iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IX **Month of publication:** September 2026

DOI: https://doi.org/10.22214/ijraset.2026.84787

Hybrid Machine Learning Framework for Electricity Theft Detection and Consumption Anomaly Analysis in Nagpur Distribution Networks

Prof. Ashish Nanotkar[1], Anjali Lakdeshwar[2], Devashree Shamkule[3], Babi Urade[4], Mansi Awachat[5]
[1]*Associate Professor, Department of IT, J D College of Engineering & Technology, Nagpur Maharashtra, India*
[2, 3, 4, 5]*U.G Student, Department of IT, J D College of Engineering & Technology, Nagpur Maharashtra, India*

Abstract: Electricity theft and electricity consumption anomalies are two of the biggest problems faced by today's distribution network systems, leading to huge losses and inefficiencies. Traditional methods of theft detection are mainly based on manual inspection and are costly and ineffective in detecting more sophisticated consumption patterns. With the emergence of smart meters, huge amount of electricity consumption data can now be analyzed with machine learning methods to detect suspicious patterns. This paper aims to design a Hybrid Machine Learning Framework for Detection of Electricity Theft and Consumption Anomaly Analysis in Nagpur Distribution Networks. The proposed system will use hybrid methods of supervised and unsupervised learning algorithms to detect both types of known and unknown electricity theft patterns. Deviation analysis using a rule-based method is first used to detect any significant deviation in the consumption pattern.
Keywords: Electricity Theft Detection, Consumption Anomaly Detection, Machine Learning, Hybrid Machine Learning, Smart Grid, Smart Meter Data, Supervised Learning, Unsupervised Learning, Random Forest, Histogram Gradient Boosting.

I. INTRODUCTION

Electrical power theft is one of the primary types of technical loss occurring in modern power distribution grids. It leads to huge monetary losses for power distribution companies and makes the overall distribution process unreliable and inefficient. With the wide adoption of smart meters, big amounts of electricity consumption data are produced, which can be used to detect any suspicious activity using data analysis methods [1].The conventional methods for detecting electrical power theft rely on manual analysis and pre-defined criteria. Such methods are usually expensive, time-consuming, and not applicable to large distribution grids. The Machine Learning (ML) techniques can automatically detect unusual consumption patterns based on the analysis of historical electricity consumption patterns [2].Various studies have shown the ability of machine learning and deep learning approaches in electricity theft and consumption anomaly detections. But a single machine learning model may not be able to detect all sorts of anomalies in the consumption pattern because the electricity theft patterns may differ considerably from one consumer to another and even from one period to another [3]. This is why a mixed approach which uses several types of learning models will increase the efficiency of detection of both familiar and unfamiliar abnormal consumption patterns. In this research work, a Hybrid Machine Learning Model for Electricity Theft Detection and Consumption Anomaly Detection in the Electricity Distribution Networks of Nagpur is proposed.

II. RELATED WORK

A lot of literature is available about electricity theft detection with the use of machine learning, deep learning, anomaly detection, and hybrid techniques. For example, Ahmed et al. discussed a taxonomy and comparative analysis of the methods for the detection of energy theft including data mining, state/network based, and other techniques along with the challenges for future research [3]. In addition, Sun et al. introduced a hybrid multi-time-scale neural network for electricity theft detection of smart grid. This technique resolved the problem arising due to imbalance of theft data and also included temporal aspects of electricity consumption [4].Sim et al. proposed a method based on CNN for detecting and localizing the location of electricity theft by using smart meter and observer meter data. This technique included temporal and spatial data of distribution system [5].

The researchers have proposed a machine learning based framework for detecting intermittent electricity theft attacks. They made use of feature engineering, LightGBM and variational Bayesian Gaussian mixture modelling for identification of theft patterns which switch between normal and malicious patterns of consumption [7].Iftikhar et al. proposed a hybrid MLP-GRU model for detection of electricity theft. They made use of pre-processing and k-means SMOTE for handling data imbalance before employing deep learning techniques for detection from smart-meter consumption [8].Khan et al. have conducted a research on deep learning based electricity theft detection with AlexNet architecture on smart-meter data. This paper showed the feasibility of CNN based approaches for electricity theft detection [9].Gündüz and Das have proposed a hybrid CNN based approach for energy theft detection based on consumption patterns. Their work is concerned about the security of smart-grids and has shown the feasibility of deep learning for electricity theft detection [10].Quasim et al. have developed an IoT based Energy Theft Prevention System by using the approaches like GRU, Grey Wolf Optimization, DDRCNN, LSTM and statistical approaches [11].Tehrani et al. have proposed an online electricity theft detection framework for smart-grid data. Their hybrid approach combined distribution-transformer information with machine learning and was designed for online and distributed processing [12].

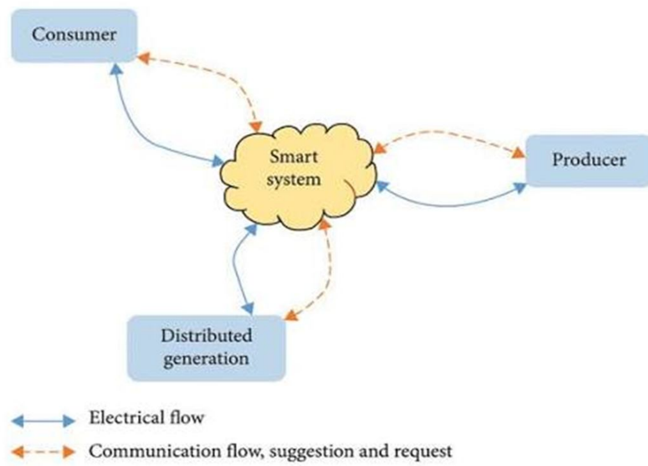


Figure 1: System Design

Figure 1: This figure shows the relationship between consumers, distributed generation, and the smart system in a smart-grid environment. Consumers stand for consumers of electrical power and those who interact with the smart system to monitor and manage their consumption. Distributed generation stands for local electricity generation systems such as solar panels and other small generation units that provide electricity to the grid while interacting with the smart system. The smart system is an interactive layer that collects information from consumers and distributed generation systems, analyzes it and controls energy flows. Arrows show two-way electricity and information exchange between the elements mentioned above.

A. Challenges and Limitations

The following limitations and challenges could be faced in the implementation of the suggested hybrid machine learning approach:

1) Lack of Real-World Implementation: The suggested approach has never been implemented and tested on a real-time electricity distribution network before. It can impact the evaluation of the practical performance of the suggested framework.
2) Quality and Availability of Data: There could be some missing values, noise, inconsistent data and a lack of theft cases in smart-meter datasets. It might affect the performance of the machine learning models.
3) Class Imbalance: Cases of electricity theft are relatively lesser than cases of normal consumption of electricity.
4) Computational Cost: The suggested system involves multiple models such as Random Forest, Histogram Gradient Boosting, Isolation Forest, and LOF. Using several models in one process increases computational cost compared to using only one model.
5) Offline Operation: The current implementation is primarily based on historical electricity consumption data, which cannot currently detect anomalies in a real-time continuous streaming manner.
6) False Positives: There may be some legitimate changes in electricity consumption that are recognized as anomalies. Seasonal variations, higher consumption by households, and consumer behavior changes could be detected as suspicious activities.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IX Sep 2026- Available at www.ijraset.com

III. INTERPRETATION

Interpretation of the consumption pattern of electricity will involve a comparison of the consumption pattern of consumers against their usual consumption behavior. Deviations in the electricity consumption will be detected through the use of rule-based deviation analysis. The proposed hybrid machine learning approach will employ Isolation Forest technique to identify any global anomaly in electricity usage while LOF will be used to detect neighborhood-based anomalies. In case of known instances of electricity theft, Random Forest and Histogram Gradient Boosting techniques will be used for supervised classification. Outputs from the two models will be merged through ensemble logic to produce an effective decision.

Table I: Models and Their Description

Model	Type	Description
Random Forest (RF)	Supervised Learning	An ensemble classification algorithm that combines multiple decision trees to classify known electricity theft patterns and improve prediction reliability.
Histogram Gradient Boosting (HGB)	Supervised Learning	A boosting-based classification model that builds multiple decision trees sequentially and is used for accurate classification of known theft cases.
Isolation Forest (IF)	Unsupervised Learning	An anomaly detection algorithm that isolates unusual observations from normal consumption data and is used to identify global consumption anomalies.
Local Outlier Factor (LOF)	Unsupervised Learning	Detects anomalies by comparing the local density of a consumer's electricity consumption with that of neighboring data points.
Rule-Based Deviation Analysis	Rule-Based	Calculates deviations from expected or normal consumption patterns and identifies significantly unusual electricity usage.
Ensemble Logic	Hybrid Approach	Combines the outputs of supervised, unsupervised, and rule-based techniques to generate a more reliable final decision and reduce false positives.

IV. FINDINGS

The literature shows that hybrid and ensemble models can improve electricity theft detection compared with individual models [15], [16].

1) Unsupervised approaches are useful for detecting unknown and previously unseen theft patterns [15], [18].
2) Deep learning methods demonstrate strong capability in learning complex electricity consumption patterns [17], [19].
3) Periodicity-based approaches can improve the detection of abnormal consumption behavior [18].
4) Combining CNN, Autoencoder, and Transformer techniques can capture multidimensional and temporal consumption characteristics [19].
5) Ensemble learning can improve the robustness of electricity theft classification by combining multiple models [16], [20].

V. FUTURE ENHANCEMENT

The proposed framework could be improved by implementing deep learning algorithms such as LSTM, CNN, and Transformers which will help to learn complex temporal and behavioral patterns in electrical consumption. The system can be further modified to include data streaming from the smart meter for better real-time monitoring of the electricity theft. IoT and cloud-based technologies can be integrated to make the process of collecting, analyzing, and deploying the model easier. Explainable Artificial Intelligence (XAI) can be implemented in the framework so as to provide understandable reasoning for any theft or anomaly in electricity consumption. Continuous model re-training could also be performed in order to make the system adaptive to changes in electricity consumption and theft. In addition, the framework could be further validated on the real world distribution network data and then used in other cities and regions in order to examine the generalizability of the proposed framework. In the future, a mobile or a web-based alerting system can be implemented to notify the electricity officials of any anomalies detected by the system.

VI. OBJECTIVES

1) To gather and pre-process electricity consumption data
2) To analyze consumption patterns and recognize deviations
3) To recognize electricity theft examples based on supervised learning
4) To recognize unknown anomalies based on unsupervised learning
5) In developing a hybrid ensemble framework

A. Project Overview

The proposed project “Hybrid Machine Learning Framework for Electricity Theft Detection and Consumption Anomaly Analysis in Nagpur Distribution Networks” is aimed at developing an automated system that will help in identifying suspicious electricity consumption patterns and detecting any electricity theft cases. The system takes electricity consumption as input, does preprocessing and deviation analysis and uses both supervised and unsupervised machine learning techniques. Random Forest and Histogram Gradient Boosting are applied to identify known patterns of theft, while Isolation Forest and Local Outlier Factor (LOF) are used to detect any unknown or never seen before anomaly. The results are then combined with the use of ensemble logic and the anomalies detected are presented through a dashboard.

Table II. Overview Of Issues Approaches, Methods

r. No.	Author(s) [Ref]	Issue Discussed	Approach & Method
1	Zhang et al. [1]	Difficulty in detecting intermittent electricity theft attacks	Machine learning-based detection framework
2	Kumar et al. [2]	Need for automated electricity theft prevention in smart grids	IoT-enabled machine learning model
3	Ahmed et al. [3]	Lack of comprehensive categorization and comparison of electricity theft detection methods	Taxonomy and comparative analysis
4	Sun et al. [4]	Theft patterns vary across different time scales	Hybrid multi-time-scale neural network
5	Sim et al. [5]	Difficulty in detecting and locating electricity theft in distribution systems	CNN-based theft detection and location
6	Grant et al. [6]	Inefficient manual electricity	Automated Machine Learning

		theft detection in developing countries	(AutoML)
7	Fang et al. [7]	Detection of intermittent electricity theft attacks	LightGBM and Gaussian Mixture Model (GMM)
8	Iftikhar et al. [8]	Imbalanced smart-meter data and complex theft patterns	GRU + MLP with SMOTE
9	Khan et al. [9]	Detection of complex theft patterns from smart-meter data	AlexNet-based deep learning
10	Gündüz and Das [10]	Energy theft and security challenges in smart grids	Hybrid CNN-based approach
11	Quasim et al. [11]	Need for IoT-based automated electricity theft prevention	IoT-enabled machine learning
12	Tehrani et al. [12]	Large-scale and online electricity theft detection	Online and distributed ML framework
13	Guarda et al. [13]	Challenges in detecting non-technical losses without hardware	Review of data-oriented, network-oriented, and hybrid methods
14	Chen et al. [14]	Difficulty in extracting meaningful features from consumption time series	Gramian Angular Field + Deep Residual Network
15	Huang and Qiu [15]	Detection of abnormal electricity usage in time-series data	LSTM-based time-series detection
16	Almalki [16]	Detection of unknown and previously unseen electricity theft	Hybrid unsupervised learning using RF, Isolation Forest, One-Class SVM, LOF and DBSCAN
17	Tsai [17]	Limitations of individual machine learning models	Ensemble electricity theft detection algorithm
18	Elshennawy et al. [18]	Detection of fraudulent electricity consumption patterns	Deep learning-based electricity theft detection
19	Wu and Wang [19]	Difficulty in detecting unseen theft patterns	Periodicity-Enhanced Deep One-Class Classification
20	Le et al. [20]	Complex temporal and multidimensional consumption patterns	Convolutional Autoencoder + Transformer with K-means SMOTE

VII. CONCLUSION

The presented Hybrid Machine Learning Framework for Electricity Theft Detection and Consumption Anomaly Detection is aimed at automatic detection of any suspicious consumption of electricity. Rule-based deviation analysis, supervised and unsupervised learning methods are combined in order to find not only known electricity theft cases, but also new consumption anomalies. Known electricity theft cases are recognized with the help of Random Forest and Histogram Gradient Boosting classifiers, while Isolation Forest and Local Outlier Factor (LOF) find global and local anomalies. It is supposed that the combination of these methods will increase the effectiveness of the detection process and will decrease the probability of false alarms. Furthermore, the developed system can help electricity companies with detecting any abnormalities in consumption via the dashboard. In such a way, there will be no need for manual detection of the anomalies.

REFERENCES

[1] Y. Zhang, L. Wang, and H. Liu, "A Machine Learning-Based Detection Framework Against Intermittent Electricity Theft Attacks," International Journal of Electrical Power & Energy Systems, vol. 149, Jul. 2023.

[2] S. Kumar, R. Singh, and A. Verma, "An IoT-Enabled Machine Learning Model for Energy Theft Prevention System," Journal of Cloud Computing, vol. 12, no. 1, 2023.

[3] R. Mehmood and K. Abbas, "Systematic Review of Household Electricity Anomaly Detection Using Machine Learning," Journal of Electrical Systems and Information Technology, 2023.

[4] M. Ahmed, A. Khan, M. Ahmed, M. Tahir, G. Jeon, G. Fortino, and F. Piccialli, "Energy theft detection in smart grids: Taxonomy, comparative analysis, challenges, and future research directions," IEEE/CAA Journal of Automatica Sinica, vol. 9, no. 4, pp. 578–600, Apr. 2022, doi: 10.1109/JAS.2022.105404.

[5] Y. Sun, X. Sun, T. Hu, and L. Zhu, "Smart grid theft detection based on hybrid multi-time scale neural network," Applied Sciences, vol. 13, no. 9, p. 5710, 2023, doi: 10.3390/app13095710.

[6] K. Sim, G. Lee, and Y. J. Shin, "Detection and location of electricity theft via convolutional neural network in distribution system," in Proc. 2023 IEEE Power & Energy Society General Meeting (PESGM), 2023, doi: 10.1109/PESGM52003.2023.10253338.

[7] L. Grant, H. Latchman, and K. Alli, "Detection of electricity theft in developing countries—A machine learning approach," Trends in Computer Science and Information Technology, vol. 8, no. 2, pp. 038–049, 2023, doi: 10.17352/tcsit.000067.

[8] H. Fang, J.-W. Xiao, and Y.-W. Wang, "A machine learning-based detection framework against intermittent electricity theft attack," International Journal of Electrical Power & Energy Systems, vol. 150, p. 109075, 2023, doi: 10.1016/j.ijepes.2023.109075.

[9] H. Iftikhar, N. Khan, M. A. Raza, et al., "Electricity theft detection in smart grid using machine learning," Frontiers in Energy Research, vol. 12, 2024, doi: 10.3389/fenrg.2024.1383090.

[10] N. Khan, M. A. Raza, D. Ara, et al., "A novel deep learning technique to detect electricity theft in smart grids using AlexNet," IET Renewable Power Generation, 2024, doi: 10.1049/rpg2.12846.

[11] M. Z. Gündüz and R. Das, "Smart grid security: An effective hybrid CNN-based approach for detecting energy theft using consumption patterns," Sensors, vol. 24, no. 4, p. 1148, 2024, doi: 10.3390/s24041148.

[12] M. T. Quasim, K. u. Nisa, M. Z. Khan, M. S. Husain, S. Alam, M. Shuaib, M. Meraj, and M. Abdullah, "An internet of things enabled machine learning model for Energy Theft Prevention System (ETPS) in Smart Cities," Journal of Cloud Computing, vol. 12, art. no. 158, 2023, doi: 10.1186/s13677-023-00525-4.

[13] S. O. Tehrani, A. Shahrestani, and M. H. Yaghmaee, "Online electricity theft detection framework for large-scale smart grid data," Electric Power Systems Research, vol. 208, p. 107895, 2022, doi: 10.1016/j.epsr.2022.107895.

[14] F. G. K. Guarda, B. K. Hammerschmitt, M. B. Capeletti, N. K. Neto, L. L. C. dos Santos, L. R. Prade, and A. Abaide, "Non-hardware-based non-technical losses detection methods: A review," Energies, vol. 16, no. 4, p. 2054, 2023, doi: 10.3390/en16042054.

[15] Y. Chen, J. Li, Q. Huang, K. Li, Z. Zhao, and X. Ren, "Non-technical losses detection with Gramian angular field and deep residual network," Energy Reports, vol. 9, pp. 1392–1401, 2023, doi: 10.1016/j.egyr.2023.05.183.

[16] J. Huang and Y. Qiu, "LSTM-based time series detection of abnormal electricity usage in smart meters," Preprints.org, 2025, doi/Manuscript: 202506.1404.

[17] **[15]** A. J. Almalki, "Unsupervised Learning With Hybrid Models for Detecting Electricity Theft in Smart Grids," IEEE Access, vol. 12, pp. 187027–187040, 2024, doi: 10.1109/ACCESS.2024.3498733.

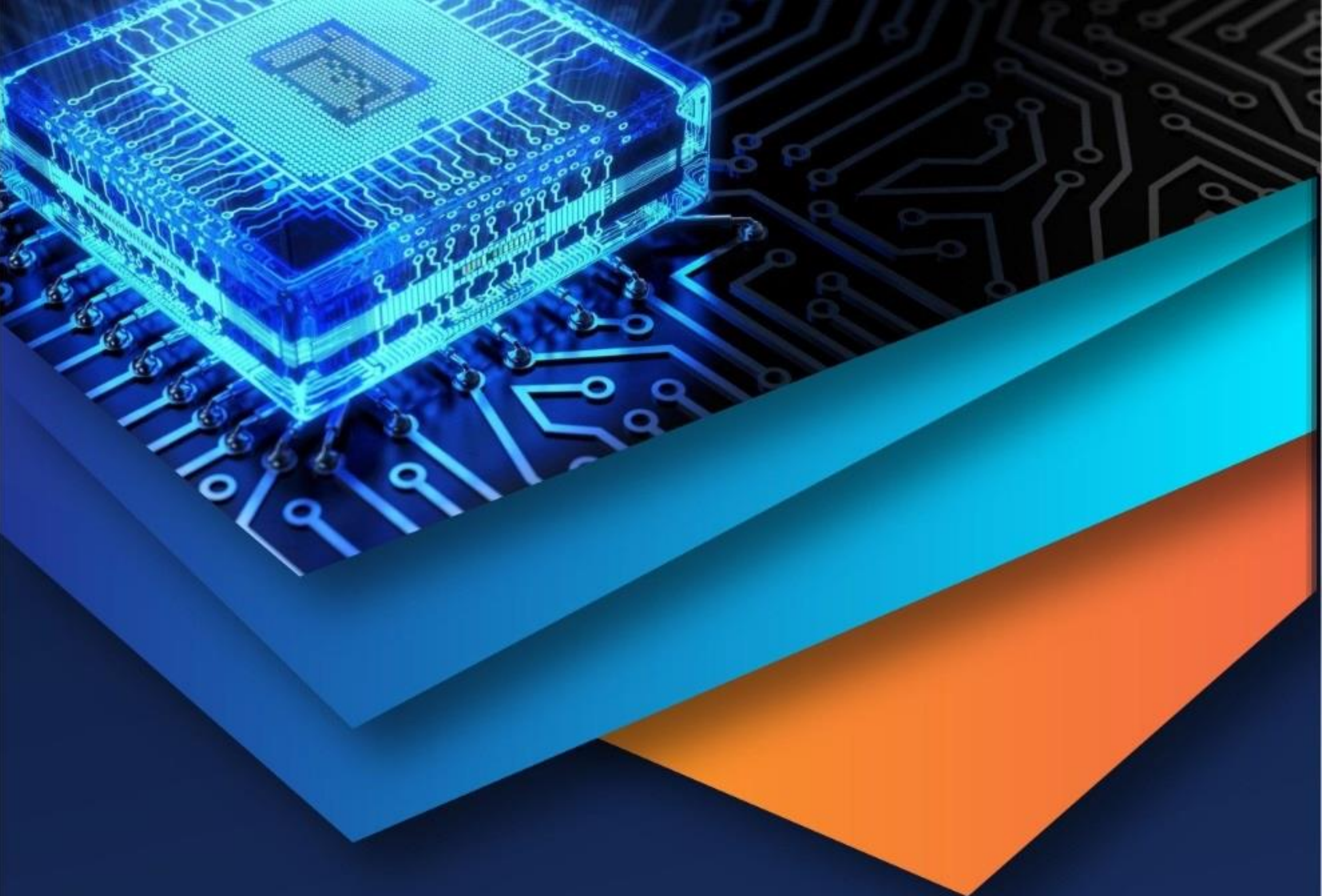
[18] C.-W. Tsai, "An effective ensemble electricity theft detection algorithm for smart grid," IET Networks, 2024, doi: 10.1049/ntw2.12132.

[19] N. M. Elshennawy, D. M. Ibrahim, and A. M. Gab Allah, "An efficient electricity theft detection based on deep learning," Scientific Reports, vol. 15, Art. no. 12866, 2025.

[20] Z. Wu and Y. Wang, "PE-DOCC: A Novel Periodicity-Enhanced Deep One-Class Classification Framework for Electricity Theft Detection," Applied Sciences, vol. 15, no. 4, p. 2193, 2025, doi: 10.3390/app15042193.

[21] T.-D. Le, N. T. M. Duy, T. H. B. Huy, P. V. Phu, H. T. Doan, and D. Kim, "Advanced deep learning-based electricity theft detection in smart grids using multi-dimensional analysis with Convolutional Autoencoder and Transformer," Engineering Applications of Artificial Intelligence, vol. 157, Art. no. 111333, 2025, doi: 10.1016/j.engappai.2025.111333.

[22] "Detecting electricity theft in smart grids using optimized machine learning ensemble techniques and eXplainable AI," Energy Reports, vol. 14, pp. 3142–3162, 2025, doi: 10.1016/j.egyr.2025.09.019.



10.22214/IJRASET

INDEX COPERNICUS

45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY