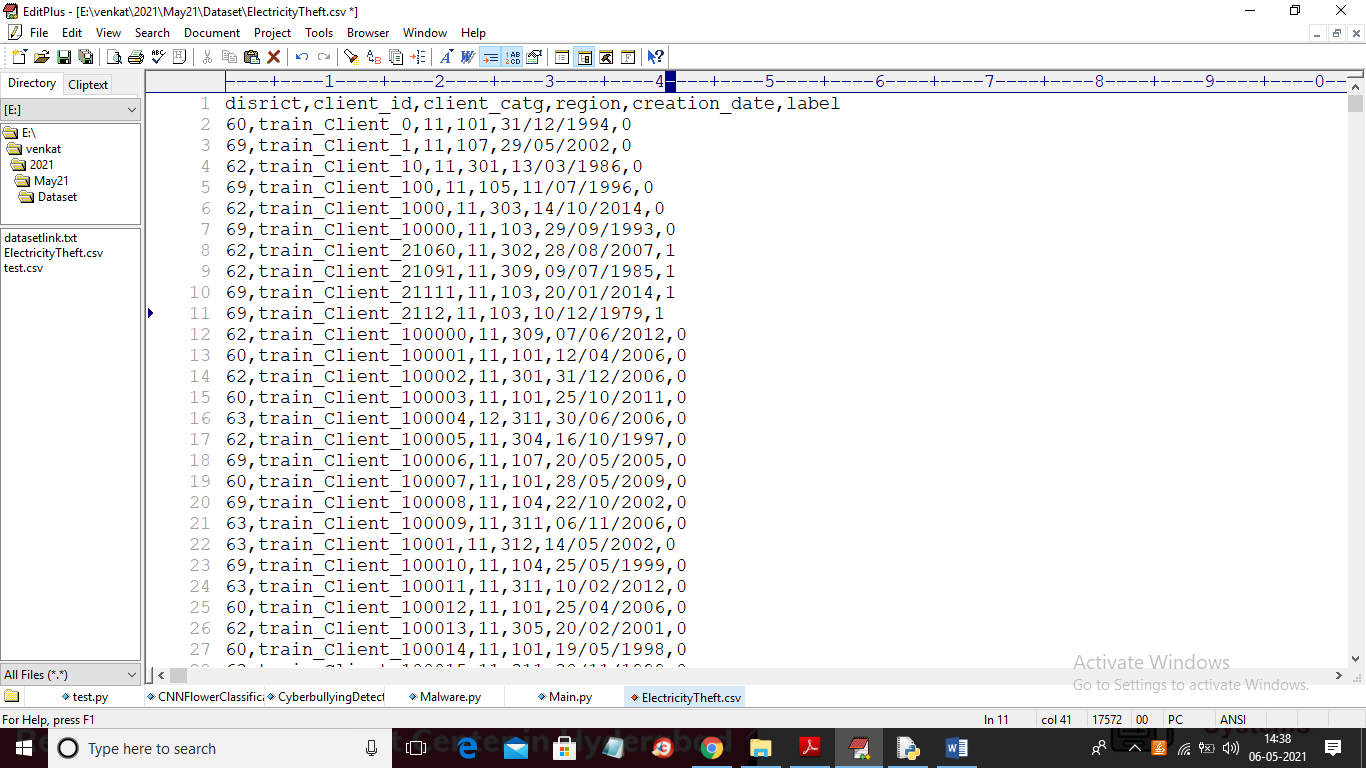
Electricity Theft Detection in Power Grids with Deep Learning and Random Forests

In this paper author is using combination of CNN (Convolution Neural Networks) and Random Forest to detect theft from electricity power grid as this theft will cause huge financial loss and disturbance in power supply. To efficiently detect theft from power grid author combining CNN and Random Forest Algorithms and after combining we are getting better prediction accuracy compare to normal algorithms. In power consumption if there is huge consumption in certain period then in dataset we will get value as 1 which indicates energy theft else we will have 0 as class label which means normal energy usage.

In propose paper author is performing following steps

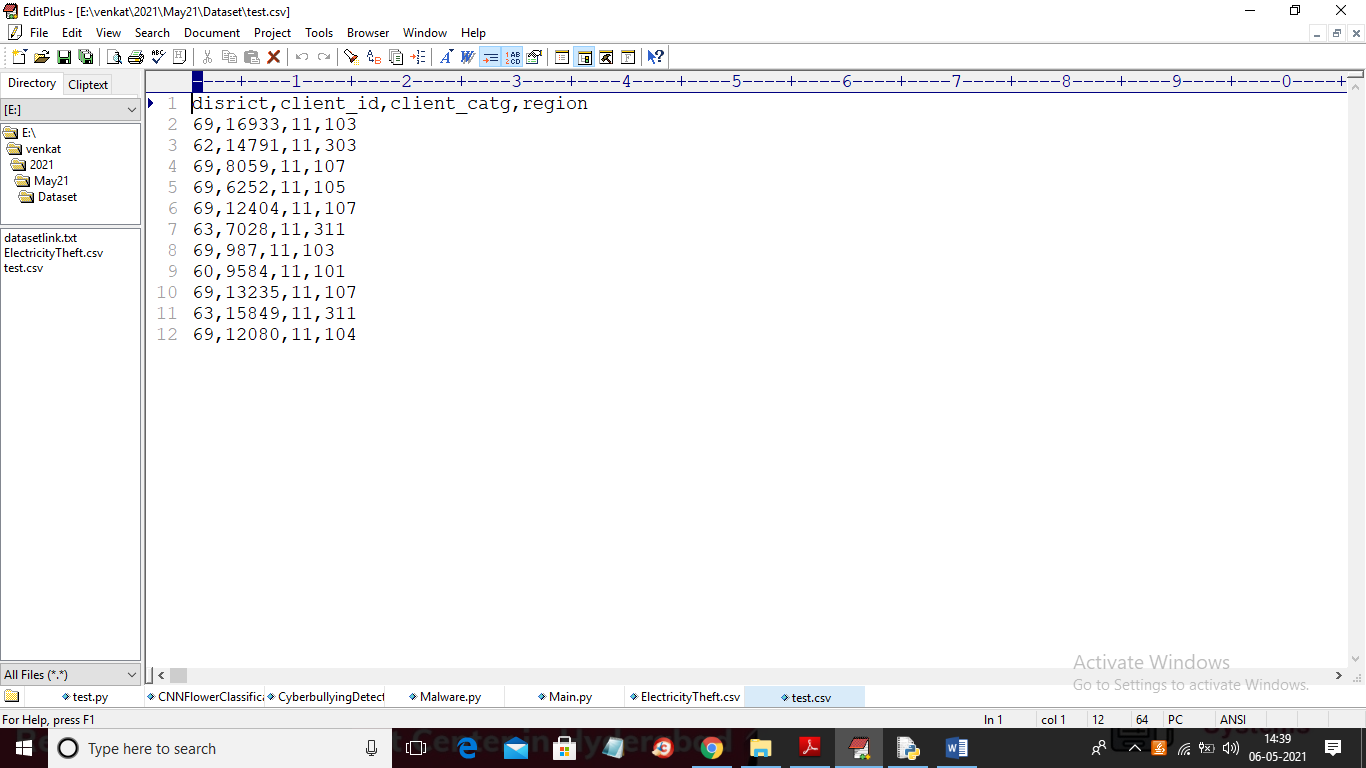
1. Reading dataset: using this module reading power consumption dataset
2. Preprocess dataset: using this module we will normalize and clean dataset by removing missing dataset
3. Train CNN Model: using this module we will train CNN with dataset and then will extract trained features from CNN and then input this trained features to random forest algorithm to build theft prediction model. To remove irrelevant features we have added DROPOUT layer.
4. Train CNN with Random Forest: using this module will train random forest with CNN features and then calculate precision, recall, FSCORE and accuracy
5. Train CNN with SVM: using this module will train SVM with CNN features and then calculate precision, recall, FSCORE and accuracy
6. Train Random Forest without CNN: Here we trained random forest on normal dataset without using CNN features and then calculate precision, recall, FSCORE and accuracy
7. Train SVM without CNN: Here we trained SVM on normal dataset without using CNN features and then calculate precision, recall, FSCORE and accuracy
8. Comparison Graph: using this we will display comparison graph between all algorithms
9. Predict Electricity Theft: Using this module we will upload test data and then CNN-RF will predict whether test records contains ENERGY THEFT or not.

Below is the dataset screen use to train all algorithms



In above dataset first row contains column names and remaining rows contains dataset values and in last column we can see values as 0 or 1 which means normal or energy theft

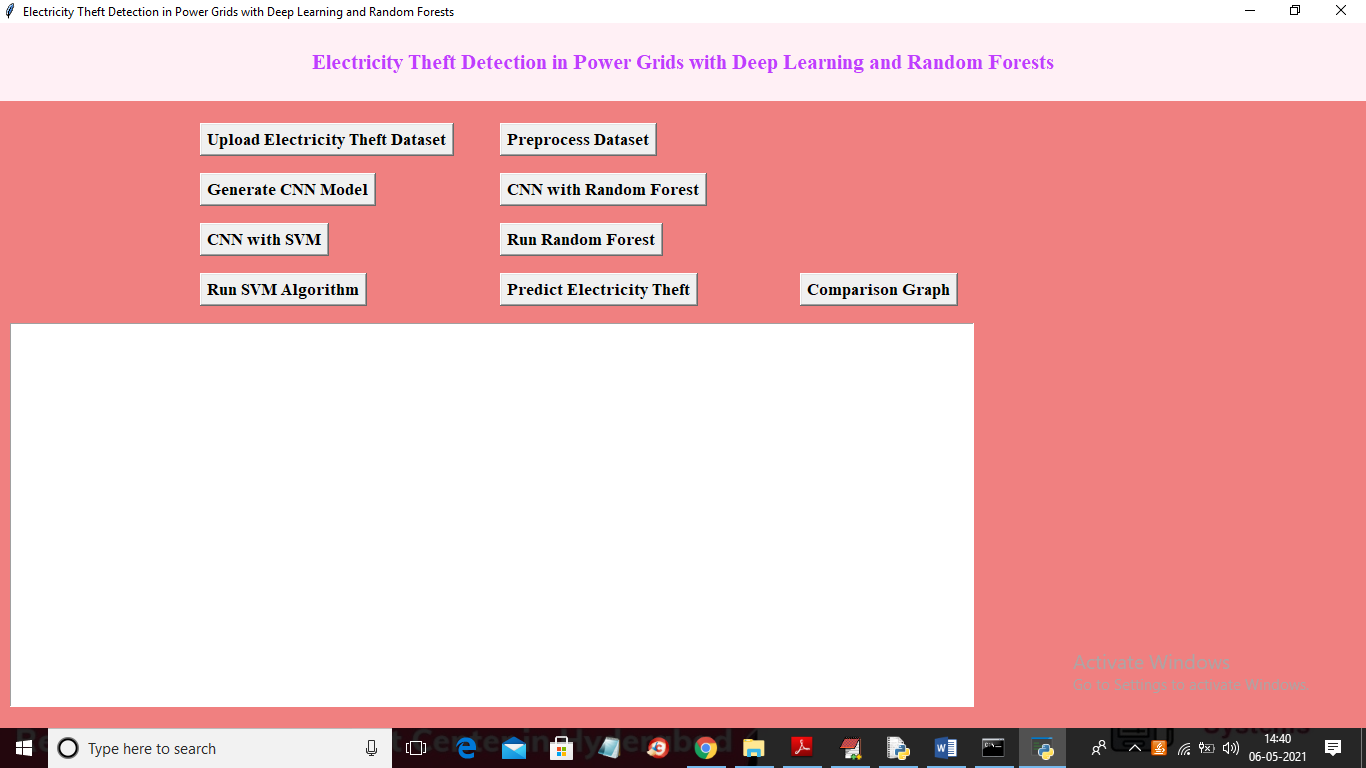
In below test data we don’t have class label as 0 or 1 and this class label will be predicted by CNN-RF



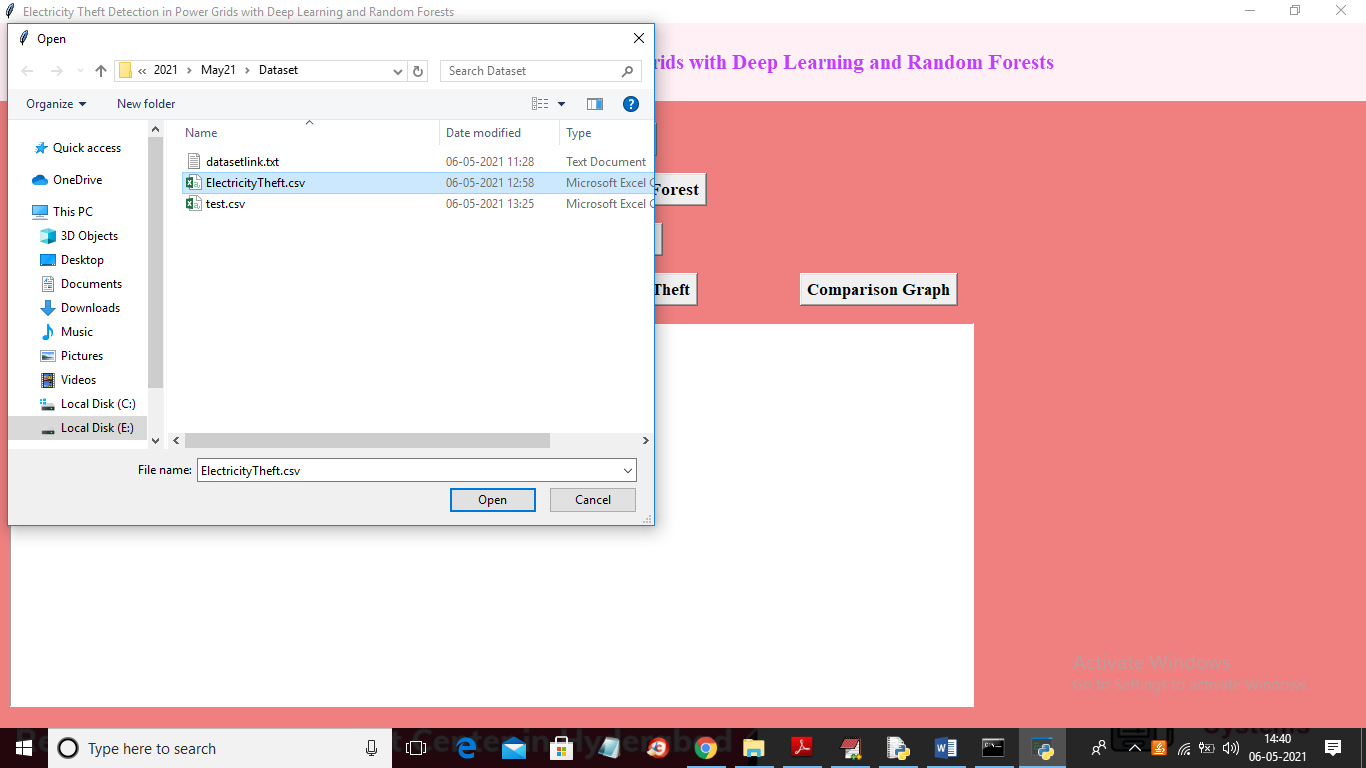
In above screen we have only energy consumption details but no class labels

SCREEN SHOTS

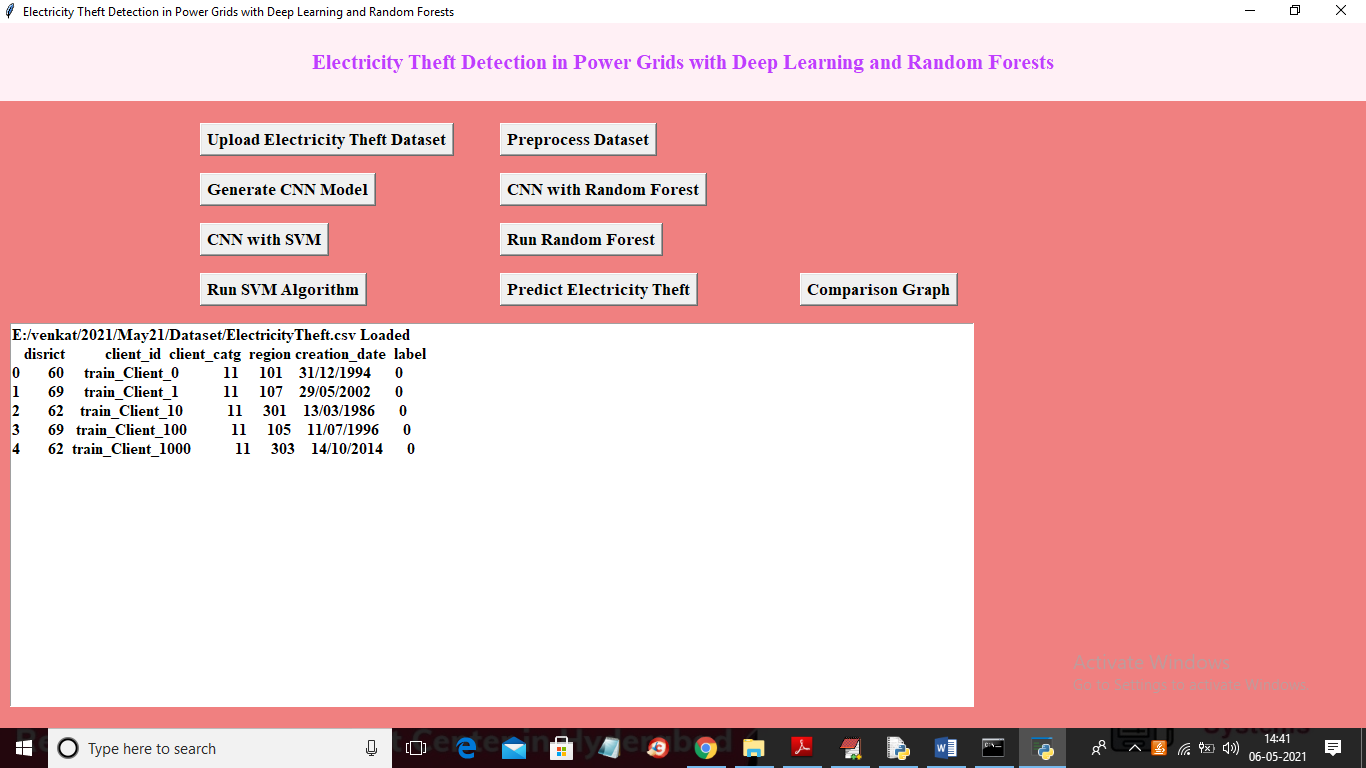
To run project double click on ‘run.bat’ file to get below screen



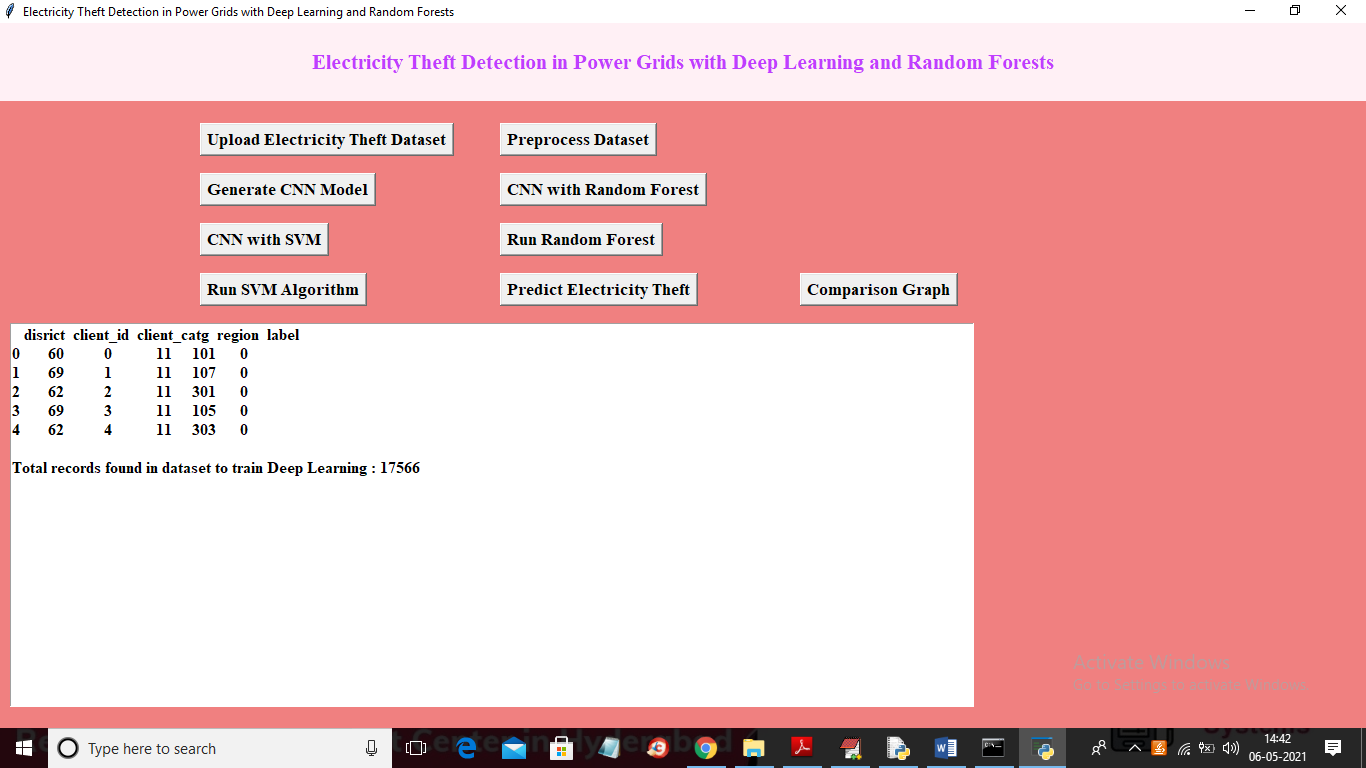
In above screen click on ‘Upload Electricity Theft Dataset’ button to upload dataset



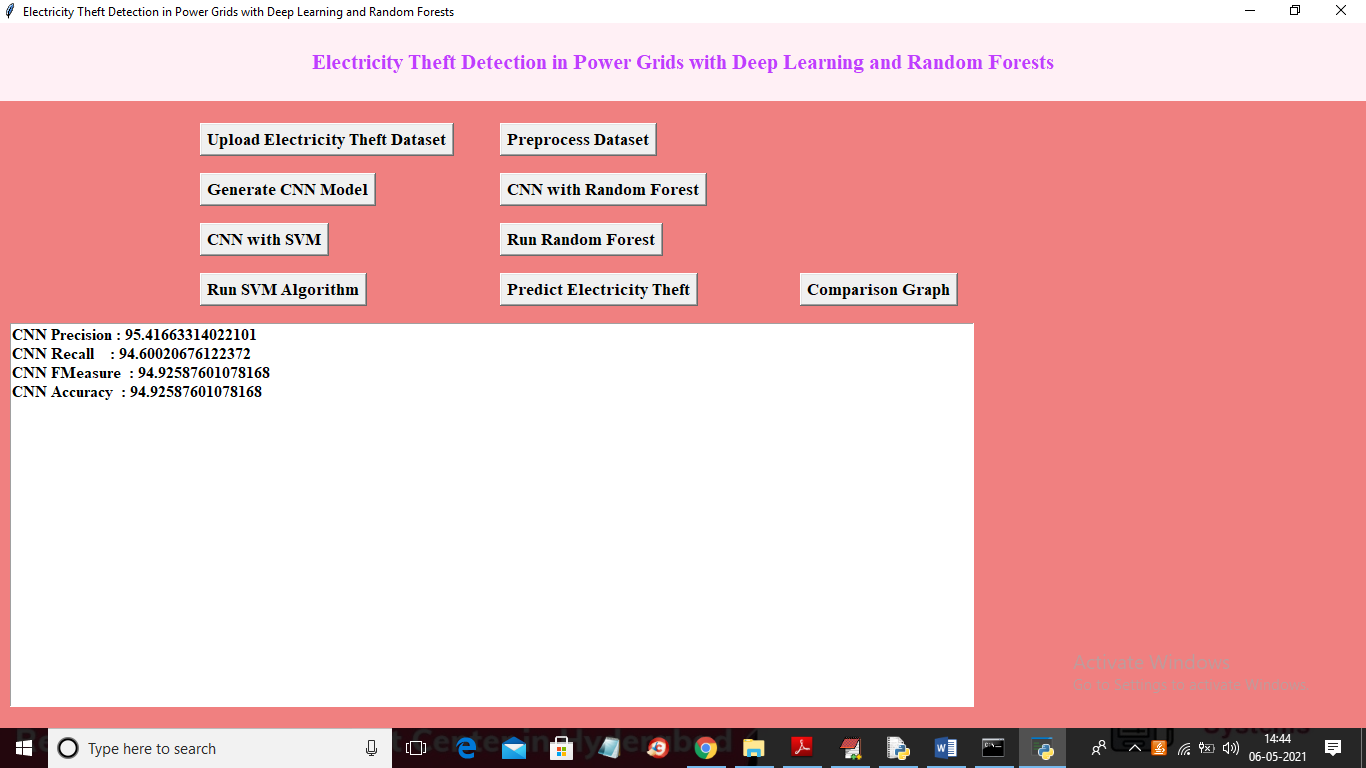
In above screen selecting and uploading ‘ElectricityTheft.csv’ file and then click on ‘Open’ button to load dataset and to get below screen



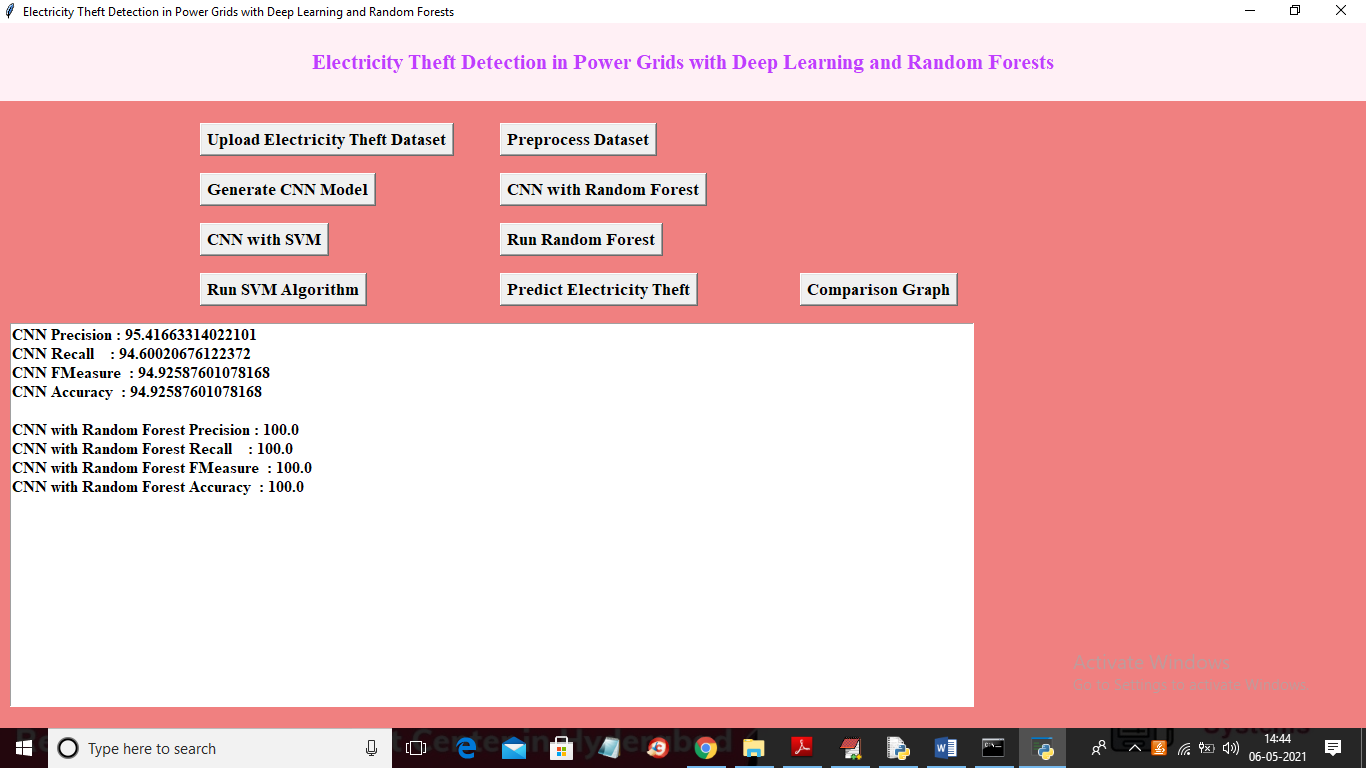
In above screen dataset loaded and displaying few records from dataset and this records contains non-numeric values and this values will not accept by machine learning so we need to convert to numeric by assigning integer ID so click on ‘Preprocess Dataset’ button to clean data



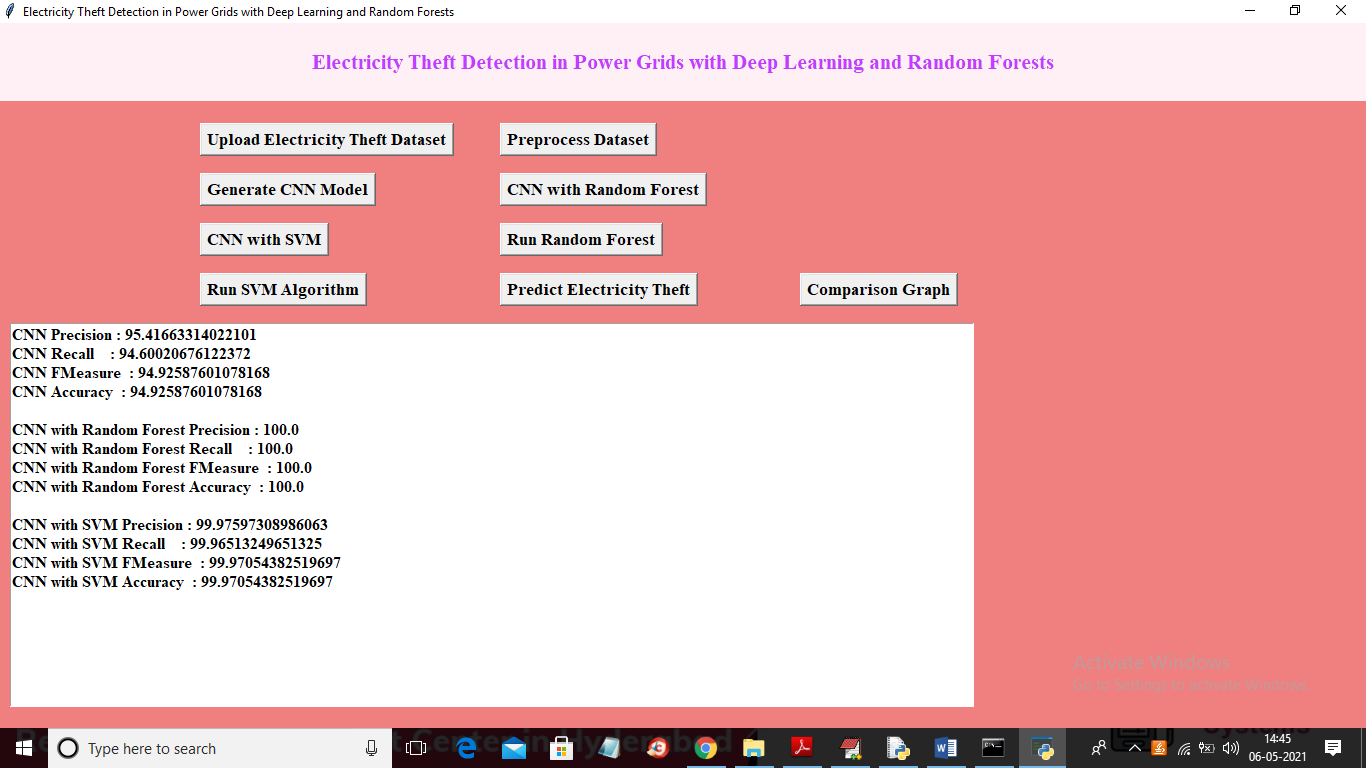
In above screen dataset converted to numeric format and now click on ‘Generate CNN Model’ button to train CNN with above dataset



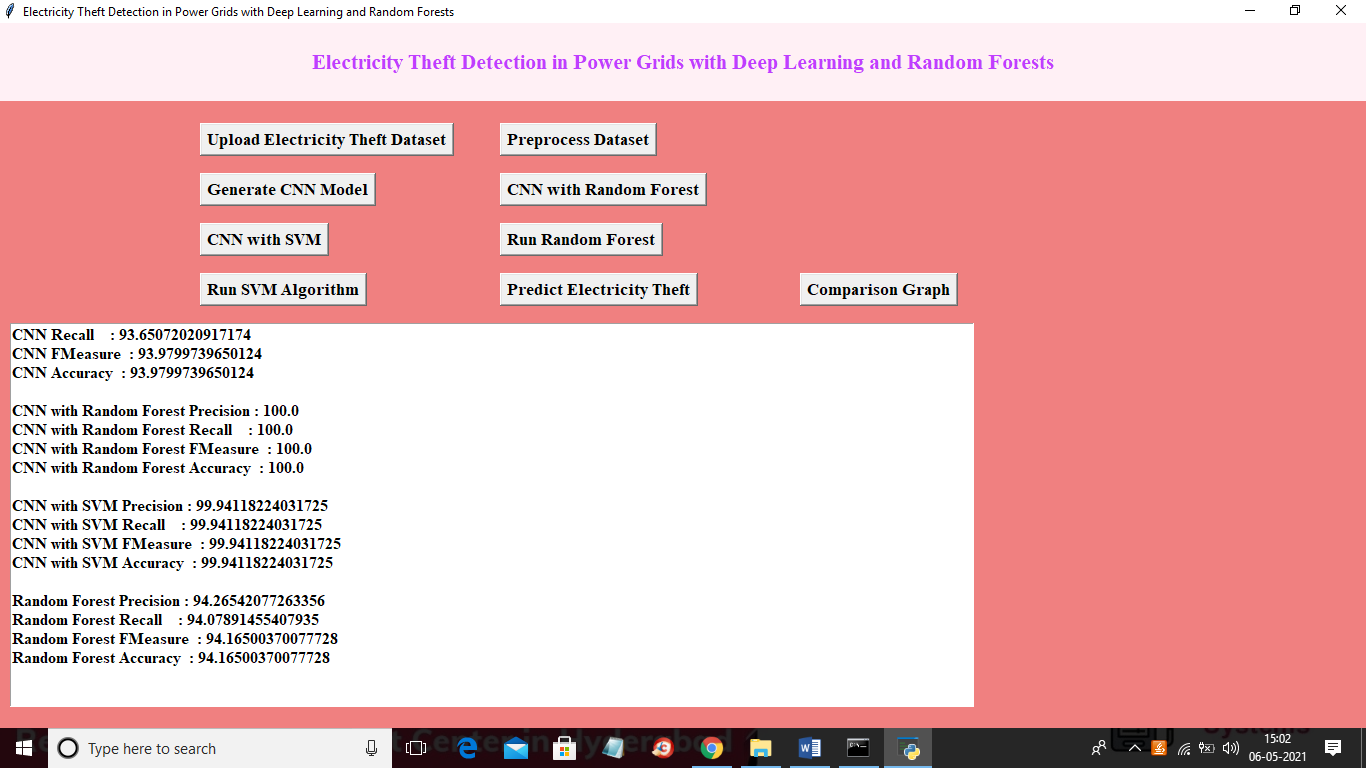
In above screen with normal CNN we got 94% accuracy and now click on ‘CNN with Random Forest’ button to train CNN with RF



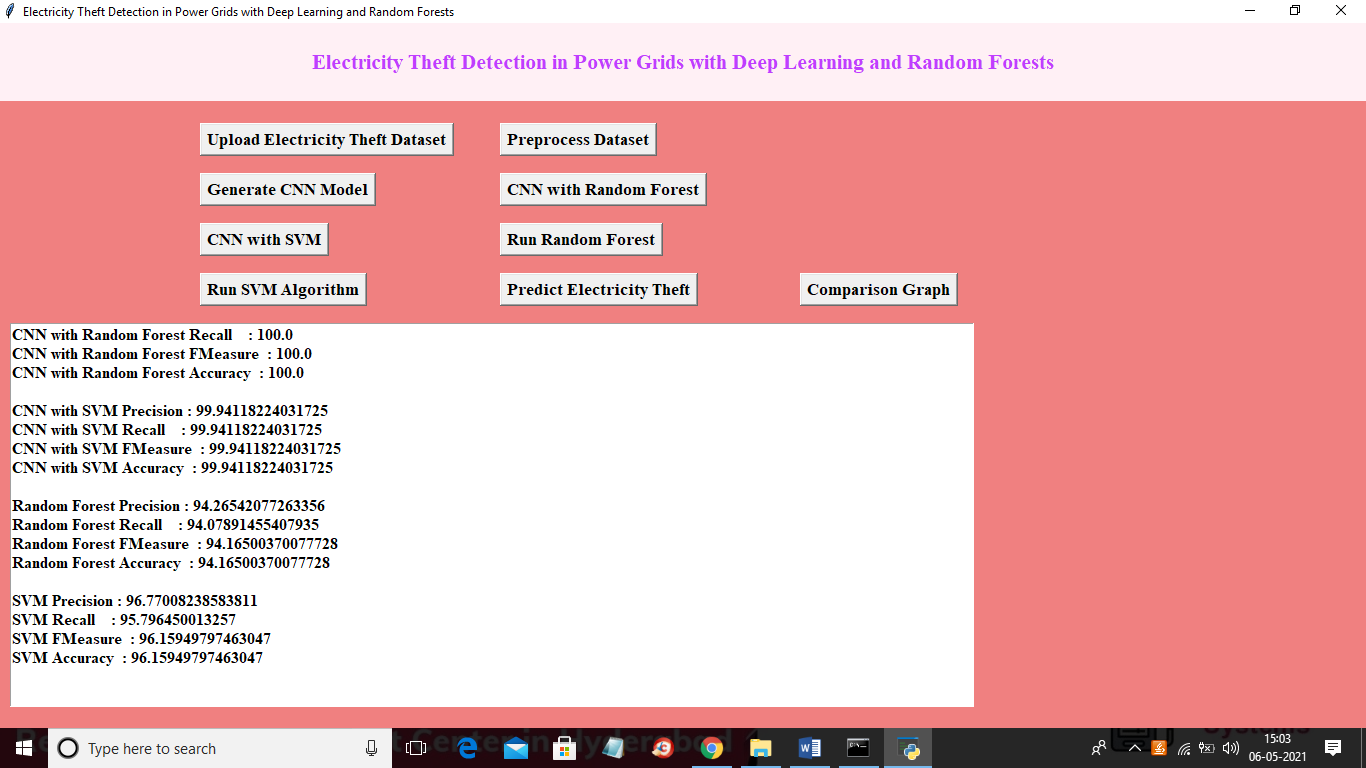
In above screen with CNN-RF we got 100% accuracy and now click on ‘CNN with SVM’ button to train dataset with CNN and SVM



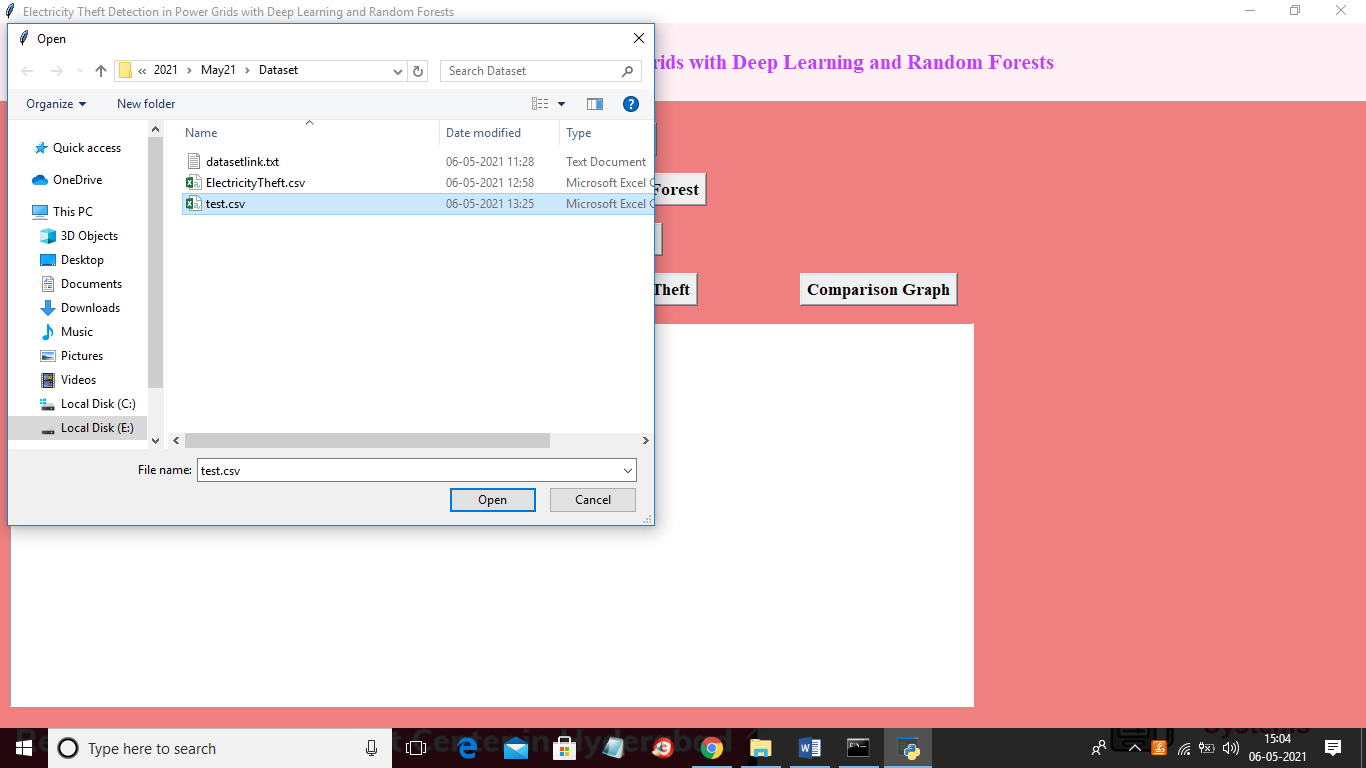
In above screen with CNN-SVM we got 99% accuracy and now click on ‘Run Random Forest’ button to train alone RF on dataset



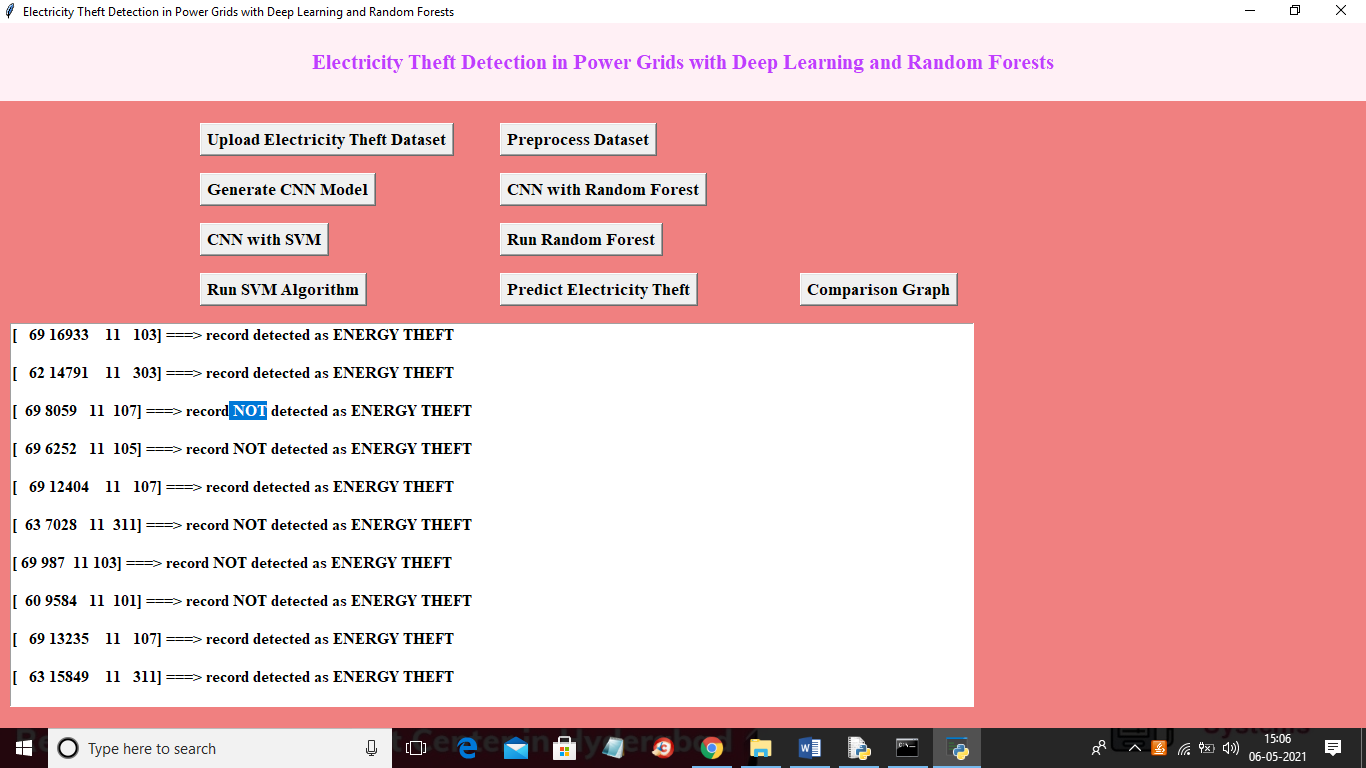
In above screen with alone Random Forest we got 94% accuracy and now click on ‘Run SVM Algorithm’ button to train alone SVM with above dataset



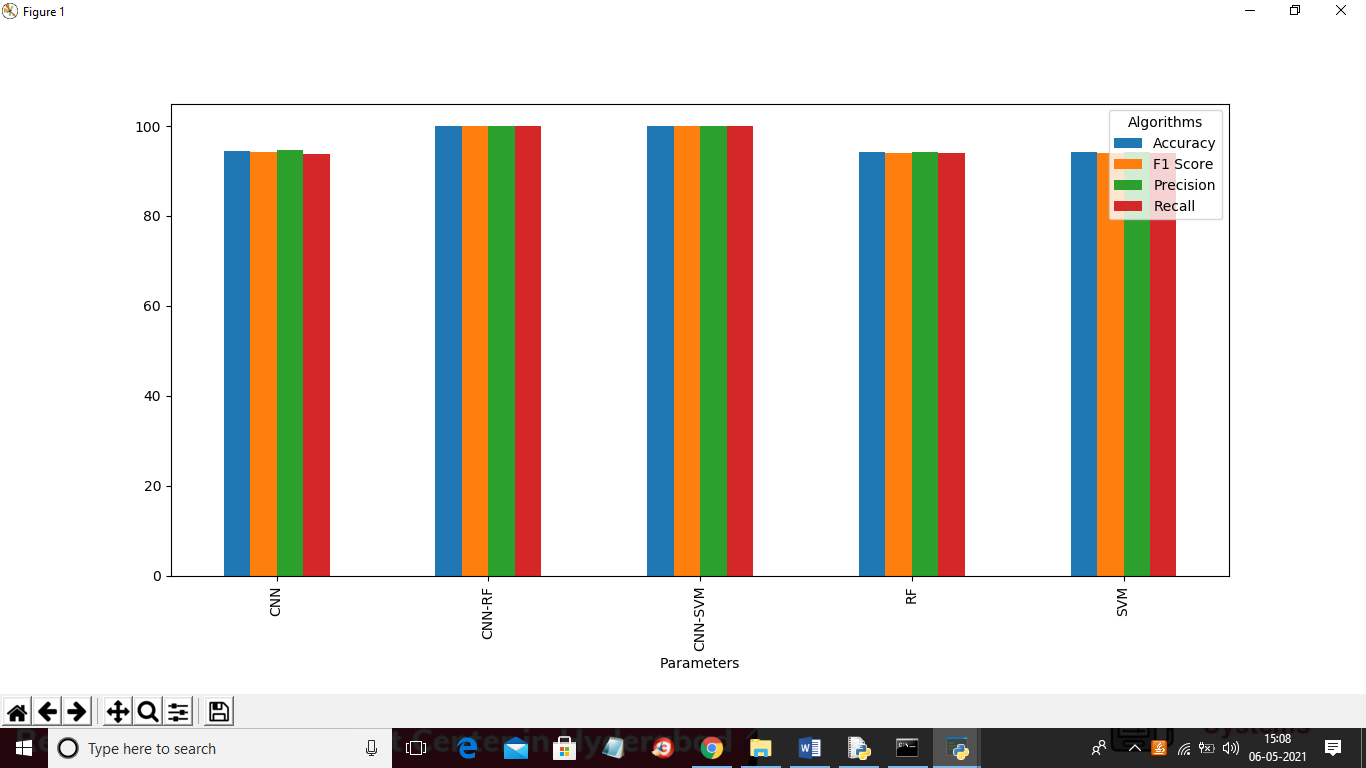
In above screen with alone SVM we got 96% accuracy and now click on ‘Predict Electricity Theft’ button to upload test data



In above screen selecting and uploading ‘test.csv’ file and then click on ‘Open’ button to load test data and to get below prediction result



In above screen in square brackets we can see test data and after square bracket we can see prediction result as ‘record detected as ENERGY THEFT’ or ‘record NOT detected as ENERGY THEFT’. Now click on ‘Comparison Graph’ button to get below graph



In above graph x-axis represents algorithm names and y-axis represents precision, recall, FSCORE and Accuracy for each algorithm and in all algorithms CNN-RF is giving 100% accuracy.