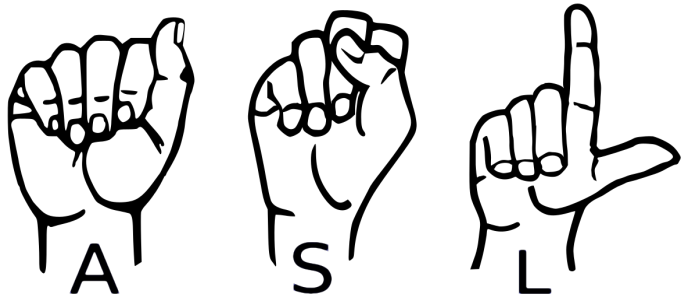
ILLINOIS INSTITUTE OF TECHNOLOGY

Developing a Classification Model for American Sign Language

Pratik Narendra Borse - A20492023
Ashlesh Khajbage - A20517913
Nikhil Singh Thakur - A20528528

Introduction



- American Sign Language (ASL), is a natural language that serves as the predominant sign language of deaf communities in the United States of America and Canada.
- ASL is a complete and organized visual language.
- The goal of developing an American Sign Language (ASL) classification model is to enable a computer to recognize and understand ASL signs automatically.
- This can be used for a variety of purposes, including
  - Assisting deaf people in communicating,
  - Translating sign language,
  - In developing educational tools for learning ASL.

Problem Statement

The objective of our project is to design and develop a highly accurate real-time classification model capable of recognizing and interpreting American Sign Language (ASL) gestures from video data.

ASL signs may involve multiple visual features, such as handshape, movement, and location.

Data Source

- All datasets are sourced from Kaggle.
- We created an additional test dataset which is similar to the testing dataset Dataset 1.

	Features
Dataset 1	- The training data set contains 87,000 images. - Dimension of images - 200x200 pixels. - There are 29 classes, of which 26 are for the letters A-Z and 3 classes for *SPACE*, *DELETE* and *NOTHING*.
Dataset 2	- This dataset contains 27000 images of the alphabet signed in ASL. - Each image is 512 x 512. - Within each set, there are 27 folders, one for each letter and an extra folder of random backgrounds. - Each training folder contains 900 examples while each testing folder contains 100 examples.

Data Preprocessing

Data Cleaning

This step involves removing any irrelevant or redundant data from the dataset.

We **removed outlier** images that contain objects but not hand-gestures.

Data Transformation

ASL signs may involve multiple visual features, such as handshape, movement, and location.

In Data transformation we converted these visual features into a numerical format that can be easily processed by machine learning algorithms.

Data Encoding

After Data Transformation, we converted images or videos into numerical arrays or tensors.

We encoded any temporal or spatial information for sequential or spatial modeling.

In our project we have implemented One-hot-encoding.

Algorithms

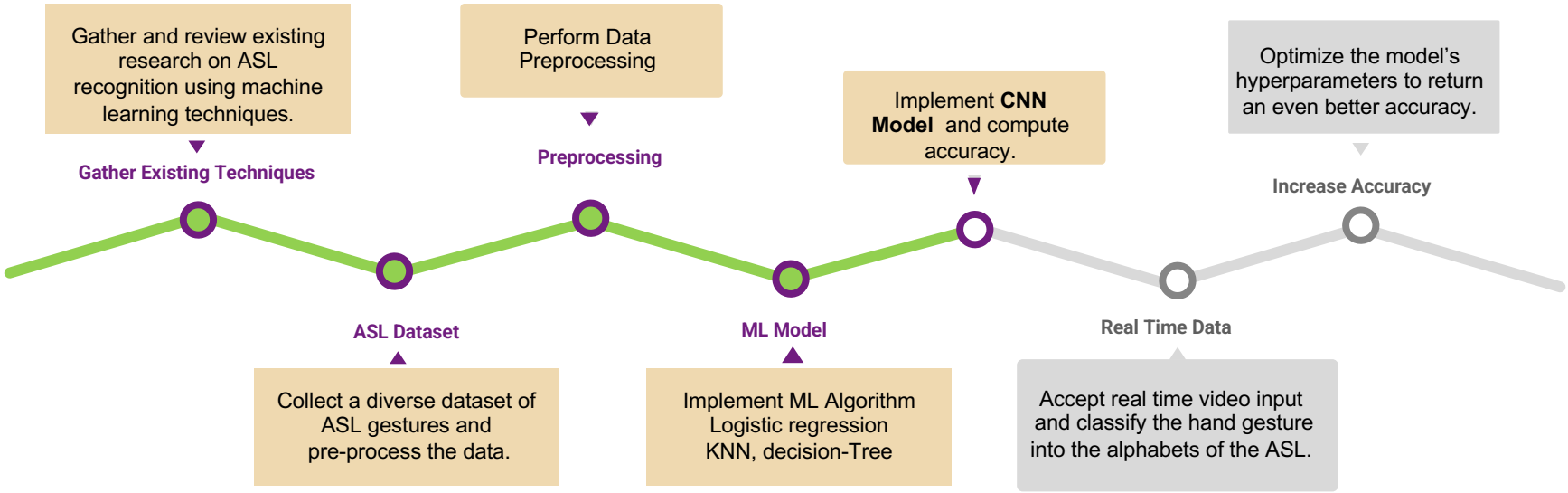
K-Nearest Neighbors

Logistic Regression

Decision Tree

CNN

Project Timeline and Workflow



Output for K–Nearest Neighbor Model

```
Accuracy of the model: 0.9204771371769384
F1 Score of the model:
[0.78787879 0.94736842 0.81481481 1.         0.81481481 0.88
 0.73333333 0.95652174 0.96296296 0.96551724 1.         0.88888889
 1.         1.         1.         1.         1.         1.
 1.         1.         0.85714286 1.         1.         0.91428571
 0.78787879 1.         1.         0.9        1.         0.88
 0.90909091 0.60606061 0.69565217 1.         1.         0.93333333]
```

Output for Logistic Regression Model

```
Accuracy of the model : 0.9860834990059643
F1 Score of the model:
[0.9375     0.97435897 1.         1.         1.         1.
 0.92857143 1.         0.96296296 1.         1.         1.
 1.         1.         1.         1.         1.         1.
 0.96551724 1.         1.         1.         1.         1.
 0.94117647 1.         1.         1.         1.         1.
 1.         1.         0.88       1.         0.95238095 0.96551724]
```

Output for Decision Tree

```
Accuracy of the model : 0.7793240556660039
F1 Score of the model:
[0.4        0.92307692 0.62068966 0.81818182 0.83333333 0.75862069
 0.625      0.91666667 0.68965517 0.55172414 0.84615385 0.90322581
 0.94117647 0.7826087  0.85714286 1.         0.93333333 0.96551724
 0.66666667 0.68181818 0.75       0.93333333 0.72       0.71428571
 0.73170732 0.95652174 0.92857143 0.875      0.66666667 0.69565217
 0.72       0.68421053 0.58823529 0.92307692 0.86956522 0.57142857]


  The Accuracy of KNN is 0.9204771371769384
  The Accuracy of Logistic Regression is 0.9860834990059643
  The Accuracy of Decision Tree is 0.7793240556660039
```

NEXT STEPS

1. Perform comparative analysis of the classification models.

2. Implement a CNN model and compare its accuracy with the previous models.

3. Get real-time video input and perform real-time classification of hand gestures.

Thank you

Any Questions