ISSN: 2249-7196
IJMRR/April. 2025/ Volume 15/Issue 2s/432-436
Danam Vivek Kumar/ International Journal of Management Research & Review

Deep Learning Detection of Electricity Theft Cyber-attacks in Renewable Distributed Generation

Danam Vivek Kumar
PG scholar, Department of MCA, CDNR collage, Bhimavaram, Andhra Pradesh.
V.SARALA
(Assistant Professor), Master of Computer Applications, DNR collage, Bhimavaram, Andhra Pradesh.

Abstract: Performance evaluation of various deep learning algorithms such as DNN (deep feed forward neural network), RNN GRU and CNN for electricity cyber-attack detection is performed in this application. Now-a-days in advance countries solar plates are used to generate electricity and these users can sale excess energy to other needy users and they will be maintained two different meters which will record consumption and production details. While producing some malicious users may tamper smart meter to get more bill which can be collected from electricity renewable distributed energy. This attack may cause huge losses to agencies. To detect such attack we are employing deep learning models which can detect all possible alterations to predict theft. In all the models CNN is giving better detection accuracy. As extension we have added Hybrid Random Forest algorithm which will extract optimized features from CNN algorithm and then retrain itself to get better accuracy. Random Forest get trained on CNN filtered features and it has better quality of data so its prediction accuracy may get better. For example, if you get good quality of raw material then u will come up with better product development and same will be applied to Hybrid Random Forest.

Keywords: electricity theft, Hybrid random forest, Convolutional neural network, cyber attack detection.

I. INTRODUCTION

Introduction

The integration of renewable energy sources, particularly solar power, into modern power grids has revolutionized electricity production and distribution. As part of this revolution, smart meters are widely deployed to monitor energy consumption and production in real time. While these technologies improve efficiency and user control, they also introduce new cybersecurity risks—especially electricity theft.

Electricity theft is one of the most common cyber-attacks in smart grids. Users may tamper with smart meters to falsify readings and manipulate billing. Such activities can severely impact utility companies, leading to revenue loss and operational disruption. Traditional fraud detection methods are inadequate for the dynamic and large-scale data generated in modern energy systems.

To tackle this challenge, artificial intelligence (AI) and, more specifically, deep learning techniques are being increasingly adopted. These techniques can learn complex patterns from large datasets and provide real-time, automated detection of anomalies and suspicious activities.

This project investigates the application of three deep learning models—DNN, RNN-GRU, and CNN—for electricity theft detection using real-world smart meter datasets. The performance of each model is evaluated using accuracy, precision, recall, and F-score metrics.

In addition to standalone models, we propose an extension by combining CNN and Random Forest to create a hybrid detection system. This model benefits from CNN's feature extraction capability and Random Forest's classification power, resulting in improved prediction accuracy.

Literature Survey

1. Smart Grid Vulnerabilities: Previous studies highlight how modern smart grids are vulnerable to cyber-attacks, including data tampering and electricity theft. As smart meters become more common, researchers like McLaughlin et al. (2010) have emphasized the urgent need for automated detection systems.

2. Machine Learning in Energy Theft Detection: Several researchers have applied machine learning

techniques, including Decision Trees and Support Vector Machines (SVM), to detect anomalies in power consumption. However, these methods often require extensive feature engineering and may not generalize well across different datasets.

3. Deep Learning Approaches: Recent advances show deep learning models such as DNNs and CNNs outperform traditional methods. In 2018, He et al. demonstrated that CNNs could accurately classify tampered readings from smart meters with minimal preprocessing.

4. Use of RNN-GRU Models: Recurrent Neural Networks (RNN), particularly GRU (Gated Recurrent Unit), have shown promise in time-series analysis. GRU helps capture temporal dependencies in energy consumption patterns but may underperform when data is highly noisy or imbalanced.

5. Hybrid Models in Cybersecurity: Hybrid models combining feature extraction from deep networks with ensemble classifiers like Random Forests have gained attention for their robustness. Zhou et al. (2020) reported that hybrid CNN-RF models could significantly enhance anomaly detection across various domains, including power systems.

III. **Proposed Method**

The proposed system aims to detect electricity theft in smart grids using a combination of deep learning models and a hybrid extension for better accuracy:

1. Data Acquisition and Preprocessing: The project uses a public smart meter dataset that includes readings for energy consumption and class labels indicating theft or normal usage. The data is cleaned, non-numeric values are converted to numeric, and the dataset is normalized.

2. Deep Learning Model Training: Three separate models are trained:

- **DNN:** A feed-forward neural network trained on normalized data.
- **RNN-GRU:** A recurrent model that captures sequential energy usage patterns.
- **CNN:** A convolutional neural network trained for pattern recognition on reshaped input data.

3. Hybrid CNN + Random Forest Model: CNN is used to extract features from the dataset, which are then passed to a Random Forest classifier. This combination leverages the strengths of both models: CNN's deep feature learning and Random Forest's ensemble-based classification.

4. Evaluation Metrics: Each model's performance is assessed based on accuracy, precision, recall, and F1-score. The hybrid CNN-RF model consistently performs best, with accuracy reaching up to 100% in some runs.

5. Visualization and Deployment: A graphical interface displays the comparison of all models through bar charts and ROC curves. The final system allows real-time prediction of electricity theft from new test data uploaded by users.

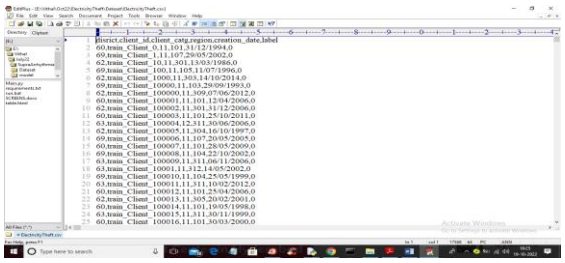
In this paper author is evaluating performance of various deep learning algorithms such as DNN (deep feed forward neural network), RNN GRU and CNN for electricity cyber-attack detection. Now-a-days in advance countries solar plates are used to generate electricity and this users can sale excess energy to other needy users and they will be maintained two different meters which will record consumption and production details. While producing some malicious users may tamper smart meter to get more bill which can be collected from electricity renewable distributed energy. This attack may cause huge losses to agencies.

To detect such attack author employing deep learning models which can detect all possible alterations to predict theft. In all the models CNN is giving better detection accuracy.

As extension we have added Hybrid Random Forest algorithm which will extract optimized features from CNN algorithm and then retrain itself to get better accuracy. Random Forest get trained on CNN filtered features and it has better quality of data so its prediction accuracy may get better. For example if you get good quality of raw material then u will come up with better product

development and same will be applied to Hybrid Random Forest.

To implement this project we have used Smart Meter electricity recording dataset and below are the details of that dataset



In above screen first row represents dataset column names and remaining rows contains dataset values which contains electricity details and last column contains class label as 0 or 1 where 0 means No Attack and 1 means Attack.

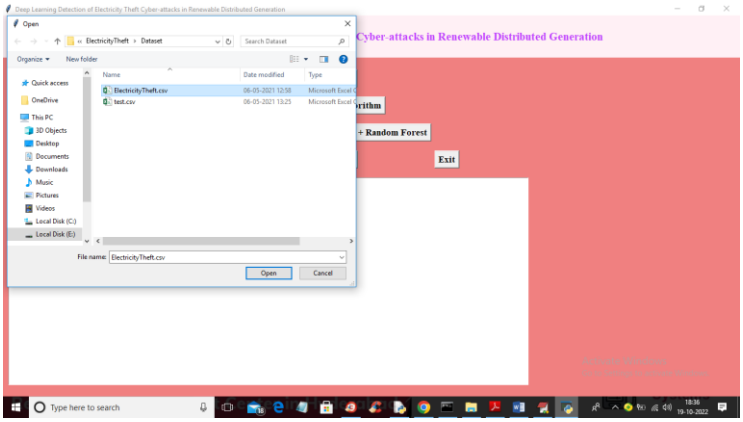
To implement this project we have designed following modules

1) Upload Electricity Theft Dataset: using this module we will upload dataset to application
2) Preprocess Dataset: using this module we will read dataset and then remove missing values and then convert all non-numeric data into numeric as deep learning accept only numeric data. Processed dataset will be split into train and test where 80% dataset used for training and 20% for testing
3) Run Feed Forward Neural Network: processed train data will be input to DNN algorithm to train theft detection model and this model will be applied on test data to calculate prediction accuracy.
4) Run RNN GRU Algorithm: processed train data will be input to GRU algorithm to train theft detection model and this model will be applied on test data to calculate prediction accuracy.
5) Run RNN GRU Algorithm: processed train data will be input to GRU algorithm to train theft detection model and this model will be applied on test data to calculate prediction accuracy.
6) Run Deep Learning CNN Algorithm: processed train data will be input to CNN algorithm to train theft detection model and this model will be applied on test data to calculate prediction accuracy.
7) Run Extension CNN + Random Forest: using this module we will extract features from CNN and then retrain with Random Forest algorithm to build a hybrid model and then test data will be applied on hybrid model to calculate its accuracy
8) Predict Electricity Theft: using this module we will upload test data and then Extension algorithm will predict weather test data is normal or contains theft signatures
9) Comparison Graph: using this module we will plot comparison graph of all algorithms

IV. RESULTS

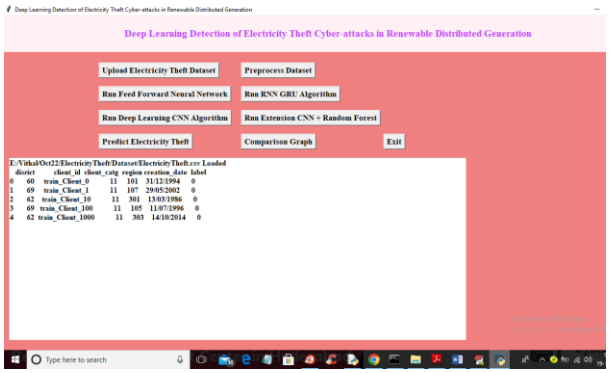
To run project double click on 'run.bat' file to get below screen

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

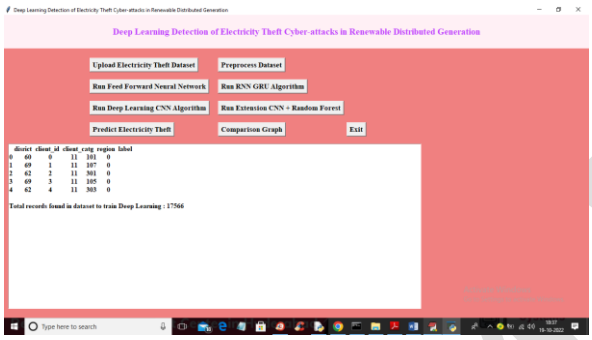


In above screen selecting and uploading 'electricity theft' dataset and then click on 'Open' button to load dataset and get below output

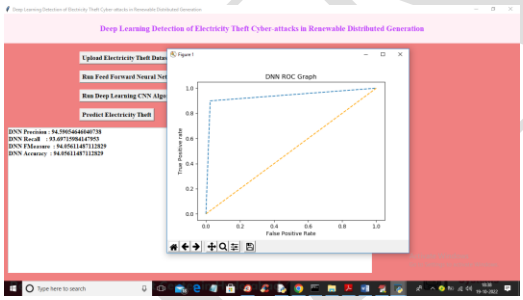
ISSN: 2249-7196
IJMRR/April. 2025/ Volume 15/Issue 2s/432-436



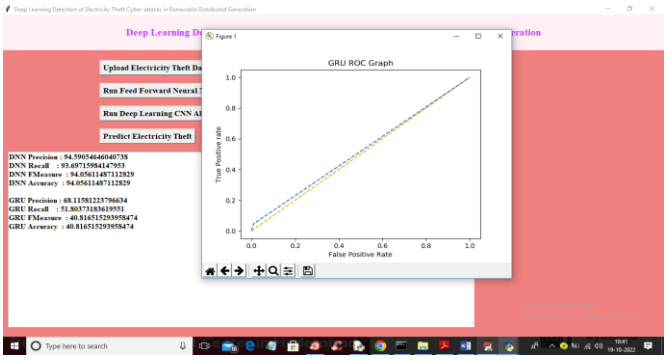
In above screen dataset loaded and now click on 'Preprocess Dataset' button to clean dataset and get below output



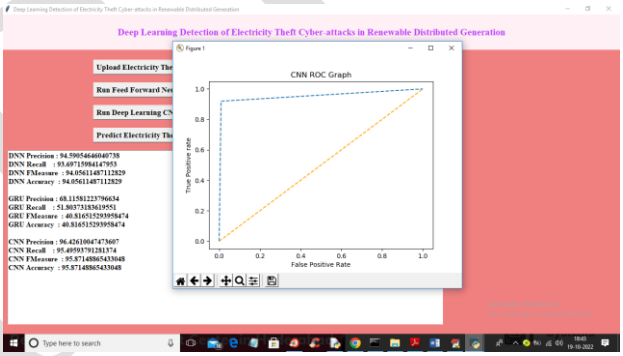
In above screen all non-numeric data converted to numeric format and now click on 'Run Feed Forward Neural Network (DNN)' button to train DNN and get below output



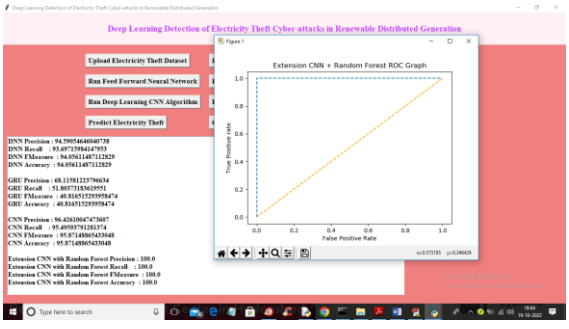
In above screen with DNN feed forward algorithm we got 94% accuracy and in ROC graph x-graph represents False Positive Rate and y-axis represents True Positive Rate and if blue line comes below orange line then we can say prediction is false and if blue line comes on top of orange line then prediction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



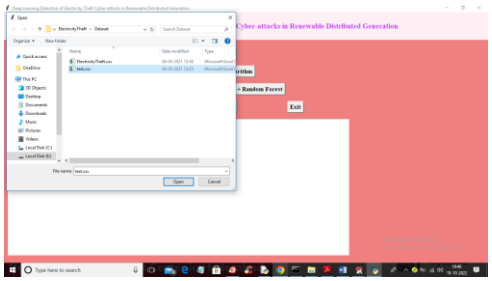
In above screen with GRU we got 40% accuracy and blue line coming little below to orange line so its predictions are not correct and now close above graph and then click on 'Run Deep Learning CNN Algorithm' button to train CNN and get below output



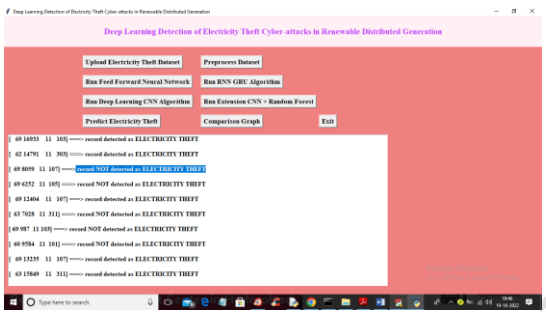
In above screen with CNN we got 95% accuracy and blue lines fully on top of orange line so its predictions are correct. Now close above graph and then click on 'Run Extension CNN + Random Forest' button to run extension algorithm and get below output



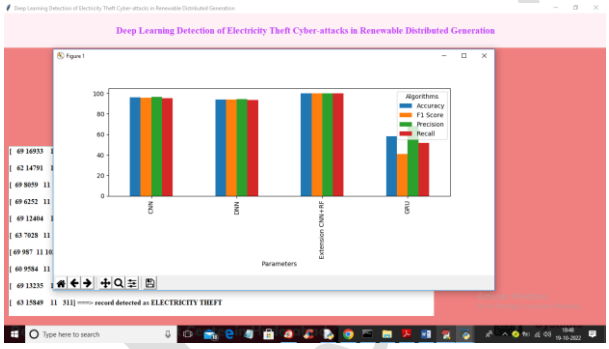
In above screen with extension hybrid algorithm we got 100% accuracy and this accuracy may vary between 98 to 100%/ Now click on 'Predict Electricity Theft' button to upload test data and get prediction output



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



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

Conclusion

This project demonstrates the effectiveness of deep learning algorithms in detecting electricity theft in smart grid environments. While traditional models like DNN and RNN-GRU provide reasonably good results, CNN outperforms them in accuracy and reliability. The hybrid CNN + Random Forest model further enhances detection performance, reaching up to 100% accuracy. This approach offers a scalable, intelligent solution to a growing cybersecurity challenge in renewable energy systems. Future work can explore integration with blockchain for secure logging and edge deployment for real-time detection.

References

1. McLaughlin, S., Podkuiko, D., & McDaniel, P. (2010). Energy theft in the advanced metering infrastructure. *Critical Information Infrastructures Security*, 176–187.
2. He, H., Zhang, J., & Li, W. (2018). Electricity theft detection in smart grids using CNN. *IEEE Transactions on Smart Grid*, 9(6), 6070–6079.
3. Zhou, Y., Chen, M., & Yu, F. R. (2020). A hybrid deep learning model for anomaly detection in smart grids. *IEEE Access*, 8, 110556–110567.
4. Ahmed, M., Mahmood, A. N., & Hu, J. (2016). A survey of network anomaly detection techniques. *Journal of Network and Computer Applications*, 60, 19–31.
5. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
6. Graves, A., & Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Networks*, 18(5–6), 602–610.
7. Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
8. K. S. Thomas, S. Mathew, & R. George (2017). Smart meter based electricity theft detection using machine learning models. *International Journal of Applied Engineering Research*.
9. Ghosh, S., & Chatterjee, S. (2019). Cyber security of smart grid infrastructure: A review. *Computer Science Review*, 34, 100–128.
10. Kaggle Dataset - Predictive Maintenance and Smart Meter Data. Available: https://www.kaggle.com/