### **Report**

#### **Summary**

OpenAI's Whisper (tiny model), the offline Vosk engine (small English model), and the cloud-based Google Speech Recognition API were used. The system, implemented in a Gradio web interface, was assessed across various audio scenarios. Key findings indicate that while the Google API excels in accuracy for clear audio, Whisper provides superior robustness in challenging conditions like background noise and fast speech. Vosk stands out as a reliable and fast offline alternative for standard use cases.

#### **System Overview**

The evaluation platform is an interactive Gradio application that accepts audio via file upload or direct microphone recording.

#### **Key Findings and Recommendations**

* For applications where conditions are ideal (clear speech, minimal noise), the **Google Speech Recognition API** is the recommended choice due to its superior contextual understanding and punctuation.
* For applications involving diverse or unpredictable audio conditions (background noise, varied speech paces, accents), **Whisper** is the most robust solution. Its performance in challenging scenarios is a significant advantage.
* **Vosk** is the ideal choice for on-device, offline applications where internet connectivity is not available or low latency is critical. While less accurate than its online counterparts in adverse conditions, its performance with clear audio is reliable.