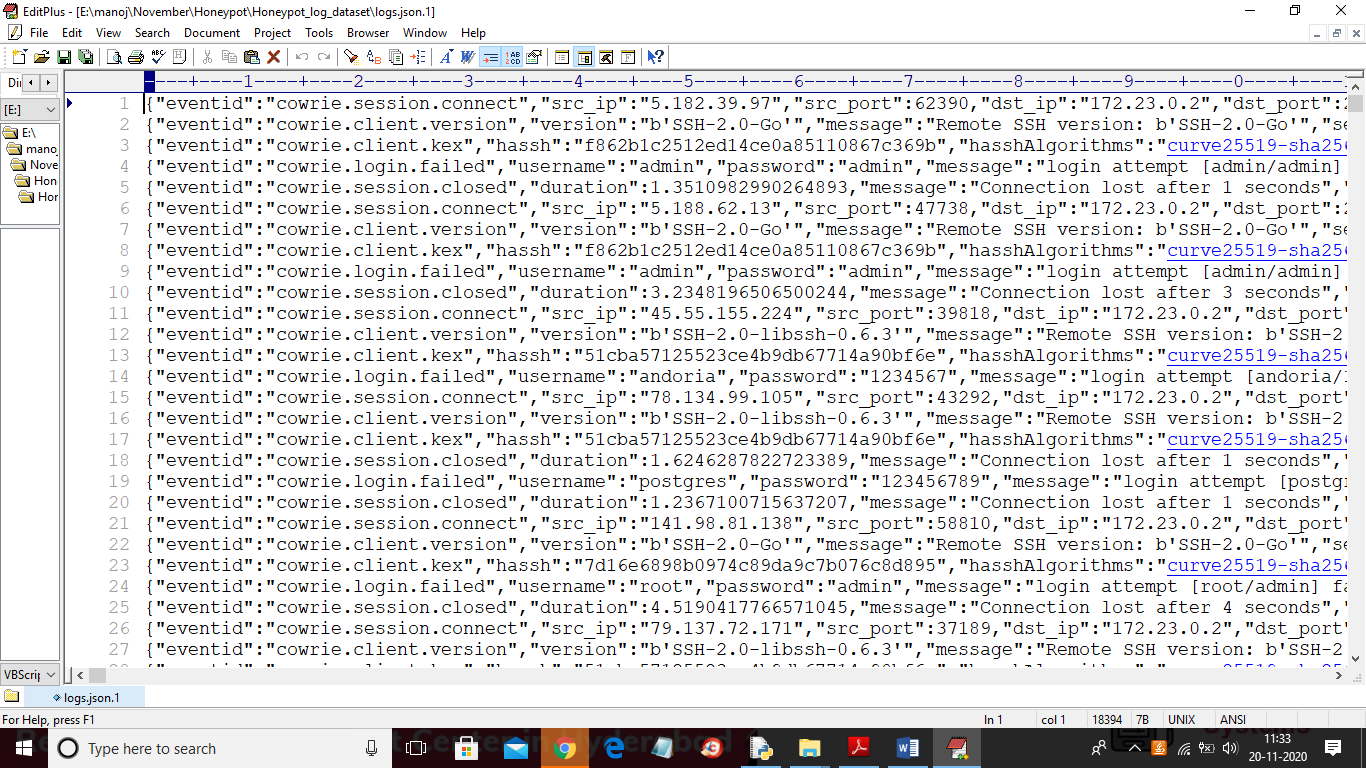
A Honeypot with Machine Learning based Detection Framework for defending IoT based Botnet DDoS Attacks

In this paper author is applying Machine Learning algorithms such as SVM, KNN, Random Forest, Decision Tree and Neural Network to detect DDOS attacks from IoT networks. IoT are small devices deployed in any environment such as battle fields, agriculture fields, healthcare hospitals etc to sense data and then send that sense data to destination server using internet connection. IoT are small devices and can be attacked by attacker to sense and send corrupted data and to provide security from such attacks heavy security algorithms cannot be installed on IoT devices so author is deploying Honeypot server which can run between centralized server and IoT networks and whenever user send any request then honeypot server will process that request and if request is genuine then it will forward that request to centralized server or IOT networks. If user send malicious request by giving wrong username and password then honeypot will server dummy response to user and try to extract more information such as IP address, mac address and then inform to centralized server and IOT network to be aware of such IP address and mac address.

In existing technique honeypot using signature based attack detection which is not efficient so author is deploying machine learning framework at honeypot server to predict whether request is normal or contains attack signature. In propose technique ML algorithms will be trained with previous data and then this trained model can be used to detect attacks from old or new request signature and this detection will solved ZERO-DAY Distributed Denial of service (DDOS) attacks.

In propose work author using honeypot server and IOT devices to capture data and this data will be used to train ML algorithms but we don’t have any IOT devices so we are using IOT dataset to trained ML algorithms. In propose work we are using SVM, Random Forest, K-Nearest Neighbours, Decision Tree and Neural Networks. In all algorithms SVM, KNN and Neural network is giving best performance. About all algorithms you can refer internet.

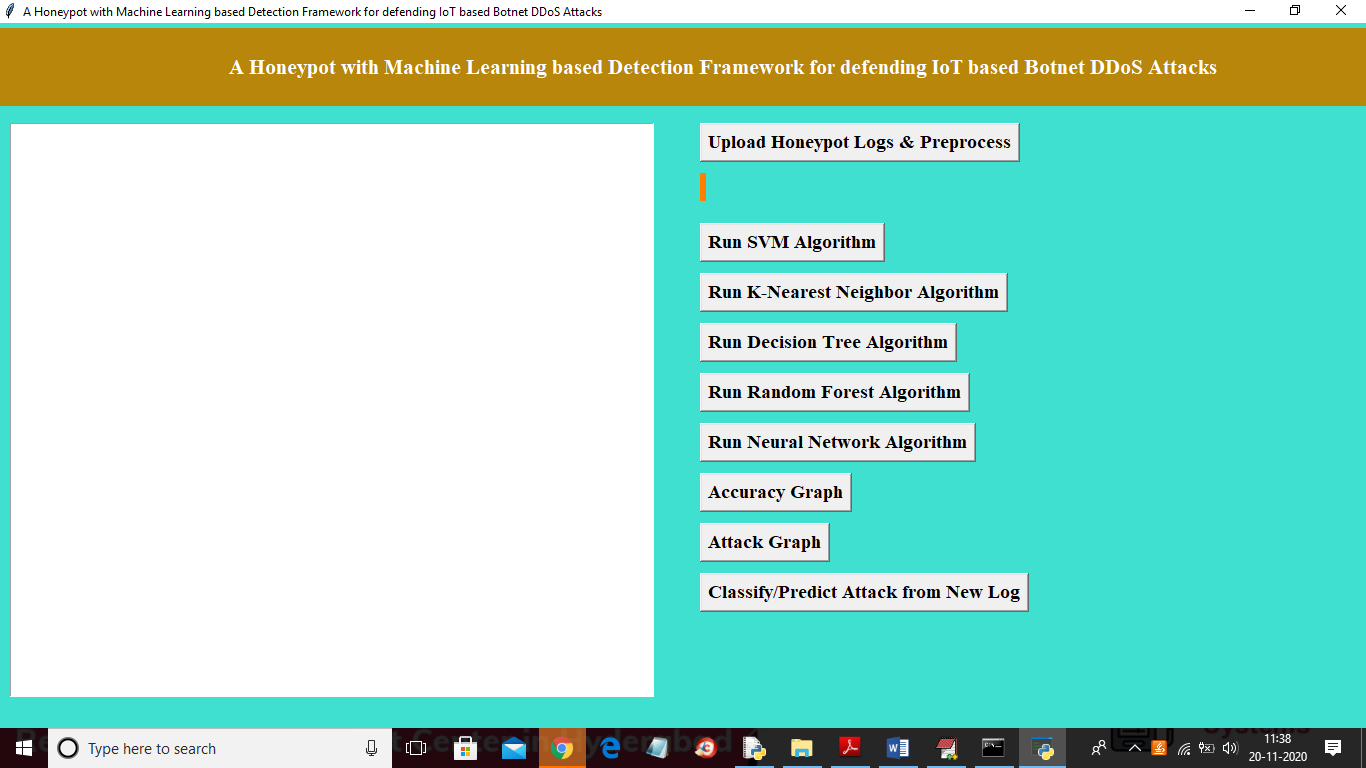
Below is the IOT dataset screen shots



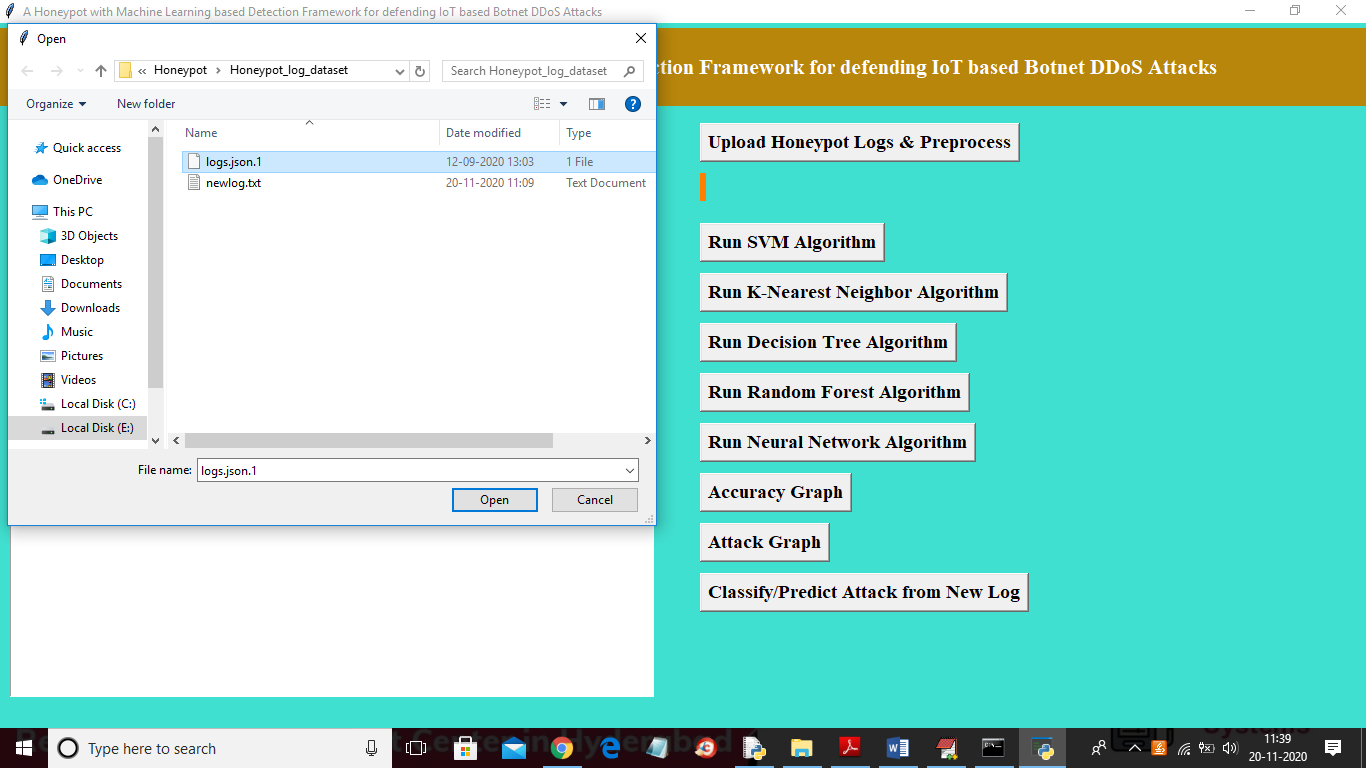
In above Honeypot IOT log dataset we have complete details such as event id, sender IP, port no, destination IP and many other signatures and using above dataset we will train machine learning algorithms. In above dataset we can see we have some request with login failed which will consider as attack and if in new request such commands appear then ML predict it as attack. After training ML with above dataset then we will upload new test data and then ML will predict whether new data contains normal or attack signature. Above dataset cannot be used to train ML so we will pre-process dataset to convert it into features. Above dataset you can see inside ‘Honeypot\_log\_dataset’ folder

SCREEN SHOTS

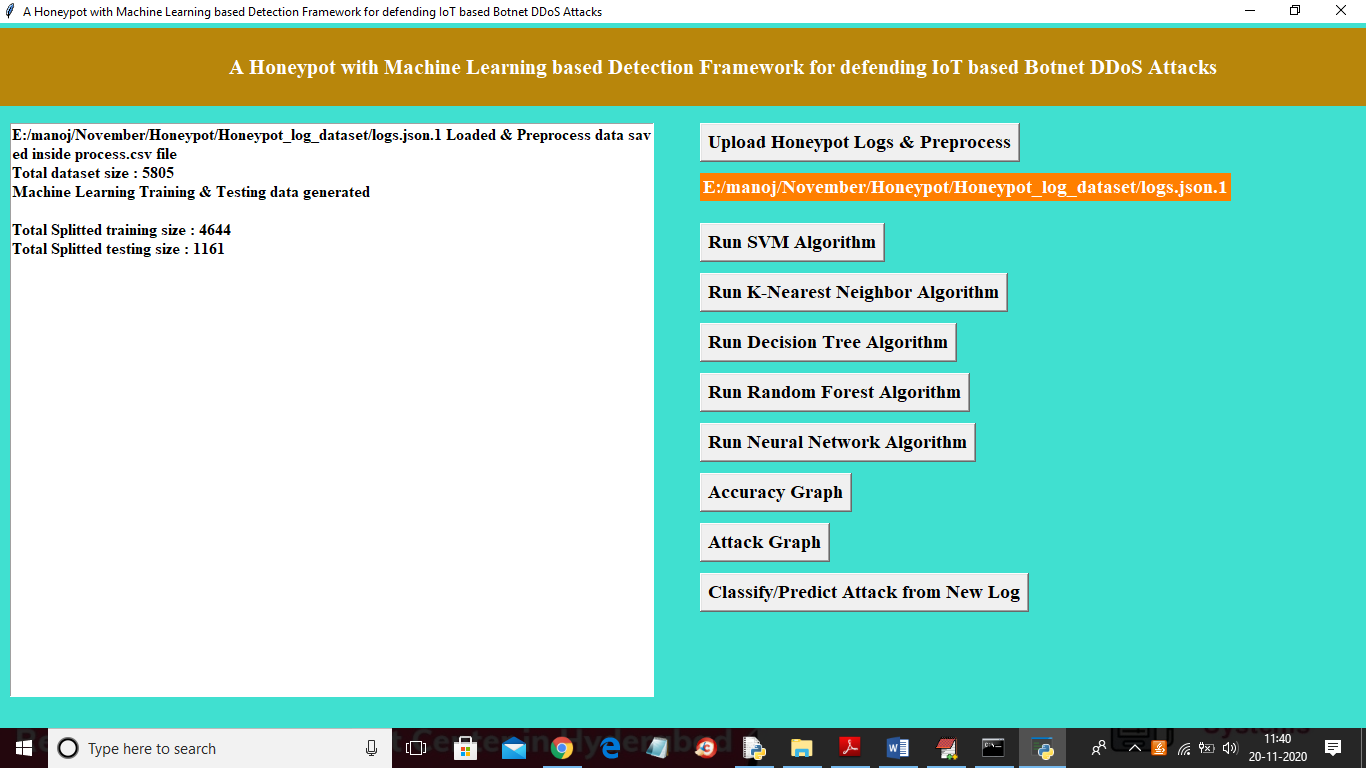
To run project double click on ‘run.bat’ file to get below screen



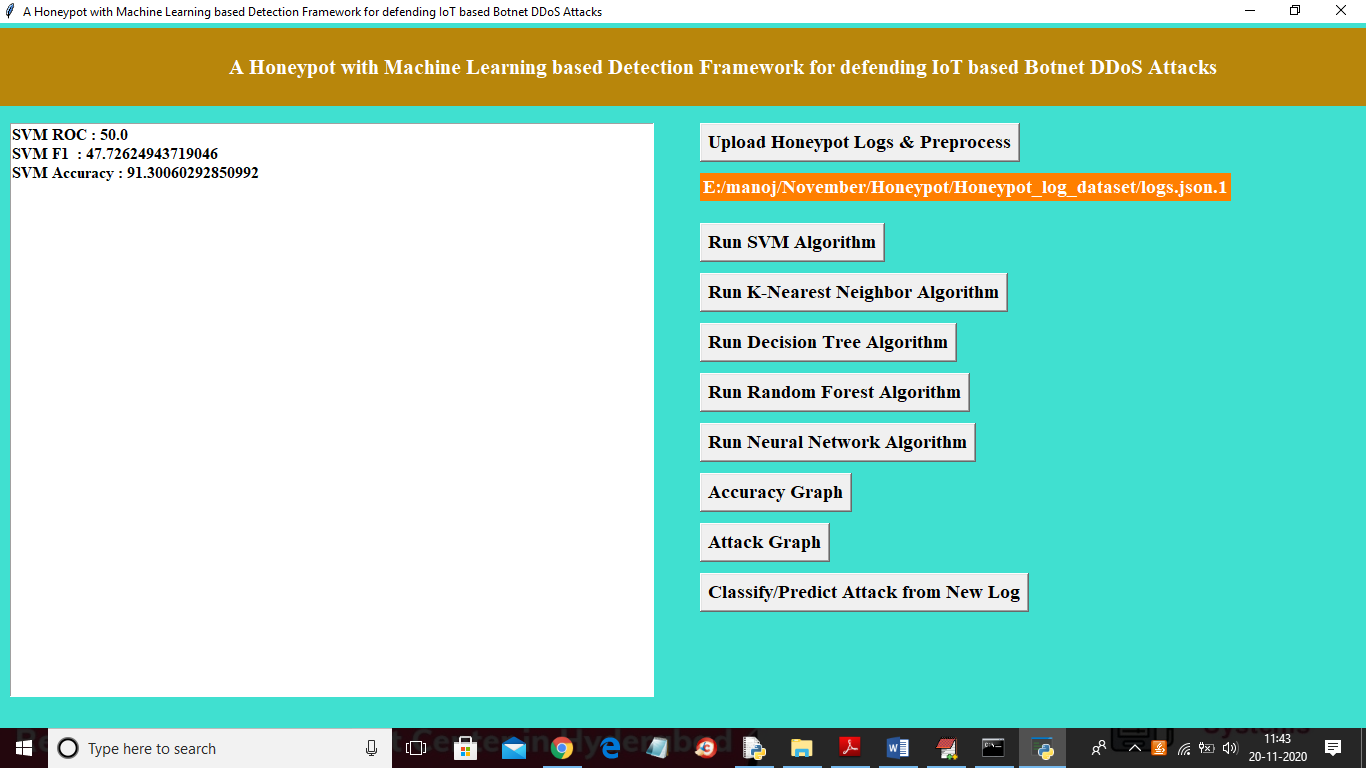
In above screen click on ‘Upload Honeypot Logs & Preprocess’ button to load log dataset and then pre-process data to convert to features



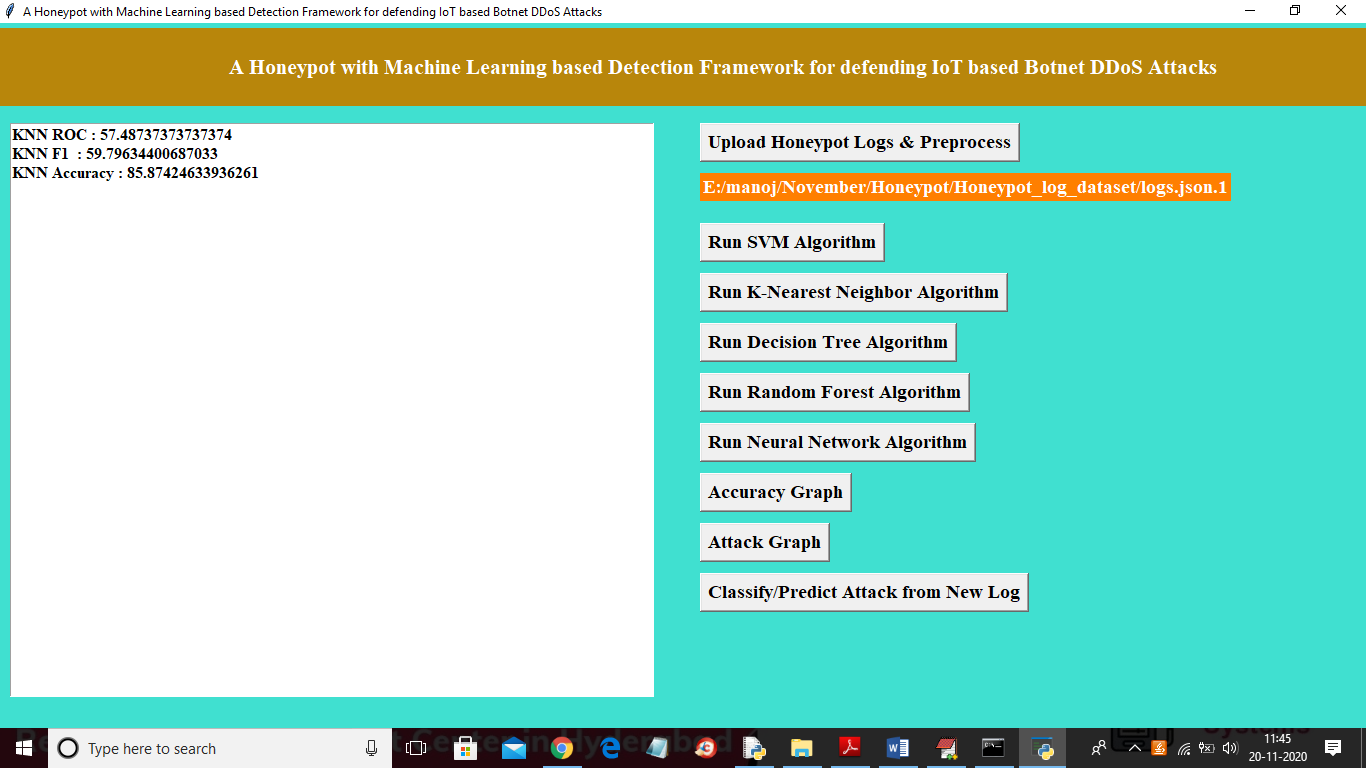
In above screen uploading ‘logs.json.1’ log file and then click on ‘Open’ button to load dataset and to get below screen



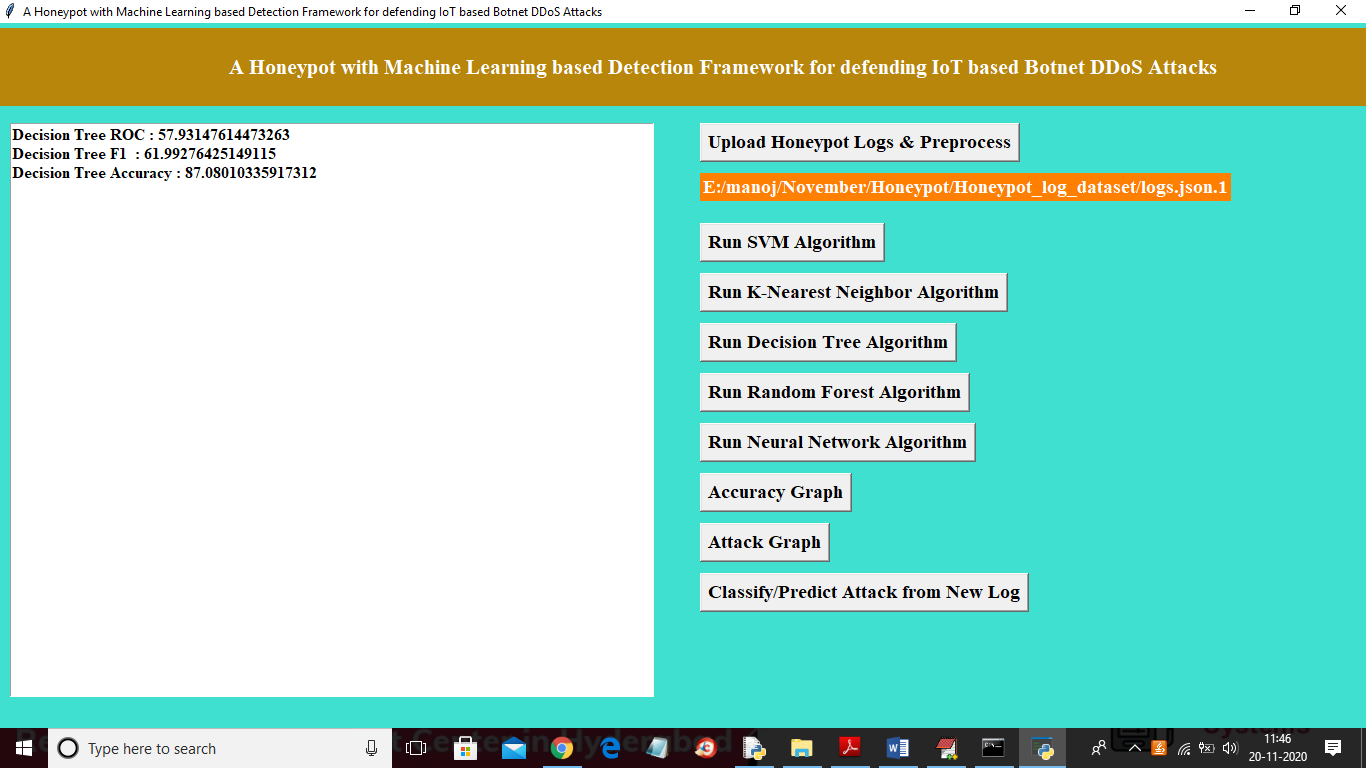
In above screen we can see dataset contains 5805 records and application split that data into train and test part and application using 4644 (80% dataset records) for training and 1161 (20% dataset records) for testing. After building model on 80% records then ML apply 20% data on trained 80% model to predict request type as normal or attack. From 20% if ML predict 18% records correctly then 18/20\*100 will give ML prediction accuracy performance. Now in above screen both train and test data is ready and now click on ‘Run SVM Algorithm’ button to train SVM model and calculate its accuracy



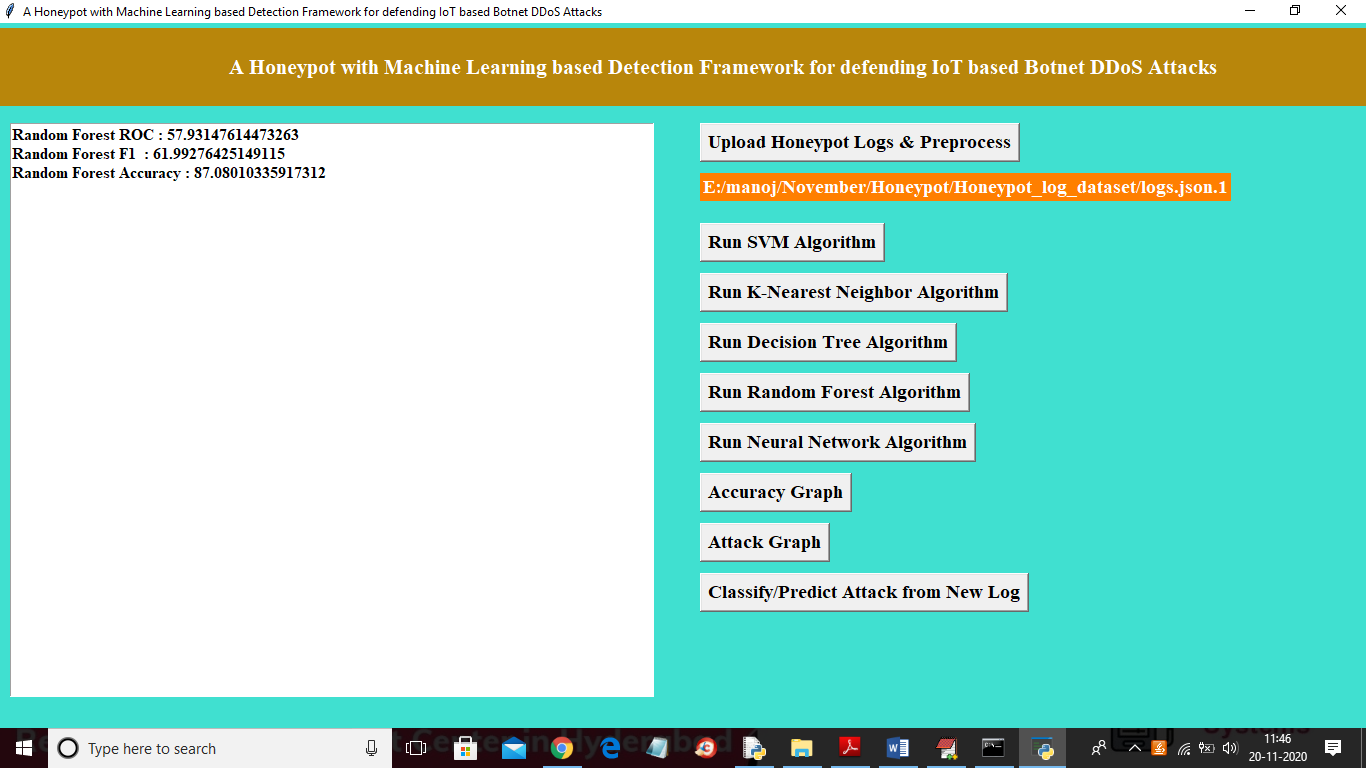
In above SVM prediction accuracy is 91% and now click on ‘Run K-Nearest Neighbor Algorithm’ button to train KNN model and calculate its accuracy



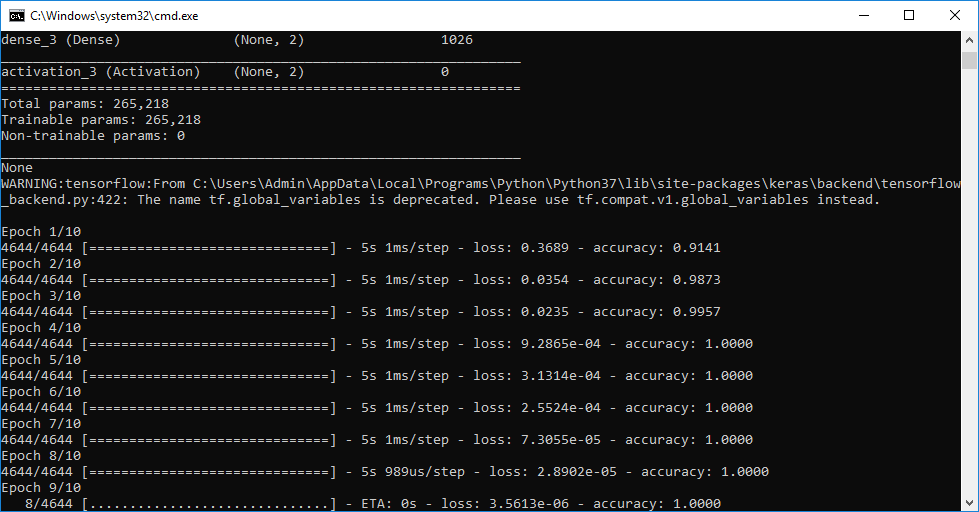
In above screen KNN got 85% accuracy and now run Decision Tree



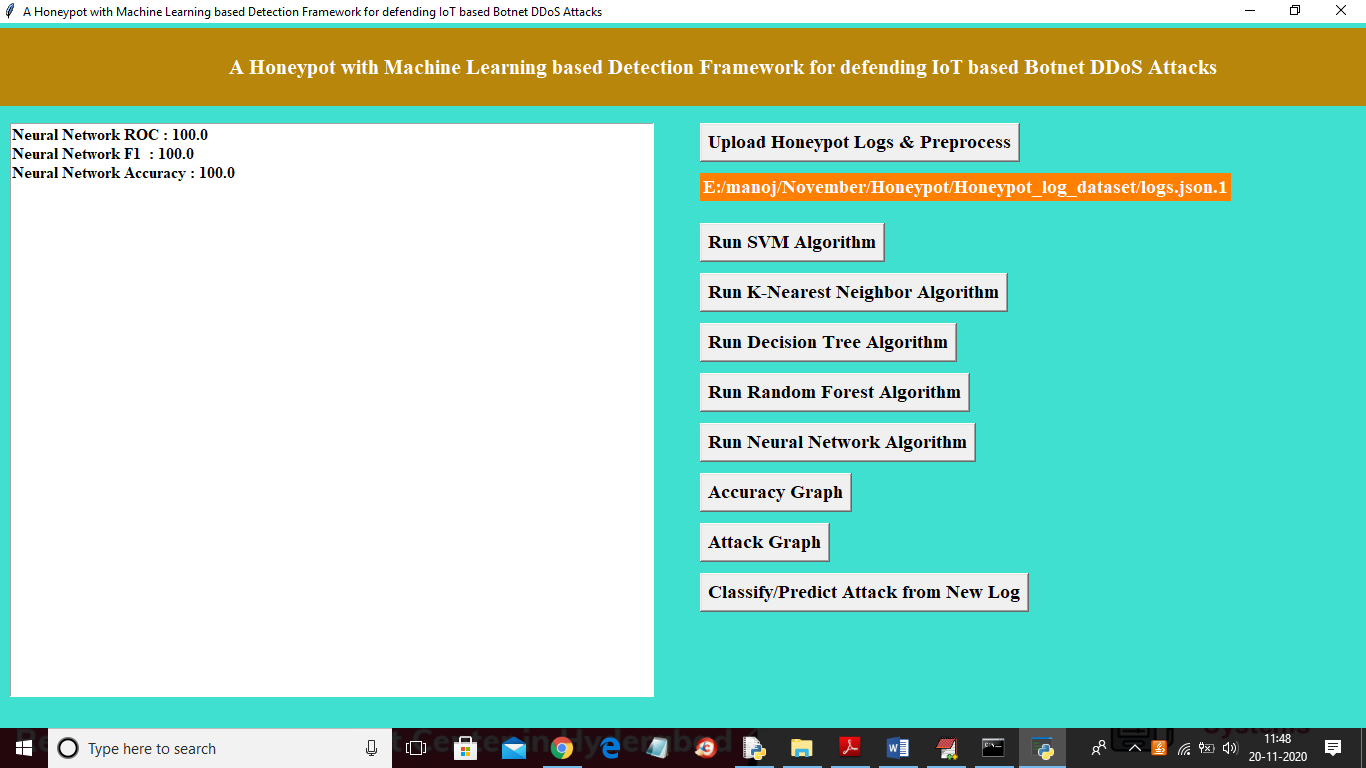
In above screen decision tree got 87% accuracy and now run random forest



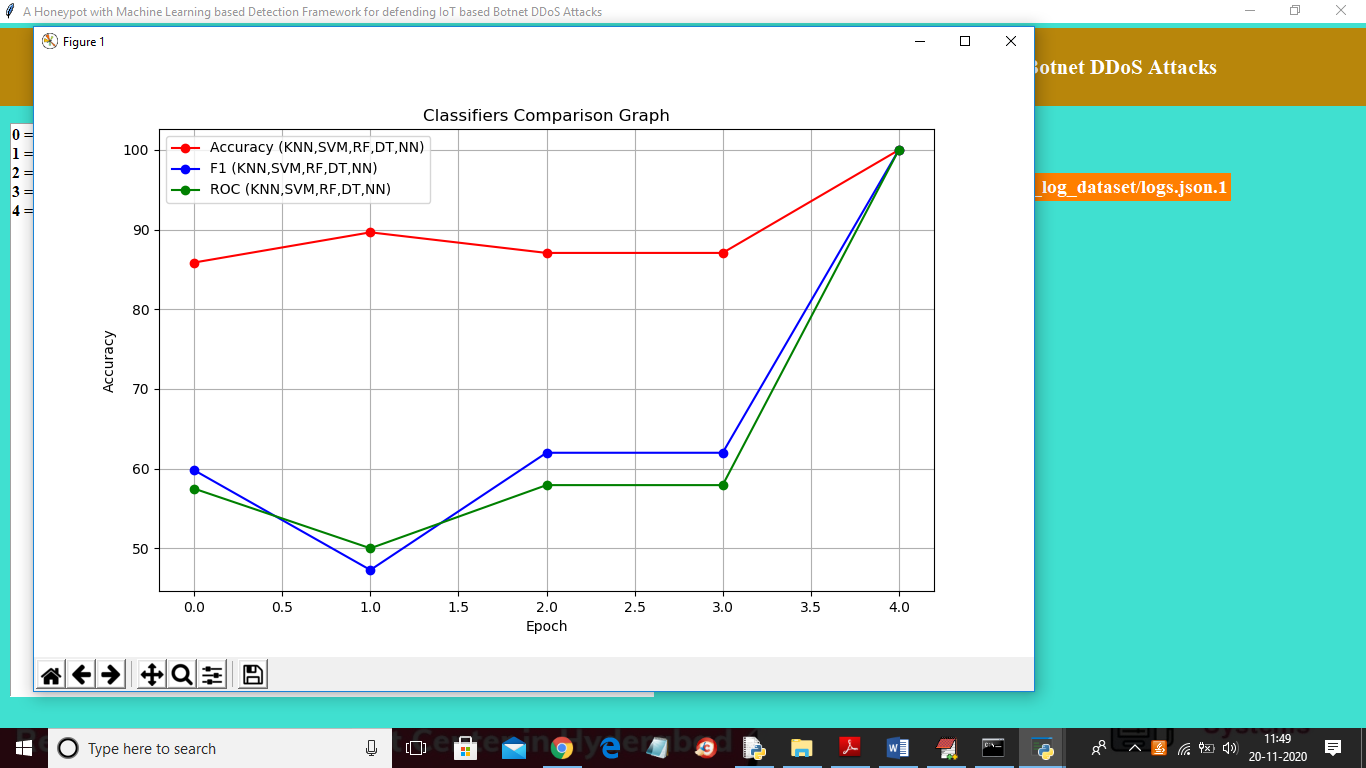
In above screen random forest also got 87% accuracy and now run neural networks



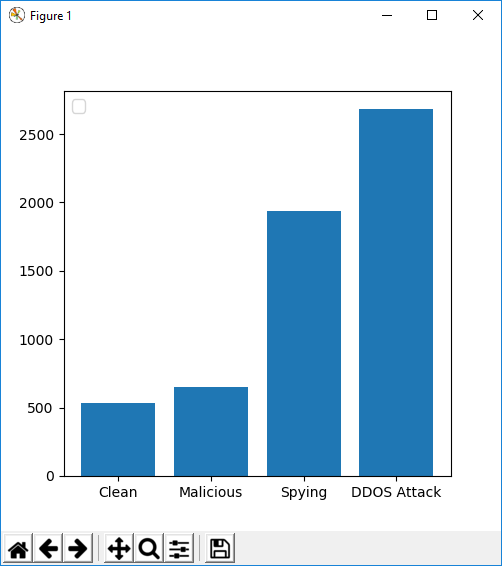
In above screen we can see neural networks start filtering dataset at each EPOC or iteration to get better prediction accuracy and in above screen we can see at first iteration accuracy is 0.91 \* 100 = 91 and at 9th iteration its increase to 100%.



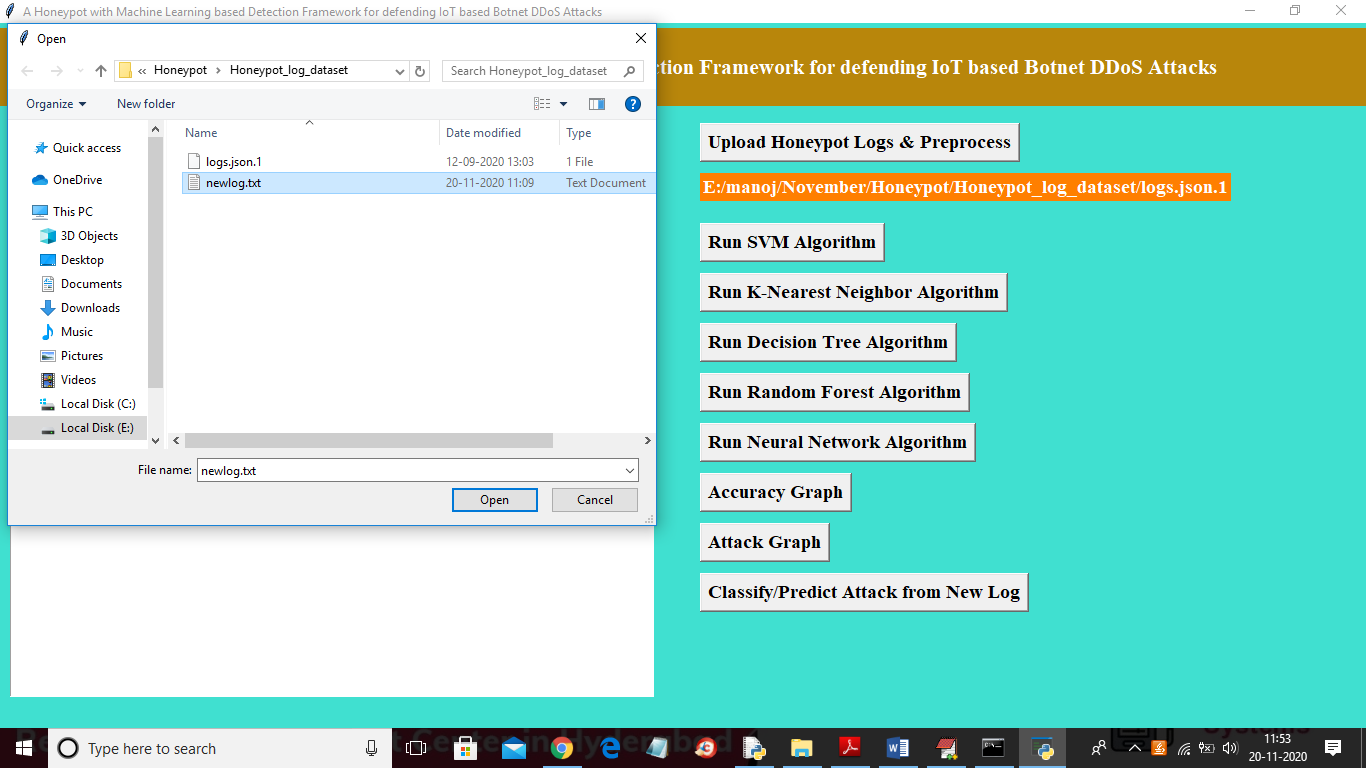
In above screen we can see neural network got 100% prediction accuracy and in all algorithms neural network is giving best performance. Now click on ‘Accuracy Graph’ button to get below graph



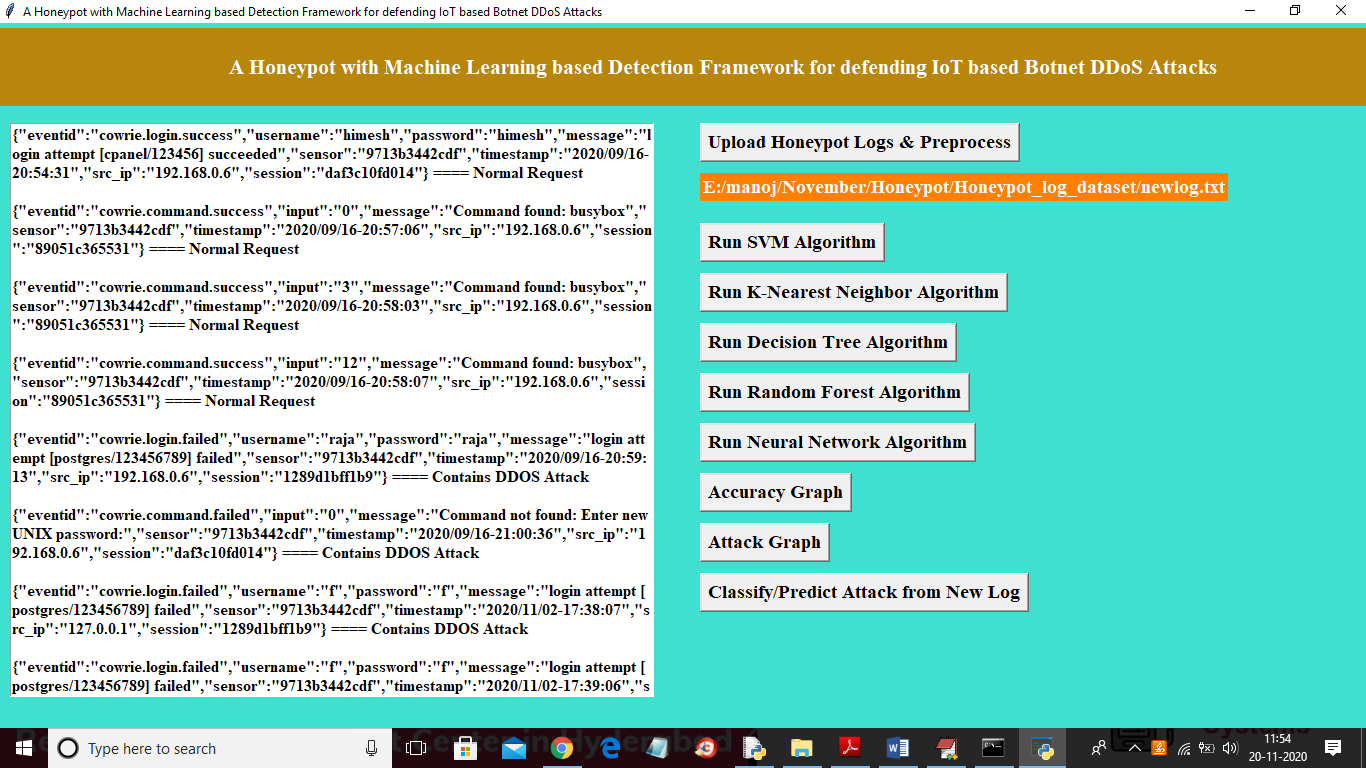
In above graph x-axis represents algorithms as KNN, SVM, RF, DT and NN and y-axis represents accuracy and in above graph red line refers to accuracy and blue line for FSCORE and green line for ROC value. In above graph each point refers value for one algorithm and last point is for NN which is having high performance. Now click on ‘Attack Graph’ button to get below graph



Above graph x-axis contains request type and y-axis contains count and this attack graph obtained from honeypot log dataset. From all request honey pot received more number of DDOs attacks. Now click on ‘Classify/Predict Attack from New Log’ button to upload new log dataset and then ML will apply on new log dataset to predict whether new log contains attack or normal request



In above screen uploading ‘newlog.txt’ file and then click on ‘Open’ button to get below prediction result



In above screen ML analyse each request and then mark that request signature as normal or DDOS attack. At each request line after equals to symbol we can see ML detection result. In above screen scroll down to view all request result

