmentation followed by a fused CNN/DNN classification architecture trained on authentic user data and artificially simulated attacks, achieved a 98% detection rate (DR) and 98.1% AUC (Emadaleslami et al., 2023).

Table 4: Summary of Contributions to Tree-Based Classifiers in ETD

Ref	Purpose (Scenario)	Main Objective	Machine learning algorithm	Solution Approach	Results Obtained	Metric	Limitation
Bohani et al., (2021)	Compare supervised classifier performance on electricity theft detection.	Evaluate four classifiers on SGCC data without data balancing.	DT, ANN, DANN, AdaBoost	Data imputation, MinMax normalization, variable train-test split ratios (90/10–60/40)	DANN: superior Recall, F1, AUC; ANN: highest mean accuracy (92.54%)	Accuracy, Precision, Recall, F1-Score, AUC	No imbalance handling applied; single dataset (SGCC)
Hussain et al., (2022)	Address NTL in smart grids via sequential ML framework.	Improve detection accuracy and reduce false positives.	KTBoost (XGBoost + Kernel functions)	Data imputation, Robust-SMOTE, Jaya optimization, classification	Accuracy: 93.38%, Precision: 95%	Accuracy, Precision, Recall	High computational complexity; validated on SGCC only
Gunturi & Sarkar, (2021)	Identify theft using fine-grained smart meter data.	Compare Bagging vs. Boosting ensembles; address class imbalance.	RF, ExtraTrees (Bagging); AdaBoost, XGB, LGB (Boosting)	Preprocessing, class imbalance handling, theft simulation, ensemble classification	Bagging (RF/ET) superior on large datasets; AUC 0.90	AUC-ROC, Precision, Recall, F1, Accuracy, CM	Boosting performance degraded with increasing users; reliance on synthetic theft data
Oprea & Bâra, (2021)	Identify fraud in regions with conventional monthly billing data.	Detect anomalies in high-frequency smart meter data.	LightGBM	Feature engineering, multivariate Gaussian anomaly detection, imbalance handling, hyperparameter tuning	ROC-AUC: 0.8606; feature engineering improved performance by 28%	ROC-AUC, Accuracy, Precision/Recall	Weak raw feature correlation; dataset contained duplicate/ conflicting meter IDs
Arif et al., (2021)	Handle data imbalance and high dimensionality in ETD.	Build accurate model with SMOTE-NM balancing and Bayesian tuning.	ResNet (feature extraction), RF (classification), Bayesian Optimizer	Interpolation, SMOTE-NM hybrid balancing, ResNet, optimized tree classifiers	RF: AUC 0.996; detected 90% of thieves and 100% of normal users	AUC, Accuracy, Macro-F1, CM	Inflexible balancing (fixed samples); computationally expensive preprocessing
Yan & Wen, (2021)	Identify power theft in AMI; minimize financial losses.	Detect abnormal consumption patterns with high accuracy.	XGBoost	Preprocessing, 6-type malicious sample generation, XGBoost classification	Outperformed 10 models; AUC 99.62%, Precision 97.53%, FPR 3.17%	Precision, Recall, FPR, AUC	Excludes climate/ regional variables; minimal real-world labeled data
Hussain et al., (2021)	Address NTL with feature-engineered detection architecture.	Enhance detection accuracy and model interpretability via CatBoost.	CatBoost, k-NN (imputation), SMOTETomek, FRESH (feature extraction), Tree-SHAP	k-NN imputation, SMOTETomek balancing, FRESH feature extraction, CatBoost training, Tree-SHAP interpretation	Accuracy 93%, DR 92%, Precision 95%; outperformed XGBoost and LightGBM	Accuracy, Recall, Precision, F1, Kappa, MCC	Feature engineering computationally expensive (600s) vs. model training (280s)

Ref	Purpose (Scenario)	Main Objective	Machine learning algorithm	Solution Approach	Results Obtained	Metric	Limitation
Appiah et al., (2023)	ETD in imbalanced smart meter data.	Develop robust tree-based ensemble with optimized hyperparameters.	ExtraTrees	Preprocessing, SMOTETomek balancing, GridSearchCV tuning, ET classification	Acc: 98%, AUC: 99.65%, Precision: 97%, F1: 98%, MCC: 95.06%	Accuracy, AUC, Precision, F1, DR, MCC	Requires labeled data; computationally intensive; limited feature engineering
Abbas et al., (2024)	ETD with reduced labeling cost via active learning.	Develop efficient detection system minimizing manual annotation.	RFAL, XGBoostAL, DTAL, GBAL, KNNAL, CBAL, LGBMAL	Pool-based AL with uncertainty sampling; 3 iterations; OEDI dataset	RFAL accuracy: 70.61%	Accuracy, F1, Precision, Recall, ROC-AUC	Lower accuracy vs. deep learning; manual labeling still required
Kawoosa et al., (2023)	Improve detection accuracy and reduce FPR using balanced synthetic data.	Detect malicious consumers using enhanced XGBoost framework.	XGBoost	6-type synthetic theft generation, PCA, weather/seasonal features, grid search tuning	DR: 96%, FPR: 3%	DR, FPR, Precision, Recall, F1, AUC	SGCC-only; synthetic attacks may not reflect real theft complexity
Zidi et al., (2023)	Benchmarking and ETD in smart grids.	Create and evaluate a public ETD dataset; compare ML classifiers.	KNN, DT, RF, Bagging, ANN	Multi-class TDD2022 dataset generation; four validation protocols	RF accuracy: 94.71% (6-class, known consumer)	Accuracy, F-measure, Kappa, AUC	Theft class 6 difficult to detect; limited to synthetic theft patterns
Norouzi et al., (2025)	ETD in regions with traditional metering and low-frequency data.	Develop accurate detection for imbalanced, asynchronous consumption data.	XGBoost + SMOTE	Monthly average conversion with seasonal adjustment, SMOTE balancing, XGBoost classification	Accuracy: 99.7%, Precision: 98.2%, Recall: 90.5%, F1: 94.2%	Accuracy, Precision, Recall, F1	Region-specific; may require retraining for evolving theft patterns
Taruna et al., (2025)	ETD for PLN postpaid customers using traditional meters in Indonesia.	Optimize field inspection targeting and subsidy allocation accuracy.	DT, RF, K-NN, LR, NB, DNN	Sequential evaluation by theft frequency; hybrid RF-KNN final detection	RF+KNN: Acc 0.89, Prec 0.83, Rec 0.98, F1 0.90, AUC 0.89	Accuracy, Precision, Recall, F1, AUC	Limited to 450 VA segment; requires up to 27 months of historical data

Table 5: Summary of Contributions to Support Vector Machines in ETD

Ref	Purpose (Scenario)	Main Objective	Machine learning algorithm	Solution Approach	Results obtained	Metric	Limitation
(Qi et al., 2022)*	Identify theft in AMI; minimize false alarms from natural low consumption.	Normalize user data using observer meters to improve unsupervised detection.	FCM Clustering	Ratio Profile (User/ Observer) computation, wavelet feature extraction, FCM clustering	AUC improved by 21%; outperformed all baselines on business and residential data	AUC, MAP@20	Cannot identify intra-day tampering; uses daily summation profiles only
(Oprea et al., 2021)	Identify NTL and theft in smart grids with big data.	Develop hybrid unsupervised-supervised anomaly detection and fraud forecasting.	SR-CNN (unsupervised), Two-Class Boosted DT (supervised), Fisher LDA	SR-CNN labels anomalous users (>15% anomalies = suspicious); Boosted DT classification	Accuracy 90%, Precision 0.875, F1 0.894, AUC 0.971	Accuracy, Precision, Recall, F1, AUC	Tested on historical data only; raw data contained timestamp errors; Martingale model lacked confidence scores
(Kong et al., 2021)	Reduce NTL and economic losses in low-voltage stations.	Propose hybrid detection with similarity measures and multi-stage classification.	DT-KSVM, WGAN (data balancing), KFCM (clustering)	Feature extraction, KFCM load profiling, Comprehensive Similarity computation, WGAN balancing, DT-KSVM classification	DR: 95.6%, FPR: 4.3%; outperformed SVM, RF, BP	DR, FPR, Accuracy	Offline historical data analysis only; no real-time detection performance evaluation
(Žarković & Dobrić, 2024)	ETD in low-voltage distribution networks.	Compare statistical and AI-based NTL detection methods across theft typologies.	ANN, ANFIS, Autoencoder, K-means, Statistical rules	5 theft scenarios defined; 6 statistical indicators extracted; trained on 15-min data from 91 consumers	ANN: 7.62% frequency error; Autoencoder: 0% detection error (no type classification); K-means: 9.26% error	Frequency error, MSE, classification accuracy	Small dataset (91 consumers); autoencoder cannot classify anomaly type; some anomalies undetectable by statistical method

Table 6: Summary of CNN solutions to Deep Learning in ETD

Ref	Purpose (Scenario)	Main Objective	Machine learning algorithm	Solution Approach	Results obtained	Metric	Limitation
Gu et al., (2022)	Address high False Positive Rate (FPR) in practical ETD deployments.	Specifically minimize FPR while preserving high detection accuracy.	1D-CNN + ResNet + PSO	Two-stage: (1) feature learning, (2) PSO-based FPR optimization	Lowest FPR (0.29%) among benchmark models; AUC 99.42%	FPR, TPR, AUC, BDR	Performance degrades on small datasets (<1,500 users) or very low sampling rates
Ullah et al., (2022)	Identify theft in high-dimensional AMI with class imbalance.	Develop hybrid DL model for feature extraction and parameter classification.	AlexNet (CNN) + AdaBoost + ABC	Preprocessing, Near-Miss balancing, AlexNet feature extraction, AdaBoost classification, ABC optimization	Accuracy: 88%, AUC: 0.91; superior to standalone CNN/SVM	Accuracy, Precision, Recall, F1, MCC, AUC	Computationally expensive (ABC); limited to consumption data (no location/demographic features)
Lepolesa et al., (2022)	Mitigate NTL causing global annual losses.	Enhance ETD performance by integrating time-frequency domain features in DNN.	DNN (Feed-forward), PCA, mRMR feature selection, Bayesian Optimizer	PCHIP interpolation, synthetic data balancing, time/frequency feature extraction, mRMR/PCA selection, optimized DNN training	AUC 97% (1% above benchmarks), Accuracy 91.8%; frequency features more critical than time features	AUC-ROC, Accuracy, Precision, Recall, F1, MCC	Historical data analysis only; no real-time detection; validated on SGCC dataset only
Ibrahim et al., (2021)	Address computational challenges of traditional algorithms in large-scale SG data.	Establish optimal sequential model configuration and feature dimensionality reduction.	CNN + Blue Monkey (BM) optimization	Synthetic data generation, sequential model configuration testing (2–4 layers), BM feature reduction, classification	Optimal config (2 layers: 128/64 nodes): Accuracy 92%; BM reduced features from 1,035 to 666; training time ~10s	Accuracy, Loss, Training Time	Relies on synthetic data; accuracy (92%) below some competing DL models
Cheng et al., (2021)	Resolve bandwidth and latency constraints of centralized system	Develop a precise, edge-deployable detection algorithm	DWMCNN (Day-Week-Month CNN) + RF + K-means	Centralized clustering (K-means) and DWMCNN training; edge-level feature extraction and RF classification	AUC 0.99, Accuracy/F1 0.97; outperformed CNN-SVM, GBDT, LR	Accuracy, Precision, Recall, F1, AUC	Edge nodes lack full model training computational capacity; intensive preprocessing required

Table 7: Summary of RNN and LSTM Applications in ETD

Ref	Purpose (Scenario)	Main Objective	Machine learning algorithm	Solution Approach	Results obtained	Metric	Limitation
Tanwar et al., (2022)	Identify NTL (energy theft) and technical losses in smart grids.	Introduce GrAb algorithm for energy usage prediction and loss categorization.	LSTM (prediction), SVM (classification), CTGAN (synthetic data generation)	Linear interpolation, LSTM load prediction + threshold calculation, SVM classification	LSTM: Avg RMSE 0.0137, MAPE 7.42% (vs. RMSE 0.47 baseline); successful synthetic theft detection	RMSE, MAPE	Dataset (OpenEI) lacked real theft data; model validated on residential data only
Lin et al., (2021)	ETD in time-series consumption data with severe class imbalance.	Develop adaptive TSRNN combined with SMOTE for accurate anomaly identification.	TSRNN, SMOTE, K-means	ARP/PEA/SHO feature selection, SMOTE balancing, adaptive TSRNN training, K-means classification	Training Accuracy 95.1%; Testing Accuracy 89.3%; 44/47 real abnormal users correctly identified (93.6%)	Accuracy, TPR, FAR	SMOTE neighbor selection uncertainty; off-peak (OPE) data eliminated as non-discriminatory
Sebastian et al., (2024)	ETD in smart grids using IoT smart meter data.	Detect electricity theft via comparative DL architectures on IoT data.	CNN, LSTM, SVM, RF, ARIMA, MLP, GRU, RNN	Cluster-wise detection, historical data classification, forecasting-based anomaly detection with statistical post-processing (MAPE, SMA, APE)	CNN: 91% accuracy; LSTM: 97% accuracy	Accuracy, Precision, Recall, F1, RMSE, MSE, MAPE	Limited real-time deployment validation; high computational requirements; assumes clean historical data
Si et al., (2025)	ETD in smart grids using cloud computing infrastructure.	Improve detection by capturing reliable long-term dependencies from complex consumption data.	ETD-SAC, WDCNN, CNN-LSTM, LFPR-DNN, HORLN	Cloud framework with ETD-SAC series decomposition, FFT-based auto-correlation, subsequence-level depende-ncy aggregation	ACC up to 0.97, FNR as low as 0.04, FPR as low as 0.04	ACC, FNR, FPR	Limited to 321 users; simulated attacks only; overfitting risk; cloud data privacy concerns

Table 7: Summary of Advanced Hybrid and Reinforcement Learning Applications in ETD

Ref	Purpose (Scenario)	Main Objective	Machine learning algorithm	Solution Approach	Results obtained	Metric	Limitation
Ullah et al., (2021)	Identify NTL via smart meter intelligent antenna integration.	Achieve accurate theft detection; address overfitting and class imbalance.	Hybrid CNN-GRU-PSO	Preprocessing, SMOTE balancing, CNN feature extraction, PSO-optimized GRU classification	Accuracy 89%, AUC 89%; PRECON Dataset: Accuracy 93%, AUC 95%	Accuracy, AUC, F1, Precision, Recall	Sequential computational complexity; reliance on synthetic/ simulated theft data
Javaid et al., (2023)	ETD in smart grids with automated hyperparameter optimization.	Improve detection using hybrid DL with automated tuning.	SSA-GCAE-CSLSTM (hybrid DL)	Cost-sensitive balancing, GCAE feature extraction, SSA hyperparameter tuning, LSTM classification	Accuracy: 92.25%, Precision: 99.45%, F1: 95.93%	Accuracy, Precision, Recall, F1, AUC	Computationally intensive; SGCC-only dataset; complex for real-time deployment
Iftikhar et al., (2024)	ETD in smart grids using real consumption data.	Develop accurate hybrid DL model for imbalanced smart meter data.	Hybrid MLP-GRU, AlexNet, GRU, BGRU, RNN	Imputation and normalization preprocessing; k-means SMOTE balanc-ing; hybrid GRU + MLP classific-ation; three train-test splits (50-50, 75-25, 90-10)	MLP-GRU: Accuracy 93.33% (90-10), Recall 95%, Precision 94%, AUC 95%	Accuracy, Precision, Recall, F1, AUC, ROC, MCC	Requires high-frequency data; no automated hyperparameter tuning; computationally intensive for large datasets
Mbey et al., (2024)	Detect electricity theft via consumption pattern analysis.	Identify fraudulent consumers using optimized SVM-PSO hybrid.	SVM + PSO	Hybrid SVM-PSO with real dataset, calendar context, co-simulation, sensitivity analysis	Precision 98.9%, AUC 0.989	Precision, Recall, F1, AUC	Smart meter data nonlinearity may affect performance; requires validation on larger, diverse datasets
Tursunboev et al., (2024)	ETD with dynamic window optimization.	Improve detection accuracy; reduce FPR; handle data imbalance; optimize detection window.	Hybrid DL (LSTM + 1D-CNN + 2D-CNN) + RMOE evolutionary optimization	Multi-objective evolutionary algorithm jointly optimizes model parameters and detection window using precision and recall as conflicting objectives	SGCC: AUC 0.823, MAP@100 0.965; Irish: AUC 0.932; London: AUC 0.880	AUC, MAP@K	High computational complexity; significant training resource requirements
Blazakis et al., (2025)	Detect power theft across consumption and net metering domains using QML.	Develop quantum-classical hybrid detector for false reading attacks.	FH-QVC-DRC (hybrid quantum-classical), LSTM, XGBoost, LightGBM, CatBoost	Angle embedding, entangling layers, data reuploading between classical layers; three-stage feature integration for net metering	Consumption: ACC 0.87; Net metering: ACC 0.977, AUC 0.983; outperforms classical methods	Accuracy, Precision, Recall, F1, AUC/ROC	Limited to 16 qubits; simulation only (no quantum hardware); slower training than classical; scalability unvalidated

Table 9: Summary of Advanced Stacked Ensembles Applications in ETD

Ref	Purpose (Scenario)	Main Objective	Machine learning algorithm	Solution Approach	Results obtained	Metric	Limitation
Khan et al., (2022)	Address big data nonlinearity challenges in smart grids.	Develop a robust stacked generalization model combining ML and DL architectures.	Stacked: SVM, RF, GBDT (base) + TCN (meta)	STLU balancing, base classifier predictions, TCN meta-classification	AUC: 98.5%, FPR: 1.03%; TCN meta-classifier accuracy: 94.6%	Accuracy, Precision, Recall, F1, AUC, CM	Not evaluated against synthetic attacks; lacks grid topology information integration
Chen et al., (2023)	ETD in smart grids with feature-engineered deep learning.	Develop accurate detection model using load/quantity feature indicators.	FLB-DCNNs (1D DenseNet + Focal Loss)	Feature engineering (load/quantity indicators), FLB-DCNNs training, statistical validation (Friedman/Holm tests)	Precision: 98.51%, Recall: 98.17%, F1: 98.16%	Precision, Recall, F1	Feature engineering-dependent; constrained by missing data (line loss, alarms)
Shehzad et al., (2023)	ETD in smart grids using deep learning-based meta-learner.	Detect NTL using two-stage stacked ensemble with MLP meta-learner.	SVM, LR, DT, RF, CatBoost, XGB, MLP	Two-stage stacked ensemble: Fold-0 (base classifiers), Fold-1 (MLP meta-learner) + AST hybrid sampling	ROC-AUC: 0.914 (imbalanced), 0.910 (balanced); PR-AUC: 0.987	ROC-AUC, PR-AUC, Accuracy, Precision, Recall, F1, FPR	Manual hyperparameter tuning; high DL computational cost; single geographic region
Massarani et al., (2025)	Zero-day ETD from high-frequency smart meter sensor data.	Detect unseen theft efficiently with low computational overhead.	OCSVM, GMM, Meta-OCSVM + PCA + K-means	PCA + K-means prototype learning, OCSVM/GMM training on prototypes, meta-OCSVM score fusion	Accuracy: 88.45%, FAR: 13.85%, data reduction: 92%, training time reduction: 76.1%	Accuracy, FAR, HD, AUC	Synthetic attack data only; parameter tuning sensitive; no real-world validation
Saqib et al., (2025)	ETD in regions without smart grid infrastructure.	Develop lightweight ensemble model for imbalanced data in non-AMI regions.	LR, RF, Voting, Stacking, ADASYN	PCA + ADASYN balancing + voting/stacking ensembles on Kaggle dataset (42,373 customers, 1,034 days)	ACC 0.94, Precision 0.95, Recall 0.94, F1 0.94, ROC-AUC 0.94, MAE 0.056	ACC, Precision, Recall, F1, ROC-AUC, MAE, MCC, Kappa	Single dataset; synthetic data may introduce noise; no real-world validation; monthly data limits real-time capability

Table 10: Summary of Reinforcement Learning Applications in ETD

Ref	Purpose (Scenario)	Main Objective	Machine learning algorithm	Solution Approach	Results obtained	Metric	Limitation
Fang et al., (2023)	Identify intermittent thieves alternating between fraudulent and legitimate consumption.	Detect intermittent attacks without theft labels.	LightGBM (supervised) + VBGMM (unsupervised clustering)	LightGBM probability scoring, Disperse Degree calculation, VBGMM clustering	AUC increased by 5.13%–9.54%; effective on low theft-rate days	AUC, F1, DR, FPR	Cannot account for users switching theft status within a single day
El-Toukhy et al., (2023)	Adaptive ETD in smart power grids against false reading attacks.	Detect attacks using RL; adapt to new consumption patterns and zero-day attacks.	DQN, DDQN, FFNN, CNN, GRU, CNN+GRU	Four scenarios: global detection (DQN/DDQN), customized models for new customers, adaptation to changed patterns, learning new cyber-attacks	Best: DDQN+CNN+GRU: ACC 97.33%, Precision 97.38%, Recall 97.33%, FA 2.06%, HD 95.27%	ACC, Precision, Recall, FA, FNR, HD, F1	High RL training cost; restricted to Irish smart meter data; privacy and secure communic-ation not addressed
Abraham et al., (2024)	ETD in smart homes using appliance-level power consumption.	Detect electricity theft using appliance-level patterns with ML/DL methods.	XGBoost, RF, MLP (supervised) MLP-AE, 1D-CONV-AE, Isolation Forest (unsupervised)	Synthetic attack data generation, data augmentation, UMAP clustering, binary classification, anomaly detection	XGBoost: AUC 98.69% (sim), 98.74% (real); Accuracy 95.54% (sim), 95.83% (real)	AUC, Accuracy, Precision, Recall, F1	Data quality dependent; assumes stable consumption patterns; privacy concerns; computational cost

Table 11: Summary of Federated Learning Applications in ETD

Ref	Purpose (Scenario)	Main Objective	Machine learning algorithm	Solution Approach	Results obtained	Metric	Limitation
Dash et al., (2021)	Predict renewable energy demand and detect theft in developing countries (India).	Dual-objective: long-term forecasting and statistical theft detection.	RNN-GBRT (hybrid) + Statistical Thresholding	RNN-GBRT forecasting; mean/standard deviation thresholding for theft detection	RNN-GBRT outperformed ARIMA/SVR; detected theft events >7.5 kW deviation	MSE, MAE, RMSE	Targets individual households; lacks smart area/network topology context
Emadaleslami et al., (2023)	ETD in smart grids with synthetic data generation and multi-source data fusion.	Detect fraudulent consumers using two-stage DL model.	CNN + DNN + K-means	User clustering, CNN theft pattern modeling, synthetic data generation, DNN classification, real-time deployment	DR: 98%, FPR: 3.76%, AUC: 98.1%	DR, FPR, AUC	Requires historical theft data; computationally intensive; lacks transformer-level area detection data
Quasim et al., (2023)	IoT-based smart city energy theft prevention.	Develop multi-objective ML-based theft prevention with real-time monitoring.	GRU + GWO + DDRCNN + LSTM + SMA	Multi-level: data collection, multi-model prediction, SMA-based decision, continuous monitoring	Higher delivery ratio, throughput, network lifetime; lower delay, energy, overhead vs. non-ETPS	Delivery ratio, throughput, delay, energy conservation, network lifetime, overhead	Requires structured historical data; simulation-based validation only
Wen et al., (2022)	Prevent theft while preserving consumer privacy in smart grids.	Enable collaborative model training without raw data sharing via federated framework.	Federated TCN	LDP, local training, Lifted EC-ElGamal encrypted aggregation, global model update	Accuracy 92.5%, AUC 0.7915; computation overhead 142ms (lower than benchmarks)	Accuracy, AUC, Computation Time	Slower convergence than centralized models; LDP noise slightly reduces accuracy with few participating stations

Comparative Synthesis

Each of the above methods was assessed based on its own parameters. In the application of the extraction framework set forth in the earlier section, when applied to all 44 studies presented in the table, there emerges a broader set of patterns that are not obvious at first glance and that relate directly to the deployment-readiness of the methods.

First, headline performance has converged while comparability has not improved. In the 44 studies analyzed, AUC is shown in 30. From 2023, however, AUC values are often high towards the upper bound of the metric's range in all four categories. This trend is largely caused by the nature of the benchmarking process where the AUC is evaluated using a small pool of data sets. Indeed, both SGCC data set and the stolen consumption profiles artificially added to it occur repeatedly in the collection of papers being discussed here, while since attack injection processes are different in every study although the consumption data may be the same, the same AUC value does not guarantee the comparability of two studies. Another problem with AUC is in the inconsistent naming of the data set utilized, Table 1 suggests.

Second, metric selection systematically favours optimistic reporting. Accuracy and AUC dominate the corpus, yet, in the case of extreme class imbalance found in the context of electricity theft, both of these measures may stay relatively high despite the rise in the false positive rate being too high to be practically implemented. Notably, only 9 out of 44 studies mentioned in the table provide information on the false positive rate. It matters because the cost of deploying an ETD system consists of the cost of conducting field inspections of the customers flagged as fraudulent by the system. In the case of Nigeria and similar settings, the cost of making a false accusation in the eyes of reputation cannot be measured by any accuracy measure.

Third, robustness and deployability stand in tension with reported accuracy. It follows from the discussion that the most effective approaches according to the benchmark metrics, namely hybrid deep and evolutionary optimized methods, require the highest costs of training and possess the lowest level of testing for adversarial or distribution-shift evaluation. On the contrary, federated and adaptive methods provide somewhat lower headline metrics but take into account all practical limitations, such as data localization, privacy, partitioning among utilities, and generalization on new attacks. As a result, two separate branches of research have emerged that either show the ability to provide high accuracy on a specific benchmark or prove robustness in changing conditions.

Overall, all of these facts lead to one single conclusion about the current state of the art: none of the considered methods combines top accuracy, efficiency in metrics used for reporting results, and validation beyond the chosen family of benchmarks. Table 12 presents the comparison of the approaches according to the discussed characteristics.

Table 12: Comparative synthesis of the reviewed corpus and implications for deployment

Dimension of comparison	Observation across the reviewed corpus	Implication for deployment
Reported performance	Of the 44 tabulated studies that report AUC, 30 and their associated clusters have reached upper-bound performance levels in publications since 2023.	Headline accuracy is mostly saturated, and further improvements upon existing benchmarks do not provide new discriminative information.
Benchmark diversity	SGCC dataset and synthetically generated theft profiles appear in all method categories, and the naming scheme of the dataset is not consistent across the reviewed publications.	Comparisons between papers are weak; equivalent numeric scores between papers do not mean equivalency.
Cost-sensitive metrics	False positive rates are reported only in 9 out of 44 tabulated studies.	The cost of field inspection, which makes up the bulk of deployment costs, is ignored by most papers.
Class-imbalance handling	Imbalance is primarily handled by means of re-sampling and generation rather than by obtaining verified theft labels.	Training imbalance is dealt with, but realism of the attacks modeled is not proven.
Robustness and adversarial evaluation	Evaluations of adversarial or distribution-shift robustness are not common and are performed mostly by adaptive, reinforcement, and federated learning studies.	Adversarial robustness is mostly claimed rather than proven.
Privacy and distribution	Federated designs report more modest headline accuracy but directly address data locality and multi-utility partitioning.	These designs are better matched to real multi-utility deployment than centralised high-accuracy models.

Dimension of comparison	Observation across the reviewed corpus	Implication for deployment
Deployment readiness	None of the reviewed approaches achieve top performance, cost sensitivity, and validation across multiple benchmark families	Cost-sensitive benchmarking is the main remaining need of the field.

Lessons Learned

Selecting Suitable ML Algorithm: The literature highlights a fundamental dichotomy in the domain of ML-based ETD, stemming from the distinct operational properties of supervised, unsupervised, and semi-supervised learning paradigms. Supervised models exhibit superior detection accuracy for known theft pattern typologies but suffer from marked performance degradation when faced with novel tampering techniques not observed during training, as static learned boundaries fail to generalise to emerging attack vectors (Ibrahim et al., 2021), (Massarani et al., 2025). On the other hand, the unsupervised models of clustering and autoencoders have an inherent ability to detect previously unknown theft patterns through their lack of reliance on existing labels, but are simultaneously burdened with high false positive rates as these models lack the ability to distinguish between genuine theft patterns and legitimate changes in consumer behaviour (Massarani et al., 2025). Semi-supervised models aim to leverage the strengths of supervised models in terms of high accuracy with the generalisation capabilities of unsupervised approaches through simultaneous exploitation of both labelled and unlabeled data; however, issues of classification still persist with respect to the accurate distinction of theft-induced anomalies from normal behavioural patterns. It is clear from the literature that the ideal approach combines shallow feature extractors, such as tree-boosting models valued for their superior parallelizability and low computational overhead, with deep neural network architectures for representing complex, high-dimensional features (Hussain et al., 2021; Kawoosa et al., 2023; Sebastian et al., 2024).

Analytical Framework for Classical ML Problem Formulation: It is imperative to formulate the classical ML problem with a well-defined set of feature engineering processes, as the univariate nature of time-series data collected from smart meters lacks the discriminatory power to distinguish theft signatures from grid noise. The literature consistently advocates the adoption of sophisticated data imputation techniques, such as K-Nearest Neighbours and linear interpolation, for filling missing values in IoT sensor data, as simplistic approaches like zero vector imputation or deletion introduce systematic bias in model training, leading to biased prediction results (Hussain et al., 2021; Yan & Wen, 2021; Kawoosa et al., 2023). There has been extensive emphasis on feature engineering techniques for identifying periodicity and seasonality in consumption patterns. Specifically, transforming one-dimensional.

Performance of Advanced Hybrid and Reinforcement Learning: The rapidly evolving nature of contemporary power grids has driven the development of advanced hybrid composite architectures that deliver superior performance across various applications. In particular, deep learning algorithms have achieved high efficacy in ETD tasks by combining them with Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) or Gated Recurrent Units (GRU) networks (Shehzad et al., 2023; Ullah et al., 2021). In these hybrid algorithms, CNN layers automate spatial feature extraction, whereas LSTM/GRU layers identify temporal sequential patterns in consumption data, thereby capitalising on the complementary representational power of these two architectural paradigms. The literature further discusses the generalisation advantages of meta-learning and stacking techniques, which systematically minimise the classifier variance and bias (Gunturi & Sarkar, 2021; Shehzad et al., 2023). Stacking is a technique in which the predictions of multiple base learners are used to train an additional, higher-order meta-classifier that achieves effective generalisation across disparate data distributions and minimises overall generalisation error. The static learning paradigm of conventional algorithms has been increasingly supplanted by dynamic reinforcement learning models, which iteratively refine detection policies through exploration and exploitation of the environment without the need for expensive offline retraining (El-Toukhy et al., 2023). The strict data privacy regulations surrounding smart grid deployments have further expedited the adoption of federated learning, which enables gradient estimation from decentralised

Training Convergence: The speed of convergence of ETD models depends highly on the effectiveness of hyperparameter tuning and the statistical properties of the training dataset. Smart meter data exhibits class imbalance in real-world deployments, with theft incidents accounting for a negligible fraction of the dataset. The resulting distributional skew leads to systematic bias during optimisation, causing gradient descent methods to converge to the minimum of the majority class, thereby hampering convergence of the decision boundary for the minority class (Saqib et al., 2025; Hussain et al., 2022). While the Synthetic Minority Oversampling Technique (SMOTE) is the most frequently employed method for minority class augmentation, improper use may impair convergence by generating synthetic data that does not adequately represent the underlying manifold of theft data. To address this problem, cost-sensitive learning has

emerged as a preferred technique, in which misclassification penalties are applied to the minority class to force convergence to the decision boundaries without resorting to oversampling (Gu et al., 2022; Javaid et al., 2023; Chen et al., 2023). To address the problem of local optima, which markedly reduces convergence rates, bio-inspired meta-heuristic evolutionary algorithms, such as Particle Swarm Optimisation (PSO) and the Artificial Bee Colony (ABC) algorithm, have yielded significantly faster gradient descent convergence in ETD applications.

Training Failure: A consistent trend in the literature is that ETD models achieve high accuracy during training but experience severe performance degradation during evaluation, mainly due to overfitting and improper data processing. The primary cause of training failure in class-imbalanced datasets is the overfitting to artificially generated theft patterns. This phenomenon is characterised by the overspecification of the feature space through synthetic theft signatures rather than learning generic discriminative features. As a consequence, deep network architectures tend to memorise noise patterns in synthetic data, resulting in poor performance during operationally realistic scenarios (Hussain et al., 2021; Hussain et al., 2022; Javaid et al., 2023). Another source of failure stems from the inability of complex algorithms to generalise to.

Optimizing ETD ML Models: The transition of ETD ML models from centralised cloud-based platforms to resource-constrained edge-based systems demands aggressive optimisation to meet stringent constraints on computational efficiency and cost (Cheng et al., 2021; Massarani et al., 2025). Dimensionality reduction via learning-based compression using Principal Component Analysis (PCA) and K-means clustering has been shown to reduce smart meter datasets by several orders of magnitude without loss of discriminatory feature information required for intermittent cyber-attack detection, thereby facilitating substantial reductions in model training time for edge-based deployment scenarios (Massarani et al., 2025). Regarding performance assessment, the literature uniformly recognises that accuracy constitutes an inadequate metric for evaluating ML-based ETD algorithms, owing to class imbalance, and recommends the use of more robust performance metrics, including F1-score and Area Under the Receiver Operating Characteristic Curve (AUC-ROC), which offer valid statistics for assessing predictive power in both majority and minority classes (Gunturi & Sarkar, 2021). A common finding in the literature is that mitigating the false-positive rate is the sole critical performance criterion in practical ETD deployments, as each false alarm necessitates a costly field inspection of suspected consumers by utility operators (Gu et al., 2022; Ullah et al., 2022; Massarani et al., 2025; Emadaleslami et al., 2023).

Challenges

Implementation Cost & Data Acquisition: One of the primary obstacles to deploying efficient and successful ETD solutions is a lack of accurate, reliably labelled examples of real theft instances, whose disclosure may be hindered by consumer privacy concerns and legal liabilities (Yan & Wen, 2021; Oprea & Bâra, 2021). Although methods like SMOTE, ADASYN, and Generative Adversarial Network (GAN) generation are commonly used to mathematically overcome imbalanced training sets (Lin et al., 2021; Khan et al., 2022; Shehzad et al., 2023), there have been recent reports that artificially generated theft data do not capture the stochastic dynamics of real-life physical tampering incidents (Gu et al., 2022; Oprea et al., 2021). Attempts to address this issue have included modelling real three-phase power consumption scenarios using physical simulations in environments such as Simulink, yet the same problem persists (Bai et al., 2024). Additionally, dynamic graph models exhibit training instability on unbalanced theft data and require complex oversampling and a cost-sensitive loss function to generalise well (Zhuang et al., 2024).

Adversarial Vulnerability & Zero-Day Evasion: The more advanced the adversary, the more they use evasive techniques like intermittent tampering, in which the adversary cycles between honest and fraudulent behaviour, making the aggregate profile statistically similar to honest usage (Fang et al., 2023). Meter tampering, meter bypass, and meter reversal can also be performed in ways that cause undetectable alterations, thereby fooling machine learning classifiers into believing no fraud is present (Ahmed et al., 2025). Furthermore, neural networks trained to classify various cyber attacks based on a fixed set of attack scenarios may not be able to classify novel attacks (Alshehri et al., 2024). Modern neural network architectures, such as CNNs and LSTMs, are highly vulnerable to zero-day exploits, in which an attack on an unseen data set causes the model to malfunction (Alshehri et al., 2024). The zero-day attack poses another challenge for detection, as (Huang et al., 2024) reports that electricity thieves exhibit less noticeable or even irregular consumption patterns compared to genuine customers.

Hardware Design & Computational Constraints: Highly accurate machine learning methods such as Transformers and Graph Neural Networks cannot be deployed on AMI due to the memory and compute limitations of IoT-capable microcontrollers (Cheng et al., 2021; Massarani et al., 2025; Gao et al., 2022). On the other hand, incorporating multidimensional power-state features such as voltage, current, and power factor (Bai et al., 2024) increases the dimensionality of the input data, requiring specialised architectures to process them efficiently. New methods, including the MixHop graph convolutional method (Zhuang et al., 2024) and the ConvLSTM (Xia et al., 2023) variants, have shown excellent performance for anomaly detection;

however, their complexity poses practical challenges when applied at scale.

Real-Time Application & Latency: Real-time, centralised detection architectures incur significant computational overhead due to continuous telemetry communication. Edge computing architecture mitigates latency problems; however, data encryption and the fusion of multiple modalities often exceed edge-computing capacity (Wen et al., 2022; Yan & Wen, 2021; Zhang et al., 2025). Extraction of global features, as well as periodic analysis over two time intervals (day-hour and day-minute), further increases computational demand, despite providing better results (Huang et al., 2024).

Reliability & Maintenance in Large-Scale Grids: Legitimate consumer behaviour that changes over time owing to the introduction of new devices, generation resources such as roof-top solar, and other modifications in consumer behaviour leads to false alarms because the model will be trained on previous consumer data, resulting in the misclassification of normal behaviour into theft (Gu et al., 2022; Liao, Takiddin, et al., 2024). Short-term legitimate variations in consumer behaviour, such as equipment maintenance or the cessation of operations, have been identified as instances of normal behaviour misclassified as tampering, resulting in increased false alarms (Gu et al., 2022). This is particularly true for distribution network segments farther from the power grid, where voltage and current variations increase with distance (Chen et al., 2025).

Open Areas

As ETD research progresses from theory to practice, some architectural and systemic limitations become apparent. Based on a synthesis of recent research, it can be stated that the following areas require further exploration:

Data Latency & Edge Integration: Further progress will require tackling the core problem of balancing the high detection accuracy of deep neural networks with data-compression approaches such as prototype learning via PCA and K-means clustering (Massarani et al., 2025). Designing Graph Attention Networks that learn inter-consumer relationships using a dynamic network construction approach (Liao, Zhu, et al., 2024), which accounts for the complex interactions among consumers, is essential for improving the accuracy of detecting fraudulent activities across varying consumer behaviour patterns.

Policy Changes Detection and Adaptability to Changing Behavioral Patterns: To address the inherent sensitivity of static models to changing consumer behaviour, it will be crucial to develop Adaptive Transfer Learning and Deep Reinforcement Learning (DRL) architectures (El-Toukhy et al., 2023; Gunduz & Das, 2024; Liao, Takiddin, et al., 2024). Future ETD systems will need to be designed to detect policy shifts independently, ensuring that decision boundaries evolve as consumer behaviour patterns change (Althobaiti et al., 2024).

Adversarial Robustness: In light of emerging context-aware malicious attacks, developing certified robust training protocols and adversarial training algorithms will be vital to defend ETD decision boundaries (Blazakis et al., 2025; Ahmed et al., 2025). Zero-day detectors will be crucial for uncovering hidden malicious activities in datasets by detecting anomalies and patterns indicative of distortion from the original baseline (Alshehri et al., 2024).

Real-World vs. Synthetic Benchmarking: An open problem is the lack of reliable real-world datasets for ETD benchmarking, which is crucial in the field (Zidi et al., 2023; Blazakis et al., 2025). Further research into integrating Digital Twin simulation frameworks and Trackable Multi-domain Collaborative GANs (TMCGAN) can facilitate the generation of realistic attack vectors that simulate actual grid manipulation (Le et al., 2025).

Explainability & Privacy: For effective regulatory compliance and enforcement of utility actions against perpetrators, opaque black-box decision boundaries must be accompanied by Explainable AI (XAI) algorithms and tools, such as LIME and SHAP (Javaid et al., 2025). SHAP-based feature importance and attention-weight visualisations, to be developed in future research, shall help generate interpretable information to facilitate auditing processes (Javaid et al., 2025; Zhang et al., 2025).

Current Trend and Research on Machine Learning-Based ETD

In the current state of affairs, ETD research trends indicate a clear progression towards higher-dimensional feature fusion and decentralised intelligence. One common trend includes the use of transformer architecture and attention mechanism for modelling long-term temporal dependencies not captured by conventional RNN approaches (Le et al., 2025; Zhang et al., 2025). New frameworks use Convolutional Autoencoders (CAEs) combined with Transformer networks to detect multivariate consumption signatures based on joint spatial and sequential feature representations (Le et al., 2025). The trend of Multi-Modal Data Fusion is gaining traction as models are starting to include not only one-dimensional

Privacy Preserving FL is currently a dominant architecture for mitigating the systemic risk of centralised data aggregation (Wen et al., 2025; Alshehri et al., 2024). Heterogeneous Federated Learning enables federated training across disparate utility partitions while overcoming class imbalance issues (Wen et al., 2025). Meanwhile, recent advancements in Contrastive Learning and NILM methods enabled the discovery of appliance-level consumption signatures that could be distinguished from legitimate lifestyle changes caused by intentional physical tampering events (Gao et al., 2023). Moreover,

there has been a recent surge in the application of QML to high-dimensional attack detection (Blazakis et al., 2025), along with the use of self-learning systems with dual neural network models to speed up computations and reduce complexity (Althobaiti et al., 2024).

CONCLUSION

The incorporation of ML algorithms, encompassing DL frameworks, into AMI systems has revolutionised the management of NTLs within contemporary smart grid infrastructures. This review paper has analysed the progression of ETD techniques from classical tree-based ensembles to convolutional and recurrent DL models, and more recently to Transformers, reinforcement learning, and federated learning. Although significant advances in feature generation and detection accuracy have been reported, several challenges persist, including imbalanced datasets, concept drift, adversarial resilience, and computational limitations encountered when deploying machine learning algorithms in edge environments constrained by limited resources.

Based on current trends in the literature, the future of ETD appears to involve integrating distributed edge computing with XAI, federated learning that protects privacy, and dynamic learning approaches that adapt to changes in consumer behaviour and novel cyber attacks. By leveraging these technologies, power grid operators will be better able to defend their infrastructure against a range of threats, including physical security risks and cyberattacks, while ensuring the privacy of consumers' information.

REFERENCES

Abbas, S., Bouazzi, I., Ojo, S., Sampedro, G. A., Almadhor, A. S., Hejaili, A. Al, & Stolicna, Z. (2024). Improving Smart Grids Security: An Active Learning Approach for Smart Grid-Based Energy Theft Detection. *IEEE Access*, *12*, 1706–1717. https://doi.org/10.1109/ACCESS.2023.3346327

Abraham, O. A., Ochiai, H., Hossain, M. D., Taenaka, Y., & Kadobayashi, Y. (2024). Electricity Theft Detection for Smart Homes: Harnessing the Power of Machine Learning With Real and Synthetic Attacks. *IEEE Access*, *12*, 26023–26045. https://doi.org/10.1109/ACCESS.2024.3366493

Ahmed, T., Saeed, M. S., Masud, M. I., Ahmad Arfeen, Z., Baloch, M., Aman, M., & Shahzad, M. (2025). Securing Smart Grids: A Triplet Loss Function Siamese Network-Based Approach for Detecting Electricity Theft in Power Utilities. *Energies*, *18*(18). https://doi.org/10.3390/en18184957

Alamu, O., Olwal, T. O., & Migabo, E. M. (2025). Machine Learning Applications in Energy Harvesting Internet of Things Networks: A Review. In *IEEE Access* (Vol. 13, pp. 4235–4266). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2024.3525263

Al-Ghaili, A. M., Ibrahim, Z.-A., Hairi, S. A. S., Rahim, F. A., Baskaran, H., Ariffin, N. A. M., & Kasim, H. (2021). A Review of anomaly detection techniques in advanced metering infrastructure. *Bulletin of Electrical Engineering and Informatics*, *10*(1), 266–273. https://doi.org/10.11591/eei.v10i1.2026

Alshehri, A., Badr, M. M., Baza, M., & Alshahrani, H. (2024). Deep Anomaly Detection Framework Utilizing Federated Learning for Electricity Theft Zero-Day Cyberattacks. *Sensors*, *24*(10). https://doi.org/10.3390/s24103236

Althobaiti, A., Rotsos, C., & Marnerides, A. K. (2024). Adaptive Energy Theft Detection in Smart Grids Using Self-Learning With Dual Neural Network. *IEEE Transactions on Industrial Informatics*, *20*(2), 2776–2786. https://doi.org/10.1109/TII.2023.3297588

Appiah, S. Y., Akowuah, E. K., Ikpo, V. C., & Dede, A. (2023). Extremely randomised trees machine learning model for electricity theft detection. *Machine Learning with Applications*, *12*, 100458. https://doi.org/10.1016/j.mlwa.2023.100458

Arif, A., Javaid, N., Aldegheishem, A., & Alrajeh, N. (2021). Big data analytics for identifying electricity theft using machine learning approaches in microgrids for smart communities. *Concurrency and Computation: Practice and Experience*, *33*(17). https://doi.org/10.1002/cpe.6316

Badr, M., Ibrahem, M., Kholidy, H., Fouda, M., & Ismail, M. (2023). Review of the Data-Driven Methods for Electricity Fraud Detection in Smart Metering Systems. *Energies*, *16*(6), 2852. https://doi.org/10.3390/en16062852

Bai, W., Xiong, L., Liao, Y., Tan, Z., Wang, J., & Zhang, Z. (2024). Detection Method for Three-Phase Electricity Theft Based on Multi-Dimensional Feature Extraction. *Sensors*, *24*(18). https://doi.org/10.3390/s24186057

Blazakis, K., Schetakis, N., Badr, M. M., Aghamalyan, D., Stavrakakis, K., & Stavrakakis, G. (2025). Power Theft Detection in Smart Grids Using Quantum Machine Learning. *IEEE Access*, *13*, 61511–61525. https://doi.org/10.1109/ACCESS.2025.3558143

Bohani, F. A., Suliman, A., Saripuddin, M., Sameon, S. S., Md Salleh, N. S., & Nazeri, S. (2021). A Comprehensive Analysis of Supervised Learning Techniques for Electricity Theft Detection. *Journal of Electrical and Computer Engineering*, *2021*. https://doi.org/10.1155/2021/9136206

Cheng, G., Zhang, Z., Li, Q., Li, Y., & Jin, W. (2021). Energy Theft Detection in an Edge Data Center Using Deep Learning. *Mathematical Problems in Engineering*, *2021*. https://doi.org/10.1155/2021/9938475

Chen, J., Guan, Z., Bai, W., Liu, J., Zhao, Y., Zhou, J., & Xiong, L. (2025). Improved DBSCAN-Based Electricity Theft Detection Using Spatiotemporal Fusion Features. *Applied Sciences (Switzerland)*, *15*(22). https://doi.org/10.3390/app152212028

Chen, J., Nanehkaran, Y. A., Chen, W., Liu, Y., & Zhang, D. (2023). Data-driven intelligent method for detection of electricity theft. *International Journal of Electrical Power and Energy Systems*, *148*. https://doi.org/10.1016/j.ijepes.2023.108948

Dash, S. K., Roccotelli, M., Khansama, R. R., Fanti, M. P., & Mangini, A. M. (2021). Long term household electricity demand forecasting based on rnn-gbrt model and a novel energy theft detection method. *Applied Sciences (Switzerland)*, *11*(18). https://doi.org/10.3390/app11188612

El-Toukhy, A. T., Badr, M. M., Mahmoud, M. M. E. A., Srivastava, G., Fouda, M. M., & Alsabaan, M. (2023). Electricity Theft Detection Using Deep Reinforcement Learning in Smart Power Grids. *IEEE Access*, *11*, 59558–59574. https://doi.org/10.1109/ACCESS.2023.3284681

Emadaleslami, M., Haghifam, M. R., & Zangiabadi, M. (2023). A two stage approach to electricity theft detection in AMI using deep learning. *International Journal of Electrical Power and Energy Systems*, *150*. https://doi.org/10.1016/j.ijepes.2023.109088

Fang, H., Xiao, J. W., & Wang, Y. W. (2023). A machine learning-based detection framework against intermittent electricity theft attack. *International Journal of Electrical Power and Energy Systems*, *150*. https://doi.org/10.1016/j.ijepes.2023.109075

Gao, A., Mei, F., Zheng, J., Sha, H., Guo, M., & Xie, Y. (2023). Electricity Theft Detection Based on Contrastive Learning and Non-Intrusive Load Monitoring. *IEEE Transactions on Smart Grid*, *14*(6), 4565–4580. https://doi.org/10.1109/TSG.2023.3263219

Gao, H.-X., Kuenzel, S., & Zhang, X.-Y. (2022). A Hybrid ConvLSTM-Based Anomaly Detection Approach for Combating Energy Theft. *IEEE Transactions on Instrumentation and Measurement*, *71*, 1–10. https://doi.org/10.1109/TIM.2022.3201569

Gu, D., Gao, Y., Chen, K., Shi, J., Li, Y., & Cao, Y. (2022). Electricity Theft Detection in AMI with Low False Positive Rate Based on Deep Learning and Evolutionary Algorithm. *IEEE Transactions on Power Systems*, *37*(6), 4568–4578. https://doi.org/10.1109/TPWRS.2022.3150050

Gunduz, M. Z., & Das, R. (2024). Smart Grid Security: An Effective Hybrid CNN-Based Approach for Detecting Energy Theft Using Consumption Patterns. *Sensors*, *24*(4). https://doi.org/10.3390/s24041148

Gunturi, S. K., & Sarkar, D. (2021). Ensemble machine learning models for the detection of energy theft. *Electric Power Systems Research*, *192*. https://doi.org/10.1016/j.epsr.2020.106904

Huang, Q., Tang, Z., Weng, X., He, M., Liu, F., Yang, M., & Jin, T. (2024). A Novel Electricity Theft Detection Strategy Based on Dual-Time Feature Fusion and Deep Learning Methods. *Energies*, *17*(2). https://doi.org/10.3390/en17020275

Hussain, S., Mustafa, M. W., Ateyeh Al-Shqeerat, K. H., Saleh Al-rimy, B. A., & Saeed, F. (2022). Electric theft detection in advanced metering infrastructure using Jaya optimized combined Kernel-Tree boosting classifier—A novel sequentially executed supervised machine learning approach. *IET Generation, Transmission and Distribution*, *16*(6), 1257–1275. https://doi.org/10.1049/gtd2.12386

Hussain, S., Mustafa, M. W., Jumani, T. A., Baloch, S. K., Alotaibi, H., Khan, I., & Khan, A. (2021). A novel feature engineered-CatBoost-based supervised machine learning framework for electricity theft detection. *Energy Reports*, *7*, 4425–4436. https://doi.org/10.1016/j.egyr.2021.07.008

Ibrahim, N. M., Al-Janabi, S. T. F., & Al-Khateeb, B. (2021). Electricity-theft detection in smart grids based on deep learning. *Bulletin of Electrical Engineering and Informatics*, *10*(4), 2285–2292. https://doi.org/10.11591/EEI.V10I4.2875

Iftikhar, H., Khan, N., Raza, M. A., Abbas, G., Khan, M., Aoudia, M., Touti, E., & Emara, A. (2024). Electricity theft detection in smart grid using machine learning. *Frontiers in Energy Research*, *12*. https://doi.org/10.3389/fenrg.2024.1383090

Javaid, N., Hasnain, M., & Ammar, M. (2025). An AI explained data-driven framework for electricity theft detection with optimized and active machine learning. *Applied Energy*, *401*. https://doi.org/10.1016/j.apenergy.2025.126632

Kawoosa, A. I., Prashar, D., Faheem, M., Jha, N., & Khan, A. A. (2023). Using machine learning ensemble method for detection of energy theft in smart meters. *IET Generation, Transmission and Distribution*, *17*(21), 4794–4809. https://doi.org/10.1049/gtd2.12997

Kgaphola, P. M., Marebane, S. M., & Hans, R. T. (2024). Electricity Theft Detection and Prevention Using Technology-Based Models: A Systematic Literature Review. *Electricity*, *5*(2), 334–350. https://doi.org/10.3390/electricity5020017

Khan, I. U., Javeid, N., Taylor, C. J., Gamage, K. A. A., & Ma, X. (2022). A Stacked Machine and Deep Learning-Based Approach for Analysing Electricity Theft in Smart Grids. *IEEE Transactions on Smart Grid*, *13*(2), 1633–1644. https://doi.org/10.1109/TSG.2021.3134018

Kim, S., Sun, Y., Lee, S., Seon, J., Hwang, B., Kim, J., Kim, J., Kim, K., & Kim, J. (2024). Data-Driven Approaches for Energy Theft Detection: A Comprehensive Review. *Energies*, *17*(12), 3057. https://doi.org/10.3390/en17123057

Kong, X., Zhao, X., Liu, C., Li, Q., Dong, D. L., & Li, Y. (2021). Electricity theft detection in low-voltage stations based on similarity measure and DT-KSVM. *International Journal of Electrical Power and Energy Systems*, *125*. https://doi.org/10.1016/j.ijepes.2020.106544

Lepolesa, L. J., Achari, S., & Cheng, L. (2022). Electricity Theft Detection in Smart Grids Based on Deep Neural Network. *IEEE Access*, *10*, 39638–39655. https://doi.org/10.1109/ACCESS.2022.3166146

Le, T. D., Duy, N. T. M., Huy, T. H. B., Van Phu, P., Doan, H. T., & Kim, D. (2025). Advanced deep learning-based electricity theft detection in smart grids using multi-dimensional analysis with Convolutional Autoencoder and Transformer. *Engineering Applications of Artificial Intelligence*, *157*. https://doi.org/10.1016/j.engappai.2025.111333

Liao, W., Takiddin, A., Tariq, M., Chen, S., Ge, L., & Yang, Z. (2024). Sample Adaptive Transfer for Electricity Theft Detection With Distribution Shifts. *IEEE Transactions on Power Systems*, *39*(6), 7012–7024. https://doi.org/10.1109/TPWRS.2024.3375939

Liao, W., Yang, D., Ge, L., Jia, Y., & Yang, Z. (2025). Electricity theft detection in integrated energy systems considering multi-energy loads. *International Journal of Electrical Power and Energy Systems*, *164*. https://doi.org/10.1016/j.ijepes.2024.110428

Liao, W., Zhu, R., Yang, Z., Liu, K., Zhang, B., Zhu, S., & Feng, B. (2024). Electricity Theft Detection Using Dynamic Graph Construction and Graph Attention Network. *IEEE Transactions on Industrial Informatics*, *20*(4), 5074–5086. https://doi.org/10.1109/TII.2023.3331131

Lin, G., Feng, H., Feng, X., Wen, H., Li, Y., Hong, S., & Ni, Z. (2021). Electricity Theft Detection in Power Consumption Data Based on Adaptive Tuning Recurrent Neural Network. *Frontiers in Energy Research*, *9*. https://doi.org/10.3389/fenrg.2021.773805

Massarani, A. H., Badr, M. M., Baza, M., Alshahrani, H., & Alshehri, A. (2025). Efficient and Accurate Zero-Day Electricity Theft Detection from Smart Meter Sensor Data Using Prototype and Ensemble Learning. *Sensors*, *25*(13). https://doi.org/10.3390/s25134111

Mbey, C. F., Bikai, J., Yem Souhe, F. G., Foba Kakeu, V. J., & Boum, A. T. (2024). Electricity Theft Detection in a Smart Grid Using Hybrid Deep Learning-Based Data Analysis Technique. *Journal of Electrical and Computer Engineering*, *2024*. https://doi.org/10.1155/2024/6225510

Norouzi, A., Ahmadi, I., Nazari, M., Pourrostami, H., & Hoseini-Kordkheili, H. (2025). A novel framework to detect electricity theft in regions with traditional metering; applying a hybrid machine learning-based approach to low-sampling-rate data. *Results in Engineering*, *25*. https://doi.org/10.1016/j.rineng.2025.104478

Oprea, S. V., & Bâra, A. (2021). Machine learning classification algorithms and anomaly detection in conventional meters and Tunisian electricity consumption large datasets. *Computers and Electrical Engineering*, *94*. https://doi.org/10.1016/j.compeleceng.2021.107329

Oprea, S. V., Bâra, A., Puican, F. C., & Radu, I. C. (2021). Anomaly detection with machine learning algorithms and big data in electricity consumption. *Sustainability (Switzerland)*, *13*(19). https://doi.org/10.3390/su131910963

Pamir, Javaid, N., Javed, M. U., Houran, M. A., Almasoud, A. M., & Imran, M. (2023). Electricity theft detection for energy optimization using deep learning models. *Energy Science and Engineering*, *11*(10), 3575–3596. https://doi.org/10.1002/ese3.1541

Qi, R., Zheng, J., Luo, Z., & Li, Q. (2022). A Novel Unsupervised Data-Driven Method for Electricity Theft Detection in AMI Using Observer Meters. *IEEE Transactions on Instrumentation and Measurement*, *71*. https://doi.org/10.1109/TIM.2022.3189748

Quasim, M. T., Nisa, K. ul, Khan, M. Z., Husain, M. S., Alam, S., Shuaib, M., Meraj, M., & Abdullah, M. (2023). An internet of things enabled machine learning model for Energy Theft Prevention System (ETPS) in Smart Cities. *Journal of Cloud Computing*, *12*(1). https://doi.org/10.1186/s13677-023-00525-4

Saqib, S. M., Mazhar, T., Iqbal, M., Almogren, A., Ghadi, Y. Y., Shahzad, T., & Hamam, H. (2025). Utilizing machine learning ensembles for effective electricity theft detection. *Energy Exploration & Exploitation*. https://doi.org/10.1177/01445987251381989

Sebastian, P. K., Deepa, K., Neelima, N., Paul, R., & Özer, T. (2024). A comparative analysis of deep neural network models in IoT-based smart systems for energy prediction and theft detection. *IET Renewable Power Generation*, *18*(3), 398–411. https://doi.org/10.1049/rpg2.12824

Shan, J., Zeng, C., Wang, Y., Ma, Z., & Shao, X. (2026). A Method for Electricity Theft Detection Based on Markov Transition Field and Mixed Neural Network. *Information (Switzerland)*, *17*(2). https://doi.org/10.3390/info17020185

Shehzad, F., Ullah, Z., Alhussain, M., Aurangzeb, K., & Aslam, S. (2023). Deep learning-based meta-learner strategy for electricity theft detection. *Frontiers in Energy Research*, *11*. https://doi.org/10.3389/fenrg.2023.1232930

Si, Z., Liu, Z., Mu, C., Wang, M., Gong, T., Xia, X., Hu, Q., & Xiao, Y. (2025). A new deep learning based electricity theft detection framework for smart grids in cloud computing. *Computer Standards and Interfaces*, *94*. https://doi.org/10.1016/j.csi.2025.104007

Stracqualursi, E., Rosato, A., Di Lorenzo, G., Panella, M., & Araneo, R. (2023). Systematic review of energy theft practices and autonomous detection through artificial intelligence methods. *Renewable and Sustainable Energy Reviews*, *184*, 113544. https://doi.org/10.1016/j.rser.2023.113544

Tanwar, S., Kumari, A., Vekaria, D., Raboaca, M. S., Alqahtani, F., Tolba, A., Neagu, B. C., & Sharma, R. (2022). GrAb: A Deep Learning-Based Data-Driven Analytics Scheme for Energy Theft Detection. *Sensors*, *22*(11). https://doi.org/10.3390/s22114048

Taruna, A. P., Arisona, G., Irwanto, D., Bestari, A. B., & Juniawan, W. (2025). Electricity Theft Detection Using Machine Learning in Traditional Meter Postpaid Residential Customers: A Case Study on State Electricity Company (PLN) Indonesia. *IEEE Access*, *13*, 7167–7191. https://doi.org/10.1109/ACCESS.2025.3526764

Tursunboev, J., Palakonda, V., & Kang, J. M. (2024). Multi-Objective Evolutionary Hybrid Deep Learning for energy theft detection. *Applied Energy*, *363*. https://doi.org/10.1016/j.apenergy.2024.122847

Ullah, A., Javaid, N., Asif, M., Javed, M. U., & Yahaya, A. S. (2022). AlexNet, AdaBoost and Artificial Bee Colony Based Hybrid Model for Electricity Theft Detection in Smart Grids. *IEEE Access*, *10*, 18681–18694. https://doi.org/10.1109/ACCESS.2022.3150016

Ullah, A., Javaid, N., Yahaya, A. S., Sultana, T., Al-Zahrani, F. A., & Zaman, F. (2021). A Hybrid Deep Neural Network for Electricity Theft Detection Using Intelligent Antenna-Based Smart Meters. *Wireless Communications and Mobile Computing*, *2021*. https://doi.org/10.1155/2021/9933111

Wen, H., Liu, X., Lei, B., Yang, M., Cheng, X., & Chen, Z. (2025). A privacy-preserving heterogeneous federated learning framework with class imbalance learning for electricity theft detection. *Applied Energy*, *378*. https://doi.org/10.1016/j.apenergy.2024.124789

Wen, M., Xie, R., Lu, K., Wang, L., & Zhang, K. (2022). FedDetect: A Novel Privacy-Preserving Federated Learning Framework for Energy Theft Detection in Smart Grid. *IEEE Internet of Things Journal*, *9*(8), 6069–6080. https://doi.org/10.1109/JIOT.2021.3110784

Wu, Z., & Wang, Y. (2025). PE-DOCC: A Novel Periodicity-Enhanced Deep One-Class Classification Framework for Electricity Theft Detection. *Applied Sciences (Switzerland)*, *15*(4). https://doi.org/10.3390/app15042193

Xia, X., Lin, J., Jia, Q., Wang, X., Ma, C., Cui, J., & Liang, W. (2023). ETD-ConvLSTM: A Deep Learning Approach for Electricity Theft Detection in Smart Grids. *IEEE Transactions on Information Forensics and Security*, *18*, 2553–2568. https://doi.org/10.1109/TIFS.2023.3265884

Yan, Z., & Wen, H. (2021). Electricity Theft Detection Base on Extreme Gradient Boosting in AMI. *IEEE Transactions on Instrumentation and Measurement*, *70*. https://doi.org/10.1109/TIM.2020.3048784

Yan, Z., & Wen, H. (2022). Performance Analysis of Electricity Theft Detection for the Smart Grid: An Overview. *IEEE Transactions on Instrumentation and Measurement*, *71*, 1–28. https://doi.org/10.1109/TIM.2021.3127649

Žarković, M., & Dobrić, G. (2024). Artificial Intelligence for Energy Theft Detection in Distribution Networks. *Energies*, *17*(7). https://doi.org/10.3390/en17071580

Zhang, K., Wang, J., Zhu, Y., Si, Y., Yin, S., & Zhang, H. (2025). Ensemble learning framework for detecting electricity theft in smart grids using weighted average method. *Engineering Applications of Artificial Intelligence*, *156*. https://doi.org/10.1016/j.engappai.2025.111111

Zhuang, W., Jiang, W., Xia, M., & Liu, J. (2024). Dynamic Generative Residual Graph Convolutional Neural Networks for Electricity Theft Detection. *IEEE Access*, *12*, 42737–42750. https://doi.org/10.1109/ACCESS.2024.3379201

Zidi, S., Mihoub, A., Mian Qaisar, S., Krichen, M., & Abu Al-Haija, Q. (2023). Theft detection dataset for benchmarking and machine learning based classification in a smart grid environment. *Journal of King Saud University - Computer and Information Sciences*, *35*(1), 13–25. https://doi.org/10.1016/j.jksuci.2022.05.007