🎧 Voice-Based Sentiment & Intent Classification Assistant

Tools Used: Python, Hugging Face Transformers, SpeechRecognition, SoundDevice, SoundFile, Wavio
Skills Demonstrated: Audio handling, speech-to-text transcription, sentiment analysis, keyword-based classification, zero-shot learning, user interaction flow

---

📌 Project Summary

This project simulates a **voice-based customer service assistant** capable of:

- Capturing and recording user voice
- Transcribing speech into text
- Analyzing sentiment
- Classifying user intent into predefined support categories
- Responding with dynamic audio prompts
- Following up based on user confirmation

This assistant integrates **speech recognition, sentiment detection, rule-based classification, and zero-shot NLP** to create an end-to-end conversational pipeline.

🧠 Key Functionalities

1. Audio Playback & Voice Capture

- Plays an introductory greeting audio using sounddevice and soundfile.
- Records 3 seconds of user voice input using the system microphone via sounddevice.
- Saves recordings as .wav files using wavio.

2. Speech-to-Text Transcription

- Utilizes the SpeechRecognition library with Google Speech API to transcribe user voice into text.
- Handles errors such as unintelligible audio or network issues gracefully.

3. Sentiment Analysis

- Applies the pre-trained MarieAngeA13/Sentiment-Analysis-BERT transformer model to detect whether the user's tone is negative, neutral, or positive.

- If a highly negative sentiment is detected (score > 0.9), flags for immediate operator intervention.

4. Intent Classification

- Uses a keyword-based few-shot classification approach to match transcribed text to one of the following categories:
    - **Account related Enquiries**
    - **Technical Enquiries**
    - **New Enquiries**
- If classification is inconclusive, plays a confused response audio.

5. Follow-up Confirmation with Zero-Shot Learning

- Records a follow-up confirmation response.
- Uses Hugging Face's facebook/bart-large-mnli model in a **zero-shot classification** setup to classify user reply as **affirmative** or **negative**.
- Based on the reply, either confirms the intent or falls back to a confused response.

6. Final Response & Closure

- Plays a closing audio message (Goodbye.mp3) based on final intent determination.

✅ **Outcomes**

- **Accurate transcription and sentiment detection** from user audio
- **Robust handling of unrecognized speech and intent confusion**
- **Dynamic, interactive audio responses**
- Designed to **simulate real-world IVR or smart assistant use-cases**

💡 **Potential Extensions**

- Add support for multiple languages using multilingual models
- Replace keyword intent detection with fine-tuned BERT or RoBERTa classifiers
- Integrate into a Flask or Streamlit web app for real-time use

- Store conversations in a database for analytics or retraining

---

🔗 Use Case Fit

This project is ideal for:

- Call center automation
- Voice bots in telecom and fintech sectors
- Customer feedback classification
- Human-in-the-loop AI systems