Noise-Robust Speech-to-Text Transcription System

Overview:

This project implements a real-time speech-to-text transcription system capable of operating in noisy environments. It captures audio from a microphone, reduces background noise using signal processing techniques, and transcribes the cleaned speech using automatic speech recognition (ASR). This system is useful for voice interfaces, accessibility tools, and customer service applications.

Technologies Used:

Python 3.8+

Speech\_recognition – For speech-to-text conversion

Sounddevice – For capturing audio

Noisereduce – For noise suppression

Numpy / scipy – Signal handling

Optional: torch + torchaudio – For advanced ASR (e.g., Wav2Vec2)

Installation:

Pip install speechrecognition sounddevice scipy numpy noisereduce

For deep learning-based transcription:

Pip install torch torchaudio

Usage:

1. Run the script:

Python noise\_robust\_transcriber.py

1. Speak into your microphone.
2. Wait for transcription results to display on the terminal.

Sample output:

Recording…

Reducing noise…

Transcribing…

Transcription: Hello, I need help with my account.

Conclusion:

This project demonstrates how to build a robust speech recognition system using Python and open-source tools. By integrating noise reduction, it significantly improves transcription accuracy in real-world environments with background interference. The system can be extended for real-time customer support, voice commands, and accessibility enhancements.