Speech Recognition for Personal Items, Furniture, and Software

-Jason Winata Kusuma 2502007483

-Kevin Praditha Giuniano 2502046175

-Justin Winata Susanto 2502039610

-Rayes Jordan Pradana 2502033102

# Overview

Proyek ini bertujuan untuk mengeksplorasi integrasi teknologi speech recognition ke dalam benda-benda sehari-hari. Hal ini bertujuan untuk memberikan layanan yang lebih mudah diakses dan nyaman bagi para pengguna. Proyek ini mencakup mengintegrasikan teknologi speech recognition yang sudah ada dan mengimplementasikannya pada penggunaan sebagai asisten yang dapat membantu. Hasil yang kami harapkan dari proyek ini adalah sebuah prototipe proof-of-concept yang menunjukkan kelayakan dan potensial teknologi speech recognition pada benda-benda pribadi maupun perabot rumah tangga. Proyek ini bertujuan untuk memberikan titik awal untuk penelitian dan pengembangan lebih lanjut di bidang teknologi speech recognition, dengan harapan agar teknologi ini dapat menjadi lebih mudah diakses oleh semua pengguna. Secara keseluruhan, proyek ini bertujuan untuk mengeksplorasi kemungkinan teknologi speech recognition pada benda-benda pribadi maupun furniture dan juga software , dengan tujuan akhir untuk membuat kehidupan kita menjadi lebih mudah dan nyaman dan menambah pengetahuan kita sebagai mahasiswa dalam membuat suatu proyek implementasi speech recognition.

# How the Product Works

Alur kerja pengenalan ucapan untuk barang pribadi, produk furnitur, software adalah sebagai berikut:

* 1. *User Input:* Pengguna memberikan perintah suara ke barang atau furnitur pribadi, seperti "nyalakan lampu" .
  2. *Input audio*: program menangkap perintah suara pengguna sebagai input audio.
  3. *Speech Recognition* : Input audio diproses oleh mesin speech recognition, yang mengubah kata-kata yang diucapkan menjadi teks.
  4. Analisis teks: Teks dianalisis untuk mengidentifikasi maksud pengguna dan mengekstrak parameter apa pun yang relevan, seperti tindakan dan objek yang diinginkan untuk dikontrol.
  5. Eksekusi perintah: progaram menjalankan perintah yang diidentifikasi, seperti menyalakan lampu.
  6. *Feedback:* Barang atau furnitur pribadi memberikan *feedback* kepada pengguna untuk mengonfirmasi bahwa tindakan yang diminta telah selesai, seperti mengatakan "lampu menyala" .

Secara keseluruhan, alur kerja produk mencakup menangkap input suara pengguna, memprosesnya melalui pengenalan suara dan analisis teks, dan menjalankan perintah yang teridentifikasi pada item atau furnitur pribadi. Tujuannya adalah untuk memberikan pengalaman pengguna yang mulus dan intuitif, di mana pengguna dapat mengontrol barang, furnitur, ataupun hal lainnya yang bersifat pribadi dan hanya dengan berbicara kepada program.

# Methodology

Workflow

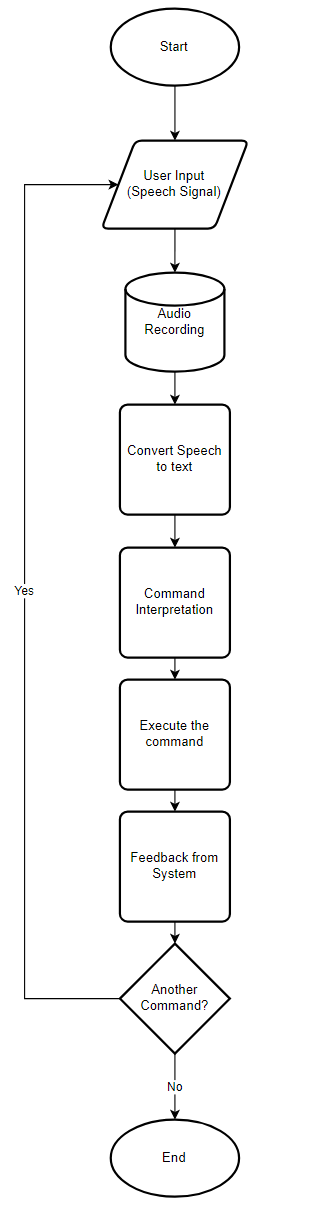


Figure 1 Flowchart of the model

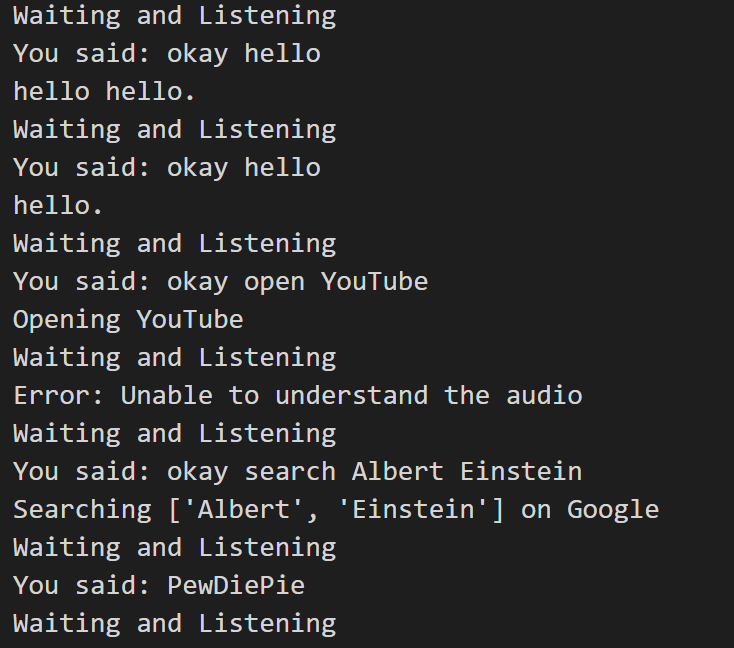
## Model

Model yang digunakan dalam projek ini adalah library speech recognition dengan Google Speech API . Library speech recognition adalah library Python yang menyediakan interface yang sederhana untuk speech recognition.[1] Hal tersebut bertindak sebagai wrapper untuk berbagai mesin speech recognition, sehingga memungkinkan untuk dengan mudah mengintegrasikan kemampuan speech recognition pada python. Speech recognition library dapat menangkap audio dari berbagai sumber seperti mikrofon atau file audio. Library ini juga bisa terkoneksi atau didukung dengan beberapa API yang sudah ada seperti Google Speech Recognition API, Sphinx, PocketSphinx, dan lainnya.[2] Library ini bisa fokus pada tugas tingkat yang lebih tinggi dalam menangkap audio, mengirimkannya untuk dikenali, dan mengambil teks yang dikenali.   
  
 Google Speech Recognition API adalah layanan web yang disediakan oleh Google Cloud Platform (GCP).[3] API ini memanfaatkan advance machine learning technique dan juga termasuk deep neural networks untuk mengubah bahasa lisan menjadi teks tertulis. API ini memiliki kemampuan speech recognition berkualitas tinggi dan akurat. Hal Ini mendukung berbagai format audio seperti FLAC, WAV, dan MP3, dan dapat memproses audio streaming real-time dan file audio pra-rekaman ].Lalu kemudian dapat mengirim data audio ke API untuk diproses dan akan dikembalikkan dengan teks yang ditranskripsi sebagai response.

# Result

This section displays the result of your project.

* Interface of the product: command console



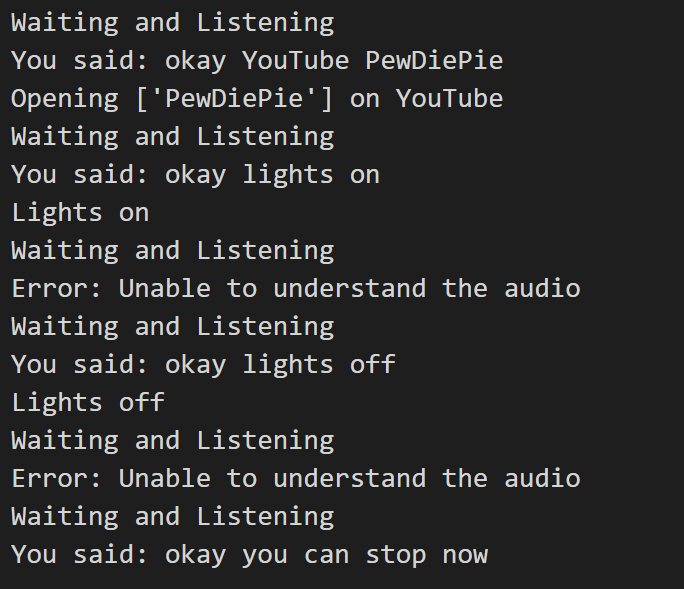


Figure 2 Result of the model

References

[1] “Speech recognition in python using google speech API,” GeeksforGeeks, <https://www.geeksforgeeks.org/speech-recognition-in-python-using-google-speech-api/> (accessed Jun. 18, 2023).

[2] Real Python, “The ultimate guide to speech recognition with python,” Real Python, <https://realpython.com/python-speech-recognition/> (accessed Jun. 18, 2023).

[3] “Speech-to-text: Automatic speech recognition google cloud,” Google, <https://cloud.google.com/speech-to-text> (accessed Jun. 18, 2023).