**Problem Statement**

Create a Volume Control using hand landmark detection and sign detection using Mediapipe framework and OpenCV. The project will use Google’s Hand Landmark model.

**Main packages to be used for the project**

Mediapipe

MediaPipe is a Framework for building machine learning pipelines for processing time-series data like video and audio. This is a platform Framework that works in Desktop, Mobile phones (ios and Android) and embedded devices and IoT devices. Unlike power-hungry machine learning Frameworks, MediaPipe requires only minimal resources. It is so small and efficient that embedded IoT devices also can run it.

MediaPipe is a deep learning and computer vision-based library for human skeletal posture detection. It employs machine learning (ML) to infer 21 3D landmarks of a hand from a single frame. MediaPipe offers ready-to-use but customizable Python solutions as a prebuilt package.

OpenCV

OpenCV-Python is a library of Python bindings designed to solve computer vision problems.

**Why hand movement recognition?**

The ability to perceive the shape and motion of hands is an important component in improving user experience across various technological domains and platforms. It can be the basis for sign language understanding and hand gesture control and can also enable the overlay of digital content and information on top of the physical world in augmented reality. While coming naturally to people, real-time hand perception is a challenging computer vision task, as hands often occlude themselves or each other (e.g. finger/palm occlusions and handshakes) and lack high contrast patterns.

**Hand Landmark Model of Google**

After the palm detection over the whole image, the subsequent hand landmark model performs precise keypoint localization of 21 3D hand-knuckle coordinates inside the detected hand regions via regression (direct coordinate prediction). The model learns a consistent internal hand pose representation and is resilient to partially visible hands and self-occlusions also.

To better cover the possible hand poses and provide additional supervision on the nature of hand geometry, google provides a high-quality synthetic hand model over various backgrounds and map it to the corresponding 3D coordinates.

