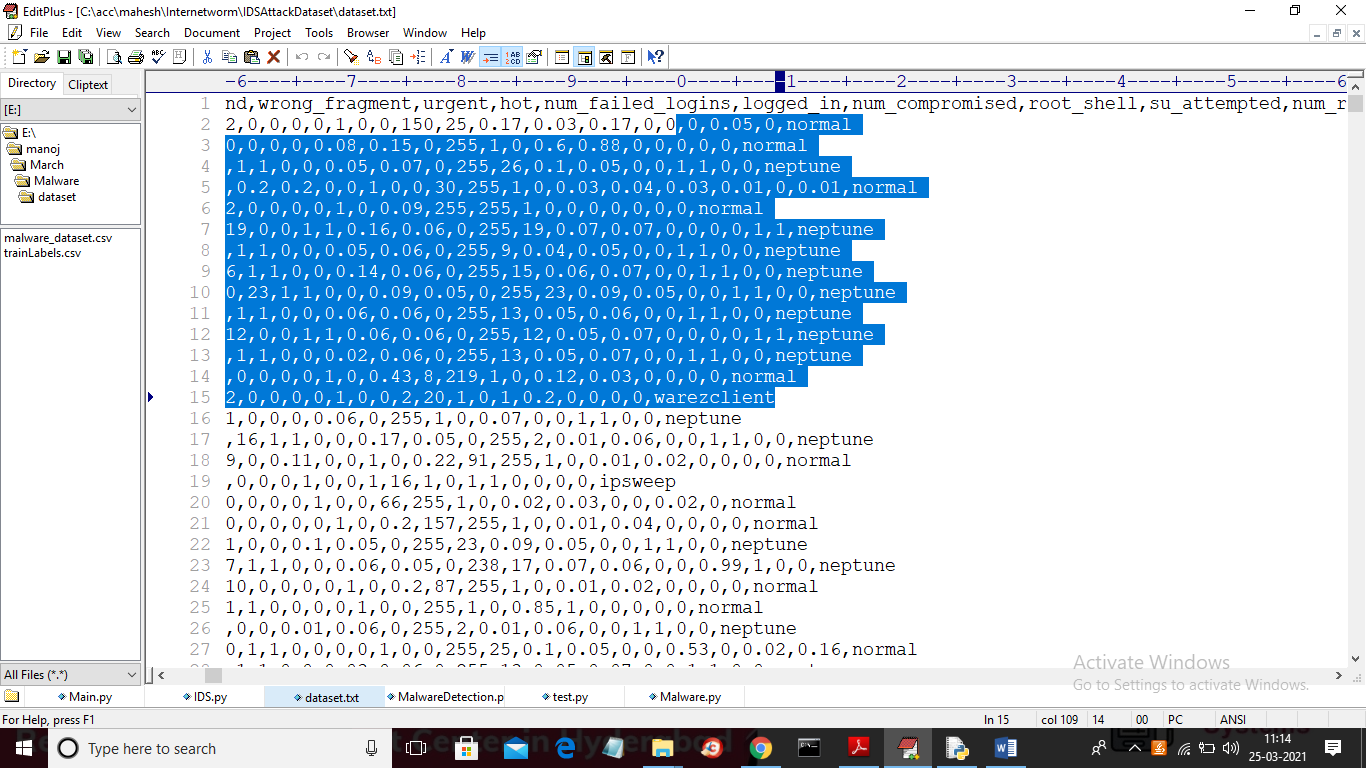
Internet Worms and its Detection

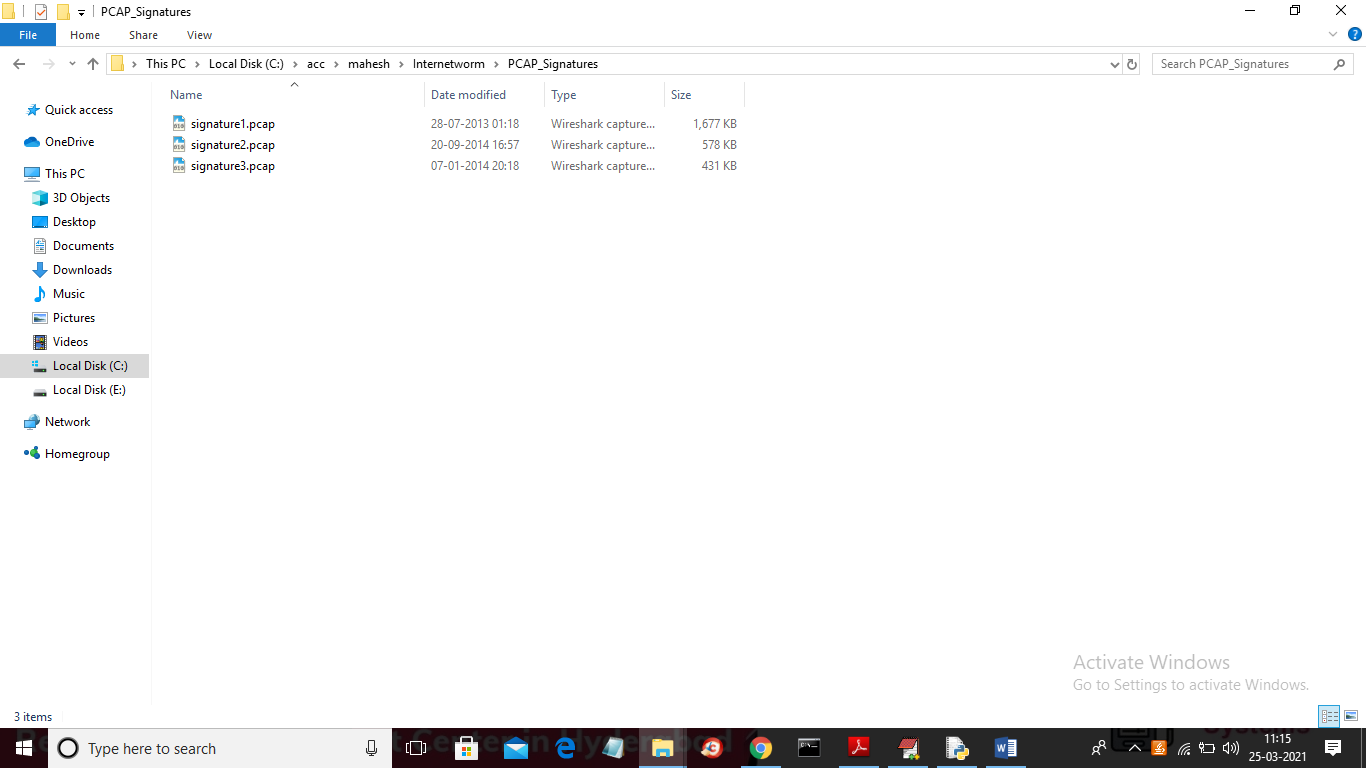
In this paper author is describing and giving brief details on different worms and how much they caused loss to computer society. Internet worms are malicious program which get downloaded in to user’s computer through internet and get installed and then start corrupting user’s files or steal user information and send to attacker. To avoid such detection various techniques has been introduced such as

1. Signature based detection which analyse internet traffic signature and then matched with predefine rules to identify whether traffic contains normal or attack signature and to analyse this signature we can use PCAP (traffic capture) files.
2. Detection through Honeypot logs: Honeypot is a server which sits between server and user request and if user send any malicious request then Honeypot will log all such request and later this logs will get inspected to block all such attackers IP ADDRESS
3. Netflow Based: This technique also inspect UDP and TCP signatures and then verify whether request contains normal or attack signatures
4. New Approach Classification Models: Here using previous datasets machine learning algorithms such as Random Forest, Decision Tree and Bayesian Networks will get trained and this trained model will use to predict attack or normal signatures from newly coming requests.
5. Network Based Detection: In this technique also intrusion detection (IDS) dataset will get trained with machine learning algorithms such as Random Forest, Decision Tree and Bayesian Networks and then this model will be used to predict attack or normal label from newly coming requests.

To detect attacks with above techniques I am using PCAP (packet capture) dataset and intrusion dataset and below screen showing dataset of IDS



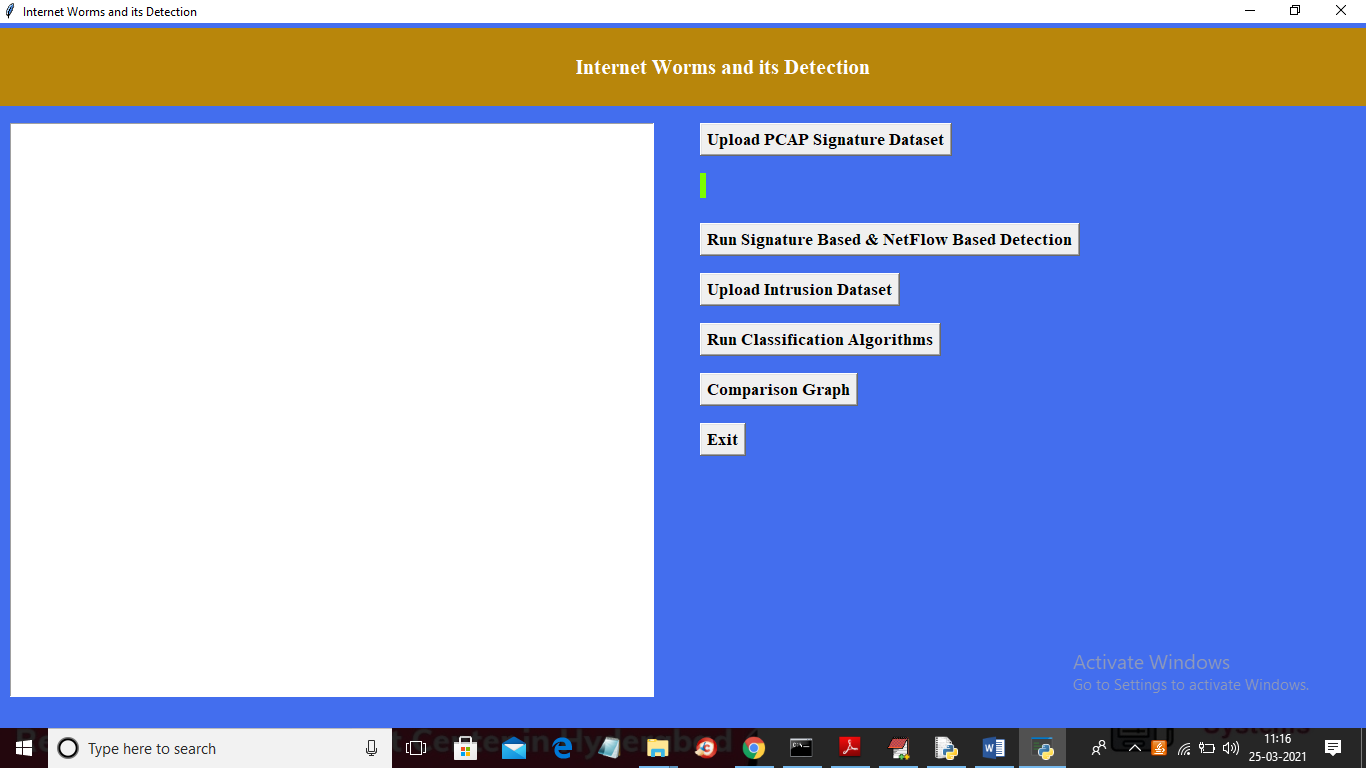
In above IDS dataset we have traffic signatures and each signature is associated with one label as normal or Neptune attack or warezclient etc. Below is PCAP dataset



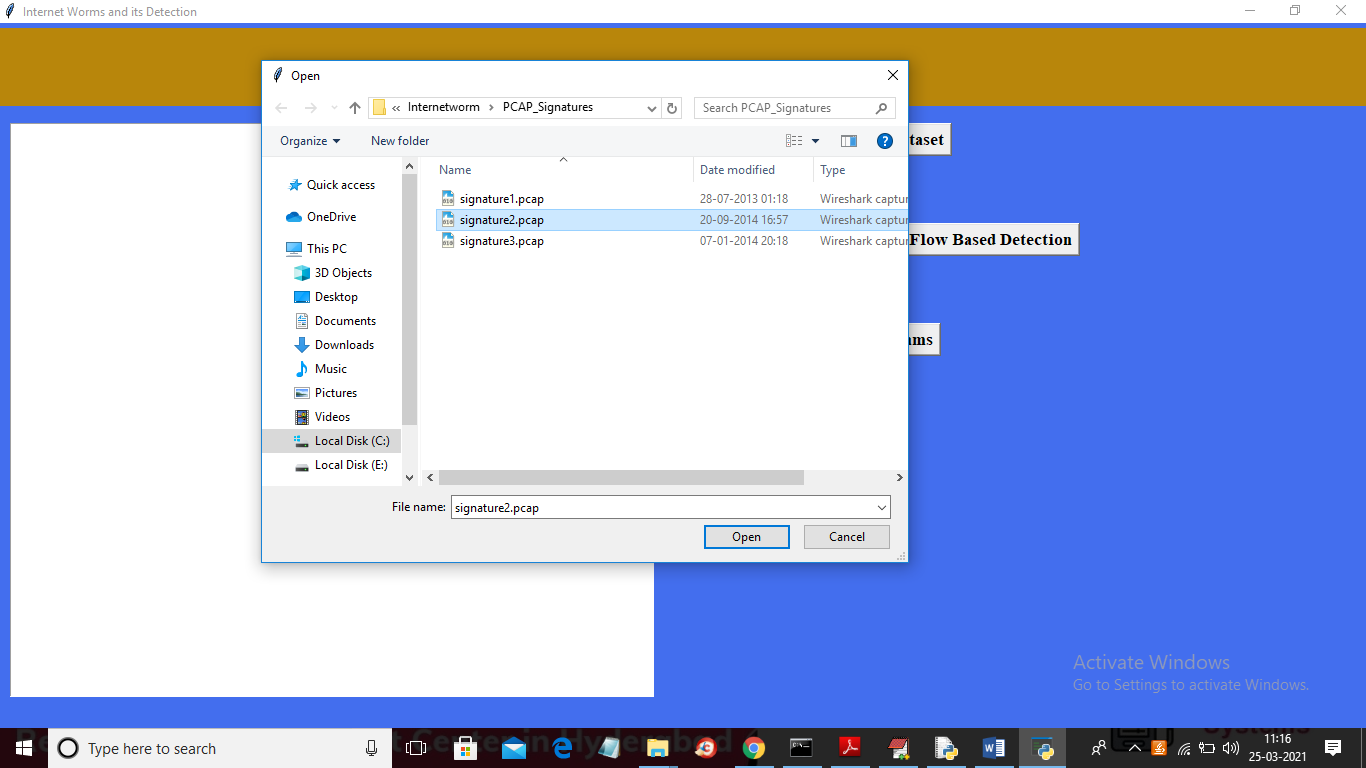
Now we will used above dataset to detect attacks

SCREEN SHOTS

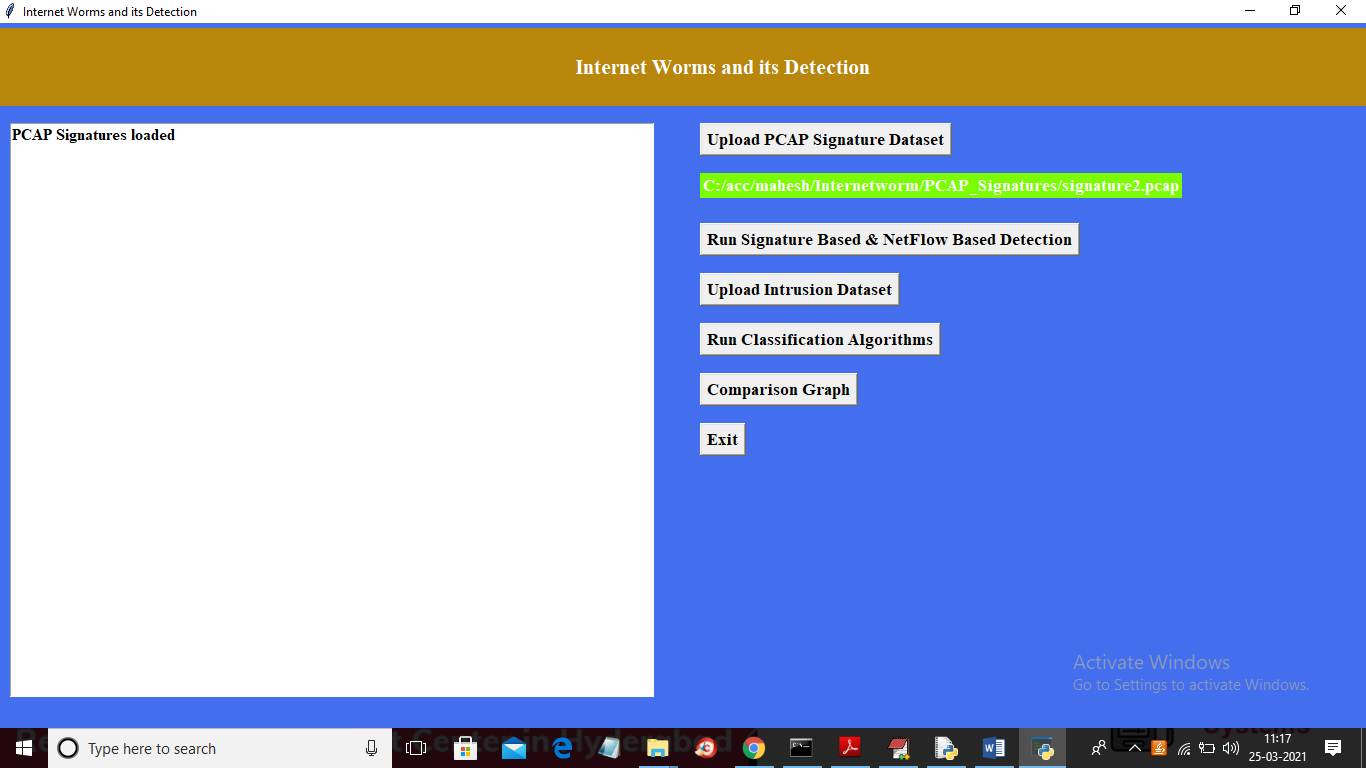
To run project double click on ‘run.bat’ file to get below screen



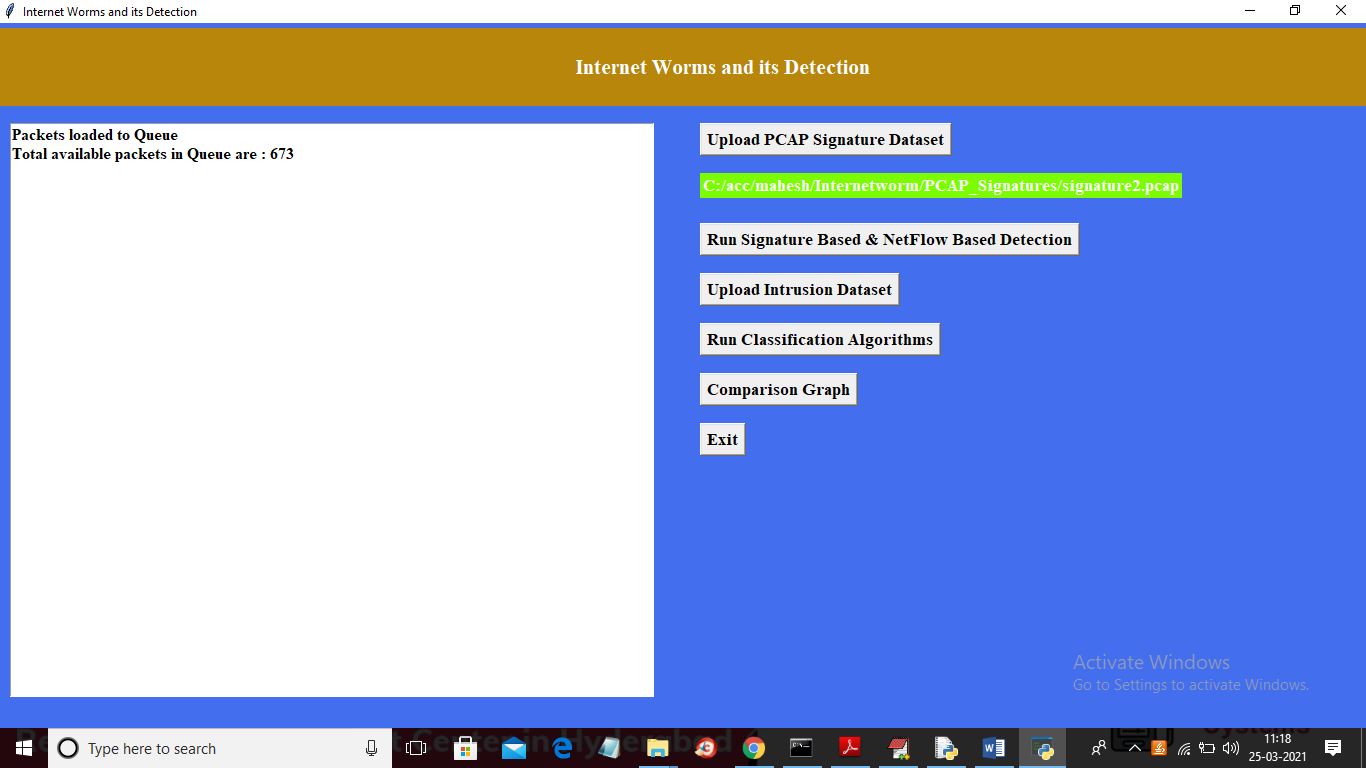
In above screen click on ‘Upload PCAP Signature Dataset’ button to upload signature



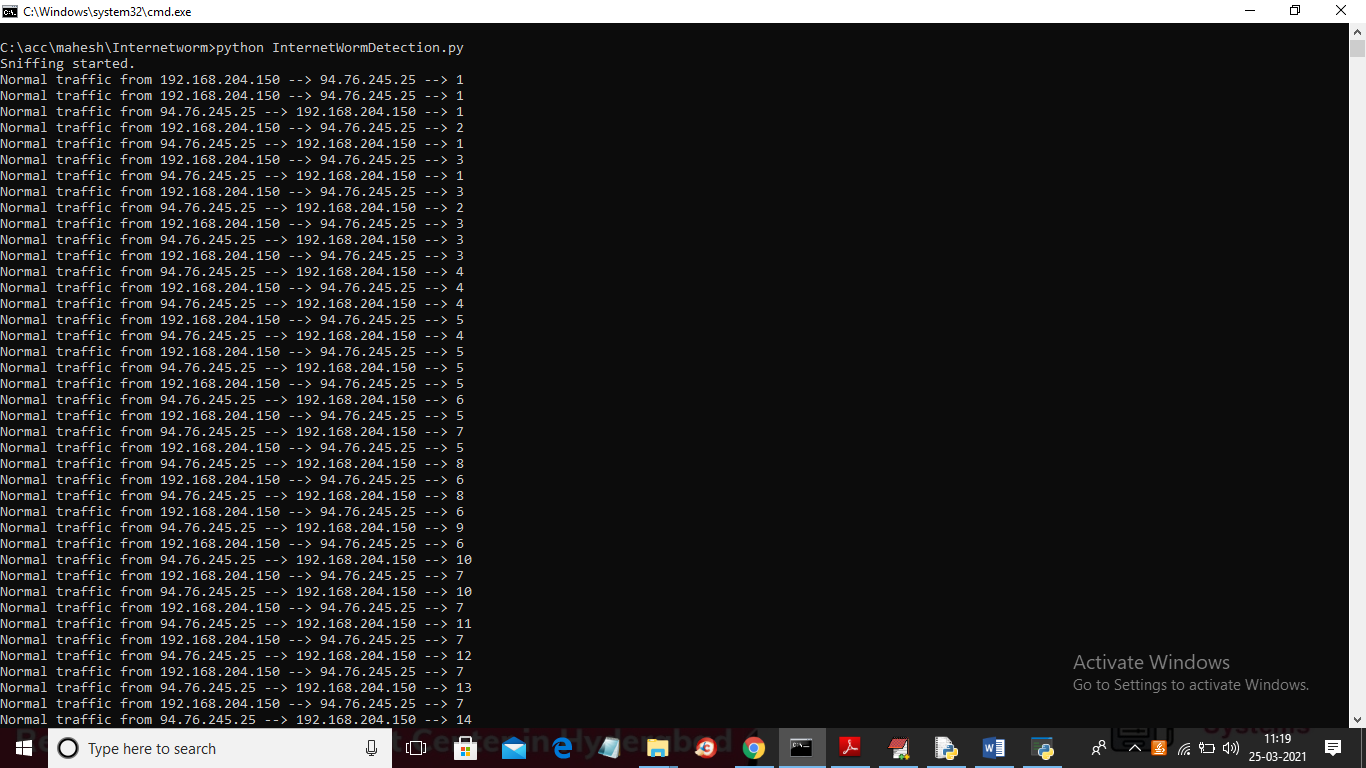
In above screen selecting and uploading ‘signature2.pcap’ file and then click on ‘Open’ button to load PCAP file and to get below output



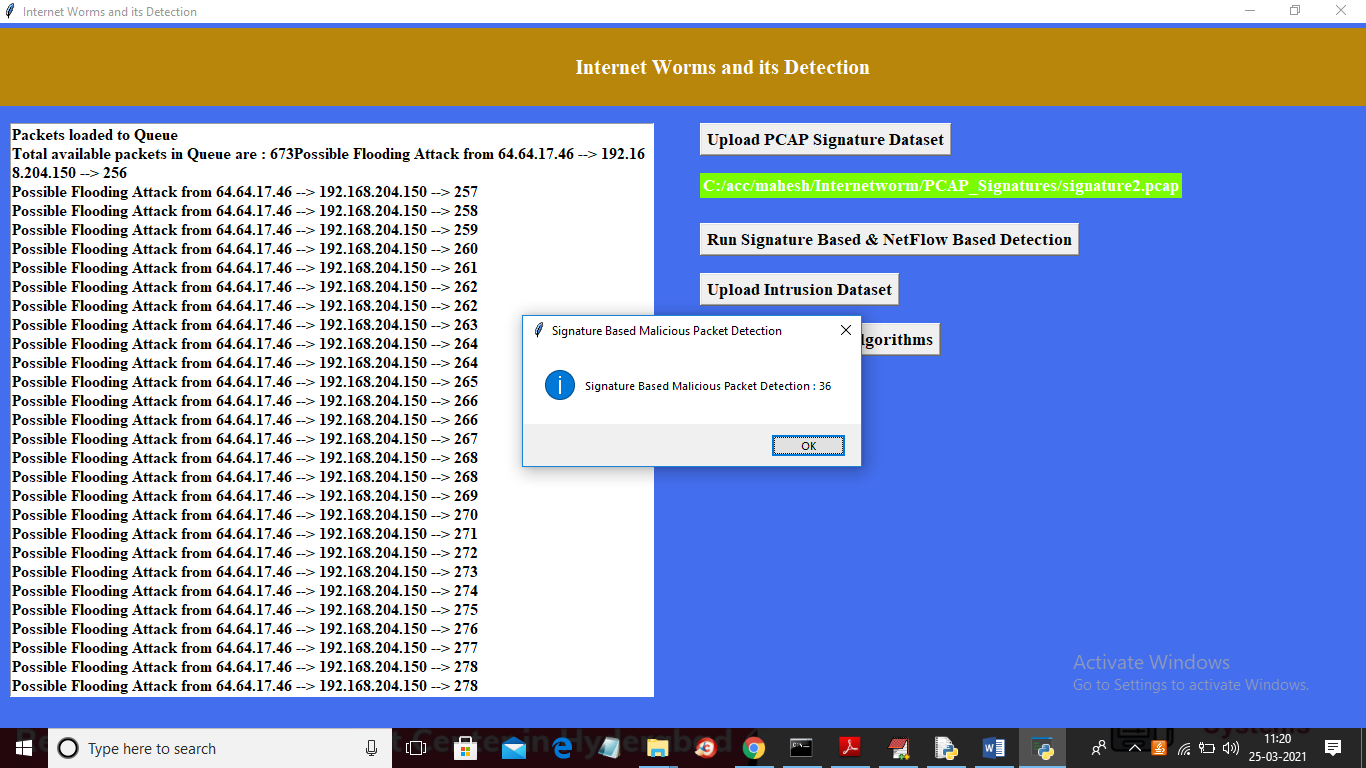
In above screen PCAP file loaded and now click on ‘Run Signature Based & NetFlow Based Detection’ button to start analysing PCAP file to detect attacks



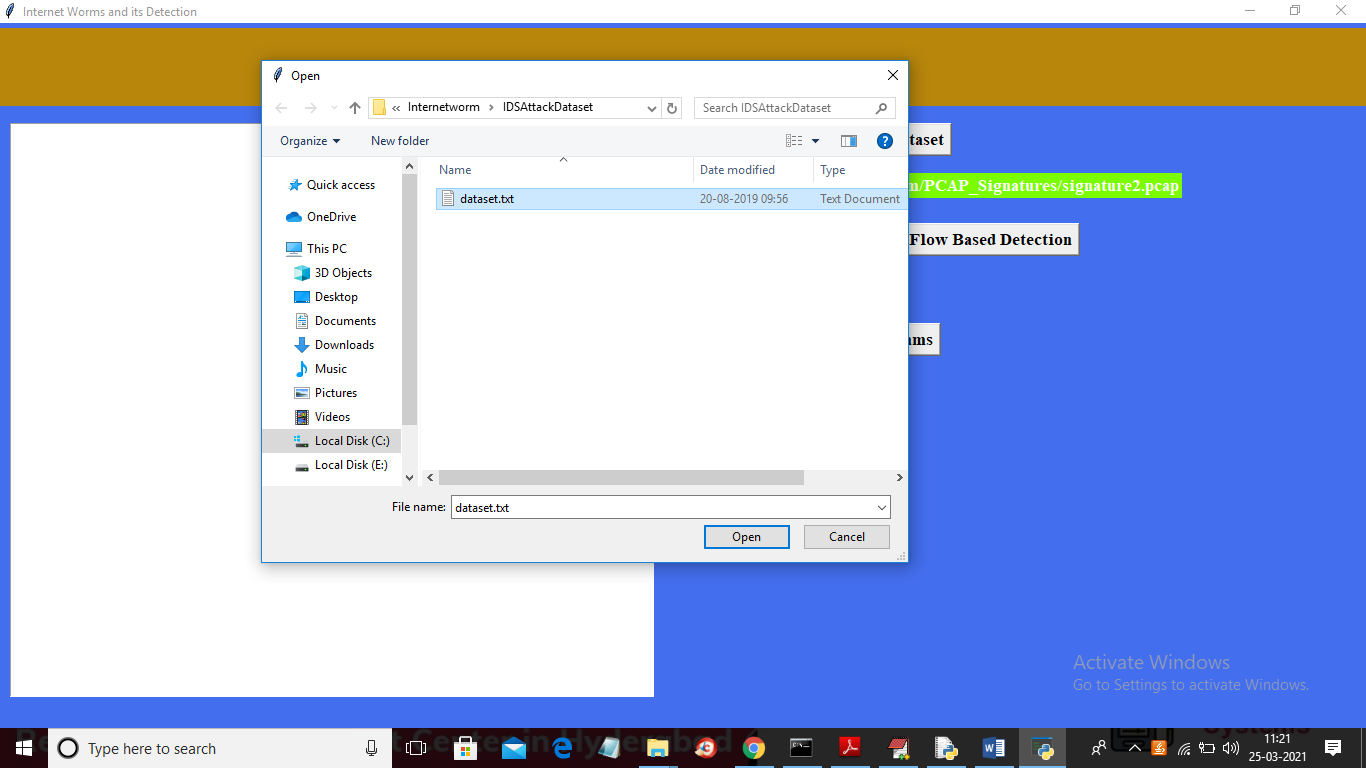
In above screen in loaded file we got total 673 packets and in below console application start inspecting each packet



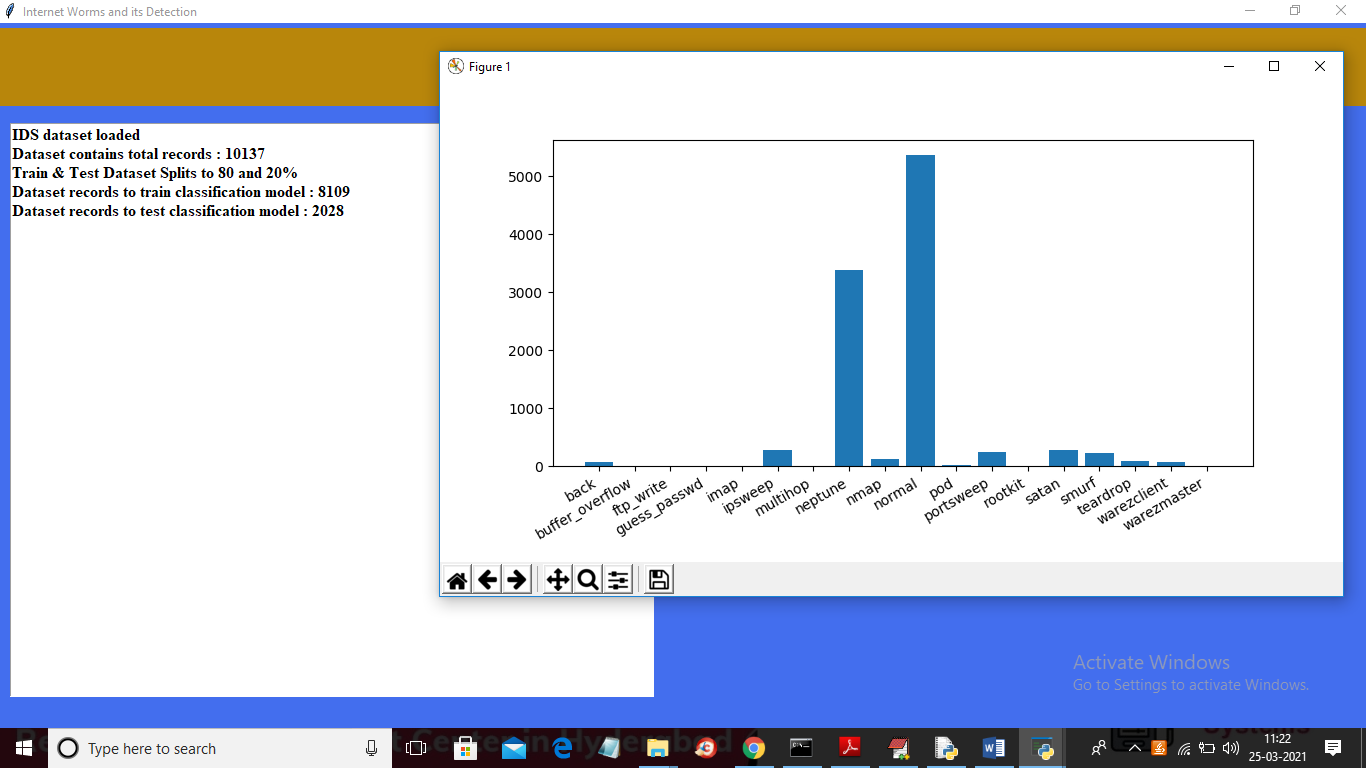
In above screen application inspect packets from different IP and then identifying whether signature contains normal packet or flood attack packet and after processing all packets will get below result



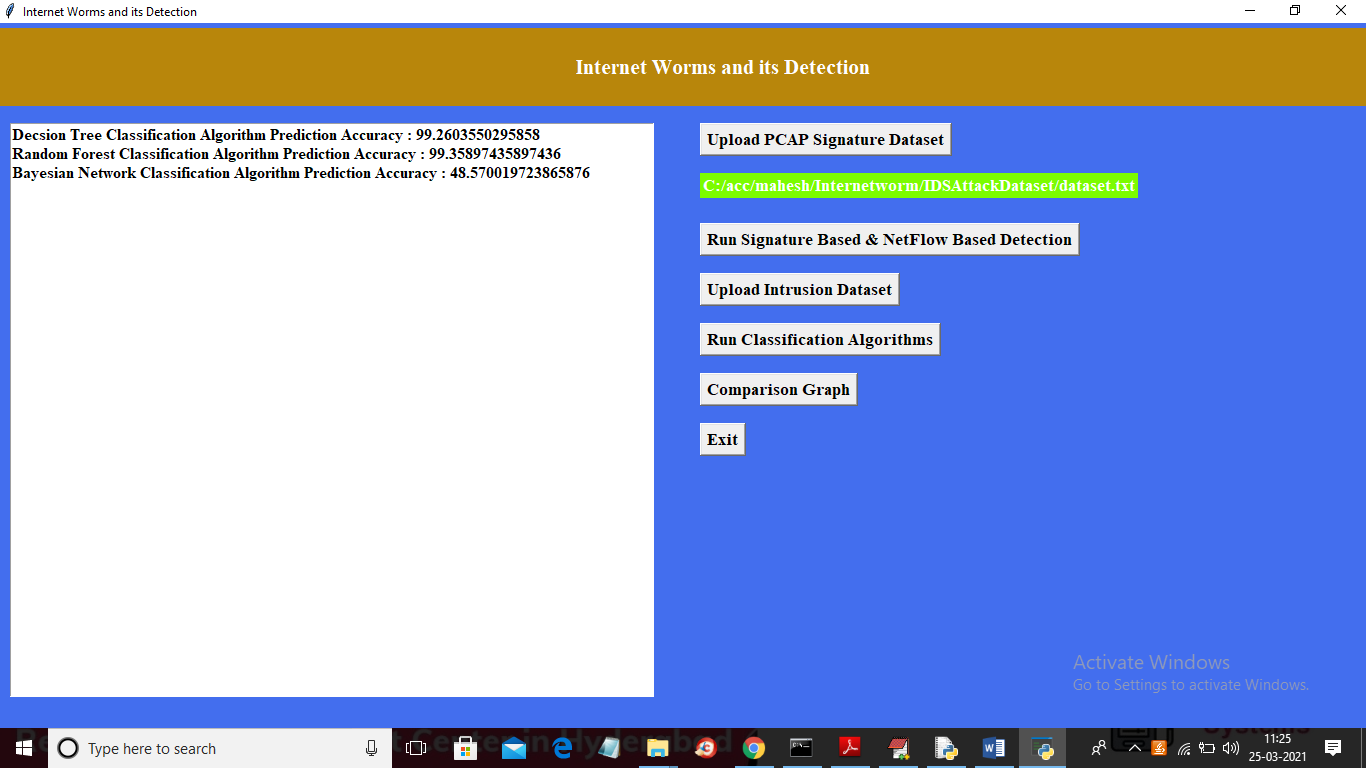
In above screen displaying total 36 malicious attack packets detected in loaded dataset out of 673 packets and we IP address and port number from which attack happen and now click on ‘Upload Intrusion Dataset’ button to upload IDS dataset



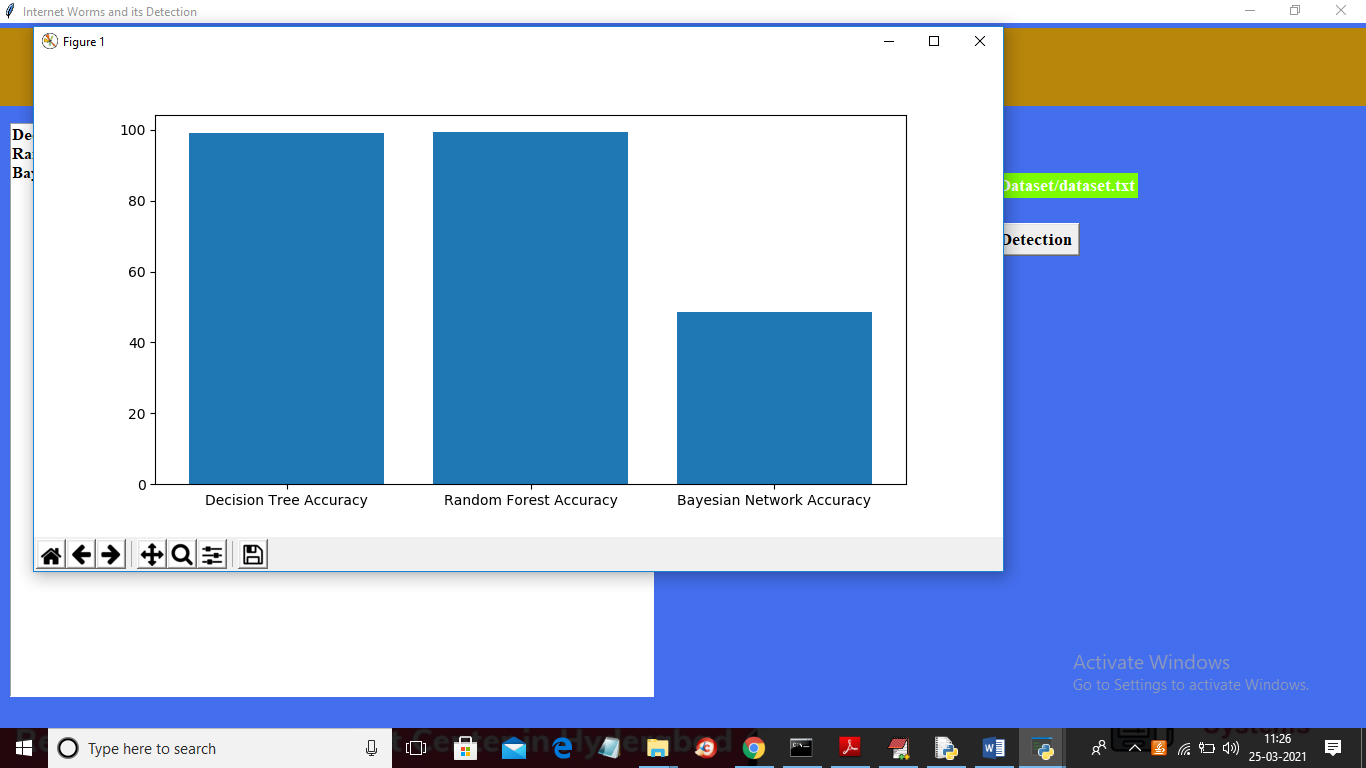
In above screen selecting and uploading IDS ‘dataset.txt’ file and then click on ‘Open’ button to get below output



In above graph we can see attack names in x-axis and total packets count from that attack is available in y-axis and in above screen in text area we can see total records in dataset and then splitting and using 80% records from dataset to train classification model and 20% records for testing classification model prediction accuracy. Now train and test dataset is ready and now click on ‘Run Classification Algorithms’ button to train various classification model and then calculate its accuracy



In above screen decision tree and random forest accuracy is 99% and Bayesian network accuracy is 48% and we can use either decision tree or random forest model to detect future attacks and now click on ‘Comparison Graph’ button to get below graph



In above graph x-axis represents algorithm name and y-axis represents accuracy of those algorithms and from above graph we can conclude that Random forest and decision tree is giving better prediction accuracy.

Note: we don’t have honeypot server or dataset to implement this technique