Setting Up Packages for Speech Recognition

Step 1: Ensure Your Environment is Active

- Open **Anaconda Prompt** (or **Terminal** if you're using macOS).

- Confirm that you are in the correct environment. You should see the environment name (e.g., (speech_env)) in parentheses.

- If you're not in the correct environment, activate it with:

conda activate speech_env

Step 2: Install Audio Processing and Visualization Packages

1. **Install librosa** (for audio processing):

   o Run the following command in **Anaconda Prompt** (or **Terminal** if you're using macOS):

   pip install librosa

2. **Install SpeechRecognition** (for converting spoken language to text):

   o Enter the following command:

   pip install SpeechRecognition

   o **Note**: Be careful with the capitalization and spacing—SpeechRecognition must be typed as a single word.

3. **Install jiwer** (for evaluating speech recognition accuracy):

   o Type the following command:

   pip install jiwer

4. **Install matplotlib** (for visualizations):

   o Use this command to install:

   pip install matplotlib

5. **Install Google Text-to-Speech (gTTS)** (for converting text into speech):

   o Run this command:

Step 3: Prepare for Whisper Installation

Whisper by OpenAI requires some additional setup to handle multimedia data effectively. Follow these steps to ensure everything is in place:

1. **Install PyTorch**:

   o Go to the PyTorch website and select the **"Get Started"** option.

   o Under **"Start locally"**, configure the following:

      ▪ **Stable** version

      ▪ **Your operating system** (e.g., Windows, macOS)

      ▪ **Package**: pip

      ▪ **Language**: Python

      ▪ **Compute platform**: Select a CUDA version if you have a GPU or select CPU.

   o Copy the provided installation command and paste it into Anaconda Prompt, then press Enter.

2. **Install Chocolatey (Windows only)**:

   o Go to the Chocolatey website and click **Install**.

   o Choose **Individual** installation and copy the provided code.

   o Open **Windows PowerShell** as an administrator, paste the code, and run it.

Note: macOS users can use **Homebrew** instead of Chocolatey for the following steps.

3. **Install FFmpeg**:

   o In Anaconda Prompt (or PowerShell with Chocolatey), type:

   choco install ffmpeg

   o Confirm with "Y" when prompted. FFmpeg is essential for Whisper to handle various audio formats.

Step 4: Install Whisper AI for Speech Recognition

- With the environment activated in **Anaconda Prompt**, install Whisper using the following command:

   pip install -U openai-whisper

- **Troubleshooting**: If you receive an error stating that Whisper isn't found after installation, you can also try running the above pip install command in **Command Prompt** instead of Anaconda Prompt.

Step 5: Verify Whisper Installation

- To confirm Whisper is correctly installed, type:

pip show openai-whisper

  - If installed, this command will display Whisper's details.

You're all set! With all packages installed, you're ready to dive into speech-to-text exercises.