Smart Device Gesture Interaction

A multimodal application for smart interaction of deaf and mute people with smart devices

Course: Multimodal Interaction
Academic Year: 2019-2020
Professor: Maria De Marsico

Full Name: Anxhelo Diko
Matricola: 1867167

Content

1. Problem Description
2. Proposed Solution
3. Implementation
4. Demo

Problem Description

Note: This problem addresses those smart devices that has as their primary communication channel the sound (for input and output)

- Smart devices usually takes only voice commands

  What about mute people?

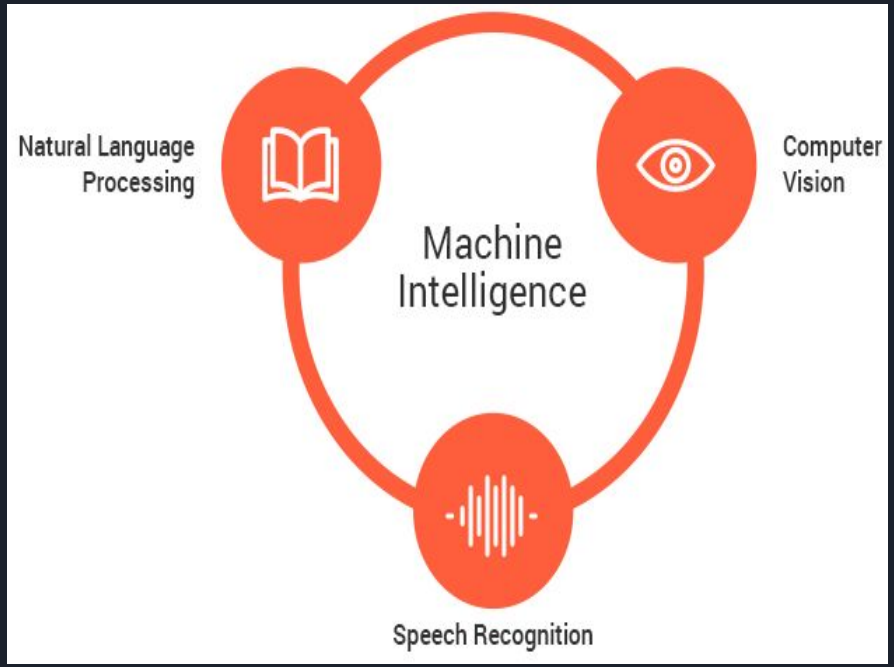
- Smart devices (without graphical interface) reply in spoken language

  What about deaf people?

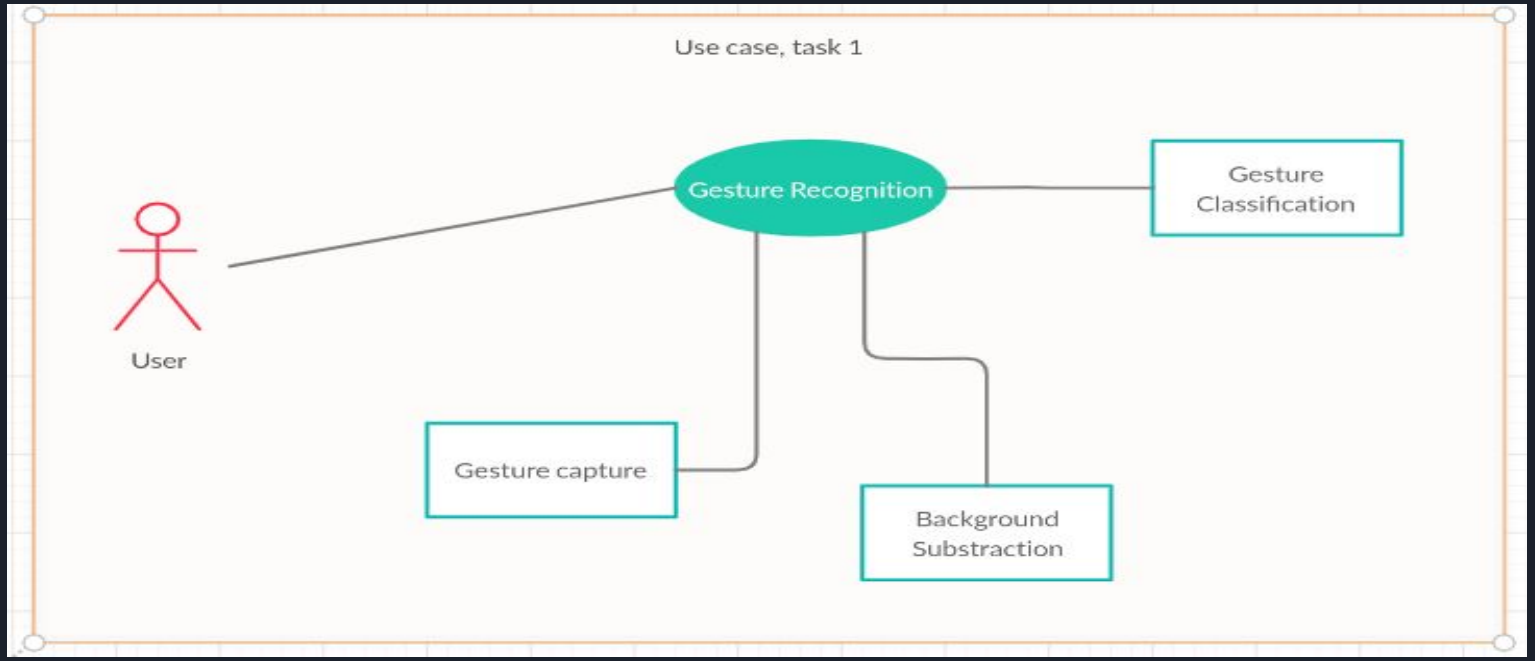
Proposed Solution

- The proposed solution consist in a software that tries to capture gestures, translates them in commands for the smart device and then it recognizes the spoken output of the smart device and shows it as text.

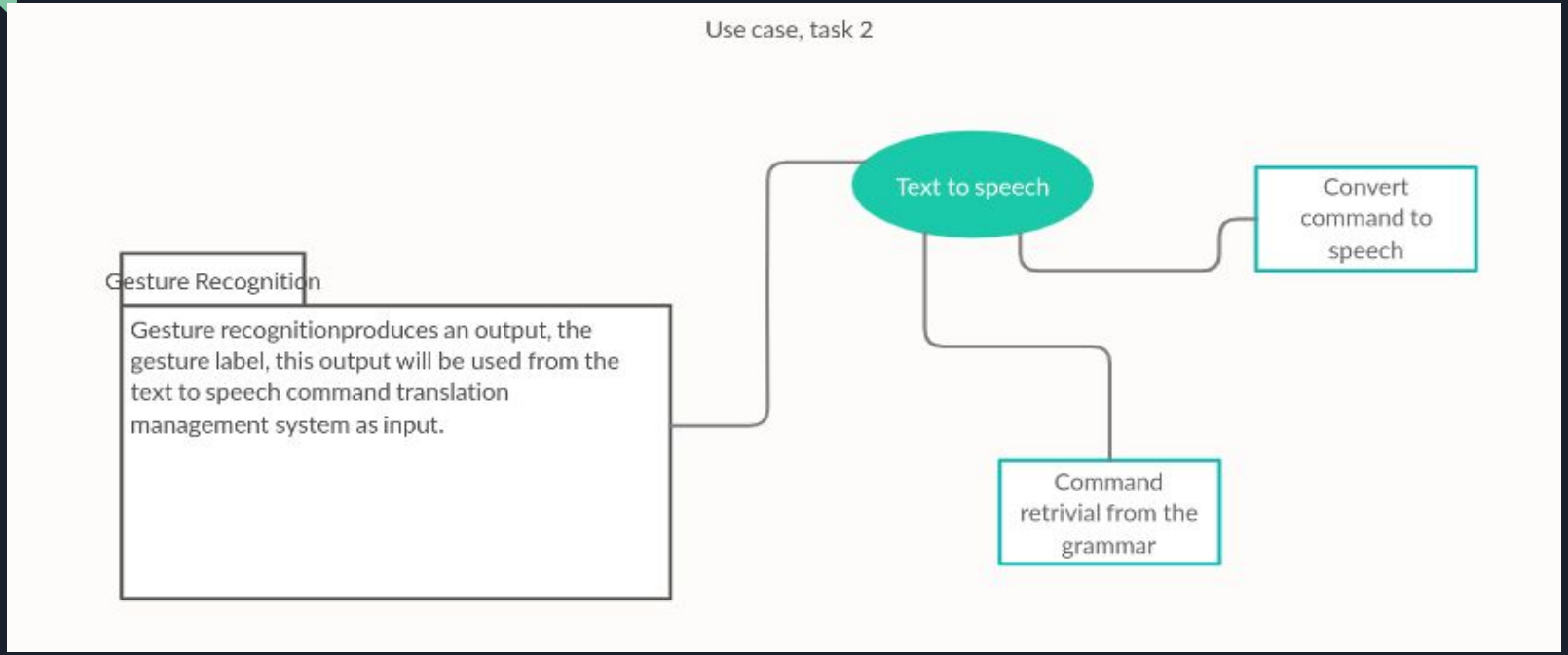
- A modest solution to make possible that mute and deaf people communicate with fancy smart devices in a fancy way.



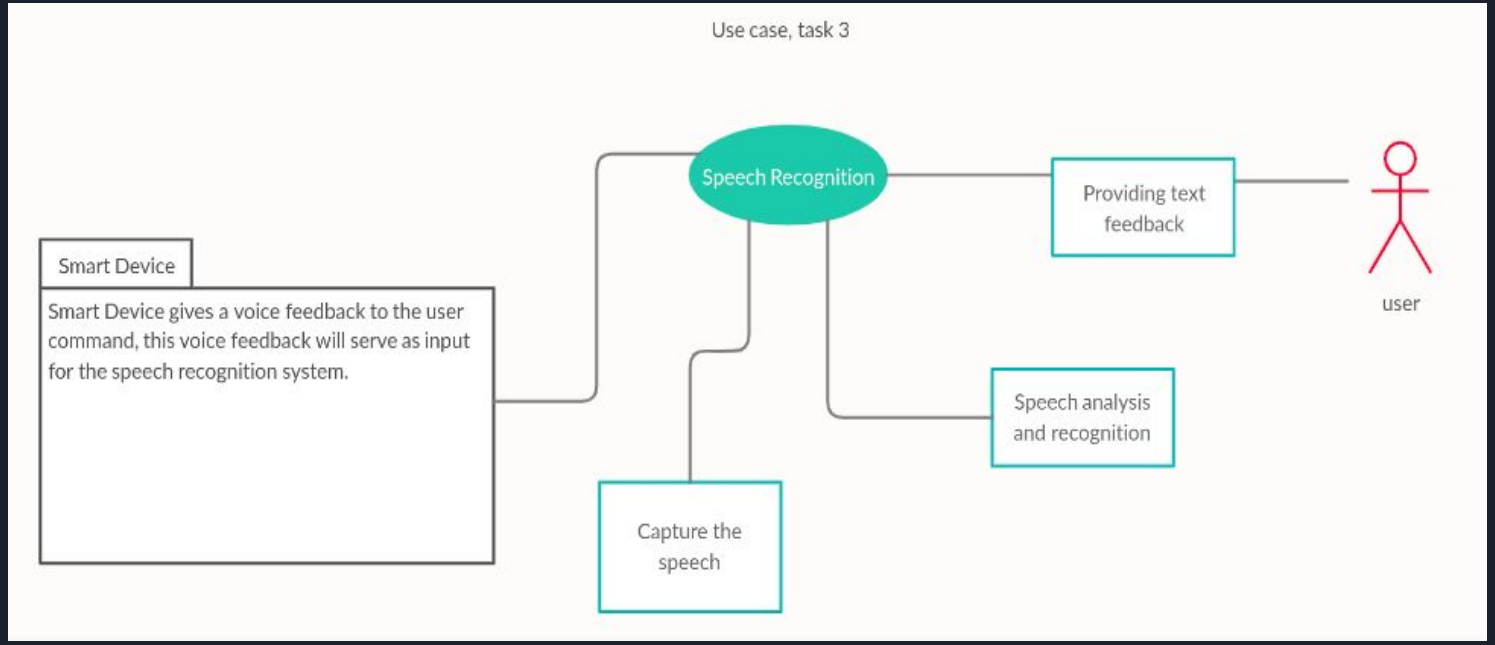
Use cases


Use case, task 1
Gesture Recognition
Gesture Classification
User
Gesture capture
Background Substraction

Use cases


Use case, task 2
Text to speech
Convert command to speech
Gesture Recognition
Gesture recognitionproduces an output, the gesture label, this output will be used from the text to speech command translation management system as input.
Command retrivial from the grammar

Use cases


Use case, task 3
Speech Recognition
Providing text feedback
user
Smart Device
Smart Device gives a voice feedback to the user command, this voice feedback will serve as input for the speech recognition system.
Speech analysis and recognition
Capture the speech

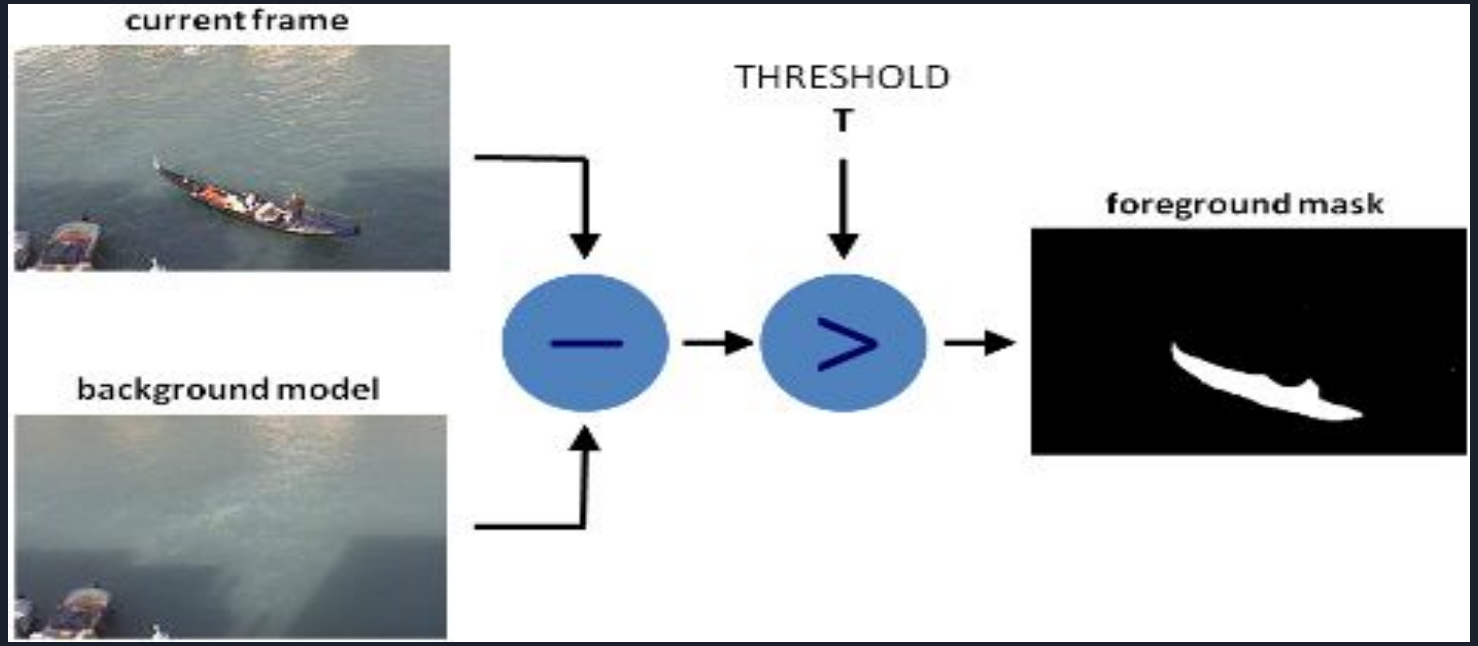
Implemenation

- Computer Vision
- Background subtraction
- Gesture recognition
- Text to Speech
- Speech recognition

Implementation: Computer Vision

- Vision is our primary input channel.
- We use a camera to capture frames that is further analysed
- Background subtraction is used to extract the gestures

Implementation: Background Subtraction



Implementation: Background Subtraction

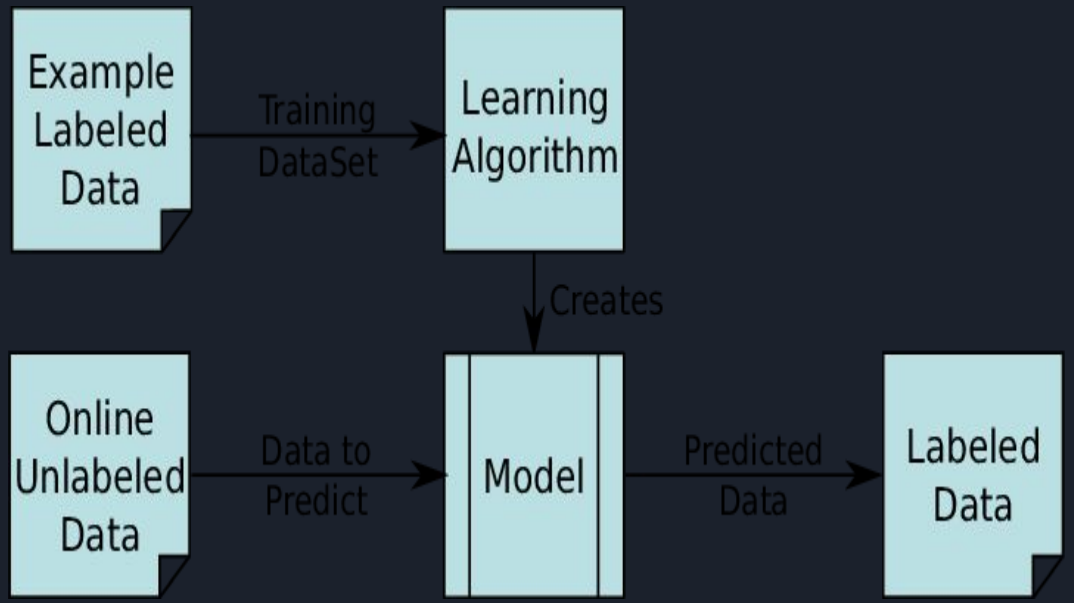
Background

Frame

Extracted Gesture

Implementation: Gesture Recognition

- After the gesture is extracted, it is passed as an input to a gesture classifier.
- The classifier consist of a deep Convolutional Neural Network (AI Based)



Implementation: Gesture Recognition

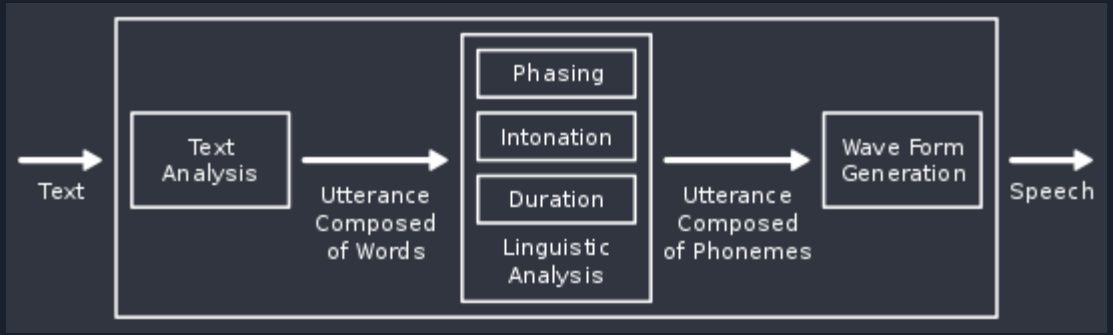
- After the gesture is recognized, the output points to a command found on the grammar.
- Each gesture has a different command on the grammar.

Example

Gesture: Peace
Command:Turn on the light

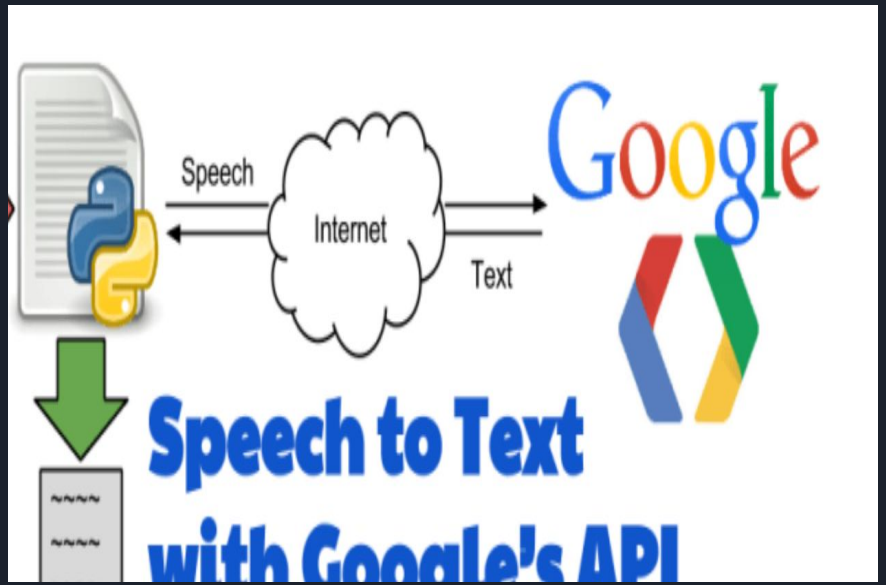
Implementation: Text to Speech

- Google Speech Recognition API for Python
- Sound eXchange (sox)
- Translates the command from text to sound to be recognized by the smart device

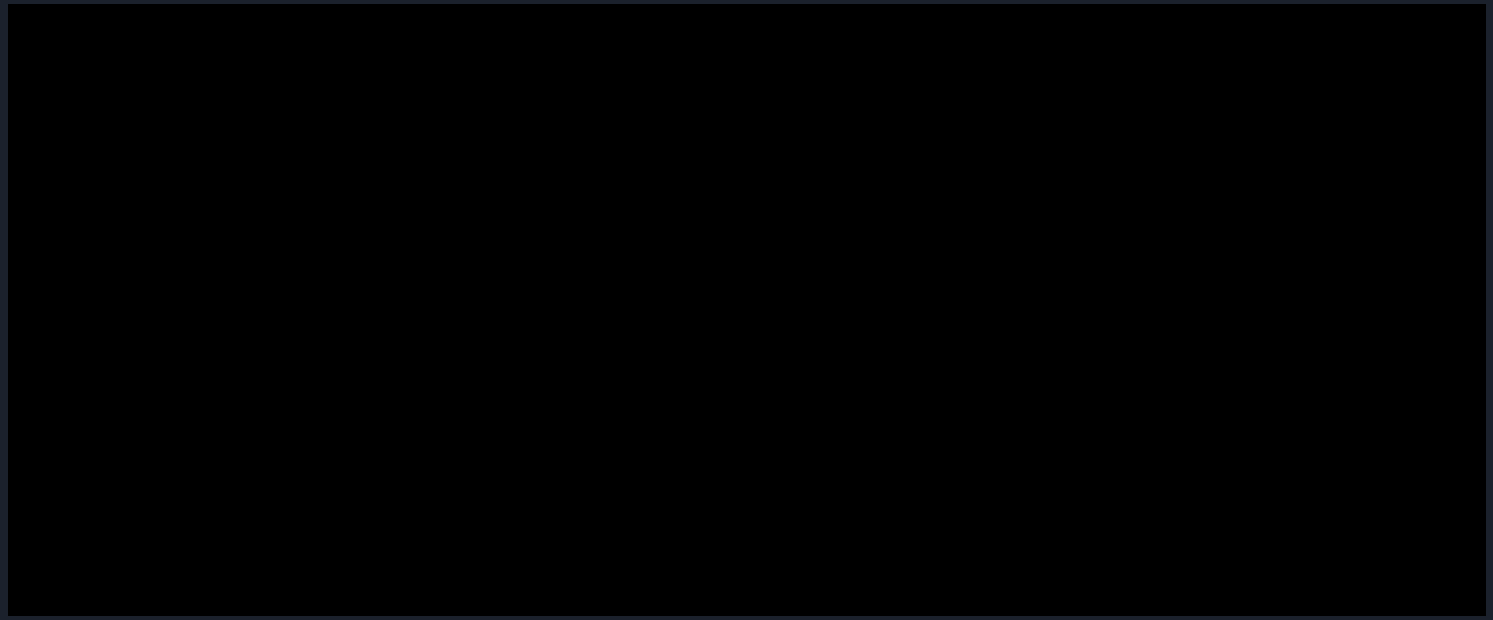


Implementation: Speech Recognition

- Google Speech Recognition API
- We capture the smart device feedback and feed it to google speech recognition engine
- Natural Language processing is also part of this step
- Note: Internet connection is a must



Demo



THANK YOU FOR YOUR ATTENTION