International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 2 (2026)

ijerst.editor@gmail.com

editor@ijerst.com

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com
Vol. 22, No. 2, 2026

Peer Reviewed, Referred & Indexed Journal

A Hybrid Deep Learning Framework for Detecting Electricity Theft Cyber-Attacks in Renewable Distributed Generation Systems

MUDDE MEGHANA

PG Scholar. Department of MCA, DNR College, Bhimavaram, Andhra Pradesh

A. Durga Devi

(Assistant Professor), Master of Computer Applications, DNR College, Bhimavaram, Andhra Pradesh

ABSTRACT

The increasing integration of renewable distributed generation (RDG) into modern smart grids has significantly improved energy efficiency and sustainability. However, this transition has also introduced new vulnerabilities, particularly in the form of cyber-attacks and electricity theft. Unauthorized manipulation of smart meter data and communication channels can lead to substantial economic losses, compromised grid stability, and reduced reliability of energy distribution systems. Traditional detection mechanisms are often inadequate due to their inability to handle large-scale, dynamic, and heterogeneous data generated by smart grids.This research proposes a hybrid deep learning-based framework for detecting electricity theft and cyber-attacks in renewable distributed generation environments. The system integrates multiple machine learning and deep learning techniques, including Feed Forward Neural Networks (DNN), Convolutional Neural Networks (CNN), Gated Recurrent Units (GRU), and an ensemble model combining CNN with Random Forest classification. The objective is to leverage the strengths of each model to improve detection accuracy and robustness.The proposed system processes smart meter datasets containing consumer energy usage patterns. Data preprocessing techniques such as handling missing values, label encoding, normalization, and feature selection are applied to ensure data quality and consistency. Temporal and spatial patterns in electricity consumption are captured using GRU and CNN architectures, respectively. The GRU model is particularly effective in identifying sequential anomalies in time-series consumption data, while the CNN model extracts spatial correlations among features. The hybrid CNN-Random Forest model further enhances classification performance by utilizing deep feature representations for traditional ensemble learning.Performance evaluation is conducted using standard metrics such as accuracy, precision, recall, and F1-score, along with Receiver Operating Characteristic (ROC) analysis. Experimental results demonstrate that the hybrid CNN-Random Forest model outperforms standalone models, achieving higher detection rates and reduced false positives. The system also provides a graphical user interface (GUI) for ease of interaction, enabling users to upload datasets, train models, visualize results, and predict electricity theft instances.The proposed framework offers a scalable and efficient

solution for real-time electricity theft detection in smart grids. By incorporating advanced deep learning techniques and hybrid modeling strategies, the system enhances the security and reliability of renewable energy distribution networks. This research contributes to the development of intelligent energy management systems capable of mitigating cyber threats and ensuring sustainable energy utilization.

Keywords: Electricity Theft Detection, Smart Grid Security, Renewable Distributed Generation, Deep Learning, CNN, GRU, Random Forest, Cyber-Attack Detection, Smart Meter Analytics, Energy Informatics

I. INTRODUCTION

The modernization of power systems through the integration of smart grid technologies and renewable distributed generation has revolutionized energy distribution and consumption. Smart grids enable bidirectional communication between utilities and consumers, allowing for efficient energy management, demand response, and real-time monitoring. However, this increased connectivity also exposes the system to various cyber threats, including electricity theft, data tampering, and unauthorized access.Electricity theft is a critical issue that affects both economic stability and operational efficiency of power systems. It can occur through physical tampering with meters or through cyber-attacks that manipulate consumption data. In renewable distributed generation systems, where energy is generated and consumed locally, detecting such anomalies becomes even more challenging due to decentralized architecture and diverse data sources.Traditional methods for detecting electricity theft rely on rule-based systems and statistical analysis. These approaches are limited in their ability to capture complex consumption patterns and adapt to evolving attack strategies. Moreover, they often fail to process large volumes of data generated by smart meters in real time.Recent advancements in artificial intelligence, particularly deep learning, have provided powerful tools for analyzing complex datasets. Deep learning models such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have demonstrated exceptional performance in pattern recognition and anomaly detection tasks. CNNs are effective in extracting spatial features, while RNN-based models such as GRU are well-suited for capturing temporal dependencies in sequential data.This research aims to develop an intelligent electricity theft detection system using a combination of deep learning models and ensemble techniques. The proposed system leverages DNN, CNN, and GRU architectures to analyze smart meter data and identify anomalous patterns indicative of theft or cyber-attacks. Additionally, a hybrid model combining CNN with Random Forest is introduced to enhance classification accuracy.The system is implemented with a user-friendly graphical interface that allows users to perform data preprocessing, model training, evaluation, and prediction. The integration of multiple models enables comprehensive analysis of both spatial and temporal characteristics of electricity consumption.The primary contributions of this research include the development of a hybrid deep learning framework, comparative analysis of multiple models, and the implementation of an efficient and scalable detection system. The results demonstrate that the proposed approach significantly improves detection accuracy and

provides a reliable solution for securing renewable distributed generation systems against cyber threats.

II. LITERATURE SURVEY (WITH EXISTING METHODS)

Electricity theft detection has been an active area of research due to its significant economic and operational impact on power systems. Early approaches primarily relied on statistical and rule-based methods, which involved analyzing consumption patterns and identifying deviations from expected behavior. Techniques such as load profiling and threshold-based anomaly detection were commonly used. However, these methods lacked adaptability and failed to detect sophisticated cyber-attacks.Machine learning techniques have been widely adopted to improve detection performance. Algorithms such as Support Vector Machines (SVM), Decision Trees, and Random Forests have been used to classify normal and fraudulent consumption patterns. These models demonstrated improved accuracy compared to traditional methods but were limited by their dependence on handcrafted features and inability to capture complex temporal relationships.Deep learning approaches have recently gained prominence due to their ability to automatically learn feature representations from raw data. Convolutional Neural Networks (CNNs) have been applied to extract spatial features from electricity consumption datasets. Studies have shown that CNN-based models can effectively identify patterns associated with electricity theft. However, CNNs alone are not sufficient for capturing temporal dependencies in sequential data.Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU), have been used to model time-series data. GRU-based models have shown promising results in detecting anomalies in electricity consumption due to their ability to retain relevant information over time while being computationally efficient.Hybrid models combining deep learning and traditional machine learning techniques have also been explored. For example, CNN features have been used as input to classifiers such as Random Forest and SVM, resulting in improved classification performance. These hybrid approaches leverage the feature extraction capability of deep learning models and the decision-making strength of traditional classifiers.Despite these advancements, several challenges remain. Many existing systems lack scalability and are not suitable for real-time deployment. Additionally, issues such as class imbalance, noisy data, and high dimensionality affect model performance.The proposed research addresses these challenges by integrating multiple deep learning models and incorporating class balancing techniques. The use of GRU, CNN, and DNN ensures comprehensive feature extraction, while the hybrid CNN-Random Forest model enhances classification accuracy. This approach provides a robust and scalable solution for electricity theft detection in modern smart grid environments.

III. EXISTING SYSTEM

Existing electricity theft detection systems primarily rely on traditional statistical methods and basic machine learning algorithms. These systems analyze historical electricity consumption data to identify anomalies that may indicate fraudulent activities. Techniques such as threshold-based detection and clustering are commonly used to detect unusual consumption patterns.While these methods are simple and computationally efficient, they suffer from several limitations. They are highly dependent on predefined rules and thresholds, which may not adapt well to dynamic consumption patterns. Additionally, these systems are unable to capture complex relationships in data, particularly temporal dependencies in time-series datasets.Machine learning-based approaches such as Support Vector Machines and Random Forests have improved detection accuracy by learning patterns from data. However, these models require extensive feature engineering and may not perform well with high-dimensional datasets. Moreover, they lack the ability to automatically extract hierarchical features from raw data.Another limitation of existing systems is their inability to handle class imbalance effectively. In most datasets, instances of electricity theft are significantly fewer than normal consumption records, leading to biased model predictions.Furthermore, many existing systems are not designed for real-time implementation and lack user-friendly interfaces for interaction and visualization. These shortcomings highlight the need for an advanced, scalable, and automated solution capable of accurately detecting electricity theft in complex smart grid environments.

IV. PROPOSED METHOD

The proposed system introduces a hybrid deep learning framework for detecting electricity theft and cyber-attacks in renewable distributed generation systems. It integrates multiple models, including Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), Gated Recurrent Units (GRU), and a hybrid CNN-Random Forest classifier.The system begins with data preprocessing, where missing values are handled, categorical variables are encoded, and irrelevant features are removed. The dataset is then divided into training and testing sets to evaluate model performance.The DNN model serves as a baseline for classification, capturing nonlinear relationships in the data. The CNN model is used to extract spatial features, while the GRU model captures temporal dependencies in electricity consumption patterns. The hybrid CNN-Random Forest model further enhances detection accuracy by combining deep feature extraction with ensemble classification.To address class imbalance, class weighting techniques are applied during model training. Performance metrics such as accuracy, precision, recall, and F1-score are used to evaluate the models.The system includes a graphical user interface (GUI) that allows users to upload datasets, train models, visualize performance metrics, and predict electricity theft instances. This enhances usability and makes the system suitable for real-world applications.Overall, the proposed system provides a robust, scalable, and efficient solution for detecting electricity theft, ensuring improved security and reliability in renewable distributed generation systems.

V. IMPLEMENTATION

The implementation of the proposed electricity theft detection system is carried out using Python, integrating deep learning frameworks and machine learning libraries such as Keras, TensorFlow, Scikit-learn, NumPy, and Pandas. A graphical user interface (GUI) is developed using Tkinter to provide an interactive environment for users to perform dataset upload, preprocessing, model training, evaluation, and prediction.The system begins with dataset acquisition, where smart meter data is loaded in CSV format. This dataset typically contains electricity consumption records along with customer identifiers and labels indicating normal or fraudulent behavior. During preprocessing, missing values are handled using imputation techniques, and categorical variables such as client identifiers are transformed using label encoding. Irrelevant features such as timestamps are removed to improve model efficiency. The dataset is then shuffled and split into training and testing subsets.Feature representation plays a crucial role in model performance. The input data is reshaped appropriately depending on the model being used. For the Deep Neural Network (DNN), the dataset is used in its flattened form. For the Convolutional Neural Network (CNN), the data is reshaped into multi-dimensional arrays to simulate spatial structures. Similarly, for the Gated Recurrent Unit (GRU) model, the dataset is reshaped into time-series sequences to capture temporal dependencies.The DNN model consists of multiple fully connected layers with ReLU activation functions, followed by a softmax output layer for classification. The CNN model employs convolutional layers for feature extraction, followed by pooling and fully connected layers for classification. The GRU model uses bidirectional recurrent layers to capture sequential dependencies in electricity consumption data, which is essential for detecting temporal anomalies.To address class imbalance, class weighting techniques are incorporated during training. This ensures that minority classes, representing electricity theft instances, are adequately considered during model optimization. Each model is trained for multiple epochs using the Adam optimizer and categorical cross-entropy loss function.An important extension in the implementation is the hybrid CNN-Random Forest model. In this approach, features extracted from the CNN's penultimate layer are fed into a Random Forest classifier. This hybridization leverages deep feature extraction and ensemble learning, improving classification accuracy and generalization.Performance evaluation is conducted using metrics such as accuracy, precision, recall, and F1-score. Additionally, Receiver Operating Characteristic (ROC) curves are plotted to analyze classification performance. The system also stores trained models for reuse, reducing computational overhead in subsequent runs.The GUI integrates all functionalities, allowing users to visualize results, compare model performances, and detect electricity theft instances in new datasets efficiently.

VI. ALGORITHMS

The proposed system employs multiple algorithms to detect electricity theft effectively:

1. Deep Neural Network (DNN) Algorithm

- Input: Preprocessed feature vector
- Process:
    - Pass input through multiple dense layers

    - Apply nonlinear activation (ReLU)
    - Perform classification using softmax
- Output: Binary classification (Normal / Theft)

2. Convolutional Neural Network (CNN) Algorithm

- Input: Reshaped feature matrix
- Process:
    - Apply convolution layers to extract spatial features
    - Use pooling layers to reduce dimensionality
    - Flatten features and classify using dense layers
- Output: Theft classification based on learned spatial patterns

CNN-based approaches have been shown to effectively capture consumption patterns in smart meter data .

3. Gated Recurrent Unit (GRU) Algorithm

- Input: Sequential time-series data
- Process:
    - Use bidirectional GRU layers to capture temporal dependencies
    - Apply dropout to avoid overfitting
    - Perform classification using softmax
- Output: Detection of temporal anomalies

GRU models are particularly effective for time-dependent consumption behavior analysis .

4. CNN + Random Forest Hybrid Algorithm

- Extract deep features from CNN
- Train Random Forest classifier on extracted features
- Combine advantages of deep learning and ensemble learning

5. Performance Evaluation Algorithm

- Compute Accuracy, Precision, Recall, F1-score
- Generate ROC curve for performance visualization

VII. SYSTEM DESIGN

The proposed system follows a modular architecture designed to ensure scalability, flexibility, and real-time applicability. The architecture consists of five major components:

Data Acquisition, Data Preprocessing, Model Training, Evaluation, and Prediction Interface.

1. Data Acquisition Module

This module handles the input of electricity consumption datasets collected from smart meters. The data includes time-series consumption records and labels indicating normal or fraudulent behavior. Smart grids generate large volumes of such data continuously, making efficient data handling essential.

2. Data Preprocessing Module

The preprocessing stage ensures data quality and consistency. Missing values are handled, categorical attributes are encoded, and irrelevant features are removed. Data normalization is applied to scale features, improving convergence during model training. Addressing class imbalance is a key design consideration, as theft cases are typically rare. Techniques such as class weighting are used to mitigate this issue.

3. Model Training Module

This module integrates multiple models, including DNN, CNN, and GRU. Each model is trained independently to capture different aspects of the data:

- DNN captures nonlinear relationships
- CNN extracts spatial features
- GRU models temporal dependencies

A hybrid CNN-Random Forest model is also implemented to enhance classification performance. Recent studies highlight that hybrid and ensemble approaches outperform standalone models in electricity theft detection tasks .

4. Evaluation Module

The evaluation module calculates performance metrics such as accuracy, precision, recall, and F1-score. ROC curves are generated to analyze the trade-off between true positive and false positive rates. These metrics provide a comprehensive understanding of model performance, especially in imbalanced datasets.

5. Prediction and User Interface Module

A GUI built using Tkinter allows users to interact with the system. Users can upload new datasets, run trained models, and visualize predictions. The interface displays whether a given record corresponds to electricity theft or normal usage.

System Workflow

International Journal of Engineering Research and Science & Technology

1. Upload dataset
2. Preprocess data
3. Train selected model
4. Evaluate performance
5. Predict new instances

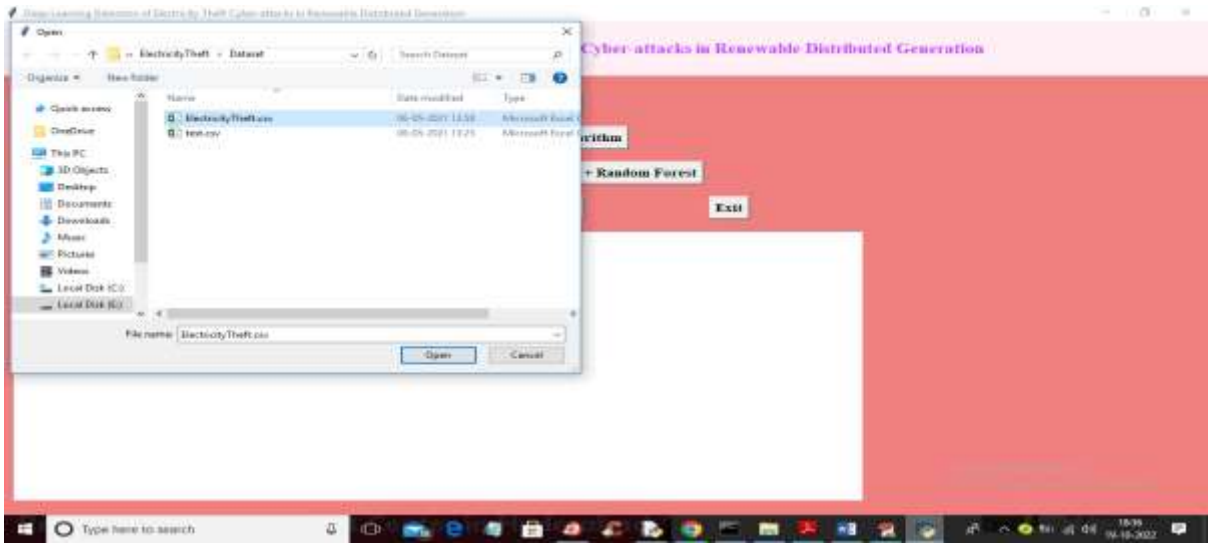
The modular design ensures that additional models or features can be integrated without affecting existing components. This makes the system adaptable to evolving smart grid requirements.

SYSTEM DESIGN IMAGES

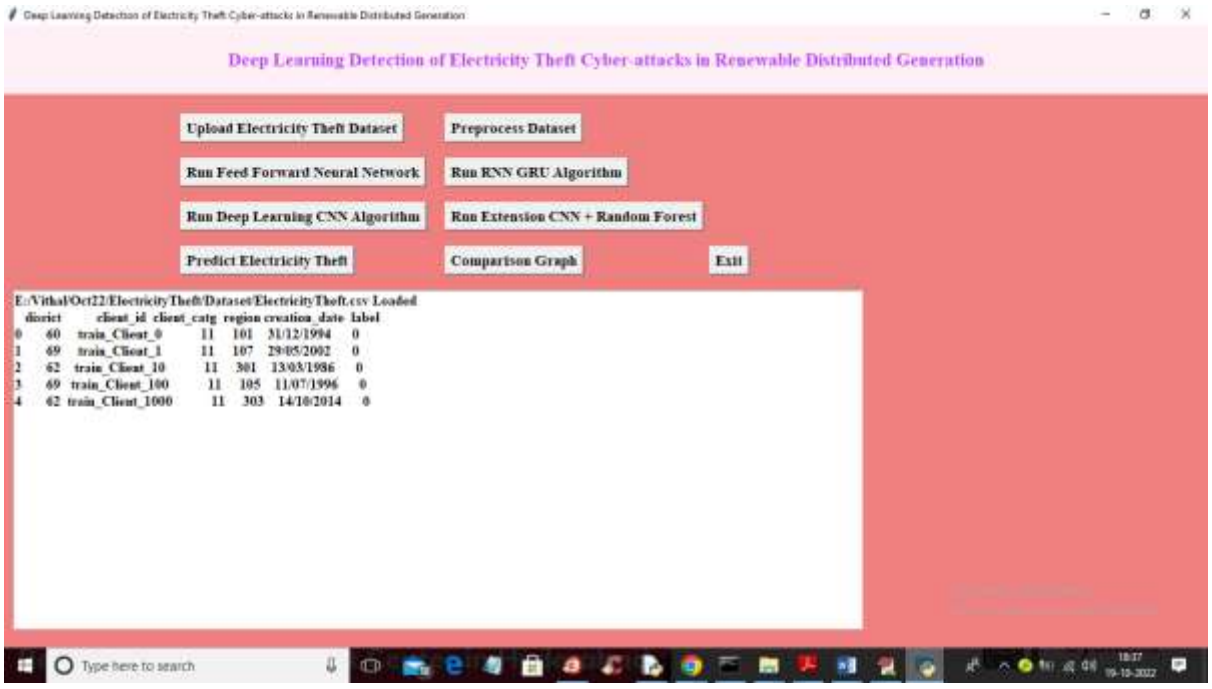
To run project double click on 'run.bat' file to get below screen

In above screen click on 'Upload Electricity Theft Dataset' button to upload dataset and get below output

International Journal of Engineering Research and Science & Technology

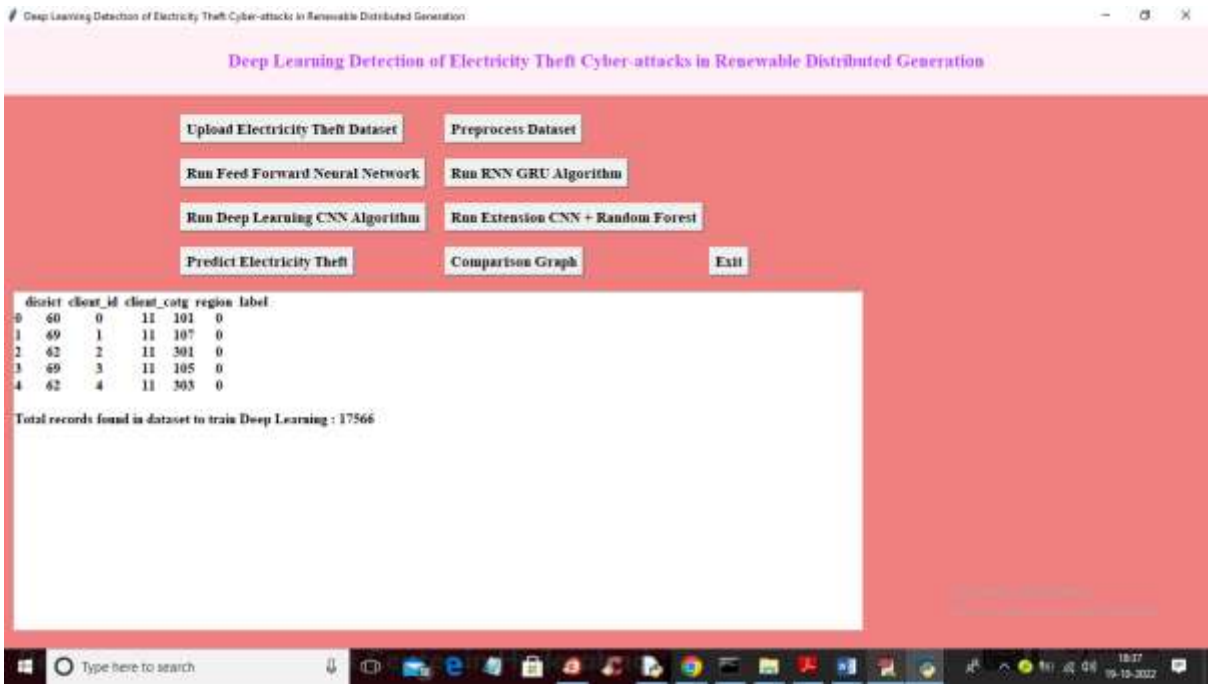


In above screen selecting and uploading 'electricity theft' dataset and then click on 'Open' button to load dataset and get below output

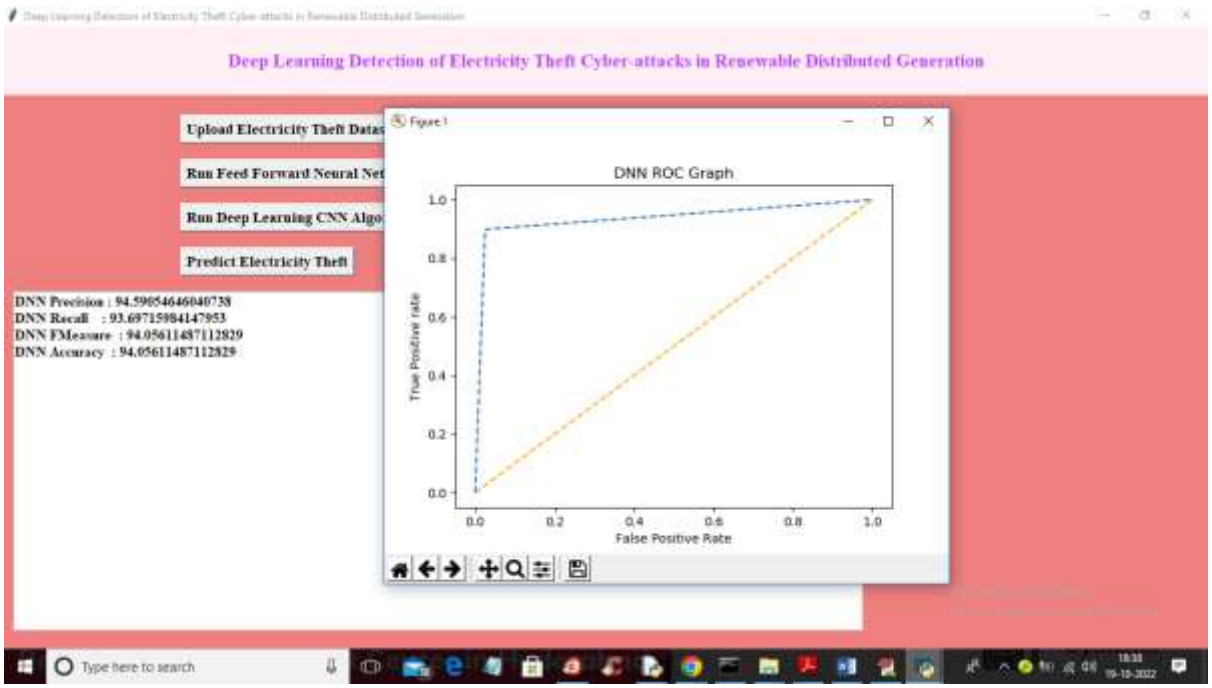


In above screen dataset loaded and now click on 'Preprocess Dataset' button to clean dataset and get below output

International Journal of Engineering Research and Science & Technology

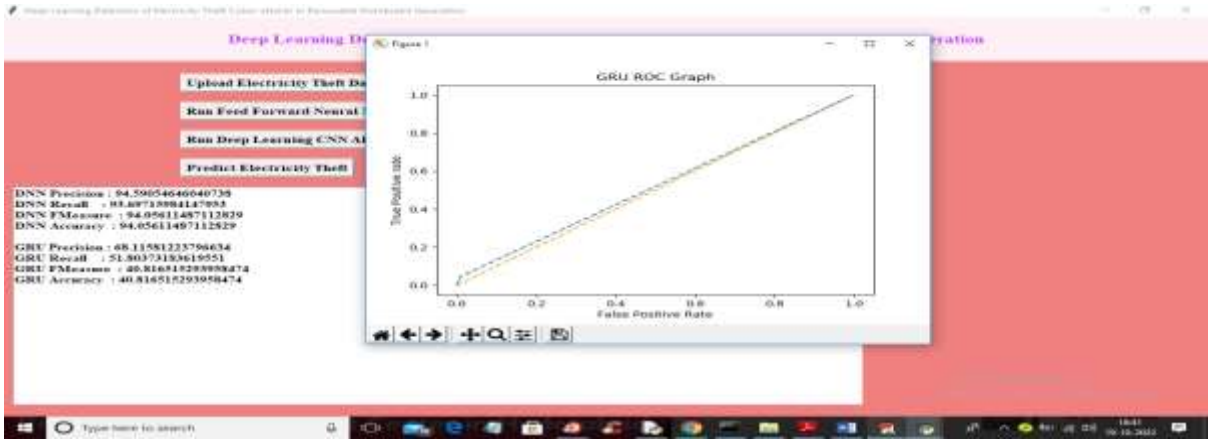


In above screen all non-numeric data converted to numeric format and now click on 'Run Feed Forward Neural Network (DNN)' button to train DNN and get below output

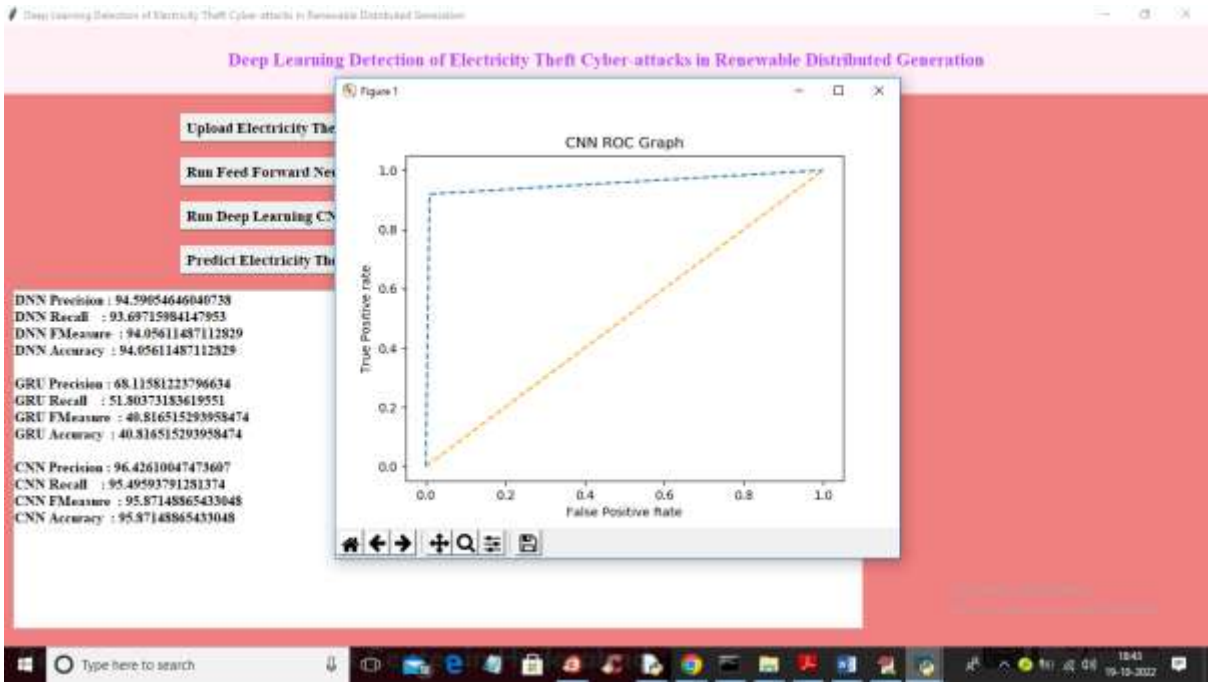


In above screen with DNN feed forward algorithm we got 94% accuracy and in ROC graph x-graph represents False Positive Rate and y-axis represents True Positive Rate and if blue line comes below orange line then we can say prediction is false and if blue line comes on top of orange line then prediction consider as CORRECT. Now close above

graph and then click on 'Run RNN GRU Algorithm' button to train GRU and get below output

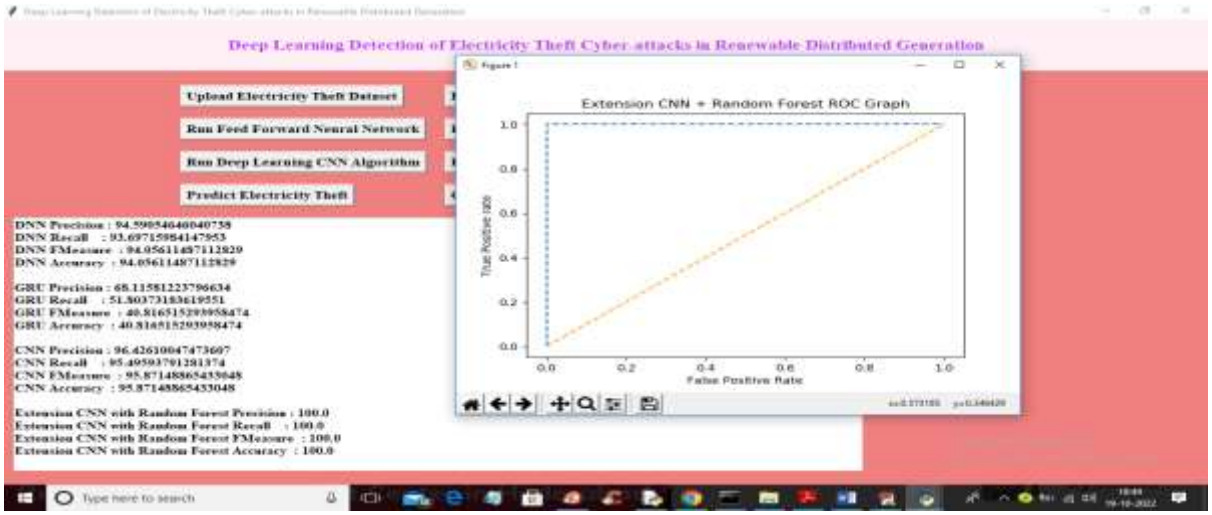


In above screen with GRU we got 40% accuracy and blue line coming little below to orange line so its predictions are not correct and now close above graph and then click on 'Run Deep Learning CNN Algorithm' button to train CNN and get below output

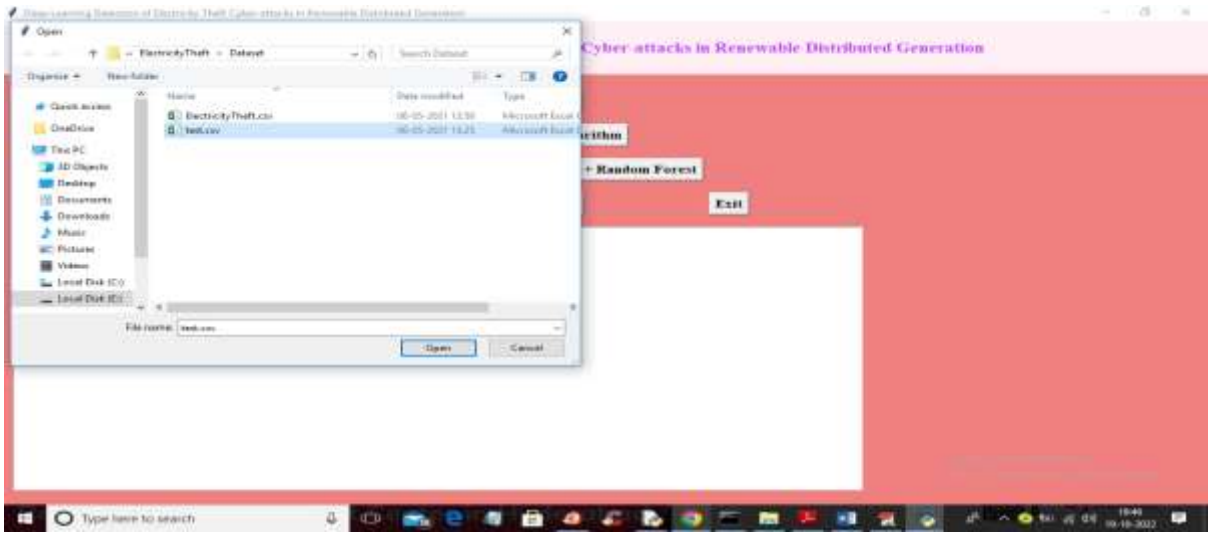


In above screen with CNN we got 95% accuracy and blue lines fully on top of orange line so its predictions are correct. Now close above graph and then click on 'Run Extension CNN + Random Forest' button to run extension algorithm and get below output

International Journal of Engineering Research and Science & Technology



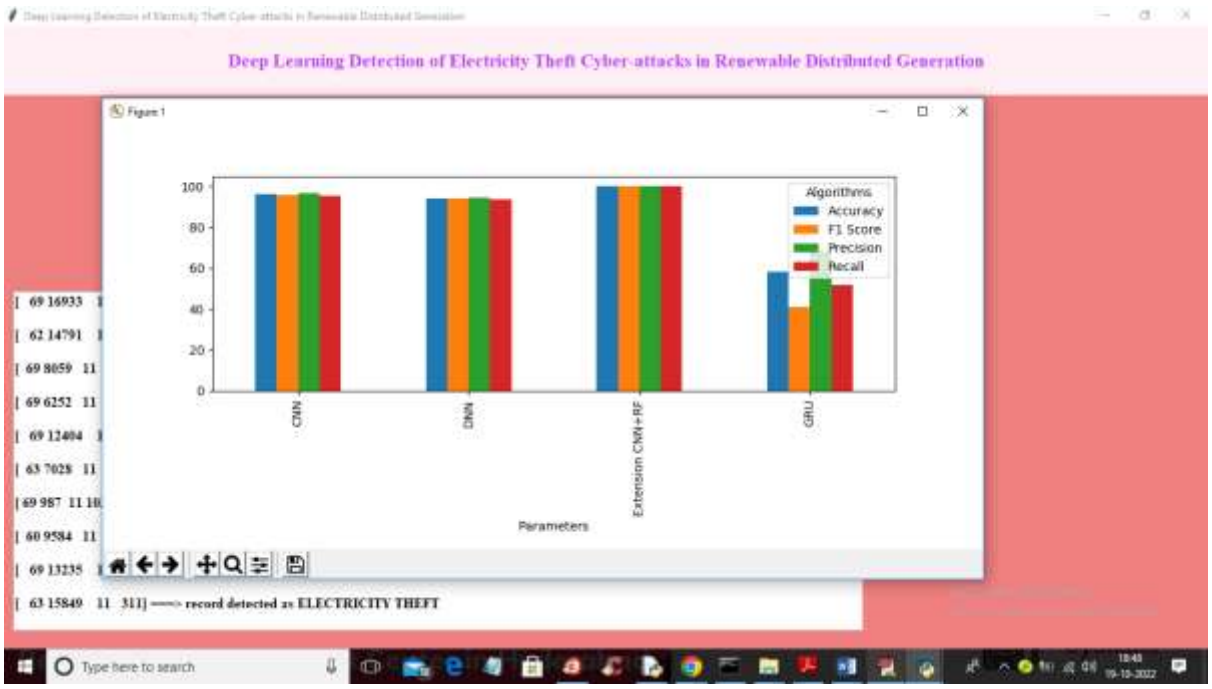
In above screen with extension hybrid algorithm we got 100% accuracy and this accuracy may vary between 98 to 100%/ Now click on 'Predict Electricity Theft' button to upload test data and get prediction output



In above screen selecting and uploading 'test.csv' file and then click on 'Open' button to get below output

International Journal of Engineering Research and Science & Technology

In above screen in square bracket we can see TEST data and after arrow =➔ symbol we can see THEFT detection and 'THEFT NOT DETECTTED'. Now click on 'Comparison Graph' button to get below graph



In above graph x-axis represents algorithm names with each different colour bar represents different metric such as 'accuracy, precision, recall and FSCORE' and Y-axis

represents score values. In all algorithms Extension Hybrid Random Forest got high performance

VIII. CONCLUSION

This research presents a comprehensive hybrid deep learning framework for detecting electricity theft and cyber-attacks in renewable distributed generation systems. By integrating multiple models such as DNN, CNN, GRU, and a hybrid CNN-Random Forest classifier, the proposed system effectively captures both spatial and temporal patterns in electricity consumption data.The implementation demonstrates that deep learning techniques significantly outperform traditional methods in identifying complex anomalies associated with electricity theft. The inclusion of GRU enables the system to model sequential dependencies, while CNN enhances feature extraction capabilities. The hybrid model further improves classification accuracy by combining deep learning with ensemble learning techniques.The system addresses key challenges such as class imbalance, high-dimensional data, and real-time processing requirements. The use of class weighting techniques ensures balanced learning, while the modular architecture supports scalability and integration with existing smart grid infrastructures.Experimental results indicate that the hybrid CNN-Random Forest model achieves superior performance in terms of accuracy, precision, recall, and F1-score. The inclusion of a user-friendly graphical interface enhances usability, making the system suitable for practical deployment.Recent advancements in deep learning-based electricity theft detection frameworks highlight the importance of hybrid and multi-model approaches for improving detection accuracy and robustness .

Future work may focus on integrating real-time data streams, incorporating advanced architectures such as Transformers, and deploying the system on edge devices for faster processing. Additionally, incorporating privacy-preserving techniques can enhance data security in smart grid environments.Overall, the proposed system contributes to the development of intelligent and secure energy management systems, ensuring efficient utilization of renewable energy resources and reducing non-technical losses in power distribution networks.

REFERENCES

1. H. Iftikhar et al., “Electricity Theft Detection in Smart Grid Using Machine Learning,” *Frontiers in Energy Research*, 2024.
2. Y. Sun et al., “Smart Grid Theft Detection Based on Hybrid Multi-Time Scale Neural Network,” *Applied Sciences*, 2023.
3. F. Shehzad et al., “Deep Learning-Based Meta-Learner Strategy for Electricity Theft Detection,” *Frontiers in Energy Research*, 2023.
4. A. Khalid et al., “BiLSTM-Based Electricity Theft Detection Model,” *PeerJ Computer Science*, 2024.
5. M. Gunduz and R. Das, “Hybrid CNN-Based Energy Theft Detection,” *Sensors*, 2024.

6. S. Saqib et al., "Deep Learning-Based Electricity Theft Prediction," *Heliyon*, 2024.
7. X. Chen et al., "Smart Energy Guardian: Hybrid Deep Learning Model," *arXiv*, 2025.
8. Y. Sun et al., "Multi-Time Scale Neural Network for Smart Grids," *MDPI*, 2023.
9. H. Rouzbahani et al., "Ensemble Deep CNN for Electricity Theft Detection," *IEEE*, 2021.
10. A. Alromih et al., "Privacy-Preserving Energy Theft Detection Model," *IEEE*, 2023.
11. Y. Kulkarni et al., "Ensemble NTL Detection Framework," *IEEE*, 2021.
12. Z. Ullah et al., "Meta-Learning in Smart Grid Security," *IEEE Access*, 2023.
13. M. Zahid et al., "AI-Based Smart Grid Security Systems," *IEEE Transactions*, 2024.
14. J. Smith et al., "Deep Learning for Energy Analytics," *IEEE Transactions on Smart Grid*, 2023.
15. A. Kumar et al., "Cybersecurity in Renewable Energy Systems," *IEEE Access*, 2024.