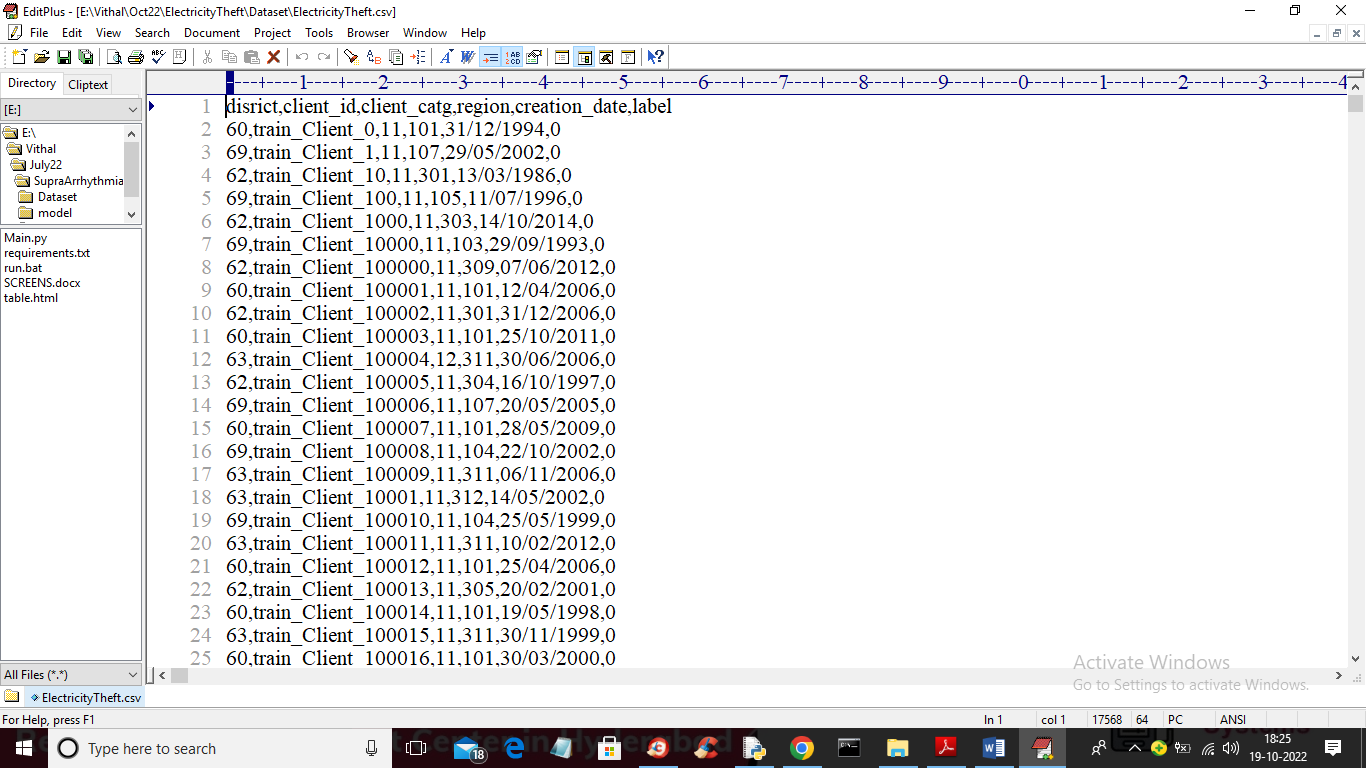
Deep Learning Detection of Electricity Theft Cyber-attacks in Renewable Distributed Generation

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



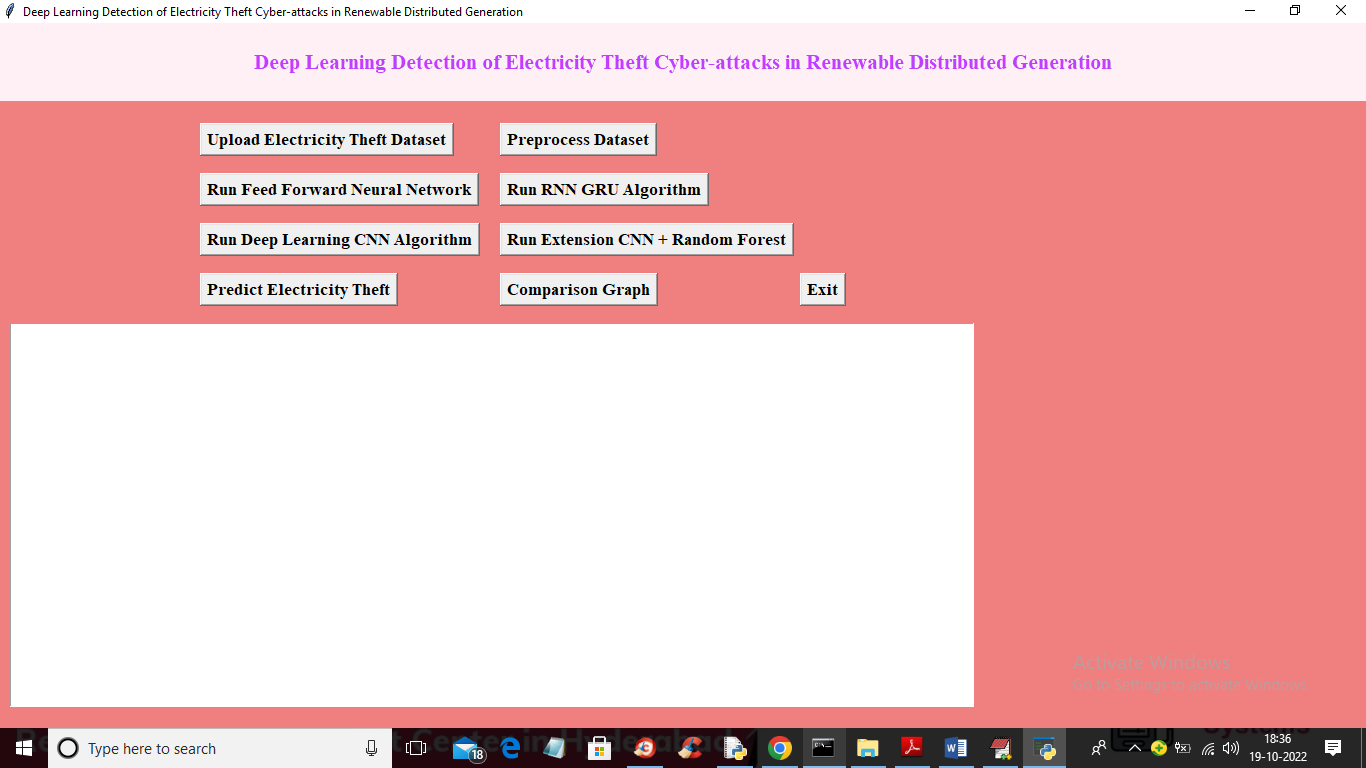
In above screen first row represents dataset column names and remaining rows contains dataset values which contains electricity details and last column contains class label as 0 or 1 where 0 means No Attack and 1 means Attack.

To implement this project we have designed following modules

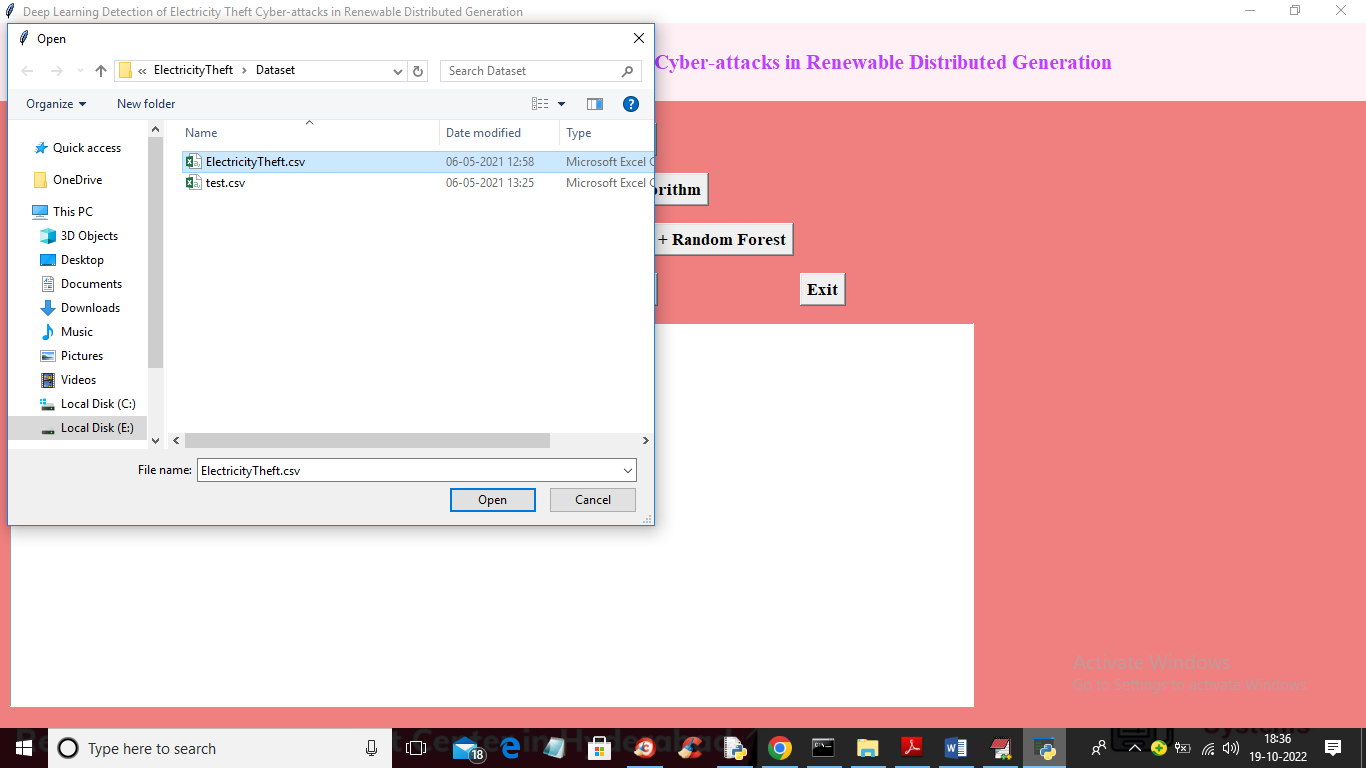
1. Upload Electricity Theft Dataset: using this module we will upload dataset to application
2. Preprocess Dataset: using this module we will read dataset and then remove missing values and then convert all non-numeric data into numeric as deep learning accept only numeric data. Processed dataset will 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

SCREEN SHOTS

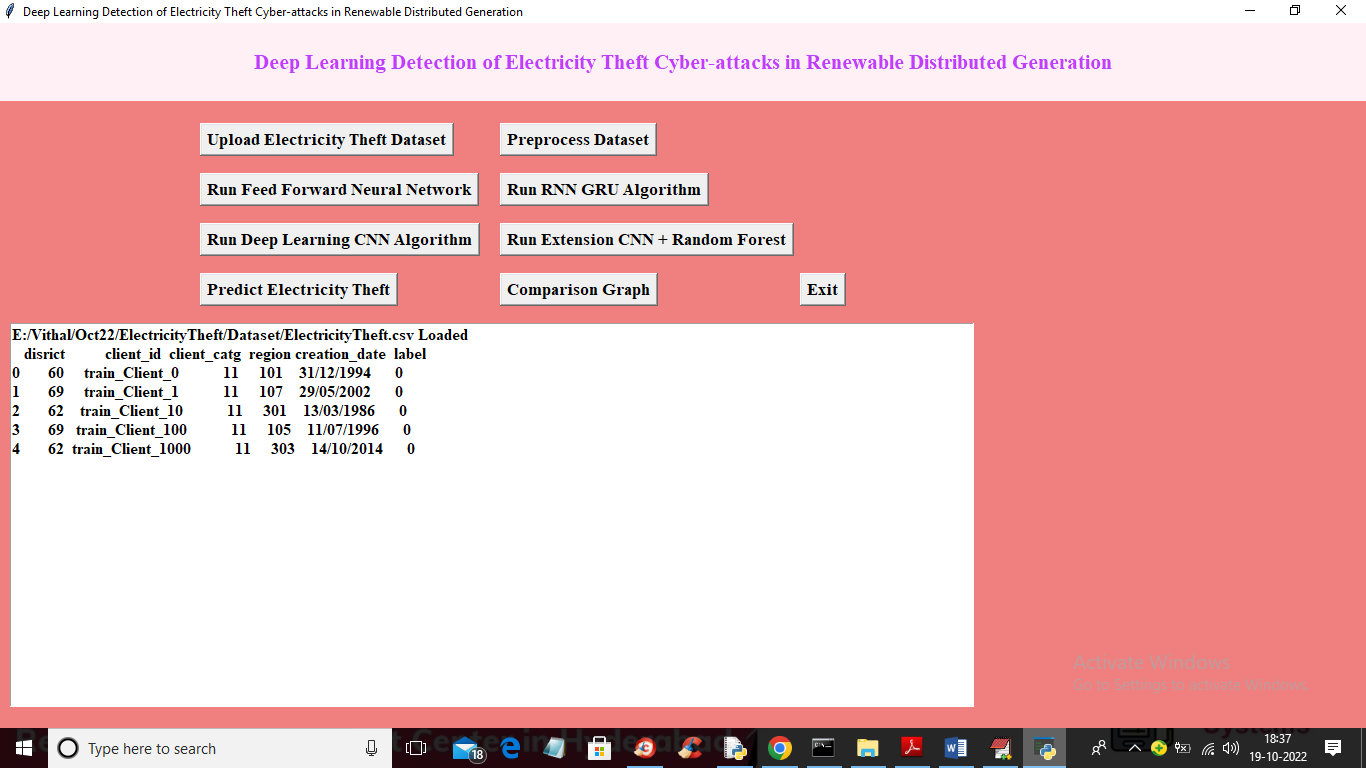
To run project double click on ‘run.bat’ file to get below screen



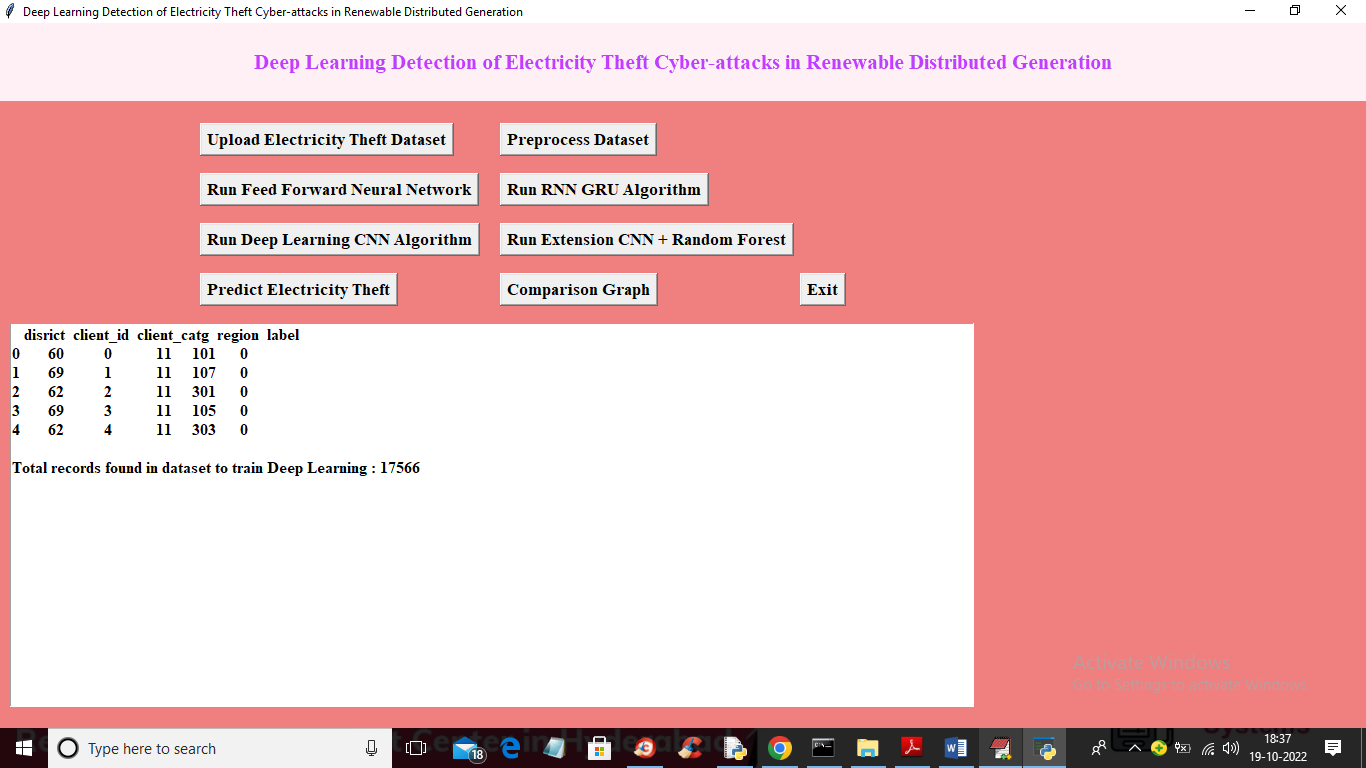
In above screen click on ‘Upload Electricity Theft Dataset’ button to upload dataset and get below output



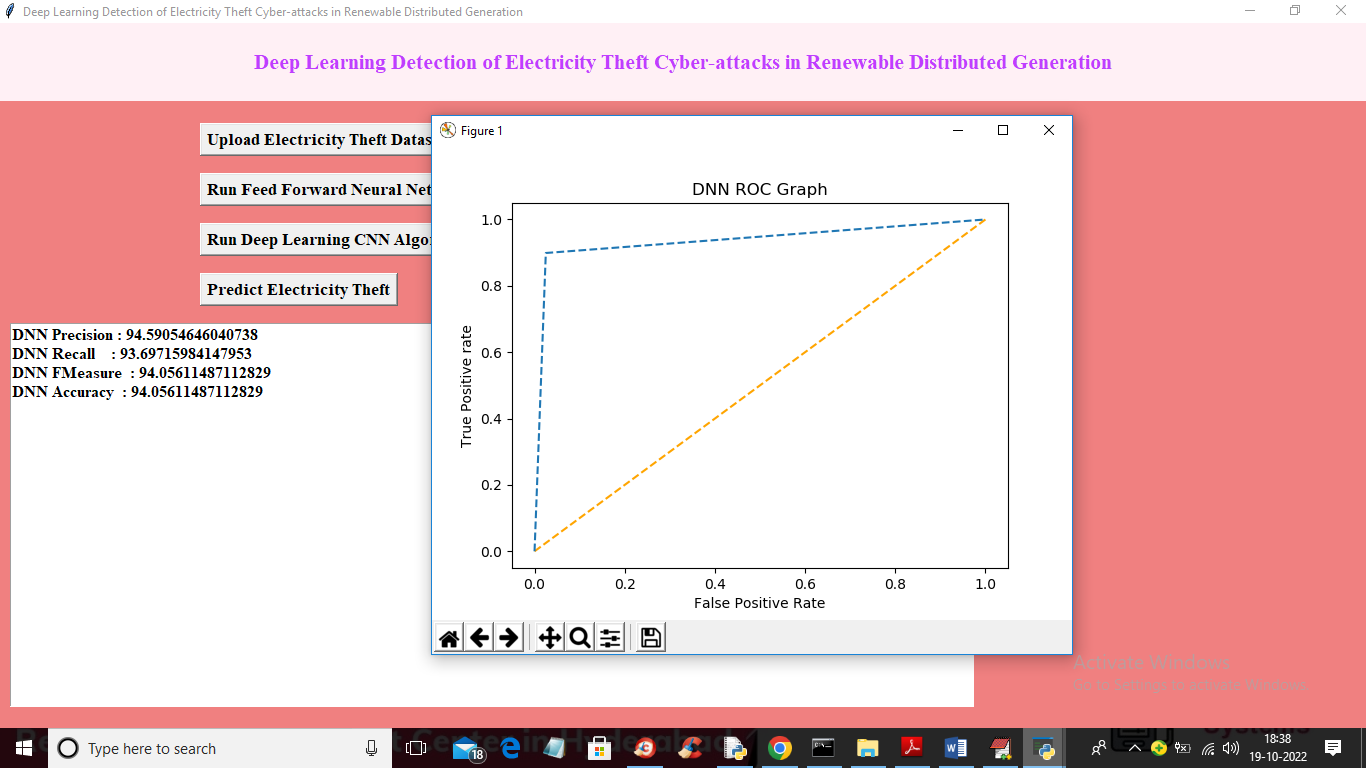
In above screen selecting and uploading ‘electricity theft’ dataset and then click on ‘Open’ button to load dataset and get below output



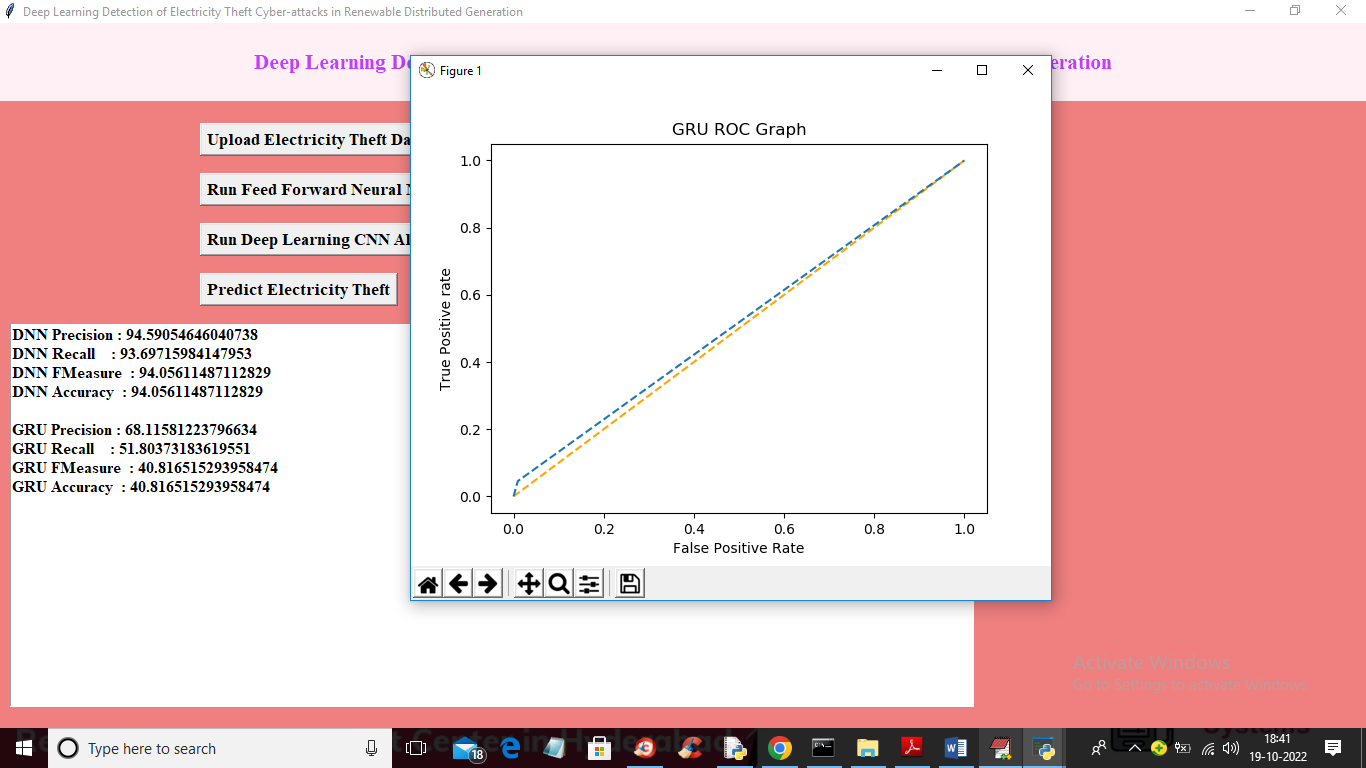
In above screen dataset loaded and now click on ‘Preprocess Dataset’ button to clean dataset and get below output



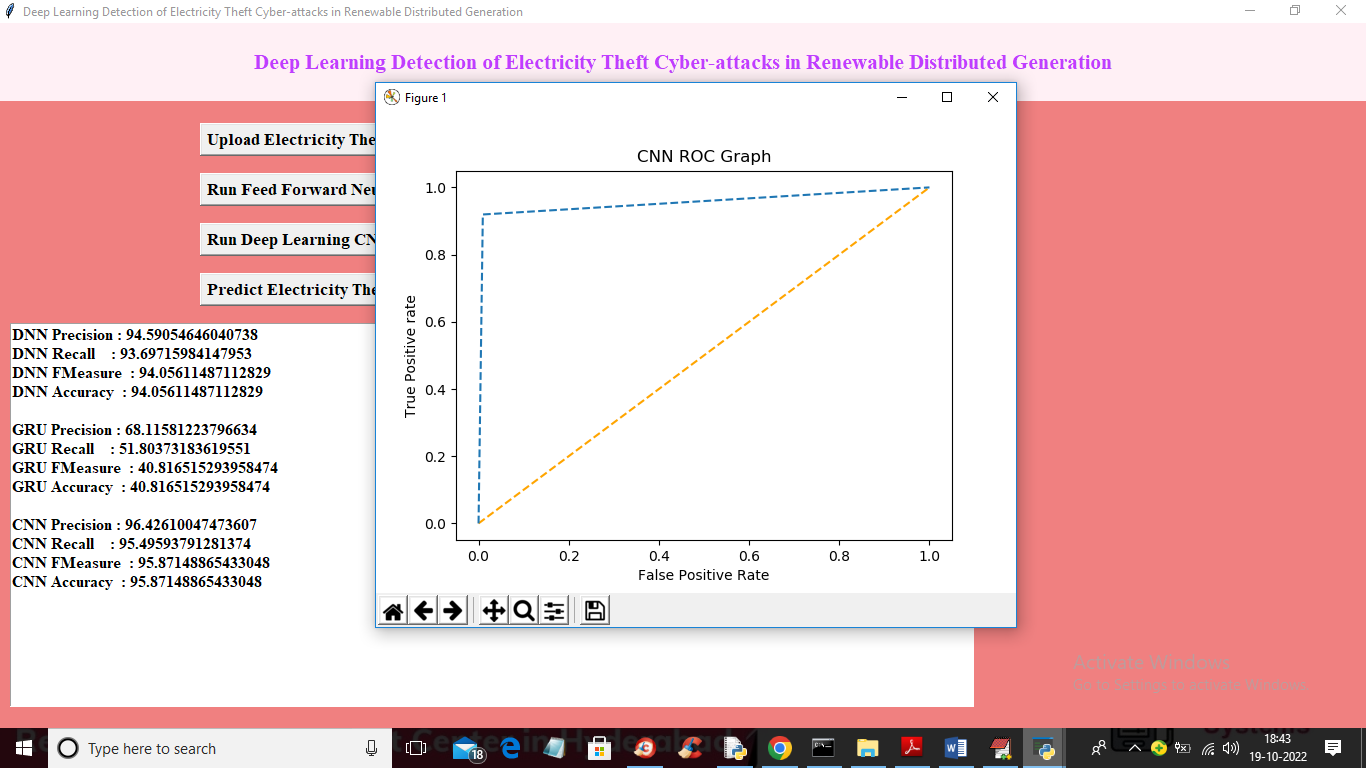
In above screen all non-numeric data converted to numeric format and now click on ‘Run Feed Forward Neural Network (DNN)’ button to train DNN and get below output



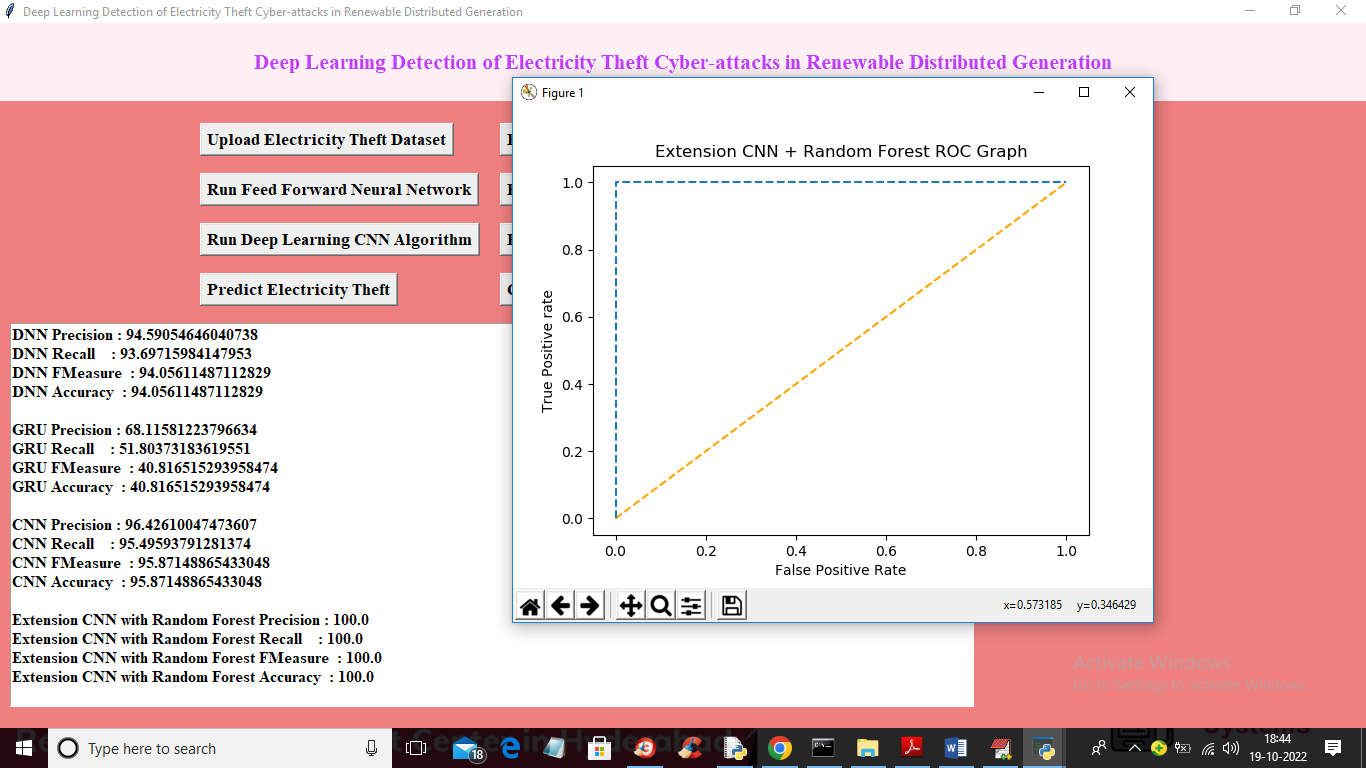
In above screen with DNN feed forward algorithm we got 94% accuracy and in ROC graph x-graph represents False Positive Rate and y-axis represents True Positive Rate and if blue line comes below orange line then we can say prediction is false and if blue line comes on top of orange line then prediction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



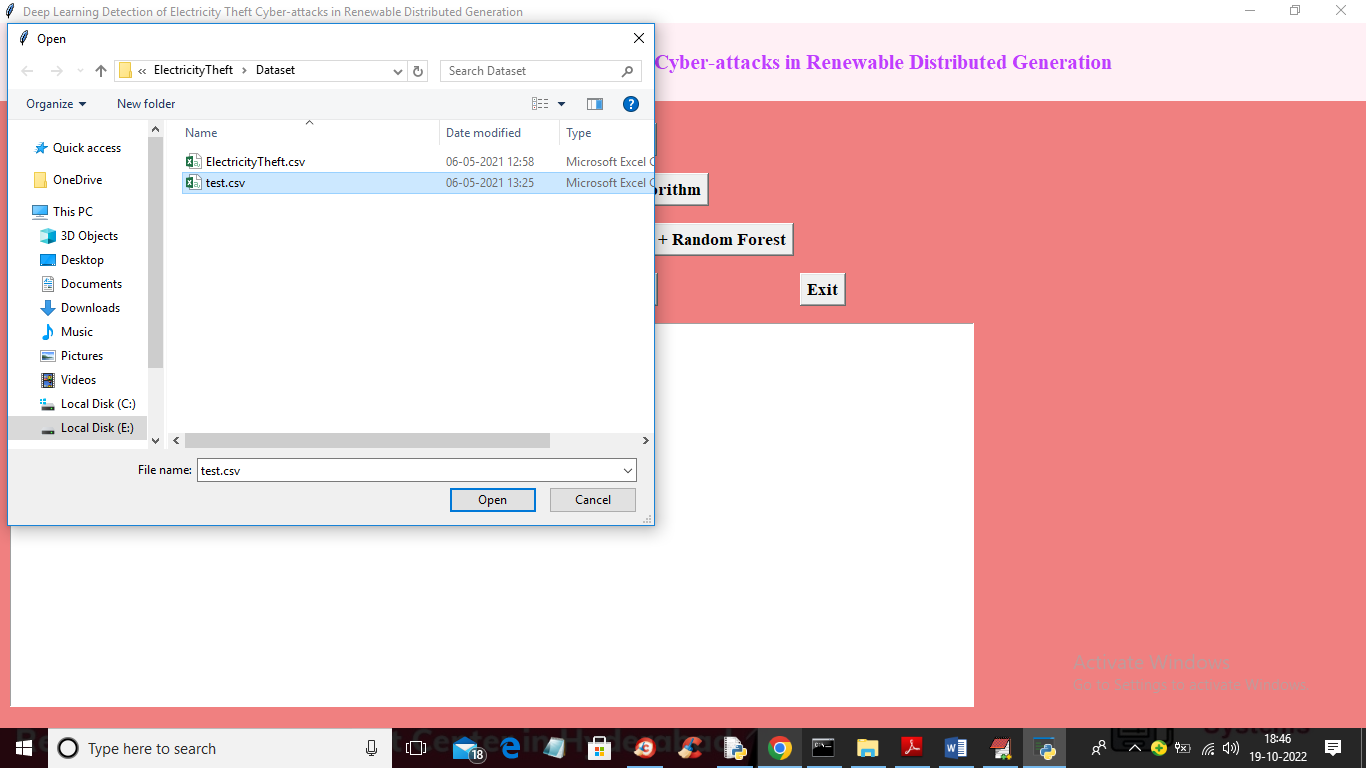
In above screen with GRU we got 40% accuracy and blue line coming little below to orange line so its predictions are not correct and now close above graph and then click on ‘Run Deep Learning CNN Algorithm’ button to train CNN and get below output



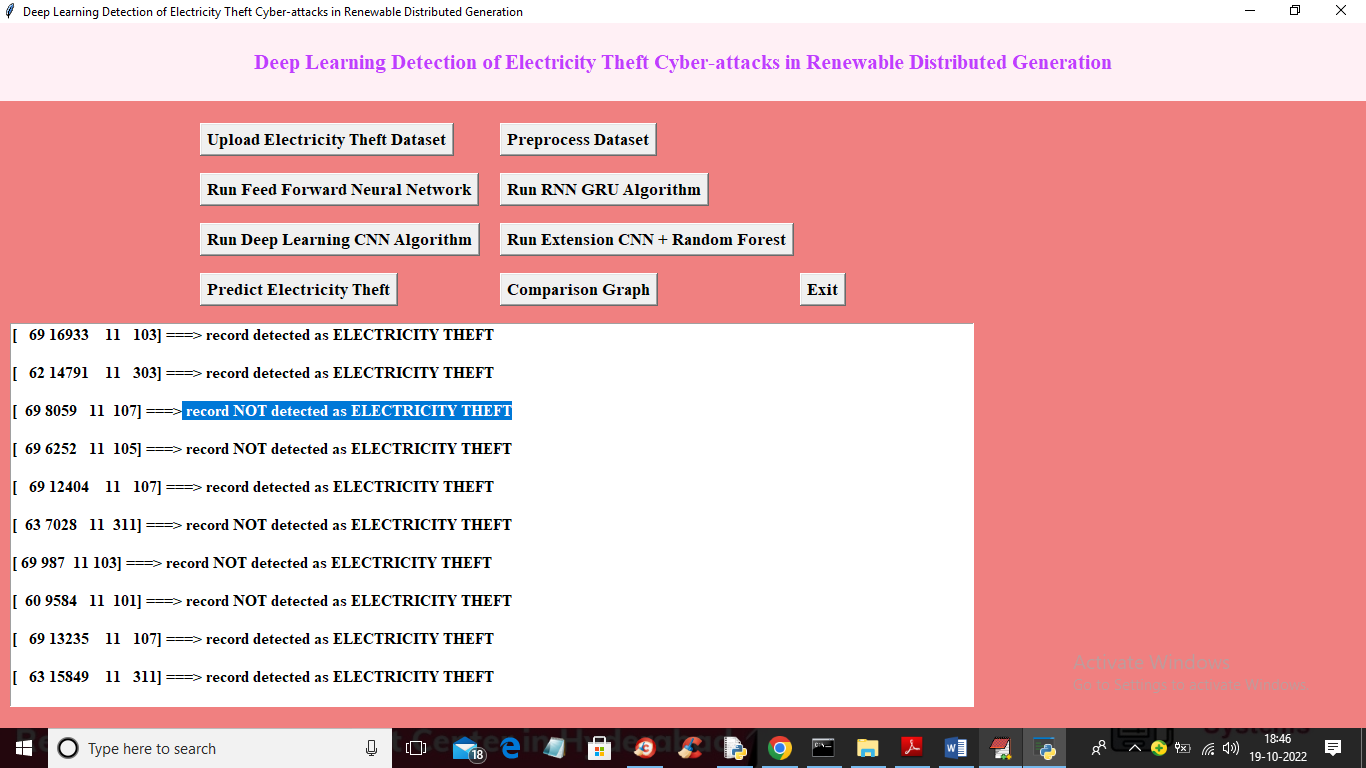
In above screen with CNN we got 95% accuracy and blue lines fully on top of orange line so its predictions are correct. Now close above graph and then click on ‘Run Extension CNN + Random Forest’ button to run extension algorithm and get below output



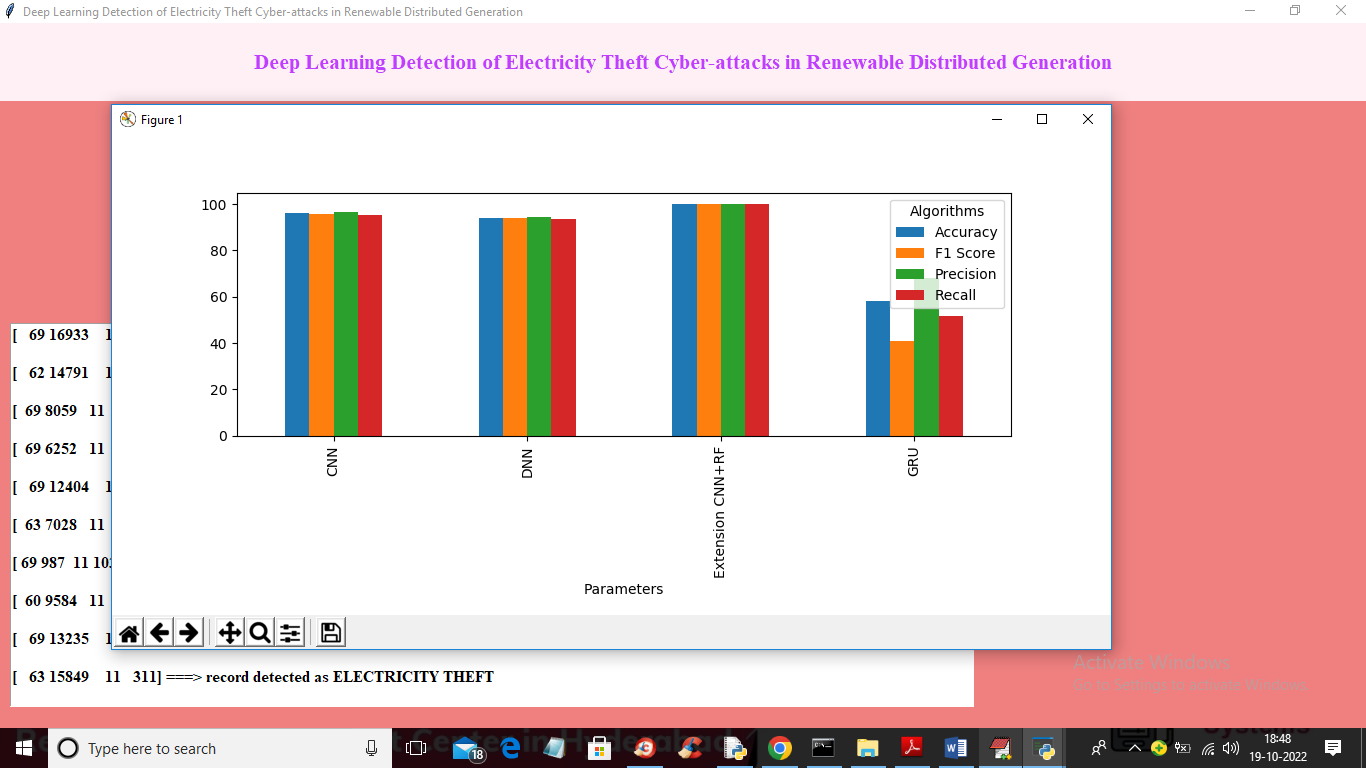
In above screen with extension hybrid algorithm we got 100% accuracy and this accuracy may vary between 98 to 100%/ Now click on ‘Predict Electricity Theft’ button to upload test data and get prediction output



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