**Theft detection for benchmarking and machine learning based classification in a smart grid environment**

### A PROJECT REPORT

***submitted by***

**CB.EN.U4EEE23105 Aryan Jaljith**

**CB.EN.U4EEE23103 Anmol Govindarajapuram Krishnan**

**CB.EN.U4EEE23116 Karthik K**

**CB.EN.U4EEE23149 Namitha Madhu**

***in partial fulfillment of***

# BACHELOR OF TECHNOLOGY

# IN

# ELECTRICAL AND ELECTRONICS ENGINEERING



**AMRITA SCHOOL OF ENGINEERING, COIMBATORE**

# AMRITA VISHWA VIDYAPEETHAM

# COIMBATORE- 641112

# April 2025

**AMRITA SCHOOL OF ENGINEERING, COIMBATORE**

**AMRITA VISHWA VIDYAPEETHAM**

COIMBATORE CAMPUS



**BONAFIDE CERTIFICATE**

This is to certify that the project entitled “**Theft detection for benchmarking and machine learning based classification in a smart grid environment**” submitted by

**CB.EN.U4EEE23105 Aryan Jaljith**

**CB.EN.U4EEE23103 Anmol Govindarajapuram Krishnan**

**CB.EN.U4EEE23116 Karthik K**

**CB.EN.U4EEE23149 Namitha Madhu**

in partial fulfillment of the requirements for the subject **23EEE213 -INTRODUCTION TO MACHINE LEARNING** in **Bachelor of Technology** in **ELECTRICAL & ELECTRONICS ENGINEERING** is a bonafide record of the work carried out under my guidance and supervision at the Department of Electrical and Electronics Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore Campus.

**Mr. Nakkeeran M**

Assistant Professor (OC),

Department of Electrical and Electronics Engineering,

Amrita Vishwa Vidyapeetham, Coimbatore Campus,

India.

This project report was evaluated by us on (date)

INTERNAL EXAMINER EXTERNAL EXAMINER

**Contents**

[List of Tables iv](#_Toc194789235)

[List of Figures v](#_Toc194789236)

[List of Abbreviations vi](#_Toc194789237)

[Introduction 1](#_Toc194789238)

[Related Works 3](#_Toc194789239)

[Proposed Approach 5](#_Toc194789240)

[Process Flow Diagram, Proposed Framework 5](#_Toc194789241)

[5](#_Toc194789242)

[3.1 Dataset 6](#_Toc194789243)

[3.2 Approach 8](#_Toc194789244)

[Methodology 9](#_Toc194789245)

[4.1 Techniques 9](#_Toc194789246)

[4.2 Data Source 9](#_Toc194789247)

[4.3 Implementation Steps 9](#_Toc194789248)

[5.1 Algorithms 12](#_Toc194789251)

[5.2 Conclusion 14](#_Toc194789252)

[5.3 Comparison 15](#_Toc194789253)

[Result Discussion 21](#_Toc194789255)

[Conclusion and Scope for further Research 22](#_Toc194789256)

[References 24](#_Toc194789257)

# List of Tables

5.1 Output from Hyperparameter tuning for different models 13

5.1 Comparison table between different models 15

# List of Figures

3.1 Process Flow Diagram 5

3.2 Violin Plots  6

3.2 Correlation Heatmap  7

5.1 Comparison chart 3d bar graph for ensemble models 16

5.2 Comparison Bar plot for individual models16

5.3 Comparison Chart using radar chart17

5.4 Confusion Matrix - Optimized Stacking Classifier  
(XGBoost + LightGBM + RF + Logistic Regression) 18

5.5 Confusion Matrix - Optimized Stacking Classifier  
(XGBoost + LightGBM + RF + Logistic Regression) with OPTUNA 19

5.6 Confusion Matrix - Optimized Stacking Classifier  
(XGBoost + LightGBM + RF + Extra Trees) with OPTUNA 20

# List of Abbreviations

ML – Machine Learning

RF – Random Forest

ET – Extra Trees

XGBoost – Extreme Gradient Boosting

LightGBM – Light Gradient-Boosting Machine

DL – Deep Learning

SMOTE – Synthetic Minority Oversampling Technique

**Abstract:**

Electricity theft means using electrical energy without authorization, including through illegal connections, false meters, or by simply not paying for energy consumption, which results in non-technical losses for the utility companies. This problem is averse to the economy of a country through lowering the profits of the utilities, expanding the costs of operation, and consequently increasing prices for the actual customers. It also keeps off investors from the country because energy is a vital factor that supports economic growth, and its instability puts off investors. Furthermore, electricity theft puts a lot of stress on the power grid, which results in power leakage, voltage drops, and fast degradation of the infrastructure, which in turn results in low performance and high costs of maintenance. This paper also reveals that besides the economic and technical implications, it encourages corruption, undermines the institution of governance, and adversely affects the environment through encouraging the consumption of energy in a wrong manner, that is, through the emission of more greenhouse gases. This paper concludes that fighting electricity theft is vital for energy security, financial sustainability, and power distribution stability.

Machine learning (ML) is a contemporary approach of using data analysis to fight and reveal electricity theft. Other models used include Random Forest, XGboost, LightGBM, Logistic Regression and Extra Tress that analyse electricity consumption patterns to detect abnormalities and classify theft cases accurately. These methods assist in distinguishing between genuine and potentially fraudulent transactions, thus minimizing the number of false positives and enabling early identification of theft. In their application, pattern recognition and anomaly detection methodologies support the concept of ML in enhancing the security of the grid, protecting revenues, and improving the utilization of energy resources. The application of ML-based frameworks in power systems enhances the operational robustness and thereby guarantees the delivery of reliable and sustainable energy infrastructure.

**Keywords:** *Electricity theft, non-technical losses, Machine learning, Anomaly detection, Power grid security, Fraud detection.*

**Chapter 1**

# Introduction

Non-technical losses (NTLs), which can account for as much as 40% of total transmission and distribution losses, are a major problem for power systems and power transmission stations around the world. These losses cause significant financial burdens for utilities, network instability, and higher electricity rates for legitimate consumers. Traditional detection methods, such as manual inspections and rule-based techniques, have not been able to keep up with the increasingly sophisticated techniques used by electricity thieves, so it is critical to develop more effective and flexible detection strategies. Large volumes of power consumption data are now available due to the growing use of Smart Grids and Advanced Metering Infrastructure (AMI), which makes it possible to use Machine Learning (ML) and Deep Learning (DL) techniques to identify irregularities linked to electricity theft, but issues like unbalanced data, hostile attacks, and changing fraudulent practices call for the creation of more resilient and flexible models.

Theft of electricity results in considerable financial detriment and instability within the system, while current detection methodologies face challenges related to data imbalance, the evolution of fraudulent techniques, and elevated rates of false positives. This study presents a hybrid machine learning architecture aimed at improving detection accuracy, adaptability, and real-time efficiency in theft detection.

This project aims to improving electricity theft detection by using optimizing machine learning models such as Random Forest, XGBoost, LightGBM and logistic regression. The main focus in the process lies on decreasing the false positives and false negatives through refined feature selection, high-quality datasets, and advanced pattern recognition. Usage of adaptive learning mechanisms like these will make sure that the existing models evolve with emerging newer fraud patterns, while enhanced anomaly detection using statistical and machine learning methods will strengthen accuracy of the system. All these together help in ensure practicality of the application.

Conventional and modern theft detection methods, including rule-based approaches, statistical models, and machine learning techniques are assessed and analysed. The effectiveness of the overall is computed by comparing various ML algorithms, such as Decision Trees, Random Forest, XGBoost

Despite having new advancements in the field of machine learning and deep learning for detecting electricity theft, the research gap that stands out for the formulation of models are:

* Efficiency: Able to work effectively within the limited computational resources available in a developing nation.
* Adaptability: Able to achieve high precision and recall levels under data constraints and unique electricity consumption patterns of non-smart grid environments.

Existing complex models have always demanded very heavy computational power and huge detailed datasets, thereby rendering them useless in areas with a not-so-well-developed infrastructure for application. A challenge here is to do the reverse: design simplified techniques that smartly use resources to locate electricity theft in these situations.

Inclusion of lightweight model development by reducing number of layers to reduce computational complexity while maintaining accuracy heads in this direction of work. In addition to this, extraction of maximum information from limited data, ensuring efficient and meaningful input for machine learning models using smart feature engineering can be employed. By choosing the most pertinent features and removing unnecessary ones, dimensionality reduction approaches can assist reduce computational load. To ensure that theft detection models successfully detect fraudulent cases without being biased toward majority-class data, resampling techniques can also be used to resolve imbalanced datasets.

Thus, the objective would be to create a solution that defines the sweet spot between model complexity and applicability in real life-a working tool for utility providers in developing countries to effectively combat electricity theft. The major contributions of this project include a comprehensive model comparison, systematically evaluating boosting techniques such as XGBoost and LightGBM alongside Random Forest to identify the most effective methods for pattern recognition and anomaly detection in electricity theft. The project also explores ensemble models, to uncover intricate consumption patterns and enhance the detection of sophisticated theft techniques. To ensure real-world applicability, robustness testing will be conducted by introducing noise and data perturbations, evaluating model performance under realistic conditions. Finally, ensemble learning will be utilized to combine multiple models, enhancing detection accuracy, improving robustness, and reducing overfitting, thereby creating a more reliable and efficient electricity theft detection framework.

The report is structured in a chapter-wise basis with each chapter focusing on a different aspect of the project. Chapter 2 (Related Works) is a brief overview of existing studies in the field of the research topic, highlighting key findings, identifying gaps in each of them, and justifying the need for the new proposed approach whereas Chapter 3 (Proposed Approach) mainly focusses on the research framework, and consists of the process flow diagram, dataset details, and the chosen methodology. Chapter 4 (Methodology) explains the various strategies, techniques, algorithms, data sources, and evaluation metrics used to assess performance while Chapter 5 (Implementation Details & Results) tries to present the experimental setup, preprocessing steps, hyperparameter tuning, and results using tables and graphs, comparing different algorithms. Chapter 6 (Result Discussion) discusses the findings from the project, their implications, and realizes the study limitations present. Finally, Chapter 7 (Conclusion & Future Works) is a summary of the key contributions, highlights major takeaways, and suggests potential improvements and future research directions. This structured organization ensures a logical flow, making it easier to understand the research process and findings.

**Chapter 2**

# Related Works

This section reviews prior research on electricity theft detection in smart grids. Below is a summary of ten key papers:

Proposed SVM-based classification for detecting irregular load profiles. Focused on load profile analysis and data mining for suspect detection [9]. Used ANN models to detect diverse energy fraud activities. Focused on analysing consumer energy consumption behaviour using smart meter data [7]. Introduced AlexNet-based CNNs for detecting electricity theft with multi-dimensional feature extraction [10]. Provided an overview of AI-based methods like SVM, ANN, and hybrid techniques for detecting non-technical losses [5]. Implemented SVM-based classification to distinguish between normal and theft cases using consumption data from AEDC [8]. Focused on three-phase electricity theft using multi-dimensional feature extraction with deep learning [11]. Proposed a graph-based deep learning model for anomaly detection in smart grid consumption data [6]. Focused on applying deep learning for theft detection in non-smart grid settings, using time-series data [12]. Addressed the detection of adversarial attacks in smart grids using deep learning models [4]. Employed ANN techniques to detect both fraudulent reporting and unauthorized usage in smart grids [7].

For detecting electricity theft, the majority of studies use machine learning and deep learning models including SVM, CNN, ANN, and ensemble techniques. Finding Non-Technical Losses (NTLs) through fraudulent consumption habits utilizing smart meter data is a common area of study. Several methods, including SMOTE, class weighting, and hybrid feature engineering, are used to address the common problem of data imbalance. The suggested work prioritizes real-time theft detection using machine learning using stacking algorithms and class weights, in contrast to previous studies that frequently focus on adversarial attacks, cyber threats, or non-smart grid fraud. While some previous studies depend on CNNs (like AlexNet), ANNs, or different stacking approaches, the modelling approach combines machine learning and consumption pattern analysis with stacking techniques and class weighting. Additionally, the suggested study employs imputation methodologies while purposefully avoiding PCA, whereas traditional studies use PCA, clustering, or encoding approaches for data preprocessing. Many rely on a single framework, while this project use stacking for better accuracy. Insufficient Consumption Trend Analysis

Many models detect outliers but fail to analyze broader behavioral patterns. Inadequate Multi-Dimensional Feature Extraction Over-Reliance on Standard Data Balancing. Existing methods focus on SMOTE or PCA[12]; this project also explores class weight method. By combining powerful models like XGBoost, Random Forest, and LightGBM and utilizing their individual advantages, ensemble learning in particularly stacking improves theft detection accuracy. The method is more efficient when class weighting is used in place of SMOTE since it manages data imbalance and lowers computational expenses, the organized architecture of the dataset probably prevented deep learning models like CNNs and transformers from providing significant gains. These results confirm that machine learning-based ensembles offer a high-performing and computationally efficient method for detecting electricity theft in smart grid settings.

**Chapter 3**

# Proposed Approach

## Process Flow Diagram, Proposed Framework

## 

Figure 3.1 Process Flow Diagram

## 3.1 Dataset

The dataset contains 560,655 samples and 13 columns, including 10 numerical features associated with their energy consumption, 2 categorical labels, and an index. The dataset reflects the electricity usage pattern among various consumer types, with multi-class theft classification being the target variable consisting of 7 different categories. The numerical features are meant to describe consumption behaviour in terms of quantifiable aspects, such as facility energy use, fans, cooling, heating, internal lighting, and equipment, all painting different imageries of energy consumption. The categorical variables include the consumer class, which denotes the type of consumer, while the theft label specifies certain types of electricity theft.

Before any preprocessing, the dataset exhibited very imbalanced class distributions, with normal consumption cases hugely outnumbering theft ones. To balance the classes, class weighting was used as the main approach; that is, to weigh the influence of each class during model training to better generalize across all categories. An alternative approach called SMOTE (Synthetic Minority Over-sampling Technique) was tested: It generates synthetic theft cases to balance the dataset, this approach was later disregarded because it produced unrealistic feature values that would pose a threat to the validity of the model output predictions, while simultaneously increasing the complexity.

The violin plots given in figure 3.2 shows us the distribution of different energy consumption measures on an hourly basis with log transformed values, to keep up with skewed distributions. Through the given graphs, it is evident that the use of different components shows different patterns according to their usage. Facility electricity has a wide distribution with several peaks, suggesting unsteady consumption levels, whereas fans and interior lights have more organized distributions, thus showing cyclic usage patterns. Cooling electricity shows a clear peak, while heating electricity has a peak that is narrow and sharp at the lower values, suggesting that consumption is less. These insights highlight the different energy consumption trends, which can be useful for optimizing energy efficiency, detecting anomalies, and improving resource management.

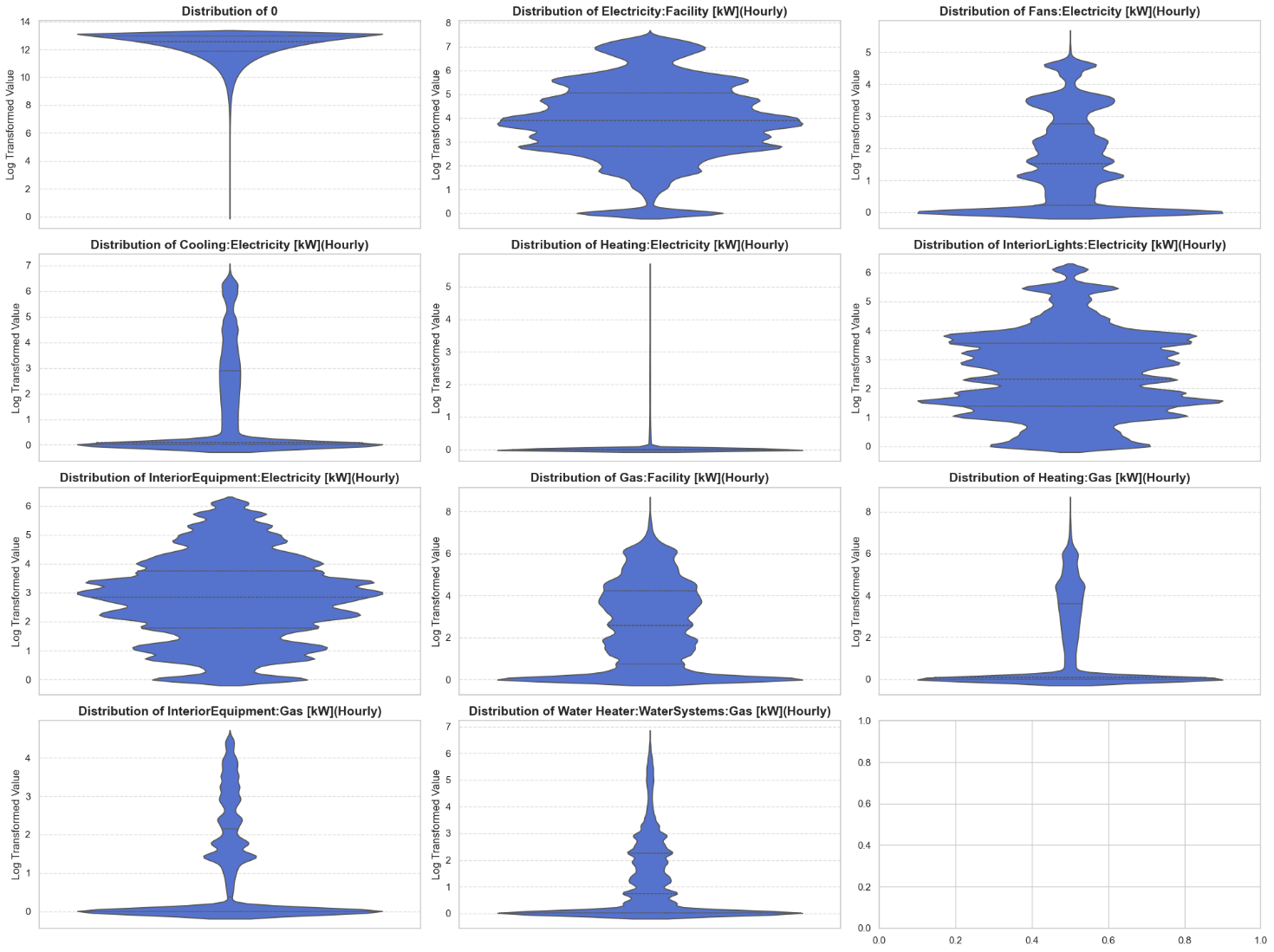


Figure 3.2 Violin Plots

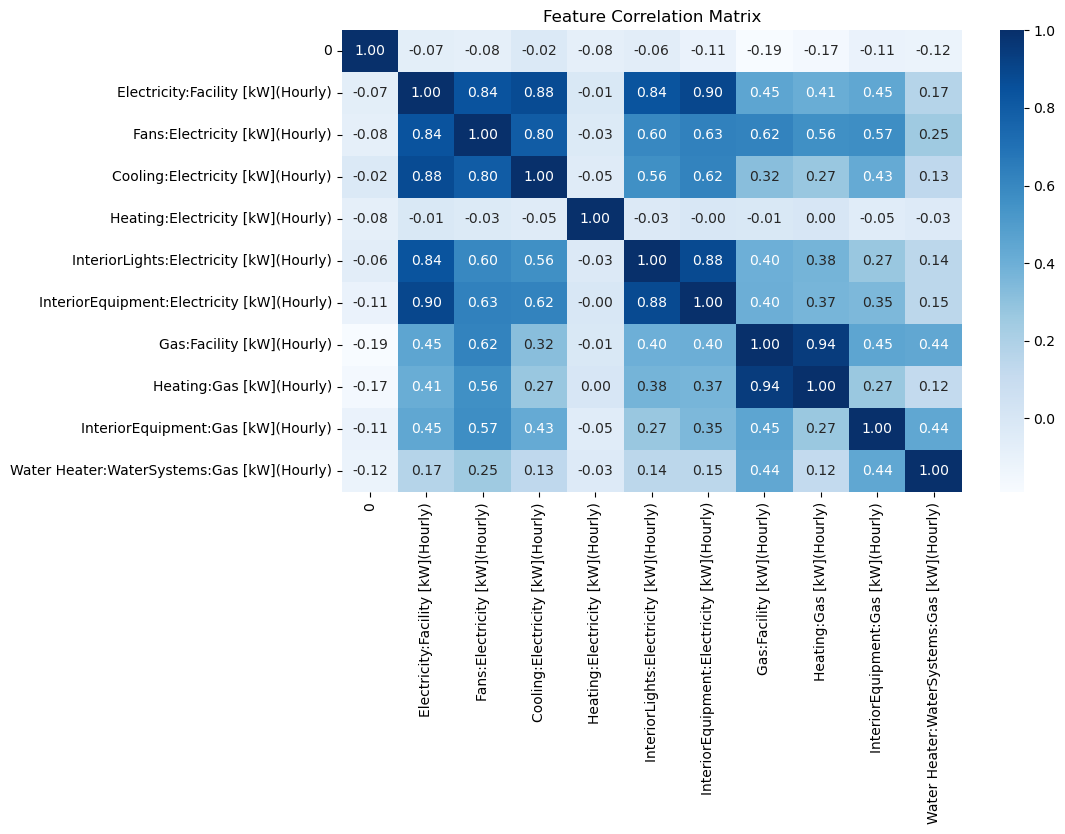


Figure 3.3 Correlation Heatmap

Figure 3.3 is a heatmap of the feature correlation matrix for various energy consumption metrices in a facility, in kWh. Strong Positive correlations include electricity used for fans and cooling (0.88) consumtion is highly related, which makes sense as these cooling system require fans for their functioning. InteriorEquipment: Electricity and Electricity: Facility (0.90) suggests that interior equipment usage significantly contributes to the facility’s overall electricity consumption. Gas:Facility and Heating:Gas (0.94) have a strong correlation, which indicates that gas consumption in the facility is mainly driven by heating.

Moderate correlations include Cooling:Electricity & Electricity:Facility (0.88) and Cooling:Electricity & Electricity:Facility (0.88) suggests that cooling and lighting&fans play a major role in the overall energy consumption.

Weak or Negative Correlations include Heating:Electricity & most other electricity-based variables (-0.01 to -0.08) which indicates that heating via electricity does not strongly relate to other electrical systems and Gas-based systems & electrical systems (-0.19 to 0.45), suggests that gas-powered heating and water systems operate somewhat independently from electrical usage.

## 3.2 Approach

In our upcoming paper, we have implemented machine-learning frameworks such as LightGBM and Stacking to compare their performance against Random Forest and Bagging. Additionally, we will explore SMOTE for imbalance handling, XGBoost and lightGBM for boosting and feature selection techniques.

Our aim over here will be to implement simplified techniques that can smartly detect electricity theft with the help of limited resources, rather than utilizing existing complex models that demand very heavy computational power and huge detailed datasets, thereby rendering them useless in areas with a not-so-well-developed infrastructure for application.

Other improvements could include the use of lightweight machine-learning models with fewer, lighter layers, smart feature engineering for to extract as much information possible in the given, limited set of data, dimensionality reduction techniques and resampling techniques to make our model more stable and efficient as possible.

Therefore, the objective here would be to create a solution that defines the sweet spot between model complexity and applicability in real life-a working tool for utility providers in developing countries to effectively combat electricity theft, thus making it a more economical, efficient and accurate model to use.

**Chapter 4**

# Methodology

## 4.1 Techniques

Random Forest, XGBoost, LightGBM, Extra Trees, and other machine learning models, along with ensemble learning techniques for pattern detection.

## 4.2 Data Source

Data Collection: [Theft detection in smart grid environment - Mendeley Data](https://data.mendeley.com/datasets/c3c7329tjj/1)

## 4.3 Implementation Steps

This project begins with **baseline reproduction**, where the methodology from the base paper is implemented to replicate its results, ensuring the correctness of the developed model by evaluating performance using key metrics such as accuracy, cohen’s kappa score, AUC, ROC and f1score and comparing them to the reported findings from the base paper. Moving forward, the focus shifts to machine learning-based models, where boosting techniques such as XGBoost and LightGBM are employed to enhance classification accuracy. To add to this, a comparative analysis machine learning models is conducted to determine their effectiveness in pattern recognition[6]. Fine-tuning model coefficients helps improve accuracy and reduce overfitting, while ensemble learning is applied to combine multiple models, enhancing overall detection capabilities. Additionally, robustness testing needs to be conducted by introducing noise and data perturbations to evaluate model performance under real-world conditions, ensuring reliability[25]. Extensive experimentation and results analysisis the final step in identifying and validating the effectiveness of the proposed approaches, refining them for practical deployment in electricity theft detection systems.

**4.4** **Evaluation Metrics**

Several metrics were used to evaluate the model. Because the dataset is imbalanced, accuracy alone wouldn’t be enough because a model that just predicts the majority class can be very accurate without detecting stolen cases. Accuracy gives a broad idea of correctness but can be misleading in imbalanced datasets as the model can achieve high accuracy by giving more weight to the majority class. F1-Score is good for theft detection as it balances recall and precision, considering both false positives and false negatives. Cohen’s Kappa is a more accurate indicator of a model’s performance as it measures the agreement between predicted and actual classes considering chance agreement. AUC-ROC evaluates the model’s ability to distinguish between normal and stolen power.

The higher the AUC, the better discrimination. Confusion Matrix gives information on specific misclassifications to determine if the model is favouring one class.

All these metrics ensure that the model is evaluated for parameters beyond accuracy so we can have a better understanding of its ability to detect theft while reducing false positives and false negatives.

**Chapter 5**

## Implementation Details & Results

## 

## Setup For the initial phases of model training, Google Colab was selected due to its provision of free GPU, which is advantageous for testing individual models prior to stacking. After determining the optimal stacking ensemble, Lightning AI was employed for the final model training, taking advantage of its capabilities for enhanced GPU training, simplified model scaling, and integrated features for checkpointing, logging, and early stopping, the dataset preprocessing due to the absence of missing values the necessity for imputation was not needed, the categorical variables was transformed through label encoding, additionally StandardScaler (Z-score normalization) was utilized. For class imbalance, class weights were selected for model training, demonstrating greater effectiveness than SMOTE, which, while increasing the dataset to over 2.3 million samples, proved to be computationally intensive, dataset was divided into 80% for training) and 20% for testing using a stratified approach to maintain the proportions of the classes. The dataset was initially explored and was checked for missing values which were found to be none, class distribution was examined and was found out to be highly imbalanced[23]. Preprocessing of data involved normalizing features for consistency and encoding theft labels for machine learning models, we proceeded to handle class imbalance initially, SMOTE was tested but produced unrealistic values, instead, class weighting was applied, which worked better while preserving data integrity[24], the individual traditional ML models like Random Forest, XGBoost, and LGBM were tested[6]. Hyperparameter tuning was performed using Optuna, enabling efficient exploration of parameter spaces for each model to enhance performance , after identifying the best-performing models and hyperparameter tuning, they were combined into an ensemble to improve overall classification performance, for fine tuning, the Individual models were trained using Google Colab, while the final stack was trained with Lightning AI for efficiency. The Accuracy, F1-score, AUC, and confusion matrices were analysed to ensure the model was learning effectively rather than just memorizing patterns.

## 5.1 Algorithms

Pseudocode for Hyperparameter Tuning using Optuna for Theft Detection Models

------------------------------------------------------------------------------

INPUT:

Train and Test Data (X\_train, y\_train), list of candidate models

OUTPUT:

Optimized hyperparameters for each model

BEGIN

1. X\_train, y\_train ← Preprocessed training data

2. models ← ['lgbm', 'xgboost', 'logreg', 'rf', 'extratrees']

3. FOR each model in models DO

 a. Define an objective function to suggest hyperparameters using trial.suggest\_ methods

 b. Initialize the model with the suggested parameters

 c. Train the model on X\_train, y\_train

 d. Predict on X\_test

 e. Return evaluation metric (e.g., accuracy\_score)

4. Run study.optimize(objective, n\_trials=30) for each model

5. Store the best trial parameters from each study

6. Output the best parameters for all models

END

Output From Hyperparameter Tuning:

Table 5.1 Output From Hyperparameter tuning

|  |  |  |
| --- | --- | --- |
| **Model** | **Parameter** | **Value** |
| **LGBM** | boosting\_type | gbdt |
|  | learning\_rate | 0.0986656417 |
|  | num\_leaves | 190 |
|  | max\_depth | 11 |
|  | min\_data\_in\_leaf | 12 |
|  | max\_bin | 356 |
|  | feature\_fraction | 0.5460055791 |
| **XGBoost** | learning\_rate | 0.0997219115 |
|  | max\_depth | 11 |
|  | min\_child\_weight | 4 |
|  | subsample | 0.8696419675 |
|  | colsample\_bytree | 0.6797301712 |
|  | n\_estimators | 126 |
| **LogReg** | solver | lbfgs |
|  | max\_iter | 222 |
|  | C | 13812.3070583 |
| **RandomForest** | n\_estimators | 184 |
|  | max\_depth | 11 |
|  | min\_samples\_split | 18 |
|  | min\_samples\_leaf | 4 |
|  | max\_features | sqrt |
| **ExtraTrees** | n\_estimators | 121 |
|  | max\_depth | 12 |
|  | min\_samples\_split | 3 |
|  | min\_samples\_leaf | 9 |
|  | max\_features | None |

## 5.2 Conclusion

The optimal performance measures were achieved by the Optimized Stacking Classifier (XGBoost + LightGBM + RF + Logistic Regression) without Optuna tuning with the highest accuracy (89.02%), highest F1-score (86.62%), and highest AUC-ROC (94.74%), and it was the highest-performing model. Surprisingly, Optuna tuning reduced performance slightly for both stacking versions, meaning the untuned version already contained maximum synergy between the base learners. While XGBoost with Intelligent SMOTE had previously shown comparable performance, it was nevertheless computationally expensive due to the creation of synthetic data. Random Forest and KNN were still behind in comparison, with KNN performing poorly in high-dimensional space.

In terms of data balancing, SMOTE vs. class weights testing confirmed class weighting superiority. SMOTE inflated the dataset size from ~500K to ~2.3M, which compromised training time and exposed it to overfitting. Class weights, on the other hand, preserved the integrity of original data, improved scalability, and steered clear of overfitting, making them more suitable for real-world applications. The effectiveness of ensemble learning was apparent. The combination of a few good learners (XGBoost, LightGBM, RF, and Logistic Regression) improved generalization and robustness. Moreover, stacking allowed a meta-learner to further adjust decision boundaries to improved performance for multiclass classification and alleviate class imbalance issues through class weights instead of resampling techniques. In short, the Optimized Stacking Classifier with the incorporation of XGBoost, LightGBM, Random Forest (RF), and Logistic Regression performed better and accurately even without the optimization process utilizing Optuna. This study identifies that ensemble strategies with class weighting are better compared to oversampling techniques in aspects of accuracy, computational cost, and usability in real-world contexts.

## 5.3 Comparison

Table 5.2 Comparison table between different models.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Accuracy** | **F1-Score** | **Kappa** | **AUC-ROC** |
| **Optimized Stacking Classifier(XGBoost + LightGBM + RF + Logistic Regression)** | **0.8902** | **0.86621** | **0.81641** | **0.94739** |
| **Optimized Stacking Classifier(XGBoost + LightGBM + RF + Logistic Regression) with OPTUNA** | **0.8646** | **0.8365** | **0.7736** | **0.9355** |
| **Optimized Stacking Classifier(XGBoost + LightGBM + RF + Extra Trees) with OPTUNA** | 0.8642 | 0.8362 | 0.7729 | 0.9352 |
| XGBoost + OPTUNA | 0.8107 | 0.8162 | 0.7023 | 0.929 |
| LightGBM + OPTUNA | 0.8505 | 0.8168 | 0.7483 | 0.9276 |
| Logistic Regression + OPTUNA | 0.2286 | 0.2458 | 0.1023 | 0.6836 |
| ExtraTrees + OPTUNA | 0.784 | 0.7647 | 0.6536 | 0.907 |
| RF + OPTUNA | 0.808 | 0.7857 | 0.6868 | 0.9104 |
| KNN(Base Paper) | 0.8451 | 0.8145 | 0.7403 | 0.9239 |
| DT(Base Paper) | 0.808 | 0.8178 | 0.6964 | 0.9091 |
| RF(Base Paper) | 0.7383 | 0.7699 | 0.6058 | 0.8642 |

Figure 5.1 Comparison chart 3d bar graph

All three stacking models, including those optimized with Optuna and ExtraTrees, show similarly high performance across metrics, with AUC-ROC consistently near 1. Minor variations are observed between models, but overall optimization and extra classifiers slightly enhance robustness without major performance drops.

Figure 5.2 Comparison Bar plot between individual models

Individually optimized models (especially LightGBM + OPTUNA and XGBoost + OPTUNA) outperform the base models significantly across all metrics. Logistic Regression lags behind, highlighting the advantage of ensemble methods and tree-based algorithms after optimization.

## 

Figure 5.3 Comparison Chart using radar chart

The radar chart shows that the stacking classifier without Optuna slightly outperforms the Optuna-tuned versions across most metrics. However, performance differences are minimal, indicating that both optimization strategies (with Logistic Regression and ExtraTrees) maintain high model quality.

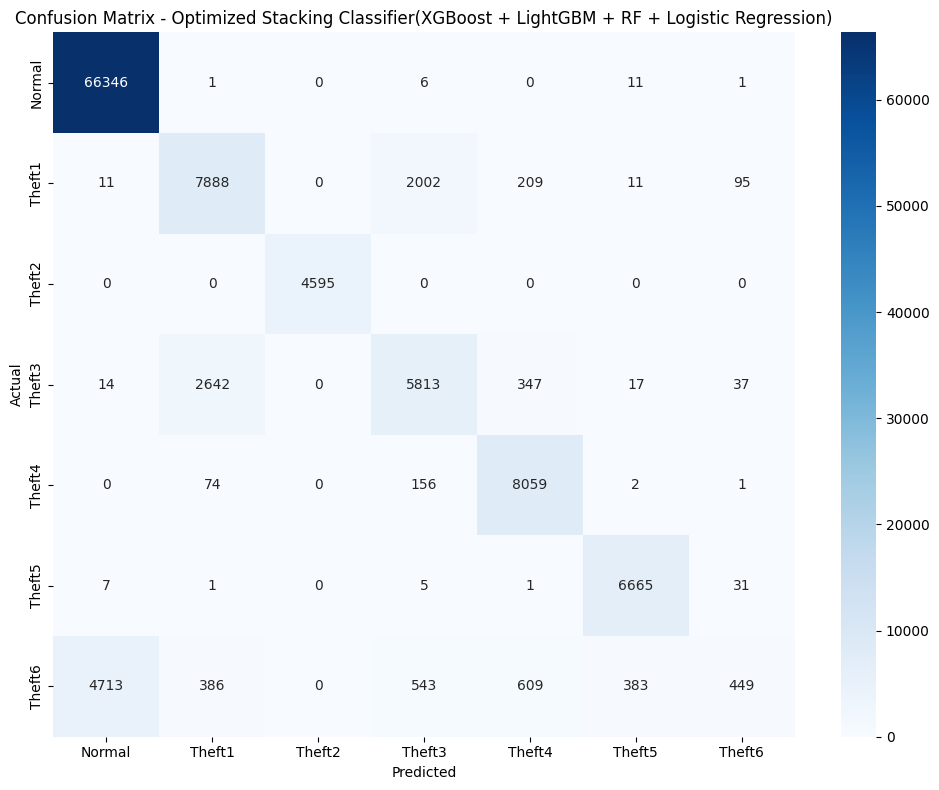


Figure 5.4 Confusion Matrix - Optimized Stacking Classifier(XGBoost + LightGBM + RF + Logistic Regression)

The model classifies 'Normal' and most 'Theft' classes well, but Theft6 has considerable misclassification into 'Normal'. Significant confusion between Theft1 and Theft3 is observed, suggesting room for improvement.

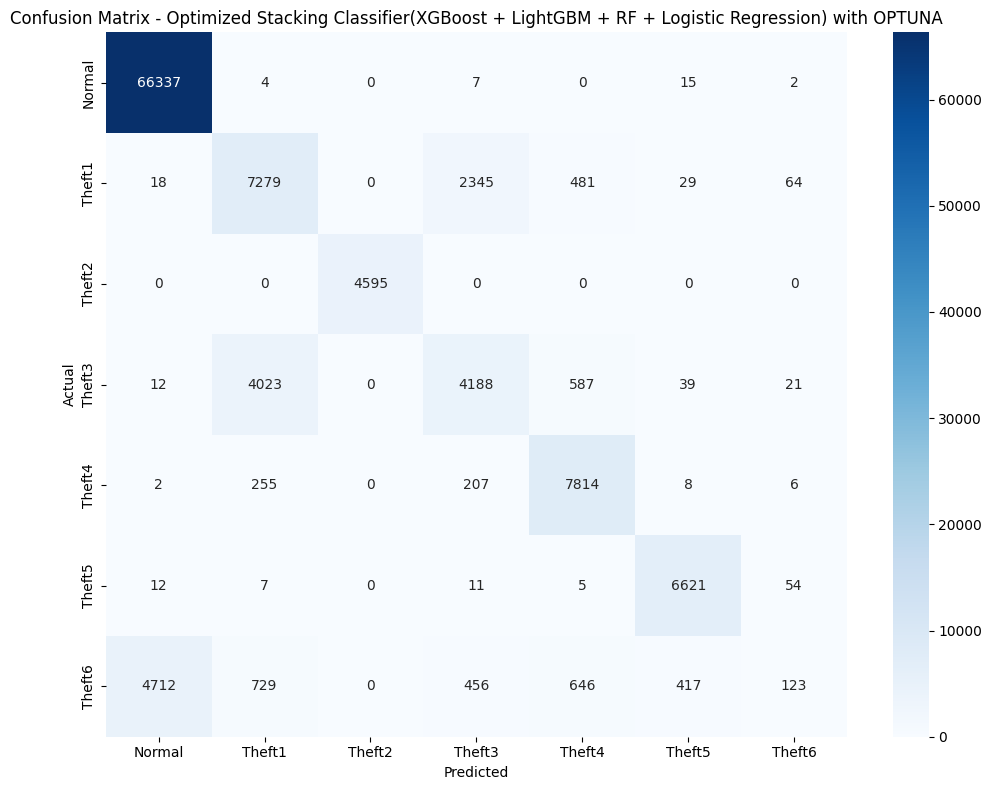


Figure 5.5 Confusion Matrix - Optimized Stacking Classifier(XGBoost + LightGBM + RF + Logistic Regression) with Optuna

OPTUNA tuning improves some correct theft classifications but increases confusion between Theft1 and Theft3. Misclassification for Theft6 into 'Normal' remains substantial and needs attention.

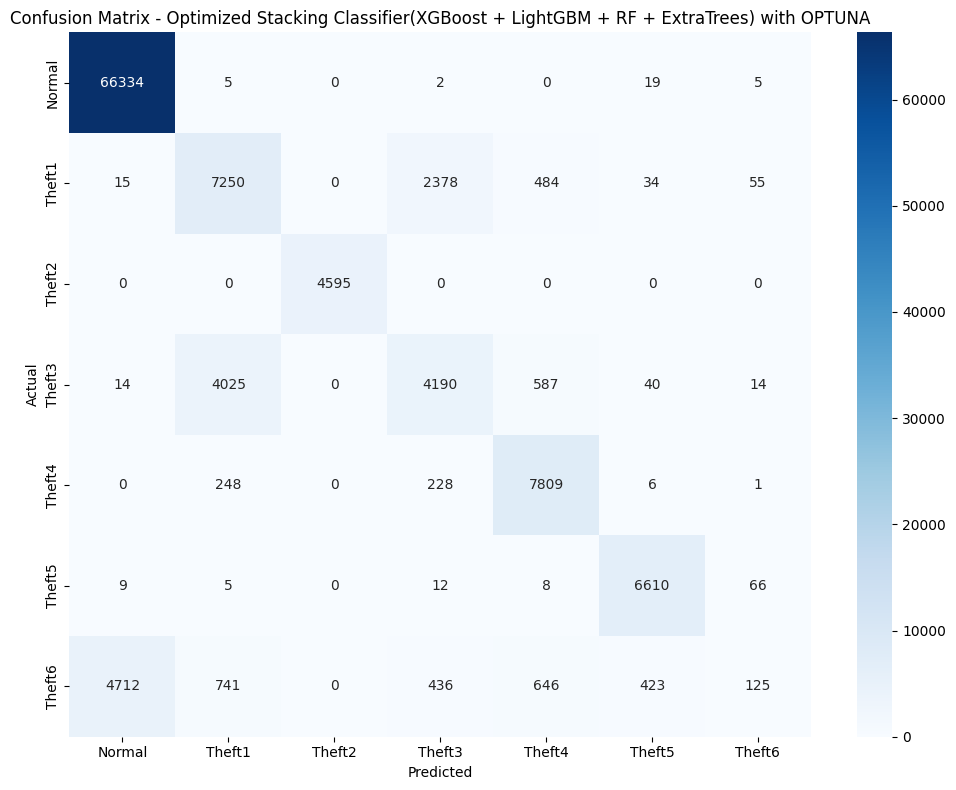


Figure 5.6 Confusion Matrix - Optimized Stacking Classifier(XGBoost + LightGBM + RF + Extra Trees­) with Optuna

Performance is similar to the Logistic Regression case but slightly better in minimizing false positives across 'Normal' and Theft classes. Theft6 misclassification persists, but the confusion between other theft classes slightly reduces.

**Chapter 6**

# Result Discussion

The results demonstrate that the Optimized Stacking Classifier (combining XGBoost, Random Forest, Logistic regression ,and LightGBM) achieved superior performance in detecting electricity theft, with an accuracy of 89.02%, an F1-score of 86.63%, and an AUC-ROC of 94.74%. The use of class weights instead of SMOTE for balancing class distribution proved effective in preserving data integrity and improving real-world applicability. Traditional machine learning models, particularly ensemble methods, gave significantly better results compared to individual models, demonstrating superior performance in terms of accuracy, robustness, and generalization

Ensemble learning (stacking) significantly enhances detection accuracy by combining the strengths of multiple models. Class weighting outperforms SMOTE by maintaining data realism and reducing noise in training. Feature selection and standardization were critical in optimizing model inputs and improving classifier performance. From the practical implications, we can conclude that a few utility companies can adopt the proposed ensemble-based system to achieve real-time, accurate detection of electricity theft. Class weighting offers a scalable solution to handle imbalanced data in real-world smart grid systems without the computational burden introduced by techniques like SMOTE. The framework can be adapted to other anomaly detection tasks in smart grids, such as load forecasting and predictive maintenance. The theoretical implications of the paper demonstrate the effectiveness of hybrid ensemble approaches over traditional techniques for data in smart grid environments. It provides evidence supporting the integration of multiple machine learning models to improve classification in highly imbalanced, multi-class datasets. Furthermore, it highlights the limitations of SMOTE in certain contexts, suggesting the need for alternative data balancing methods in similar applications.

Although ensemble learning improved performance, training complex models like XGBoost and LightGBM required significant computational resources, particularly when hyperparameter tuning was performed which is a computational constraint. While SMOTE was tested, it produced unrealistic synthetic data, which degraded model performance and increased training time. The dataset, although comprehensive, may not capture all types of electricity theft scenarios or real-world anomalies found in diverse geographic or infrastructural conditions. This also shows scalability concerns, while effective in a controlled environment, the scalability of the proposed system in live smart grid infrastructures needs further validation. Class classification proves to be difficult due to the presence of imbalanced data.

**Chapter 7**

# Conclusion and Scope for further Research

This research effectively illustrated how machine learning models in particular, ensemble learning can be used to detect electrical theft. When XGBoost, Random Forest, Logistic regression ,and LightGBM were combined, the Optimized Stacking Classifier produced the best performance metrics and the highest accuracy (89.02%) when compared to the separate models. In addressing imbalanced data, the study demonstrated the benefits of class weighting over SMOTE, guaranteeing improved generalization and practical application.

This research offers a scalable and effective theft detection framework for smart grids by utilizing feature selection, model stacking, and rigorous assessment methodologies. It was discovered that ensemble methods offered a better trade-off between accuracy and processing economy. In order to evaluate the model's performance in real-time smart grid scenarios, future work for this project will involve putting it into practice. Improving the framework's flexibility and scalability will enable it to handle various grid configurations and changing theft trends, furthermore, investigating deep learning approaches may enhance sequential anomaly detection even more. Enhancements in feature engineering will also be essential for improving the accuracy of theft classification. A more reliable and secure energy distribution system might be ensured by incorporating cybersecurity safeguards into the model, which could also aid in detecting hostile assaults and reducing cybersecurity risks in smart grids. The primary models that performed well were gradient boosting models such as LightGBM, XGBoost and CATBoost[17]. Additionally, non-gradient boosting models such as RF and Extra Trees for multi-class classification performed well. This indicates that gradient boosting models learn the dataset better.

The use of Optimized Stack Classifier by combining XGBoost, RF, Logistic Regression and LightGBM produced better results than alternative methods that relied on majorly using, KNN, DT and didn’t apply stacking methods. Usage of class weights was done rather that synthetic sampling here. A combined effect of these helped reduce the model training time and improved the overall performance of the models.

The usage of class weights over synthetic sampling proved to be less computationally straining. This helped keep down the execution time of the developed model. The features due to “Gas Facility” when removed significantly improved the classification as observed. The main challenge faced was dataset imbalance and quality of dataset. The quality of dataset majorly impacted the classification of the thefts. Even after applying class weights the dataset proved to be still imbalanced which affected the results. The reason for the removal of features due to the “Gas Facility” is for more specificity towards the domain of EEE, the main objective of this paper is to classify electricity theft whereas the dataset provides two major features one electricity production another gas production, the same features co-exist for both the production types, due to this a lack of quality is observed which enhances misclassification in thefts.

# References

1. L. Ge, J. Li, T. Du, and L. Hou, "Double-layer stacking optimization for electricity theft detection considering data incompleteness and intra-class imbalance," *Int. J. Electr. Power Energy Syst.*, vol. 165, p. 110461, Apr. 2025, doi: 10.1016/j.ijepes.2025.110461.
2. M. Hashim, L. Khan, N. Javaid, Z. Ullah, and A. Javed, "Stacked machine learning models for non-technical loss detection in smart grid: A comparative analysis," *Energy Rep.*, vol. 12, pp. 1235–1253, Dec. 2024, doi: 10.1016/j.egyr.2024.06.015.
3. Y.-C. Tsao, D. Rahmalia, and J.-C. Lu, "Machine-learning techniques for enhancing electricity theft detection considering transformer reliability and supply interruptions," *Energy Rep.*, vol. 12, pp. 3048–3064, Dec. 2024, doi: 10.1016/j.egyr.2024.08.068.
4. S. Ness, "Adversarial Attack Detection in Smart Grids Using Deep Learning Architectures," *IEEE Access*, vol. 13, pp. 16314–16323, 2025, doi: 10.1109/access.2024.3523409.
5. R. Yadav, M. Yadav, Ranvijay, Y. Sawle, W. Viriyasitavat, and A. Shankar, "AI Techniques in Detection of NTLs: A Comprehensive Review," *Arch. Comput. Methods Eng.*, vol. 31, no. 8, pp. 4879–4892, Jun. 2024, doi: 10.1007/s11831-024-10137-z.
6. S. I. Evangeline, S. Darwin, P. P. Anandkumar, and M. C. Thanu, "Anomaly detection in smart grid using a trace-based graph deep learning model," *Electr. Eng.*, vol. 106, no. 5, pp. 5851–5867, Mar. 2024, doi: 10.1007/s00202-024-02327-6.
7. V. Ford, A. Siraj, and W. Eberle, "Smart grid energy fraud detection using artificial neural networks," in *Proc. IEEE Symp. Comput. Intell. Appl. Smart Grid (CIASG)*, Dec. 2014, pp. 1–6, doi: 10.1109/ciasg.2014.7011557.
8. A. H. T. Alfarra, B. A. Attia, and C. S. M. El Safty, "Nontechnical Loss Detection for Metered Customers in Alexandria Electricity Distribution Company Using Support Vector Machine," *Renew. Energy Power Qual. J.*, vol. 16, no. 4, Jan. 2024, doi: 10.24084/repqj16.353.
9. J. Nagi, A. M. Mohammad, K. S. Yap, S. K. Tiong, and S. K. Ahmed, "Non-Technical Loss analysis for detection of electricity theft using support vector machines," in *Proc. IEEE Int. Power Energy Conf. (PECon)*, Dec. 2008, pp. 907–912, doi: 10.1109/pecon.2008.4762604.
10. N. Khan et al., "A novel deep learning technique to detect electricity theft in smart grids using AlexNet," *IET Renew. Power Gener.*, vol. 18, no. 6, pp. 941–958, Feb. 2024, doi: 10.1049/rpg2.12846.
11. W. Bai et al., "Detection Method for Three-Phase Electricity Theft Based on Multi-Dimensional Feature Extraction," *Sensors*, vol. 24, no. 18, p. 6057, Sep. 2024, doi: 10.3390/s24186057.
12. S.M. Saqib et al., "Deep learning-based electricity theft prediction in non-smart grid environments,"*Heliyon*,vol.10,no.15, p. e35167, Aug. 2024, doi: 10.1016/j.heliyon.2024.e35167.
13. S. Zidi, A. Mihoub, S. M. Qaisar, M. Krichen, and Q. Abu Al-Haija, "Theft detection dataset for benchmarking and machine learning based classification in a smart grid environment," *J. King Saud Univ. - Comput.Inf.Sci.*,vol.35no.1,pp. 13–25, Jan. 2023, doi: 10.1016/j.jksuci.2022.05.007.
14. M. Hashim, L. Khan, N. Javaid, Z. Ullah, and A. Javed, "Stacked machine learning models for non-technical loss detection in smart grid: A comparative analysis," *Energy Rep.*, vol. 12, pp. 1235–1253, Dec. 2024, doi: 10.1016/j.egyr.2024.06.015.
15. Y.-C. Tsao, D. Rahmalia, and J.-C. Lu, "Machine-learning techniques for enhancing electricity theft detection considering transformer reliability and supply interruptions," *Energy Rep.*, vol. 12, pp. 3048–3064, Dec. 2024, doi: 10.1016/j.egyr.2024.08.068.
16. S. Ness, "Adversarial Attack Detection in Smart Grids Using Deep Learning Architectures," *IEEE Access*, vol. 13, pp. 16314–16323, 2025, doi: 10.1109/access.2024.3523409.
17. R. Yadav, M. Yadav, Ranvijay, Y. Sawle, W. Viriyasitavat, and A. Shankar, "AI Techniques in Detection of NTLs: A Comprehensive Review," *Arch. Comput. Methods Eng.*, vol. 31, no. 8, pp. 4879–4892, Jun. 2024, doi: 10.1007/s11831-024-10137-z.
18. S. I. Evangeline, S. Darwin, P. P. Anandkumar, and M. C. Thanu, "Anomaly detection in smart grid using a trace-based graph deep learning model," *Electr. Eng.*, vol. 106, no. 5, pp. 5851–5867, Mar. 2024, doi: 10.1007/s00202-024-02327-6.
19. V. Ford, A. Siraj, and W. Eberle, "Smart grid energy fraud detection using artificial neural networks," in *Proc. IEEE Symp. Comput. Intell. Appl. Smart Grid (CIASG)*, Dec. 2014, pp. 1–6, doi: 10.1109/ciasg.2014.7011557.
20. A. H. T. Alfarra, B. A. Attia, and C. S. M. El Safty, "Nontechnical Loss Detection for Metered Customers in Alexandria Electricity Distribution Company Using Support Vector Machine," *Renew. Energy Power Qual. J.*, vol. 16, no. 4, Jan. 2024, doi: 10.24084/repqj16.353.
21. J. Nagi, A. M. Mohammad, K. S. Yap, S. K. Tiong, and S. K. Ahmed, "Non-Technical Loss analysis for detection of electricity theft using support vector machines," in *Proc. IEEE Int. Power Energy Conf. (PECon)*, Dec. 2008, pp. 907–912, doi: 10.1109/pecon.2008.4762604.
22. N. Khan et al., "A novel deep learning technique to detect electricity theft in smart grids using AlexNet," *IET Renew. Power Gener.*, vol. 18, no. 6, pp. 941–958, Feb. 2024, doi: 10.1049/rpg2.12846.
23. W. Bai et al., "Detection Method for Three-Phase Electricity Theft Based on Multi-Dimensional Feature Extraction," *Sensors*, vol. 24, no. 18, p. 6057, Sep. 2024, doi: 10.3390/s24186057.
24. S. M. Saqib et al., "Deep learning-based electricity theft prediction in non-smart grid environments," *Heliyon*, vol. 10, no. 15, p. e35167, Aug. 2024, doi: 10.1016/j.heliyon.2024.e35167.
25. S. Zidi, A. Mihoub, S. M. Qaisar, M. Krichen, and Q. Abu Al-Haija, "Theft detection dataset for benchmarking and machine learning based classification in a smart grid environment," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 35, no. 1, pp. 13–25, Jan. 2023, doi: 10.1016/j.jksuci.2022.05.007.
26. S. C. Yip, W. N. Tan, C. K. Tan, M. T. Gan, and K. S. Wong, "An anomaly detection framework for identifying energy theft and defective meters in smart grids," *Int. J. Electr. Power Energy Syst.*, vol. 101, pp. 189–203, 2018, doi: 10.1016/j.ijepes.2018.01.041.
27. S. Li et al., "Electricity theft detection in power grids with machine learning and random forests," *J. Electr. Comput. Eng.*, vol. 2019, pp. 1–10, 2019, doi: 10.1155/2019/4136753.
28. L. Ge, J. Li, T. Du, and L. Hou, "Double-layer stacking optimization for electricity theft detection considering data incompleteness and intra-class imbalance," *Int. J. Electr. Power Energy Syst.*, vol. 165, p. 110461, Apr. 2025, doi: 10.1016/j.ijepes.2025.110461.