REAL -TIME COMMUNICATION SYSTEM POWERED BY AI FOR SPECIALLY ABLED

SUBMITTED BY

M.SURESHPANDI

K KALAISELVI

R.DEEPAKRAJA

S.KRIPALSINGH CHIRONJE

S.SIBIN

R.DURAIKANI

S.MANIKANDAN

DEPARTMENT OF COMPUTER SCIENCE

AND ENGINEERING

UDAYA SCHOOL OF ENGINEERING

VELLAMODI

2019-2022

TEAM ID:PNT2022TMID52071

CONTENTS

INTRODUCTION

1.1 PROJECT OVERVIEW

The project aims to develop a system that converts the sign language into a human hearing voice in the desired language to convey a message to normal people, as well as convert speech into understandable sign language for the deaf and dumb. We are making use of a convolution neural network to create a model that is trained on different hand gestures. An app is built which uses this model. This app enables deaf and dumb people to convey their information using signs which get converted to human-understandable language and speech is given as output.

1.2 PURPOSE

The technology is developing day by day but no significant developments are undertaken for the betterment of these people. Communications between deaf-mute and a normal person has always been a challenging task. It is very difficult for mute people to convey their message to normal people. Since normal people are not trained on hand sign language. In emergency times conveying their message is very difficult. The human hand has remained a popular choice to convey information in situations where other forms like speech cannot be used. Voice Conversion System with Hand Gesture Recognition and translation will be very useful to have a proper conversation between a normal person and an impaired person in any language.

2.literature survey

2.1 EXISTING PROBLEM

Communication is a social process of exchanging information from one entity to another in verbal and non-verbal form. It defines our existence and it is an important instrument that connects people together. It comes naturally as a raw skill embedded in most people at birth and we acquired the ways of communication through cognitive learning. Communication is the basis, which drives the process of development in all the fields (Manohar, 2008) and it is the very core of our civilisation. The ability to communicate allows us to express emotion, feelings, convey our thoughts and ideas as well as to relate our experiences. It plays an important role in the dis- semination of information and sharing of knowledge especially in the academic arena. Research has found that human started to learn how to communicate with each other since they are born not only through spoken and written languages but also body gesture, posture, facial expression and eye contacts (Busso, et al., 2004; Cohen, Grag & Huang, 2000). Communication skill might come as a natural ability in majority of people. However, there are some people inflicted with some form of physical defects which affect their ability to communicate. One of the more severe disabilities is known as “cerebral palsy”, a congenital disorder at birth which causes abnormality in their Motor system. It affects their muscle movement and coordination, learning and Speech abilities. Their malfunctioned motor system causes an uncontrollable and involuntary movement. They are unable to control their oralfacial muscles, thus affects their ability to perform facial expression appropriately.

From the limitation of the existing tools reviewed (Novita, 2006; Macsolvers, 2009; Standup, 2006; Universiteit van Amsterdam, 2008; Crestwood, 2009; Sci-enceDaily, 2008), there is still a pressing need for more effective and efficient tools to alleviate this problem. One the possible methods are to implement a facial expres-sion recognition system to predict or determine the emotional state of a disabled person through his expression projected on his face. biometrics information system can be employed as a means to detect and classify the physiological aspect of a person in real time. Franco and Treves (2001) further support the notion that facial expression can be used for human computer interaction and usability enhancement.

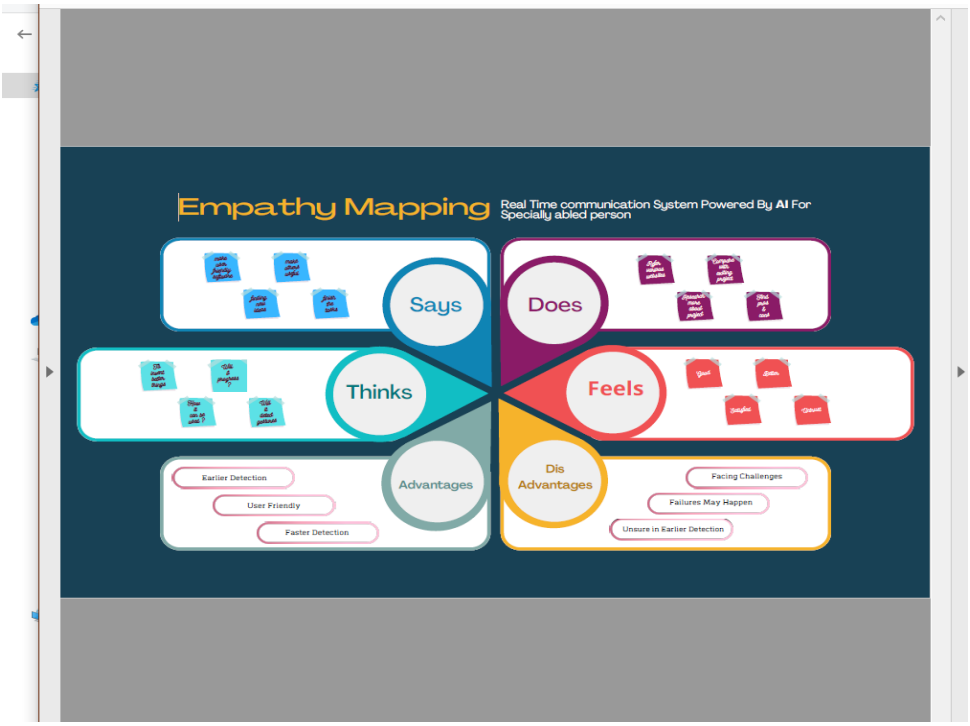
2.2 problem statement definition

Based on the problem statements deliberated above, we propose an improved real-time communication system using machine learning and computer vision. The aim is to create a communication channel between the specially abled and the society, so they can express there

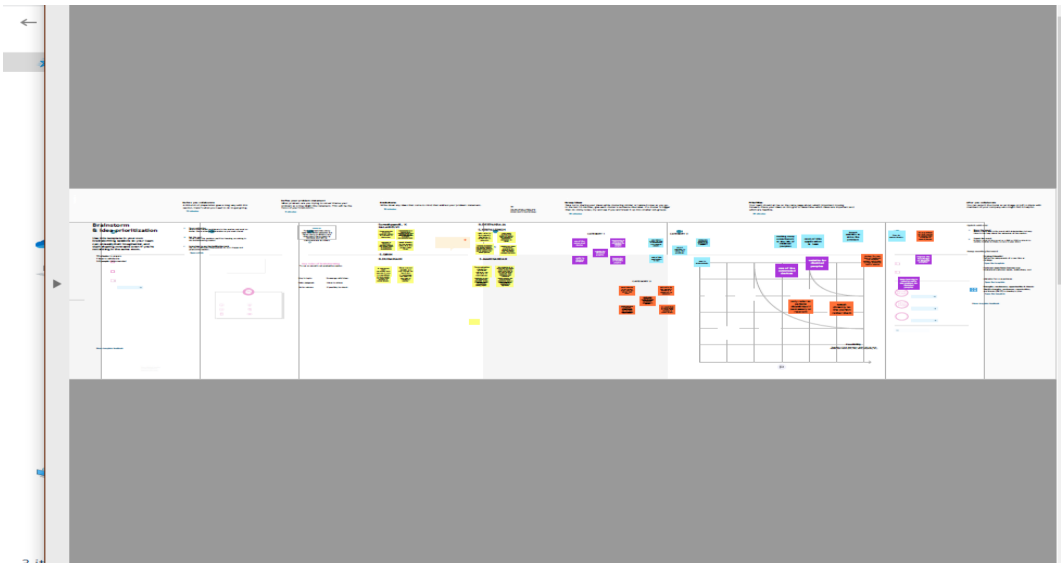
feelings, thoughts and understand other people’s feelings and thoughts through real time communication and facial expressions.

3.IDEATION & PROPOSED SOLUTION

3.1 EMPATHY MAP CANVAS



3.2 IDEATION AND BRAINSTORMING

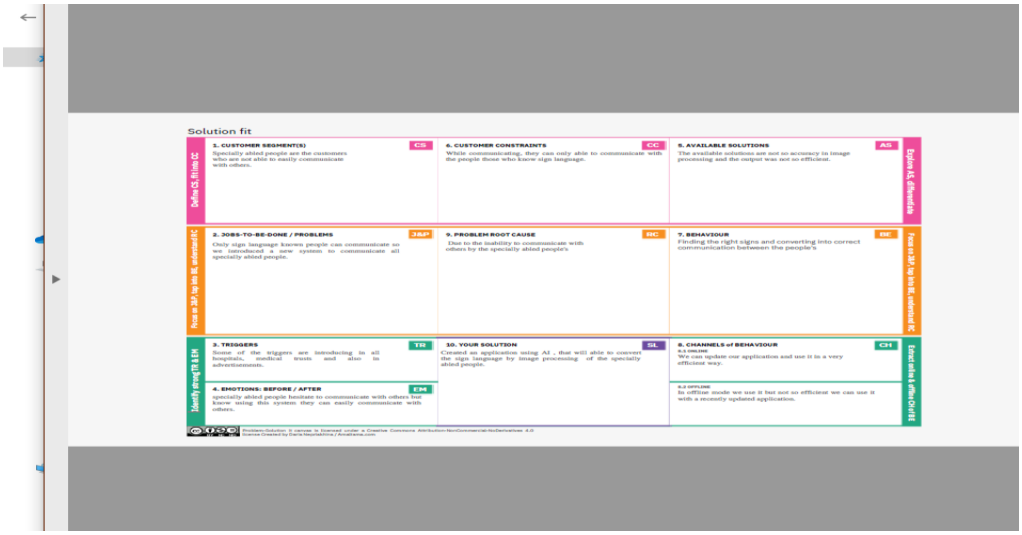


3.3 PROPOSED SOLUTION

S.No.	PARAMETER	DESCRIPTION
1	Problem Statement (Problem to be solved)	Differently able like dump and mute people can communicate only through the sign language,normal people those who do not know the sign language feels difficult to communicate with them.
2	Idea / Solution description	To overcome this problem we have an idea that an application is created to communicate with the normal

		people.
3	Novelty / Uniqueness	This process the image of the person who is using sign language and convert it into the voice by analyzing the sign used.
4	Social Impact / Customer Satisfaction	Differently able people feel free to communicate and it bring a huge difference comparing to past.
5	Business Model (Revenue Model)	There are many people in the world who is differently able,this application will become more popular among them and it will be installed by all and it will be used,and so it will produce more money.
6	Scalability of the Solution	Thus this would bring a new evolution in Real Time Communication System Powered by AI for Specially Able with less time and safe enough resources

3.4 PROBLEM SOLUTION FIT



4. REQUIREMENT ANALYSIS

4.1 FUNCTIONAL REQUIREMENTS

The **System requirements** that are required are specified below,

- Deaf/Dumb person should be able to **perform a sign that represents digit/number.**
- Deaf/Dumb person should be able to **perform a sign that represents a character.**
- Deaf/Dumb person should be able to **perform a sign, where group of characters forms a word.**
- Deaf/Dumb person should be able to **perform a sign, where group of words forms sentence.**
- Especially Deaf people especially should be able to **see the translation of signs to text format.**
- Dumb person should be able to **understand the conversion of text into voice mode.**
- **Normal users should be able to understand the corresponding information conveyed by disabled through sign language.**

Default Operation:

- Users of the app **face the camera and perform the concerned hand sign to convey information.**
- System/Desktop **analyses the sign made by the user**.
- Once analysis gets finished, then the **concerned signs together are shown as a text based and also through voice.**

4.2 NON-FUNCTIONAL REQIREMENTS

1. Users of the app **show the hand sign towards the camera**.

2. Desktop shows that the sign **is not within ROI**.

3. Still User, make sure to present his/her sign within frame.

4. At last, **Desktop finally detects the hand sign**.

Desktop indicates that the user's hand sign is not within the frame or in Region of Interest (ROI).

● Signs are not recognized

1. Excepts the signs that are trained and included in the dataset, the Desktop will never detect the sign rather than this.

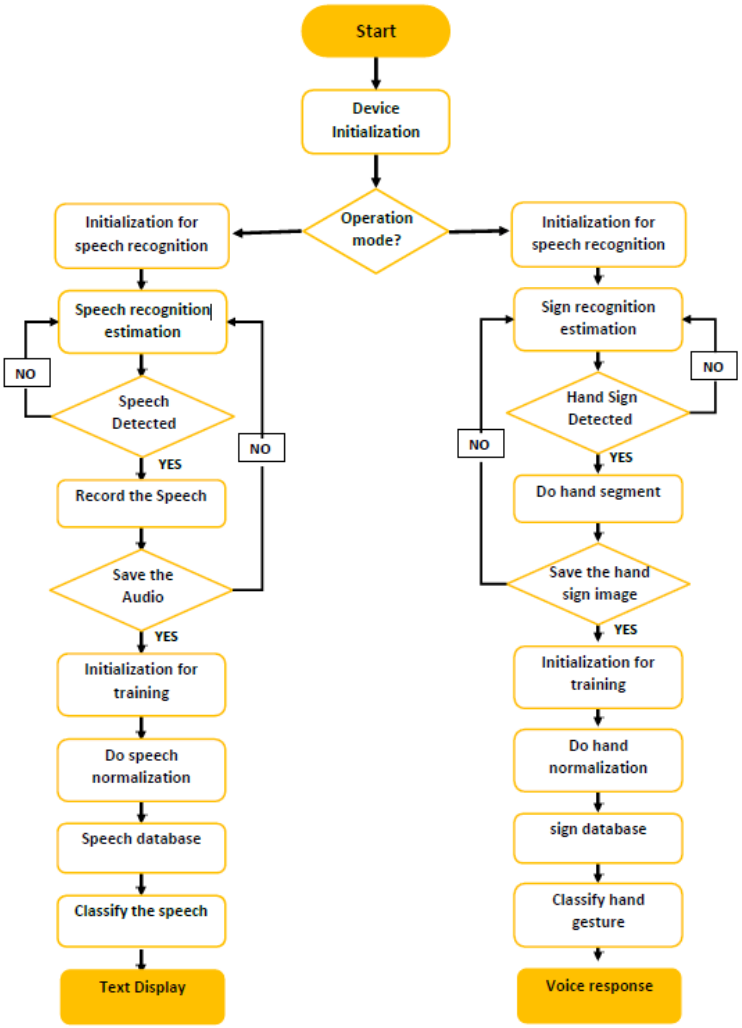
2. User Performs the sign and sees that after 50ms, **the concerned letter occupies the space of text.**

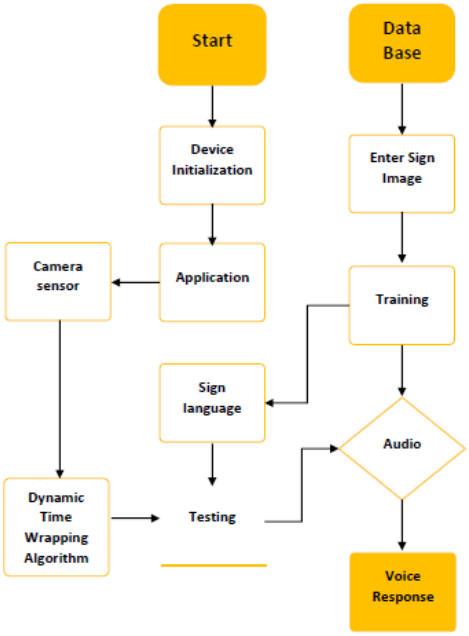
● Speech/Voice assistant is implemented

Speech assistant is to be implemented in order to **convert the output text into voice.**

5. PROJECT DESIGN

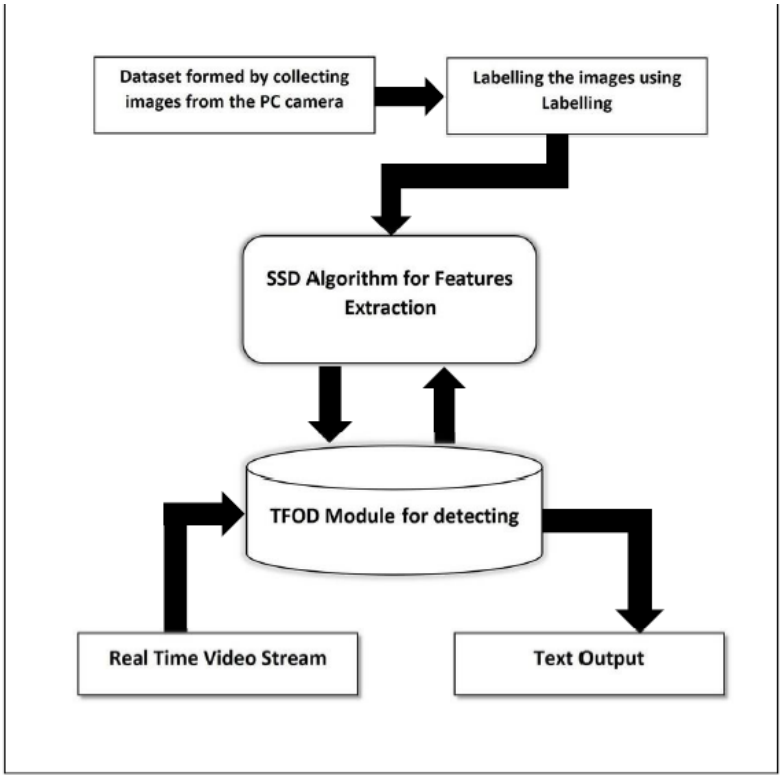
5.1 DATA FLOW DIAGRAM




Start
Data Base
Device Initialization
Enter Sign Image
Camera sensor
Application
Training
Sign language
Audio
Dynamic Time Wrapping Algorithm
Testing
Voice Response

SOLUTION AND TECHNICAL ARCHITECTURE

SOLUTION ARCHITECTURE



6.PROJECT PLANNING AND SCHEDULING

6.1 SPRINT PLANNING AND ESTIMATION

Product Backlog, Sprint Schedule, and Estimation

Sprint	Functional Requirement(Epic)	User Story Number	User Story/ Task	S P
Sprint-1	Data Collection	USN-1	Collect Dataset.	
Sprint-1		USN-2	Image preprocessing	
Sprint-2	Model Building	USN-3	Import the required libraries, add the necessary layersand compile the model	
Sprint-2		USN-4	Training the image classification modelusing CNN	
Sprint-3	Training and Testing	USN-5	Training the model and testingthe model's performance	
Sprint-4	Implementation ofthe application	USN-6	Converting the input sign languageimages into Englishalphabets	

6.2 SPRINT DELIVERY SCHEDULE

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed
Sprint-1	10	6 Days	24 Oct 2022	29 Oct 2022	8
Sprint-2	10	6 Days	31 Oct 2022	04 Nov 2022	5
Sprint-3	10	6 Days	07 Nov 2022	11 Nov 2022	7
Sprint-4	10	6 Days	14 Nov 2022	18 Nov 2022	5

8.RESULTS

8.1 PERFORMANCE METRICS

From the limitation of the existing tools reviewed (Novita, 2006; Macsolvers, 2009; Standup, 2006; Universiteit van Amsterdam, 2008; Crestwood, 2009; Sci-enceDaily, 2008), there is still a pressing need for more effective and efficient tools to alleviate this problem. One the possible methods are to implement a facial expres-sion recognition system to predict or determine the emotional state of a disabled person through his expression projected on his face. biometrics information system can be employed as a means to detect and classify the physiological aspect of a person in real time. Franco and Treves (2001) further support the notion that facial expression can be used for human computer interaction and usability enhancement.

Based on the problem statements deliberated above, we propose an improved real-time communication system using machine learning and computer vision. The aim is to create a communication channel between the specially abled and the society, so they can express there feelings, thoughts and understand other people's feelings and thoughts through real time communication and facial expressions.

10.CONCLUTION

In real life, inter personal human interaction are performed not only using speech or spoken language, but also nonverbal cues for example hand gesture, body gesture, facial expression and tone of the voice. All these cues are sometimes being used for expressing feeling and give feedback (Busso, et al, 2004; Cohen, et Al., 2000). We can see how human interact with each other using non-verbal cues every day. For example a child cries in front of his mother because he is not happy or dissatisfied with something. Other people might interpret it differently thinking that the child might be in pain.

Facial expression interaction is relevant mainly for community social life, teacher and student interaction, credibility in difference contexts, medicine and so on. Besides, facial expression recognition is useful for designing new interactive devices which offers the possibility of new ways for human computer interaction - HCI (Franco & Treves, 2001). Cohen, et al. (2000) conducted survey on their users and noticed that they have been through traditionally HCI consists of the keyboard, mouse, joystick, trackballs, data gloves and touch screen monitors

11. FUTURE SCOPE

Human can interpret and generate major facial expressions but a computer is not built with any facial recognition ability unless through the use of some software. It is even more complicated for the computer to interpret irregular facial expression, especially from those suffering from cerebral palsy. Due to their disorder, they do not have the ability to reflect their emotions like a normal typical person. Thus, a more natural and naive method has to be employed for the system to work by a manual labelling of the image captured with the emotion of the user.

12. APPENDIX

Python:

Python is an interpreted, high-level, general purpose programming language created by Guido Van Rossum and first released in 1991, Python's design philosophy emphasizes code Readability with its notable use of significant White space. Its language constructs and object oriented approach aim to help programmers write clear, logical code for small and large-scale projects. Python is dynamically type and garbage collected. It supports multiple programming paradigms, including procedural, object oriented ,and functional programming.

Keras :

Keras is a powerful and easy-to-use free open source Python library for developing and evaluating **deep learning** model **.**It wraps the efficient numerical computation libraries **Theano** and **TensorFlow** and allows you to define and train neural network models in just a few lines of code. It uses libraries such as Python, C#,C++ or standalone machine learning toolkits. Theano and TensorFlow are very powerful libraries but difficult to understand neural network.Keras is based on minimal structure that provides a clean and easy way to create deep learning models based on TensorFlow or Theano. Keras is designed to quickly define deep learning models. Well, Keras
is an optimal choice for deep learning applications.

Steps for creating a keras model:

1)First we must define a network model.
2)Compile it, which transforms the simple sequence of layers into a complex group of matrix operations.
3)Train or fit the network.
To import: from keras.models import Sequential
From keras.layers import Dense, Activation, Dropout

TensorFlow:

TensorFlow is a Python library for fast numerical computing created and released by Google. It

is a foundation library that can be used to create Deep Learning models directly or by using wrapper librarie sthat simplify the process built on top of **TensorFlow.** TensorFlow tutorial is designed for both beginner and professionals. Our tutorial provides all the basic and advanced concept of machine learning and deep learning concept such as deep neural network, image processing and sentiment analysis. TensorFlow is one of the famous deep learning frameworks, developed by **Google** Team. It is a free and
open source software library and designed in **Python** programming language, this tutorial is designedin such a way that we can easily implements deep learning project on TensorFlow in an easy andefficient way. Unlike other numerical libraries intended for use in Deep Learning like **Theano**,**TensorFlow** was designed for use both in research and development and in production systems. It canrun on single CPU systems, GPUs as well as mobile devices and largescale distributed systems ofhundreds of machines.

Numpy:

NumPy is a Python library used for working with arrays. It also has functions for working in domain oflinear algebra, Fourier transform, and matrices. Numpy which stands for Numerical Python, is a libraryconsisting of multidimensional array objects and a collection of routines for processing those arrays.Using NumPy, mathematical and logical operations on arrays can be performed. This tutorial explains
the basics of NumPy such as its architecture and environment. It also discusses the various array functions, types of indexing, etc. It is an opensource project and you can use it freely.
NumPy stands for Numerical Python. NumPyaims to provide an array object that is up to 50x faster than traditional Python lists. The array object inNumPy is called **ndarray**, it provides a lot of supporting functions that make working with **ndarray** very easy. Arrays are very frequently used in data science, where speed and resources arevery important.

Pillow:

Pillow is a free and ope nsource library for the Python programming language that allows you to easily create &s manipulate digital images. Pillow is built on top of PIL (Python Image Library). PIL is one of the important modules for image processing in Python. However, the PIL module is not supported since 2011 and does n't support python 3.
Pillow module gives more functionalities, runs on all major operating system and support for python
3. It supports wide variety of images such as "jpeg", "png", "bmp", "gif", "ppm", "tiff". You can do almost anything on digital images using pillow module. Apart from basic image processing functionality, including point operations, filtering images using built-in convolution kernels, and color space conversions.

Tkinkter:

Tkinter is the standard **GUI library** for Python. Python when combined with Tkinter provides a fast and easy way to create **GUI applications**. Tkinter provides a powerful object-oriented interface to the Tk GUI toolkit.We need to import all the modules that we are going to need for training our model. The Keras library already contains some datasets and MNIST is one of them. So we can easily import the dataset through Keras. The mnist.load_data() method returns the training data, its labels along with the testing data and its labels.

Jupyter Notebook:

Jupyter Lab is a web-based interactive development environment for Jupyter notebooks, code, and data. JupyterLab is flexible: configure and arrange the user interface to support a wide range of workflows in data science, scientific computing, and machine learning.JupyterLab is extensible and modular: writeplugins that add new components and integrate with existing ones.

Machine Learning:

Machine learning is a method of data analysis that automates analytical model building. It is a branch of artificial intelligence based on the idea that systems can learn from data, identify patterns and make decisions with minimal human intervention.

Deep Learning:

Deep learning is an artificial intelligence (AI) function that imitates the workings of the human brain in processing data and creating patterns for use in decision making. Deep learning is a subset of machine learning in artificial intelligence that has networks capable of learning unsupervised from data that is unstructured or unlabeled. Also known as deep neural learning or deep neural network.

Neural Networks:

A neural network is a series of algorithms that endeavors to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates. In this sense, neural networks refer to systems of neurons, either organic or artificial in nature.

GitHub link
https://github.com/IBM-EPBL/IBM-Project-36314-1660294095