Name : Abdul Rehman Section :B

Course : AI&ES Roll No : CT-22052  
---------------------------------------------------------------------------------------------------------------------  
 **Lab # 6**

# Impact of Different Accents or Languages on Transcription

Speech recognition models are often trained on large datasets of specific languages and accents. If an accent deviates significantly from the training data, errors increase. For instance, an American- trained model may struggle with strong Scottish or Indian accents, leading to incorrect transcriptions. Similarly, if a model isn't trained on a language, it may fail to recognize words or interpret them as gibberish.

# Effect of Background Noise on Accuracy

Yes, background noise can significantly reduce accuracy. If there’s constant noise (e.g., fans, traffic), the model may pick up unintended sounds, leading to misinterpretations. If the noise level is high compared to the speaker's voice (low Signal-to-Noise Ratio, SNR), words might be missed entirely.Advanced models use noise reduction techniques, but in extreme cases, noise still degrades performance.

# Performance with Different Audio Files ("Eagle" vs. "Elephant")

Shorter, distinct words (like "Eagle") are easier to recognize than longer words (like "Elephant") if noise or distortion is present. Phonetically similar words (e.g., "Eagle" vs. "Legal") can be misinterpreted. If audio quality is poor or the speaker mumbles, recognition struggles more with complex or less frequent words.

# Recognition Accuracy with Different Voice Characteristics ("Shrill" vs. "Grave")

Models trained on diverse voice datasets generally perform well, but extreme pitch variations (very high shrill voices or deep grave tones) can cause issues. Some models struggle with high-pitched children's voices or very deep voices, leading to missed words or incorrect transcriptions.

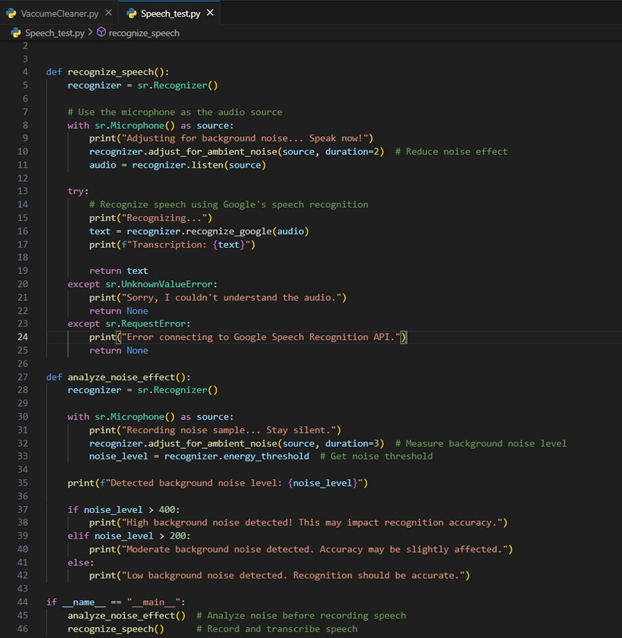
# Self-Introduction and Background Noise Analysis

If I were to introduce myself:

"Hello, this is Abdul Rehman."

In a quiet environment, this would be accurately transcribed. However, if background noise is present:

* + **Mild noise (like a fan)** → Slightly reduced accuracy but mostly correct.
  + **Moderate noise (people talking in the background)** → Some words may be altered.
  + **Loud noise (construction, music, or heavy traffic)** → High chance of missed or incorrect words.



**Output:**