**Group 1:**

*Group Members:* *Pawandeep Singh, Ayush Pratap Singh, Ansh Kumar, Shivanshu Rawat*

**Abstract for Project Title: Sign Language to Audio Pipeline**

This project delves into the realm of deep learning with a focus on real-time interpretation of sign language, transforming visual gestures into corresponding text or audio outputs. The system employs a straightforward Long Short-Term Memory (LSTM) neural network for both training and inferencing, implemented in Python. Key libraries utilized include ***OpenCV*** for image processing, the ***Keras*** API with TensorFlow backend for neural network development, and ***Mediapipe*** for precise hand landmark recognition.

The project unfolds through a structured process:

**1. Data Collection:**

- Manual gathering of hand landmark data encompassing a diverse range of sign language expressions.

**2. Data Preprocessing:**

- Normalization of collected data points to facilitate consistent and effective training.

**3. Neural Network Construction:**

- Development of an LSTM model using the Keras framework, tailored for sign language interpretation.

**4. Training Phase:**

- Utilization of the constructed model for training on the preprocessed dataset, refining the system's ability to interpret various sign language gestures.

**5. Inferencing:**

- Deployment of the trained model for real-time inferencing, enabling the system to interpret sign language from live footage.

The dataset structure is formatted as ***(N\_SAMPLES, TIMESTEPS, FEATURES)***, reflecting the temporal nature of sign language expressions. The output is a softmax calculation of ***(N\_SIGNS)***, providing a probabilistic distribution of potential sign language interpretations.

To enhance user interaction, a user interface (UI) will be developed using Tkinter or, preferably, the more modern web-based package, Gradio. Gradio offers an intuitive and user-friendly interface for real-time testing and interaction with the sign language interpretation system.

Additionally, the project incorporates a ***tortoise*** codebase dedicated to Text-to-Speech (TTS) functionality. This component adds a layer of naturalness to the audio outputs, ensuring a more human-like and intelligible experience for end-users. The TTS module complements the overall system, enhancing the accessibility and usability of the Sign Language to Audio Pipeline.