International Journal of Creative and Open Research in Engineering and Management
ISSN: **3108-1754 (Online)**
Volume 02 Issue 04 April-2026 | Impact Factor: 3.5

Machine Learning Based Detection of Electricity Theft using Smart Meter Data

K Naresh[1], Pandikunta Sreelatha[2]

[1]Assistant Professor, Department of MCA, Annamacharya Institute of Technology and Sciences, Tirupati, Andhra Pradesh, India.

[2]Postgraduate, Department of MCA, Annamacharya Institute of Technology and Sciences, Tirupati, Andhra Pradesh, India.

How to Cite this Article:

Sreelatha, P. (2026). Machine Learning Based Detection of Electricity Theft using Smart Meter Data. International Journal of Creative and Open Research in Engineering and Management, 02(04). https://doi.org/10.55041/ijcope.v2i4.074

https://doi.org/10.55041/ijcope.v2i4.074

Abstract: Energy fraud and electricity theft are serious problems that impact power systems' dependability and result in large financial losses for power distribution corporations. Because smart meters provide a lot of data about electricity consumption, it is challenging to manually detect such fraudulent operations. Using data on electricity consumption, this study suggests a machine learning-based method for identifying power theft. The technology looks for unusual behavior that can point to fraudulent activities by analyzing customer usage trends. Based on their patterns of electricity usage, a number of machine learning techniques, such as Logistic Regression, Decision Tree, Random Forest, and Support Vector Machine (SVM), are used to categorize customers as normal or fraudulent. To increase prediction accuracy and dependability, the dataset is preprocessed and examined. Python and the Django framework are used to create the system as a web-based application that enables users to upload consumption data and obtain fraud detection findings. Machine learning approaches can successfully identify suspicious consumption behavior and increase the accuracy of fraud detection, according to experimental investigation. Power utility firms can monitor electricity use, lower revenue losses, and increase the effectiveness of energy management systems with the help of the suggested solution.

Keywords: Electricity Theft Detection, Energy Fraud Detection, Machine Learning, Smart Meter Data Analysis, Electricity Consumption Analysis, Fraud Classification, Data Mining, Smart Grid Security.

1. Introduction

One of the most vital resources for contemporary society is electricity, and economic growth depends on effective energy distribution. However, energy fraud and electricity theft have grown to be significant problems for power distribution corporations worldwide. Unauthorized electricity use lowers the stability and dependability of power systems and causes large financial losses. Conventional techniques for identifying electricity theft mostly rely on rule-based monitoring and manual inspections, both of which are laborious and frequently ineffective. A significant amount of data on electricity consumption is produced every day due to the quick development of smart meters and digital energy monitoring devices. Unusual consumption patterns that can point to fraudulent activity can be found by analyzing this data using sophisticated computer tools. By examining past consumption data and spotting unusual usage trends, machine learning offers a practical way to identify electricity theft. These methods are able to automatically identify trends in massive datasets and discern between typical and questionable electricity use patterns. The technology can detect anomalous patterns that can point to fraudulent behavior by looking at fluctuations in usage trends. In this research, consumer electricity usage data gathered from smart meters is analyzed to create a machine learning-based electricity theft detection system. In order to identify anomalous consumption habits that deviate from typical usage patterns, the system processes and analyzes the data. Using Python and the Django framework, it is constructed as a web-based application that enables users to upload data on electricity consumption and effectively obtain fraud detection findings. The suggested method makes it easier for electricity distribution firms to spot questionable consumption trends. It enhances the dependability and effectiveness of power management systems, promotes improved monitoring of electricity usage, and lowers revenue losses brought on by unauthorized consumption. Furthermore, by automating the fraud detection process, the solution lets energy providers make decisions more quickly by eliminating the need for manual inspections.

2. Literature Review

Due to the growing losses that power distribution companies around the world are experiencing, electricity theft detection has emerged as a significant study subject. Conventional techniques for detecting electricity theft mostly rely on rule-based monitoring systems and manual inspections. These methods are frequently ineffective, time-consuming, and unable to manage the massive amount of data on electricity use produced by contemporary smart grids. Utilizing data-driven methods to increase the precision and effectiveness of electrical fraud detection has been the subject of recent research. Large volumes of data on electricity consumption can be evaluated to find unusual usage patterns thanks to the advancement of smart meters and sophisticated data collection technology. Numerous intelligent algorithms that examine past energy usage data to identify anomalous behavior that might point to power theft have been proposed by researchers. Numerous studies emphasize how crucial it is to examine consumer consumption trends in order to identify questionable activity. These systems can detect anomalous consumption patterns that deviate from typical user behavior by analyzing changes in daily and monthly electricity usage. To increase the quality of the dataset and boost the effectiveness of fraud detection algorithms, data preprocessing and feature analysis are frequently employed. The introduction of automated detection systems coupled with web-based platforms, which enable energy providers to monitor electricity usage more effectively, is another recent development. Utility providers can prevent electricity theft by using this technology to promptly identify suspect consumers and take appropriate action**.** Overall, prior studies show that, as compared to conventional methods, intelligent data analysis techniques can greatly enhance the detection of electricity theft. Better energy management, enhanced electricity usage monitoring, and lower financial losses for power distribution firms are all facilitated by these solutions.

3. Methodology

By examining user electricity consumption data and spotting unusual usage trends, the suggested solution detects electricity theft. Data gathering, preprocessing, model training, and fraud prediction are some of the steps in the methodology.

A. Data Collection

The system makes use of a dataset file (data.csv) that includes data on various consumers' electricity consumption. A consumer number (CONS_NO), fraud

label (FLAG), and daily electricity consumption values over time are all included in each record. The consumer's normal status or involvement in electricity theft is shown by the FLAG column.

B. Data Preprocessing

Using Python data processing libraries, the dataset is initially loaded. Unnecessary characteristics, including customer identification numbers, are eliminated during preprocessing as they don't aid in the prediction process. The remaining consumption values have been arranged and are ready for examination. To preserve the consumption chronology, the data columns that reflect dates are formatted appropriately and placed in a sequential manner.

C. Data Splitting

The dataset is separated into training and testing datasets following preprocessing. The prediction model is constructed using the training data, and the system's performance is assessed using the testing data. By taking this step, the model's ability to generalize to new and unseen data of electricity usage is ensured.

D. Model Training

The processed dataset is used by the system to train a classification model. In order to distinguish between typical and questionable usage patterns, the model uses historical data on electricity consumption. In order to detect potential fraudulent activity, the training procedure examines differences in patterns of electricity consumption.

E. Fraud Prediction

The learned model is saved and utilized for prediction when the training phase is over. Data on electricity consumption can be uploaded by users using the online interface. After processing the uploaded data, the algorithm determines if the consumption pattern suggests typical usage or potential electricity theft.

4. Results & Analysis

A. System Implementation

Python and the Django framework were used to create a web-based application for the suggested electricity theft detection system. The system evaluates the dataset's electricity consumption statistics to determine whether consumer behavior suggests typical electricity use or potential electricity theft. Through the interface of the online application, customers can upload data on electricity consumption and receive prediction results.

B. Consumption Pattern Analysis

In order to distinguish between typical and fraudulent consumption behavior, the system examines consumer electricity usage patterns. The technology may identify anomalous trends that can point to electricity theft by looking at usage values across several days.

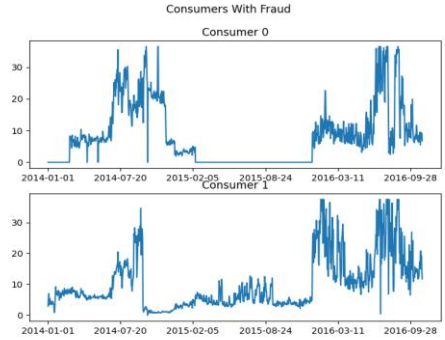


Figure 1: Consumers Without Fraud Analysis

The typical consumer's electricity use pattern is depicted in this figure. With just minor variations over time, the graph shows comparatively steady electricity use behavior.

C. Fraudulent Consumption Pattern

When compared to typical consumer behavior, fraudulent power consumption typically manifests as irregular patterns or abrupt changes in electricity usage.

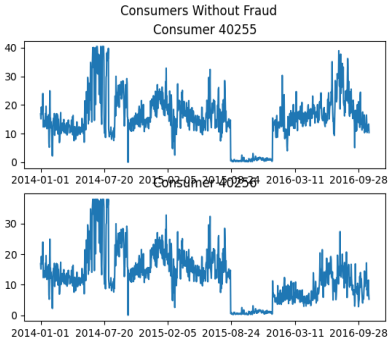
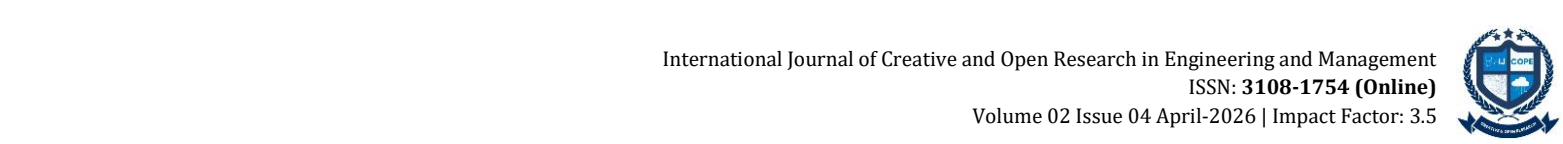


Figure 2: Consumers With Fraud Analysis



The trends of fraudulent consumers' power consumption are depicted in this figure. Unusual fluctuations in electricity consumption are highlighted in the graph, which aids in spotting questionable consumption patterns.

D. Correlation Analysis

Understanding the relationship between various factors related to energy use and spotting trends related to electricity fraud are made easier with correlation analysis.

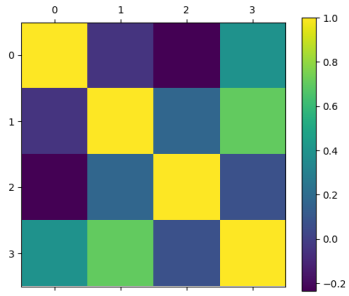


Figure 3: Correlation Analysis using Fraud Detection

A correlation heatmap illustrating the relationships between variables related to power use is shown in this graphic. The graphic aids in spotting trends that can point to unusual electricity use.

E. Weekly Consumption Comparison

Finding anomalous usage fluctuations is made easier by comparing electricity consumption over several time periods.

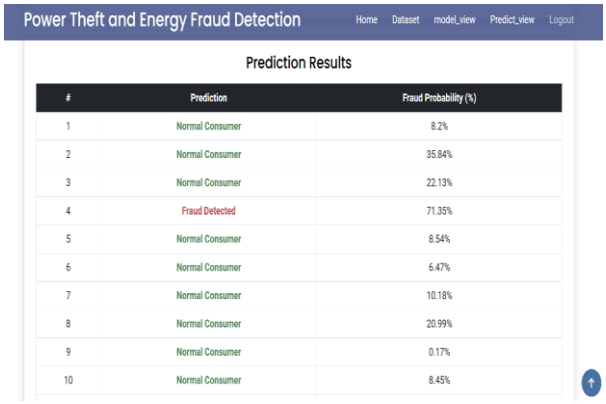


Figure 4: Four-Week Comparison of Electricity Use

This graph shows how much electricity was used during a four-week period. These comparisons aid in spotting anomalous usage patterns that can point to electricity theft.

F. Statistical Analysis

Data on power use can be statistically analyzed to reveal typical patterns of electricity use.

In order to better understand the distribution of power usage values, this figure provides statistical analysis of patterns of electricity consumption for typical consumers.

The findings show that anomalous usage patterns linked to electricity theft can be successfully identified by examining data on electricity consumption. Through the online interface, the developed system effectively evaluates records of electricity use, finds suspicious trends, and presents forecast results. This method makes it easier for power distribution firms to monitor electricity usage and detect electricity theft.

5. Conclusion

Power distribution firms suffer large financial losses because of electricity theft, which also compromises the stability of power systems. When dealing with massive amounts of data on electricity consumption, traditional detection techniques are frequently ineffective and time-consuming. In order to monitor consumer electricity usage patterns and spot suspect activity, a machine learning-based electricity theft detection system was created for this project. After analyzing data on electricity consumption, the system determines whether it is fraudulent. Python and the Django framework were used to develop the system as a web-based application that lets users upload data and get prediction results. The findings demonstrate that examining patterns of electricity consumption can assist power distribution firms in lowering electricity theft and enhancing energy monitoring by identifying unusual usage behavior.

References

[1] S. Hasan, M. Hossain, and M. Islam, "Electricity Theft Detection in Smart Grids Using Machine Learning Techniques," *IEEE Access*, vol. 10, pp. 45678–45690, 2022.

[2] R. Kumar, P. Singh, and A. Sharma, "Detection of Energy Theft in Smart Meter Data Using Data Mining Techniques," *International Journal of Electrical Power & Energy Systems*, vol. 141, pp. 108–116, 2022.

[3] L. Wang, Y. Zhang, and J. Li, "Smart Grid Electricity Theft Detection Based on Data Analytics and Machine Learning," *IEEE Transactions on Smart Grid*, vol. 14, no. 2, pp. 1250–1261, 2023.

[4] M. Ahmed and K. Raza, "Electricity Fraud Detection in Smart Grids Using Consumption Pattern Analysis," *IEEE Access*, vol. 11, pp. 23987–23998, 2023.

[5] P. Sharma and D. Gupta, "Intelligent Energy Theft Detection System Using Data Analysis Techniques," *IEEE Internet of Things Journal*, vol. 11, no. 4, pp. 6125–6134, 2024.

[6] T. Nguyen, H. Tran, and P. Le, "Smart Meter Data Analytics for Electricity Theft Detection Using Machine Learning," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 1, pp. 314–322, 2024.

[7] A. Verma and S. Patel, "Data-Driven Electricity Fraud Detection in Smart Grid Systems," *IEEE Systems Journal*, vol. 19, no. 1, pp. 410–418, 2025.