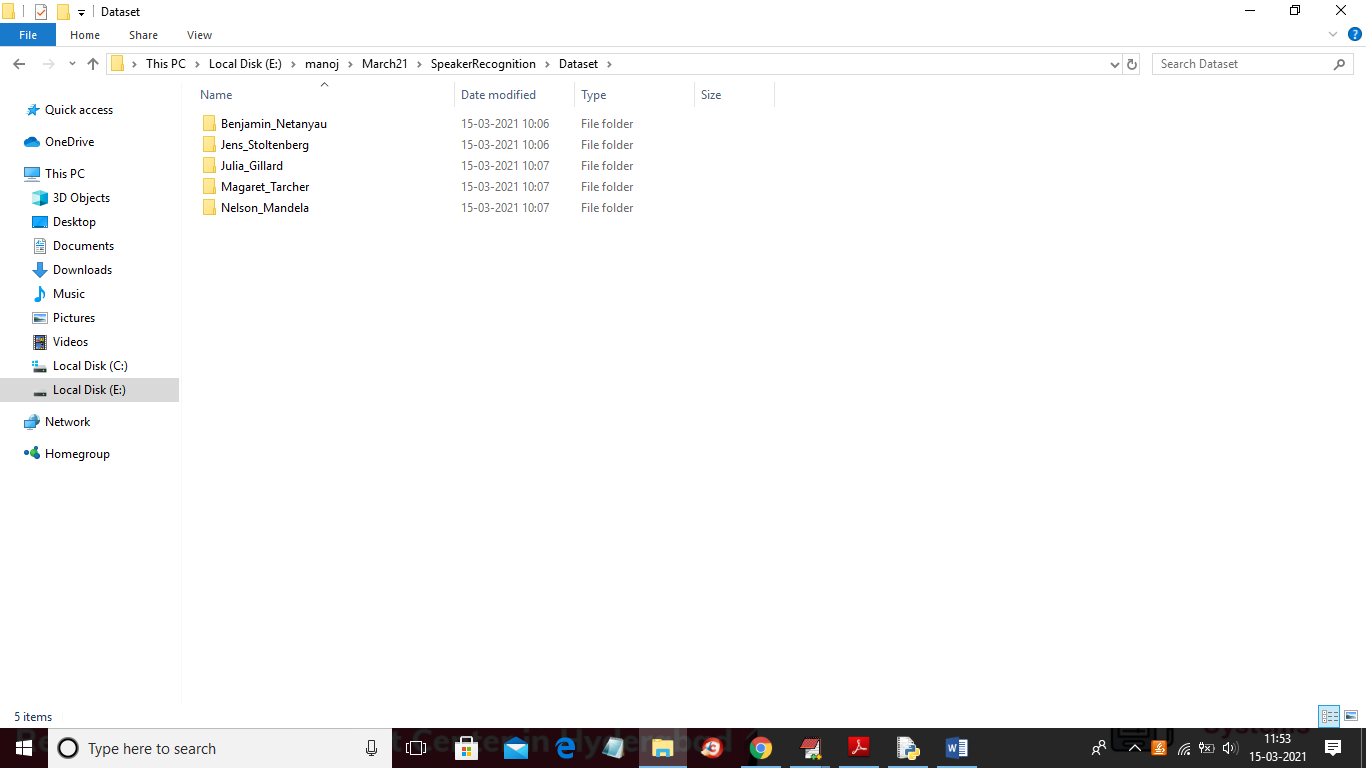
CNN BASED SPEAKER RECOGNITION IN LANGUAGE AND TEXT-INDEPENDENT SMALL-SCALE SYSTEM

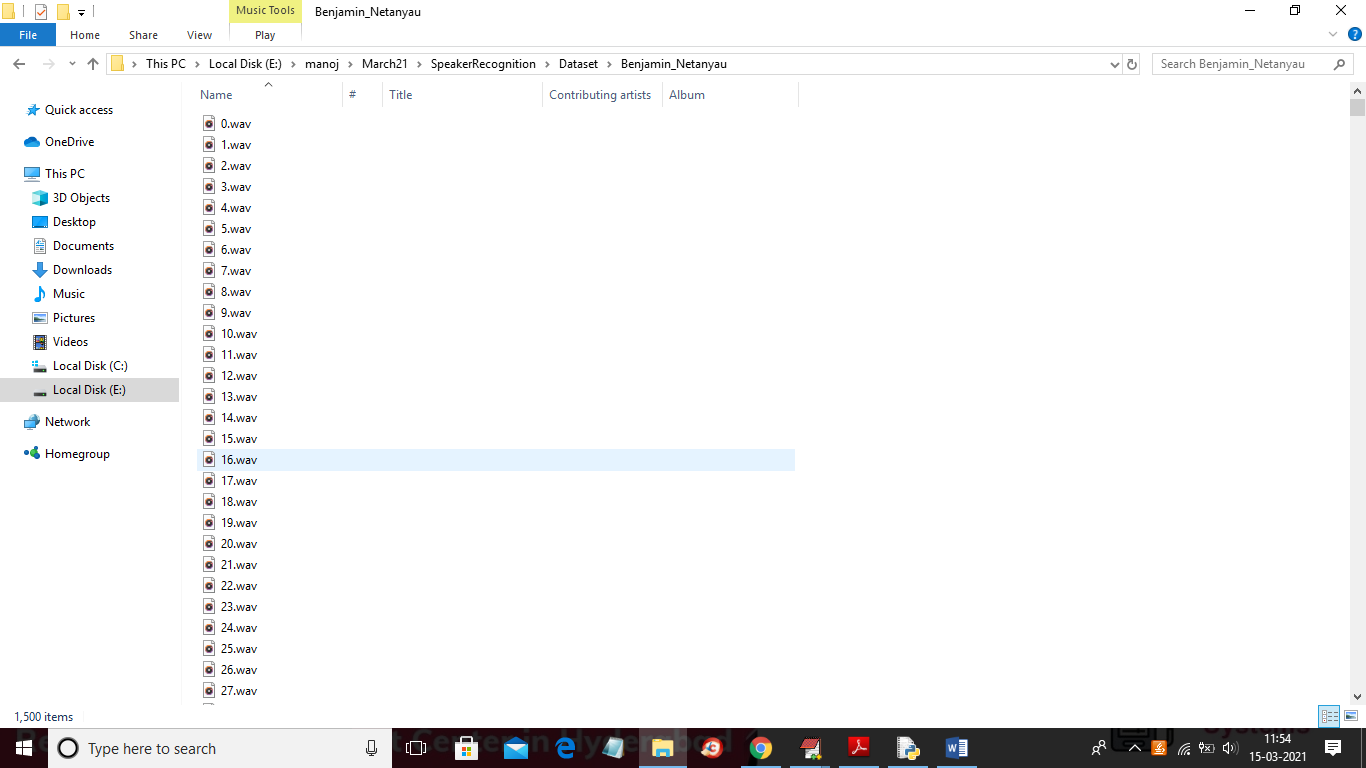
In this paper requirement student has said CNN based recognition in title and in paper he mention about to use Random Forest so I implement both algorithms and comparing accuracy of both algorithms. To implement this project we have designed following modules

1. Upload Dataset: Using this module we will upload speech recordings of five different persons
2. MFCC Processing: Using this module we are removing noise from audio and then applying MFCC algorithm to extract speech features and this features will be used to train both CNN and Random forest algorithms
3. runCNN: using this module we will apply above MFCC features on CNN algorithm to build CNN training model and then apply test data on that trained model to predict or recognize persons
4. runRandomForest: Above MFCC features will be used to train random forest algorithm and then apply test data on random forest to calculate prediction accuracy
5. Graph: using this module we will plot CNN accuracy and loss graph
6. Predict: Using this function we will upload test speech data and then CNN will recognize person from that speech.

To implement above project we have used speech audio of five different persons and this dataset saved inside ‘Dataset’ folder. Below screen shots showing all five persons folders



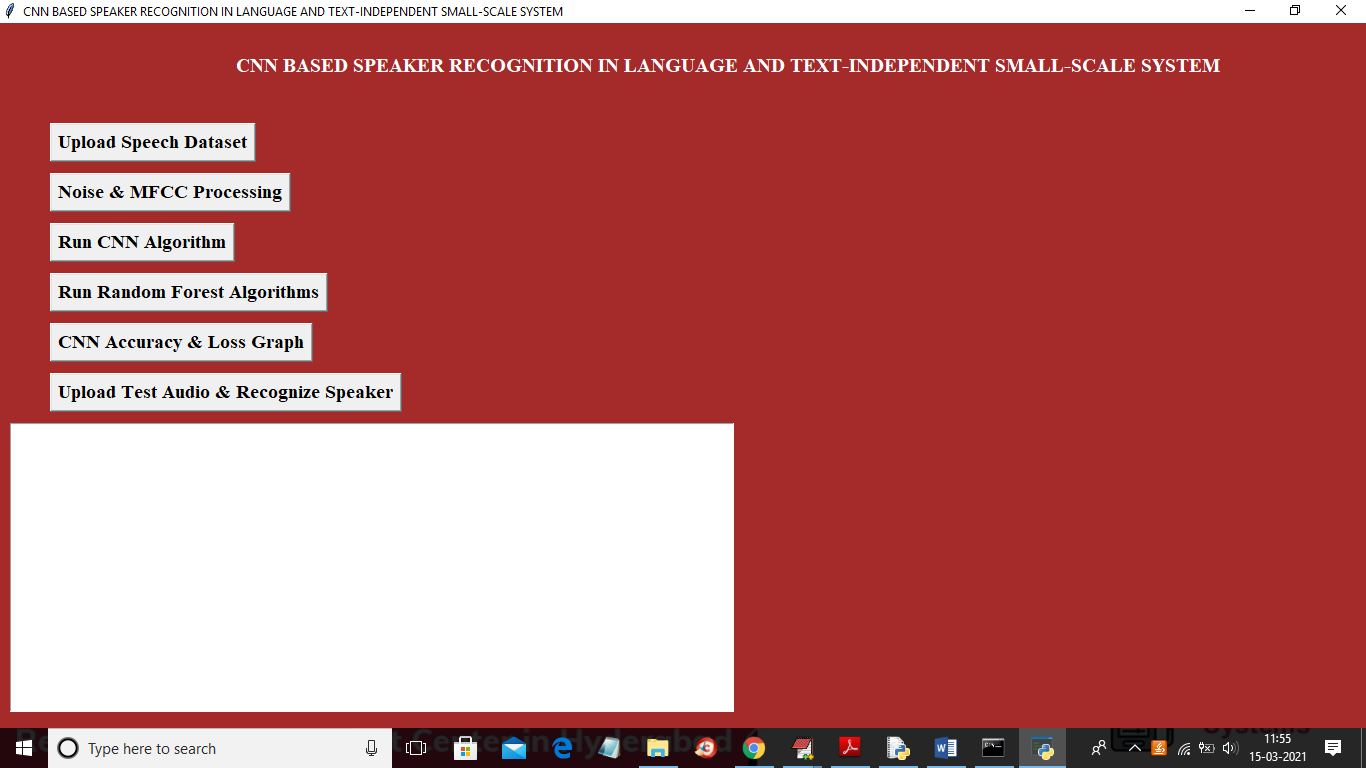
In above screen go inside any folder to see speech audio files of that person like below screen



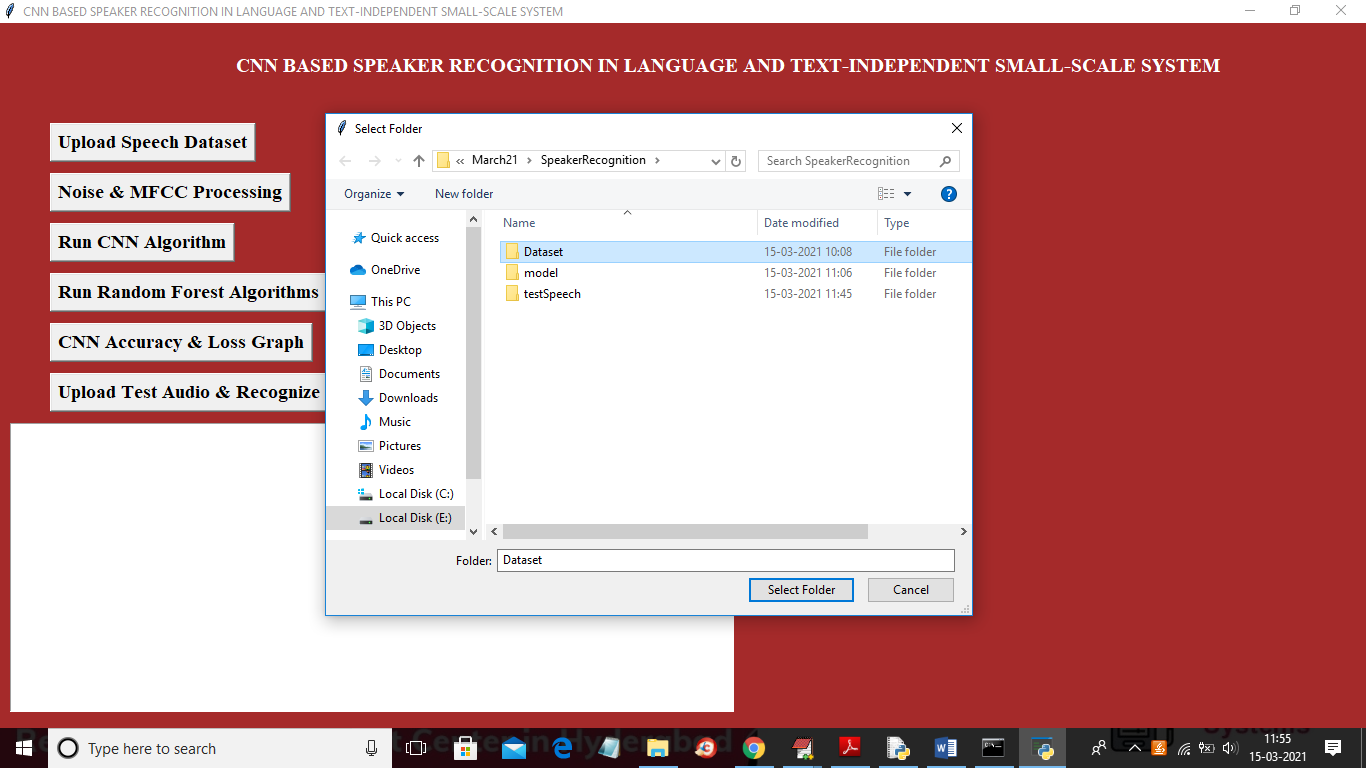
In above screen showing all audio files dataset

SCREEN SHOTS

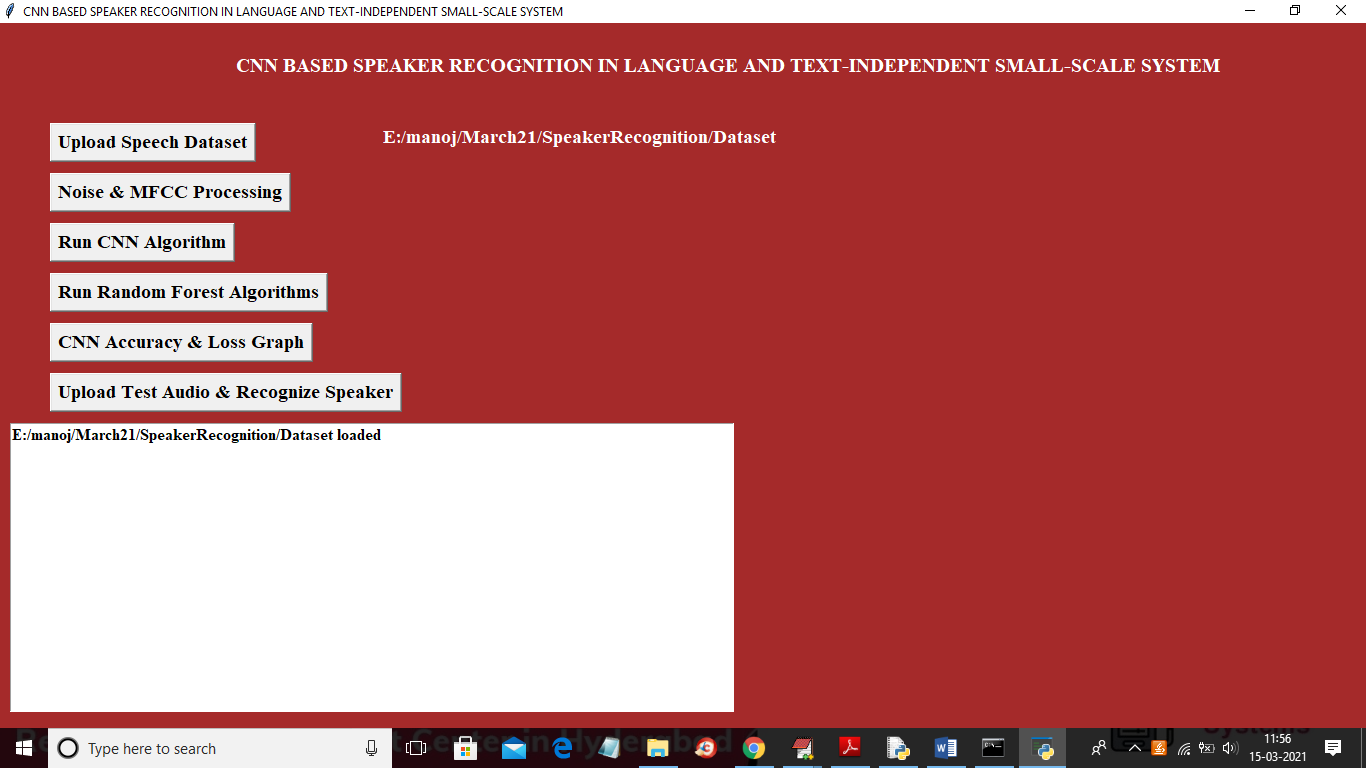
To run project double click on ‘run.bat’ file to get below screen



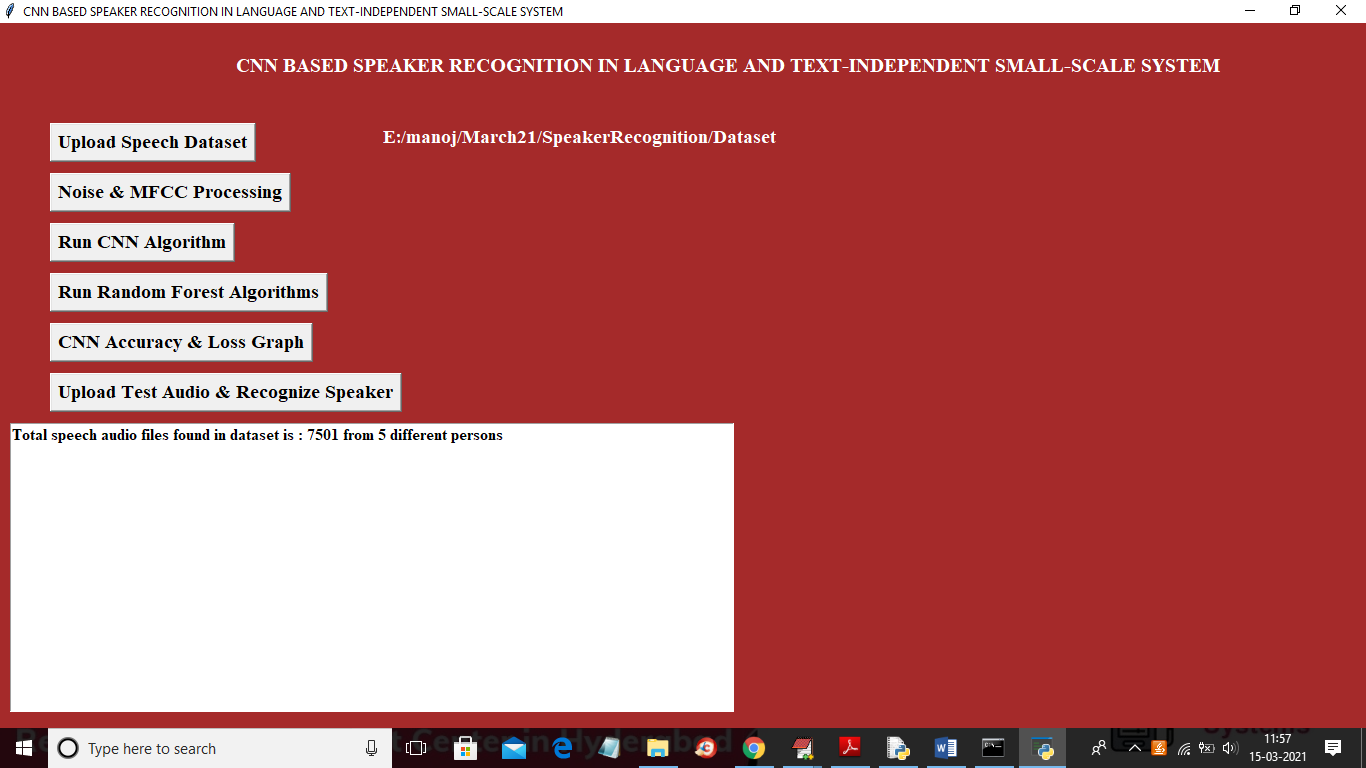
In above screen click on ‘Upload Speech Dataset’ button to upload dataset



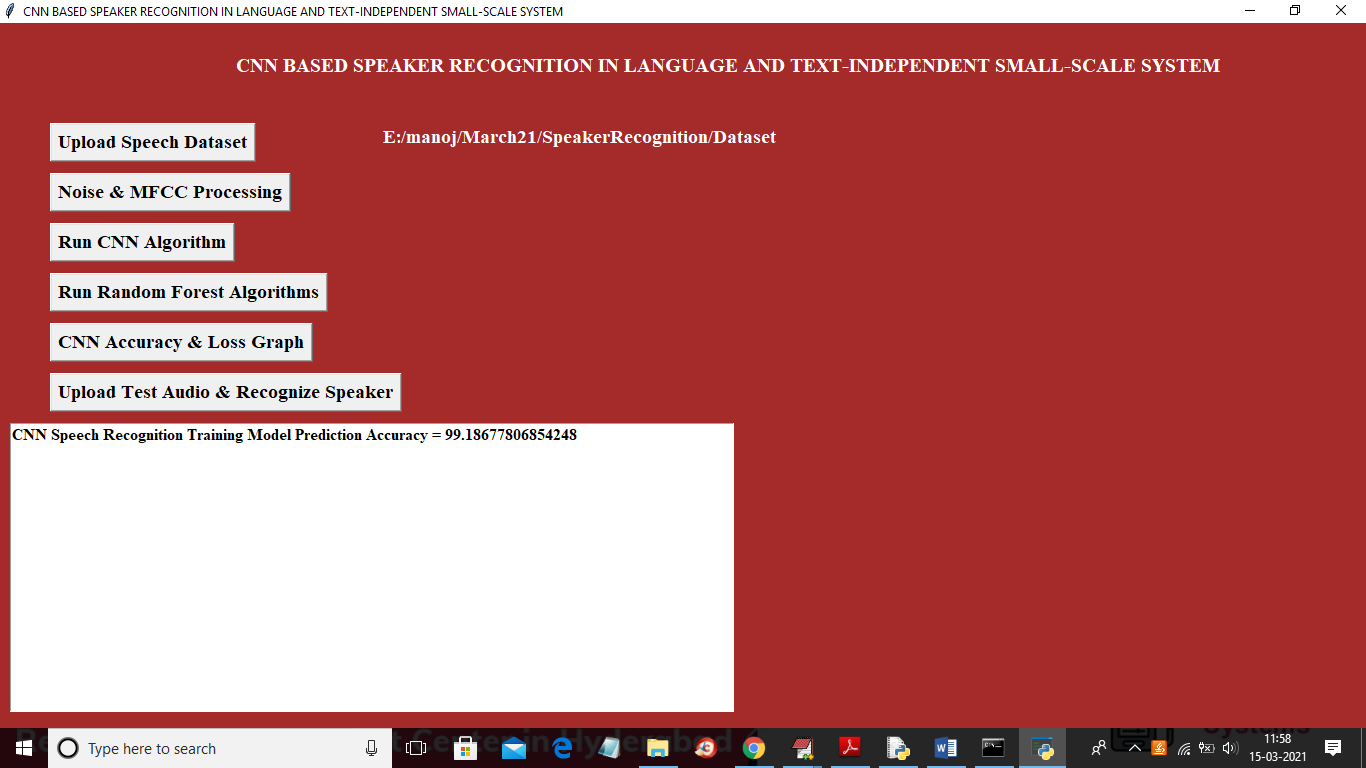
In above screen selecting and uploading ‘Dataset’ folder and then click on ‘Select Folder’ button to load dataset and to get below screen



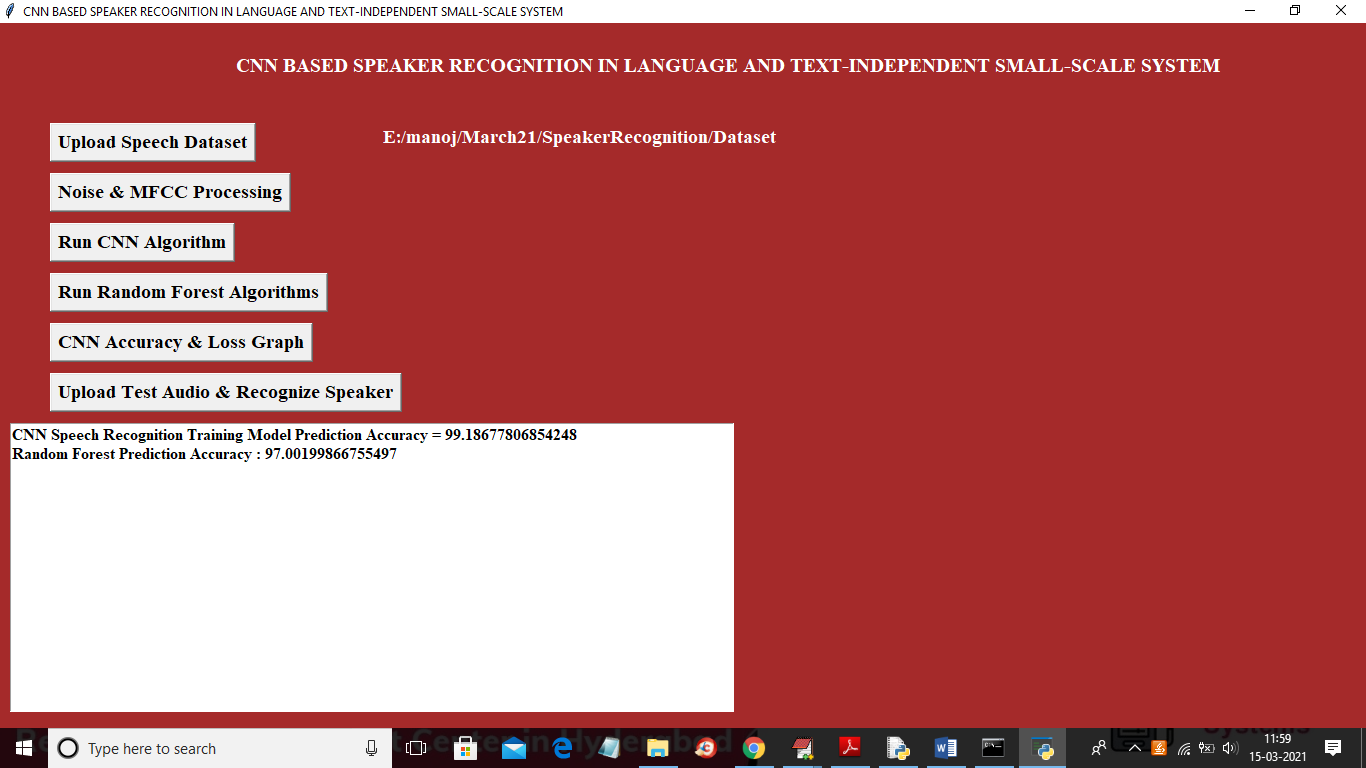
In above screen dataset loaded and now click on ‘Noise & MFCC Processing’ button to read all audio files and then remove noise and apply MFCC to extract features



In above screen application extracted features from 7501 audio files belonging to 5 different persons and now dataset features are ready and now click on ‘Run CNN Algorithm’ button to train CNN with above dataset



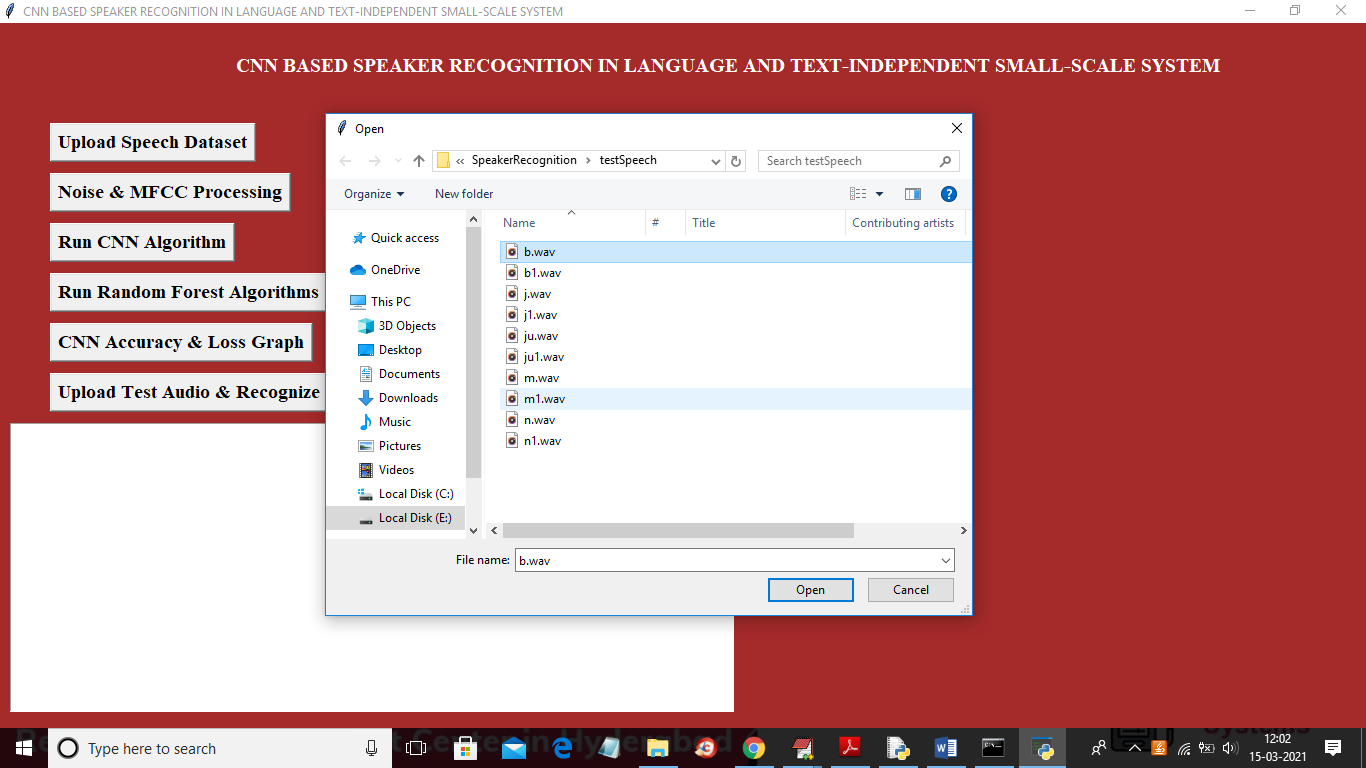
In above screen CNN training process completed and we got its accuracy as 99.18% and now click on ‘Run Random Forest Algorithm’ button to train random forest with above dataset and then calculate its accuracy



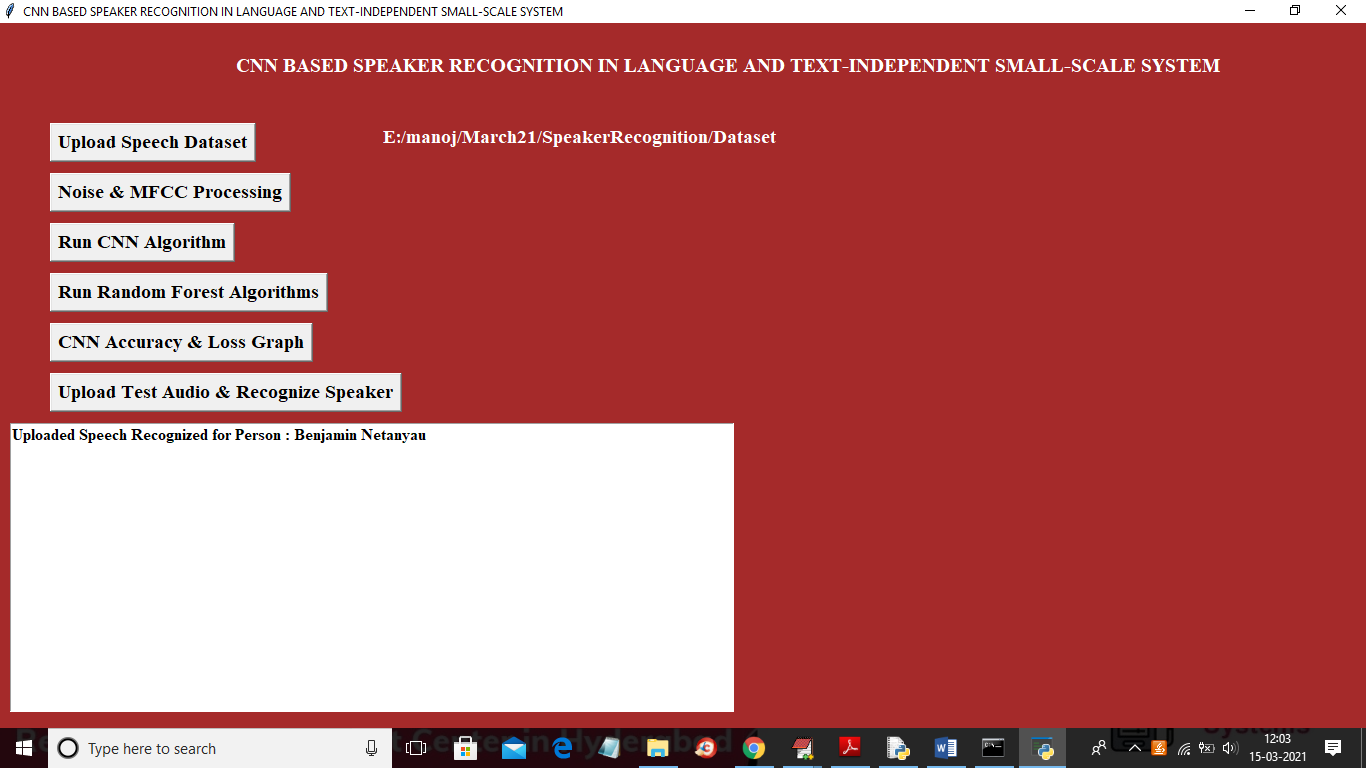
In above screen we got random forest accuracy as 97% and CNN got more accuracy than random forest and now click on ‘CNN Accuracy & Loss Graph’ button to get below graph



In above graph x-axis represents epoch/iterations and I took 10 epoch and in above graph while increasing epoch we can see accuracy get increase and loss get decrease to 0 which means model built with accurate prediction. In above graph yellow line represents accuracy and blue line represents loss. Now click on ‘Upload Test Audio & Recognize Speaker’ button to upload test speech and then CNN will recognize person from that speech



In above screen I am selecting and uploading ‘b.wav’ file and then click on ‘Open’ button to predict person from that file like below screen



In above screen in text area we can see person recognized as ‘Benjamin Netanyan’ and similarly you can upload other files and recognize and after recognize then that file will play voice