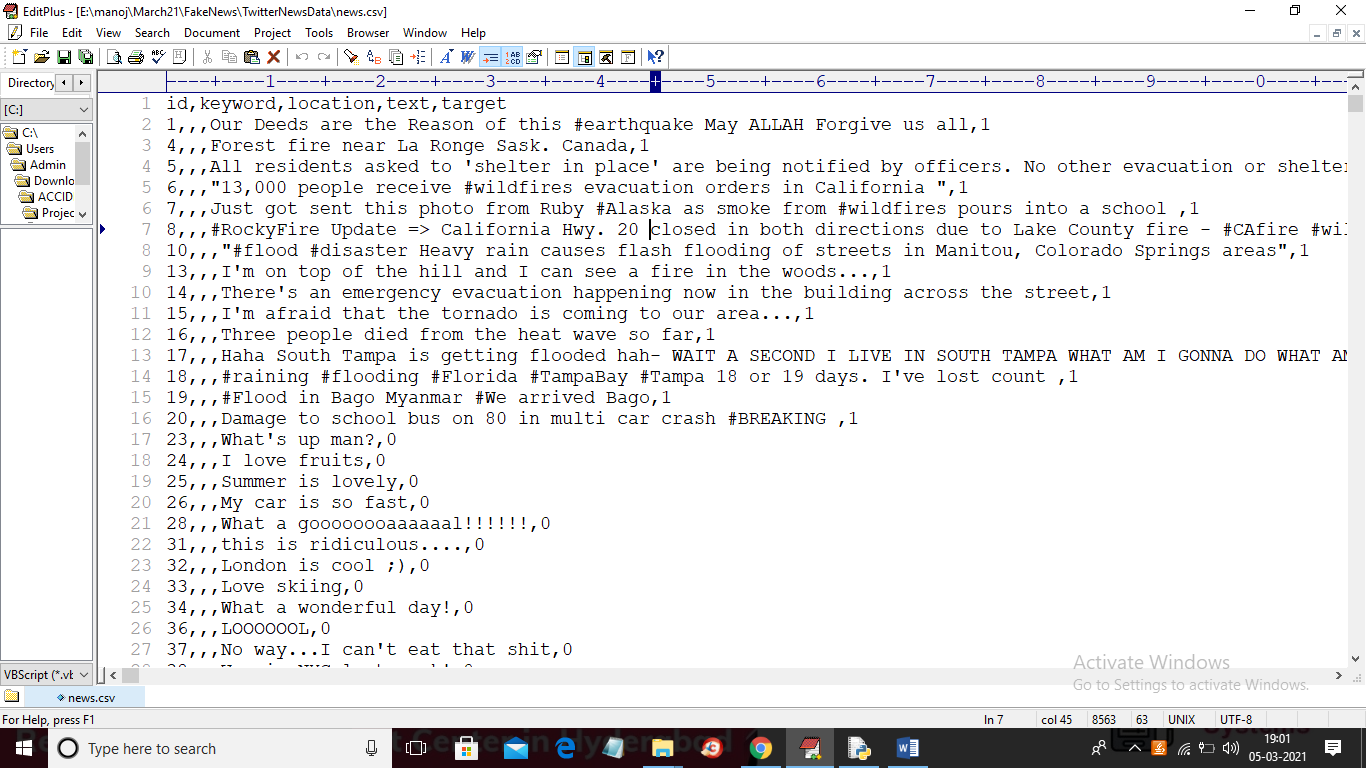
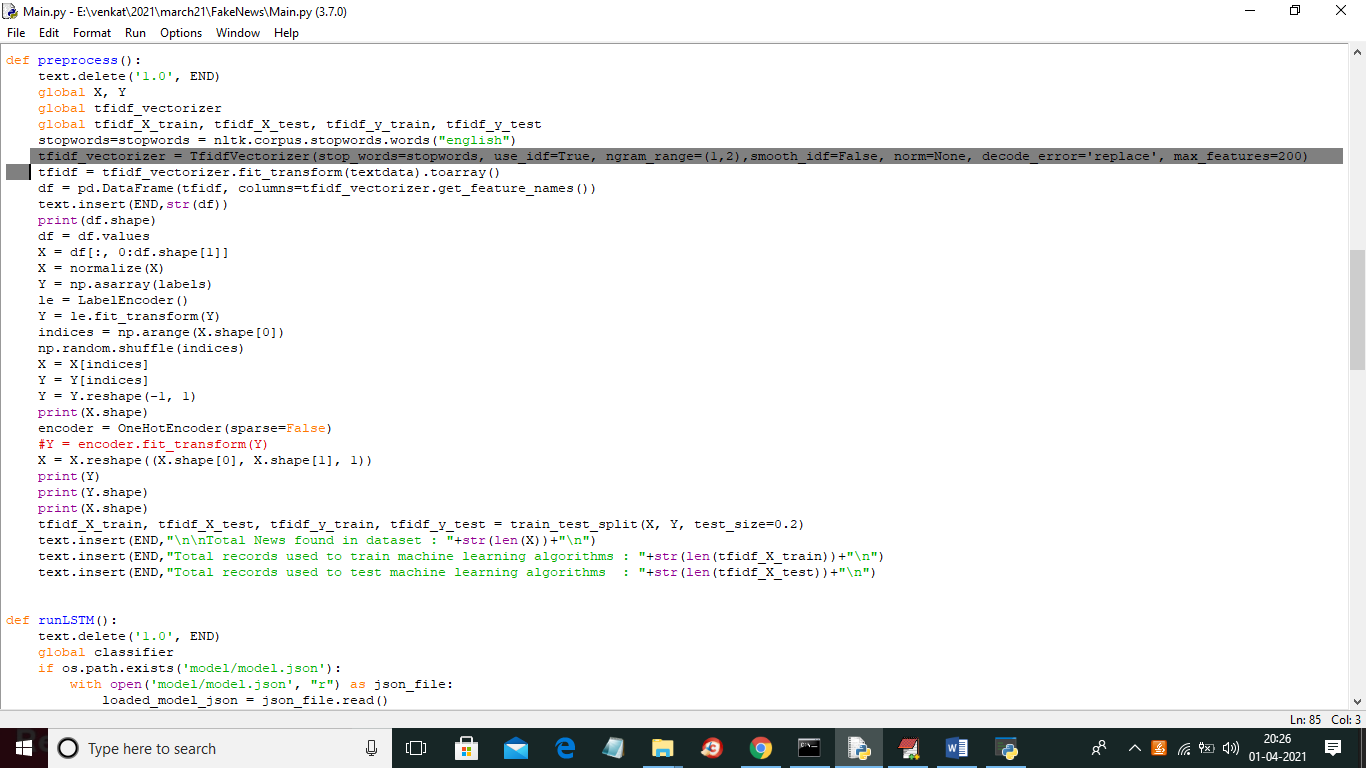
**Classifying fake news articles using natural language processing to identify in article as a supervised le**

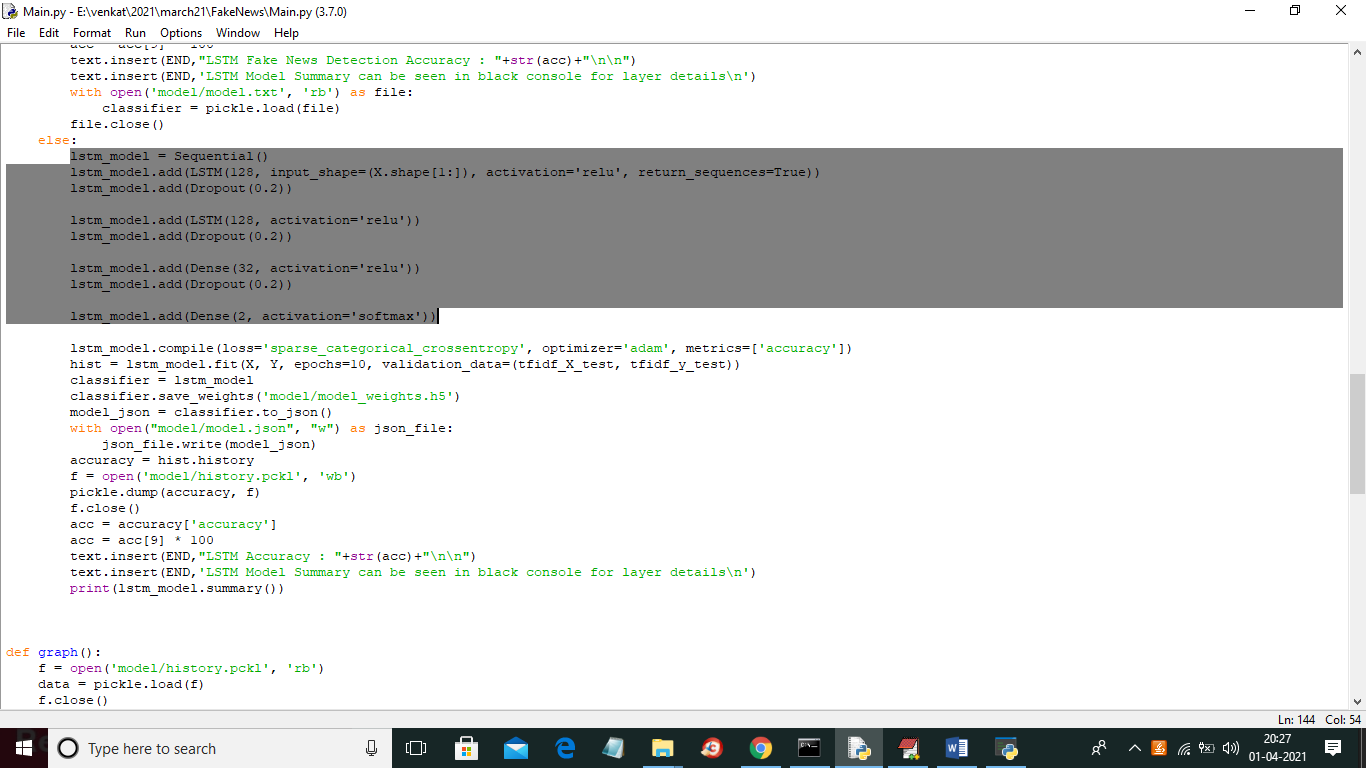
In this project we are using LSTM (Long Short Term Memory) Recurrent Neural Network to predict fake news as huge amount of fake news is gathering in all types of media such as social media or news media and to detect fake news author is training LSTM neural network with past news data label as ‘Genuine’ and ‘Fake’. We downloaded available twitter FAKE NEWS tweets from internet and below is the dataset screen shots



In above dataset we can see last two columns contains news text and target value where text contains news data and target contain class label as 0 and 1 where 0 means that news is GENUINE and 1 means that news is FAKE. In below screen we are showing code that using NGRAM to convert news data to TF-IDF vector

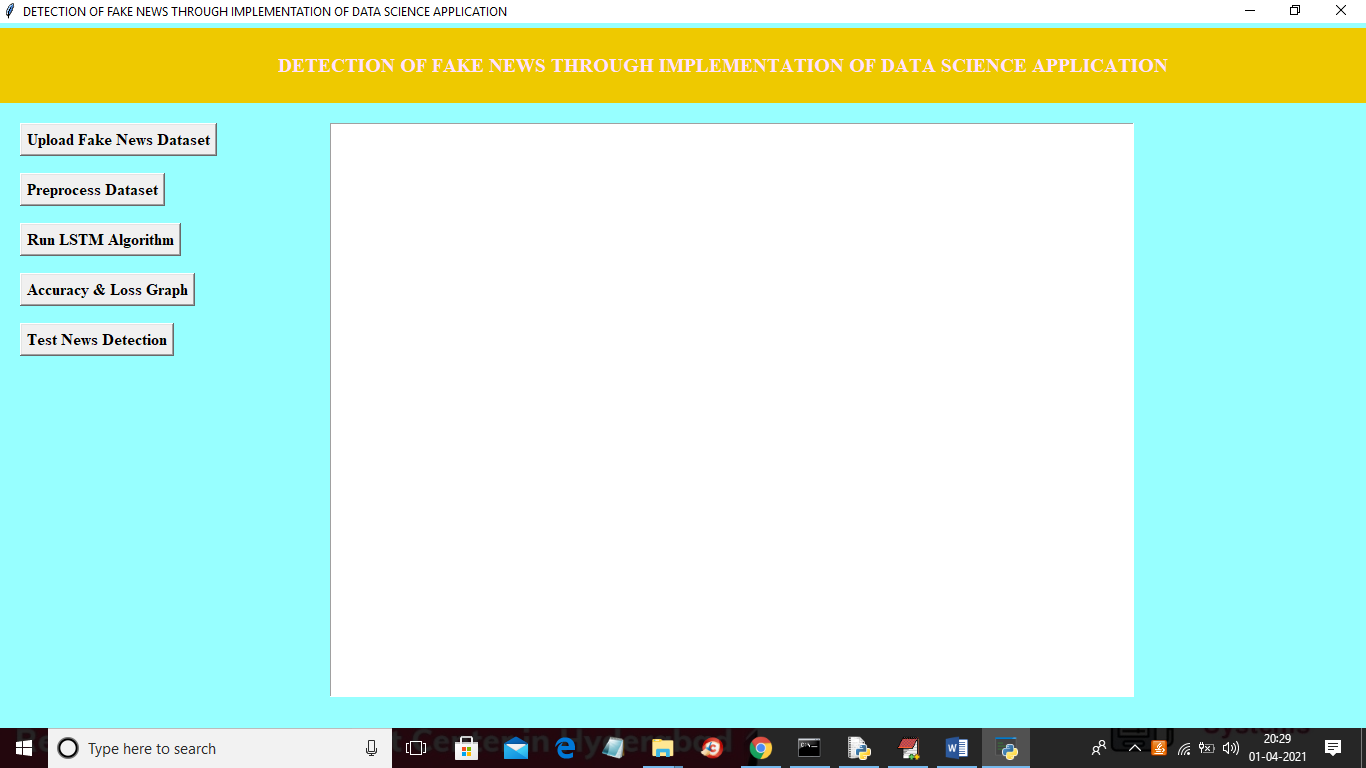


In above screen in selected line we are creating TFIDF object with NGRAM and then in next line we are applying TFIDF object on text data (news dataset) to convert news data to numeric vector which contains count of each word. In below screen showing LSTM neural network creation

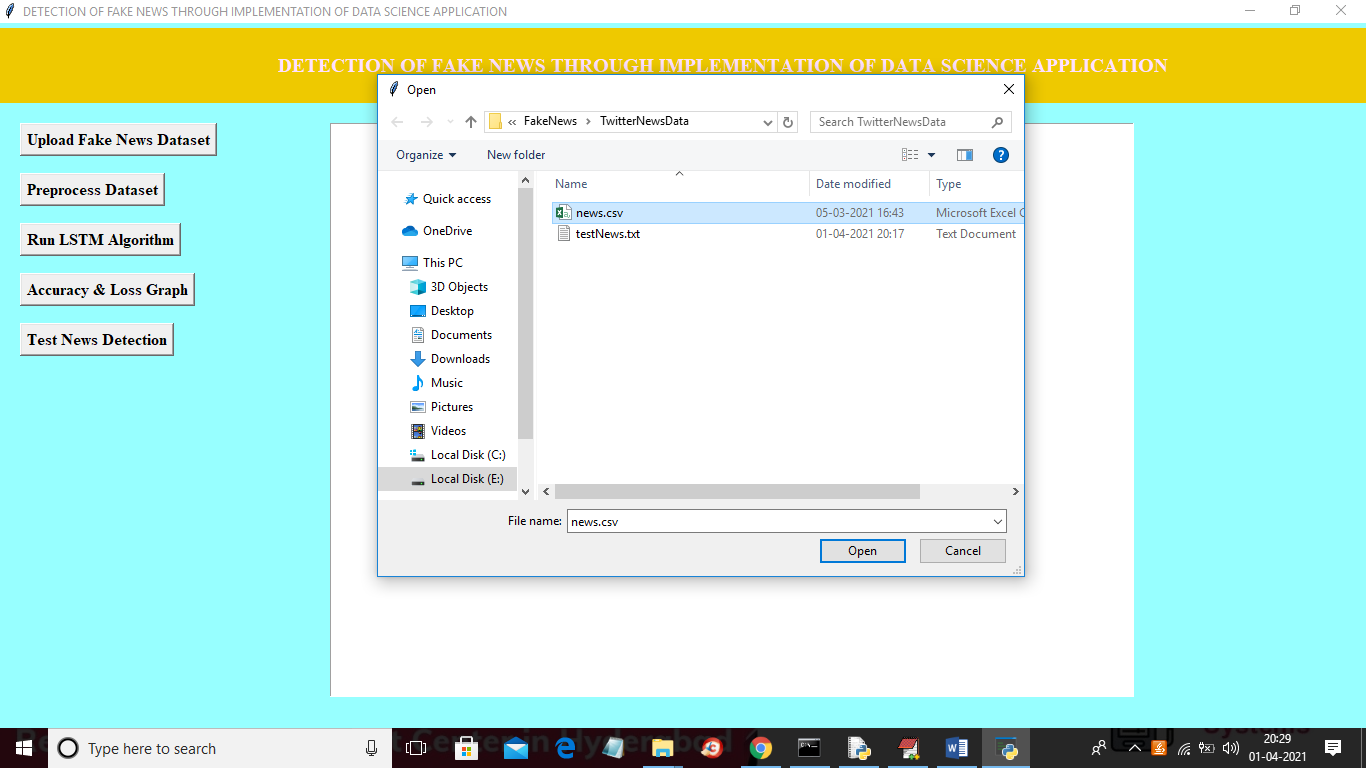


SCREEN SHOTS

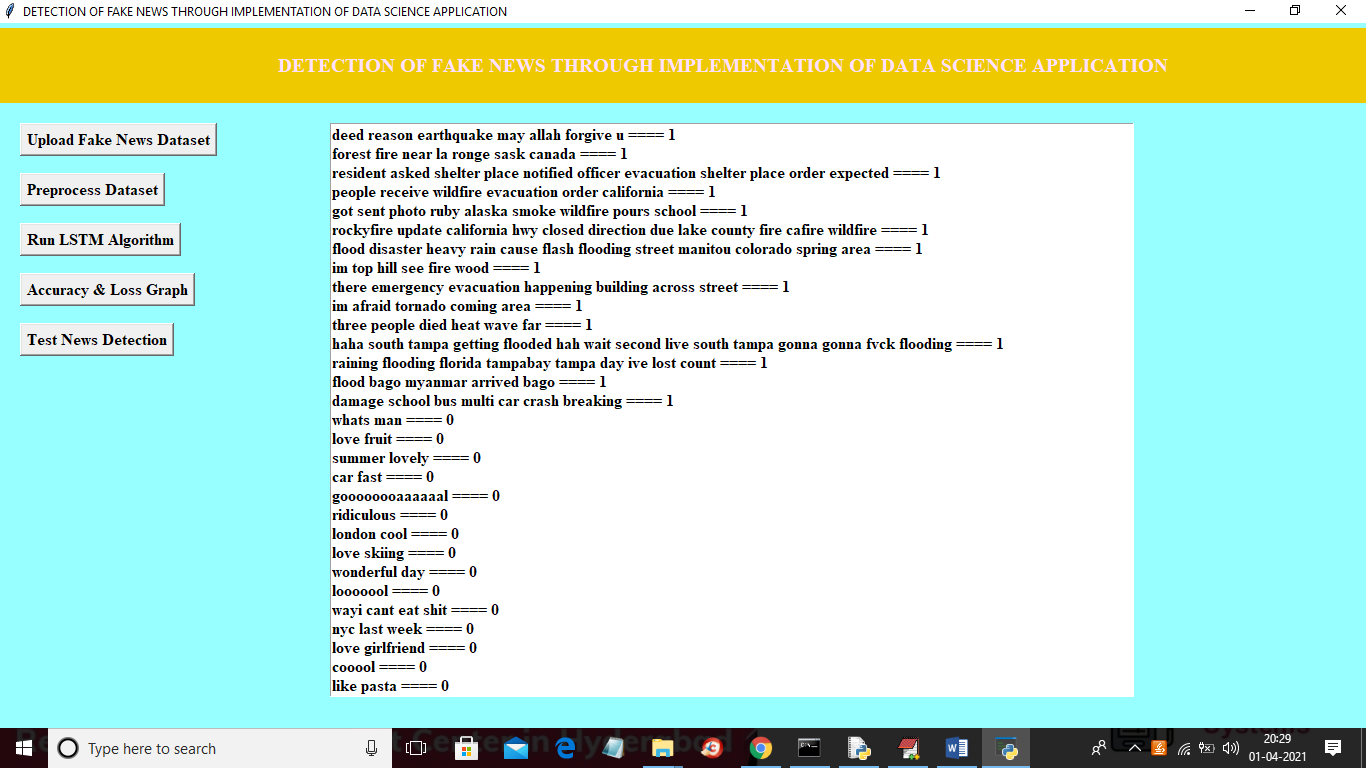
To run project double click on‘run.bat’ file to get below screen



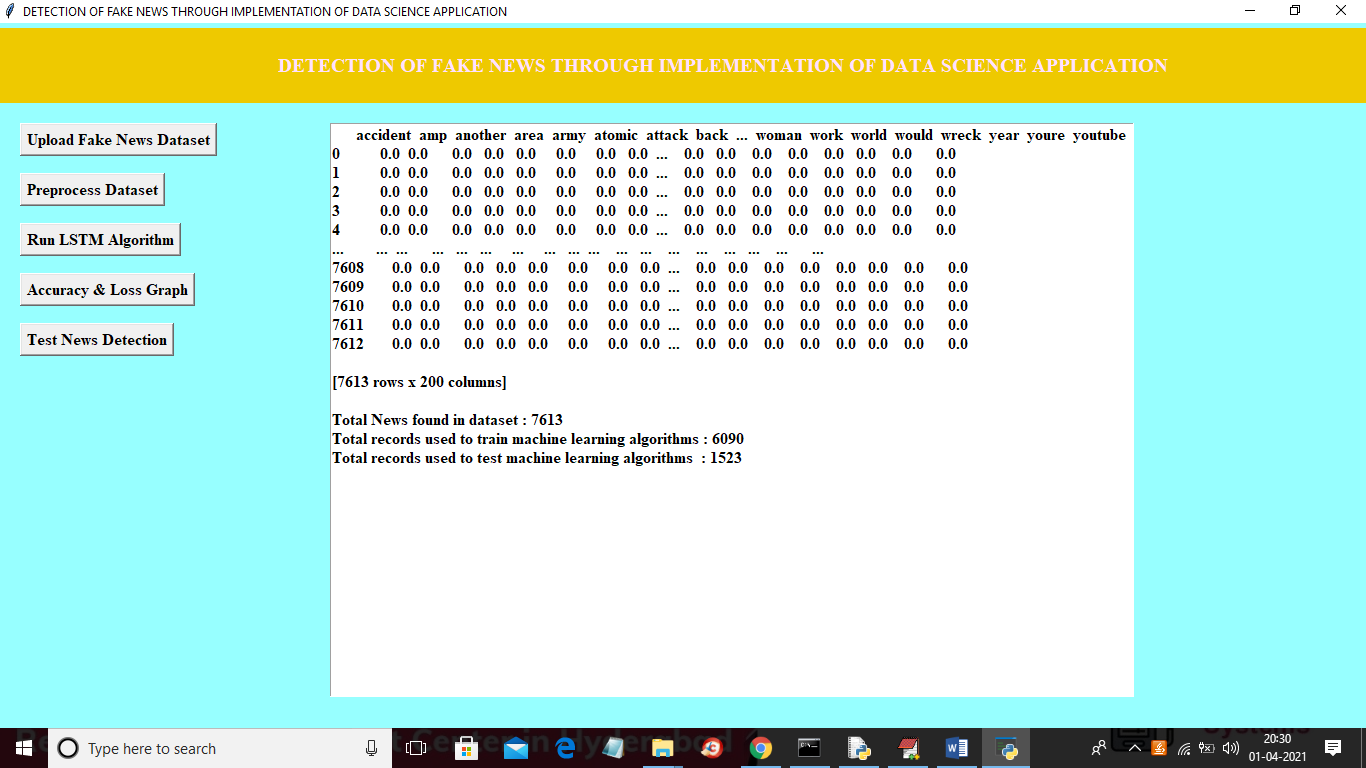
In above screen click on ‘Upload Fake News Dataset’ button to upload dataset



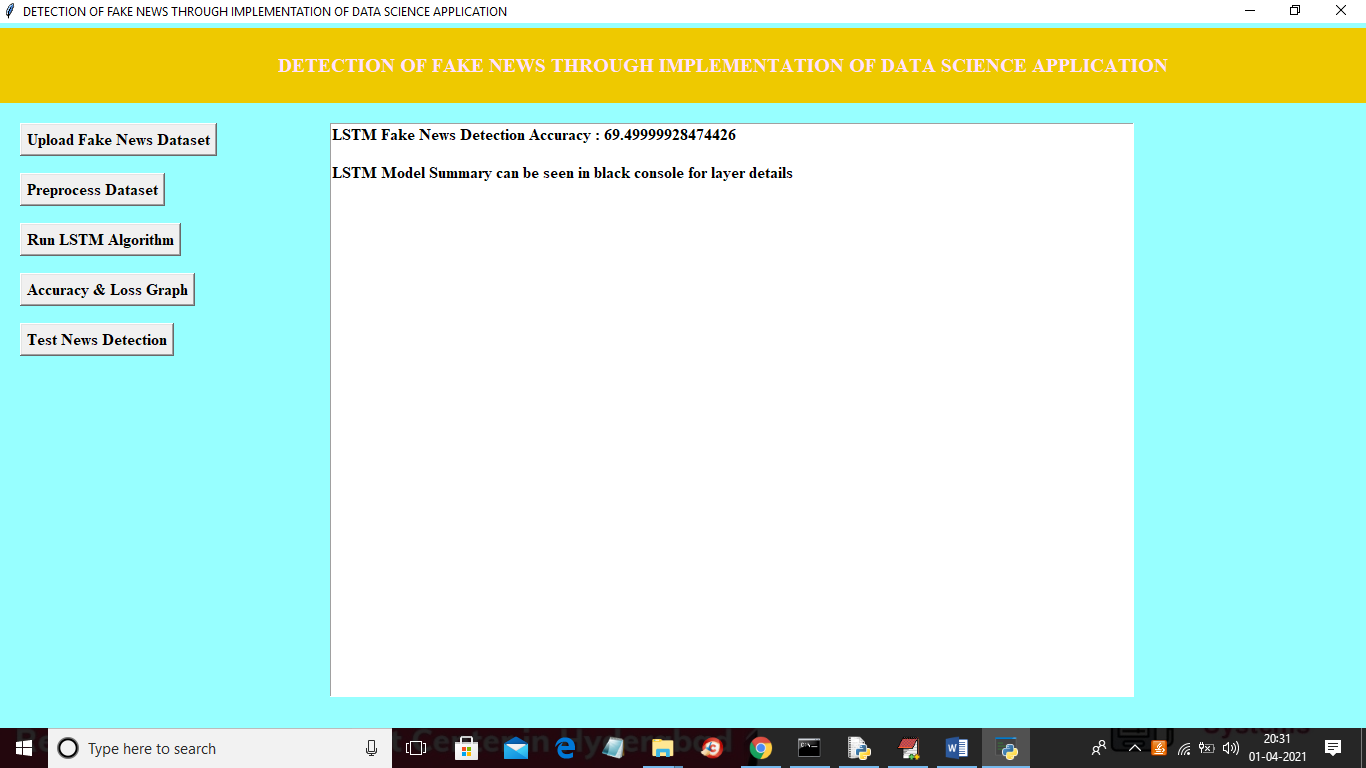
In above screen selecting and uploading ‘news.csv’ file and then click on ‘Open’ button to load dataset and to get below screen



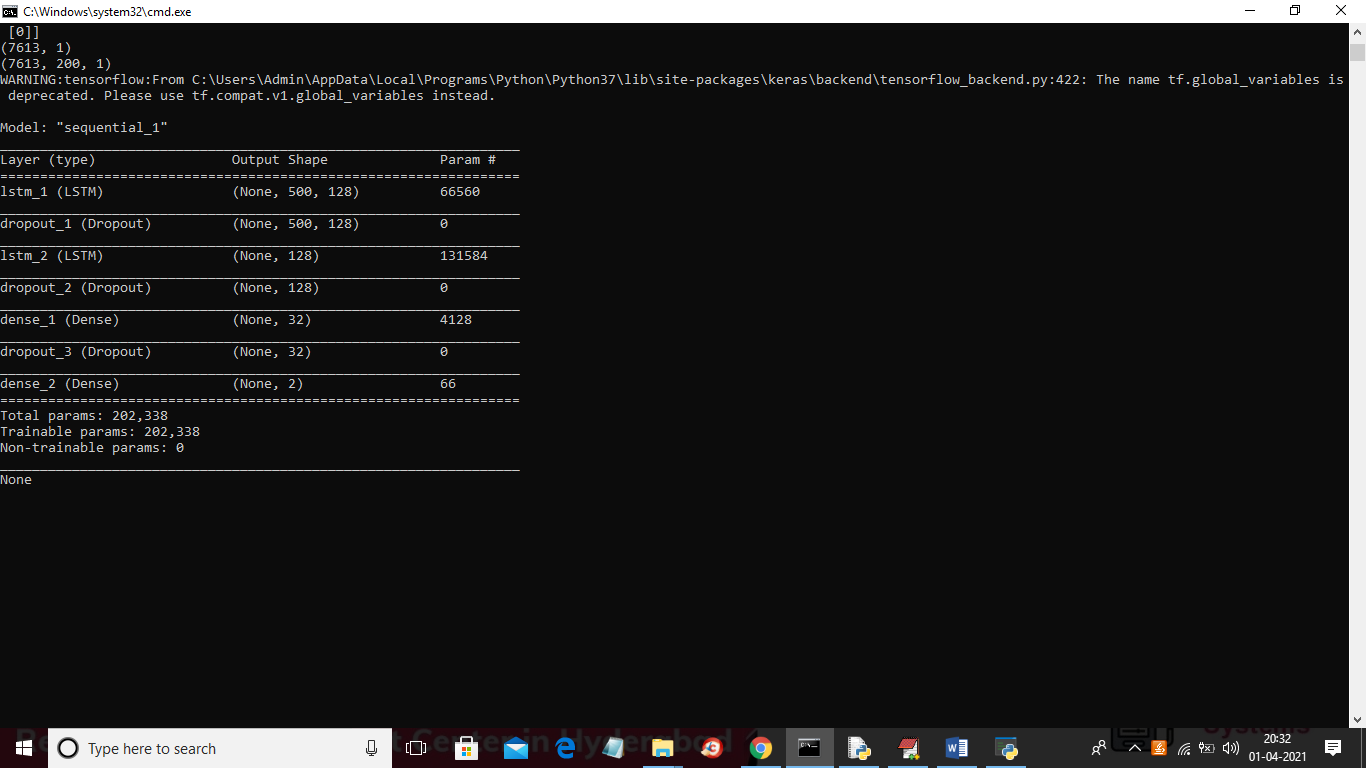
In above screen dataset loaded and then in text area we can see all news text with the class label as 0 or 1 and now click on ‘Preprocess Dataset & Apply NGram’ button to convert above string data to numeric vector and to get below screen



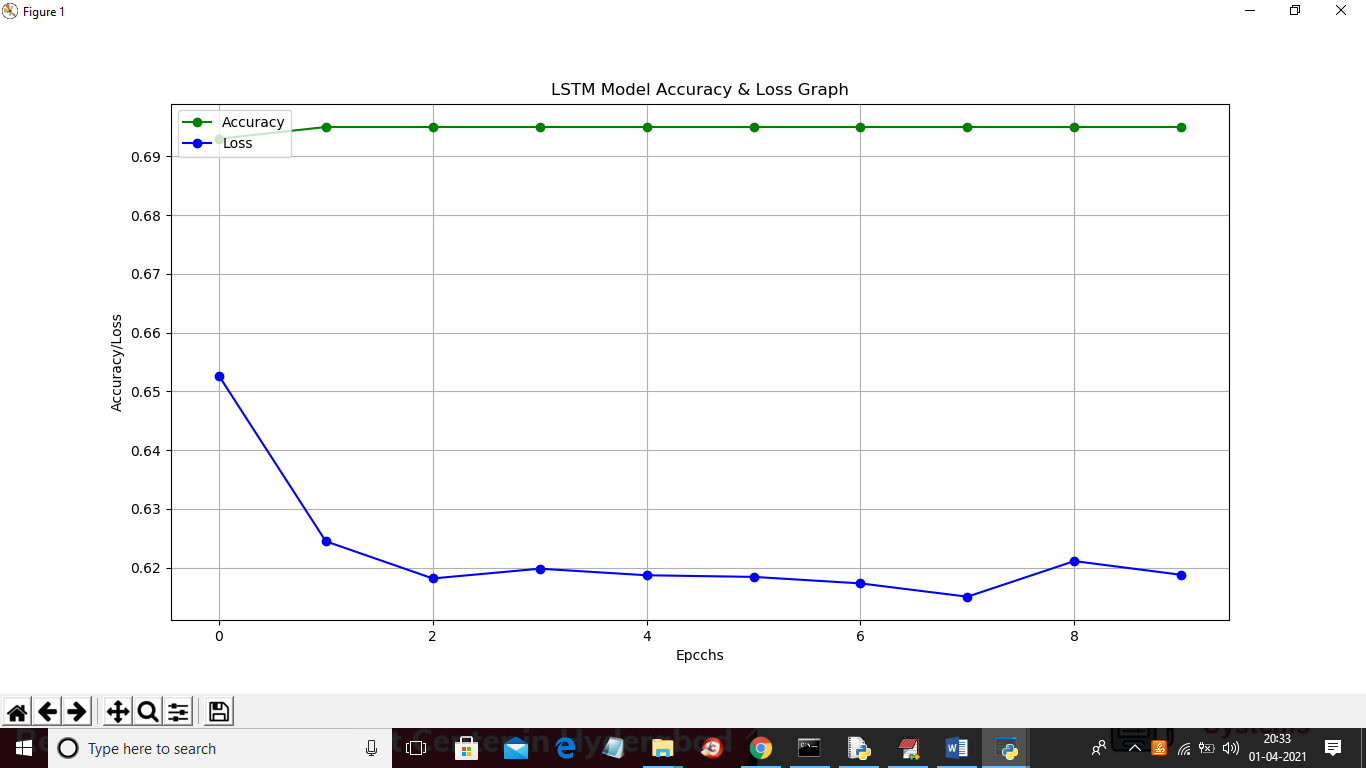
In above screen all news words put in column header and if that word appear in any row then that rows column will be change with word count and if not appear then 0 will be put in column. In above screen showing some records from total 7612 news records and in bottom lines we can see dataset contains total 7613 records and then application using 80% (6090 news records) for training and then using 20% (1523 news records) for testing and now dataset is ready with numeric record and now click on ‘Run LSTM Algorithm’ button to train above dataset with LSTM and then build LSTM model and then calculate accuracy and error rate



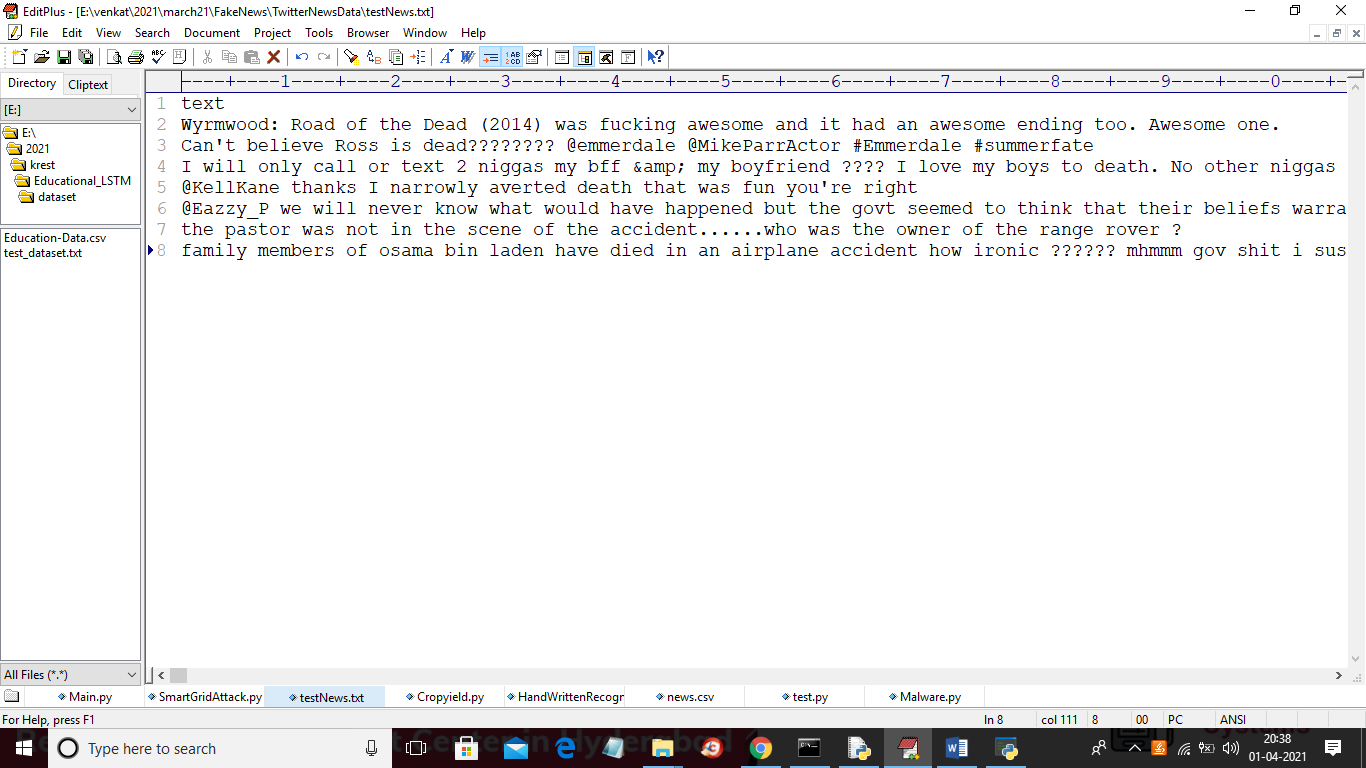
In above screen LSTM model is generated and we got its prediction accuracy as 69.49% and we can see below console to see LSTM layer details



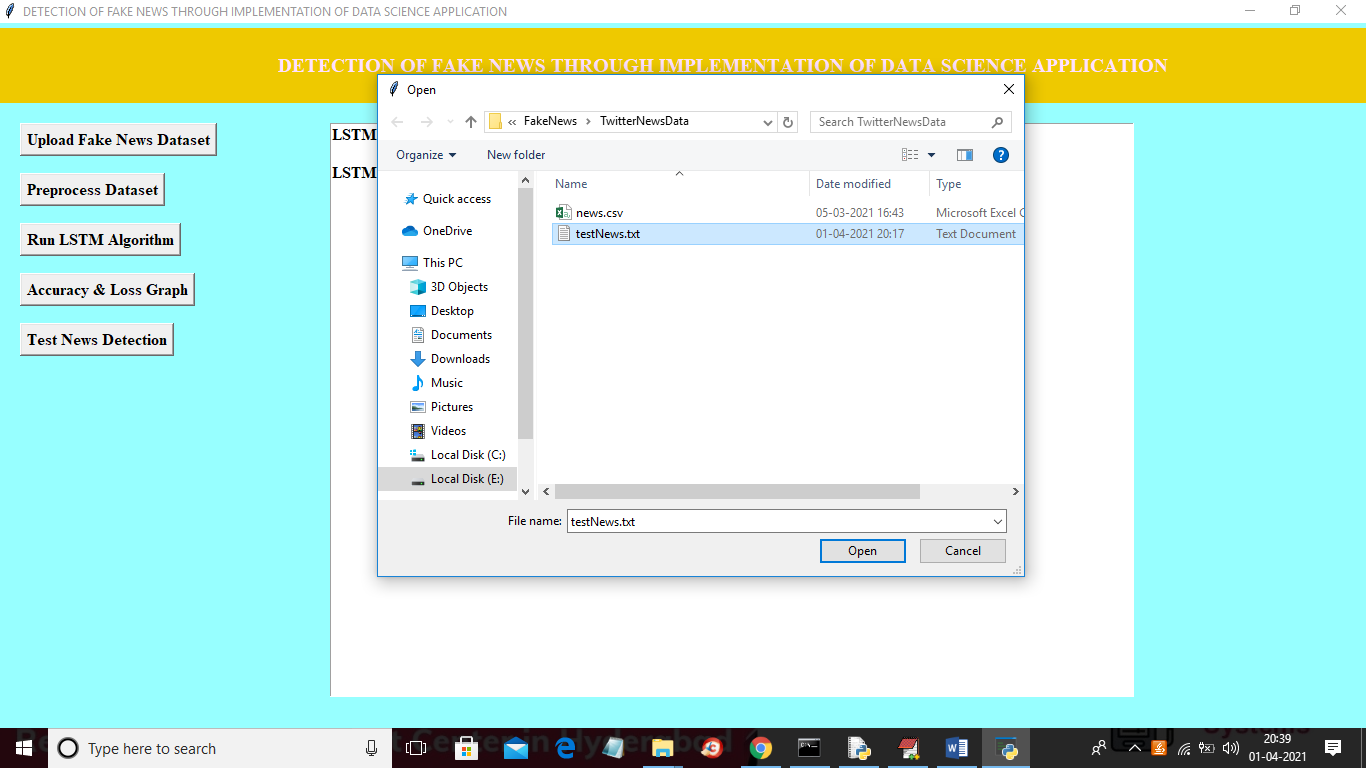
In above screen different LSTM layers are created to filter input data to get efficient features for prediction. Now click on ‘Accuracy & Loss Graph’ button to get LSTM graph



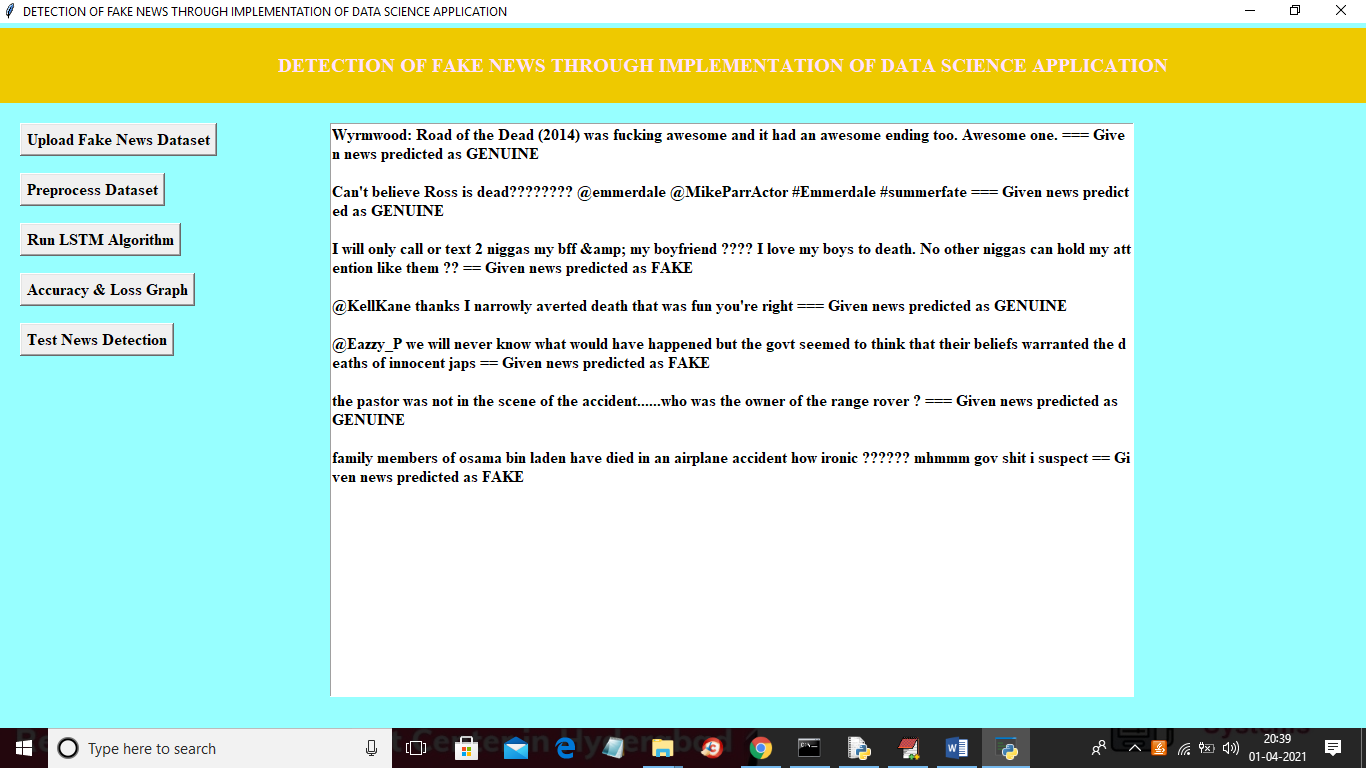
In above graph x-axis represents epoch/iterations and y-axis represents accuracy and loss value and green line represents accuracy and blue line represents loss value and at each increasing epoch loss values get decrease and accuracy reached to 70%. Now click on ‘Test News Detection’ button to upload some test news sentences and then application predict whether that news is genuine or fake. In below test news dataset we can see only TEXT data no class label and LSTM will predict class label for that test news



In above screen in test news we have only one column which contains only news ‘TEXT’ and after applying above test news we will get prediction result



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



In above screen before dashed symbols we have news text and after dashed symbol application predict news as ‘FAKE or GENUINE’. After building model when we gave any news text then LSTM will check whether more words belongs to genuine or fake category and whatever category get more matching percentage then application will predict that class label.