**📄 Milestone 1 Documentation: Audio File Cleaning**

**1. Objective**

The goal of Milestone 1 is to prepare raw audio recordings for further speech processing tasks (transcription, diarization, summarization). This involves:

* Taking an input audio file.
* Checking its technical properties.
* Cleaning and converting it into a standardized format (mono, 16kHz).
* Saving the cleaned version for later use.

**2. Tools and Libraries Used**

* **Google Colab** → for coding environment.
* **Google Drive** → for storing and accessing audio files.
* **Python wave library** → to inspect audio properties (channels, sample rate, duration).
* **FFmpeg** → for cleaning and converting audio format.

**3. Process Steps**

**Step 1: Mount Google Drive**

Connect Google Drive to Colab so that audio files can be accessed and saved.

from google.colab import drive

drive.mount('/content/drive')

**Step 2: Set File Paths**

Define the path for the raw audio and cleaned audio file.

File\_path = "/content/drive/MyDrive/audio.wave

**Step 3: Inspect Original Audio Properties**

Use the **wave** library to print key properties.

import wave

wf = wave.open(file\_path, "rb")

print("Original File Properties:")

print(" - Channels:", wf.getnchannels())

print(" - Sample rate:", wf.getframerate())

print(" - Duration (seconds):", round(wf.getnframes() / wf.getframerate(), 2))

wf.close()

✔ Output Example:

Original File Properties:

- Channels: 2

- Sample rate: 44100

- Duration (seconds): 12.53

**Step 4: Install FFmpeg**

FFmpeg is used for audio conversion and cleaning.

!apt-get install -y ffmpeg

**Step 5: Clean and Convert Audio**

Convert the raw audio into **mono (1 channel)** and **16kHz** sample rate.

!ffmpeg -i "/content/drive/MyDrive/auido.wav" -ac 1 -ar 16000 "/content/drive/MyDrive/auido\_mono.wav"

file\_path = "/content/drive/MyDrive/auido\_mono.wav"

**Step 6: Verify Cleaned Audio Properties**

Check the cleaned file again.

wf = wave.open(file\_path, "rb")

print("\nCleaned File Properties:")

print(" - Channels:", wf.getnchannels())

print(" - Sample rate:", wf.getframerate())

print(" - Duration (seconds):", round(wf.getnframes() / wf.getframerate(), 2))

wf.close()

✔ Output Example:

Cleaned File Properties:

- Channels: 1

- Sample rate: 16000

- Duration (seconds): 12.53

**4. Results**

* The original audio was **stereo (2 channels) with 44100 Hz sample rate**.
* The cleaned audio is now **mono (1 channel) with 16000 Hz sample rate**.
* Duration remains the same (ensuring no trimming).
* The cleaned file is saved as:
* /content/drive/MyDrive/audio\_clean.wav

**5. Conclusion**

In Milestone 1, we successfully:

* Loaded raw audio from Google Drive.
* Inspected its properties.
* Converted it into a clean format suitable for speech-to-text tasks.
* Verified and saved the cleaned audio file.