



A Mid Defense Presentation On

FISPAN

*(Finger Spelling Communication Platform for NSL
using AI-Neural Network)*

Presentation By:

Dipen Boyaju	(750313)
Jatin Bhusal	(750315)
Nir Ratna Shakya	(750318)
Sonia Dhaubhadel	(750343)

December 7, 2022

Department Of Computer Engineering

OVERVIEW

- **Background**
- **Problem Statement**
- **Objectives**
- **Scope**
- **Literature Review**
- **Project Management**
- **Tools And Technologies**
- **Methodology**
- **Result And Progress**
- **Remaining Tasks**
- **References**

BACKGROUND

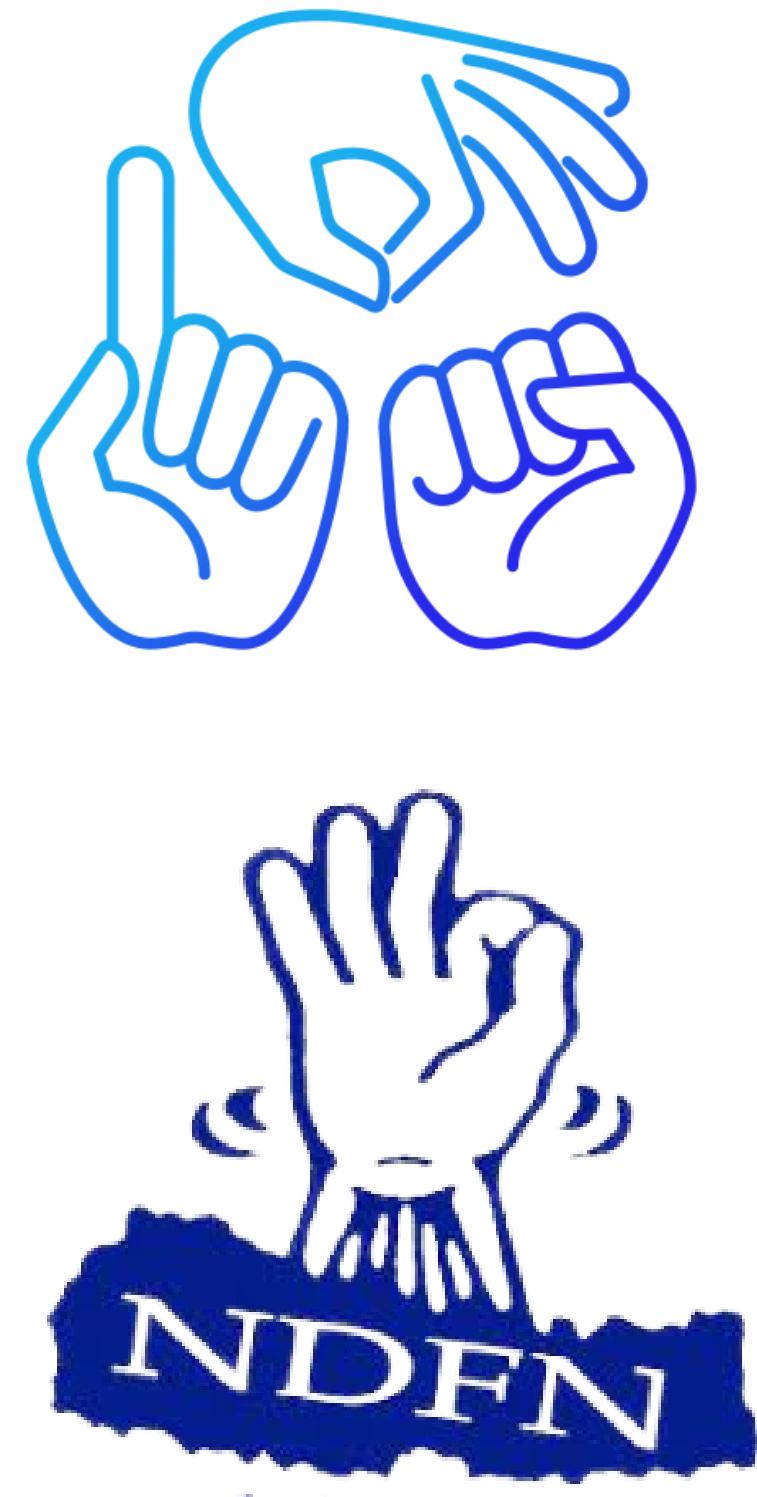
278 Listening impairment cases in 2005

466 million

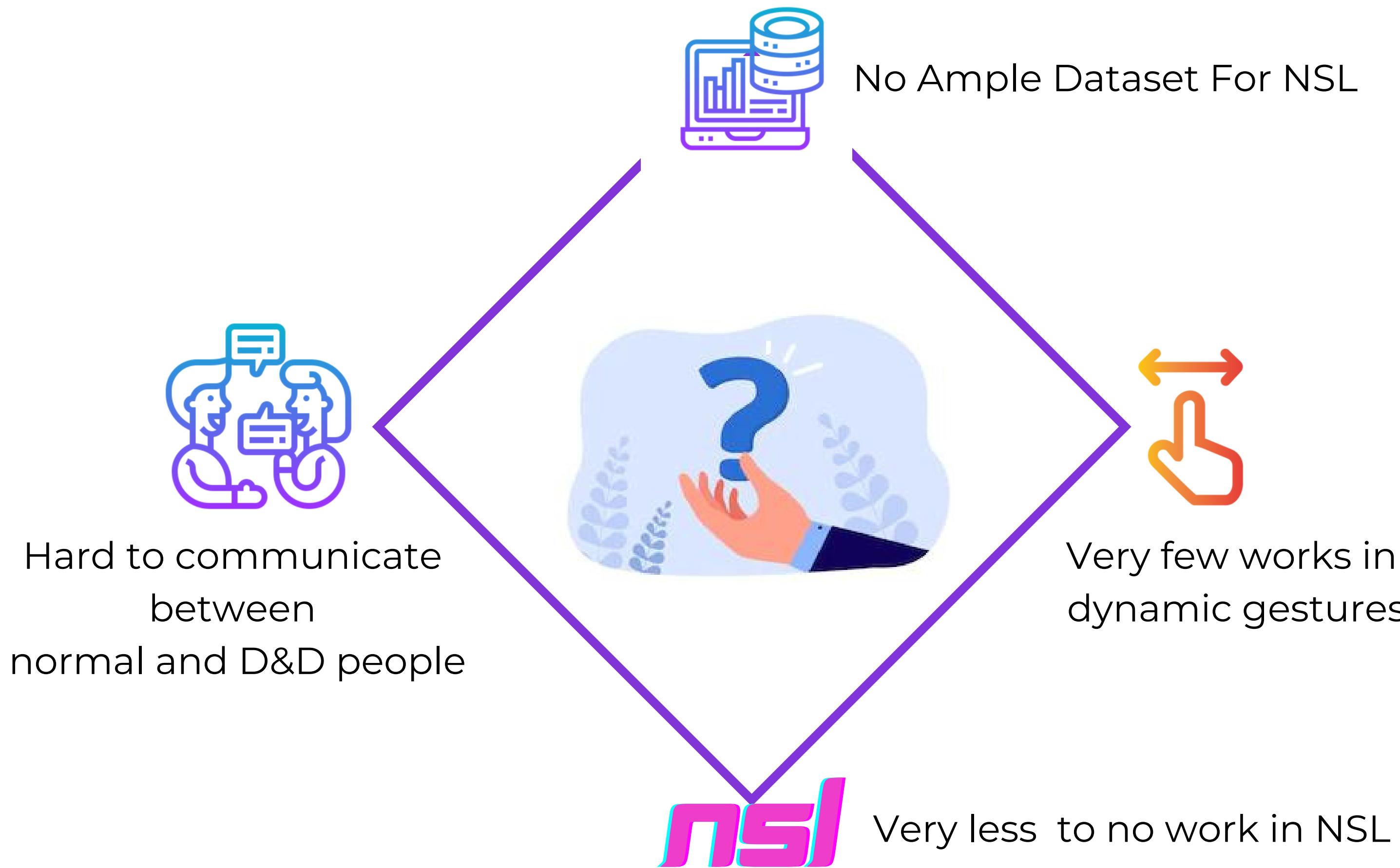
Listening impairment cases by 2015

400 million

New cases shall be added by 2050

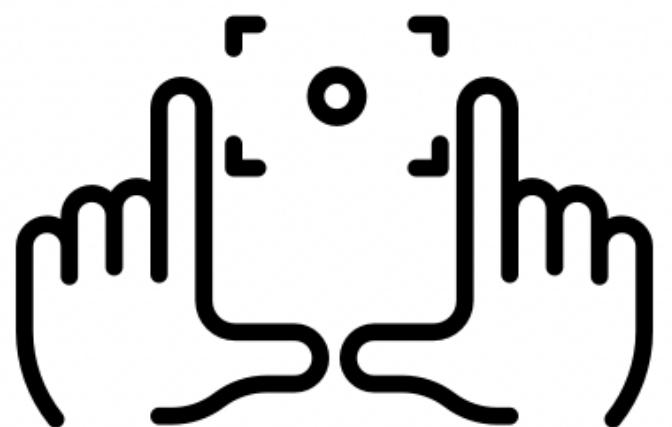


PROBLEM STATEMENT



OBJECTIVE

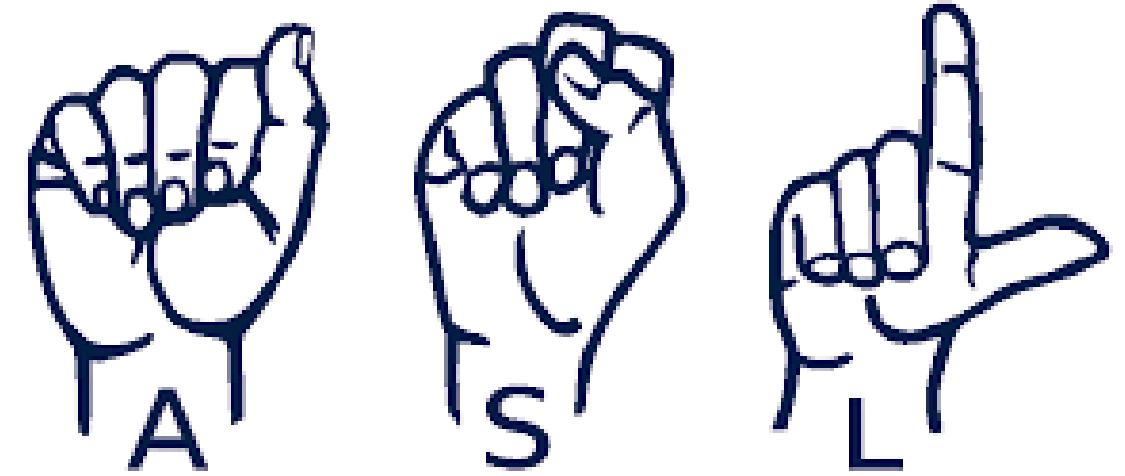
- Initiating the collection of ample and reliable datasets for Nepali Sign Language.
- Focusing on both static and dynamic Sign recognition.



SCOPE

- This project can be used by D&D people to communicate with normal people that detects basic gestures, Nepali alphabets and numbers.

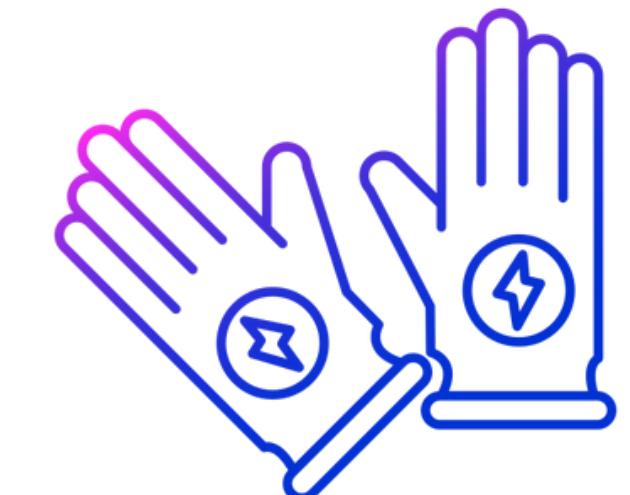
LITERATURE REVIEW



- Static Signs-based Recognitions
- Dynamic Signs-based Recognitions.

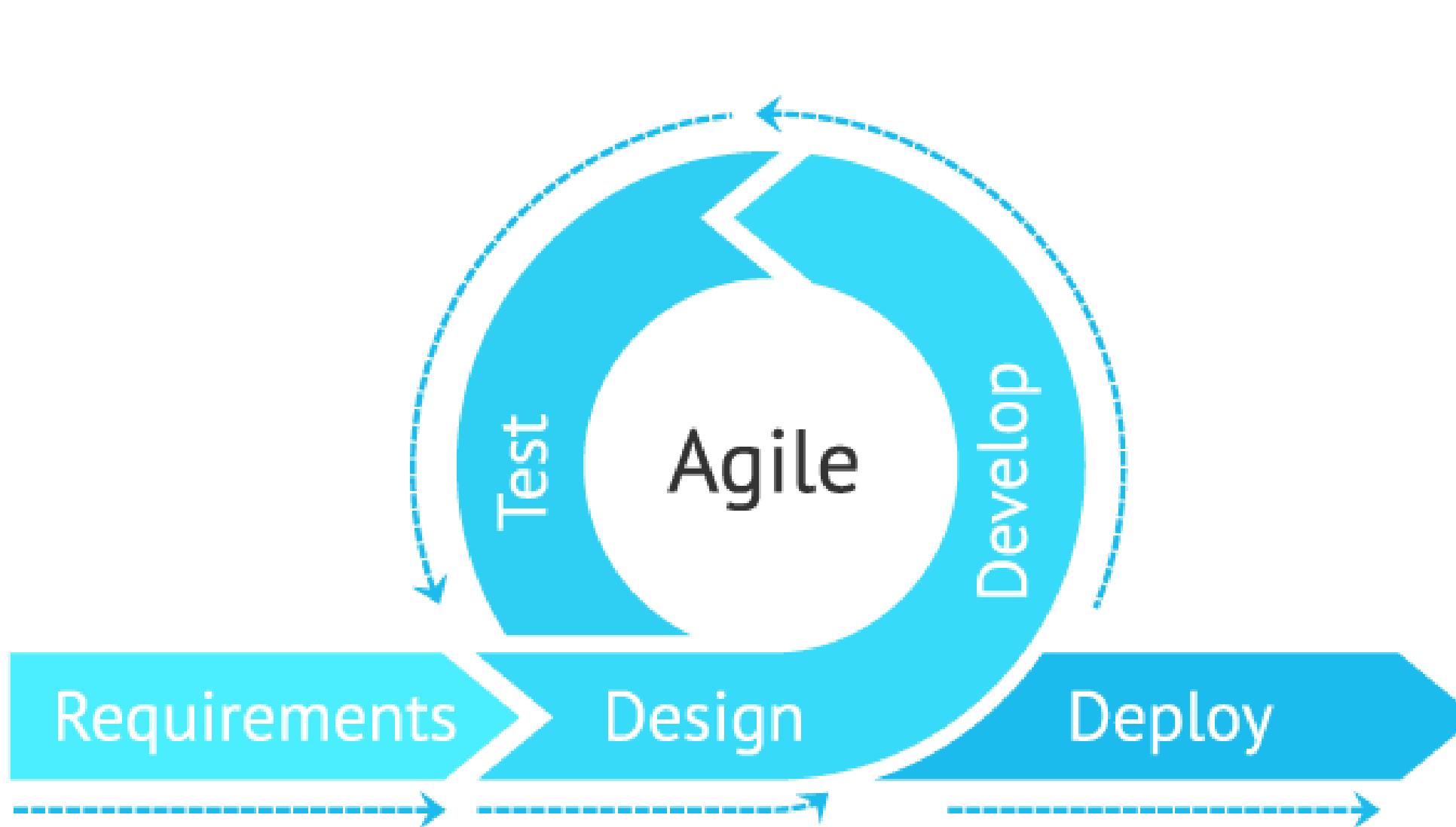


Gesture Based



Sensor Based

PROJECT MANAGEMENT



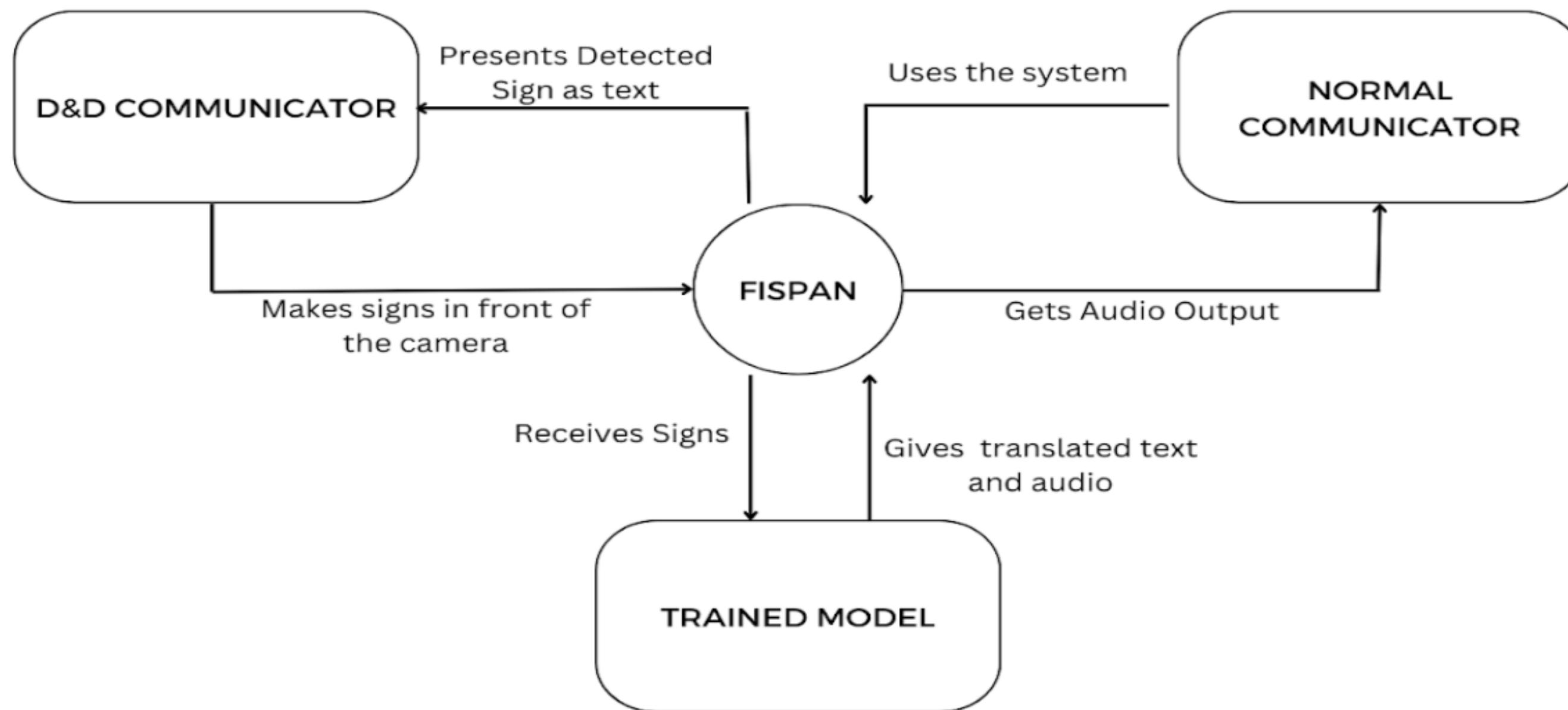
- Data Acquisition,
- Pre-Processing,
- Classification
- Prediction

PROJECT MANAGEMENT(2)



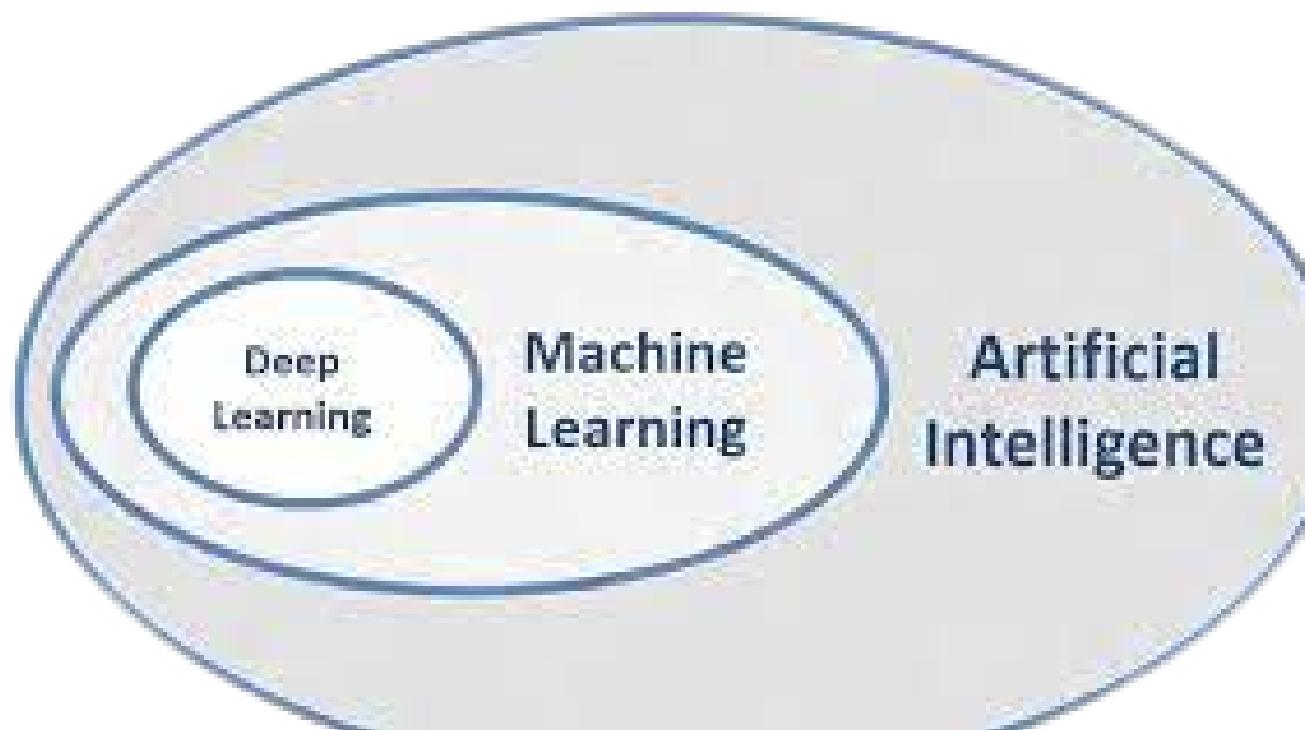
Gantt chart

PROJECT MANAGEMENT(3)

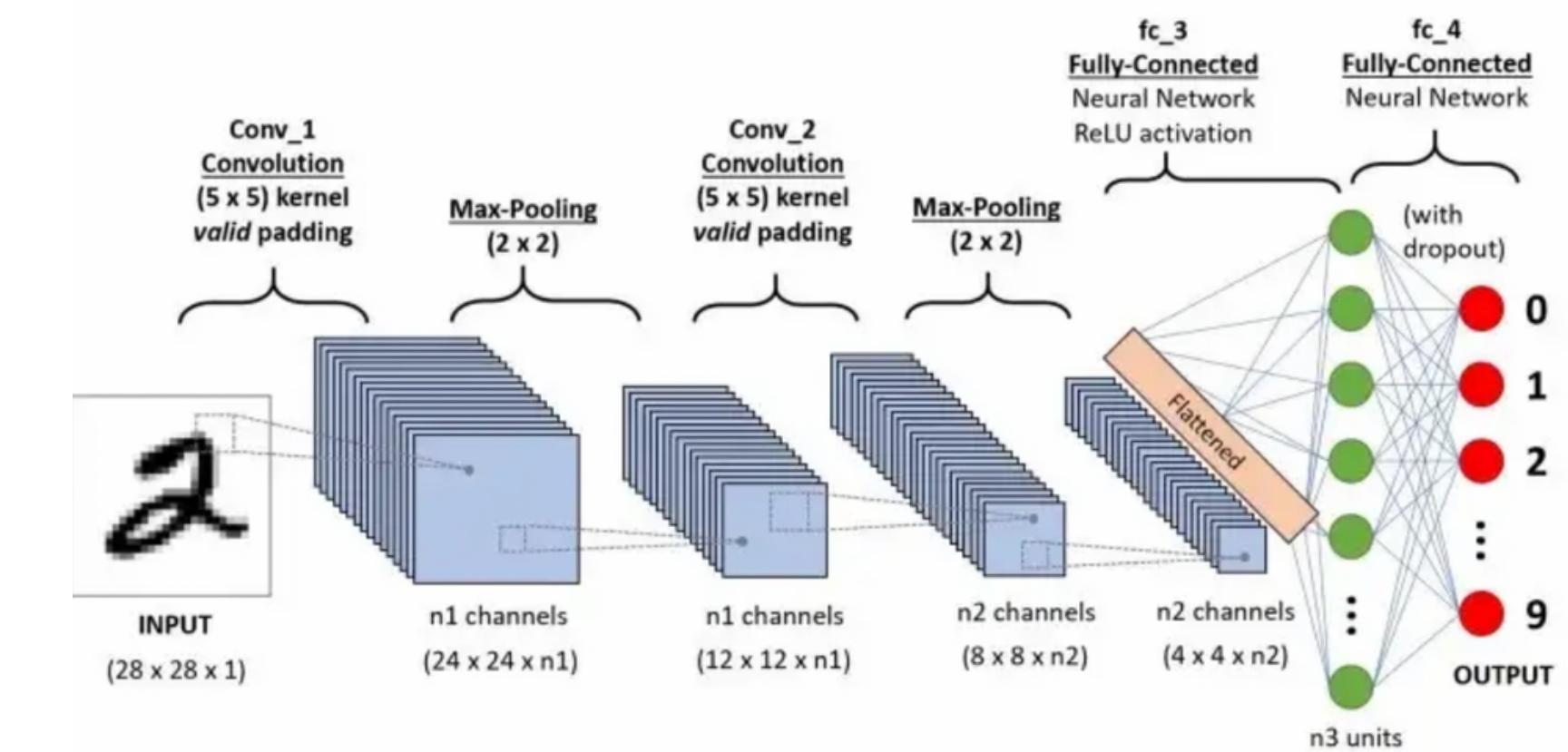


Context Diagram

TOOLS AND TECHNOLOGIES

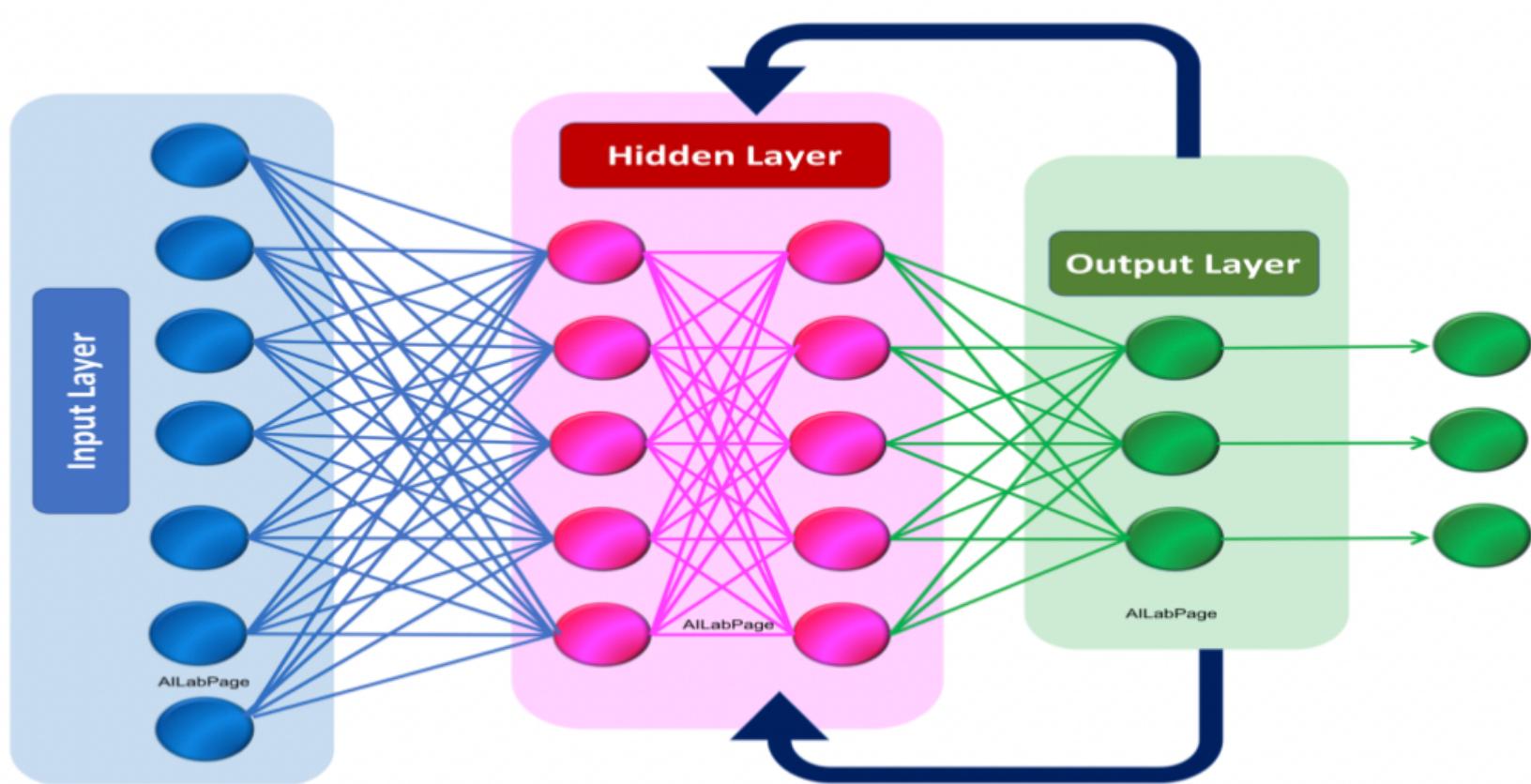


Relationship between
AI, ML and DL

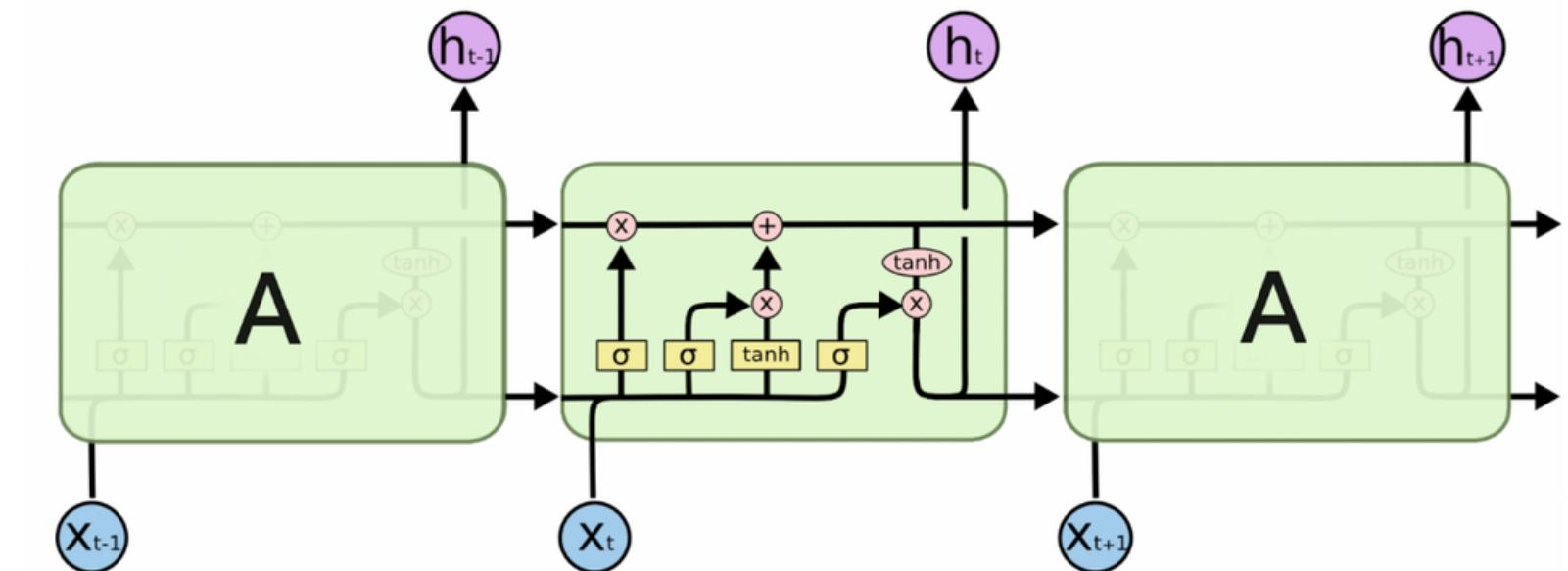


Convolutional Neural
Network

TOOLS AND TECHNOLOGIES(2)



Recurrent Neural
Network



Long Short Term
Memory Network

TOOLS AND TECHNOLOGIES(3)

Tensorflow

Tensorboard

Keras

Numpy

Open CV

**Adam
Optimizer**

Python

**Jupyter
Notebook**

METHODOLOGY

Image Capturing and Preprocessing

Manual collection of 1000 hand signs of each of the first five Devanagari alphabet.

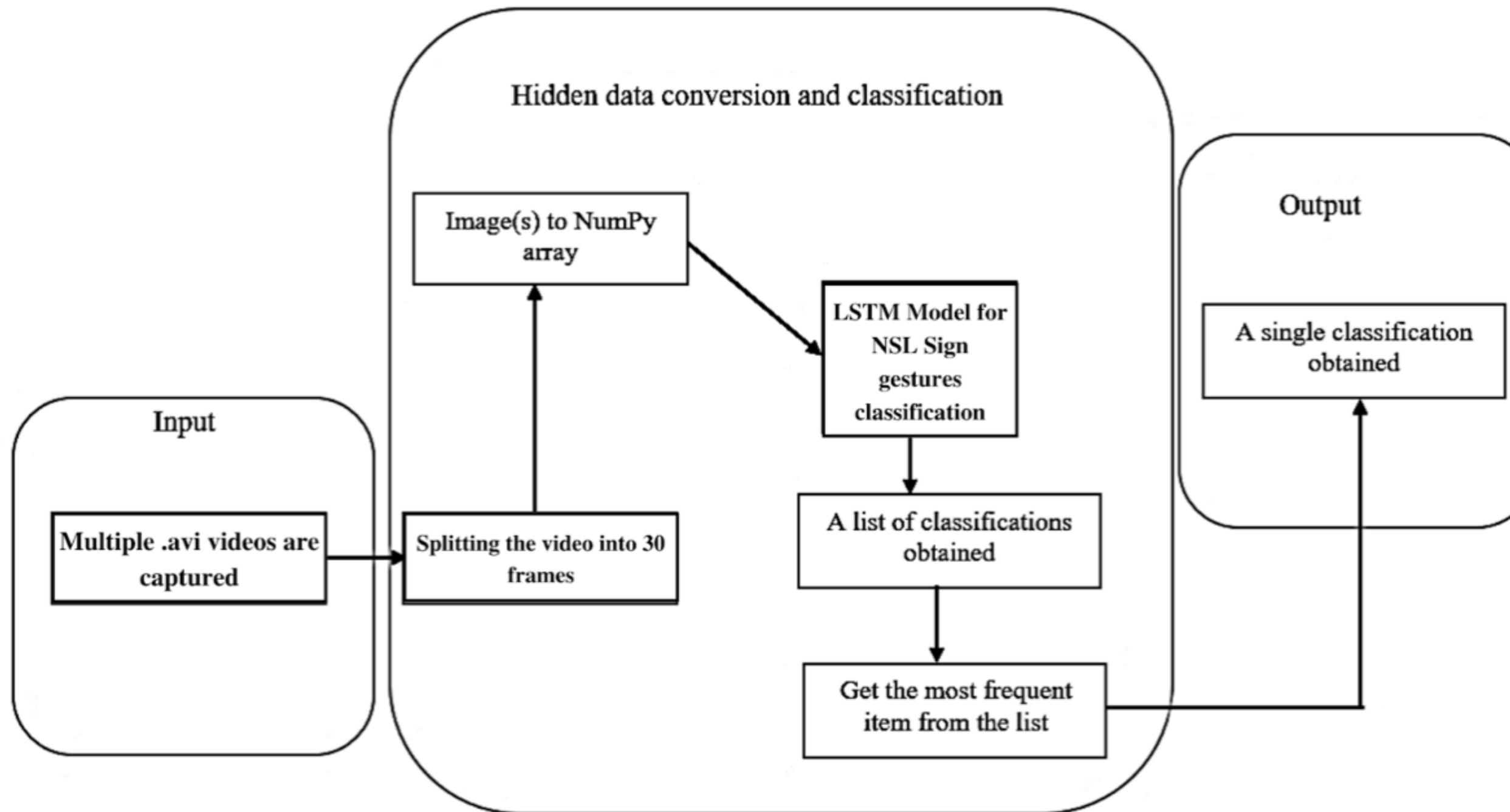
Converted data into Gray-scale image till now.

Collection of 30 frames worth of data for ‘क०’, ‘Sunday’, ‘Red’, ‘Father’ and ‘1’ using “Video capture and split frame” technique for dynamic signs.

X_data and y_data creation for labelling,train-test split, and obtained (X_train,y_train) and (X_test, test) as our training and testing NumPy array in the ratio of 80:20.

METHODOLOGY(2)

Classification



Video Capture Split frame and Classify Method

METHODOLOGY(3)

Classification

```
model.fit(X_train, y_train, epochs=2000, callbacks=[tb_callback])  
  
Epoch 1/2000  
5/5 [=====] - 1s 823ms/step - loss: 2.3234 - categorical_accuracy: 0.1972  
Epoch 2/2000  
5/5 [=====] - 1s 116ms/step - loss: 5.0821 - categorical_accuracy: 0.1761  
Epoch 3/2000  
5/5 [=====] - 1s 110ms/step - loss: 4.0086 - categorical_accuracy: 0.2394  
Epoch 4/2000  
5/5 [=====] - 1s 113ms/step - loss: 2.9205 - categorical_accuracy: 0.2254  
Epoch 5/2000  
5/5 [=====] - 1s 108ms/step - loss: 3.8699 - categorical_accuracy: 0.2113  
Epoch 6/2000  
5/5 [=====] - 1s 110ms/step - loss: 4.9504 - categorical_accuracy: 0.2606  
Epoch 7/2000
```

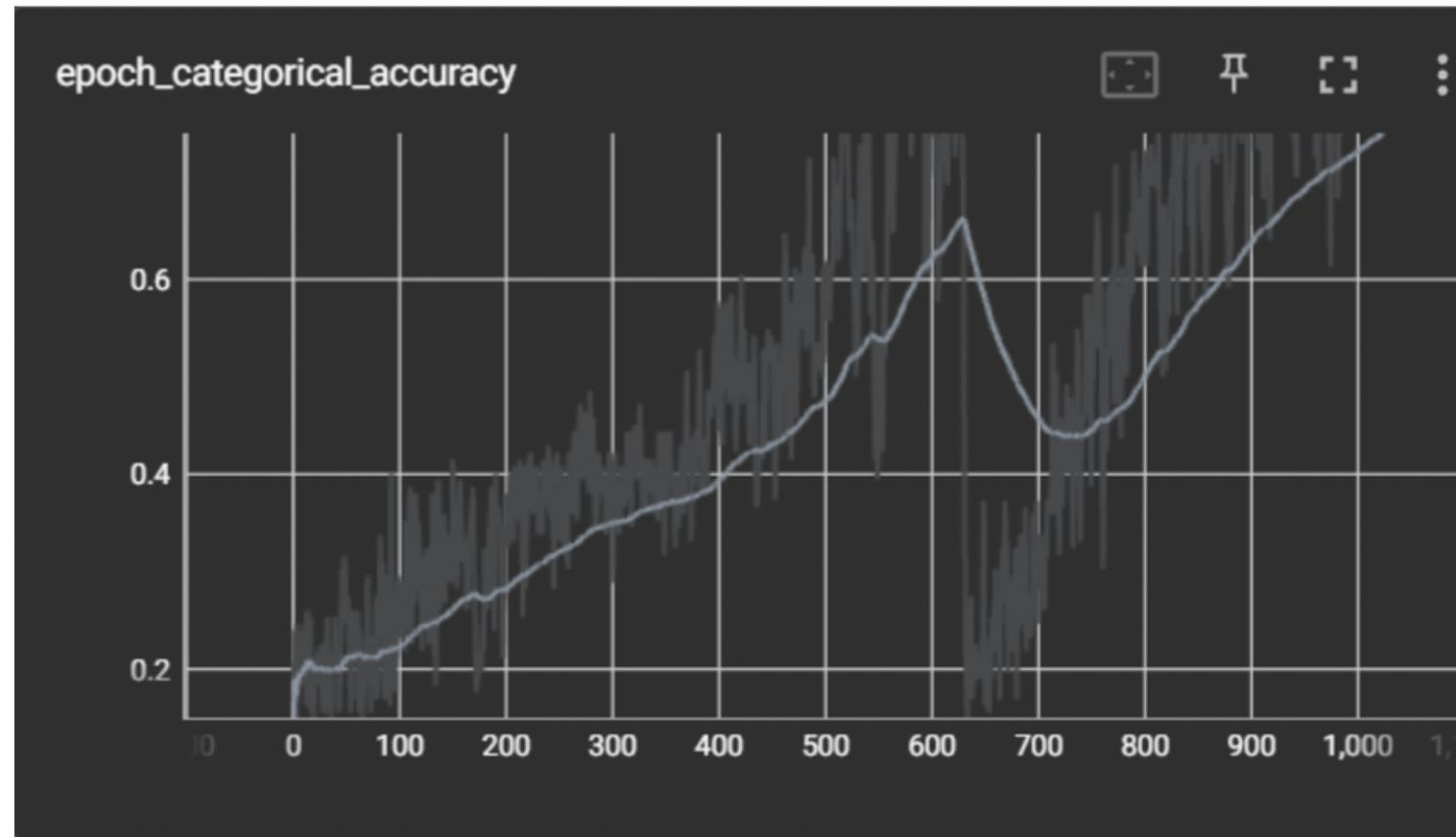
Adam's Epoch

```
model.summary()  
  
Model: "sequential"  
  
Layer (type)          Output Shape         Param #  
=====              ======           ======-----  
lstm (LSTM)          (None, 30, 64)      442112  
lstm_1 (LSTM)        (None, 30, 128)     98816  
lstm_2 (LSTM)        (None, 64)          49408  
dense (Dense)        (None, 64)          4160  
dense_1 (Dense)      (None, 32)          2080  
dense_2 (Dense)      (None, 5)           165  
=====  
Total params: 596,741  
Trainable params: 596,741  
Non-trainable params: 0
```

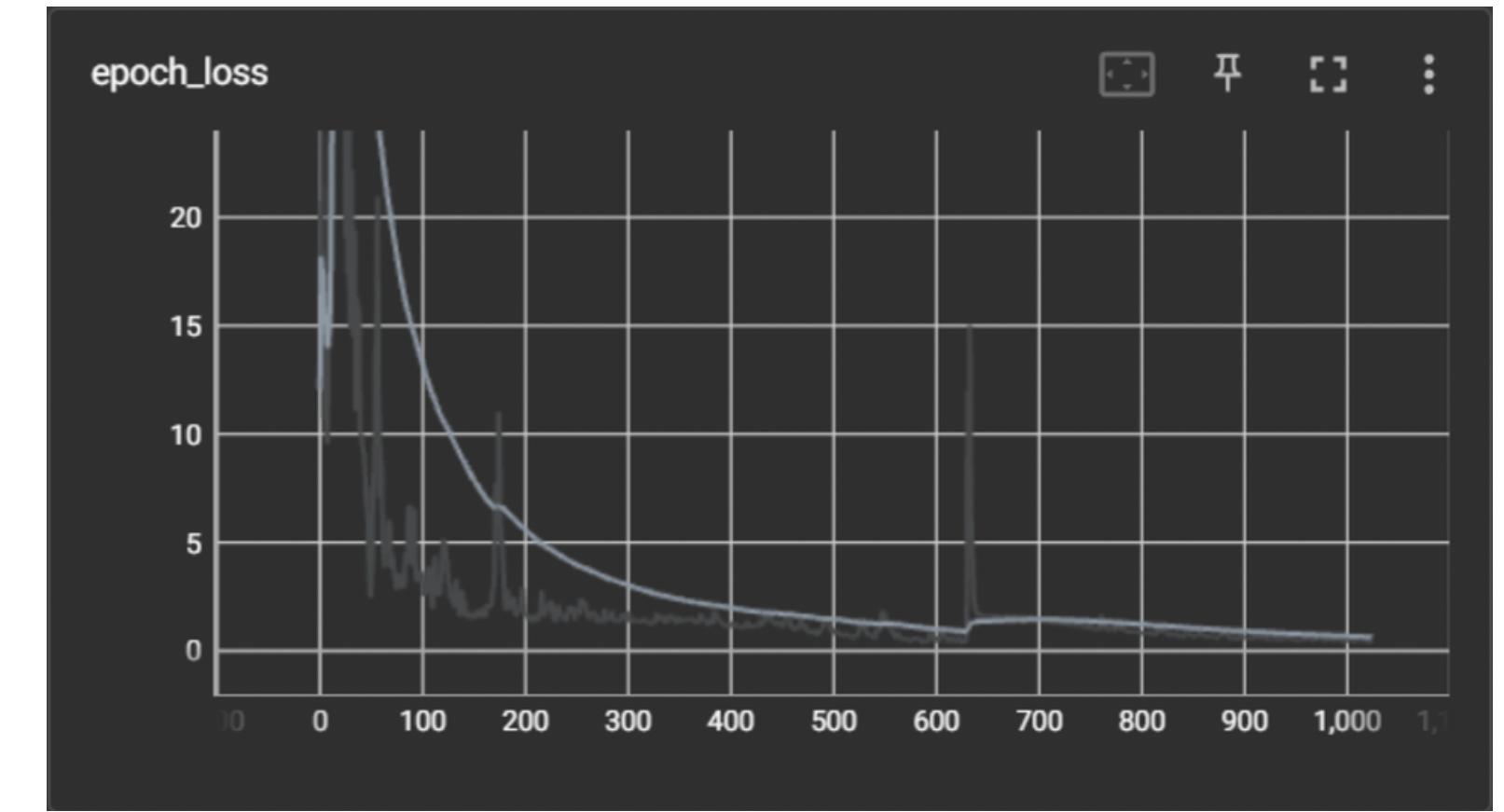
LSTM Execution

METHODOLOGY(4)

Prediction



Epoch Categorical Accuracy



Epoch Loss

METHODOLOGY(5)

Prediction

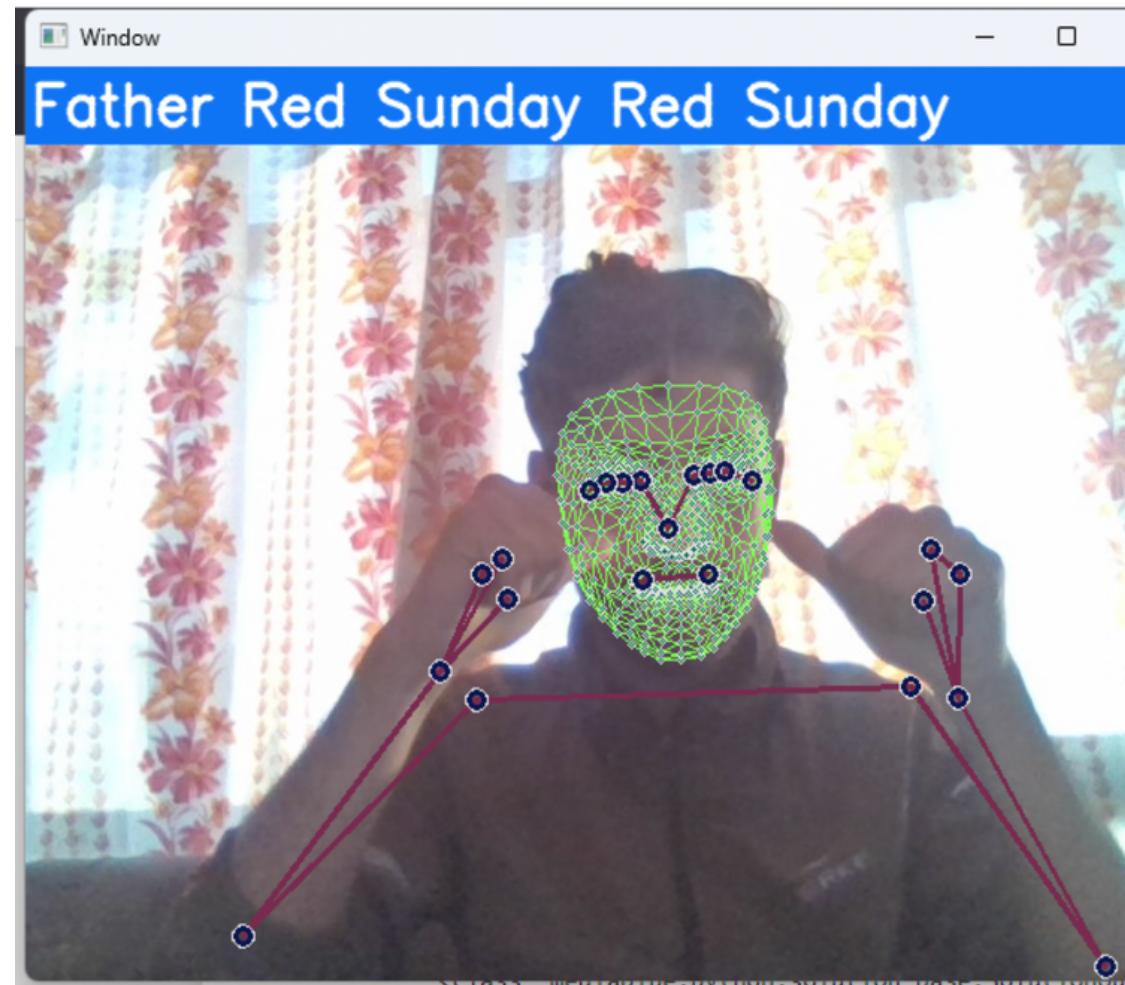
```
In [115]: ► accuracy_score(ytrue, yhat) #ACCURACY TEST: GOT 75%
Out[115]: 0.75
```

Train Data Accuracy

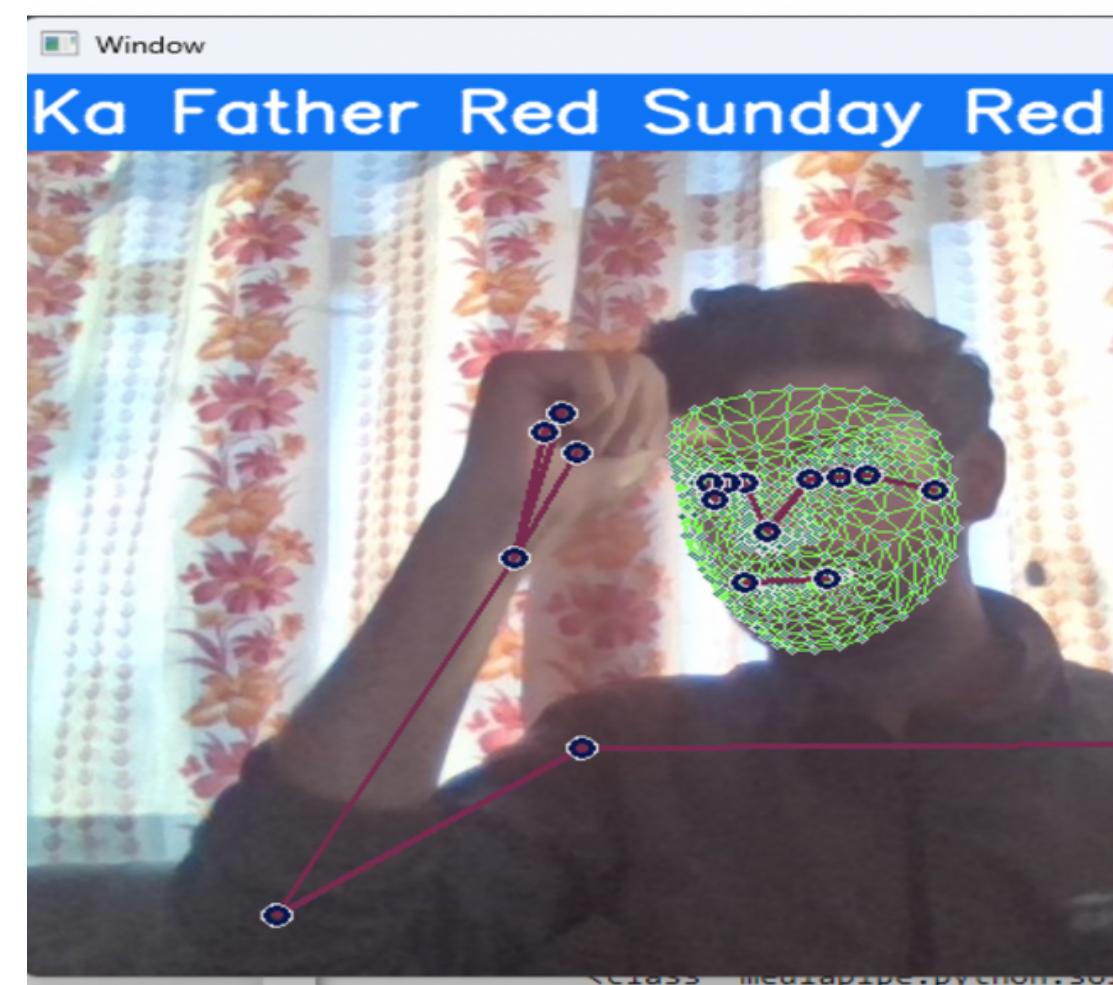
```
In [110]: ► accuracy_score(ytrue, yhat) #ACCURACY TRAIN: GOT 75% |
Out[110]: 0.6619718309859155
```

Test Data Accuracy

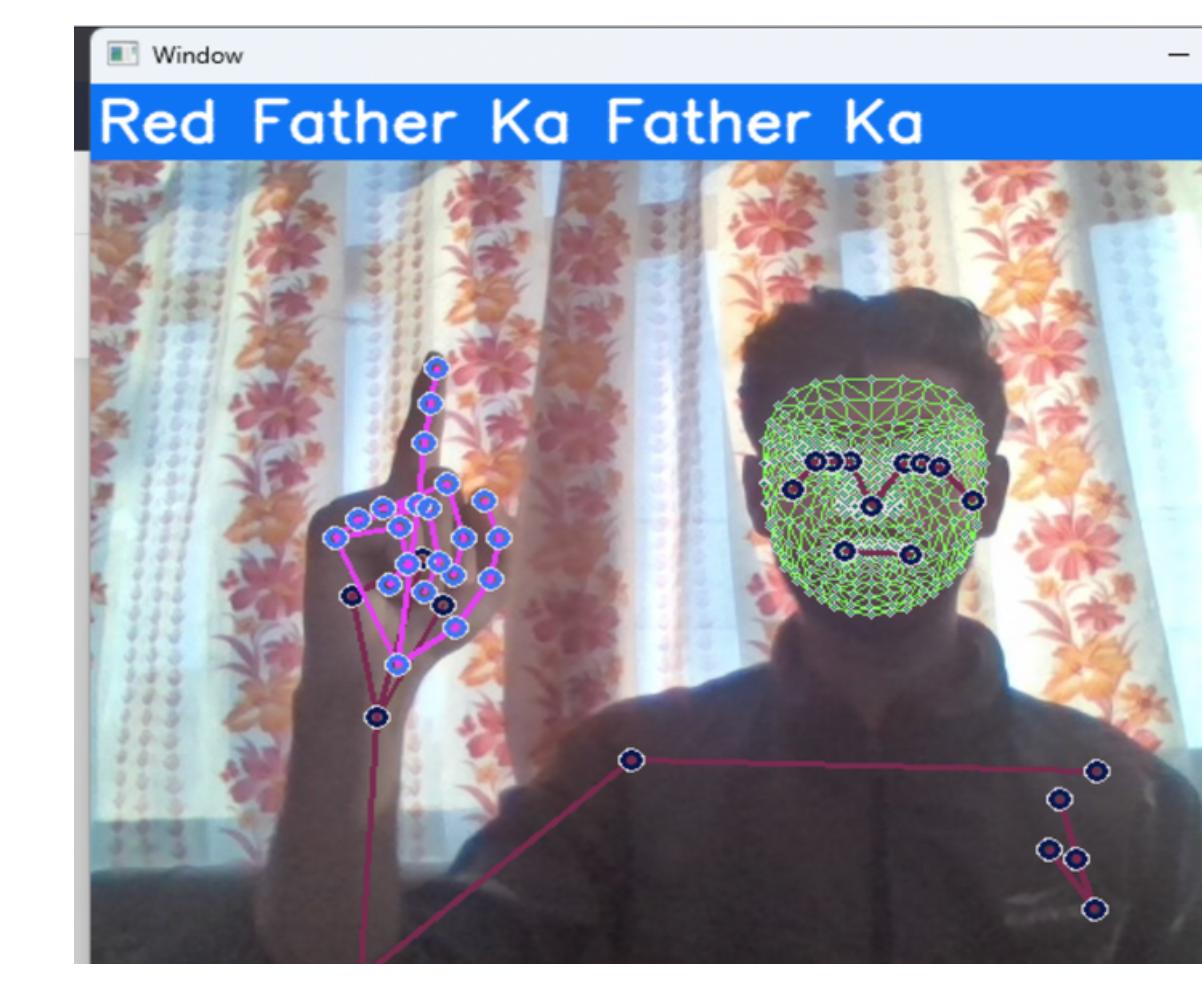
RESULT AND PROGRESS



Detection of 'Sunday'



Detection of 'Red'



Error in detecting 'I'

REMAINING TASKS

Though we have manually collected 1000 images per alphabet, we are yet to process those images and use them to train our model.

Also, we are yet to collect data for dynamic gestures.

Since we have used only 30 data per gesture, we have used LSTM as our model. BUT after we will process all our data, we shall be using VGG-19 model for training.

We aim to add text to speech feature in our project.

The accuracy is by far very low from the threshold. We aim to extend the accuracy above 95%.

BIBLIOGRAPHY

- [1] A. THAKUR, P. BUDHATHOKI, S. UPRETI, S. SHRESTHA, AND S. SHAKYA, "REAL TIME SIGN LANGUAGE RECOGNITION AND SPEECH GENERATION," JOURNAL OF INNOVATIVE IMAGE PROCESSING, VOL. 2, NO. 2, PP. 65–76, JUN. 2020, DOI: 10.36548/JIIP.2020.2.001.
- [2] J. ZHENG ET AL., "AN IMPROVED SIGN LANGUAGE TRANSLATION MODEL WITH EXPLAINABLE ADAPTATIONS FOR PROCESSING LONG SIGN SENTENCES," COMPUTATIONAL INTELLIGENCE AND NEUROSCIENCE, VOL. 2020, PP. 1–11, OCT. 2020, DOI: 10.1155/2020/8816125.
- [3] M. M. RAHMAN, M. S. ISLAM, M. H. RAHMAN, R. SASSI, M. W. RIVOLTA AND M. AKTARUZZAMAN, "A NEW BENCHMARK ON AMERICAN SIGN LANGUAGE RECOGNITION USING CONVOLUTIONAL NEURAL NETWORK," 2019 INTERNATIONAL CONFERENCE ON SUSTAINABLE TECHNOLOGIES FOR INDUSTRY 4.0 (STI), 2019, PP. 1-6, DOI: 10.1109/STI47673.2019.9067974.
- [4] V. I. PAVLOVIC, R. SHARMA, AND T. S. HUANG, "VISUAL INTERPRETATION OF HAND GESTURES FOR HUMAN-COMPUTER INTERACTION: A REVIEW," IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 19, NO. 7, PP. 677–695, JUL. 1997, DOI: 10.1109/34.598226.
- [5] JIYONG MA, W. GAO, JIANGQIN WU, AND CHUNLI WANG, "A CONTINUOUS CHINESE SIGN LANGUAGE RECOGNITION SYSTEM," PROCEEDINGS FOURTH IEEE INTERNATIONAL CONFERENCE ON AUTOMATIC FACE AND GESTURE RECOGNITION (CAT. NO. PR00580), 2000, DOI: 10.1109/AFGR.2000.840670.
- [6] C. VOGLER AND D. METAXAS, "HANDSHAPES AND MOVEMENTS: MULTIPLE-CHANNEL AMERICAN SIGN LANGUAGE RECOGNITION," GESTURE-BASED COMMUNICATION IN HUMAN-COMPUTER INTERACTION, PP. 247–258, 2004, DOI: 10.1007/978-3-540-24598-8_23.
- [7] E. ABRAHAM, A. NAYAK, AND A. IQBAL, "REAL-TIME TRANSLATION OF INDIAN SIGN LANGUAGE USING LSTM," 2019 GLOBAL CONFERENCE FOR ADVANCEMENT IN TECHNOLOGY (GCAT), OCT. 2019, DOI: 10.1109/GCAT47503.2019.8978343.
- [8] WORLD HEALTH ORGANIZATION, WHO. [ONLINE]. AVAILABLE:
[HTTPS://WWW.WHO.INT/NEWS-ROOM/FACT-SHEETS/DETAIL/DEAFNESS-ANDHEARING-LOSS](https://www.who.int/news-room/fact-sheets/detail/deafness-andhearing-loss)
- [9] S. SAHA, "A COMPREHENSIVE GUIDE TO CONVOLUTIONAL NEURAL NETWORKS—THE ELI5 WAY," TOWARDS DATA SCIENCE, DEC. 15, 2018. [HTTPS://TOWARDSDATASCIENCE.COM/A-COMPREHENSIVE-GUIDE-TO-CONVOLUTIONAL-NEURAL-NETWORKS-THE-ELI5-WAY-3BD2B1164A53](https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53)
- [10] O. BOUFELOUSSEN, "SIMPLE EXPLANATION OF RECURRENT NEURAL NETWORK (RNN)," THE STARTUP, SEP. 05, 2020. [HTTPS://MEDIUM.COM/SWLH/SIMPLE-EXPLANATION-OF-RECURRENT-NEURAL-NETWORK-RNN-1285749CC363](https://medium.com/swlh/simple-explanation-of-recurrent-neural-network-rnn-1285749cc363)
- [11] C. OLAH, "UNDERSTANDING LSTM NETWORKS -- COLAH'S BLOG," GITHUB.IO, AUG. 27, 2015. [HTTPS://COLAH.GITHUB.IO/POSTS/2015-08-UNDERSTANDING-LSTMS/](https://colah.github.io/posts/2015-08-Understanding-LSTMs/)
- [12] JASON BROWNLEE, "GENTLE INTRODUCTION TO THE ADAM OPTIMIZATION ALGORITHM FOR DEEP LEARNING," MACHINE LEARNING MASTERY, JUL. 02, 2017. [HTTPS://MACHINELEARNINGMASTER.COM/ADAM-OPTIMIZATION-ALGORITHM-FOR-DEEP-LEARNING/](https://machinelearningmastery.com/adam-optimization-algorithm-for-deep-learning/)

THANK YOU!