AI Powered Real Time Speech Translation for Multilingual Content

Description

This project focuses on building a real-time speech-to-speech translation system that transforms live commentary or spoken content from English and Hindi into more than twelve languages. The core idea is to integrate this solution seamlessly within OTT digital platforms, ensuring that users can enjoy live content in their preferred language without any noticeable delay. By leveraging Azure OpenAI and Azure Speech-to-Text services, the project aims to deliver accurate translations with minimal latency. The result is an inclusive, accessible media experience that can cater to global audiences across diverse linguistic backgrounds.

Tech Stack

- Azure Speech-to-Text for real-time audio recognition.
- Azure OpenAI Services for advanced natural language processing and translation.
- Python for implementing scripts, data preprocessing, and model handling.
- VS Code for experimentation, testing, and iterative development.

Team List

Shashank Raj

S. Venkata Sai Kumar Raju

Alluri Pavani

Bitra Praveeth Sai

Haaswith Sai

Gaurav Chhajer

Swedha

Abichellam

Vattikutti Ajay

S.Mushfira Mehek

How to run the Code

Prerequisites
Python 3.8+ installed and on PATH.
Git installed.
Visual Studio Code (VS Code).
VS Code extensions: Python, Pylance, Jupyter (install from Extensions view).
Azure Speech resource and key (Speech-to-Text): you will need the key and region.

Step-by-step procedure to run the project on your system:
1. Clone the repository using the command:

git clone <repo_link>

2. Upgrade pip and install dependencies

python -m pip install --upgrade pip
pip install -r requirements.txt

3. Configure Azure credentials

AZURE_SPEECH_KEY=your_azure_speech_key
AZURE_REGION=your_azure_region

4. Add **.env** to **.gitignore.**

5. Run Python scripts

Use the integrated terminal to run scripts from src/. Examples:

python src/recognize_once.py
python src/continuous_recognition.py

View milestone1_report.docx