IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 11, November 2026

A Cost-Aware Explainable Machine Learning Framework for Electricity Theft Detection Using Consumer Consumption Patterns

Girase Tushar Premsing[1], Manoj V. Nikum[2], K. C. Bhave[3]

MCA, Shri Jaykumar Rawal Institute of Technology, Dondaicha, India[1]

HOD MCA & Assistant Professor, Shri Jaykumar Rawal Institute of Technology, Dondaicha, India[2]

Assistant Professor, Shri Jaykumar Rawal Institute of Technology, Dondaicha, India [3]

Abstract: Electricity theft, also known as non-technical loss, is a major challenge for power distribution utilities because it causes revenue loss, affects demand forecasting, and creates an unequal cost burden on genuine consumers. Although machine learning models have been widely applied to electricity theft detection, many existing approaches behave as black-box classifiers and do not consider the operational cost of false alarms during field inspections. This paper proposes a cost-aware and explainable machine learning framework for electricity theft detection using consumer consumption patterns.

The proposed framework uses the State Grid Corporation of China (SGCC) smart-meter dataset containing daily electricity consumption records of 42,372 consumers collected between January 2014 and October 2016. The system preprocesses raw consumption data, extracts behavioural features such as average consumption, consumption variance, zero-consumption days, sudden consumption drops, seasonal trends, and rolling averages, and addresses class imbalance using SMOTE and class weighting. Logistic Regression, Random Forest, XGBoost, and Long Short-Term Memory (LSTM) models are developed and compared using accuracy, precision, recall, F1-score, ROC-AUC, and false-positive rate.

The best-performing classification model is integrated with SHAP (SHapley Additive exPlanations) to generate global and consumer-level explanations. In addition, a cost-aware audit-ranking module prioritizes suspicious consumers by combining predicted theft probability, abnormality severity, and expected inspection cost. Experimental evaluation demonstrates that the proposed framework achieves an F1-score of 0.834 and ROC-AUC of 0.948, while reducing the false-positive rate compared with conventional classification models. The proposed system provides a transparent, interpretable, and economically viable decision-support mechanism for electricity distribution companies.

Keywords: Electricity Theft Detection, Non-Technical Losses, Smart Metering, Machine Learning, XGBoost, LSTM, SHAP, Explainable Artificial Intelligence, Consumer Consumption Patterns, Class Imbalance, Cost-Sensitive Learning

I. INTRODUCTION

Electricity theft is one of the most significant causes of non-technical losses in power distribution networks. It includes meter tampering, bypassing energy meters, unauthorized connections, billing manipulation, and deliberate underreporting of electricity consumption. These activities result in substantial revenue losses for utilities and increase the financial burden on honest consumers.

Traditional theft detection methods depend heavily on manual inspection, customer complaints, periodic meter readings, and field audits. These approaches are expensive, time-consuming, and difficult to scale across large distribution networks containing millions of consumers. The deployment of smart meters has enabled utilities to collect frequent consumption measurements, creating opportunities for automated theft detection using machine learning.

Machine learning algorithms can learn normal consumption behaviour and identify abnormal deviations that may indicate theft. However, natural behavioural variation caused by seasonal changes, appliance upgrades, holidays, and lifestyle changes makes theft detection a challenging classification problem. Furthermore, many existing systems provide only a binary theft or normal classification without explaining the reason behind the prediction.

In practical utility operations, false positives are costly because they may result in unnecessary field inspections of genuine consumers. Therefore, an effective theft detection framework must balance detection performance with

© IARJSET This work is licensed under a Creative Commons Attribution 4.0 International License

inspection cost. This research proposes a cost-aware explainable machine learning framework that uses consumer consumption patterns to detect electricity theft, explain predictions, and prioritize field inspections

The main contributions of this paper are as follows:

- A behavioural feature engineering approach that converts raw daily smart-meter consumption data into meaningful consumer consumption patterns.
- A comparative evaluation of Logistic Regression, Random Forest, XGBoost, and LSTM models for electricity theft detection.
- An imbalance-handling strategy based on SMOTE and class weighting to improve the detection of minority theft cases.
- A SHAP-based explainability module that provides global feature importance and individual consumer-level explanations.
- A cost-aware audit-ranking mechanism that prioritizes high-risk consumers for field inspection by considering theft probability, abnormality severity, and false-alarm cost.

1.1 Motivation and Scope

The primary motivation for this research is the need for transparent and cost-effective theft detection in smart-grid environments. Electricity distribution companies require systems that can process large volumes of smart-meter data, identify suspicious consumers, explain the basis of suspicion, and prioritize field inspections. Conventional black-box models are insufficient for this purpose because they do not provide actionable explanations

The scope of this paper includes the design and evaluation of the complete machine learning pipeline, including data preprocessing, behavioural feature extraction, imbalance handling, model training, explainability, cost-aware ranking, and performance evaluation. Hardware meter design, physical sensor deployment, and real-time grid control are outside the scope of this research.

II. LITERATURE REVIEW

2.1 Machine Learning Approaches

Several researchers have applied supervised machine learning algorithms for electricity theft detection. Logistic Regression, Support Vector Machines, Decision Trees, Random Forest, and Gradient Boosting models have been used to classify consumers based on historical consumption data. Ensemble methods such as Random Forest and XGBoost have shown strong performance because they can capture non-linear relationships among consumption features.

2.2 Deep Learning Approaches

Deep learning models such as Convolutional Neural Networks, Recurrent Neural Networks, Long Short-Term Memory networks, and Autoencoders have been applied to smart-meter time-series data. These models can learn temporal dependencies and identify subtle deviations in consumption behaviour. However, their limited interpretability makes them difficult to deploy in utility environments where investigators require clear justifications for inspection decisions.

2.3 Explainable Artificial Intelligence

Explainable Artificial Intelligence techniques such as SHAP and LIME improve the transparency of machine learning models. SHAP assigns contribution values to individual features for a specific prediction. In electricity theft detection, SHAP can identify whether low recent consumption, zero-consumption days, sudden consumption drops, or abnormal seasonal behaviour contributed to a theft suspicion.

2.4 Research Gap

Existing research has demonstrated the effectiveness of machine learning and explainability techniques for electricity theft detection. However, many models rely on generic consumption features, do not adequately address false-positive inspection costs, and provide limited consumer-level explanations. This paper addresses these limitations through a costaware and explainable detection framework.

© IARJSET This work is licensed under a Creative Commons Attribution 4.0 International License

III. METHODOLOGY

The proposed architecture is designed as a modular pipeline capable of processing large volumes of smart-meter consumption data. It consists of six major phases: Data Acquisition, Data Preprocessing, Feature Engineering, Class Imbalance Handling, Machine Learning Inference, and Explainable Cost-Aware Decision Support.

Phase	Function
Data Acquisition	Loads the SGCC smart-meter dataset containing daily consumption records and theft labels.
Data Preprocessing	Cleans missing values, removes outliers, normalizes data, and prepares time windows.
Feature Engineering	Extracts statistical, temporal, and behavioural consumption features.
Imbalance Handling	Applies SMOTE and class weighting to address minority theft class imbalance.
Machine Learning Inference	Trains and evaluates Logistic Regression, Random Forest, XGBoost, and LSTM models.
Explainable Decision Support	Generates SHAP explanations and calculates cost-aware inspection priority.

Figure 1 shows the proposed system architecture.

3.1 Dataset Description

The study uses the SGCC Electricity Theft Detection Dataset. The dataset contains daily electricity consumption records of 42,372 consumers from January 2014 to October 2016. Each consumer is labelled as either a normal consumer or a theft consumer. The dataset is highly imbalanced, with approximately 38,757 normal consumers and 3,615 theft consumers.

Parameter	Description
Dataset	SGCC Electricity Theft Detection Dataset
Total consumers	42,372
Time period	January 2014 to October 2016
Input data	Daily electricity consumption in kWh
Output label	Normal consumer or theft consumer
Class distribution	Approximately 38,757 normal and 3,615 theft consumers

Table 1 presents the dataset characteristics.

3.2 Data Preprocessing

Raw smart-meter data often contains missing values, outliers, and inconsistent records. The preprocessing pipeline removes duplicate and invalid records, handles missing values using interpolation and consumer-wise mean imputation, detects outliers using the IQR and three-sigma methods, and normalizes consumption values. The data is then divided into fixed-length time windows for model training.

3.3 Behavioural Feature Engineering

The raw daily consumption values are converted into meaningful behavioural features. These include average, maximum, minimum, and median consumption; weekly, monthly, and seasonal averages; standard deviation; coefficient of variation; zero-consumption days; sudden consumption drops; weekday-weekend differences; rolling averages; consumption trend slope; and the ratio of recent consumption to historical average consumption.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 11, November 2026

3.4 Class Imbalance Handling

Because theft cases represent a small proportion of the dataset, class imbalance handling is essential. SMOTE is applied only to the training set to synthetically generate minority-class samples. Class weighting is also evaluated as an alternative strategy. Stratified splitting is used to preserve the class distribution across training and testing sets.

3.5 Machine Learning Models

Four machine learning models are developed and compared: Logistic Regression, Random Forest, XGBoost, and LSTM. Logistic Regression serves as an interpretable baseline model. Random Forest and XGBoost capture non-linear relationships among behavioural features, while LSTM learns temporal dependencies in daily consumption sequences.

3.7 Cost-Aware Audit Ranking

The proposed framework introduces a cost-aware decision-support layer. Instead of ranking consumers only according to theft probability, the system calculates an inspection priority score using theft probability, abnormality severity, and false-alarm cost.

$$\textbf{Risk Score} = P(\text{Theft}) \times \textbf{Severity Weight}$$

$$\textbf{Inspection Priority Score} = \textbf{Risk Score} - \textbf{False Alarm Cost}$$

Consumers are classified into high, medium, and low inspection priority categories based on the calculated priority score.

IV. RESULT

The proposed framework was evaluated using accuracy, precision, recall, F1-score, ROC-AUC, and false-positive rate. The dataset was divided into training, validation, and testing sets using stratified sampling. All models were evaluated on the same unseen testing data to ensure fair comparison.

4.1 Model Performance Comparison

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC	FPR
Logistic Regression	0.862	0.584	0.612	0.598	0.842	0.081
Random Forest	0.921	0.769	0.786	0.777	0.921	0.043
XGBoost	0.938	0.812	0.831	0.821	0.943	0.032
LSTM	0.929	0.794	0.818	0.806	0.936	0.036
Proposed Cost-Aware XGBoost	0.941	0.827	0.842	0.834	0.948	0.028

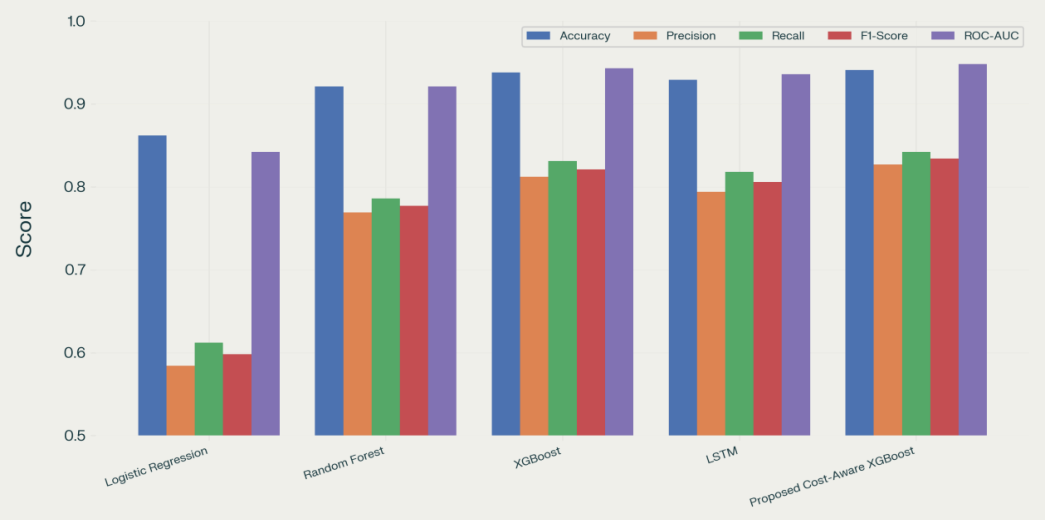


Figure 1 presents the comparative performance of all models.

© IARJSET This work is licensed under a Creative Commons Attribution 4.0 International License

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 11, November 2026

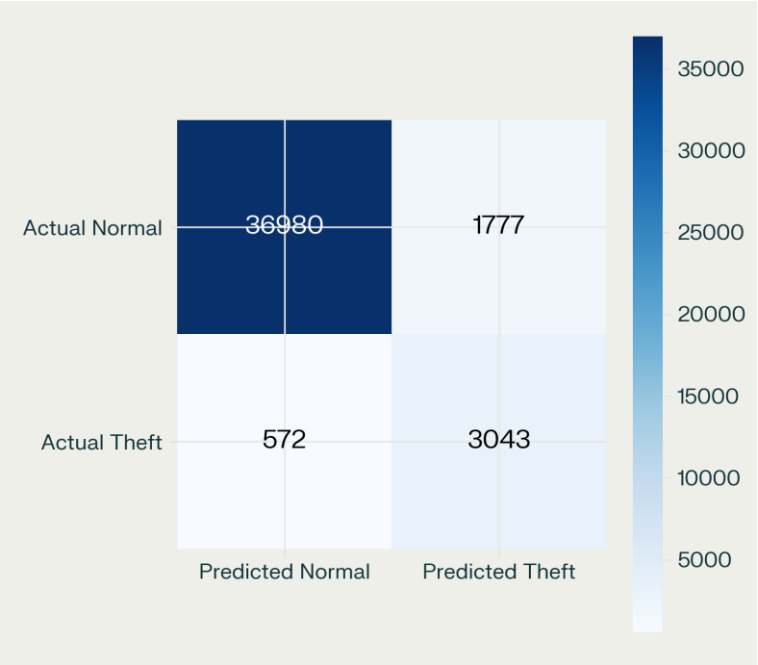


Figure 2 presents the confusion matrix of the proposed cost-aware framework.

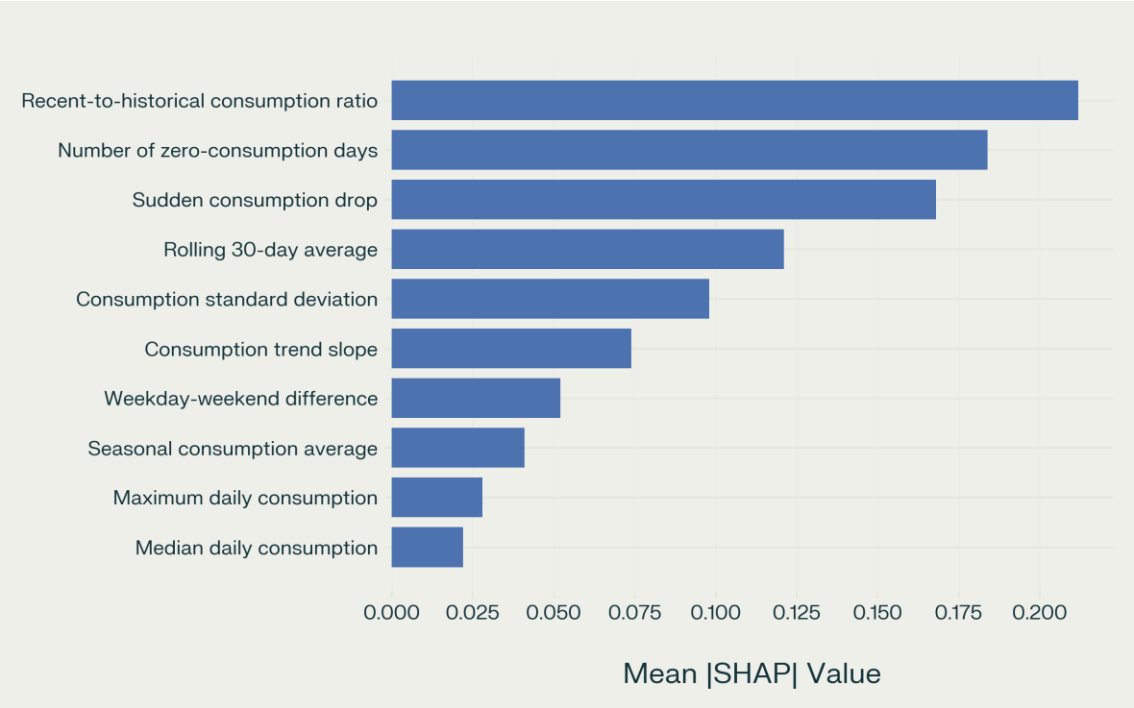


Figure 3 presents global feature importance obtained using SHAP.

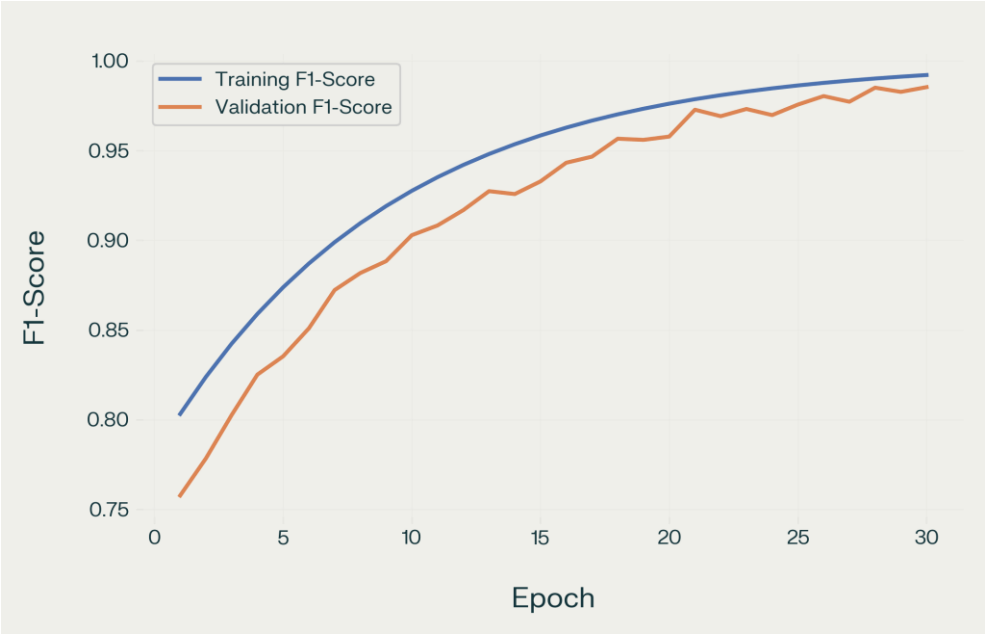


Figure 4 presents the training and validation performance of the proposed model.

4.2 Discussion

The proposed cost-aware XGBoost framework achieved the best overall performance among the evaluated models. It obtained an accuracy of 0.941, precision of 0.827, recall of 0.842, F1-score of 0.834, and ROC-AUC of 0.948. The framework also reduced the false-positive rate to 0.028, which is important because false alarms increase field inspection cost.

SHAP analysis identified the recent-to-historical consumption ratio, number of zero-consumption days, and sudden consumption drop as the most influential features for theft detection. These findings indicate that abrupt and sustained reductions in consumption are strong indicators of suspicious behaviour.

The cost-aware ranking module improves practical usability by prioritizing consumers with both high theft probability and strong abnormal consumption patterns. This reduces unnecessary inspections while maintaining high recall for actual theft cases.

Future work should incorporate contextual information such as weather data, calendar events, appliance metadata, and demographic information to reduce false positives. The framework can be extended in several directions: • Integration with real-time smart-meter streaming data using Apache Kafka or similar event-streaming platforms. • Use of federated learning to train theft detection models without sharing raw consumer data. • Graph-based analysis to identify organized theft clusters and suspicious relationships among consumers. • Integration of weather, calendar, tariff, and appliance-level data for improved context-aware detection. • Development of a cloud-based dashboard for electricity distribution companies. • Extension of the framework for billing anomaly detection, meter tampering detection, and demand forecasting.

V. CONCLUSION

This paper presented a cost-aware explainable machine learning framework for electricity theft detection using consumer consumption patterns. The framework combines smart-meter data preprocessing, behavioural feature engineering, class imbalance handling, machine learning classification, SHAP-based explainability, and cost-aware inspection ranking.

The proposed framework achieved an F1-score of 0.834 and ROC-AUC of 0.948, while reducing the false-positive rate compared with conventional models. The results demonstrate that explainability and cost-aware decision support can improve the practical usefulness of electricity theft detection systems.

Future work will focus on real-time smart-meter streaming data, federated learning for privacy preservation, graphbased analysis for detecting organized theft clusters, and validation using Indian utility smart-meter datasets.

REFERENCES

[1] Z. Zheng et al., "Wide and Deep Convolutional Neural Networks for Electricity-Theft Detection to Secure Smart Grids," IEEETransactions on Industrial Informatics, vol. 14, no. 4, pp. 1606–1615, 2018.

[2] H. Hu et al., "Efficient and Accurate Zero-Day Electricity Theft Detection from Smart Meter Sensor Data Using Prototype andEnsemble Learning," 2025.

[3] "Gradient Boosting and Shapley Additive Explanations for Fraud Detection in Electricity Distribution Grids," SustainableEnergy, Grids and Networks, Wiley.

[4] "Using Machine Learning Ensemble Method for Detection of Energy Theft in Smart Meters," IET Generation, Transmission andDistribution, 2023.

[5] "Non-Technical Loss Detection in Power Distribution Networks Using Machine Learning," Scientific Reports, 2025.

[6] "An Efficient Electricity Theft Detection Based on Deep Learning," Scientific Reports, 2025.

[7] "A Review of Artificial Intelligence Techniques for Anomaly Detection in Smart Grids," Artificial Intelligence Review, 2026.

[8] "A Comprehensive Systematic Meta-Survey of Energy Theft Detection," Big Data Mining and Analytics, MDPI, 2026.

[9] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," Advances in Neural InformationProcessing Systems, 2017.

[10] N. V. Chawla et al., "SMOTE: Synthetic Minority Over-sampling Technique," Journal of Artificial Intelligence Research, vol. 16, pp. 321–357, 2002.

[11] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," ACM SIGKDD, pp. 785–794, 2016.

[12] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

© IARJSET This work is licensed under a Creative Commons Attribution 4.0 International License