A MID-AIR WORD GESTURE AND VOICE ALERT SYSTEM FOR PHYSICALLY CHALLENGED USING MACHINE LEARNING

[1]Mr. Anand M, [2]Sona Ganesh G, [3]Sushmashree B S, [4]Namana Prasanna

[1]Assistent Professor, [2]Student, [3]Student, [4]Student
[1]Information Science and Engineering,
[1]GSSS Institute of engineering and technology for women, Mysuru, Karnataka

Abstract : Motion gestures provide a complimentary modality for general human–computer interaction. Motion gestures are meant to be simple so that a user can easily memorize and perform them. However, motion gestures themselves are not expressive enough to input text for motion-based control. We define "air-writing" as writing letters or words in a free space. Air-writing is especially useful for user interfaces that do not allow the user to type on a keyboard. To develop a Machine Learning based system that uses gestures to perform functions. There are a lot of gesture-based applications existing in today's world but every time we want to use a gesture-controlled application, we need to learn the predefined gestures and the functionality is also limited to the defaults provided with it. For this project, we aim to build a system which can be trained to recognize the gestures we make and perform the dedicated function we decide for it. We will be demonstrating this by training the system to recognize letters by the gestures we make in air.
To build this system, we will be using an Arduino board interfaced with an accelerometer. The device can be attached to the user's hand. The accelerometer will provide input to the microcontroller about the hand's coordinates. The algorithm will pick up this data and maintain a database to recognize each gesture differently. Once we train the system with the same gesture multiple times it will gather enough data to have an estimate of what the gesture should look like. The inputted letters are then converted to voice.

IndexTerms **– Machine Learning, Virtual keyboard, GPS, Training data**

I. INTRODUCTION

Motion gestures provide a complimentary modality for general human–computer interaction. Motion gestures are meant to be simple so that a user can easily memorize and perform them. However, motion gestures themselves are not expressive enough to input text for motion-based control. We define "air-writing" as writing letters or words in a free space. Air-writing is especially useful for user interfaces that do not allow the user to type on a keyboard. To develop a Machine Learning based system that uses gestures to perform functions. To build a system which can be trained to recognize the gestures we make and perform the dedicated function we decide for it. These applications demonstrate this by training the system to recognize letters by the gestures we make in air. An Arduino board interfaced with an accelerometer device can be attached to the user's hand. The accelerometer will provide input to the microcontroller about the hand's coordinates. The algorithm will pick up this data and maintain a database to recognize each gesture differently. After training the system with the same gesture multiple times it will gather enough data to have an estimate of what the gesture should look like. The inputted gesture is converted into voice and after a full stop gesture it will send that gestured message to nurse's mail. In order to alert a nurse, the default message saying "there is a new message" is sent to the respective nurse via SMS Using GSM Model.

II. Literature Review

A literature survey represents a study of previously existing material on the topic of the report. This includes existing theories about the topic which are accepted universally. Books written on the topic, both generic and specific. Research done in the field usually in the order of oldest to latest. Challenges being faced and ongoing work, if available.

In [1] **"Hand gesture keyboard using machine learning**", Sam sunny, Sindhu, ASIET 2020.
The SVM classifier will classify the information content data according to the trained data set and predicts he corresponding character. Different hand gesture methods have been reviewed and analyzed.

In [2] **"Visual Gesture Recognition for Text writing in Air"**, Saria Beg, M. Fahad Khan and Faisal Baig, IEEE 2018

This device is a gesture tracking device which uses machine learning's support vector machine model to convert accelerometer data to a sequence of alphabets. To build a device using an Arduino pro micro that translates gestures into words wirelessly.

In [3] **"Hand Gesture Recognition with Depth Images", J. P.** Wachs, M. Kölsch, H. Stern, and Y. Edan, IEEE 2017

This section describes the types of gestures and postures classified by papers reviewed, and the classification methods used. The system provides students a common and user-friendly platform for participating in different events.

In [4] **Implementation of Gesture Keyboard Using Machine Learning"**, Bharathi M, Bhavani R, Keerthana Priya B, R. Paavaikarasi M.E, IJIRSET 2019
This project proposes a handy keyboard that is used to type in character by recognizing hand gesture keyboard are currently the most universally accepted computer input device. The system uses a virtual keyboard that recognizes the letters based on the gesture provided by the users.

In [5] **"A gesture air writing tracking system"**, Pengcheng wang, Junyang lin, IEEE 2020.
The alpha beta filter and RTS smoothing algorithm. SIMO radar system was used to solve the problem of drawing characters in a planar area.

In [6] **"Gesture Recognition Using Laser-Based Tracking System"** Stephane Perrin, Alvaro Cassinelli and Masatoshi Ishikawa, IEEE 2020.
A system aimed at tracking a finger without the use of any special device for gestural character input.

In [7] **"A Study of Human Preferences in Usage of Gesture Types for HCI"**, Roland Aigner Daniel Wigdor Hrvoje Benko, MSR-TR-2020
In this paper we present the results of a study of human preferences in using mid-air gestures for directing other humans.

In [8] **"Implementation of Gesture based voice and Language Translator"**, L Anusha, Y Usha Devi ASIET 2020
K nearest neighbors' algorithm is used to classify new cases after sorting all available cases. The hand gesture symbol may vary.

In [9] **"Finger Recognition and Gesture based virtual Keyboard",** Chinnam Datta Sai Nikhil, Chukka Uma Someswra Rao, T Anandhi, IEEE 2020
The Haar cascade calculation is utilized to prepare information. Machine learning algorithms are organized by python programming language. Arduino pro micro and Bluetooth hc-06 is used. The vision-based hand motion may not be accurate in dark places.

In [10] **"Individual Soldier Gesture Intelligent Recognition System",** Depeng Zhu, Weida Zhan, Ranran Wei, Ziqiang Hao, IEEE 2019
The entire recognition system algorithm includes data normalization and endpoint detection DTW algorithm computational similarity and gesture recognition. When gesture is complicated result is more accurate

III. System Design

Requirements

Functional Requirements

Functional requirement defines a function of a system or its component, where a function is described as a specification of behavior between outputs and inputs. Functional requirements may involve calculations, technical details, data manipulation, processing and other specific functionality that define what a system is supposed to accomplish. The functional requirements of "A mid-air word gesture and voice alert system for physically challenged using machine learning" are:

Us e r/Community

User can give the inputs by waving their hands in air. User can write the letters in air and the written letters are displayed on the monitor. The displayed message will be sent to the nurse's mail. By sending a default SMS it keeps nurse alert.
.

Non-Functional Requirements

Non-functional requirement is a requirement that specifies criteria that can be used to judge the operation of a system, rather than specific behaviors. They are contrasted with functional requirements that define specific behavior or functions. The main Non-functional requirements of "A mid-air word gesture and voice alert system for physically challenged using machine learning" are:

Performance and Response time: The system should have high performance rate when executing user's input and should be able to provide feedback or response within a short time span.

Availability: This system should always be available for access at 24 hours, 7 days a week also in the occurrence of any major system malfunctioning□

Error handling: Error should be considerably minimized and an appropriate error message that guides the user to recover from an error should be provided.□

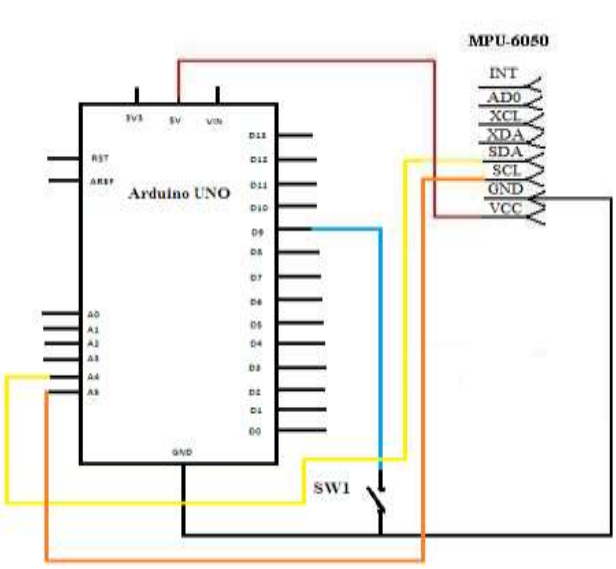
Usability: The developing application is going to be used by the user which helps to communicate between blind people.

Efficiency: It reduces the complications when information has several functionalities, thus increases the efficiency. Provide feedback or response within a short time span.

IV. System analysis and design

Circuit Diagram

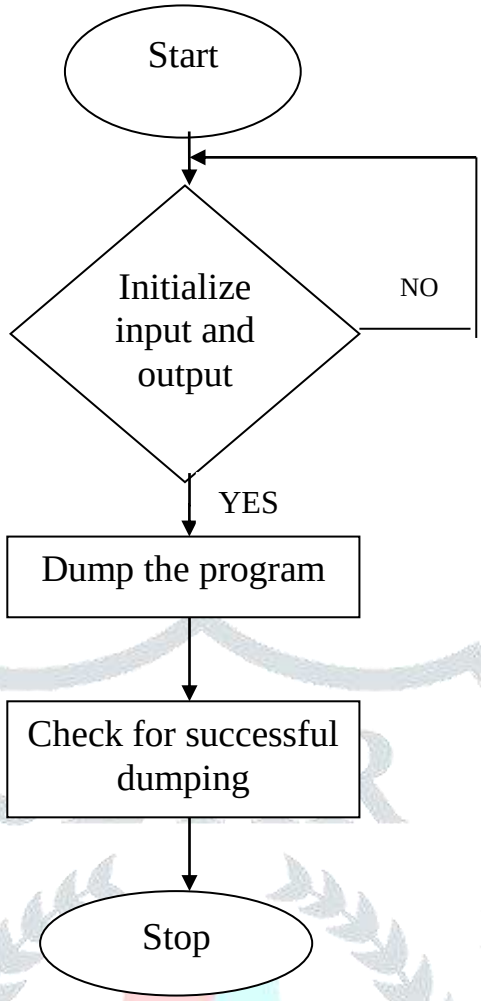
Architecture



Gesture Keyboard is a library used to convert accelerometer data to a sequence of characters and sentences. Gesture recognition typically involves a limited vocabulary set. It is relatively easy to collect sufficient data of each gesture and straight forward to model each gesture directly from its own recordings. However, the vocabulary of air-writing can easily be thousands of words, and it is difficult to collect enough data for every word in the vocabulary. In order to get the accelerometer data, build a module using an Arduino, a MPU-6050 as accelerometer, the module starts to send accelerometer data to the PC. InPC the library is written in Python and uses Scikit-learn "Support Vector Machine algorithm" to classify the signals into letters.

- **Arduino UNO**

Arduino is an open-source platform that offers both hardware and software platform functionalities for interactive projects that have the control and sensing capability providing an extension to connect to other such controls. Arduino board uses a micro-controller board along with 14 digital input and output configuration pins and 6 analog inputs are used. The output of the accelerometer sensor is given to three analog inputs of the Arduino board. The inbuilt ADC of the board converts the values to digital and sends the data to the Wireless or Bluetooth module through Serial Port. Arduino can be programmed through its IDE software USB cable. It can be powered by a USB cable. In our project this plays an important role which provides a platform to run a program.



- **Accelerometer module-MPU 6050**

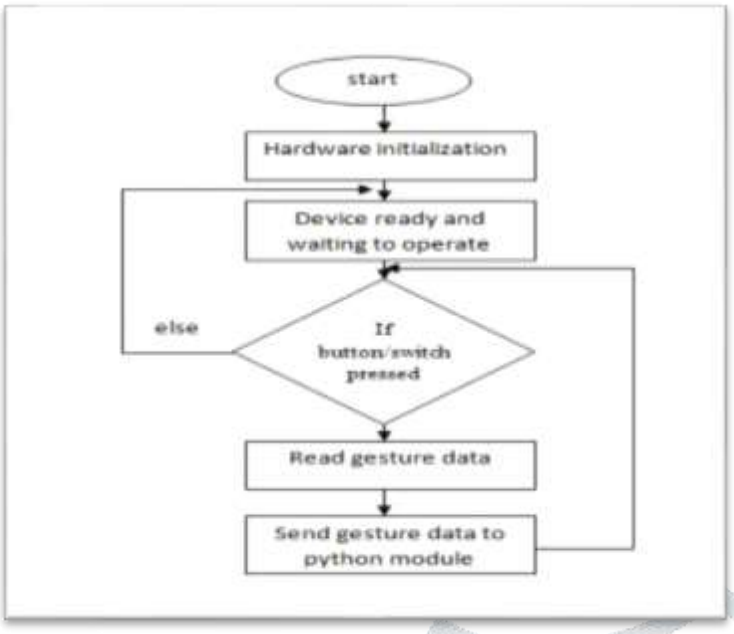
An accelerometer is a micro electromechanical sensor (MEMS) that measures acceleration. It is also helpful in detecting tilt or orientation with respect to the earth of a device it is attached to. This application uses its property to detect the tilt of the hand to move the mouse pointer. The continuous data streams are divided into individual gesture according to the button pressing label. An accelerometer can sense keenly the acceleration data of three spatially placed orthogonal axes in each gesture in a given sampling frequency. A gesture can be denoted as:

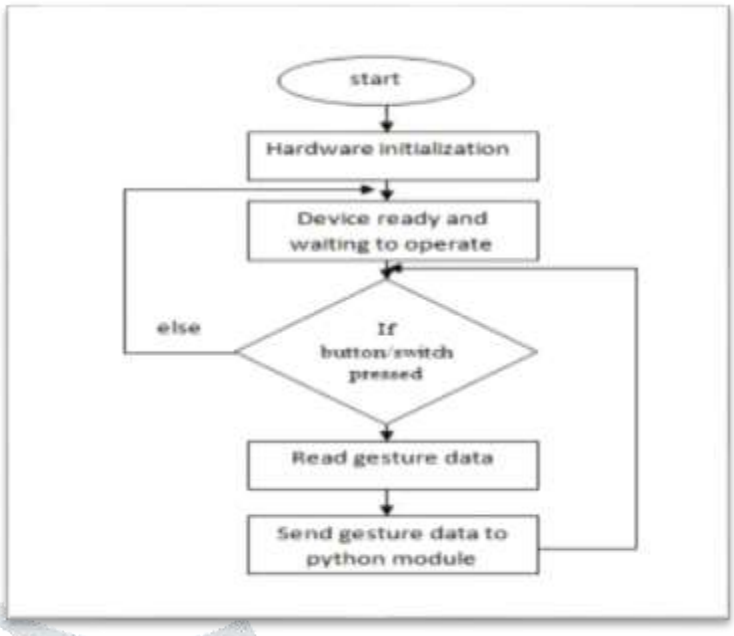
G = (Ax, Ay, Az)

Here, **A**x, **ay**, Az is the acceleration vector of an axis and *L* is the length of the temporal sequence. To describe the whole gesture but distinguish the periods from each other, we divide a gesture into *N*+1 segment identical in length, and then every two adjunct segments make up a frame. For the feature extraction of each gesture, firstly, it is divided into *N* frames. The feature vector is eventually put into a classifier in order to train a classification model or retrieve a recognized gesture type. More specifically, a gesture can be represented as: Intuitively, more frames a gesture is broken up into, more details we know about the gesture. However, it may lead to the over-fitting problem if the frame number *N*ecomes large. It will also increase the dimension of the feature space, which increases computational complexity. We will conduct an experiment to determine the optimal frame number *N*later.

Algorithm for gesture recognition with accelerometer

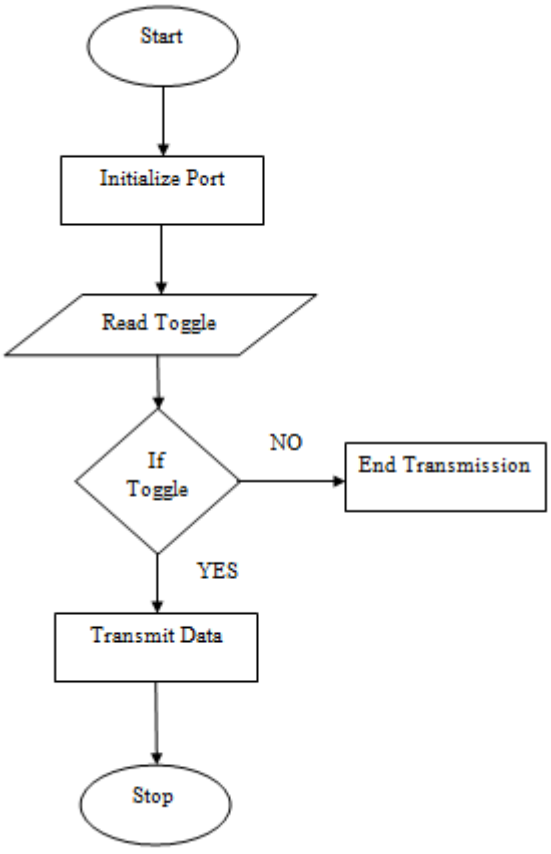
Step1: Start the Transmitter by providing power supply to the module.
Step2: Once the supply is given, Arduino uno starts communication with MPU-6050 by sending the slave address.
Step3: After the slave address has been sent, some registers of MPU 6050 have to be programmed.
Step4: Now, the registers of Accelerometer (X, Y, Z-axis registers) are read.
Step5: These values are sent to the Processor through USB.





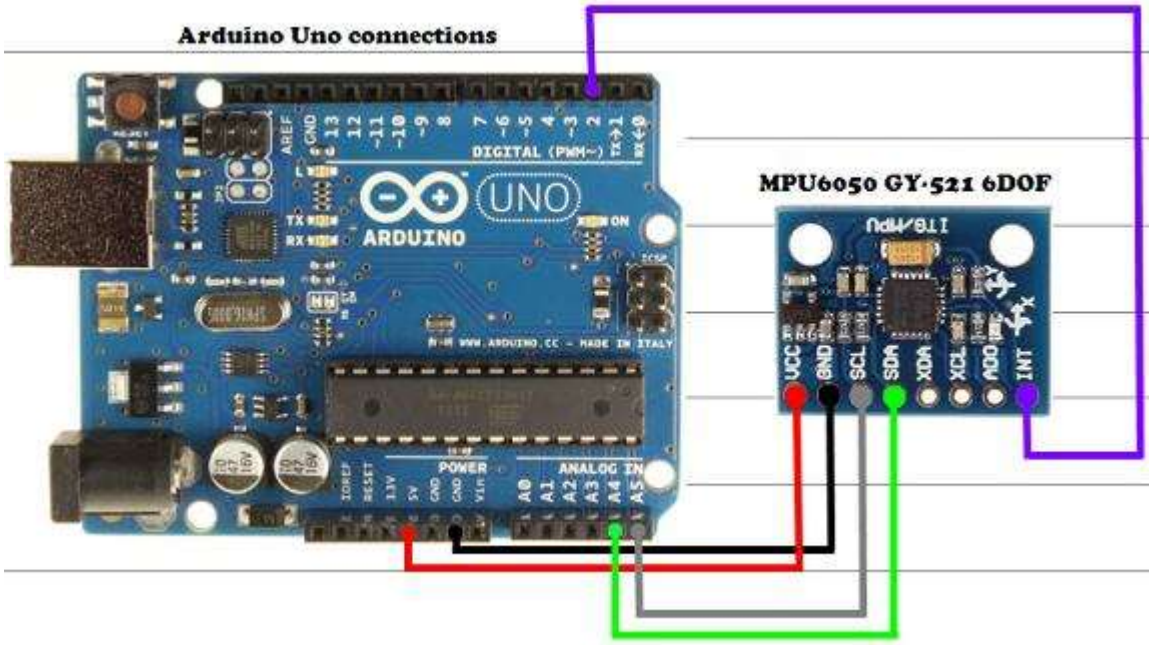
- **Switch**

The switch is used to control device operation for toggling activation of the sensor. When someone wants to signal and record gestures, we first press the button, and hence the module starts to send accelerometer data to the computer connected. At the moment when the button is released, the transmission stops immediately.



Arduino uno and Accelerometer Connection

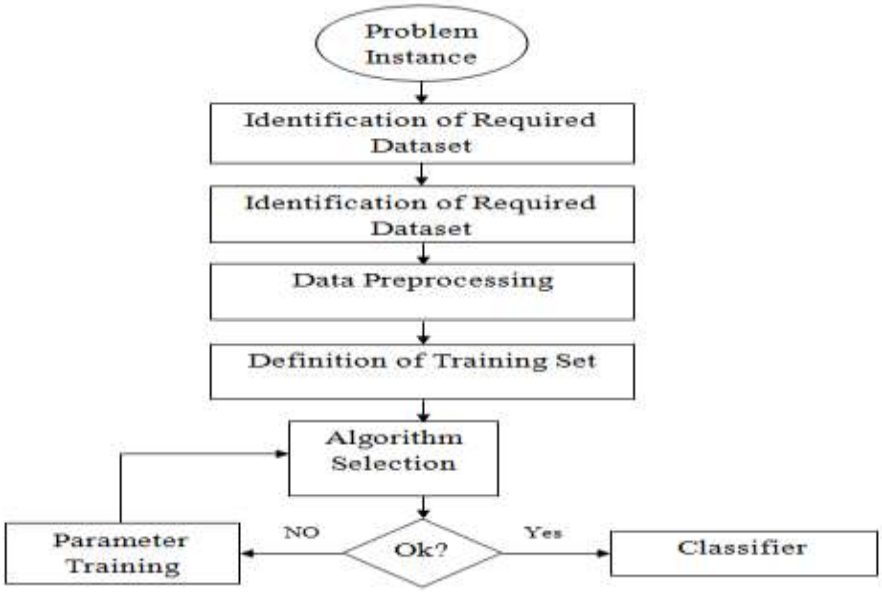
The circuit connection between MPU-6050 and Arduino Uno is connected as shown in the figure below. The operating range of the MPU-6050 is 2.375-3.4V is as stated in the datasheet. Thus, we will directly use the ready 5V voltage source that the Arduino is able to provide. The Ground (GND) and AD0 are connected to both the ground pin on Arduino. The AD0 pin functions to defining the default I2C address and thus we need to ground it for the default evaluation purposes. SDA and SCL on the sensor module are then connected to the dedicated I2C communication on the Arduino which is the A4 and A5 analogue pin. As for obtaining the raw sensor values, the MPU-6050 outputs a digital signal from the ADC for each of the sensors. The sensors values are represented by 16-bit 2's complement format. Take example for a ±2g full scale range accelerometer the bits that are able to represent the scale are 2^{16}-1 which is a total of 65535 bits. The coding of Arduino to use MPU-6050 is as referred from open source. The main processing of all sensor data happens in the microcontroller which will be the main components of the project.



The initial result as from the Arduino sketch. The sensor is place on a table with no movements to test the validity of the sensors. The baud rate is set to 9600 which can be monitored on Arduino.

Creation of gesture recognition software

Scikit-learn is a programming library in machine learning for the programming in Python. It highlights distinctive characterization, backslide and clustering calculations counting reinforce vector machines, self-assertive forests, slant boosting and is aiming to interoperate with the Python numerical and coherent libraries. Gesture Recognition or any other machine learning tasks that uses supervised learning goes through a process of stages before the final classifier produces an efficient output. All the stages are indecent of each other and play a unique role in contributing to the accuracy of classification.



- **Required dataset**

The first step is to fabricate the data attributes that are crucial in determining the goal of the output. All values must have a constraint or an upper and lower bound limits to initiate remission of boundary limits. The dataset is then made available to be identified in terms of different parameters that can further extracted. The set is then made available to be identified in terms of different parameters that can further extracted.

- **Data Pre-Processing**

Depending on the type of problem, the researchers have concluded that the pre-processing of data may vary on disparate elements of attributes. For example, instance selection is not only used to handle noise but to cope with the infeasibility of learning efficiently from heavy datasets. Instance selection in such datasets is an optimization problem that aims to maintain the mining quality while attempts to degrade the sample size. It reduces data and enables a data classification algorithm to function and work effectively with very large or heavy datasets. There is a variety of procedures for sampling instances from a large dataset.

- **Training Set**

Once a clear input variable is made available on a global scope with logistic and constraint satisfying training inputs can be fetched at every input cycle, the similar categories of elements are classified together as a distinct entity. This is called as the

training set and serves as the basis of classification space to make experiments and observe trends. The set consists of attributed key-value pairs of abundance different instances that cumulatively make the set.

- **Algorithm Selection**

In this step, the documents are split into training and testing documents, the training documents are used to train the system (i.e., learn the system) to recognize different patterns of categories, the testing documents are used to evaluate the system, the process of categorization depends on the algorithm used. The choice of which specific learning algorithm we should use is a critical step. Once we complete the preliminary testing and if achieved as to be satisfactory, the classifier that maps the unlabeled instances into classes is available for the routine to use. The classifier's evaluation is most often based on prediction accuracy (the percentage of correct prediction divided by the total number of predictions). There are three techniques used to calculate a classifier's accuracy which is essential One technique is to split all the training set by using two-thirds of it for training and the other third of it for estimating their performance. The average of the error rate of each subset is roughly an estimate of the error rate of the classifier.

Classification Algorithms for Gesture Recognition

In this section, we first elaborate on our data set and target values and then begin to evaluate the performances of algorithm for training of gestures.

- **Gesture Dataset**

To improve precision, we will create a new data set and evaluate their trends. Each time an accelerometer position is moved or a different accelerometer is used, the device becomes less accurate. So, it is important to keep adding the training data to learning set in added time sequences. The algorithm can feature a maximum of 40different gestures, such that each gesture is associated with a single letter of the English alphabet. The first step is to begin recording a new batch of data inputs for gestures by pressing the online button, the library records the data from movements of accelerometer using the following markup Target: Argument tells module it wants to record new gesture samples

a: It is the one that characterizes the unique gesture of length 1

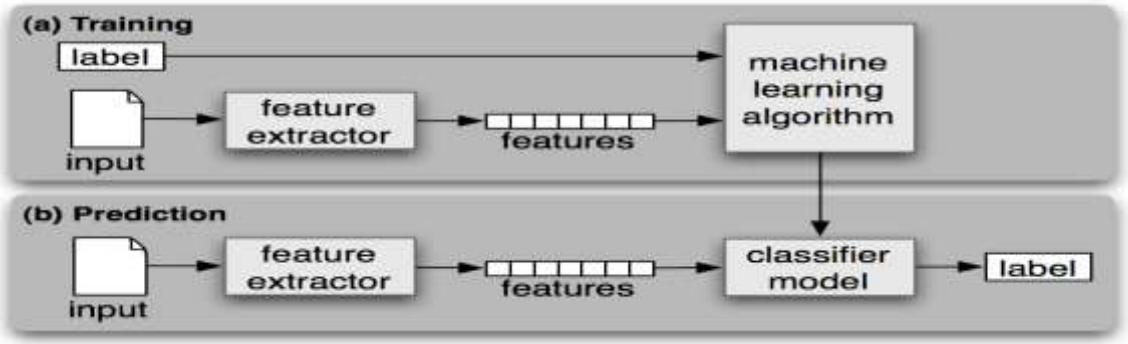
0: Batch Number; It must be different each time we begin to register a new batch in order to avoid overriding

Port: The port represents the serial port connection to the Micro-controller/Arduino

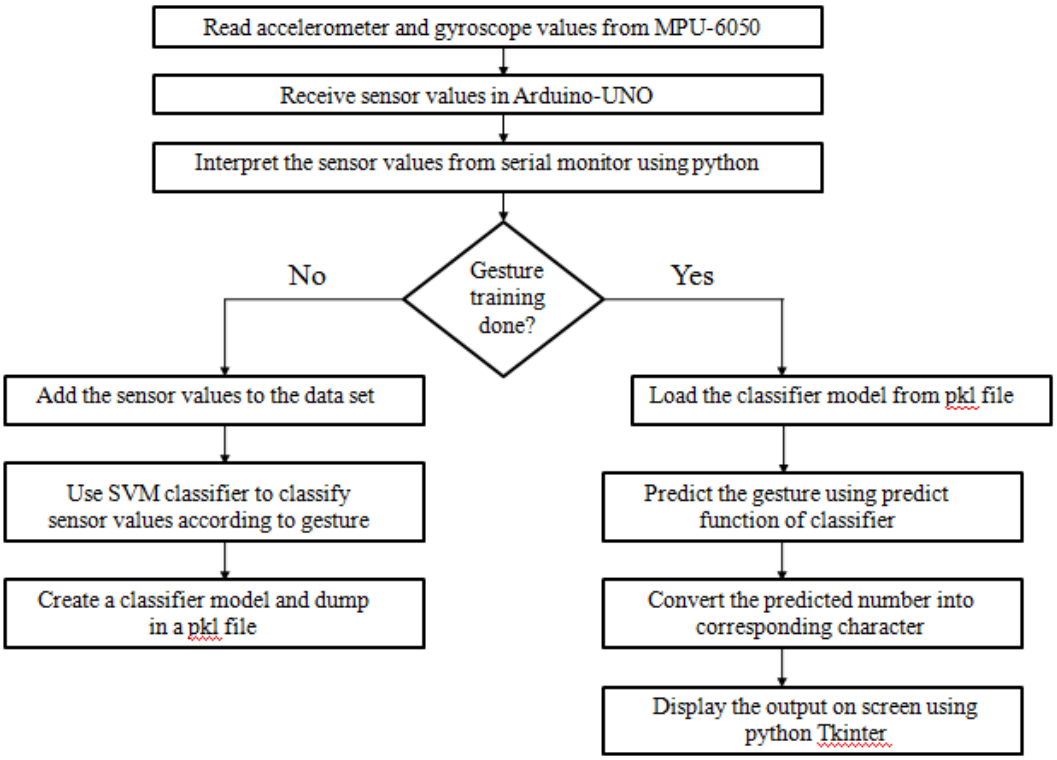
Ideally, each sample must be recorded and saved as a different file in the data folder and is later evaluated. Once the dataset is ready a suitable model is selected to train these inputs. The choice of a perfect algorithms can be a daunting and challenging task; hence we aim to compare 2 of the leading classification techniques on our dataset and evaluate the best model.

- **Classification Algorithms**

Classification algorithms use a set of variables such as "feature variable" or "independent variable" or "predictor variable" for making predictions and "class variable or dependent variable" for holding outcome of prediction. These algorithms are driven by "class variable" for the given data that can be either "categorical" or "continuous". "Classification" techniques are used in cases where the class label is categorical and "Regression" techniques are used in cases where the class label is continuous. The classifier selection is based on prediction accuracy of algorithm which is measured by a parameter called "Confusion matrix". Confusion matrix is used to assess percentage of test data that has been correctly labeled by the classifier.



Gesture Keyboard Work Flow



Working of the system initially the machine won't know anything. The machine has to read the gestures which are made by the hardware. The gestures are read through the MPU 6050 sensor attached to the Arduino. The sensor values provide the continuous location of the hand used to make the gesture. A combination of the sensor values will be used to estimate the path taken by the hand while making the gesture. The sensor values will be used to form a dataset. The sensor values will be mapped to the corresponding character which was specified by the user while training the machine. This mapping will be used to predict the character from the gesture made in thin air. For example, making an 'a' in the air with our hand can be mapped to print the letter 'a' on the computer screen. Initially the user has to feed the preferred gesture to the program and after doing that the user has to train the machine by testing it continuously. This way the machine comes to know the gestures required for a particular operation.

V. Testing

SYSTEM TESTING

Before implementing the new system operation or functionality, a test run of the system is done for removing the bugs. It is an important phase of any successful system. After codifying the whole programs of the system, a test plan should be developed and run on a set of test data. The output of the test run should match the expected results. System testing is considered as a part of implementation process.

Testing involves a variety of things, but most importantly it measures the quality of the software we are developing. This view will find out the defects in the software waiting to be discovered and this view is rarely disproved or even disputed. Several factors contribute to

the importance of making testing a high priority of any software development effort. These include

- Development cost of the program is reduced.
- Assuring that the application behaves exactly as we explain to user for the majority of programs, unpredictability is the least desirable consequences of using an application.

The total cost of ownership is reduced. By providing software that looks and behaves as shown in the documentation, fewer hours of training and less support from product experts is sufficient for customers.

Developing customer loyalty.

System testing is the process of checking the objective and requirements. It is a very critical element of software quality assurance (SQA) and represents the review of specification, design and coding. Testing represents an interesting anomaly i.e deviation from what is actually required for the software. Thus, a series of tests are performed for the proposed system before the system is ready for customer acceptance. Some of the various test cases used to test the system are as follows

- The test cases are written for testing against requirements of the unit being tested.
- If the unit modifies the database, test for the integrity of the database after the operation.
- Test cases for the path or branch coverage should be done.
- Test cases for data flow coverage should be done.

- Test cases based on the experience such as testing for boundary conditions minimum, maximum and off by one boundary.

TEST CLASSIFICATION

UNIT TESTING

Unit testing, a testing technique in which individual modules are tested to determine if there are any issues and/or bugs by the developer. It is concerned with functional correctness of the individual modules. The main aim is to isolate each unit of the system to identify, analyze and fix the defects.

INTEGRATION TESTING

Upon completion of unit testing, the units or modules are to be integrated which gives raise to integration testing. The purpose of integration testing is to verify the functional, performance, and reliability between the modules that are integrated.

FUNCTIONAL TEST

Functional Testing is a testing technique that is used to test the features/functionality of the system or Software, should cover all the scenarios including failure paths and boundary cases.

SYSTEM TEST

System Testing is usually carried out by a team that is independent of the development team in order to measure the quality of the system unbiased. It includes both functional and Non-Functional testing.

WHITE BOX TESTING

White box testing is a testing technique that examines the program structure and derives test data from the program logic/code. The other names of glass box testing are clear box testing,open box testing, logic driven testing or path driven testing or structural testing.

BLACK BOX TESTING

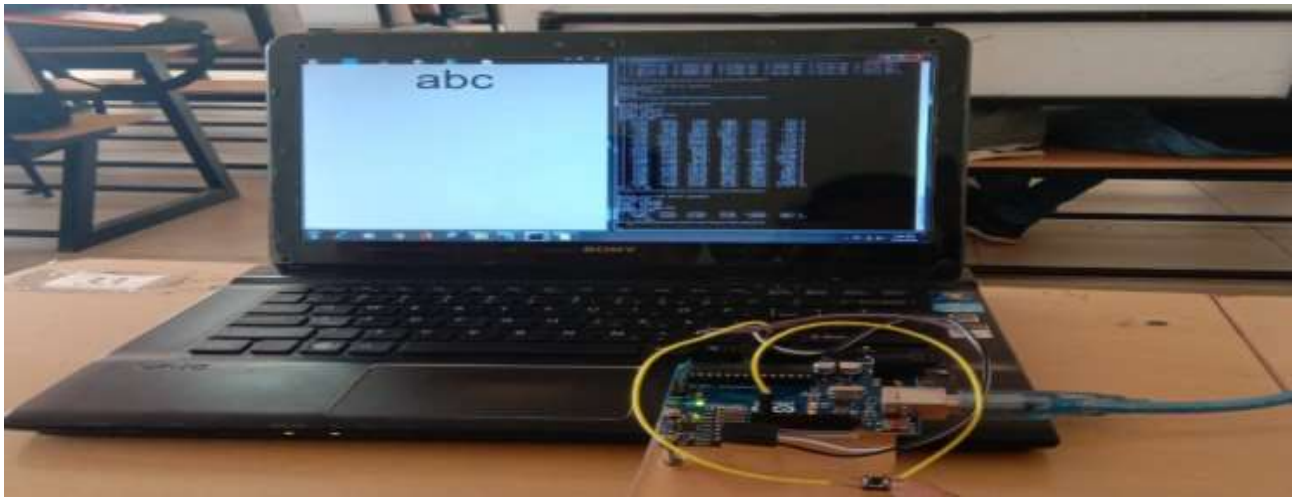
Black-box testing is a method of software testing that examines the functionality of an application based on the specifications. It is also known as Specifications based testing. Independent Testing Team usually performs this type of testing during the software testing life cycle.

ACCEPTANCE TESTING

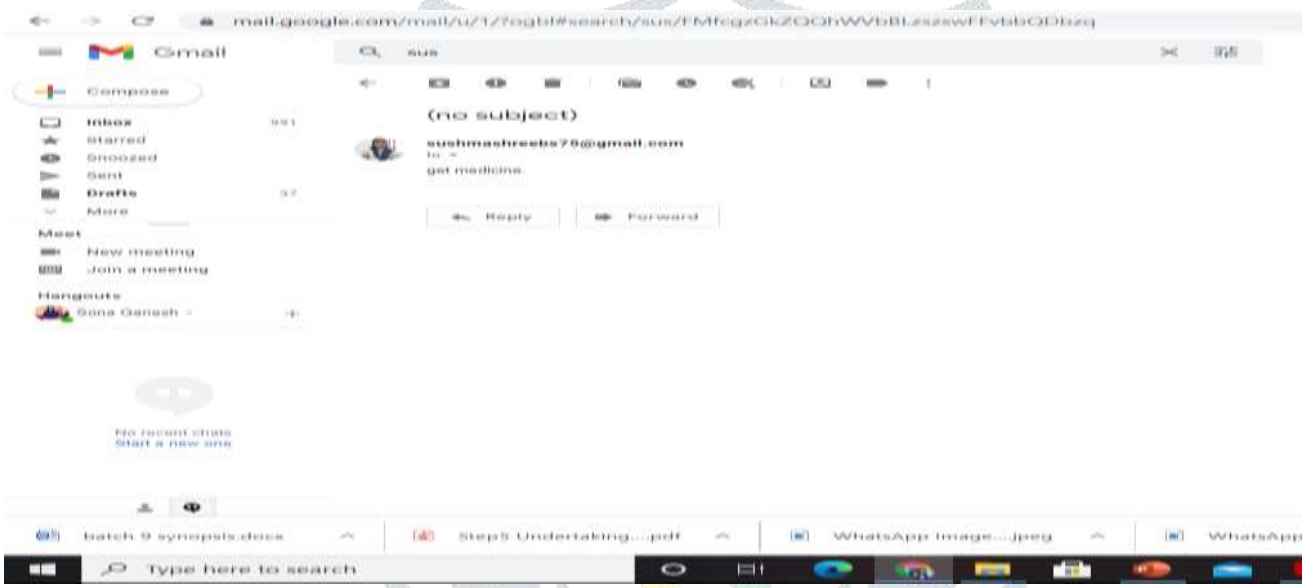
Acceptance testing, a testing technique performed to determine whether or not the software system has met the requirement specifications. The main purpose of this test is to evaluate the system's compliance with the business requirements and verify if it is having met the required criteria for delivery to end users

VI. Results and Discussion

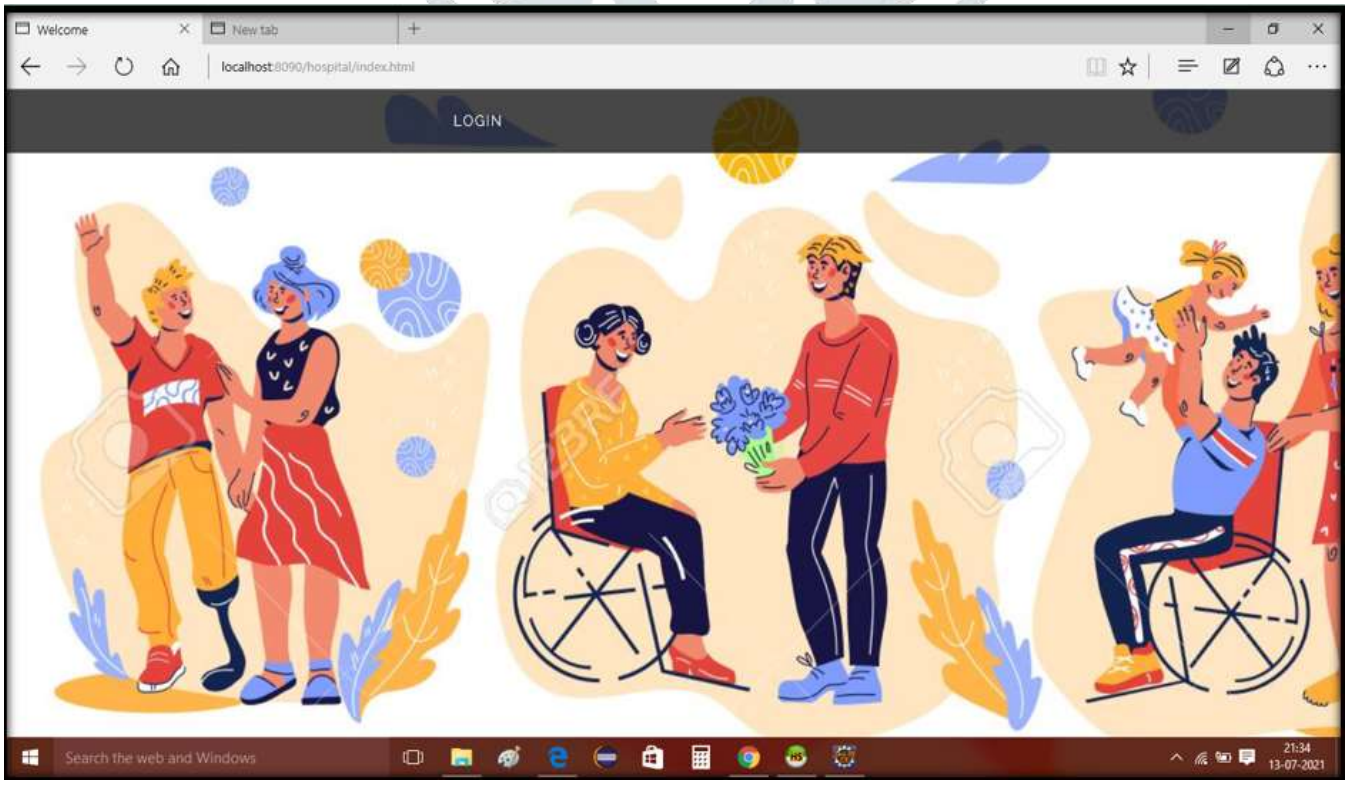
The project is being tested and is working properly. It has the ability to identify the letters drawn through a gesture in air with the assistance of a input device. The strategy proposed here effectively made an acknowledgement framework, that can perceive in which motion is performed by the user and precisely play out the usefulness related with it. Motion letters are handled similar to motion gestures. The former is proven to be sufficient for word-based word recognition. In this project we have used very less components, so it is cost effective and less components. Unfortunately, a set of gestures which can be recognize in this way is very limited. Nevertheless, they are useful for extension of communication channel between a human and a machine. In this context the obtained results are preliminary and show that recognition of some class of gestures is possible by using very simple range finders.



In this project we have used very less components, so it is cost effective and less components. Unfortunately, a set of gestures which can be recognize in this way is very limited. Nevertheless, they are useful for extension of communication channel between a human and a machine. In this context the obtained results are preliminary and show that recognition of some class of gestures is possible by using very simple range finders.



The gestured message by the user will be sent to nurse's mail by adding punctation at the end.



The login page for the admin and nurse

VI. CONCLUSION AND FUTURE SCOPE

Conclusion

This project portrays a framework that have the ability to identify the character drawn through a gesture in air with the assistance of a input device. The strategy proposed here effectively made an acknowledgment framework, that can perceive which motion is performed by the user and precisely play out the usefulness related with it. The project serves as an efficient recognition of sentence/words/alphabets/numbers by just waving their hands in air with specific gesture sign. Helps the blind people, who communicate by sign language.

Future Research Development

The present framework gives best outcomes in a plain foundation and henceforth put certain imperatives on the user for effective working. The future work will incorporate usage of extra signals which will empower the client to perform more capacities easily. Moreover, foundation subtraction calculation can be utilized for a more compelling execution. The proposed framework utilizes just the correct hand to perform signals. Henceforth, upgrade of the procedure proposed, is conceivable utilizing the two hands for performing diverse PC activities. Examinations should be done on a bigger scale with the goal that outcomes can be more exact.

VII. References

[1] P. Cooper (ed.), The Open GIS Abstract Specification - Topic 2: Spatial referencing by coordinates, version 4.0. Open Geospatial Consortium, Wayland, MA, USA (2010).

[2] M. Fowler, “Inversion of control containers and the dependency injection pattern” (2004).http://www.martinfowler.com/articles/injection.html

[3] S. Winter and A. Frank, “Topology in Raster and Vector Representation”, Geo Informatica, Springer-Verlag, Berlin,Heidelberg, Germany, 4 (1), pp. 35–65 (2000).

[4] A. Cary, Z. Sun, V. Hristidis and N. Rishe, “Experiences on Processing Spatial Data with Map Reduce”, Lecture Notes inComputer Science, Springer-Verlag, Berlin, Heidelberg, Germany,5566, pp. 302–319 (2009).

[5] N. Golpayegani and M. Halem, “Cloud Computing for Satellite Data Processing on High End Compute Clusters”, IEEE International Conference on Cloud Computing, 2009.

[6] R. Giachetta, “AEGIS – A state-of-the-art spatio-temporal framework for education and research”, OSGeo Journal, OSGeo Foundation, Beaverton, Oregon, USA, 13, pp. 68–77, (2014).

[7] L. Li, Y. Chen, X. Yu, R. Liu and C. Huang, "Sub-pixel flood inundation mapping from multispectral remotely sensed images based on discrete particle swarm optimization," *ISPRS Journal ofPhotogrammetry and Remote Sensing, vol. 101, pp. 10-21, Mar.*2015.

[8] S. Martinis, A. Twele and S. Voigt, "Towards operational near real-time flood detection using a split-based automatic thresholding procedure on high resolution TerraSAR-X data,“ *Natural Hazards and Earth System Science, vol. 9, no. 2, pp. 303*- 314, Mar. 2009.

[9] K. E. Joyce, S. E. Belliss, S. V. Samsonov, S. J. McNeill and P. J. Glassey, "A review of the status of satellite remote sensing and image processing techniques for mapping natural hazards and disasters," *Progress in Physical Geography, vol. 33, no. 2, pp.* 183-207, Apr. 2009.

[10] K. Zhang, L. Zhang, M.-H. Yang and D. Zhang, "Fast tracking via spatio-temporal context learning," *arXiv preprint arXiv:1311.1939, Nov. 2013.*

[11] J. Kittler and J. Illingworth, "Minimum error thresholding,“ *Pattern recognition, vol. 19, no. 1, pp. 41-47, 1986.*

[12] G. Bai, "China environmental and disaster monitoring and forecasting smallsatellite— HJ-1A/B," *Aerospace China, vol. 5,* pp. 10-15, 2009.

[13] “Difference of Alzheimer’s disease sub-groups using two features from intensity size zone matrix.” Seung-Hak Lee, Hyunjin Park.
“Multiscale Dispersion Entropy for the Regional Analysis of Resting-state Magnetoencephalogram complexity in Alzheimer’s Disease.” Hamed Azami1, Eli Kinney-lang.