Scientific Journal of Impact Factor (SJIF): 5.71

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 5, Issue 05, May -2018

SPEECH CONTROLLED ROBOTIC VEHICLE USING IOT AND MFCC ALGORITHM

Bharati Thawali[1], Apurwa Dhase[2], Neha Biradar[3], Onkar Sonawane[4]

[1] *Prof, Department of Electronics & Telecommunication, JSPM's RSCOE Engineering College, Pune*
[2] *U.G. Student, , Department of Electronics & Telecommunication, JSPM's RSCOE Engineering College, Pune*
[3] *U.G. Student, Department of Electronics & Telecommunication, JSPM's RSCOE Engineering College, Pune*
[4] *U.G. Student, , Department of Electronics & Telecommunication, JSPM's RSCOE Engineering College ,Pune*

Abstract - *Speech recognition is recognized as one of the most advanced and reliable biometric technologies, commonly used in automatic personal identification systems. There are various applications such as multimedia browsing tool, access center, security and finance. This paper includes the study of recognition of speech from human voice using Mel-frequency cepstrum coefficients (MFCC) features. Commands are used for recognition of speech with the help of 16 mel-scale warped cepstral coefficients independently. The growth in IOT technology and mobile devices has sustained the development of speech recognition system.*

KEYWORDS *-- MEL-FREQUENCY CEPSTRAL COEFFICIENTS (MFCC); INTERNET OF THINGS TECHNOLOGY(IOT); SPEECH RECOGNITION*

I. INTRODUCTION

One of the most important manners of communication to exchange feelings and information is Speech. Without knowing the meaning of speech, the computer based speech recognition systems simply converts the speech into textual representation. This type of system has various applications such as security devices, household appliances, cellular phones, ATM machines and computers. Speech processing mainly includes Speech recognition and Speaker recognition. The main aim of automatic Speaker recognition is to extract, characterize and recognize the information about speaker identity to identify a speaker by his or her voice.

Speech is very useful and efficient way to train a machine. Also it can be used for communicate with the machines. This feature can also be used for biometric authentication for verifying the identity of a person in applications like banking by telephone and voice mail. Speech recognition system is evolving as big need in the robotics field. To make the robot capable to follow commands through voice, speech recognition is most important.

If we consider human-robot interaction, there are various options for human to communicate with robot, which are from visual to touch to voice. There are various ways through which interaction is possible, whether the interaction is human to human or human to robot.

In this process, initially we are going to capture the voice signal using microphone. Then by applying algorithm, we are going to reduce noise interference and silence suppression. Then the output signal is then processed to extract the features. To extract the features MFCC algorithm is used. Then features are used as input to Artificial Neural Network system. ANN is the most efficient and accurate for training and testing of speech samples.

In this process, initially we are going to capture the voice signal using microphone. Then by applying algorithm, we are going to reduce noise interference and silence suppression. Then the output signal is then processed to extract the features. To extract the features MFCC algorithm is used. Then features are used as input to Artificial Neural Network system. ANN is the most efficient and accurate for training and testing of speech samples.

II. RELATED WORK

While Internet of Things gradually become a hot topic of research and business and has been everywhere used in many industries. IOT with speech processing definitely will result in great human help related to many industrial as well as surveillance work. In this paper, we firstly we captured the speech signal and then extracted its features using MFCC algorithm. Then, using IOT technology and Wi-Fi, communication between human and robot takes place.[2]

III. LITERATURE SURVEY

Pialy Barua et al [1] worked on neural network based recognition of speech using MFCC features. This paper shows the results from primary study to recognize the speech from human voice using mel-frequency cepstrum coefficients (MFCC) features. The 16 mel-scale cepstral coefficients were warped and used independently for reorganization of speech from two Bangla commands of our native language. Result of matching features in neural network shows that MFCC features work significantly to recognize the speech.

Chin Kim On et al [2] worked on mel-frequency cepstral coefficient analysis in speech recognition. Speech recognition is considered as one of the most popular and reliable technologies used in automatic personal identification systems. In this paper, the zero crossing extraction and energy level detection are applied to the recorded speech signal for voiced/unvoiced area detection. Then detected voice signals are inputted for segmentation. Then further this MFCC signals are processed for ANN.

Amab Pramanik et al [3] proposed the automatic speech recognition system using correlation analysis. The growth in wireless communication has supported with mobile devices with speech recognition systems. In this paper, a simple algorithm has been developed for matching patterns to recognize the speech. The mel-frequency cepstral coefficient is used as feature of the recorded speech. This is implemented using principle of correlation. All the simulation experiments carried out with MATLAB software which produced relatively good results.

Bacha Rehman et al [4] implemented hardware for speech recognition using MFCC and Euclidean distance. In this paper, Digital signal processor based speech recognition system has been developed with improved performance in terms of recognition accuracies and computational cost. This paper gives an approach of isolated speech recognition by Digital Signal Processor using MFCC and Euclidean distance.

IV. BLOCK DIAGRAM

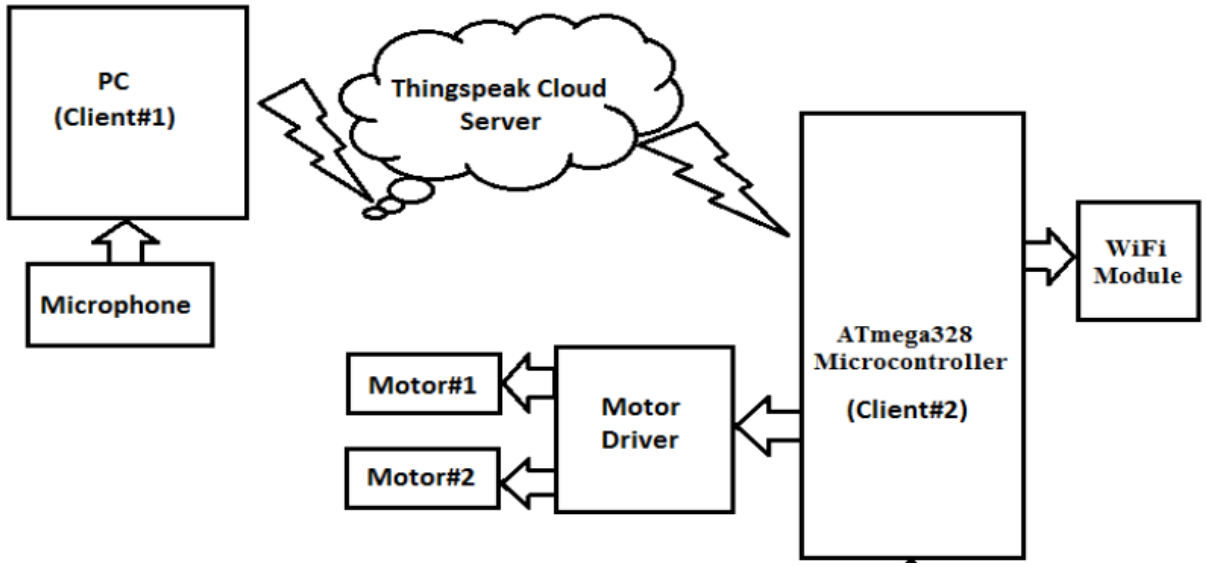


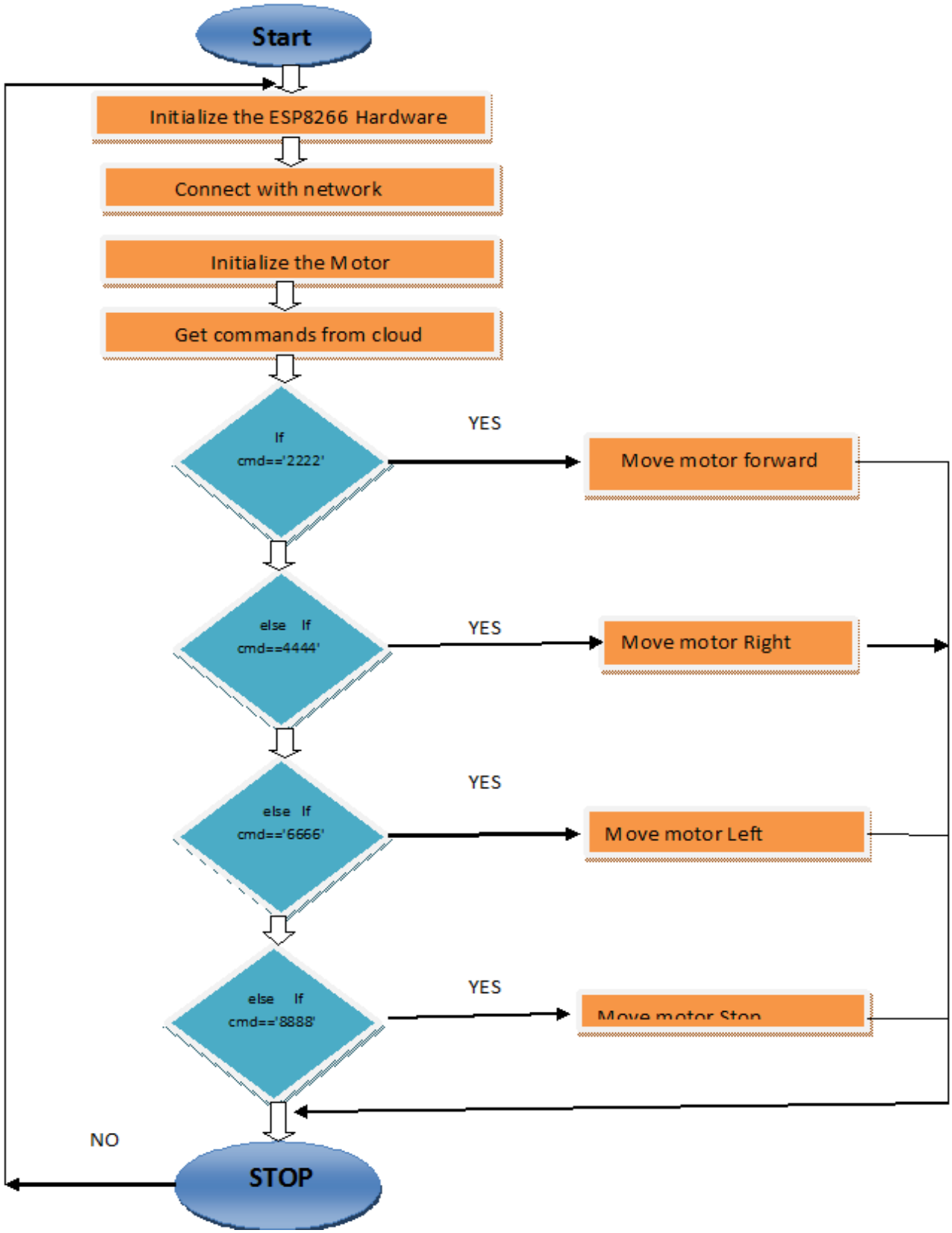
Fig1- Block Diagram

V. THE OPERATION OF THE SYSTEM

In our proposed system we are using MFCC algorithm of artificial neural network. This algorithm is used as a feature extraction. Features extracted from MFCC is given to ANN. So that we can recognize commands to control the vehicle. The proposed system also leverage of Internet of Things that is used to control the vehicle remotely from any part of the world. ThingSpeak cloud server is the communication server between PC and vehicle. Speech recognition is done by PC. It sends commands to ThingSpeak cloud server. WIFI is used by the vehicle to connect with the internet. Mainly TCP/IP communication protocol is used for communication purpose. Training of commands that to be used is included in speech recognition. After complete training, we can send commands to control the vehicle. Robotic vehicle operates as per the commands received. For this purpose ATmega328 microcontroller is used to operate the vehicle. Microphone is used to transmit the data. The user can send commands like 'FORWARD','RIGHT','LEFT','STOP'.
After receiving the commands, microcontroller then works on it to control the motor in given directions. Microcontroller is connected to the motor driver circuit to drive the robotic vehicle. Now we can control the vehicle remotely from any part of the world that is the main advantage of this system.

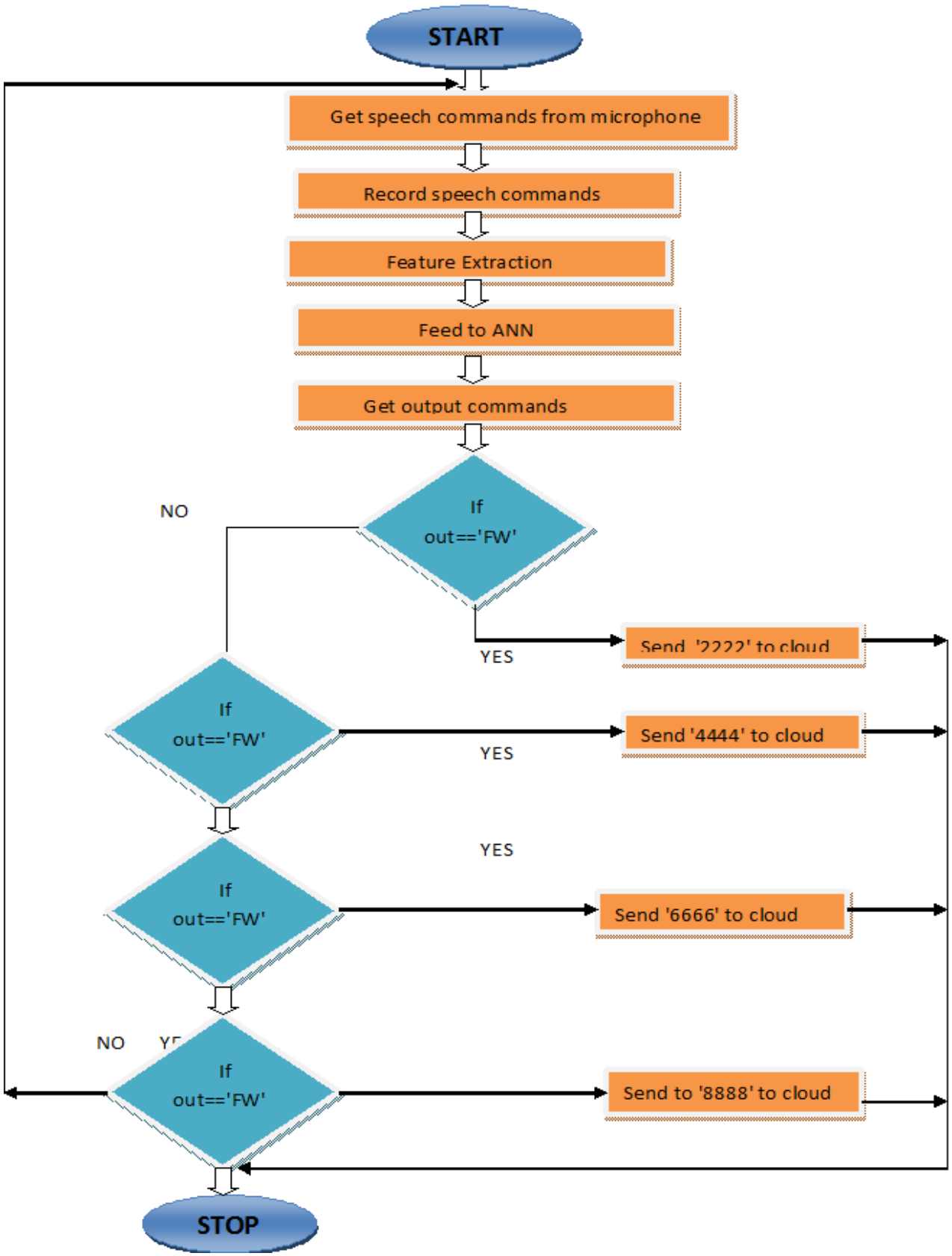
Flow Chart-

ROBOT SIDE



FLOW CHART-

Speech Processing Side



VI. CONCLUSION

Training and testing is carried out in planned manner, it has more benefits and is therefore worth considering. This is achieved by implementing MFCC and ANN. Testing and training is fast and accurate .In this project Arduino is used with atmega328 microcontroller is used to construct robot vehicle in order to recognize the spoken word. Internet of thing technology is used to access in wireless mode i.e. we can access robotic vehicle from any part of the world remotely.

VII. RESULT

RESULT OBTAINED ON SYSTEM

Output obtained on hardware

On hardware part the input is voice from microphone is fed to PC in terms of voice signal. Role of PC is to provide output according to program the commands will be sent to microcontroller.

For robot application, here DC motor is used. So motor driver is required due following two reasons.

- Microcontroller output gives max 30 to 40 mA of current but motor needs 1Amp of current.
- Microcontroller output gives max 5 V of voltage but motor needs 12 V of voltage.

Output obtained on software

The input is fed to microphone it will be asked to record.

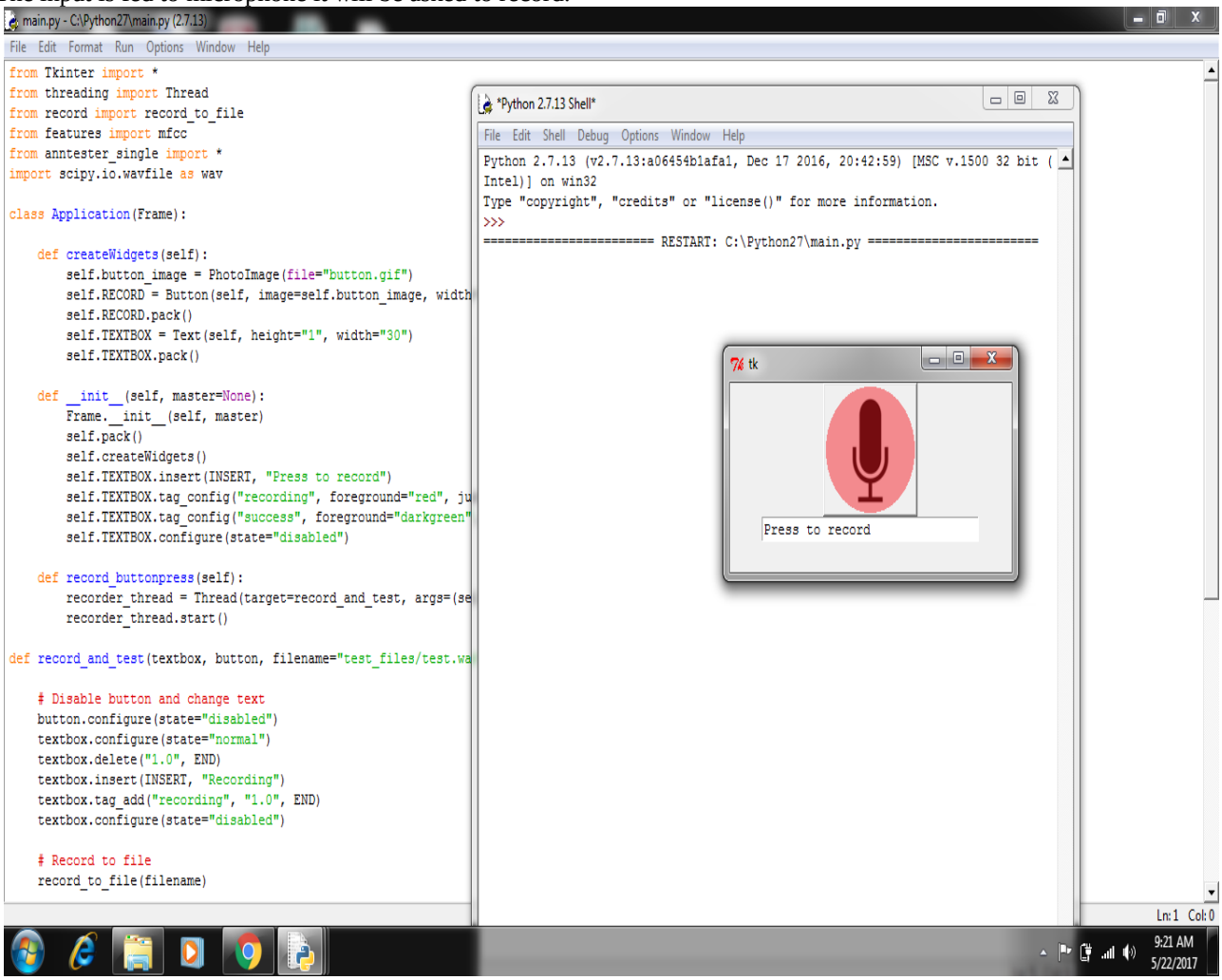


Fig2- Asked to provide input

The input is given as voice signal is obtained and into keep separately in train folder.

Fig 3- Result obtained after training.

Once input is fed again with a voice command will be there in test folder.

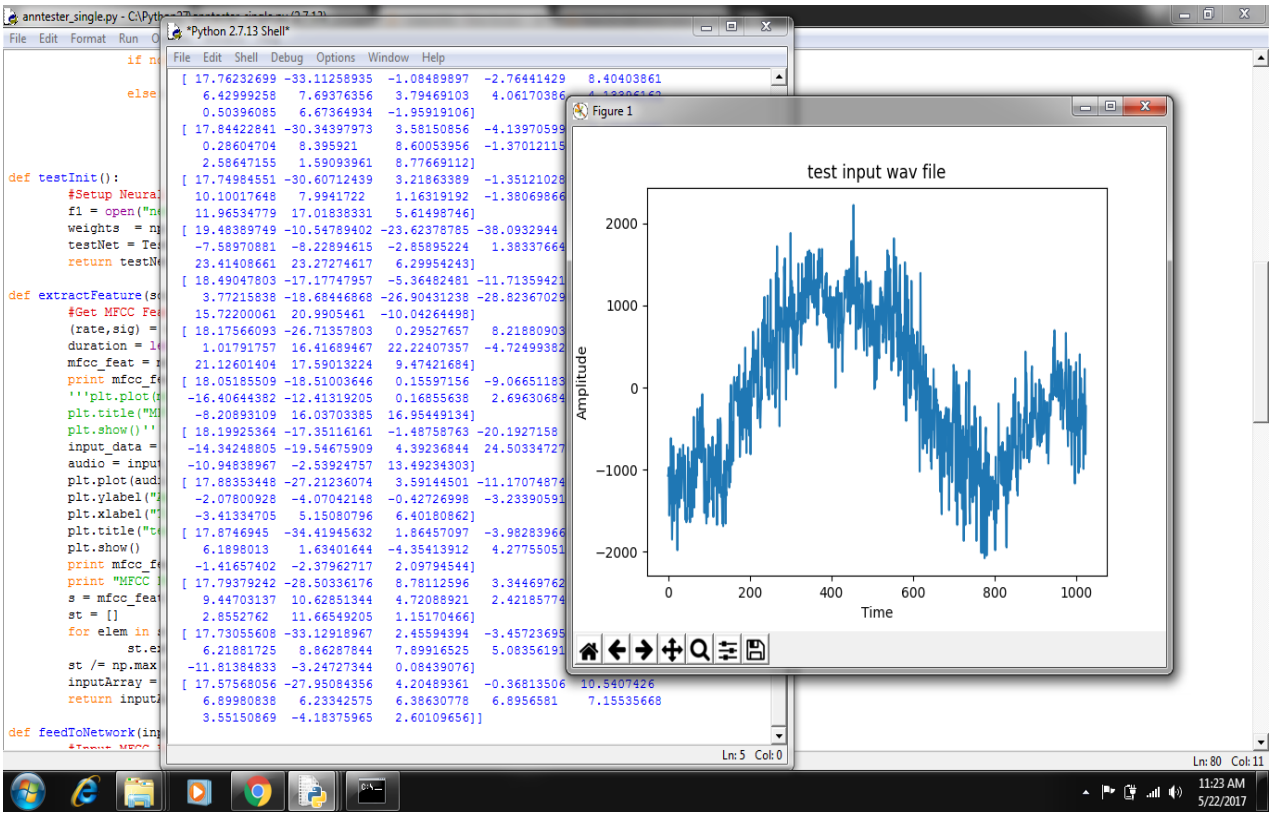


Fig 4- A Sample Test File

The testing file will be worked out for feature extraction technique. Here MFCC is used.

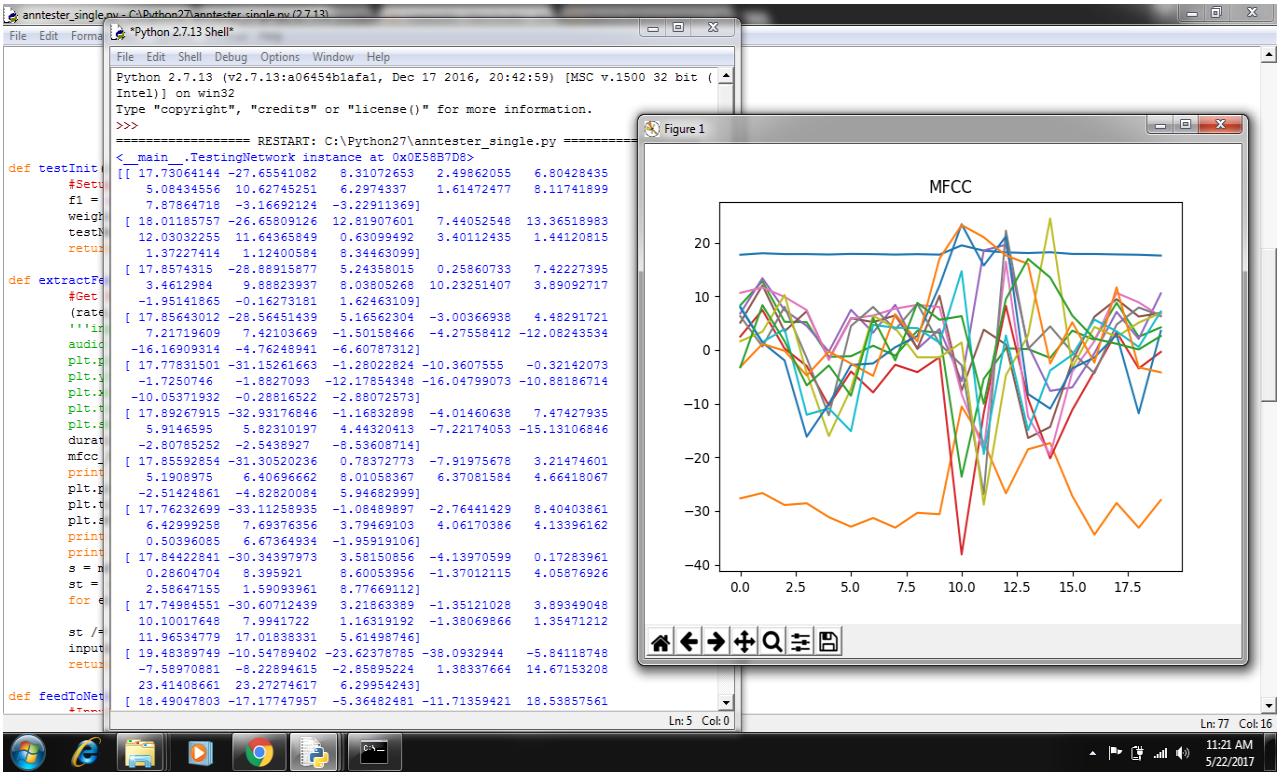


Fig 5- Feature extraction using MFCC

Detection of features is compared with the train and test database from that recognition of voice command will be detected.

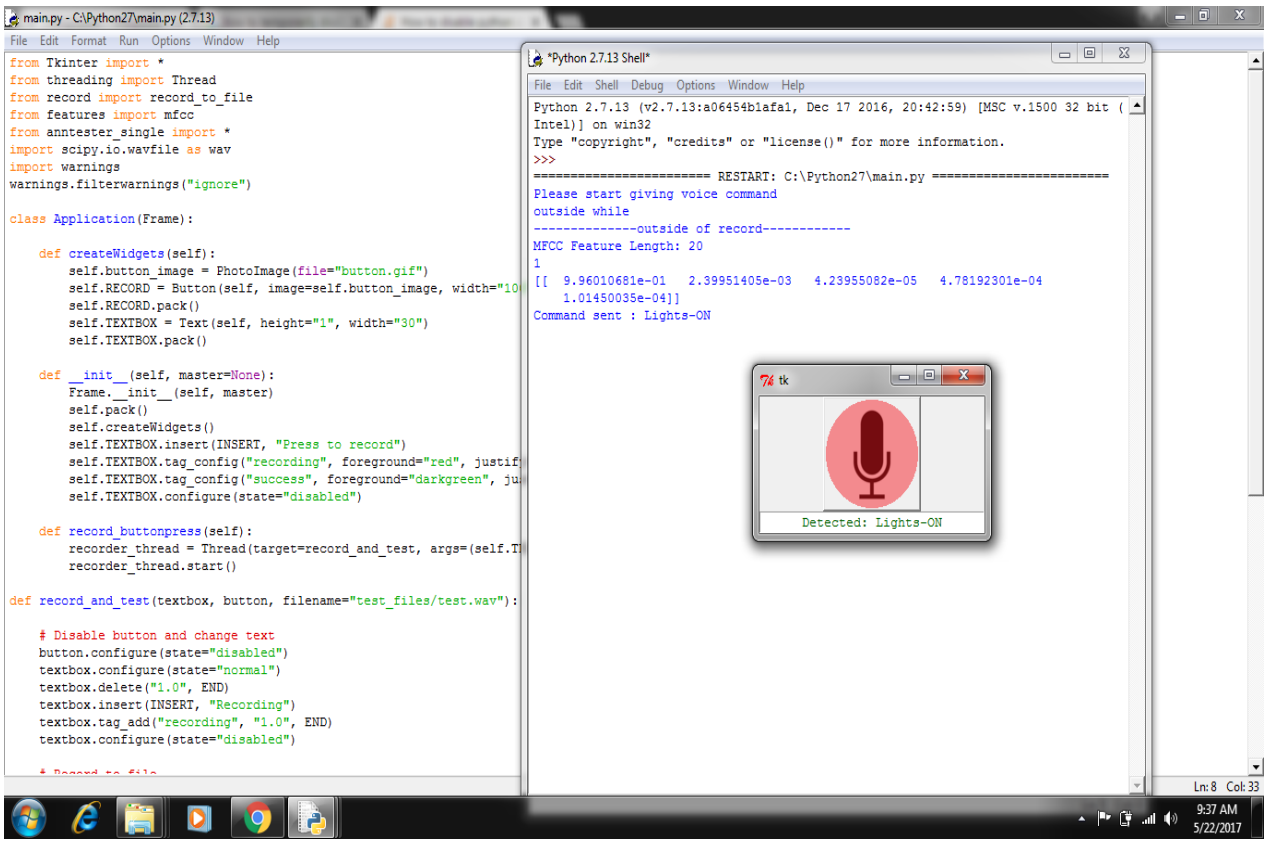


Fig 6- Detection of command.

Once command is detected is sent to cloud. Thingspeak cloud is used to transmit the information.

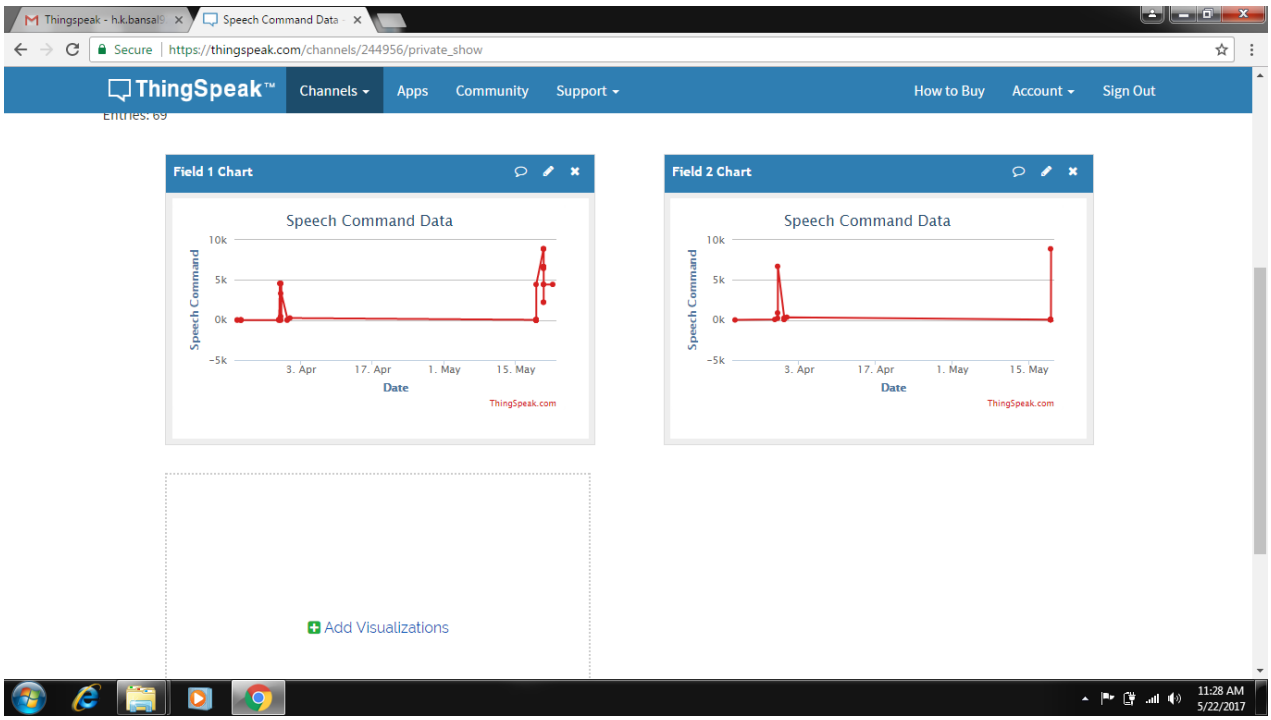


Fig 7- ThingSpeak Cloud Platform

Result Table-

Speech	**Actual Result**	**Test Result/Command Sent to Cloud**
FORWARD	The command is "2222"	"2222"
RIGHT	The command is "4444"	"4444"
LEFT	The command is "6666"	"6666"
STOP	The command is "8888"	"8888"

REFERENCES

1) Pialy Barua,Kanij Ahmad,Ainul Anam Shahjamal Khan,Muhammad Sanaullah, **"*Neural Network Based Recognition of Speech Using MFCC Features*"** 3rd International Conference on Informatics, Electronics & Vision ,IEEE ,2014.

2) Umarani J.Suryawanshi, Prof. Dr. S. R. Ganorkar, "***Hardware Implementation of Speech Recognition Using MFCC and Euclidean Distance*"** IJAREEIE, Vol. 3, Issue 8, August 2014, pp 11248-11254**.**

3) Arnab Pramanik, Rajorshee Raha**, "*Automatic Speech Recognition using Correlation Analysis*"** World Congress on Information and Communication Technologies, IEEE 2012, pp.670-674.

4) Chin Kim On, Paulraj M. Pandiyan, Sazali Yaacob, Azali Saudi, "***Mel-Frequency Cepstral Coefficient Analysis in Speech Recognition*"** ICOCI, IEEE 2006.

