🎙️ Speech Intelligence Application

Overview

The Speech Intelligence Application is a Python-based project that integrates **speech processing and Natural Language Processing (NLP)**. It converts spoken input into text, generates concise summaries, and translates text between languages. The system intelligently switches between **online and offline speech recognition** based on internet availability.

Features

- 🎧 **Speech to Text** (Online & Offline)
- **Abstractive Text Summarization** using Transformer models
- 🎴 **Multilingual Text Translation**
- **Automatic Network Detection**
- 🏓 **Menu-driven Command Line Interface**

Technologies & Models Used

NLP & Deep Learning

- **BART (facebook/bart-large-cnn)** – Text summarization
- **MarianMT (Helsinki-NLP/opus-mt-*)** – Machine translation

Speech Processing

- **Google Speech Recognition API** – Online speech-to-text
- **Vosk (vosk-model-en-us-0.42-gigaspeech)** – Offline speech-to-text

---

Requirements

Create a `requirements.txt` file with the following content:

```
torch
transformers
sentencepiece
SpeechRecognition
vosk
```

```
pyaudio
ping3
```

Installation

Install all dependencies using:

```
pip install -r requirements.txt
```

Additional Setup

- Download the Vosk offline model: `vosk-model-en-us-0.42-gigaspeech`
- Place the model folder in the project root directory

Library Usage

torch

Used as the backend for running transformer-based NLP models.

transformers

Loads and manages BART and MarianMT models for summarization and translation.

sentencepiece

Required tokenizer support for MarianMT models.

SpeechRecognition

Handles microphone input and online speech recognition.

vosk

Provides offline speech recognition without internet access.

pyaudio

Captures audio input from the microphone.

ping3

Checks internet connectivity to switch recognition modes automatically.

How It Works

1. The system checks for internet connectivity.
2. Speech input is captured via microphone.
3. Online or offline speech recognition is selected automatically.
4. The recognized text can be:
5. Displayed as text
6. Summarized using BART
7. Translated using MarianMT

Notes

- Internet is required for first-time model downloads.
- Offline speech recognition requires manual download of the Vosk model.
- PyAudio installation may require OS-specific setup.
- CPU-only systems are supported, but sufficient RAM is recommended.
- Ensure microphone permissions are enabled.

Use Cases

- Voice-based note summarization
- Multilingual speech processing
- Offline-capable voice assistants
- Academic NLP and AI mini-projects

Author

Developed as an NLP and Speech Processing integration project using Python and Transformer models.