Speech Interface for

Specially Abled People (Mute People)

Final Year Project Report

By

Ahmad Ali

Muhammad Huzaifa

Muhammad Sarosh

In partial fulfillment of the requirements for the degree

Bachelor of Electrical Engineering (BEE) &

Computer Science (CS)

School of Electrical Engineering and
Computer Science National University of
Sciences and Technology Islamabad, Pakistan

(2021)

DECLARATION

We hereby declare that this project report entitled **“Speech Interface for disabled people”** submitted to the “Department of Electrical Engineering and computer Science”, is a record of an original work done by us under the guidance of Advisor “**Sir Latif Anjum”** and Co-Advisor **“Sir Majid Maqbool”** and that no part has been plagiarized without citations. Also, this project work is submitted in the partial fulfillment of the requirements for the degree of Bachelor of Electrical Engineering and Computer Science.

Team

Member:

Muhammad Huzaifa

Ahmad Ali

Muhammad Sarosh

Advisor: Dr. Latif Anjum **Signature:**

Co-Advisor: Sir Majid Maqbool **Signature:**

Date: 3rd June. 2021

DEDICATION

To Allah the Almighty,

To our Parents and Faculty

&

To everyone who we met along the way

ACKNOWLEDGMENTS

We are immensely grateful to our respected advisor **Sir Latif Anjum** and co-advisor **Sir Majid Maqbool** for guiding us throughout in our final year project. Your guidance, skills, advice, motivation, and support enabled us in achieving our goals.

We greatly appreciate everyone else who was involved in this project which includes but not limited to parents, faculty, friends who helped us in making us succeed.

Table of Contents

Table of Contents

List of Figures:

List of abbreviation:

Terminologies Used	Non-abbreviated Form
ML	Machine Learning
PSL	Pakistan Sign language
Flex	Use to detect bending of fingers
Gyro	Acceleration of hand
STS	Sign to Speech

Abstract

About 5% of the world's population is facing hearing loss. Deaf people feel isolated due to the huge communication gap. Due to lack of communication between the disabled and the normal people the transfer of original content become impossible.

To reduce the communication gap between deaf people and others this project represents the study and design of a device for native sign speakers. Instead of focusing on the recognition of isolated signs, this project presents a system that incorporates complete sign language into speech. There are many steps of interpretation which is done for communication. When a disabled person address to the normal person at that time signed values processed and convert into the speech.

Introduction

Disabled people cannot talk with the normal person properly as there is some gap of communication between them. There is always need a for sign language through which they talk to each other. It is very important for deaf people in the present as well as the people who were in the past. According to the WHO, there are almost 1.5 billion people out of which around 400 million people use the devices to make their hearing system better. It is assumed that by the end of 2050 there will be around 2.5 billion people who have this disability. To reduce this loss, almost 700 million people will require the device to make their hearing system better.

Of course, choosing a specialist or psychologist with the resources available to function successfully for those people who have hearing problems is not a straightforward task. While deaf or disabled people have managed to learn and adapt too many situations but there will always be difficulties. Although many deaf people do not want to hear and accept deafness to their own unique culture, they are still viewed with pity by the wider world of hearing. The truth is, most people who are deaf or hearing disabled do not seek sympathy, and they only want to be handled with dignity.

Some people required such devices which can reduce their communication barrier with other people. In past, they use only signs which can be understandable for other deaf people only. A normal person cannot understand these signs. Sign to speech interface is required to reduce the barrier between people.

We made a smart glove on which flex and gyro sensors are integrated to obtain the data of the movement of fingers and the hand, respectively. All the data which comes from the sensors is evaluated and processed using the machine learning and microcontroller. Finally, the recognized gesture is transformed into a text and then the output of speech is delivered using the Bluetooth speaker. In this way, this process offers the ability for the disabled users to interact with the normal people.

Language	Legal recognition and where spoken natively by significant population	Ethnologic estimate
American Sign Language	Native to United States, Canada and Guatemala. Used in varying degrees in Philippines, Singapore, Hong Kong, Côte d'Ivoire, Burkina Faso, Ghana, Togo, Benin, Nigeria, Chad, Gabon, Democratic Republic of the Congo, Central African Republic, Mauritania, Kenya, Madagascar and Zimbabwe.	≈ 500,000 in the USA
Indo-Pakistani Sign Language	No legal recognition. Native to India, Pakistan and Bangladesh.	2,700,000 in India (2003)
Saudi Sign Language	Native to Saudi Arabia.	720,000 (2010)
Chinese Sign Language	Native to China. Also spoken in Malaysia and Taiwan .	unknown, maybe 1,000,000–20,000,000 (no date)
Brazilian Sign Language	Legally recognized by law 10.436, April 24, 2002[2] - Native to Brazil.	3,000,000 (no date)

According to an estimate of WHO, almost 5% of the world's population suffers from disabling hearing loss. Among these unfortunate people, there are some even more unfortunate ones who lose this blessing at a very early age and must resort to sign language as a means of communication. The lack of sign language interpreters creates a huge barrier in communication and results in extreme isolation of millions of deaf people all over the world.

Therefore, it is very important for the people as more people understand each other more will benefit occur from it. Deaf people want to share their thoughts and feeling with other people. They also need some to communicate with the people to meet their daily requirements. Hence, we must make such a device which is helpful for them.

Problem Statement

The main challenge for the people is that how they can transfer their thoughts and feeling to some other person if he does not understand their language. There is a need for such a device that can reduce the communication barrier between dumb people and normal people. This problem can be solved by using different types of sensors and microcontrollers to control the input and manage the output according to them. A glove on which all the flex and gyro sensors are attached and can be used to monitor the movement of the hand and then give the output regarding the input.

Motivation

Verbal communication is the most genetic and expressive way to learn for people with auditory difficulties. People, who are not deaf, do not ever try to learn sign language to communicate with deaf people. This tends to lead to the isolation of the deaf. Studies have shown that autistic people are twice as probable from mental illnesses such as depression and panic attacks. Research suggests that this is due to feelings of isolation. The most appropriate cure for these types of issues is probably talking to a mental health professional. People with hearing disabilities face numerous challenges. They undergo and connect to the world a lot differently than those with normal hearing. Public address systems remind us what is happening all the time, but a hearing disabled person still would not get a signal. When someone acknowledges that they meet a hearing-impaired person, they often turn to a slower form of communication.

Purpose

The purpose of these flex sensors is to measure the bending of the fingers and the gyro sensor is used to measure the rotation as well as the acceleration of the hand. Both sensors are embedded on the glove with the help of the PCB board. These sensors relate to the microcontroller which is raspberry pi 3b+ in this case. We use machine learning to improve the efficiency of the system. In this machine learning, all the data which is obtained from the sensors are converted into the datasheet which is used to predict the output of the different signs. After training the specific sign a minimum of 100 times the machine learning will give the best-predicted output after the sign will be performed.

By using the techniques and the sensors we can obtain the best output of the device. Machine learning is used for training the model of the glove. This device is used for improving the communication between the dumb and the normal people. The gap between these people decreases and people more understand each other. Through this device, the problems of the deaf will be solved, and they will not feel more degraded by society. For performing the signs, we use the official Pakistan Sign Language (PSL) from their official website.

Objectives

Our aim is to develop a wearable device that convert sign language into speech by enabling them to hear the people around them in their own way, resulting in a more inclusive and accessible society. The interpretation of the project is given below:

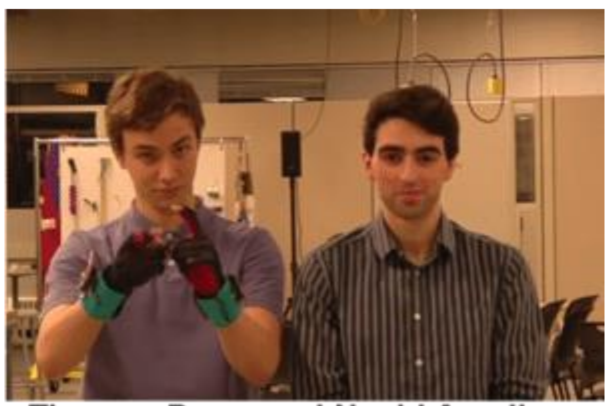
Sign to Speech Conversion:

By using a smart glove on which different sensors attached we will implement the machine learning algorithms to predict the specific output and this output will be in the form of voice by attaching blue tooth speaker.

Literature Review

A lot of research has been done in the field of sign to a speech by using different types of techniques, but the implementation of these projects is not done properly in the past.

Navid Azodi and Thomas Pryor [1]: Two students at the University of Washington came up with a research of developing gloves for the deaf people. They believe that communication is a gift from nature and is a fundamental right of every individual. A combination of gloves can interpret hand movements that suit terms and phrases in the American Sign Language. Every glove incorporates sensors that monitor hand location and movement and send data wirelessly through Bluetooth to a laptop computer.



Thomas Pryor and Navid Azodi
University of Washington

Figure 1

Sign School [2]: It is a popular American Sign Language online learning site. Sign School's mission is to promote meaningful contact between the Deaf and the listening community by offering a high-quality digital forum for learners to discover American Sign Language. They have a multicultural workforce of content creators from different parts of the US. At Sign School, they demonstrate the signs that are more widely found in the Deaf culture. People use these different types of signs to express their thoughts with the normal people. By using different techniques these signs can be translated into the speech. People use different kind of sensors such as flex and gyro to measure the acceleration, rotation, and temperature of the environment around the sensor. There are many other techniques which used for implementing the sign-to-speech interface.

Figure 2

Real time Hand gesture recognition [3]: The real-time hand gesture recognition technique is very important which includes the processing of the image in real-time as the movement of the hand is captured and then analyze in the real time and then show the desired output. This technique is widely used as the processing is done in real time and the desired output can be seen on the output screen. In this technique, there is a need of camera to capture the image of the gesture and then process this data to find out the most likely output.

Figure 3

Gesture recognition system

Static hand gesture technique [4]: This hand gesture recognition techniques include the usage of the flex and gyro sensors. We also used this technique to implement the sign-to-speech interface for disabled people. All the sensors and devices are attached on the glove and then the data is collected by using different types of sensors and stored in the microcontroller. Which processed the data come from the sensors and then give the desired output. We read almost more than 10 research papers and many other articles about this sign-to-speech interface and now come with the advanced solution by adding few more and new features in it.

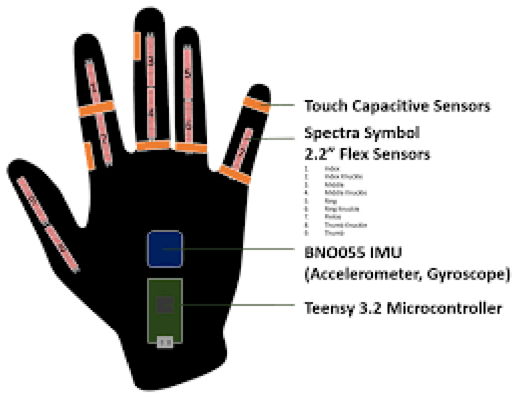


Figure 4

Intended Audience

This document of Sign to Speech interface is intended for the professors of SEECS National University of Science and Technology (NUST) Dr. Latif Anjum and the Majid Maqbool and other professors of this group knowledge. This document is also made for the people who are interested in our project and want to work with us. All the information related to the project is explained such as the working plan, process, and future goals. Many companies related to the technology can contact us and we will be very excited to end our documentation to them.

Project Scope

Automated multi-sensors board:

Gesture recognition system

- This will help the people to collect real-time data of the movement of the hand by using different types of sensors.
- Collect data by using flex and gyro sensor.
- Collected data is processed in the microcontroller such as raspberry pi 3b+
- Use the PCB board to join the sensors with the microcontroller.

Machine Learning-based efficient data

- By using machine learning the efficiency of output increase
- Use the python language to implement the ML.
- The communication gap between the normal and dumb people decreases.
- Monitor the movement of the hand and then give the instructions accordingly.

Detailed System Design and Architecture

Design Methodology:

Moving forward we will start our discussion with the overall design of the prototype of the wearable aid for native signers. We will closely look at all the utilized resources to build the required system which will convert speech into Pakistan sign language (PSL) almost in real time. Our system will require multiple components to achieve the required objectives:

- The user will wear the glove on which the flex and the gyro sensors are embedded to take data.
- The product will be able to detect sign language performed by a disabled person.
- Real time sign generated by the user will be detected by our machine learning code and its meaning is extracted accurately.
- All the processing is done by using the microcontroller named as Raspberry pi 3b+.
- The intelligent glove will also be connected to Bluetooth speaker to read loud output so that the result to be heard.
- Our project will give output in English language.

System Overview:

During the design of this system, our team must face many hurdles. The system needed a brain i.e. a micro-controller which can coordinate and control all the movements whose input would be provided by the user using a micro-computer. For the controller, raspberry pi model 3 B+ was selected as it provided the required features for our system design. A Bluetooth speaker was selected to give output in the form of speech which is already saved in the microcontroller. The components and the sensors which are used in this project are explained in detail below:

Flex Sensor:

We use 4 flex sensors in our project. Our objective was to measure the bending of the fingers. After researching all the wearable sensors, we came up with flex sensor. It fulfills all our requirements. Flex is simply a resistor or a potentiometer whose resistance changes as we bend the strip. Flex sensor has 2 terminals just like a resistor and output from this sensor is in analog. Flex sensor is available in two sizes that is 2.2 inch or 4.2 inch. We prefer using 2.2-inch flex sensor because it meets the glove finger's size requirement. The flat resistance of flex sensor is 25k ohms which varies from 45k to 125k ohms based on the bending. We use flex sensor as a potential divider by connecting another resistor in series. It is best illustrated by the following image We use flex on all the fingers of glove except last finger.

Flex Sensor

Flex Sensor Pinout

Figure 5

Pin Number	Description
P1	Usually connected to positive of power source.
P2	Usually connected to ground.

Gesture recognition system

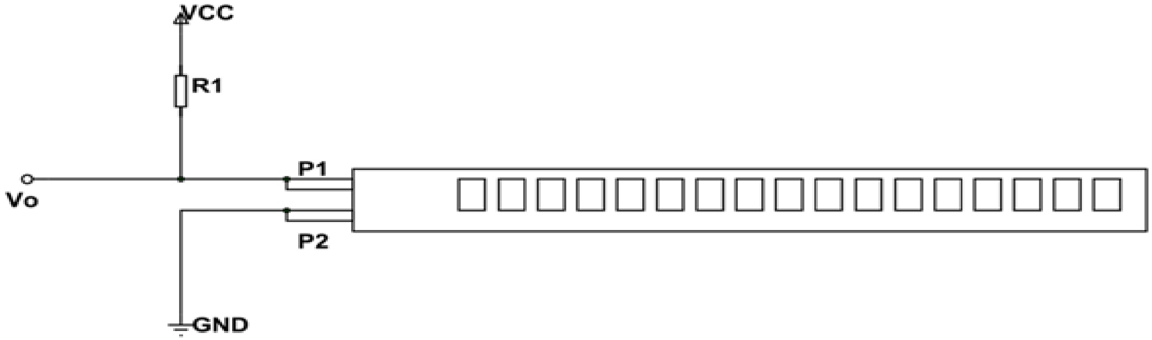


Figure 6

Vo is then directly connected to one of the terminals of mcp3008 ADC IC which is then fed to raspberry pi digital input.

MCP 3008:

There is no analog input pin on raspberry pi. But the output from flex sensor is in the form of analog signal, so we need something that can convert analog signal to digital signal. MCP 3008 prove to be the right choice. It has eight channels means it can convert 8 different analog signals to digital signal in run time. It is a 10-bit ADC IC. It measures analog voltage and after conversion it communicate to microcontroller through SPI communication.

MCP3008 8-Channel 10-bit ADC IC

MCP3008 Pinout

Figure 7

Gesture recognition system

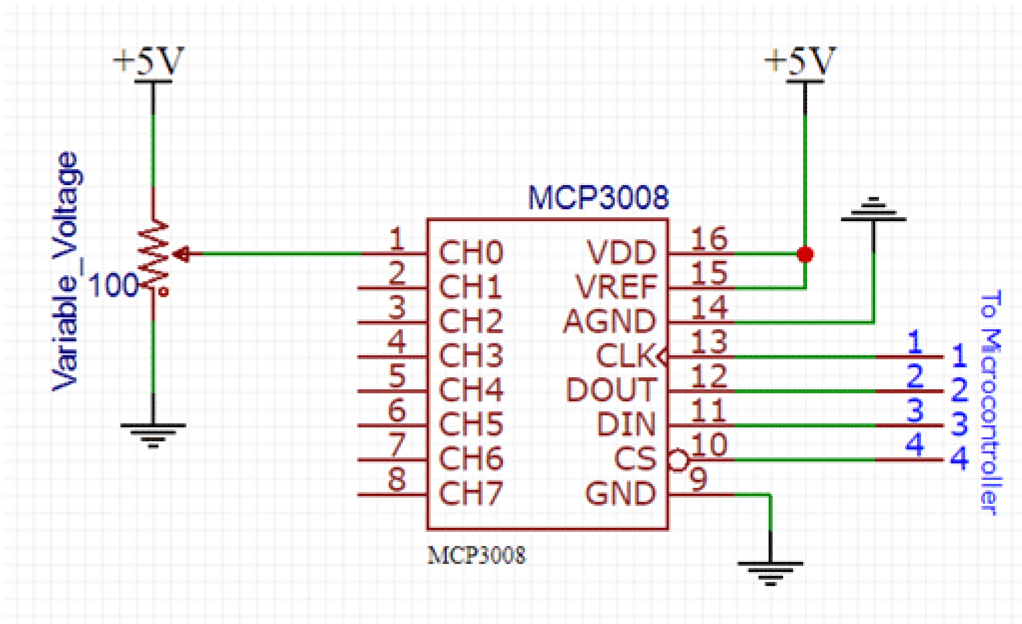


Figure 8

The formulae to convert this digital data back into voltage is:

$$\textbf{Analog Voltage Measured} = \textbf{ADC Reading} \times \frac{\textit{System Voltage}}{\textit{Resoltuion of ADC}}$$

In the above circuit diagram, the system voltage is 5V and the resolution of ADC is 102.

Gyro Sensor (MPU 6050):

To continuously monitor the acceleration of glove we use gyro sensor (MCP 6050). It measures the acceleration in x, y and z direction. It has measure rotation in 3 axis and temperature as well. It can be connected directly to raspberry pi as it has its own in build analog to digital converter. It uses I2C bus to communicate between different sensors and microcontroller.

It consists of

- DMP (Digital Motion Processor)
- 3-Axis Accelerometer
- On chip Temperature sensor

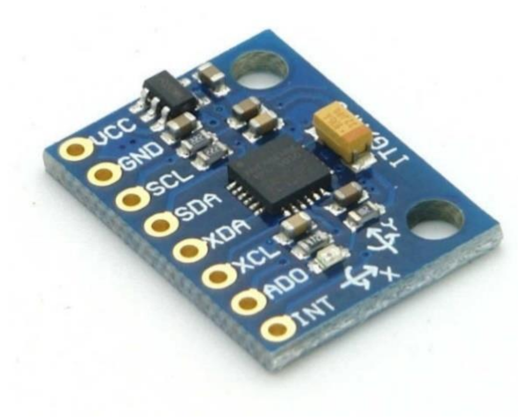


Figure 9

Raspberry Pi 3 Model B+:

The system utilizes a **Raspberry Pi 3B+** [4] micro-computer to serve as the controller which will take input and perform speech conversion along with the generation of Pakistan sign language (PSL) models. It has a Quad-Core ARM processor running at 1.4 GHz with 1GB of DDR2-RAM with on board wireless, Bluetooth, and Ethernet connections. It runs Linux-based kernel (Raspbian) and can connect with touch screens and monitors with HDMI ports. It has 4x USB ports to interface with serial devices.

Figure 10

Bluetooth Speaker:

Bluetooth speakers are those speakers which convert the audio file into the voice which can be audible loudly. These speakers use the Bluetooth technology to perform their action. The

Bluetooth of the Raspberry Pi 3B+ relates to the Bluetooth of the speaker which give the signals to the speaker to play the specific output. These speakers can be of different types.

Figure 11

Raspberry Touch Screen:

To display the results of the machine learning, 3.5-inch Raspberry Pi touchscreen was used which could be easily configured. This Raspberry Pi touchscreen will display the gift models based on the text and the voice recording through which the deaf will hear the speech in forms of Pakistan sign language (PSL). It is connected through the HDMI port.

Figure 12

Pakistan Sign Language (PSL):

In Pakistan sign language (PSL) there are a lot of signs which cover all the basic and the daily need of the people to communicate with other persons. In our project we include 16 signs which are copied from the website of the Pakistan sign language (PSL). All these signs are saved in the microcontroller along with their outputs in the form of voice notes. Machine learning is applied to get the most likely output. The list of these signs is given below:

- Beautiful
- Broad
- Class
- Cell number
- Book
- Bed
- Better
- Alarm clock
- Cheap
- Dumb
- Early
- Blank
- Nearby
- Energetic
- Sad
- Empty

Overall Description

Product Perspective:

The glove will be used for conversion from sign to speech and it will include:

- A glove with embedded sensors.
- Sign performed by a person is recorded.
- The recorded input will map to corresponding output.
- The output is shown on the screen.
- The output will also be hearable from the speaker.

Gesture recognition system

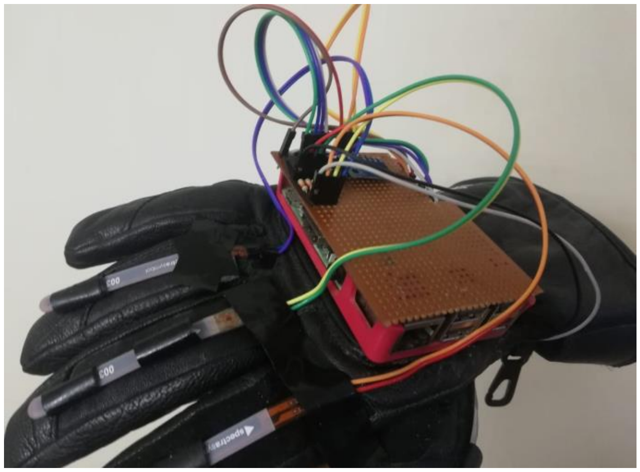


Figure 13

The microcontroller and sensors will be powered by a power-bank connected with the raspberry pi. The readings and input of all sensors, flex and gyro sensor combined will be sent to the raspberry pi and our python script running the hardware will record and save these values in a csv file in the form of a list or data frame.

Product Functions:

Hardware:

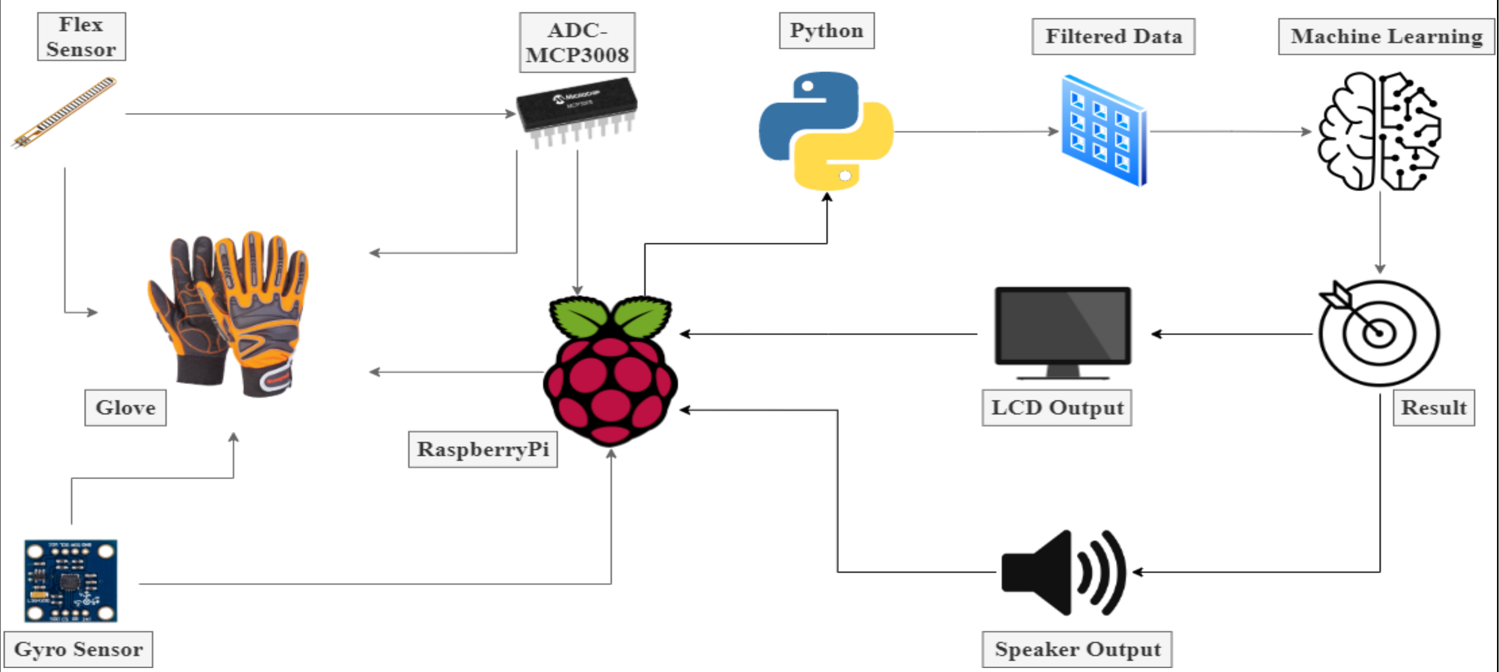


Figure 14

1. The user will wear the glove on which the raspberry pi and the sensors are attached for the ease of use.
2. The product will be able to detect sign language performed by a disabled person.
3. Real time sign generated by the user will be detected by our machine learning code and its meaning is extracted accurately.
4. The intelligent glove will also be connected to LCD to display output and to speaker to read the result loudly which is to be heard.
5. Our project will give output in English language.

Software:

Data Generation

1. The data is in the form of arithmetic values for both flex and gyro sensors.
2. Flex sensors generate voltage values which will change respectively if we bend the finger or not.
3. Gyro sensors generate acceleration values with acceleration in all three dimensions x-axis, y-axis, and z-axis.
4. The code contains a loop and each time it runs will generate value of sensors. In other words, you can say that sensors will give input data for each execution of the loop for a limited amount of time the loop is running, and data is appended into a list.
5. The lists of data generated from the loop will then be converted to a data frame and is pushed to a csv file containing all the previous records of data generated.

Machine Learning Implementation:

1. To generate the result and output data accurately we are implementing machine learning classification model.
2. The machine learning model is trained on the sample training data. This sample data will come from the previously generated csv file.
3. The machine learning code will first open the file, read its content, and append it to data frame in the specified variables for training.
4. The scikit learn library is used to import classification model function and give it the training values data set.

5. After the training we can also give test values to a prediction function imported from the same scikit learn library and give it the test values.
6. The model will predict accurate output based on the given testing sample data and display the output on the screen as well as on the speaker.

Operating Environment:

For our hardware part we purchased and used the following equipment and sensors:

- Raspberry Pi 3B+ (1x)
- Flex Sensors FS-L-0095 or equivalent (4x)
- Gyro Sensor MCP 6050 (1x)
- Leather Glove (1x)
- Jumper Wires (30x)
- MCP 3008 Analog to Digital converter (1x)
- PCB Board (1x)
- Small LCD screen attached to Raspberry Pi (1x)
- Power Bank (1x)
- USB Data Cables (2x)
- HDMI Cable (1x)

Software component and technologies used:

- Raspbian Operating System
- Python3
- Thorny Python compiler and editor
- PSL Pakistan sign language website for signs dictionary
- Teams Viewer for remotely accessing Raspberry pi.
- CSV file
- Python libraries used to get sensors data readings:
  - RPi.GPIO
  - time
  - Adafruit_MCP3008

    - smbus
    - numpy
    - pandas
    - math
- Python libraries used for implementing machine learning:
    - numpy
    - glob
    - pandas
    - sklearn

External Interface Requirement

Data Collection

The data is generated from gyro sensor and four flex sensors. There are almost 70 data points generated when we run the project for about 5 seconds. The data points can also be increased by reducing the sleep time. The user had to perform the sign in 5 seconds for the project to distinguish between signs.

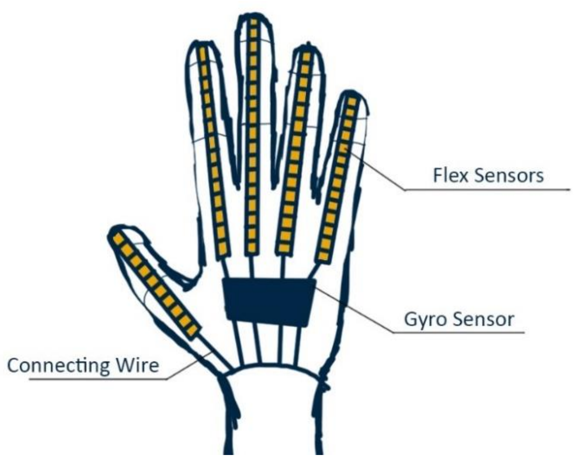


Figure 15

Data Storage:

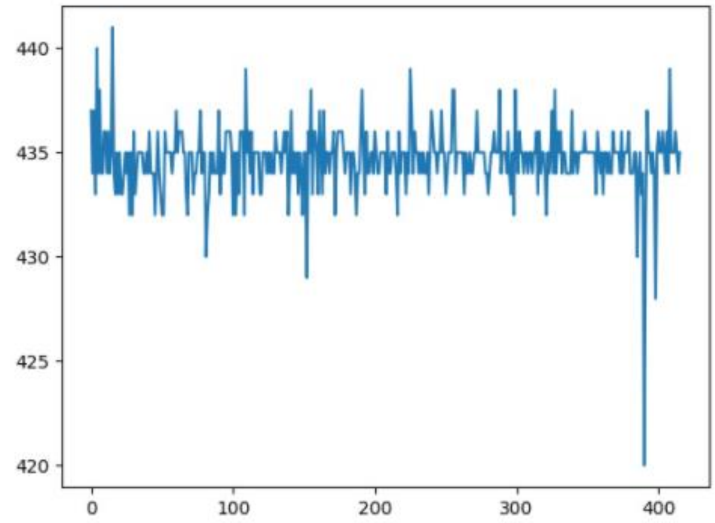
The data is stored in csv file. There are two csv file used in our project. The first csv file used to store the training data sets. Now this file contains dataset of more than 100 samples for each dataset. The data collected for each sample is initially in the form of rows and columns. We then convert the data in a single row and assign the relevant sign at the last column. The next csv file is for testing purpose. The real time data is stored in the file for comparison. The data is overwriting every time you run the program.

AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV	AW	AX	AY	AZ	BA	BB	BC
71	18	436	453	24	6	-65	435	453	28	42	-35	434	466	-8	69	20	461	438	-4	61	29	456	442	beautiful				
47	-11	432	441	-14	52	34	461	441	-1	65	25	456	392	30	-12	-58	411	433	26	-33	-46	444	435	beautiful				
34	-32	452	461	31	59	-1	455	451	-12	67	19	459	446	23	11	-65	446	441	25	-30	-49	441	453	beautiful				
58	-9	455	437	-4	83	5	459	437	36	-25	-43	440	444	29	-32	-44	439	450	26	-36	-43	363	442	beautiful				
53	-15	457	444	-11	75	10	460	441	50	21	-32	444	437	46	-31	-27	442	445	41	-36	-29	442	442	beautiful				
45	-10	447	452	-18	58	26	460	439	16	72	8	452	443	49	-20	-34	455	443	44	-32	-30	451	442	beautiful				
54	-16	451	448	-11	73	13	460	439	59	-11	-29	439	445	48	-32	-23	446	449	47	-32	-24	441	442	beautiful				
57	5	448	445	-18	66	16	466	439	46	36	-22	437	401	37	-36	-33	441	448	34	-33	-38	442	442	beautiful				
35	-23	435	452	-12	53	34	460	439	14	60	26	446	443	36	-30	-40	441	444	36	-30	-39	444	455	beautiful				
52	-4	432	454	1	84	6	460	441	12	65	22	449	444	45	-21	-38	436	442	35	-32	-38	445	442	beautiful				
73	5	450	449	-7	65	23	460	438	44	-19	-40	443	440	49	-32	-23	438	439	44	-34	-27	443	440	beautiful				
67	15	459	437	43	45	-12	444	428	31	-28	-45	449	450	42	-32	-31	443	441	47	-28	-30	443	435	beautiful				
2	-8	454	458	33	-24	-47	445	440	25	-34	-46	453	451	25	-33	-47	447	443	25	-34	-45	445	444	better				
-9	-10	459	468	32	-26	-47	447	444	21	-34	-48	444	442	21	-35	-47	439	440	20	-34	-49	445	443	better				
-3	-6	437	457	38	-20	-45	451	446	28	-33	-44	366	444	33	-33	-40	444	372	32	-33	-41	443	442	better				
3	-7	467	453	85	1	-5	450	450	45	-33	-26	446	457	43	-37	-25	442	446	44	-38	-22	447	448	better				
3	3	446	457	72	17	-6	451	410	47	-31	-26	447	440	45	-32	-28	446	466	44	-32	-29	445	446	better				
44	2	460	452	79	11	-3	462	455	55	-30	-17	446	442	42	-35	-29	449	435	42	-37	-26	456	445	better				
0	0	460	449	78	-11	-5	458	531	54	-19	-30	448	445	40	-36	-29	445	457	41	-37	-27	446	440	better				
-11	-15	453	454	85	-5	0	455	467	84	-6	3	461	455	53	-33	-17	446	445	53	-33	-15	447	445	better				
-4	-13	466	455	79	10	-4	462	453	50	-22	-32	446	446	33	-39	-34	452	444	33	-41	-32	450	445	better				
69	-11	461	456	73	16	5	460	456	28	-36	-41	444	442	39	-40	-26	444	438	37	-40	-28	444	445	better				
56	-14	457	456	73	12	12	463	457	24	-39	-42	434	447	31	-41	-34	443	448	30	-43	-32	437	447	better				
5	4	460	462	69	15	15	462	460	81	-8	-4	454	462	34	-40	-31	445	444	33	-41	-32	453	446	better				
-1	-1	463	459	72	10	15	460	460	89	0	-1	456	450	32	-44	-30	442	451	30	-45	-30	443	460	better				
-2	3	459	457	76	13	-5	459	464	85	3	4	447	457	36	-36	-34	445	447	29	-41	-35	448	470	better				
-1	0	460	456	75	9	11	477	458	41	-25	-39	444	444	32	-39	-35	446	451	32	-42	-32	447	445	better				
7	-11	459	454	78	10	-6	460	450	36	-31	-39	446	444	24	-41	-39	445	444	23	-42	-39	446	445	better				
26	8	455	444	71	18	4	556	456	82	6	5	457	460	37	-33	-36	445	446	29	-41	-35	447	437	better				
0	3	460	454	81	9	0	463	458	53	-14	-34	447	448	23	-39	-43	446	445	25	-39	-41	447	446	better				
19	-18	445	449	70	-19	8	460	454	72	17	-3	450	461	32	-43	-30	444	446	35	-46	-23	443	449	better				
-1	-9	456	461	75	14	2	458	455	66	-10	-22	442	444	35	-38	-32	440	446	34	-39	-32	444	443	better				
4	-11	460	452	84	3	5	415	457	71	-18	-4	441	452	36	-42	-27	445	453	34	-45	-26	447	450	better				
70	4	452	457	68	19	-9	455	457	19	-34	-50	445	446	28	-39	-38	449	446	27	-38	-40	444	446	better				
-1	-17	457	455	80	9	-5	460	451	56	-16	-29	447	448	26	-36	-43	446	444	26	-37	-42	455	446	better				
70	5	457	444	77	-5	-12	460	453	79	11	-1	460	458	24	-25	-54	447	440	24	-38	-42	447	445	better				
84	5	461	445	18	63	-19	452	449	45	45	-3	470	458	73	17	3	461	455	43	-24	-38	460	446	better				
7	-6	450	451	73	6	17	460	456	57	28	-16	450	367	29	-32	-44	452	441	29	-37	-39	442	436	better				
2	-21	453	446	-6	55	-35	460	456	69	20	-4	456	516	19	-37	-47	447	447	22	-41	-41	438	446	better				
32	-17	459	448	-8	77	9	464	434	74	9	-13	458	454	73	16	-3	457	448	36	-27	-42	446	443	better				
62	-11	461	456	67	23	-2	460	451	28	-35	-43	437	365	21	-40	-43	446	445	23	-42	-40	440	444	better				
29	53	458	435	83	-6	-3	433	447	78	-10	6	460	456	31	-2	-59	439	449	26	-41	-38	445	453	better				
2	66	457	457	42	2	48	457	459	60	1	30	454	457	70	-3	-20	438	449	69	-6	-20	438	451	better				
30	10	449	453	53	1	37	457	457	74	6	15	452	453	55	-6	34	449	458	69	-7	20	457	457	better				
10	19	454	455	59	7	30	455	458	63	-6	26	457	458	56	-3	34	457	452	81	-6	6	457	452	better				
3	32	456	459	51	-6	39	460	463	57	-1	33	456	459	56	-2	34	442	458	58	-3	32	451	449	better				
0	43	458	459	68	-5	21	454	458	51	0	39	440	459	62	-5	27	451	458	71	-7	-18	446	472	better				

Figure 16

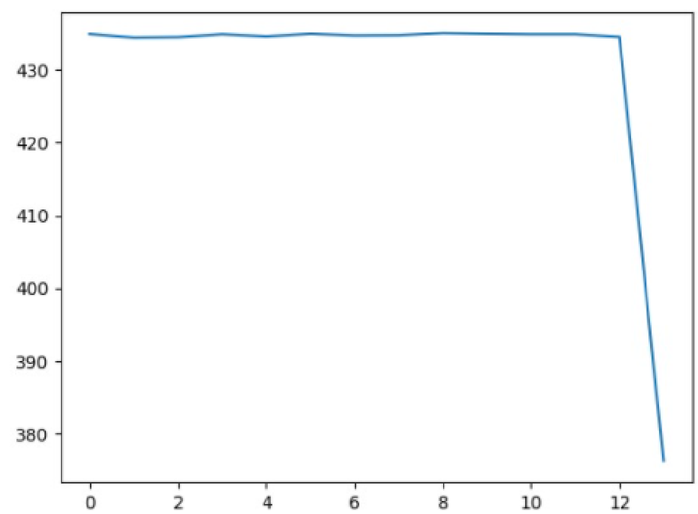
Data cleaning:

One the important process of implementation of efficient machine learning is data cleaning. We use “Median algorithm” to erase unnecessary data. In this algorithm we divide the dataset in 12 equal parts with 30 samples each. Now, we arrange the data in ascending order and remove 5 samples from both extreme. We now have data of 20 sample, take the average and come up with a single digit.



nonfiltered

Figure 17



filtered

Figure 18

Data visualization:

Initially we perform different test and experiment to analyze the graph. We perform same signs repeatedly to see weather they are somewhat similar or not. After that we also perform two different signs, and we see a clear difference in the graph.

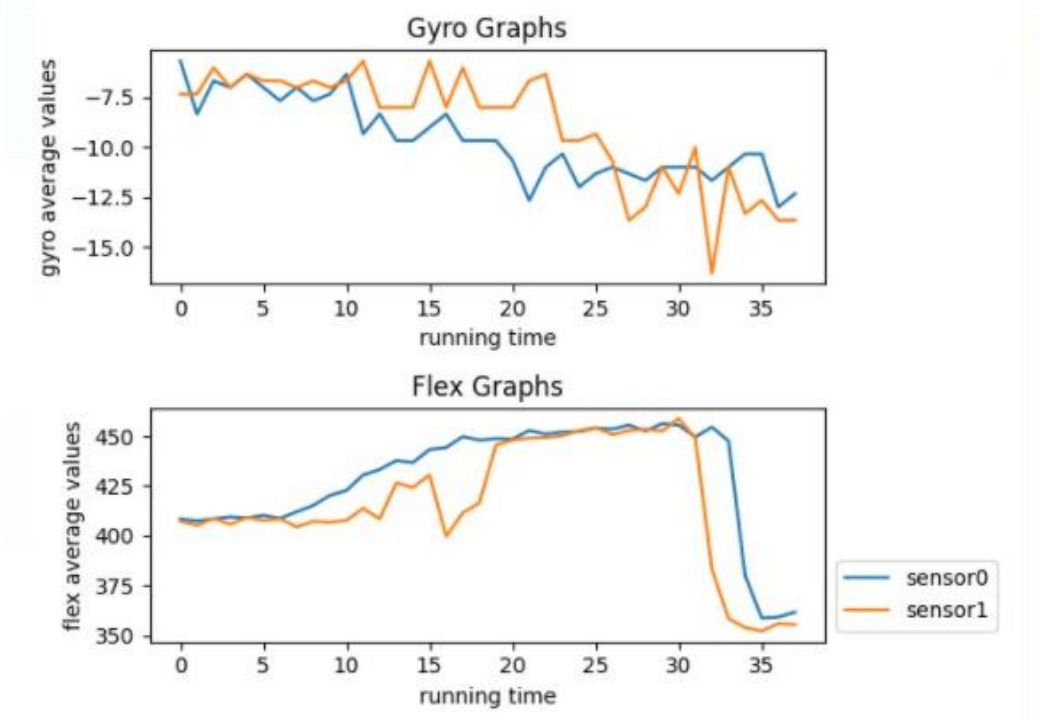


Figure 19 (Graph of same signs performed twice.)

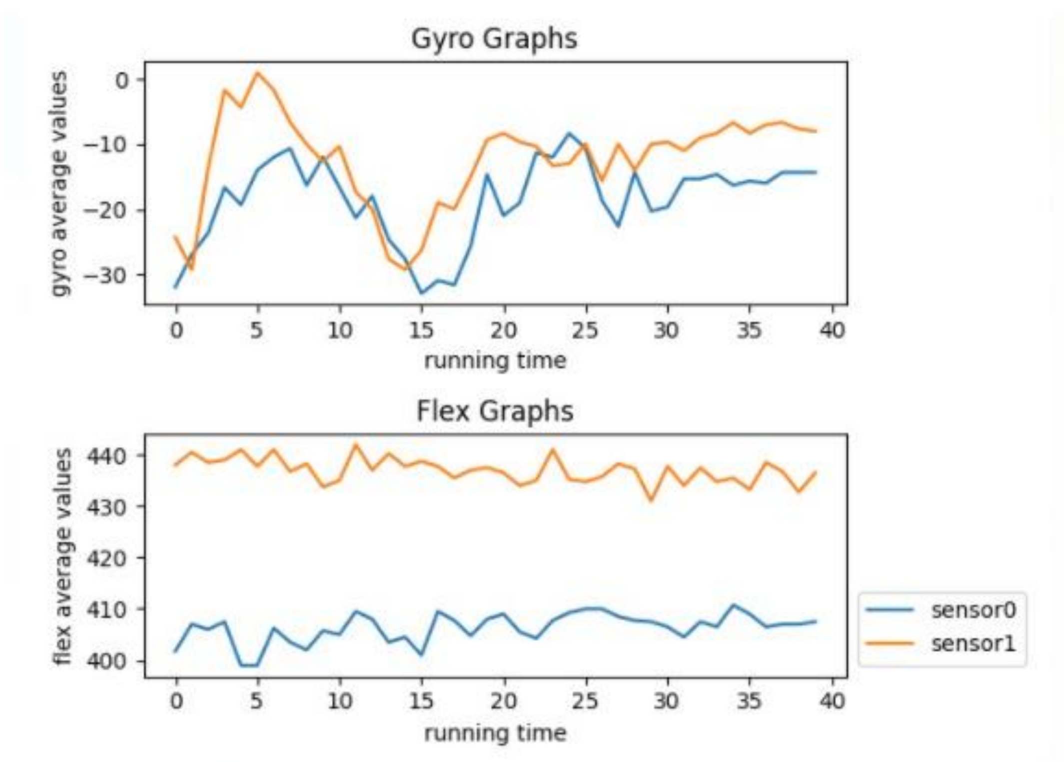


Figure 20 (Graph of two different signs.)

Hardware Interface:

The hardware consists of raspberry pi 3b+ as a processor, 4x flex and 1x gyro sensor. The flex sensor measures the bending of finger and the gyro sensor measure the acceleration in x, y and z direction. There is no analog input pin available on raspberry pi, so we use MCP3008 IC to convert the analog signal from 4 flex sensor to digital signal. We use 10k resistor in series with flex sensor and use flex sensor as a potential divider. There is inbuild analog to digital converter in gyro sensor, so we don’t need to convert it by ourselves.

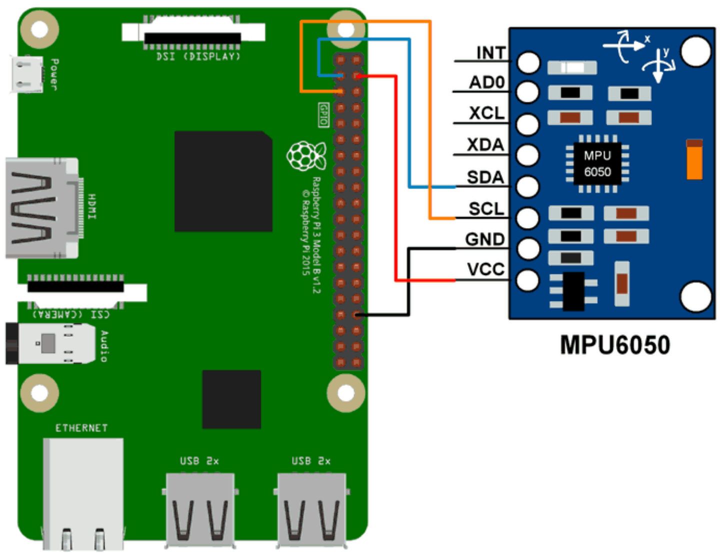


Figure 21 (Gyro sensor interfacing with raspberry pi.)

Gesture recognition system

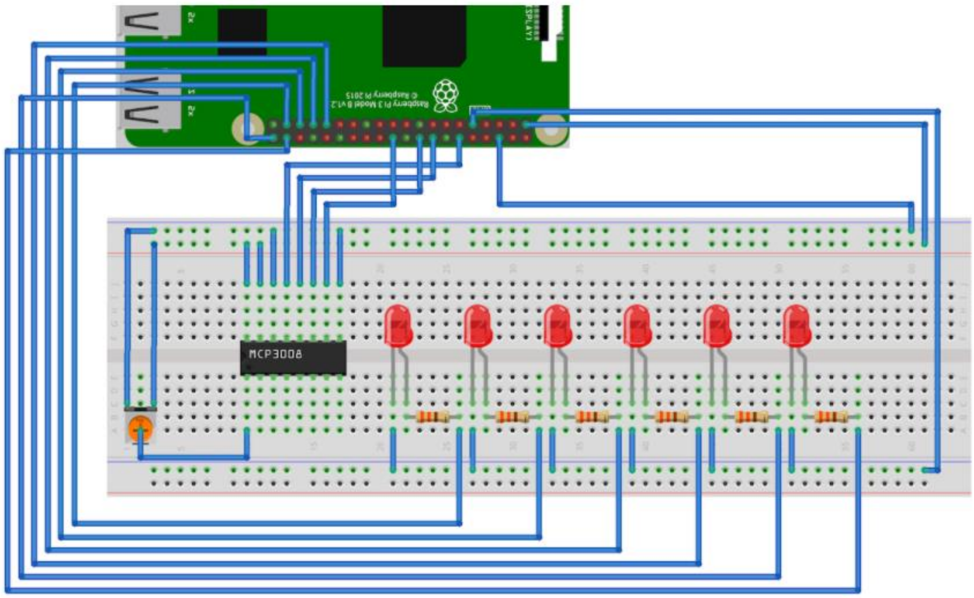


Figure 22 (Raspberry pi interfacing with MCP3008 IC)

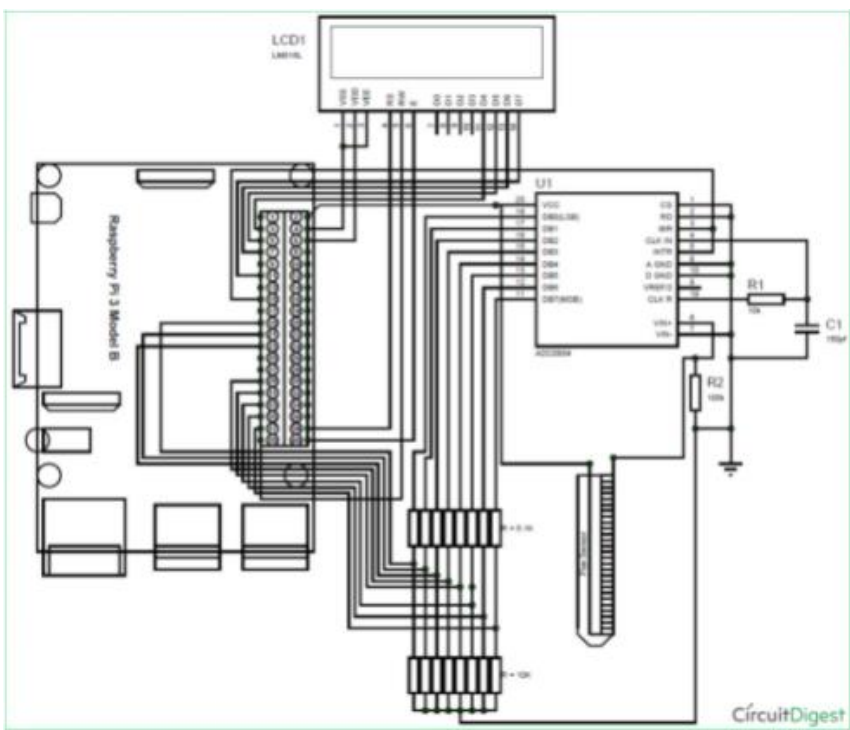


Figure 23 (Flex sensors interfacing with the raspberry pi.)

Software Interface:

The project is programmed using python programming language. Training data is stored in csv file and machine learning algorithm is deployed for high accuracy and efficiency. Tree classifier is used as our main machine learning algorithm. It is predicting correctly among all the signs. “Thonny Python IDE” is used to program our project. It is the in-build raspberry pi python IDE. “Sklearn” library is used for machine learning algorithms. We used the “Decision tree classifier” for implementing the machine learning into the project. This tree classifies the data by sorting the tree from the node to the terminal. We also import mcp6050 and mcp3008 library in our project. Bluetooth speaker will be integrated with the project to receive the output in the form of speech.

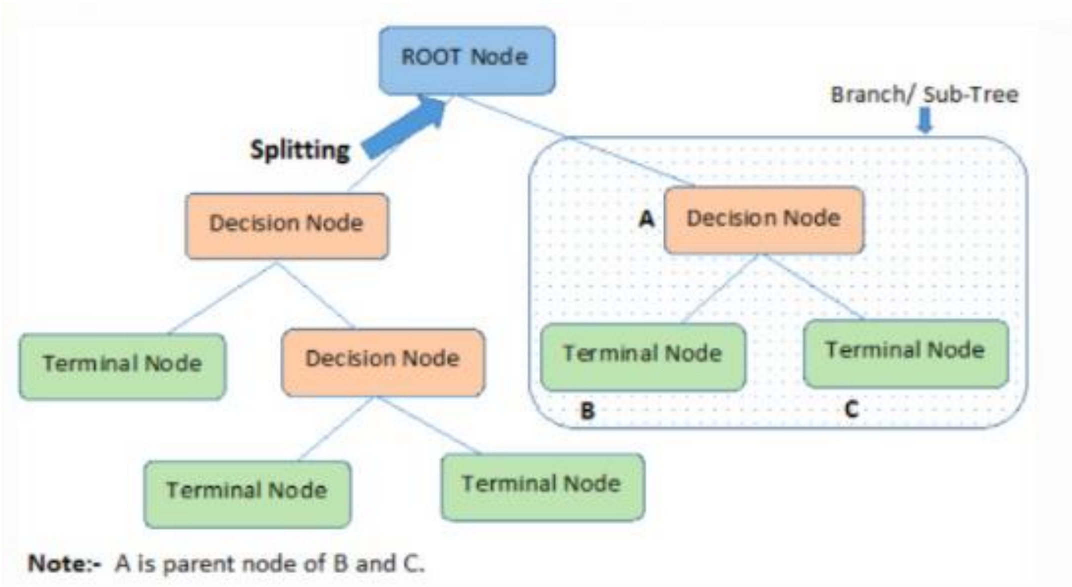


Figure 24

Final Product Shape:

The shape of the final product for the sign to speech for disabled people is shown in the picture below. All the wires and the raspberry pi which is used as a microcontroller is integrated with each other and the final output is displayed on the screen as well as on the Bluetooth speaker.

Figure 25

System Feature

System features in terms of hardware

1. The glove is easily wearable.

Description:

The glove is made up leather and is easily wearable. Raspberry pi and power bank is embedded on glove. The glove is light weight and easy to use, this feature is high priority since the glove is light weight which is much convenient to use.

Stimulus/response sequence:

The purpose of making the glove light weight is that the user does not feel difficult to use and can perform signs with ease.

Functional Requirement:

In order to make the glove light weight and wearable we use leather glove and replace the breadboard with small PCB chip.

2. Real time data collection is done.

Description:

The real time data from sensor is collected and stored in csv file in rows and columns which is then converted to a single row.

Functional Requirement:

For real time data collection, we use csv file and panda's library to convert it in a single row.

3. The product is cost friendly.

Description:

The project cost is under $200 and can be reduce if we use cheap microcontroller. In the project after raspberry pi, flex sensors are costly. The price could reduce if we bought this in bulk and use less flex sensor as well.

Stimulus/response sequence:

Cost is another important factor that customer will notice before taking the leap of faith and buying the product. Along with the many system features, the user will be compelled to think twice about not buying the product.

Functional Requirement:

To drive the price down, we have researched for the best possible components and balanced the quality and price parameter of the product.

4. The product requires no prior knowledge.

Description:

If you sign from Pakistan Sign language website, you are good to use our product. It only takes gesture as an input and convert to relevant speech.

Stimulus/response sequence:

The product is also designed for school going special children. They don’t have much knowledge of how technology works. The plus point is the simplicity of our project make it easier for them to use the product.

Functional Requirement:

The project uses both English and Urdu language so that the listener can understand it clearly.

System features in terms of software

5. Web portal to display results.

Description:

You can access the project code and implementation if the raspberry pi is connected to Wi-Fi.

Stimulus/response sequence:

If there is a need to upgrade or update the code, you can always do that by just connecting the raspberry pi to Wi-Fi and make the required changes on web browser.

Functional Requirement:

The project is connected to IP address where it shows the output, accuracy, and code details.

6. Anyone can use the product if they know the credentials.

Description:

Anyone can use the project it only requires the customer to buy the product there is a user manual available which will provide the step-by-step guidance to the product. The user only needs to know the login credentials.

Stimulus/response sequence:

Illiterate children have less knowledge about the technology therefore they can easily understand the guide and play with the product. The only thing they need to know is the relevant signs.

Functional Requirement:

Our system has plug and play capabilities which makes installation and admiration easy.

7. Output on Bluetooth speaker.

Description:

The output of the project is displayed on LCD as well as on speaker in the form of speech.

Stimulus/response sequence:

It is difficult and not convenient for the other person to come and see the output on LCD. So, it is desired to achieve the output in the form of speech on wireless speaker.

Functional Requirement:

We are correctly working to achieve the output on speaker. We will be using Bluetooth speaker to achieve our desire output.

NON-FUNCTIONAL REQUIREMENTS

Scalability:

Under excessive use like performing signs continuously and for a longer time period it will give output all the same regardless of this excess use.

Portability:

Easy to carry and move from one place to another by just simply carrying hardware contained box and glove with care as to not damage the wires or disrupt connections.

Performance:

Raspberry pi is a powerful hardware tool with a lot more RAM and storage than other microcontroller like Arduino and it shows results in a matter of seconds.

Compatibility:

The software tools used are up to date with the latest version and can be updated further if new versions come out without disrupting the working of project. Whatsoever software updates will cause no issues.

Reliability:

Under normal usage conditions like performing pre-defined signs from specific sign language library will cause no problems but performing different sign with expectation of getting same output will result in failure.

Maintainability:

The maintenance part may take a lot of time depends on the functionality changing ranging from 90% chance per day to 10% chance per day as hardware replacement and software programming changes can take a lot of time for building, testing and implementing.

Availability:

The availability of the product relies heavily on its reliability and maintainability as described above the availability can become an issue if maintenance could take more time and will not be accessible for a large period.

Business Plan:

The market size of our project in Pakistan is very large as around 5 to 6% of the population of Pakistan is deaf this earns more people need such devices which can give them relaxation and they can easily communicate with the other people. Many people want to talk with other people and try to tell them their feeling and thoughts but due to this gap, they can’t do this. Our device will make them communicate with people easier.

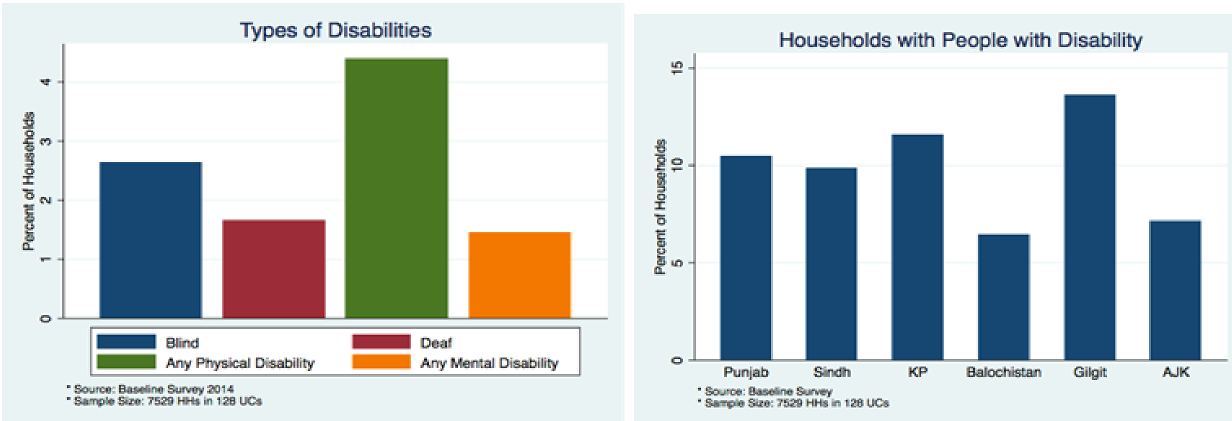


Figure 26

Service or Product line:

We will give them a smart hardware-based whole system with reduced components on the glove. It will include machine learning on the sensors to work them properly. All the organizations included government as well the private can use this project to improve efficiency.

Marketing and sale:

We will market our project in different cities and places to make the maximum customers of our project. We want to decrease the difficulties of deaf people by giving the hand glove device with the minimum rate to the organizations.

Methods of Income:

There will be two methods used to generate revenue for our product. These two methods are discussed in detail below:

- **Sales of our project:**

We will try to generate revenue by providing the basic model of our hand glove to the organization and large companies who are interested in our project and will take benefit from this deal. All the data which is being trained will be given to the interested companies. This will make the organizations to the product in the large amount as no one will question them and they can take profit from this project as many as they want.

- ### Data sell by marketing:

We will do marketing of our product to make them more and more customers and will sell them directly by making the more such devices. Through this way, we can earn more profit than by selling our product to the organization or the company only. We will also try to convince the government that what is the value of our product if they made contact with us then it will be more beneficial for us.

Competitive advantage:

We are on the top when comparing our product with the same product of other organizations. We will try to sell our product by half of the price of the other competitive company. The sensors embedded in the glove and the microcontroller which is raspberry pi 3b+ are not officially launched in the market of Pakistan. All these factors will make the company more financially strong. Furthermore, there is no such big competition for our product in the market which is also a positive point.

Testing and Results:

The performance of the designed model is analyzed by one class to other class basis and using all the 16 words. Different types of detection methods are implemented to find the precision and the accuracy of all the different types of the alphabets. We used the configure matrix for determining the accuracy and precision of all the signs which we used in our project. Through this method the performance and the efficiency of the overall system can be determine. The final results of the configure matrix are attached below:

Confusion Matrix

	alarmclock	beautiful	bed	better	blank	book	broad	cellnumber	cheap	class	dumb	early	empty	energetic	sad
alarmclock	0.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.01	0.00
beautiful	0.00	0.90	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.06	0.00
bed	0.00	0.00	0.89	0.00	0.00	0.08	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.01
better	0.00	0.09	0.00	0.87	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
blank	0.00	0.11	0.02	0.03	0.75	0.02	0.00	0.00	0.00	0.00	0.02	0.00	0.03	0.02	0.00
book	0.00	0.04	0.00	0.00	0.00	0.87	0.00	0.00	0.00	0.00	0.00	0.06	0.04	0.00	0.00
broad	0.00	0.00	0.00	0.06	0.00	0.00	0.83	0.00	0.00	0.00	0.10	0.00	0.02	0.00	0.00
cellnumber	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
cheap	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.77	0.00	0.21	0.00	0.02	0.00	0.00
class	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.89	0.00	0.00	0.00	0.00	0.00
dumb	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.81	0.00	0.00	0.02	0.00
early	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.06	0.73	0.17	0.00	0.00
empty	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.95	0.00	0.00
energetic	0.00	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.88	0.02
sad	0.00	0.00	0.00	0.11	0.00	0.25	0.00	0.00	0.00	0.00	0.06	0.00	0.00	0.00	0.58

Color Scheme	
0.00	represents 0
0.02	between 0.01 and 0.1
0.21	between 0.1 and 0.5
1.00	represent > 0.5

Figure 27

The python code which we used in our project is shown in the annexures. All the codes are written in python, and we used the visual studio for the running of code.

When someone perform any sign then the output of the system is shown on the LCD as well as in the form of voice from the Bluetooth speakers. The output of the "Energetic" sign can be shown in the figure below:

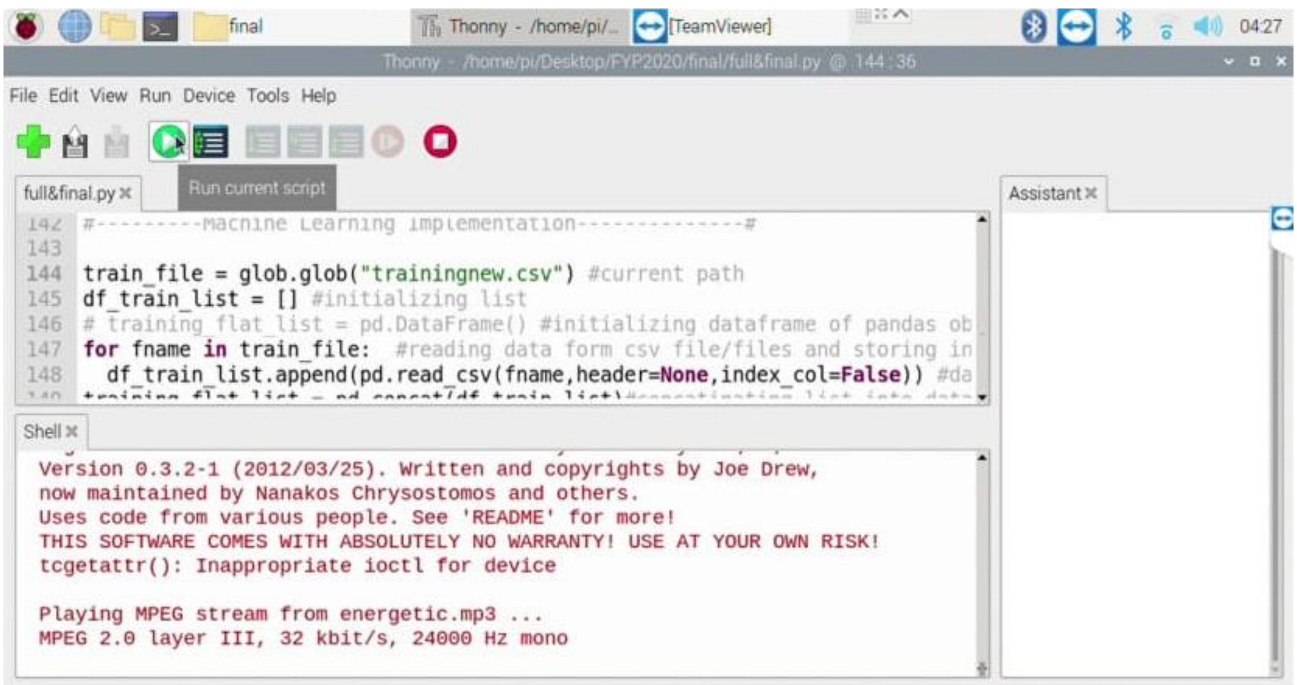


Figure 28

Cost Analysis:

One of the biggest selling point of this project is it being very inexpensive in nature. Now we will discuss in detail about the total cost of the final system.

Gesture recognition system

S/No	Product	Quantity	Price (PKR)
1	Raspberry Pi 3B+ with Accessories	1	6000
2	3.5” Raspberry Pi Touch Screen	1	2500
3	Flex sensors	4	8000
4	Gyro sensor	1	700
5	Bluetooth Speaker	1	500
6	ADC 3008 IC	1	500
7	Wires and tapes etc.	1	300
	Grand Total		18500

(Actual cost of product may vary due to changing in price in markets)

Work Distribution:

This project of sign to speech for disabled people is done by three group members and all the members participate equally to finish this project. This whole project is divided into 2 parts.

Gesture recognition system

1. Hardware
2. Software

The project is a combined effort of every team member, however the work among members is distributed as follows:

MUHAMMAD HUZAIFA	**MUHAMMAD SAROSH**	**AHMAD ALI**
Test and program flex and gyro sensors	*Implemented machine learning algorithm*	*Hardware attachment on glove*
Integration of MCP 3008 IC ADC with flex sensor	*Train Model*	*Literature Review of similar research works.*
Data Collection and testing	*Data processing and cleaning*	*Documentation*

Future Works

Current Problems and limitations:

Sign to Speech Language Translator:

a. The Raspberry Pi 3B+ used for the wearable device is a little heavy for long term usage by the deaf community.
b. The response time from speech recognition to sign language display is 1-2 seconds, which limits the extent of the usage for long conversations.
c. If the project is used by some random person whose data is not present in the train model database, then the amount of error increase drastically.
d. The project is a bit expensive because of costly flex sensor.

e. Wearing glove in summer is quite handy.

Future Improvements and Recommendations:

Sign Language Translator:

a. A dedicated processor can be deployed to improve real time response.
b. Further research is needed for a suitable gesture recognition system to reduce the response time.
c. Immense amount of training using deep neural networks is needed for including words in the dictionary.
d. Reduce the project cost by maybe using cheap processor or using less flex sensors.

Conclusion:

This project will help the mute and deaf people to communicate with normal people effectively. Pakistan sign language which is being taught in schools is a plus for us as there we can directly deploy our project. Most of the people cannot understand sign language of mute people. This project will enable them to listen and speak in their own way in real time. The products being inexpensive could be utilized by any deaf person to listen to what people are talking. This project helped us understanding and utilizing the practical end of our skills. It works decently and can be marketed so that maximum people benefit from it.

References:

Raghavendra Reddy. B, Mahender, “Speech to Text Conversion using Android Platform”, International Journal of Engineering Research and Applications (IJERA) Vol. 3, Issue 1, pp.253-258, January -February 2013.

Sangeetha, K. and Bharathi Krishna, L. “Gesture Detection for Deaf and Dumb People” International Journal of Development Research Vol. 4, Issue, 3, pp. 749-752, March, 2014.

Shanmukha Swamy, Chethan M P, Mahanta’s Gatward, “Indian Sign Language Interpreter with Android Implementation”, International Journal of Computer Applications (0975 – 8887) Volume 97– No.13, pp. 36-41, July 2014.

Sinora Ghosalkar ,Saurabh Pandey,Shailesh Padhra,Tanvi Apte, “Android Application on Examination Using Speech Technology for Blind People”, International Journal of Research in Computer and Communication Technology, Vol 3, Issue3, March-2014.

Tapas Kumar Patra, Biplab Patra, Puspanjali Mohapatra, “Text to Speech Conversion with Phonematic Concatenation”, International Journal of Electronics Communication and Computer Technology (IJECCT) Volume 2 Issue 5, pp.223-226, September 2012.

A.-C. Tsai, A. Paul, J.-C. Wang and J.-F. Wang, "Intensity gradient technique for efficient intra-prediction in H. 264/AVC", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 5, pp. 694-698, 2008.

R. Pradipa and S. Kavitha, *Hand gesture recognition-analysis of various techniques methods and their algorithms*, pp. 2003-2010.

S. Mitra and T. Acharya, "Gesture recognition: A survey", *IEEE Transactions on Systems Man and Cybernetics Part C (Applications and Reviews)*, vol. 37, no. 3, pp. 311-324, 2007.

W. Gonzalez and R.E. Woods, Eddins Digital Image Processing Using MATLAB, Third New Jersey:Prentice Hall, 2004.

G. Amalorpavam, H.T. Naik, J. Kumari and M. Suresha, "Analysis of Digital Images Using Morphlogical Operations", *International Journal of Computer Science & Information Technology*, vol. 5, no. 1, pp. 145, 2013.

Annexures:

Raspberry Pi 3b+ Code:

```
# Import libraries

import Adafruit_MCP3008

import time

import RPi.GPIO as GPIO

GPIO.setmode(GPIO.BCM)

import smbus

import math

import numpy as np

from csv import writer

import pandas as pd

import glob

import os

import matplotlib.pyplot as plt

from sklearn import datasets
```

Gesture recognition system

```
from sklearn.metrics import confusion_matrix

from sklearn.model_selection import train_test_split

from sklearn.metrics import accuracy_score

from sklearn.tree import DecisionTreeClassifier

from gtts import gTTS

#-------------------------------------------

# Some MPU6050 Registers and their Address

#-------------------------------------------

PWR_MGMT_1   = 0x6B

SMPLRT_DIV   = 0x19

CONFIG       = 0x1A

GYRO_CONFIG  = 0x1B

INT_ENABLE   = 0x38

ACCEL_XOUT_H = 0x3B

ACCEL_YOUT_H = 0x3D

ACCEL_ZOUT_H = 0x3F

GYRO_XOUT_H  = 0x43

GYRO_YOUT_H  = 0x45

GYRO_ZOUT_H  = 0x47

#-------------------------------------------

# Post-Processed MPU6050 Global Variables
```

Gesture recognition system

```
#--------------------------------------------

Ax=0

Ay=0

Az=0

Gx=0

Gy=0

Gz=0

G_x=0

G_y=0

G_z=0

A_x=0

A_y=0

A_z=0

a = 0

factor = 1.0

Mul_acc = 16384.0/factor

Mul_gyro= 131.0/factor


SAMPLING_RATE = 60

SAMPLING_TIME = 1/SAMPLING_RATE

CLK  = 18
```

Gesture recognition system

```
MISO = 23

MOSI = 24

CS   = 25

mcp = Adafruit_MCP3008.MCP3008(clk=CLK, cs=CS, miso=MISO, mosi=MOSI)

def MPU_values():

    global Ax

    global Ay

    global Az

    global Gx

    global Gy

    global Gz

    global A_x

    global A_y

    global A_z

    global Mul_acc

    global Mul_gyro

    #Read Accelerometer raw value

    acc_x = read_raw_data(ACCEL_XOUT_H)

    acc_y = read_raw_data(ACCEL_YOUT_H)

    acc_z = read_raw_data(ACCEL_ZOUT_H)

    #Read Gyroscope raw value
```

```
gyro_x = read_raw_data(GYRO_XOUT_H)

gyro_y = read_raw_data(GYRO_YOUT_H)

gyro_z = read_raw_data(GYRO_ZOUT_H)

Ax = acc_x/Mul_acc

#Ax = acc_x

Ay = acc_y/Mul_acc

#Ay = acc_y

Az = acc_z/Mul_acc

#Az = acc_z

Gx = gyro_x/Mul_gyro

Gy = gyro_y/Mul_gyro

Gz = gyro_z/Mul_gyro

x2 = Ax*Ax

y2 = Ay*Ay

z2 = Az*Az

sqr_x = math.sqrt(y2+z2)

sqr_y = math.sqrt(x2+z2)

sqr_z = math.sqrt(y2+x2)

A_x = 57.298*math.atan2(Ax,sqr_x)

A_y = 57.298*math.atan2(Ay,sqr_y)

A_z = 57.298*math.atan2(Az,sqr_z)
```

```
def MPU_Init():
    #write to sample rate register
    bus.write_byte_data(Device_Address, SMPLRT_DIV, 0)

    #Write to power management register
    bus.write_byte_data(Device_Address, PWR_MGMT_1, 1)

    #Write to Configuration register
    bus.write_byte_data(Device_Address, CONFIG, 1)

    #Write to Gyro configuration register
    bus.write_byte_data(Device_Address, GYRO_CONFIG, 0)

    #Write to interrupt enable register
    bus.write_byte_data(Device_Address, INT_ENABLE, 1)

def read_raw_data(addr):
    #Accelero and Gyro value are 16-bit
        high = bus.read_byte_data(Device_Address, addr)
        low = bus.read_byte_data(Device_Address, addr+1)
        #concatenate higher and lower value
```

```
        value = ((high << 8) | low)


        #to get signed value from mpu6050
        if(value > 32768):
                value = value - 65536
        return value
bus = smbus.SMBus(1)    # or bus = smbus.SMBus(0) for older version boards
Device_Address = 0x68   # MPU6050 device address
#plt.ion()
MPU_Init()
```

#---------Machine Learning Implementation--------------#

```
train_file = glob.glob("trainingdata.csv") #current path
df_train_list = [] #initializing list
for fname in train_file:  #reading data form csv file/files and storing in variable
  df_train_list.append(pd.read_csv(fname,header=None,index_col=False))  #data of variable is appending in list
training_flat_list = pd.concat(df_train_list) #concatinating list into dataframe
training_flat_list = pd.DataFrame(training_flat_list) #concatinating data of the list in pandas dataframe
```

Gesture recognition system

```
train = training_flat_list


X_train = train.iloc[:,:-1] #selecting all rows and all columns except last column

y_train = train.iloc[:,-1] #selecting all rows with only last column


#Fit model to DtreeClassifier

dtree = DecisionTreeClassifier(random_state=0)

dtree.fit(X_train,y_train)

#----------------------------------------------------#

timeout = 5

timeout_start = time.time()

while True:

    i = 0

    list_values = []

    testing_flat_list = []

    df_test_list=[]

    while i<=9: #run the loop 10 times to gather values from sensors

            # Read the ADC channel, normalize and round it to required decimal places

            value0 = round((mcp.read_adc(0)),3)

            value1 = round((mcp.read_adc(1)),3)

            value2 = round((mcp.read_adc(2)),3)
```

```
        value3 = round((mcp.read_adc(3)),3)

        print("{0} {1} {2}".format(value1,value2,value3))

        MPU_values()

        print("Ax = ",round(A_x,0), "| Ay = ", round(A_y,0), "| Az = ", round(A_z,0))

        Axx = round(A_x,0)

        Ayy = round(A_y,0)

        Azz = round(A_z,0)

        #adding recorded values to the list

        list_values.append(Axx)

        list_values.append(Ayy)

        list_values.append(Azz)

        list_values.append(value1)

        list_values.append(value3)

        i = i+1

        time.sleep(0.30)

    print(i)



    testing_array=np.array(list_values)

    for oneD in testing_array:

        df_test_list.append(np.array(oneD).flatten())

    df_test_list=pd.DataFrame(list_values)
```

```
df_test_list=df_test_list.transpose() #makes rows of data

#Sorting data

test = df_test_list

X_test = test.iloc[:,:]

y_test = test.iloc[:,-1]

#Predict values by DtreeClassifier

dtree_predict = dtree.predict(X_test)

print(dtree_predict)

filename=np.array_str(dtree_predict)

strfilename=''.join(dtree_predict) #using the value predicted as string for filename to be played

os.system("mpg321 {}.mp3 &".format(strfilename)) #playing data from audio file

time.sleep(2) #wait for 2 seconds for loop to run again
```

OUTPUT:

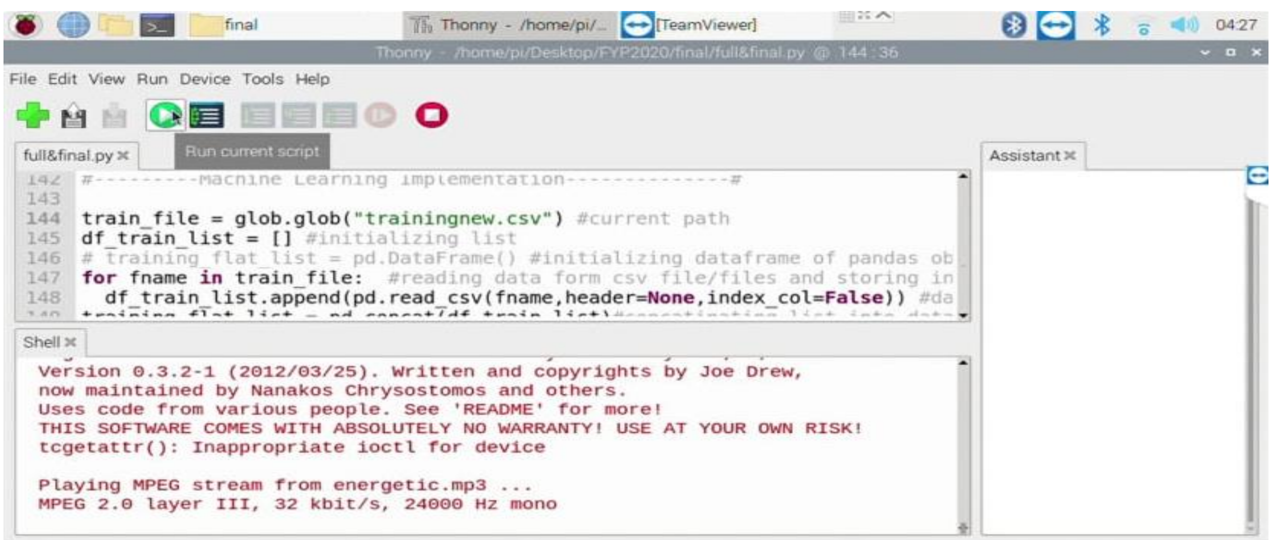


Figure 29