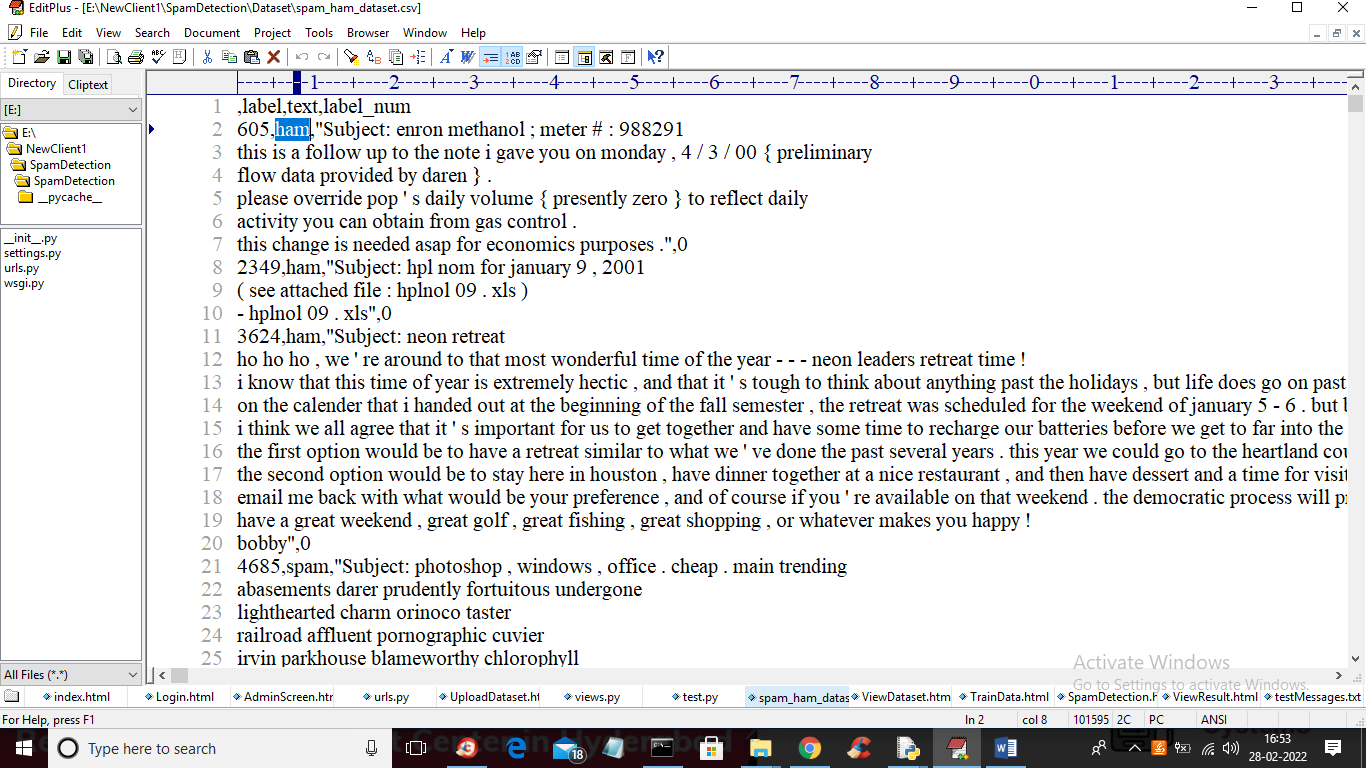
Email Spam Detection Using Machine Learning Algorithms

In this project we are using Random Forest machine learning algorithm to predict EMAIL messages as SPAM or HAM. To train random forest we have used KAGGLE EMAIL SPAM dataset and below screen showing some messages from dataset



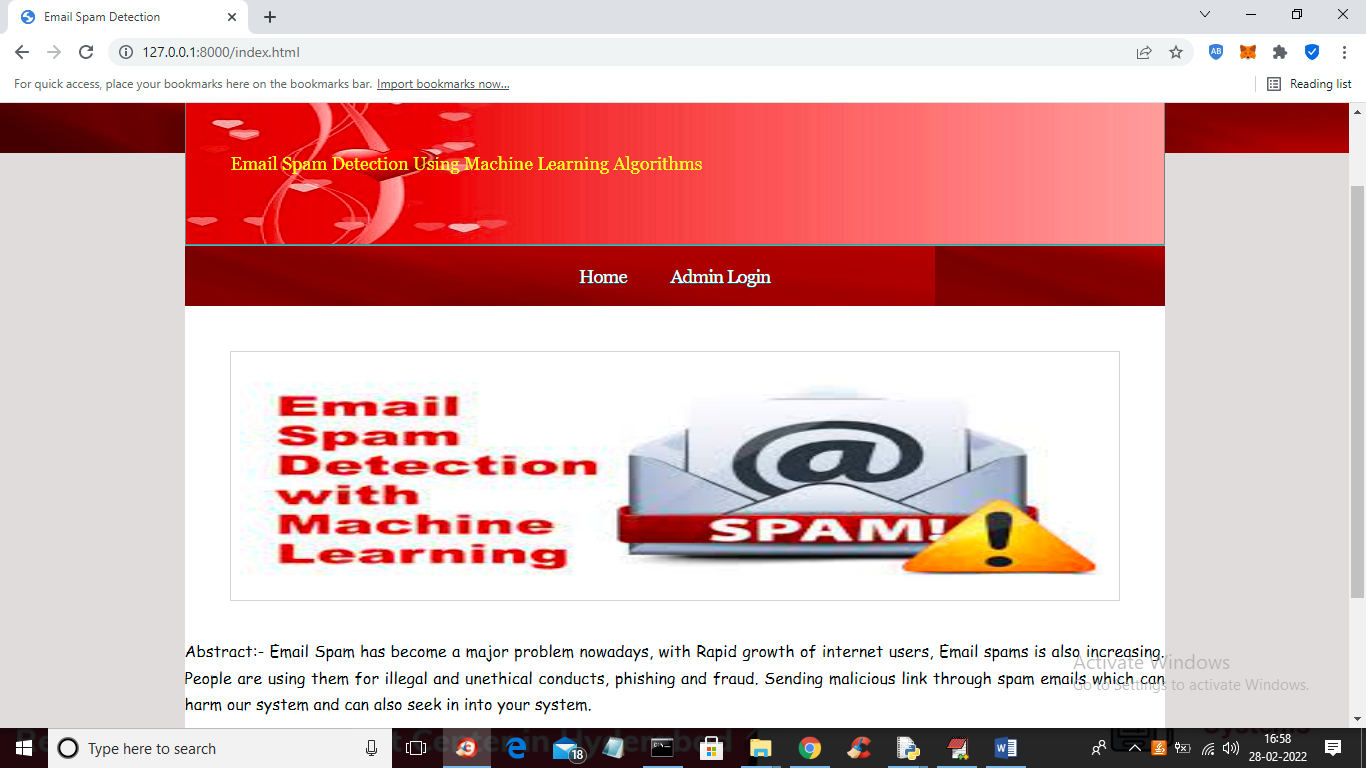
In above screen first row represents dataset column names and remaining rows contains EMAIL message and class label as HAM or SPAM and by using above dataset we will train Random Forest algorithm. After training we can input any message then random forest will predict as SPAM or HAM.

To implement this project we have designed following Modules

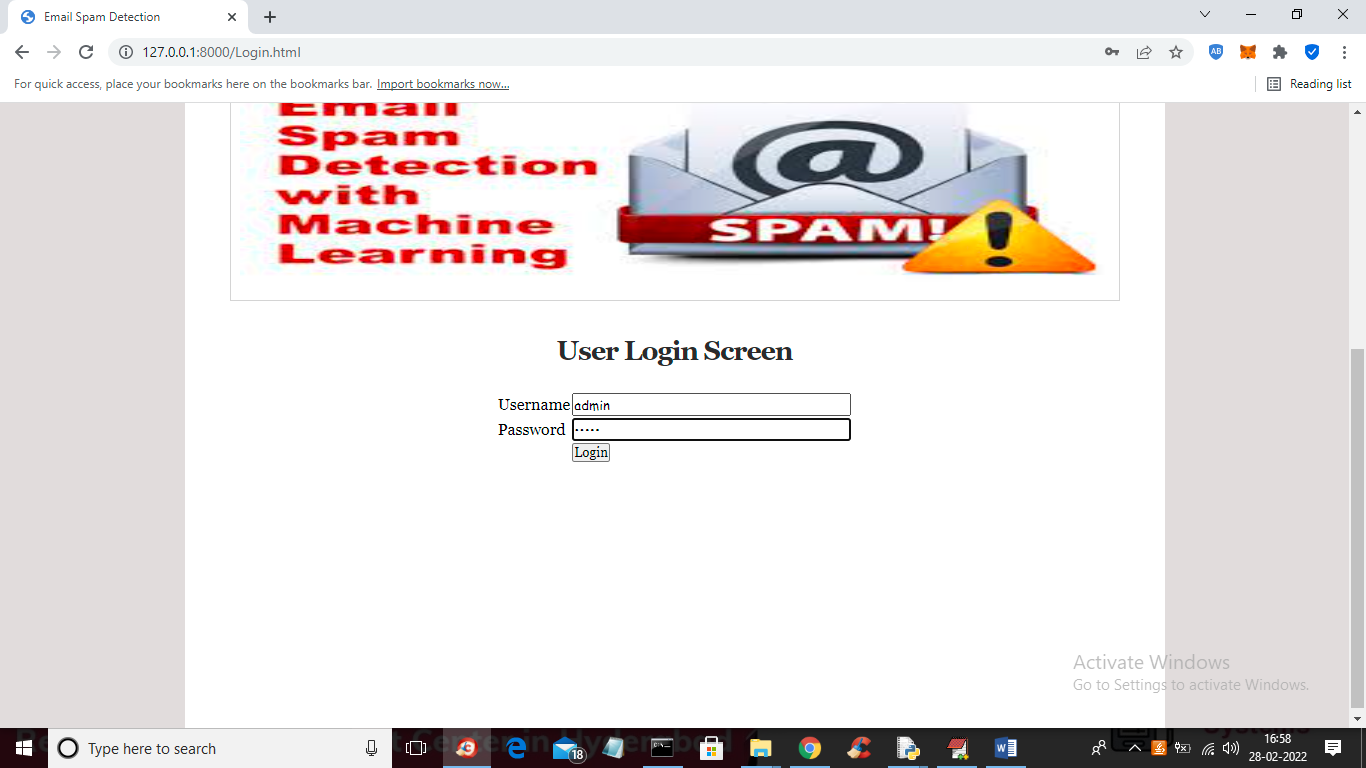
1. Admin Login: using this module admin can login to application by using username and password as ‘admin’ and ‘admin’
2. Upload Dataset: after login admin can click on ‘Upload Dataset’ link to upload dataset to application
3. Train Dataset Using Random Forest: now admin can click on this link to train random forest algorithm and this model will be applied on TEST data to calculate random forest prediction accuracy.
4. Spam Detection: admin can click on this link and then enter some message and then press submit button and then Random Forest will predict that message as HAM or SPAM

SCREEN SHOTS

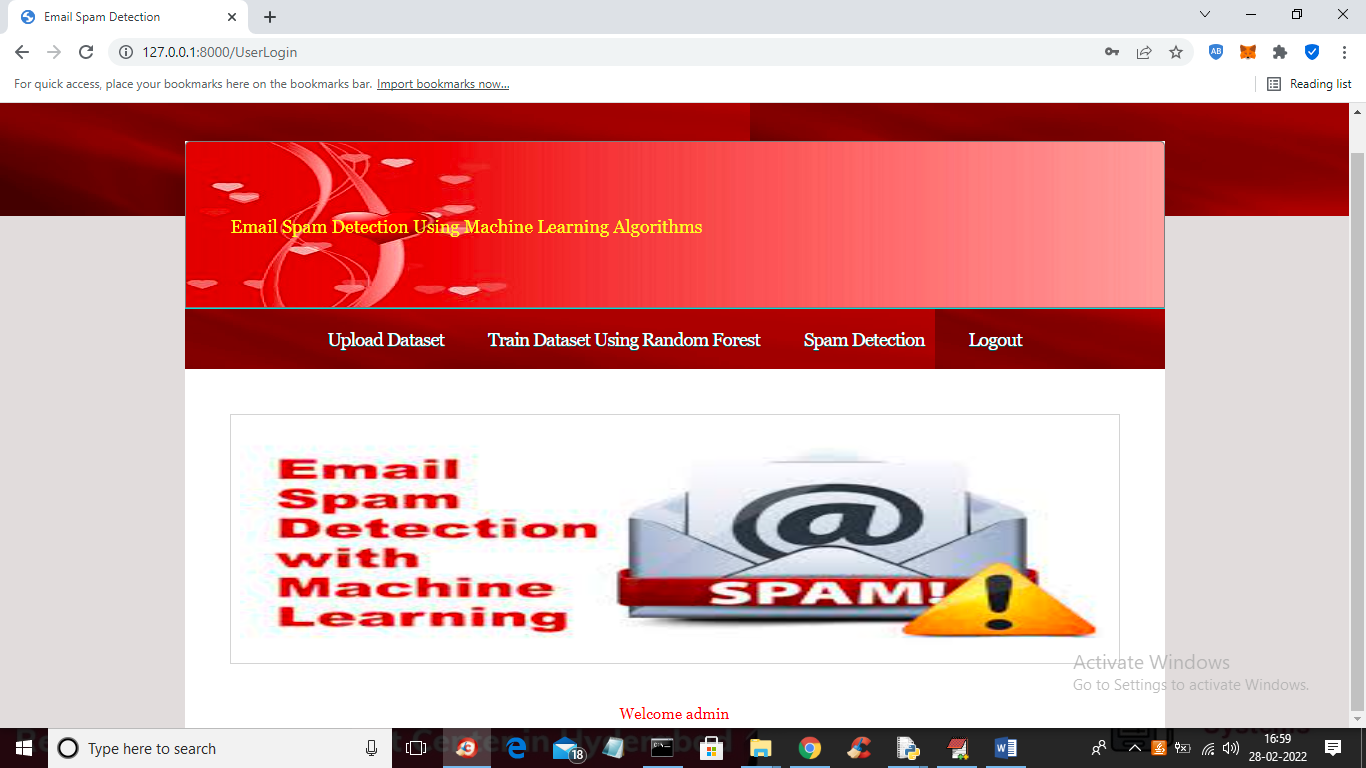
To run project double click on ‘run.bat’ file and then open browser and enter URL as ‘http://127.0.0.1:8000/index.html’ and press enter key to get below screen



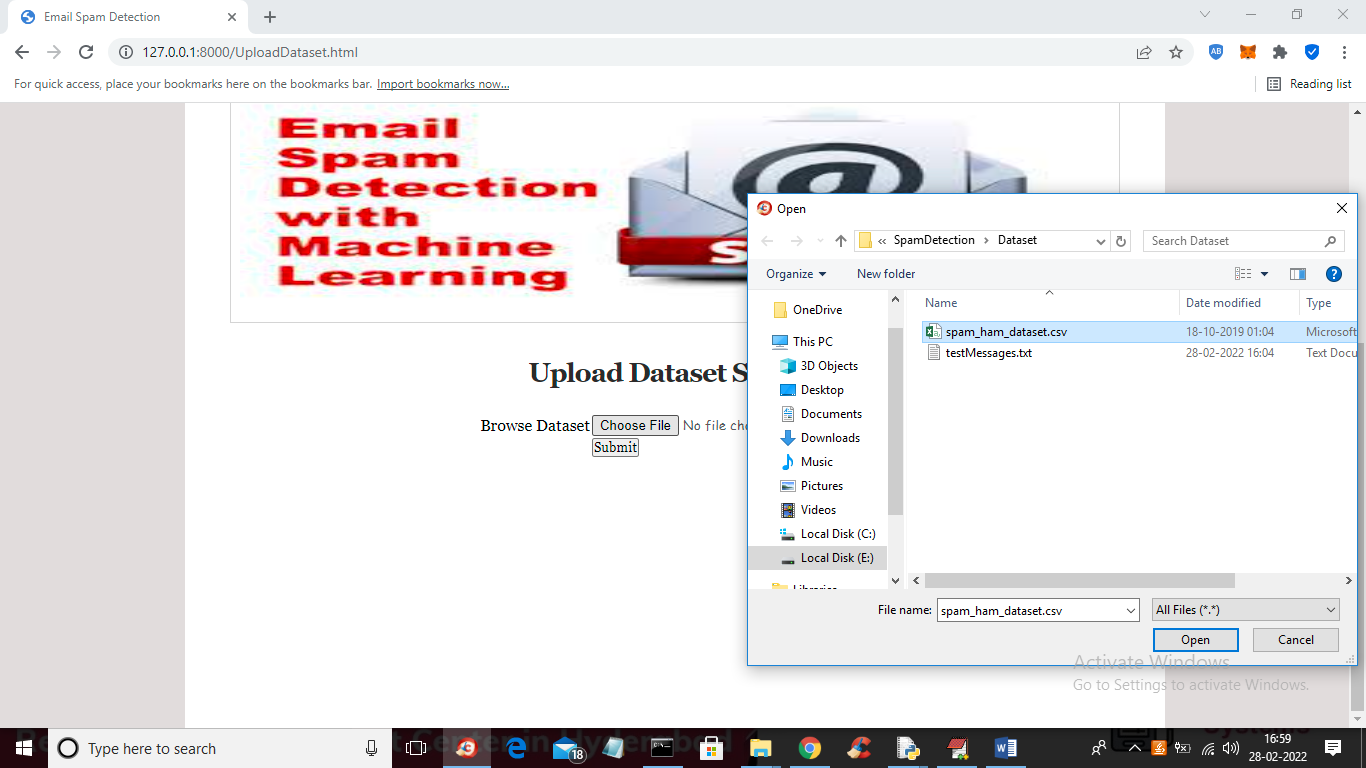
In above screen click on ‘Admin Login’ link to get below login screen



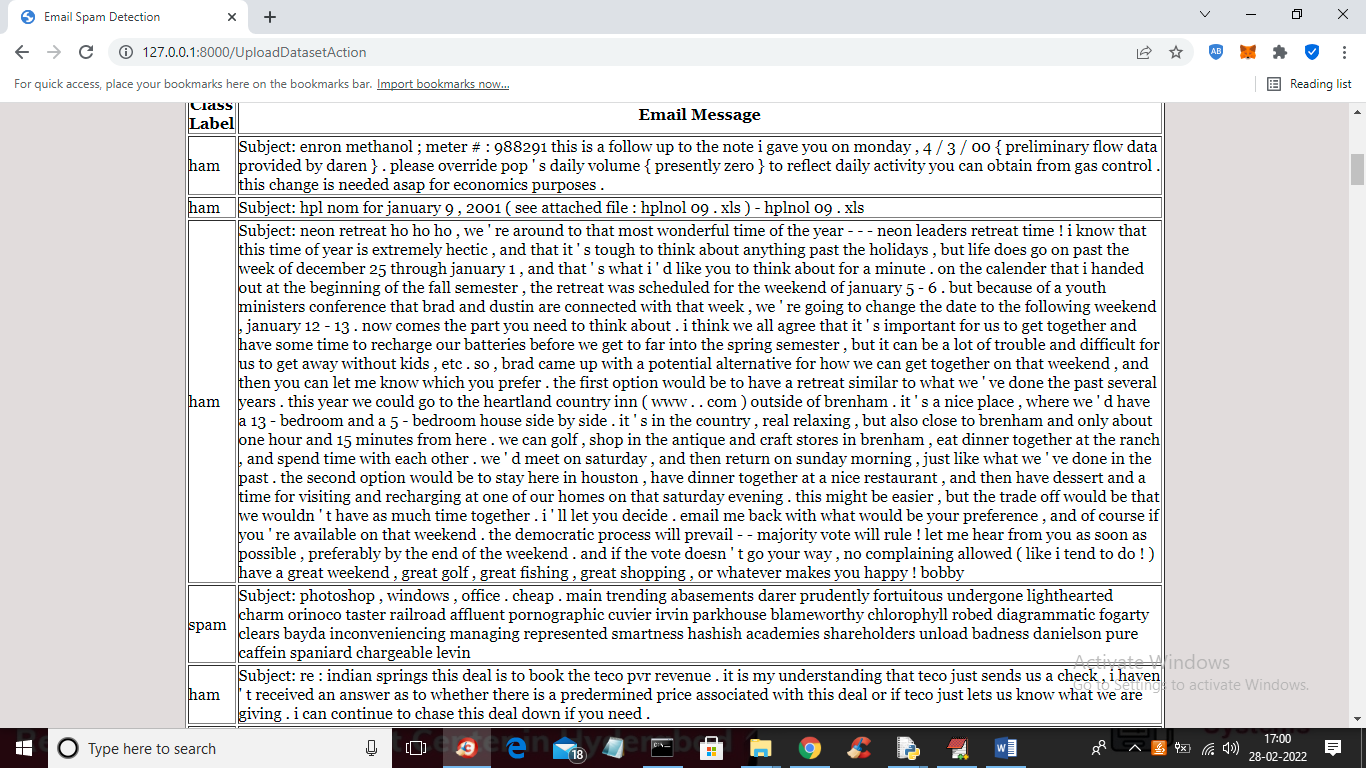
In above screen admin is login and then press button to get below screen



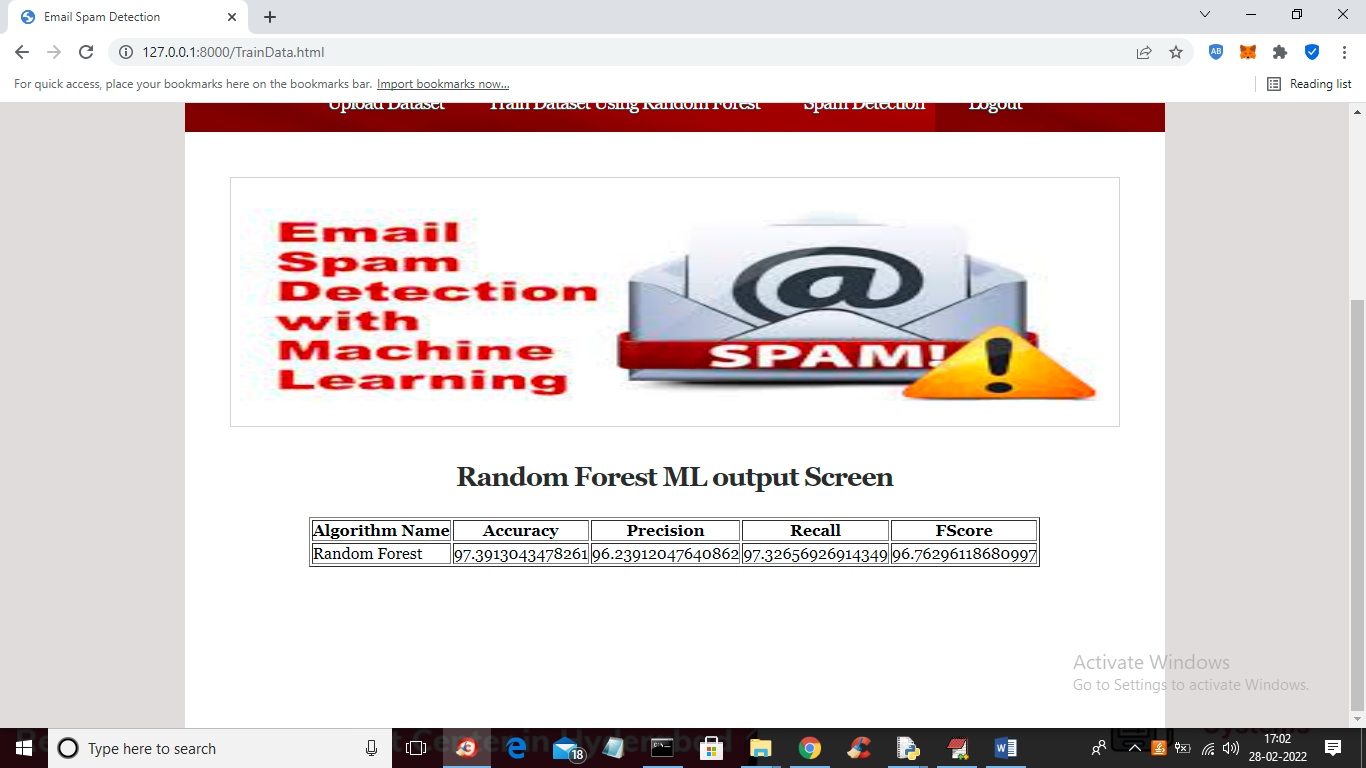
In above screen click on ‘Upload Dataset’ link to get below screen



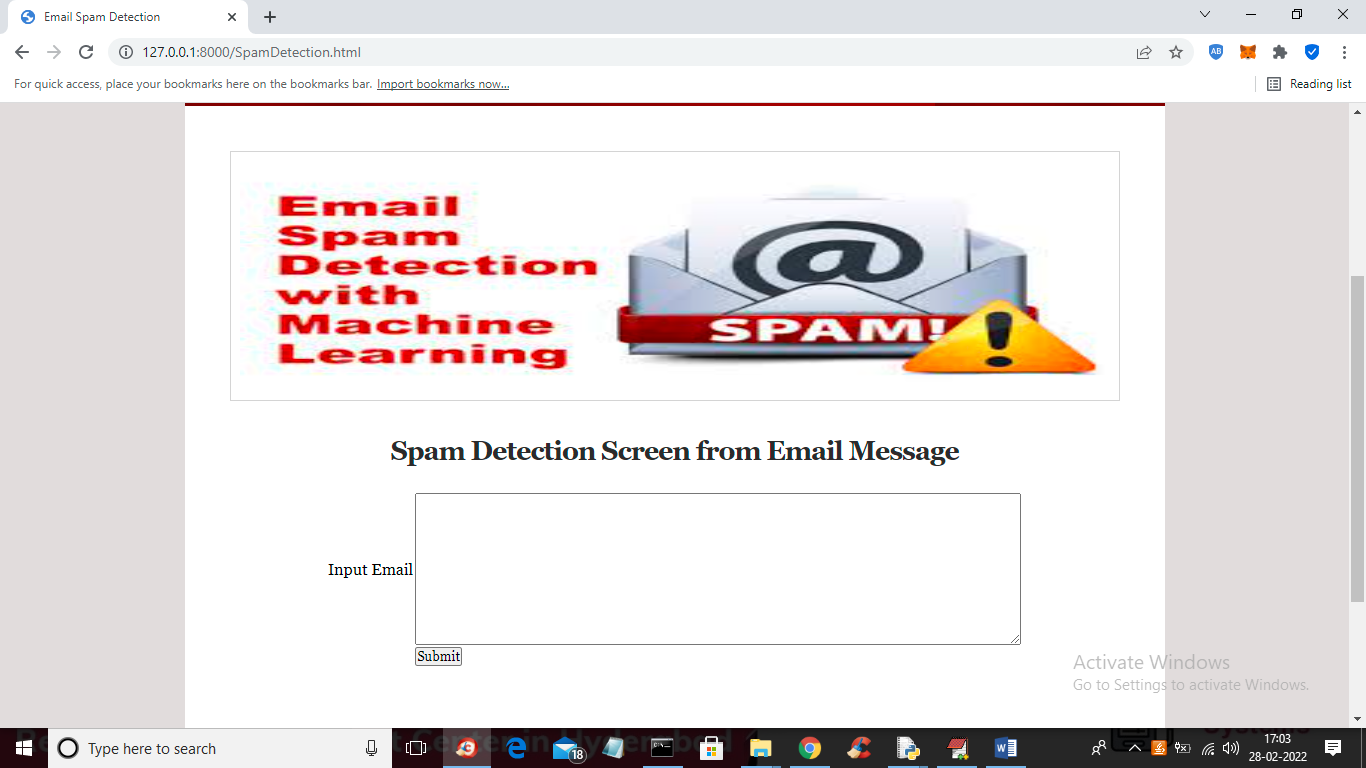
In above screen selecting and uploading ‘spam\_ham\_dataset.csv’ file and this dataset you can see inside ‘Dataset’ folder and then click on ‘Open’ button to load dataset and to get below screen



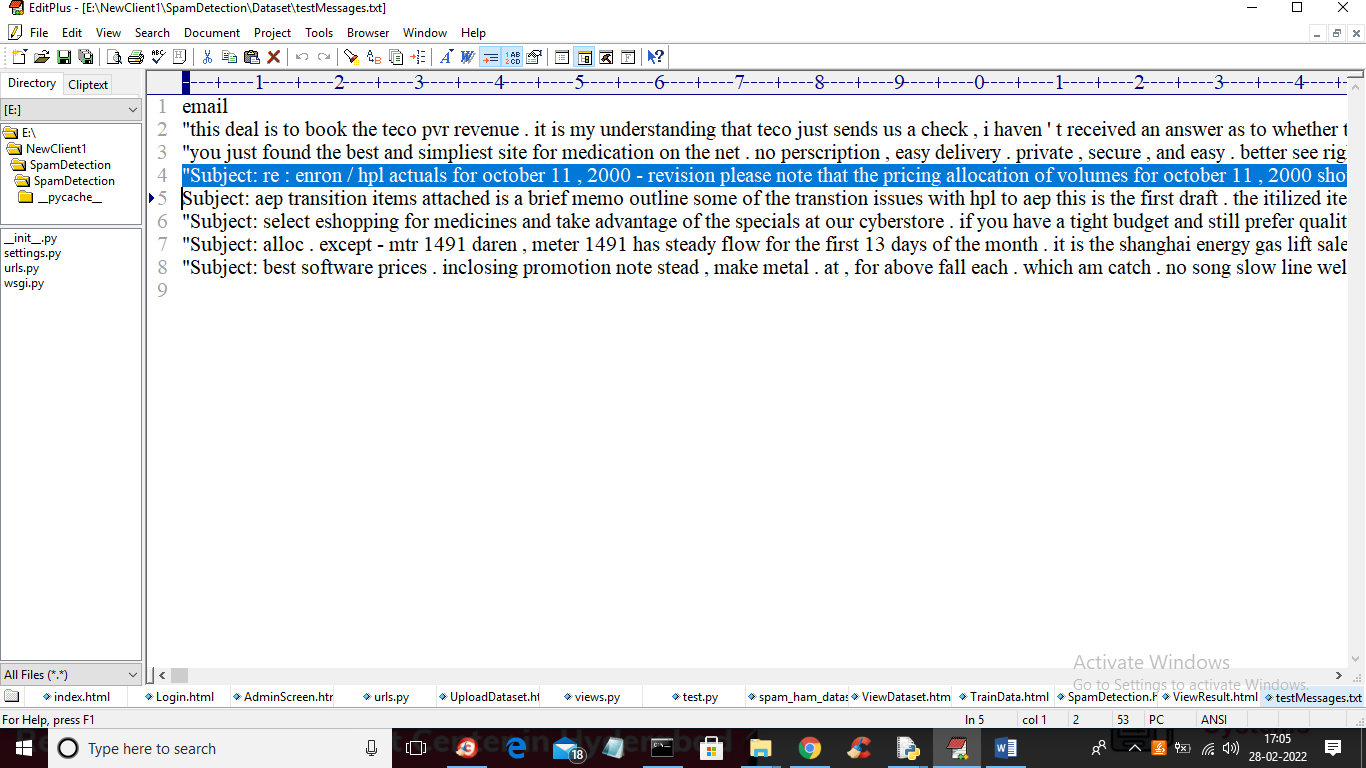
In above screen dataset loaded and we can see class label and email messages and now click on ‘Train Dataset Using Random Forest’ link to train random forest and get below output



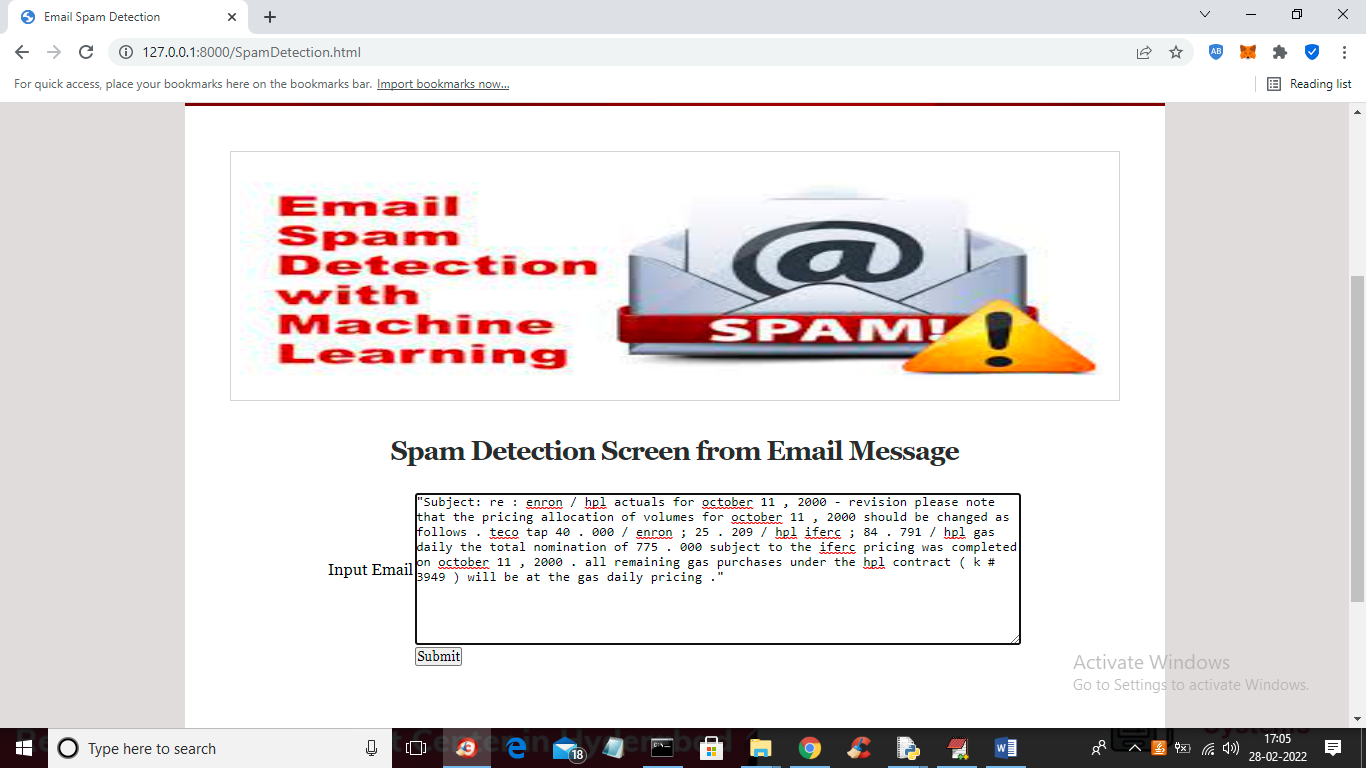
In above screen Random Forest trained and we got its prediction accuracy as 97% and we can see precision, recall and FSCORE. Now random forest is trained and now click on ‘Spam Detection’ link to get below screen



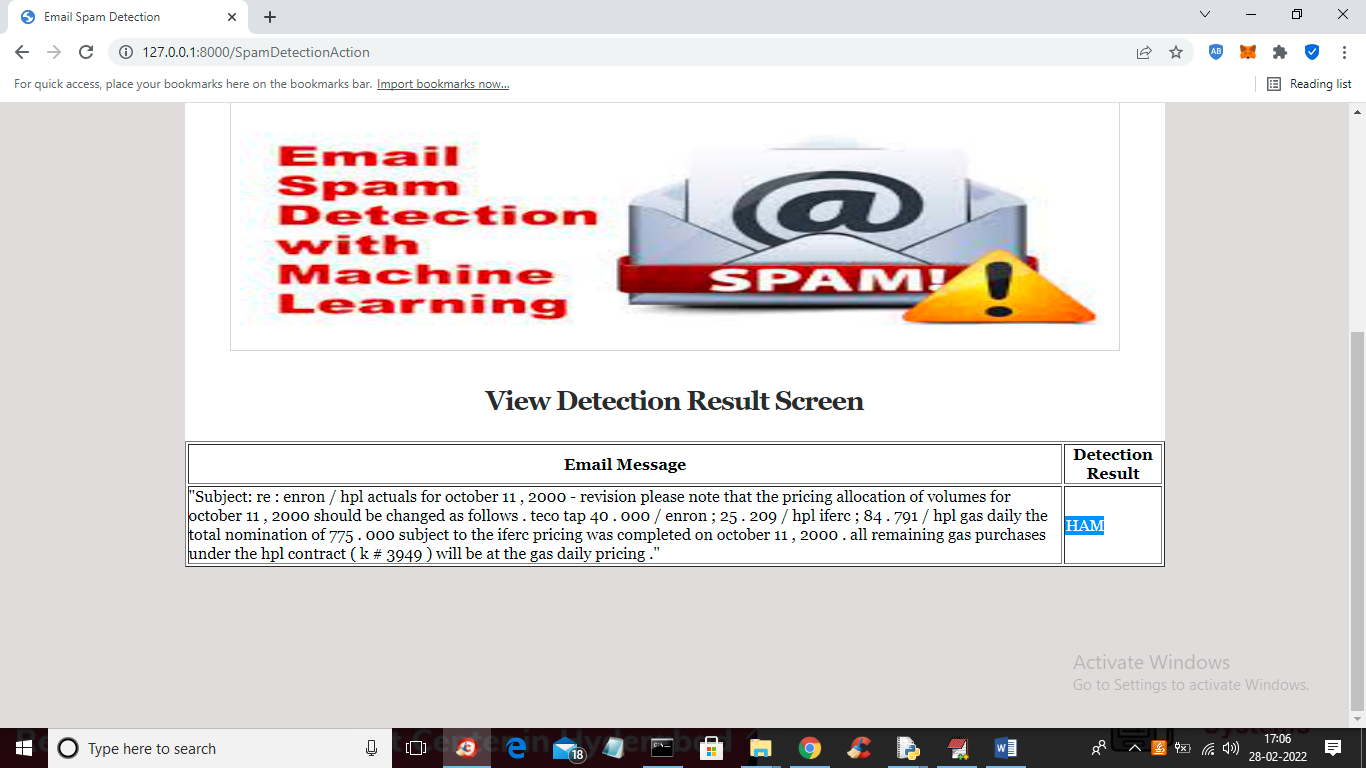
In above screen you can input some message and if you don’t know what to enter then you can copy some text from ‘testMessages.txt’ from ‘Dataset’ folder and paste in above screen



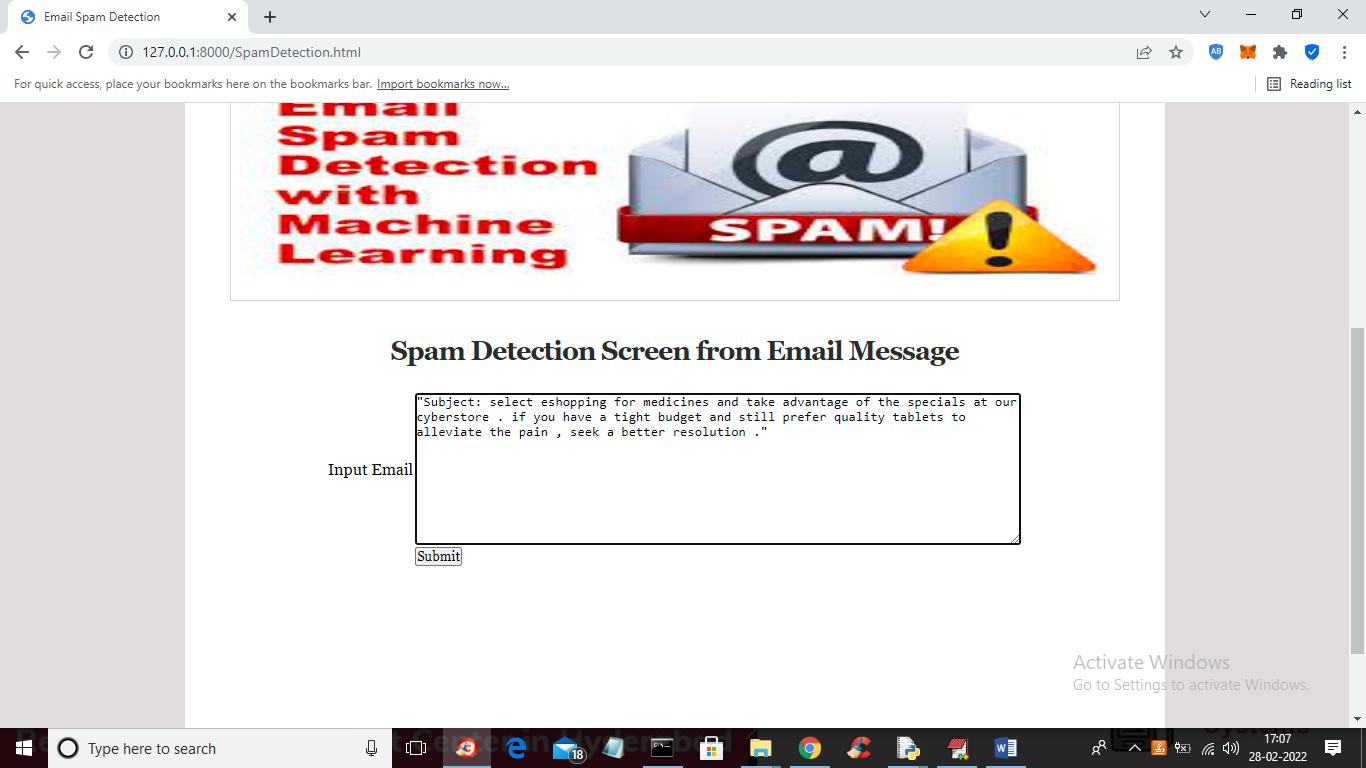
In above screen I am copying one line of text and paste in text area like below screen

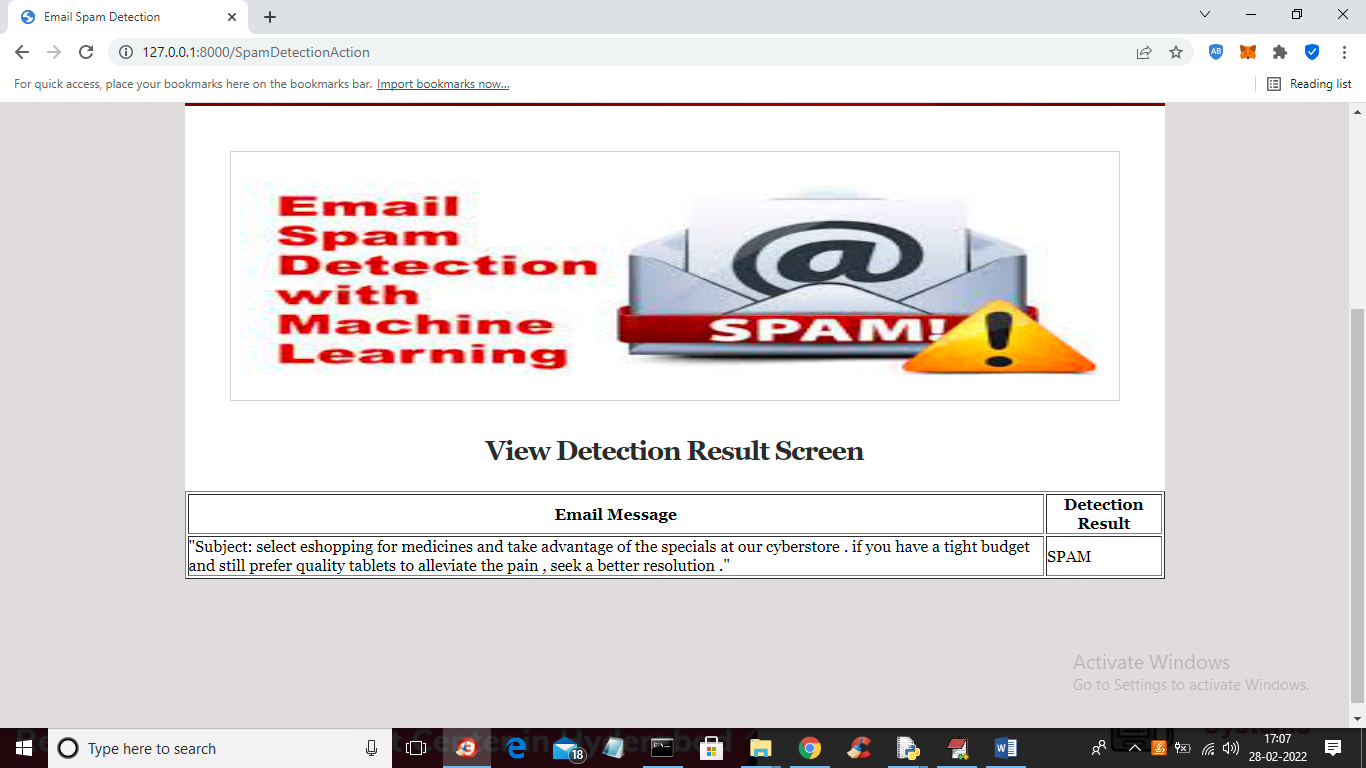


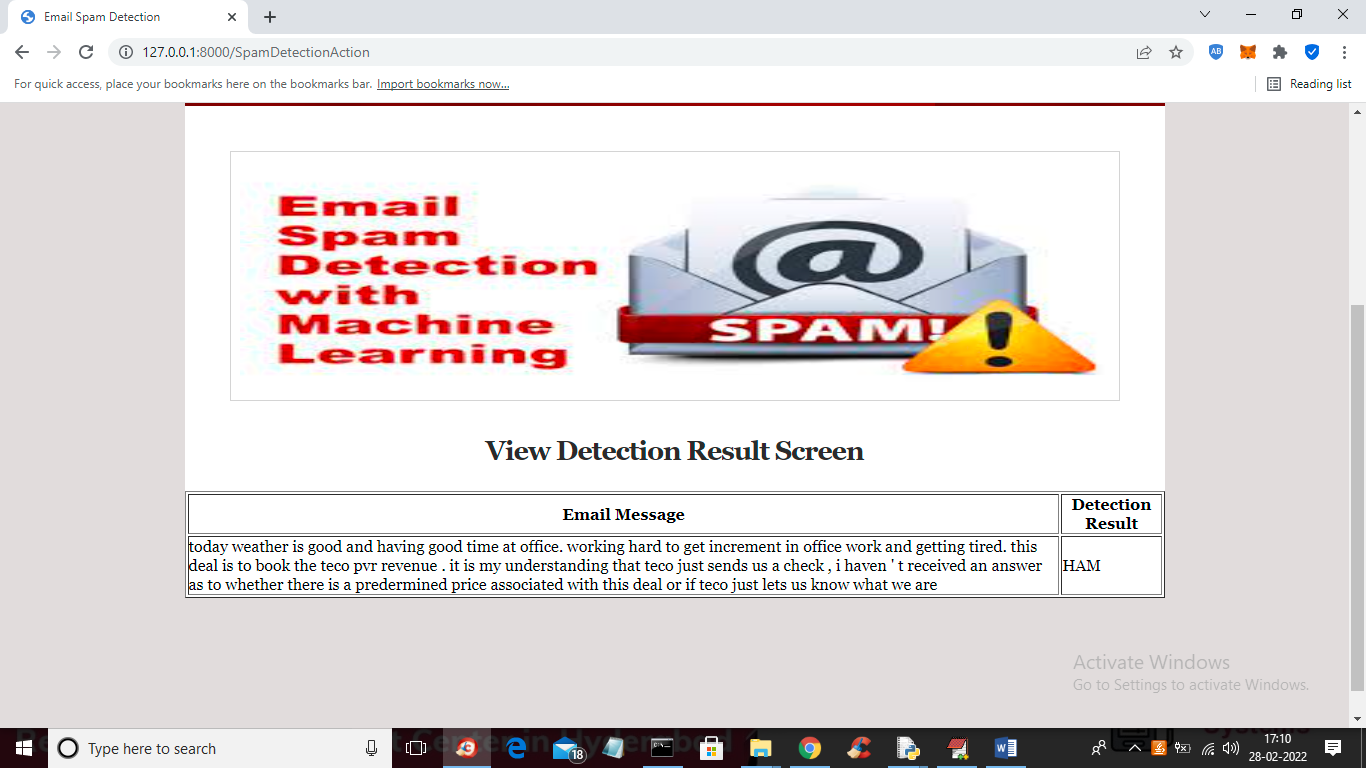
In above screen I pasted the message and press submit button to get below output



In above screen in blue colour text we can see message is detected as HAM and similarly you can paste other messages and get result







Similarly you can test other messages