IELTS Speaking Test Simulator Documentation

Overview

This project is an IELTS Speaking Test Simulator designed to allow users to practice their speaking skills for the IELTS test. The tool uses a combination of Flutter for the frontend and Python with FastAPI for the backend, integrated with an LLM (Large Language Model) for generating feedback.

Key Features:

- **Real-Time Speech-to-Text: Converts speech into text using Google Cloud Speech-to-Text API.**
- **Feedback Generation: Provides scores and detailed feedback based on IELTS criteria (Fluency, Grammar, Vocabulary, Pronunciation).**
- **Test Mode: Simulates a full IELTS Speaking Test with three parts.**
- **Practice Mode: Instant feedback for individual responses.**
- **PDF Report: Generates a downloadable PDF of feedback and transcription.**

Source Code Structure

1. Frontend (Flutter)

The frontend application is developed in Flutter and contains the following key components:

Main Components:

- ``
  - **Manages the full IELTS Speaking Test flow.**
  - **Includes real-time audio recording, speech-to-text conversion, and navigation between test parts.**
  - **Final feedback is displayed on a dedicated feedback screen.**
- ``
  - **Displays transcriptions and feedback for all parts of the test.**
  - **Includes a button to generate and download a PDF report.**

Important Flutter Packages Used:

- **google_speech: Integrates Google Speech-to-Text API for real-time transcription.**
- **flutter_sound: Handles audio recording.**
- **permission_handler: Manages microphone permissions.**
- **pdf and printing: Generate and share PDFs.**

2. Backend (FastAPI)

The backend provides RESTful APIs to handle feedback generation and scoring.

Key Components:

- **``: Contains the FastAPI application.**
- **``** Endpoint**: Accepts transcription input, generates feedback, and returns JSON response.**

LLM Integration:

- **The backend uses OpenAI's GPT-4o-mini model to simulate an IELTS examiner.**
- **Prompt Design: The transcription is analyzed using a carefully structured prompt to generate feedback and scores.**

Scoring System

The scoring is based on IELTS criteria:

1. **Fluency and Coherence**
   - **Analyzes logical flow, timing, and pauses.**
2. **Grammar**
   - **Evaluates grammatical range and accuracy.**
3. **Vocabulary**
   - **Assesses the range and appropriateness of words used.**
4. **Pronunciation**
   - **Judges clarity and phonetic accuracy.**

The LLM generates a score out of 9 for each criterion, along with detailed feedback.

APIs Used

1. Google Cloud Speech-to-Text

- **Converts spoken input into text in real-time.**
- **Used in both Practice and Test Modes.**

2. OpenAI GPT-4o-mini

- **Processes the transcription and generates feedback based on IELTS standards, leveraging the lightweight and efficient GPT-4o-mini model.**
- **Highly flexible and customizable prompt design.**

Challenges Faced and Solutions

1. Continuous Audio Recording

- **Challenge: Recording was running indefinitely, causing performance issues.**
- **Solution: Implemented a "Start" and "Stop" mechanism to control recording.**

2. Multiple API Requests During Transcription

- **Challenge: Real-time feedback caused excessive API calls.**
- **Solution: Restricted feedback requests to trigger only after recording stops.**

3. Layout Overflow

- **Challenge: Feedback display caused UI overflow on smaller screens.**
- **Solution: Moved final feedback to a dedicated screen with scrollable content.**

4. PDF Integration

- **Challenge: Formatting transcriptions and feedback into a clean PDF format.**
- **Solution: Used the pdf package with customizable widgets to generate structured PDFs.**