# Software Requirement Specification

June 19, 2024

| Project ID & Name | Speech Detection | | |
|---|---|---|---|
| Project Manager | Md. Mohoiminul Islam Chowdure | | |
| Customer Request Reference | | | |
| Prepared by | Abu Darda | Date | 24 March, 2024 |
| Reviewed by | | Date | |

# Contents

# 1 Requirement Scope Summary

The project involves developing a Python script capable of processing an audio file to detect the language spoken, convert the speech to text, translate the text to English, and convert both the detected language text and translated English text back to speech. The outputs, including detected language, generated texts, and synthesized speeches, are to be saved to disk.

# 2 Functional Requirements Identification

1. Install necessary libraries.

2. Load an audio file.

3. Detect the language spoken in the audio file.

4. Convert the speech to text.

5. Translate the text to English.

6. Convert both the detected language text and the English text back to speech.

7. Save the detected language, generated texts, and synthesized speeches to disk.

# 3 Functional Requirements Details

## 3.1 Install Libraries

- **Description:** Install the required Python libraries for audio processing, language detection, speech recognition, text translation, and text-to-speech conversion.

- **Libraries:** `pip install langdetect SpeechRecognition googletrans gtts`

## 3.2 Load an Audio File

- **Description:** Load an audio file from the disk for processing.

- **Input:** Path to the audio file.

- **Output:** Audio data loaded into the program.

## 3.3 Detect Language

- **Description:** Analyze the audio file to detect the language spoken.

- **Output:** Detected language code (e.g., 'en' for English, 'es' for Spanish).

## 3.4 Convert Speech to Text

- **Description:** Use a speech recognition library to convert the detected speech in the audio file to text.

- **Output:** Text representation of the speech.

## 3.5 Translate Text to English

- **Description:** Translate the text obtained from speech recognition into English.

- **Output:** Translated text in English.

## 3.6  Convert Text to Speech

- **Description:** Convert the detected language text and translated English text back to speech using a text-to-speech library.

- **Outputs:** Synthesized speech audio files for both the detected language and English.

## 3.7  Save Outputs to Disk

- **Description:** Save the detected language, generated texts, and synthesized speeches to disk.

- **Files:**

    - Detected language in `language.txt`
    - Generated text in detected language in `detected_text.txt`
    - Generated text in English in `translated_text.txt`
    - Synthesized speech in detected language in `detected_speech.mp3`
    - Synthesized speech in English in `translated_speech.mp3`

- **Directory:** All files saved in `audio_advanced_output/`