Extension work using EML algorithm

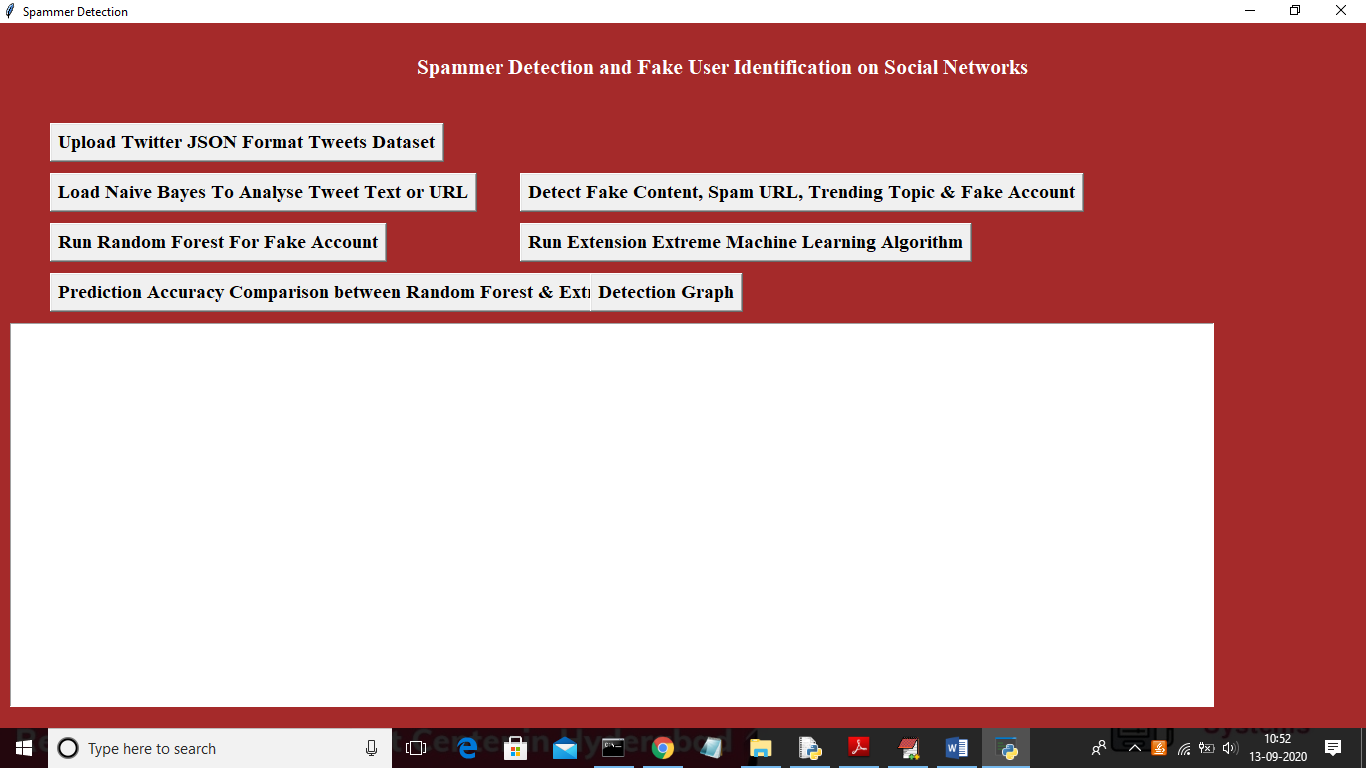
In this project as extension work I have added EML algorithm and below is the algorithm details. For base paper details you can refer previous screen shots and this contains details about extension only

Extreme machine learning (EML) are feed forward neural networks for classification, regression, clustering, sparse approximation, compression and feature learning with a single layer or multiple layers of hidden nodes, where the parameters of hidden nodes (not just the weights connecting inputs to hidden nodes) need not be tuned. These hidden nodes can be randomly assigned and never updated (i.e. they are random projection but with nonlinear transforms), or can be inherited from their ancestors without being changed. In most cases, the output weights of hidden nodes are usually learned in a single step, which essentially amounts to learning a linear model. The name "extreme learning machine" (ELM) was given to such models by its main inventor Guang-Bin Huang.

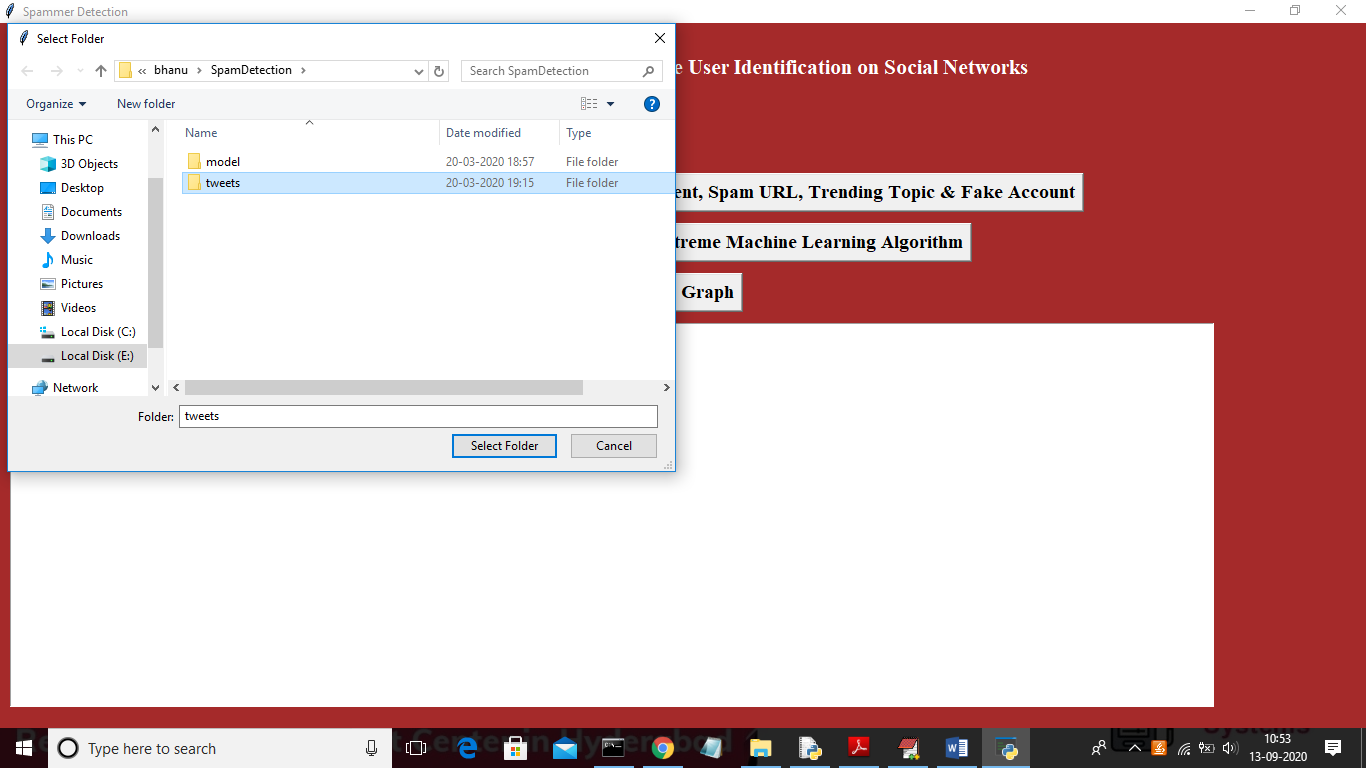
According to their creators, these models are able to produce good generalization performance and learn thousands of times faster than networks trained using backpropagation. In literature, it also shows that these models can outperform support vector machines (SVM) and other classifiers.

SCREEN SHOTS

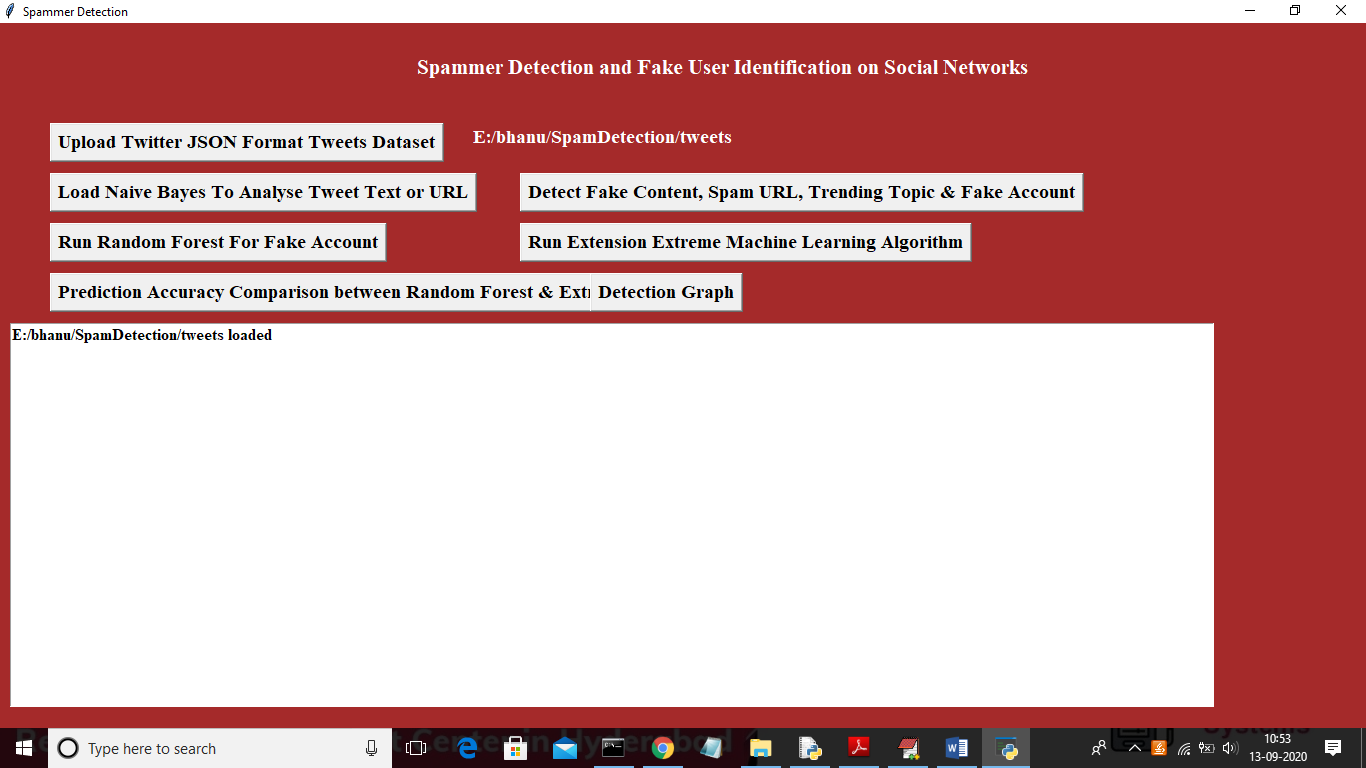
To run project double click on ‘run.bat’ file to get below screen



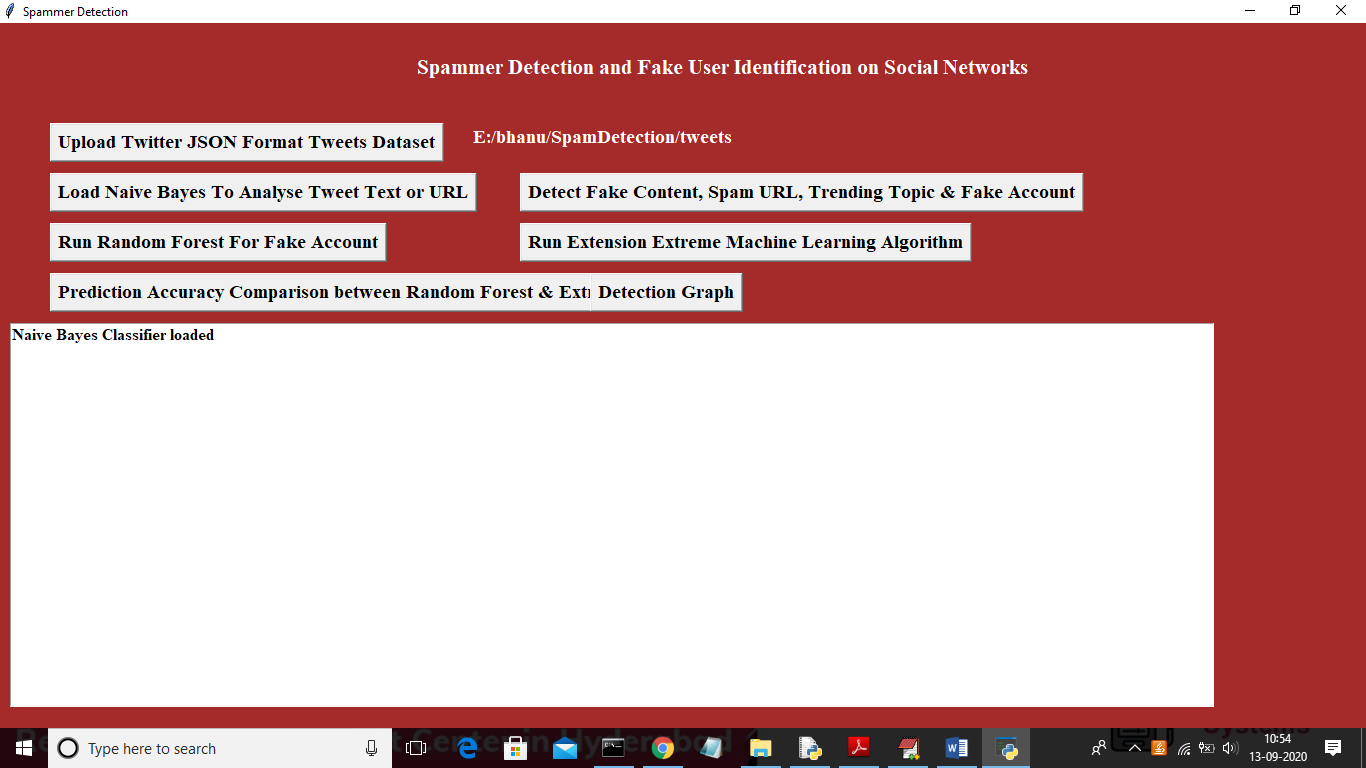
In above screen upload dataset by clicking on ‘Upload Twitter JSON Format Tweets Dataset’ button



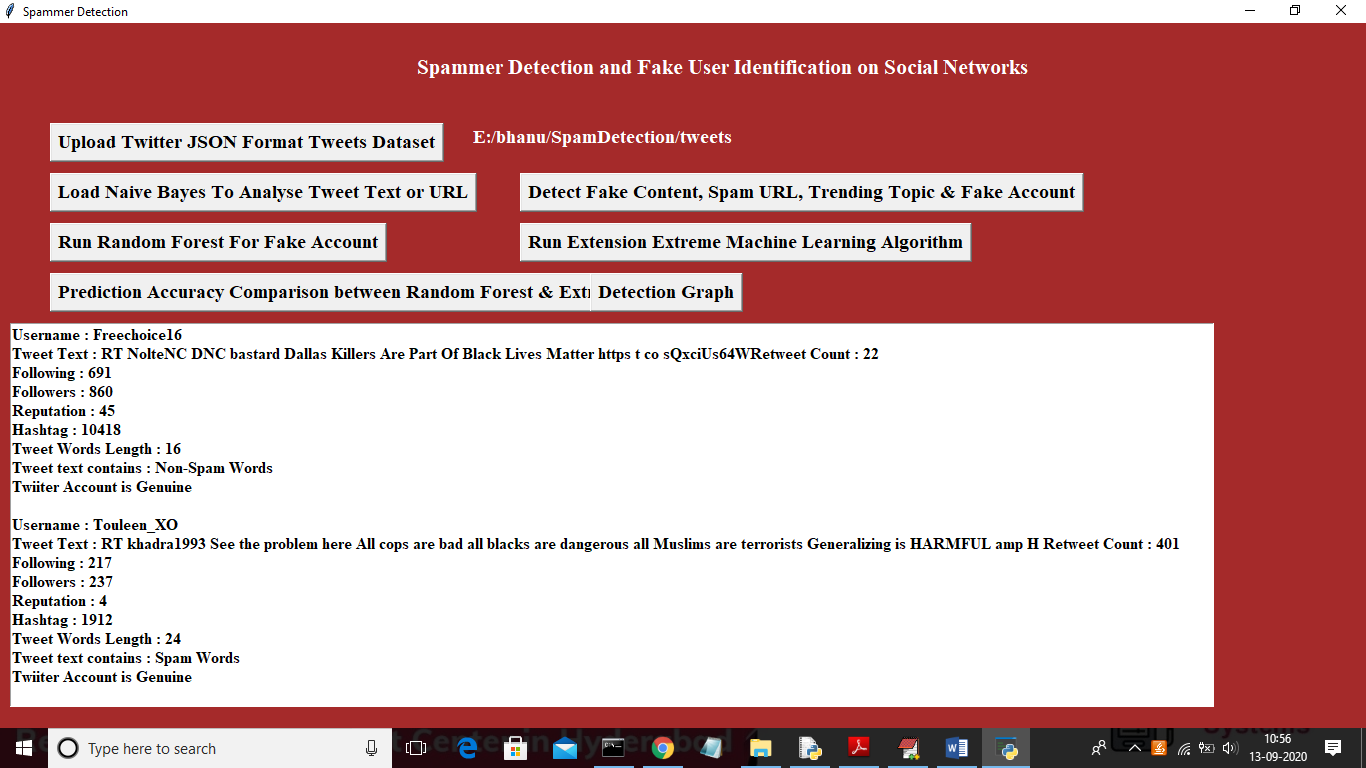
In above screen after selecting tweets folder click on ‘Select Folder’ to load tweets dataset and to get below screen



Now click on ‘Load Naïve Bayes To Analyse Tweet Text or URL’ button to analyse tweet text and generate machine learning features



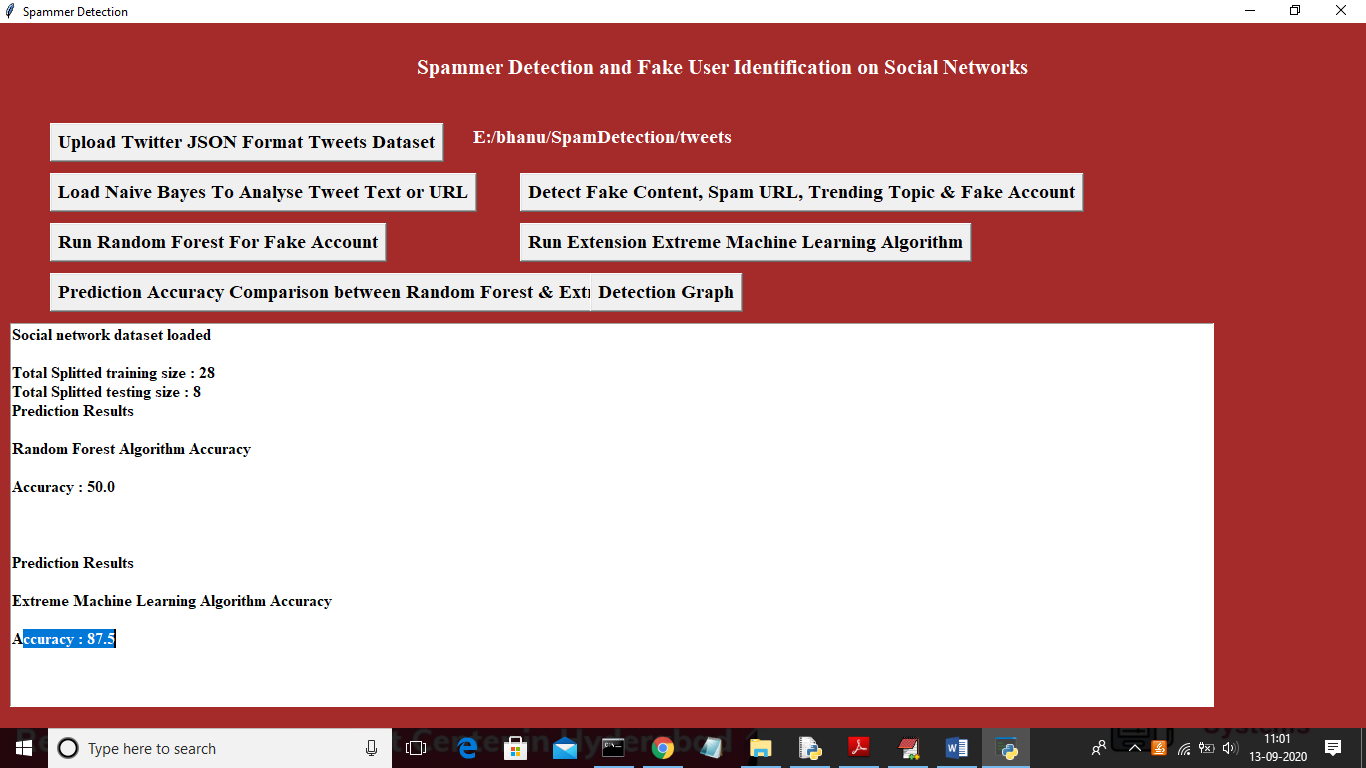
In above screen naïve bayes classifier loaded to analyse tweet text and now click on ‘Detect Fake Content, Spam URL, Trending Topic & Fake Account’ to detect spam accounts and create machine learning features



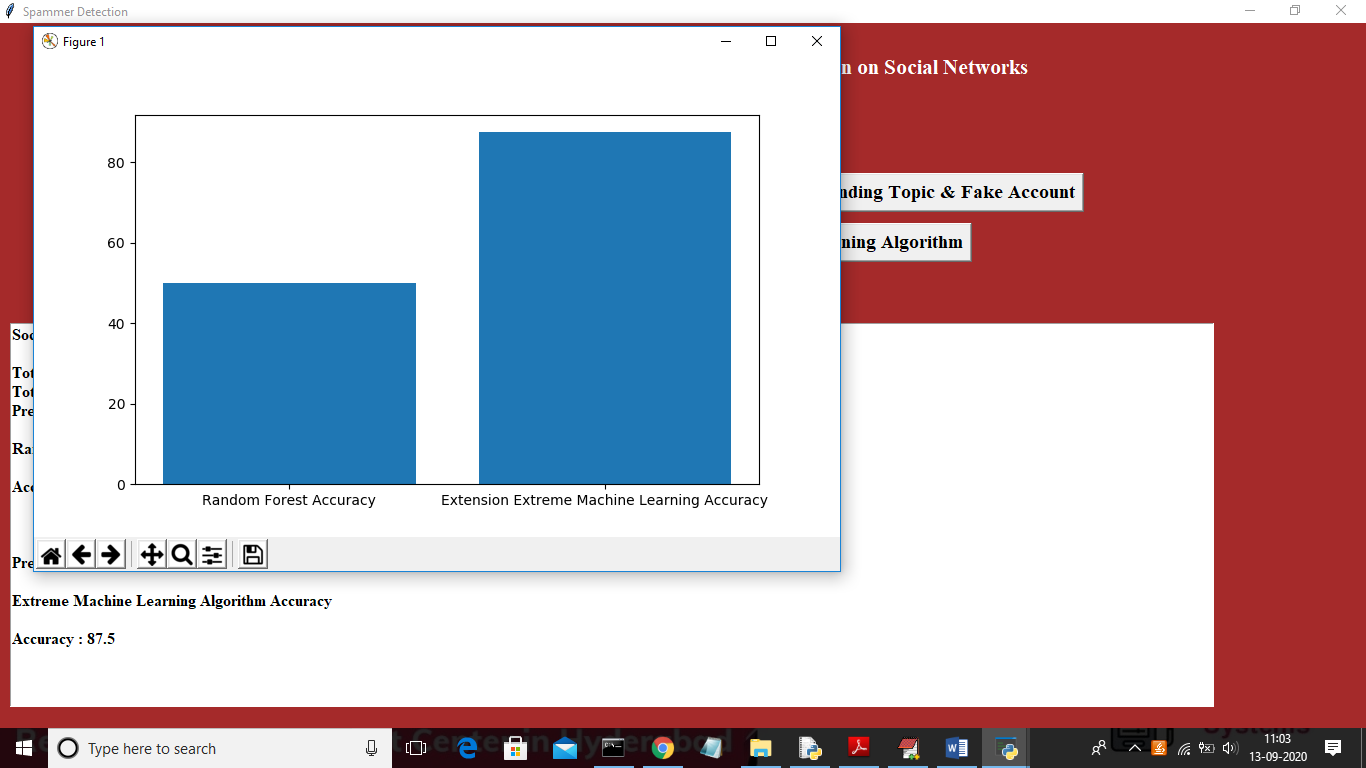
In above screen we analyse all users accounts and then identify whether account is normal or contains spam content using naïve bayes algorithm and now click on ‘Run Random Forest For Fake Account’ button to build machine learning model on above data so in future by applying new account details we can predict whether account is normal or fake.



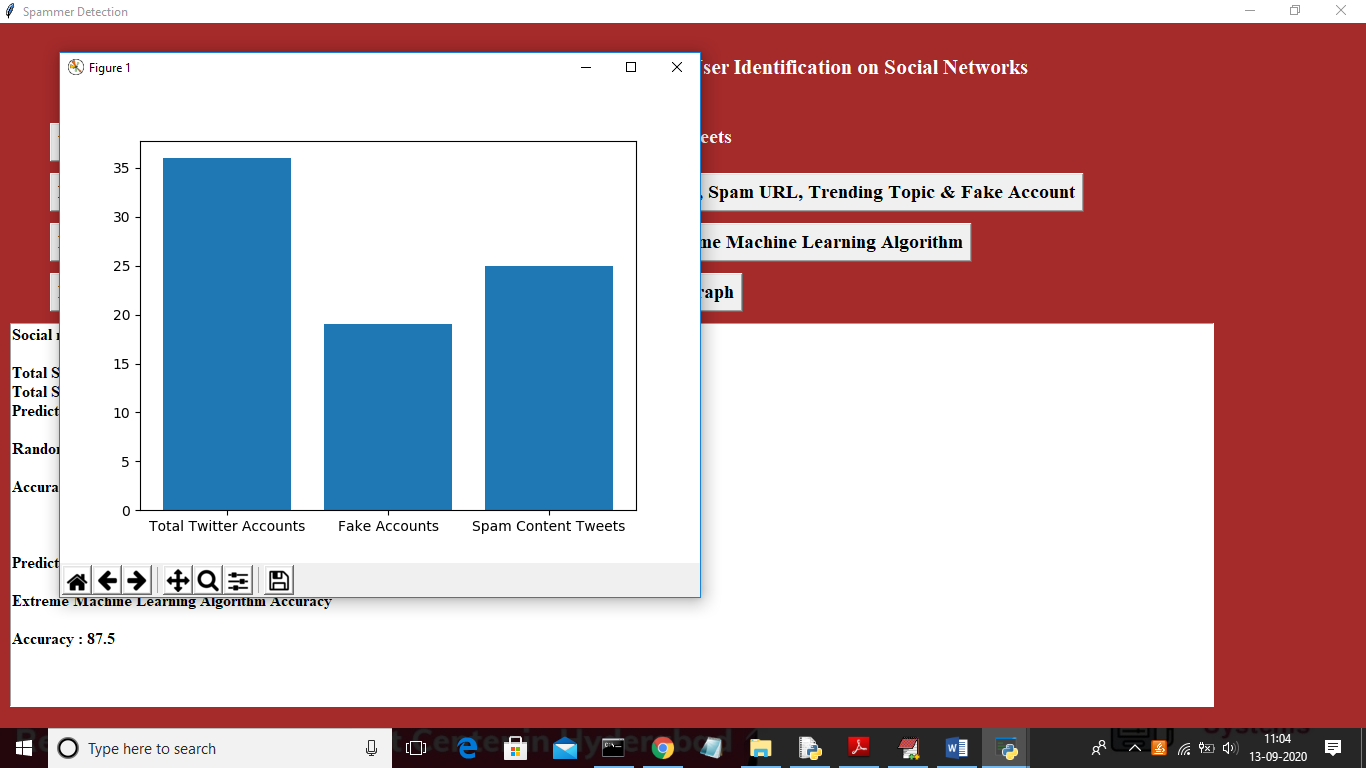
In above screen we can see total dataset contains 36 accounts and application using 80% records (28) for training and 20% records (8) for testing and with random forest we got test data prediction accuracy as 50%. Now click on ‘Run Extension Extreme Machine Learning Algorithm’ button to generate training model on above data and to get prediction accuracy on test data



In above screen with extension extreme machine learning algorithm we got 87.5% accuracy and its better in prediction fake or normal account compare to random forest. Here dataset will be splitted to training and testing part randomly so accuracy may vary for each run. Now click on ‘Prediction Accuracy Comparison between Random Forest & Extreme ML’ button to get accuracy comparison between random forest and extreme machine learning



In above graph x-axis represents algorithm names and y-axis represents accuracy of those algorithm and from above graph we can conclude that extension work get better accuracy and now click on ‘Detection Graph’ button to get total accounts, fake and genuine accounts graph



In above graph x-axis represents types of account and y-axis represent count of each type account.