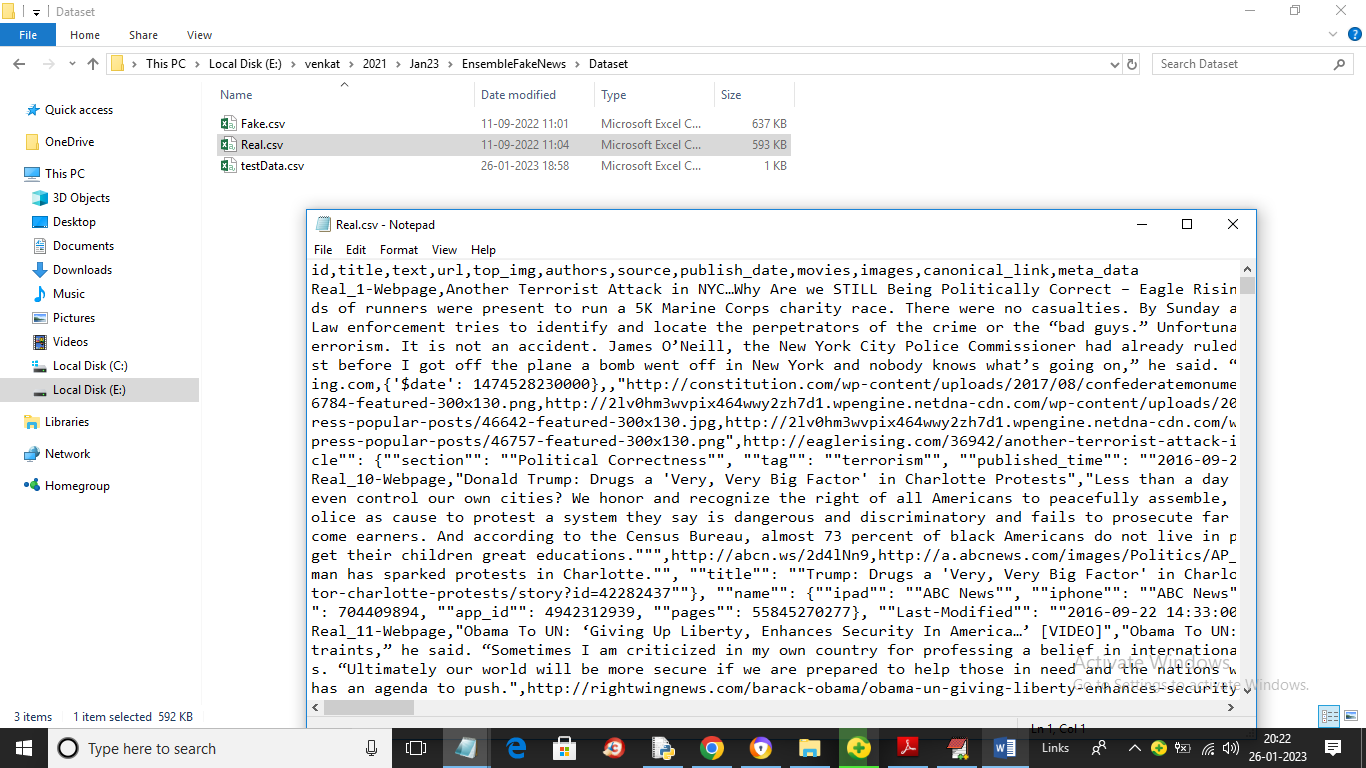
Fake Detect: A Deep Learning Ensemble Model for Fake News Detection

Due to advancement of technologies digital data is everywhere like tweets, post messages and online News. Some malicious users are taking advantage of technologies to spread fake news and this fake news can put bad influence on normal users and community. Often normal users blindly believe on news and if they believe on fake news then they may take wrong decision. Often we read products reviews before purchasing and if malicious users give fake good review on any product then user will believe as good product and purchase it.

To overcome from above issue many machine learning and deep learning algorithms are introduced but they lack of optimization so their prediction accuracy is not accurate. In propose paper author employing ensemble algorithms by combing Deep Learning and BI-LSTM-GRU where deep learning algorithm and BI-LSTM-GRU algorithm will get trained on dataset and then extract optimized features to form an ensemble model and this ensemble model will analyse NEWS and then predict it as Fake or Real.

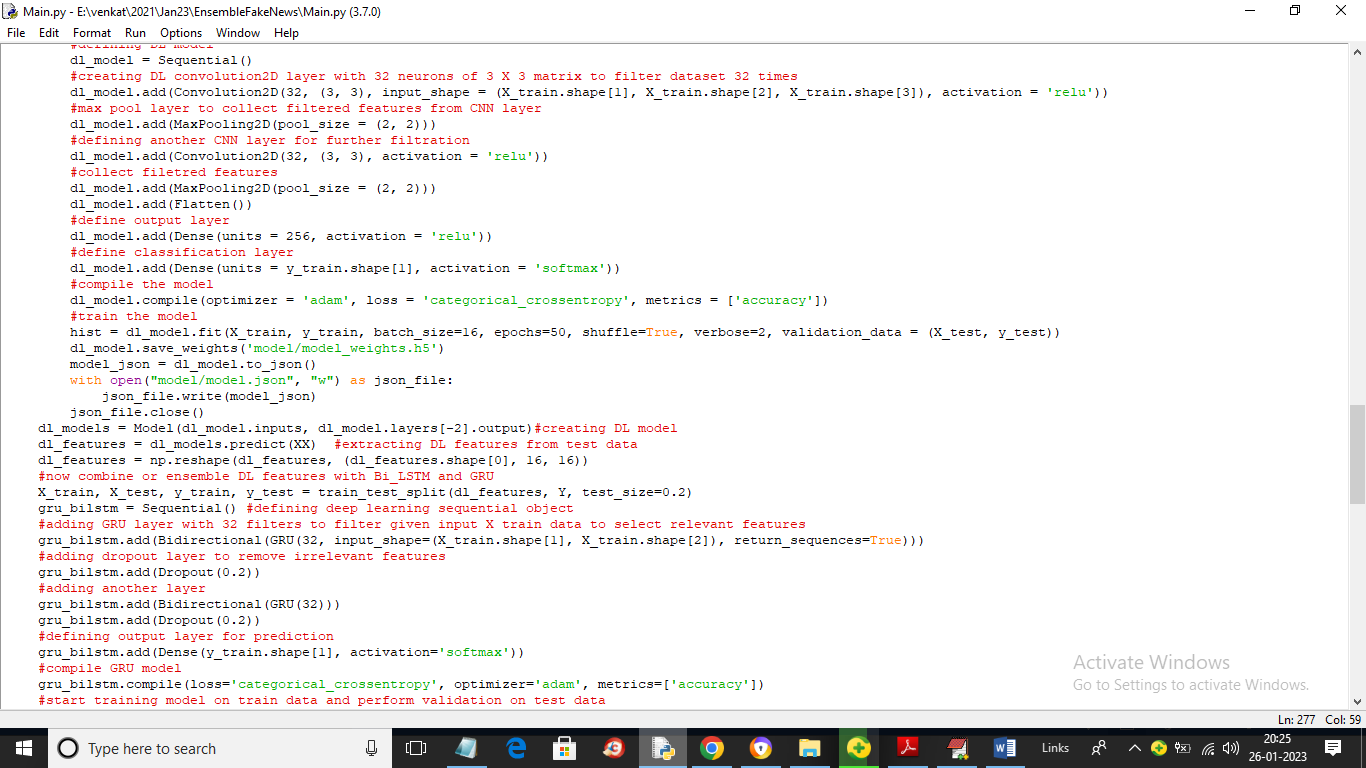
In propose work author comparing performance of ensemble DL-BI-LSTM-GRU with existing SVM, CNN and many other algorithms and in all algorithms propose ensemble DL-BI-LSTM-GRU is giving better accuracy.

To train all algorithms author using LIAR dataset and we are also using same dataset which consist of FAKE and REAL news. Below screen showing dataset details.



In above dataset screen we have two folders called FAKE and REAL which contains news text and by using above dataset will evaluate performance of all algorithms.

In below screen we are showing code for ensemble of Deep Learning CNN and BI-LSTM-GRU algorithms



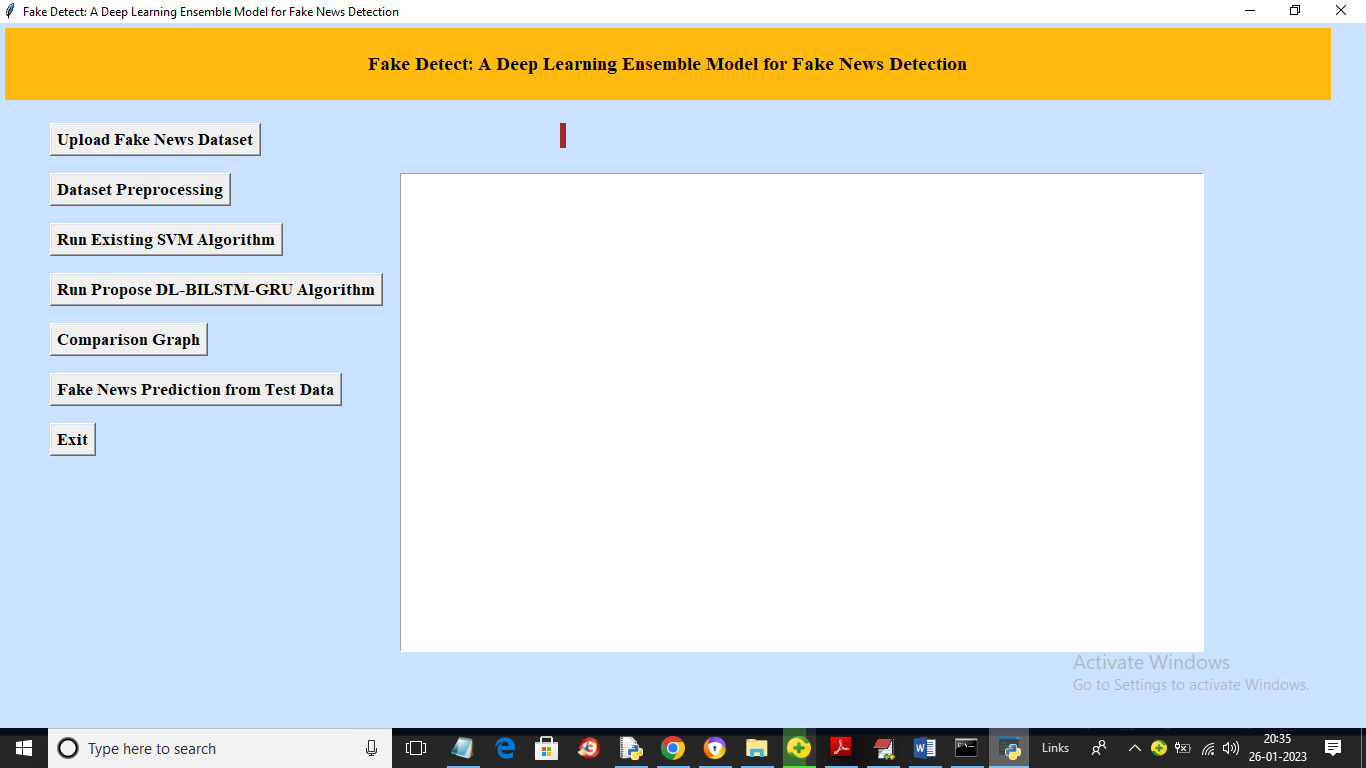
In above screen read red colour comments to know about propose Deep Learning CNN and BI-LSTM-GRU algorithms

To implement this project we have designed following modules

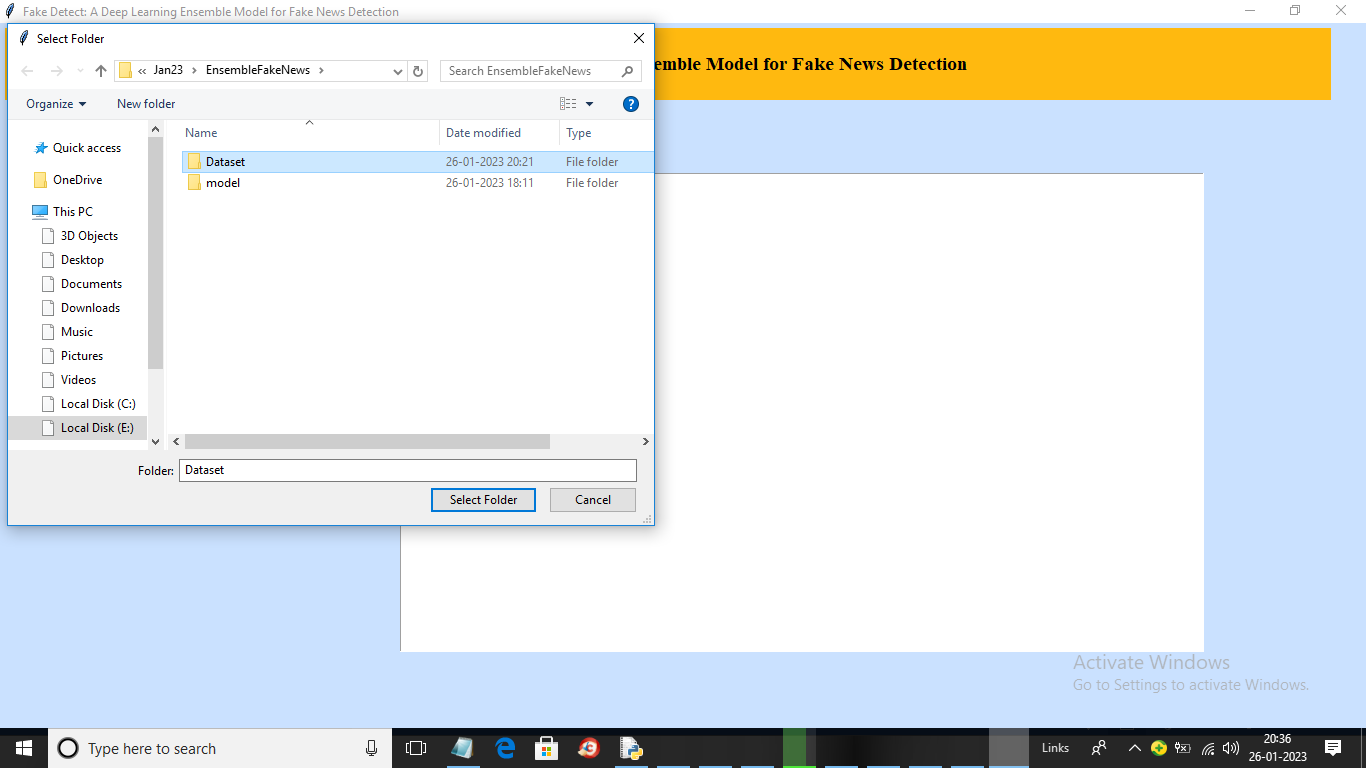
1. Upload Fake News Dataset: using this module we will upload dataset to application and then application will read all news and then plot WORD cloud graph
2. Dataset Preprocessing: using this module we will remove stop words, apply stemming and lemmatization and then convert all TEXT news into numeric vector by finding average frequency of each word
3. Run Existing SVM Algorithm: using this module we will split processed numeric vector into train and test where application using 80% dataset for training and 20% for testing. 80% dataset vector will be input to SVM to train a model and then 20% test data will be applied on trained model to calculate prediction accuracy
4. Run Propose DL-BILSTM-GRU Algorithm: using this module we will split processed numeric vector into train and test where application using 80% dataset for training and 20% for testing. 80% dataset vector will be input to Propose DL-BILSTM-GRU Algorithm to train a model and then 20% test data will be applied on trained model to calculate prediction accuracy
5. Comparison Graph: using this module we will plot accuracy comparison graph between existing SVM and propose algorithms
6. Fake News Prediction from Test Data: using this module we will input TEST news and then propose DL-BILSTM-GRU Algorithm will analyse TEST news and predict it as FAKE or REAL

SCREEN SHOTS

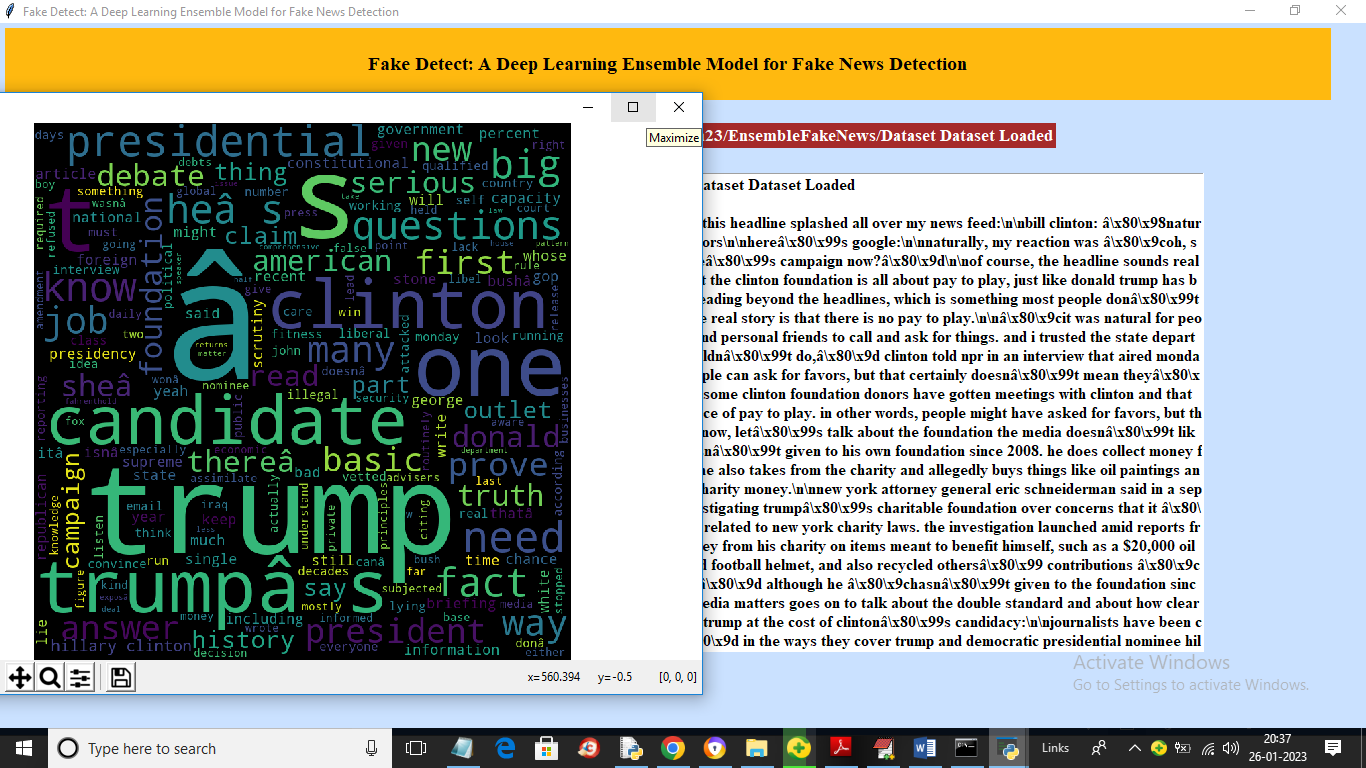
To run project double click on ‘run.bat’ file to get below screen



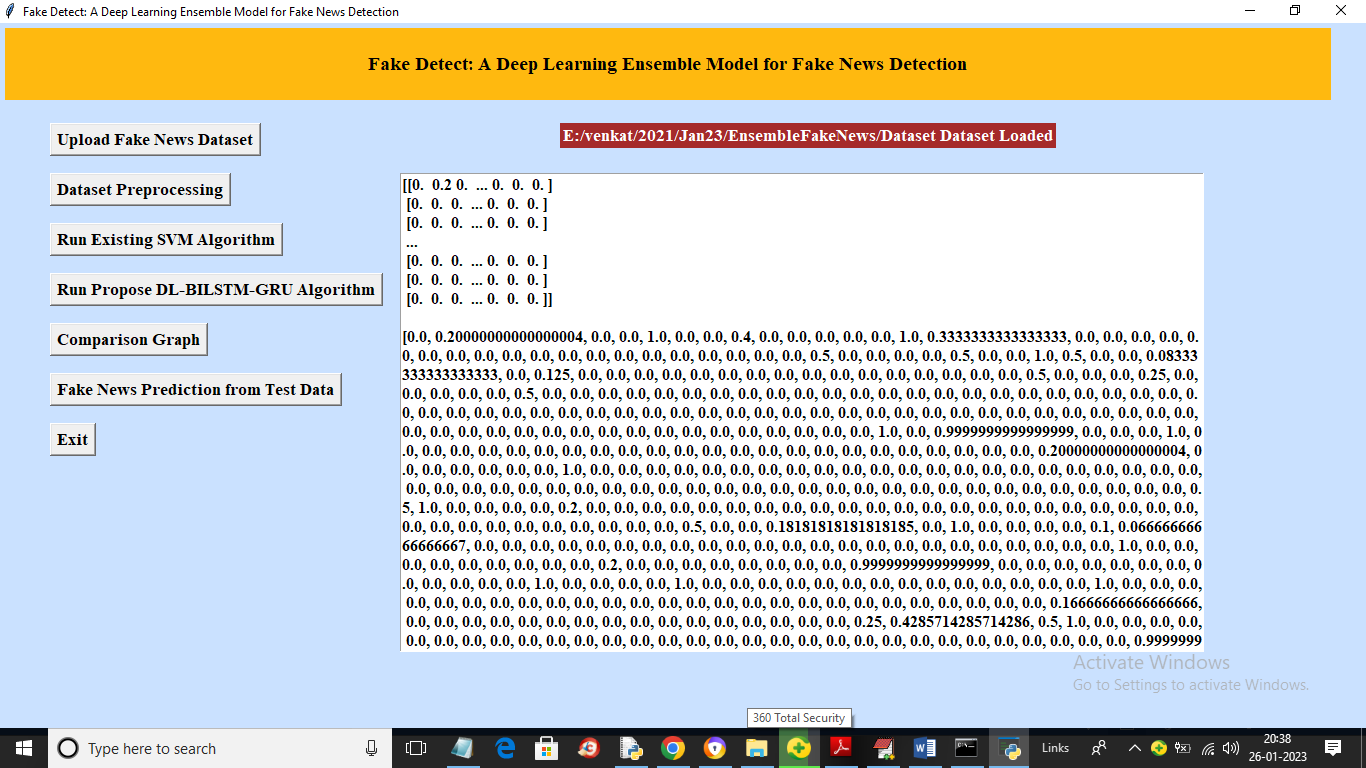
In above screen click on ‘Upload Fake News Dataset’ button to upload dataset and get below output



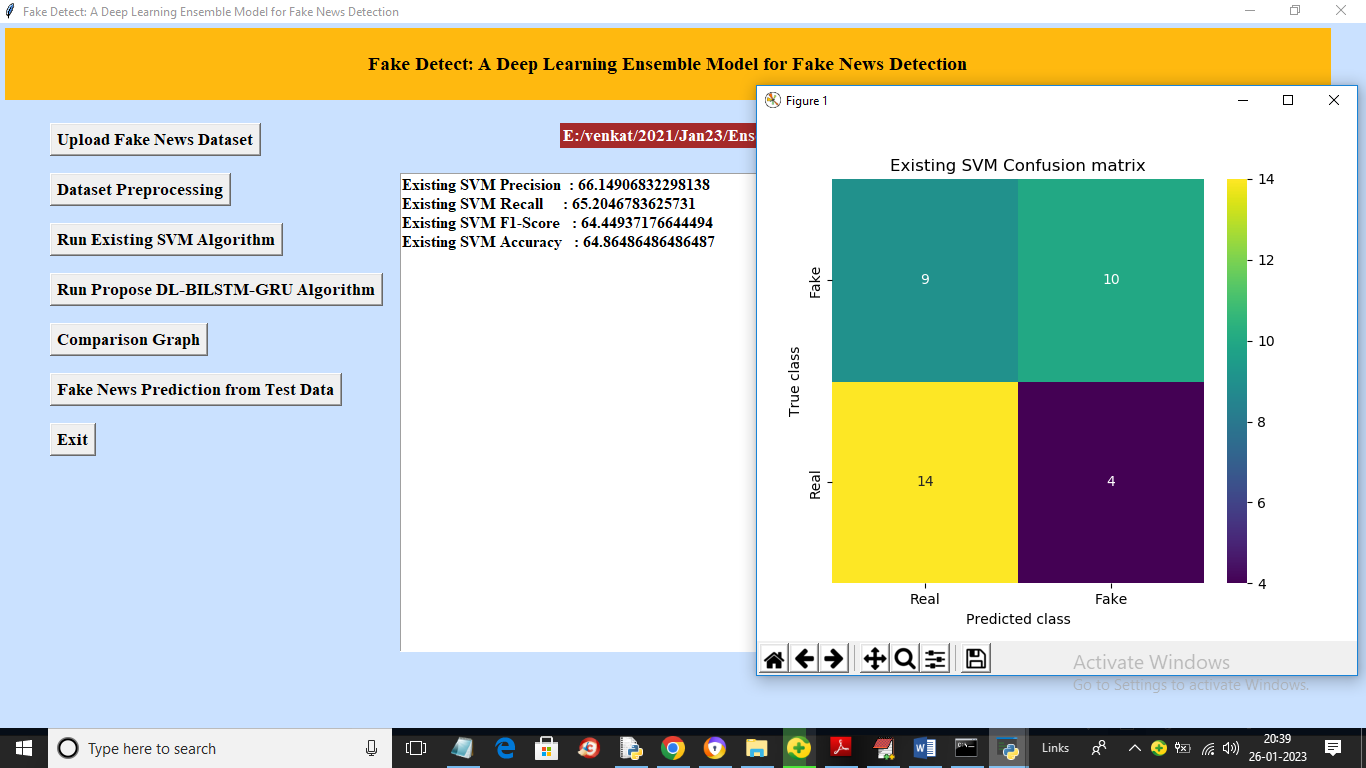
In above screen selecting and uploading entire ‘Dataset’ folder and then click on ‘Select Folder’ button to load dataset and get below output



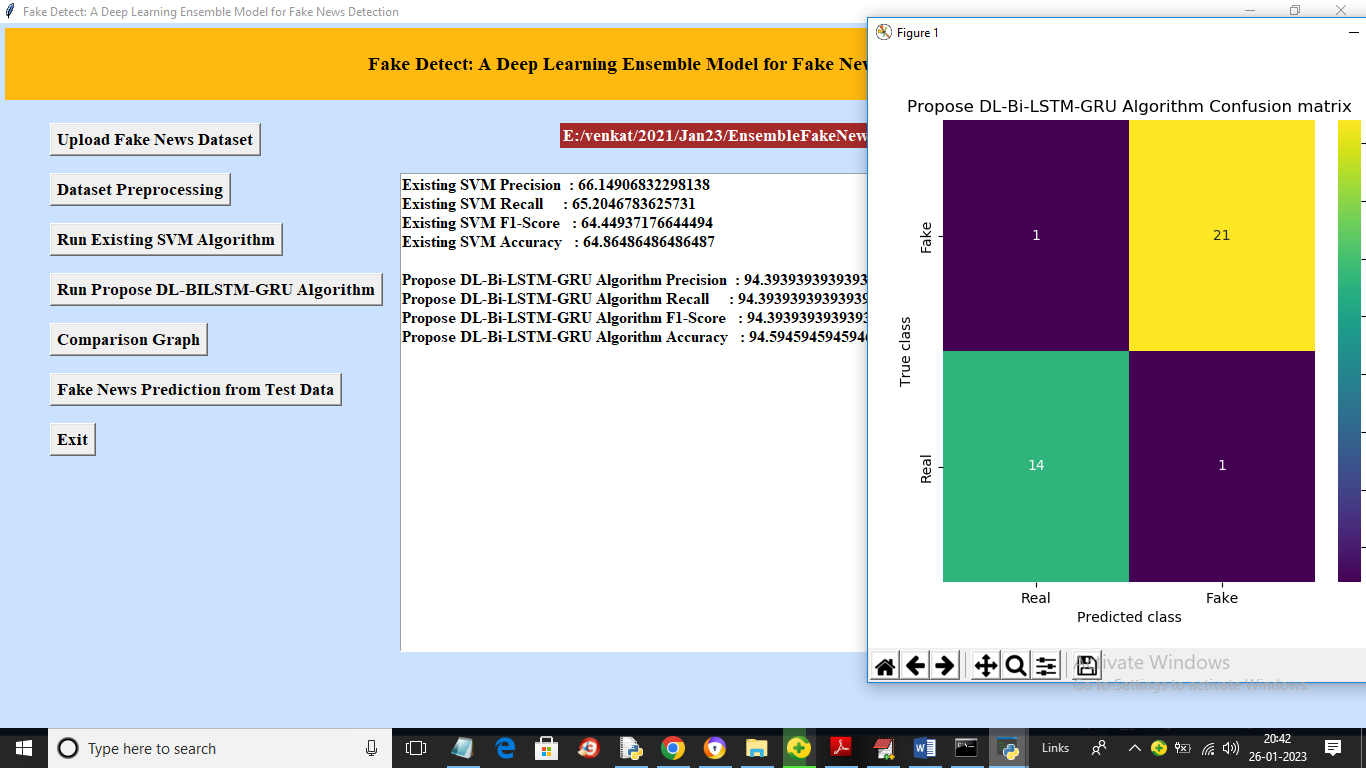
In above screen in text area we can see all text news dataset loaded and in word cloud graph whatever word has highest frequency will have more bold size and now close above graph and then click on ‘Dataset Preprocessing’ button to remove stop words, apply stemming and lemmatization and then convert all text data into numeric vector and get below output



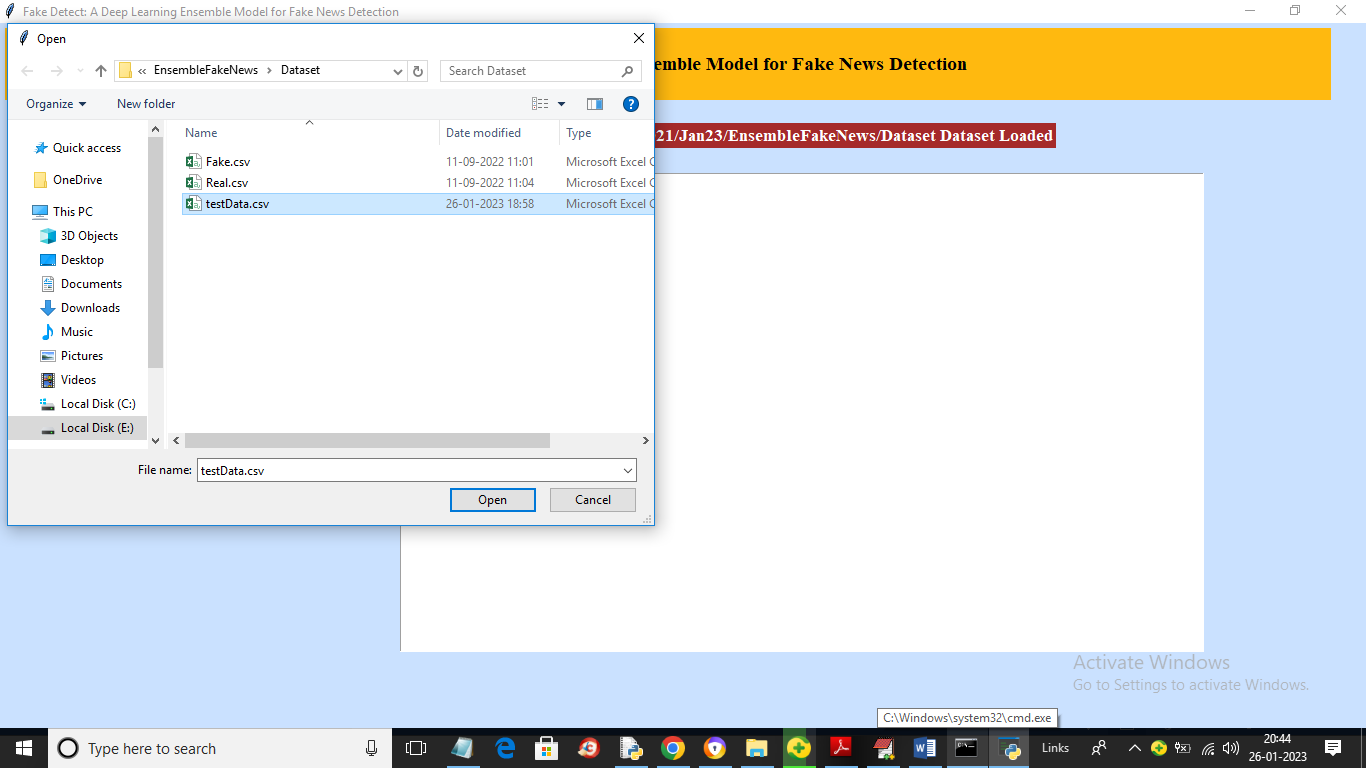
In above screen entire dataset converted into numeric vector by replacing each words with its average frequency. Now click on ‘Run Existing SVM Algorithm’ button to train SVM and get below output



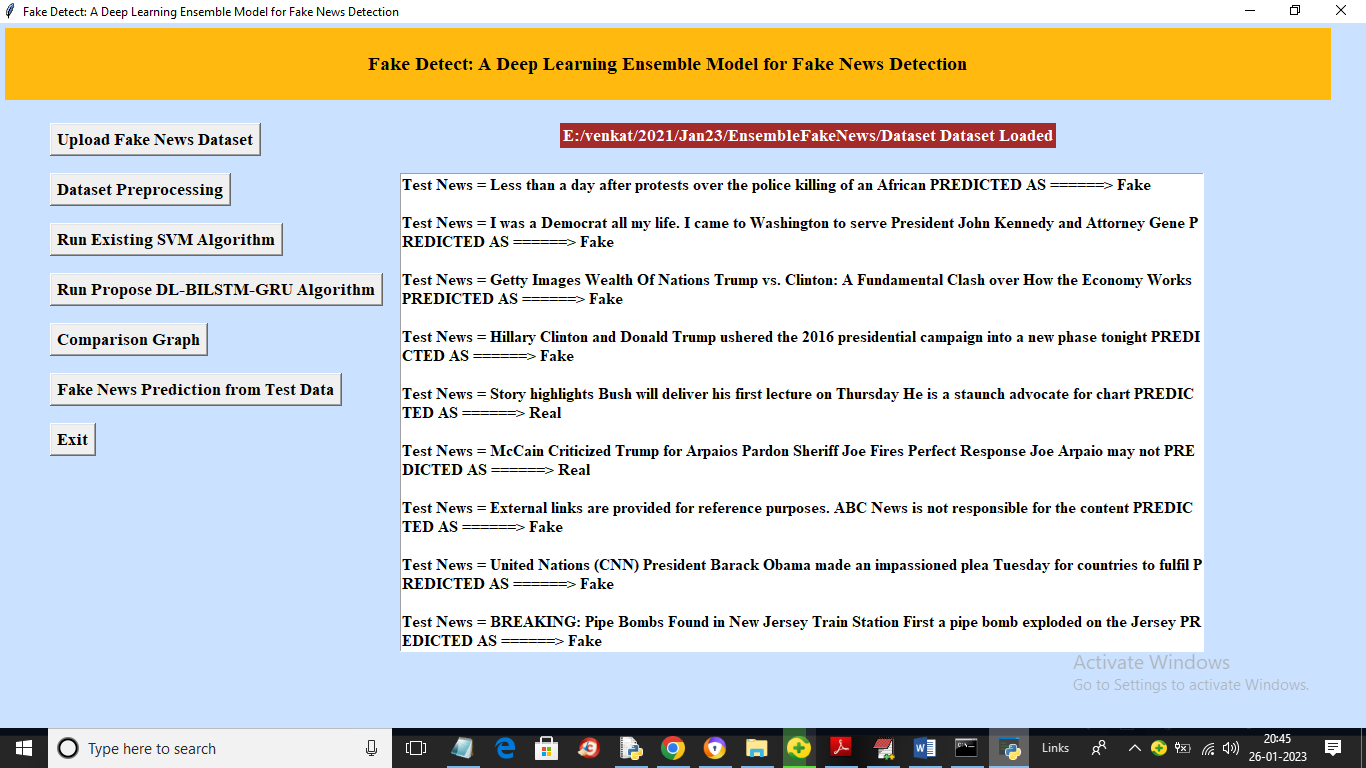
In above screen with SVM we got 66% accuracy and in confusion matrix graph x-axis represents Predicted labels and y-axis represents TRUE labels and yellow box and green box in diagnol contains correct prediction count and blue box and green box in diagnol contains incorrect prediction count which are 9 and 4. 34 and 10 are correct prediction count and now close above graph and then click on ‘Run Propose DL-BILSTM-GRU Algorithm’ button to train propose algorithm and get below output



In above screen with propose algorithm we got 94% accuracy and in confusion matrix graph blue boxes contains incorrect prediction count which is only 2 and green and yellow box in diagnol contains correct prediction count. Now close above graph and then click on ‘Fake News Prediction from Test Data’ button to upload test data and get below output



In above screen selecting and uploading ‘testData.csv’ file which contains TEST news and then click on ‘Open’ button to get below prediction output



In above screen after ‘=’ symbol we can see TEST news and after arrow symbol =🡺 we can see predicted output as ‘Fake’ or ‘Real’