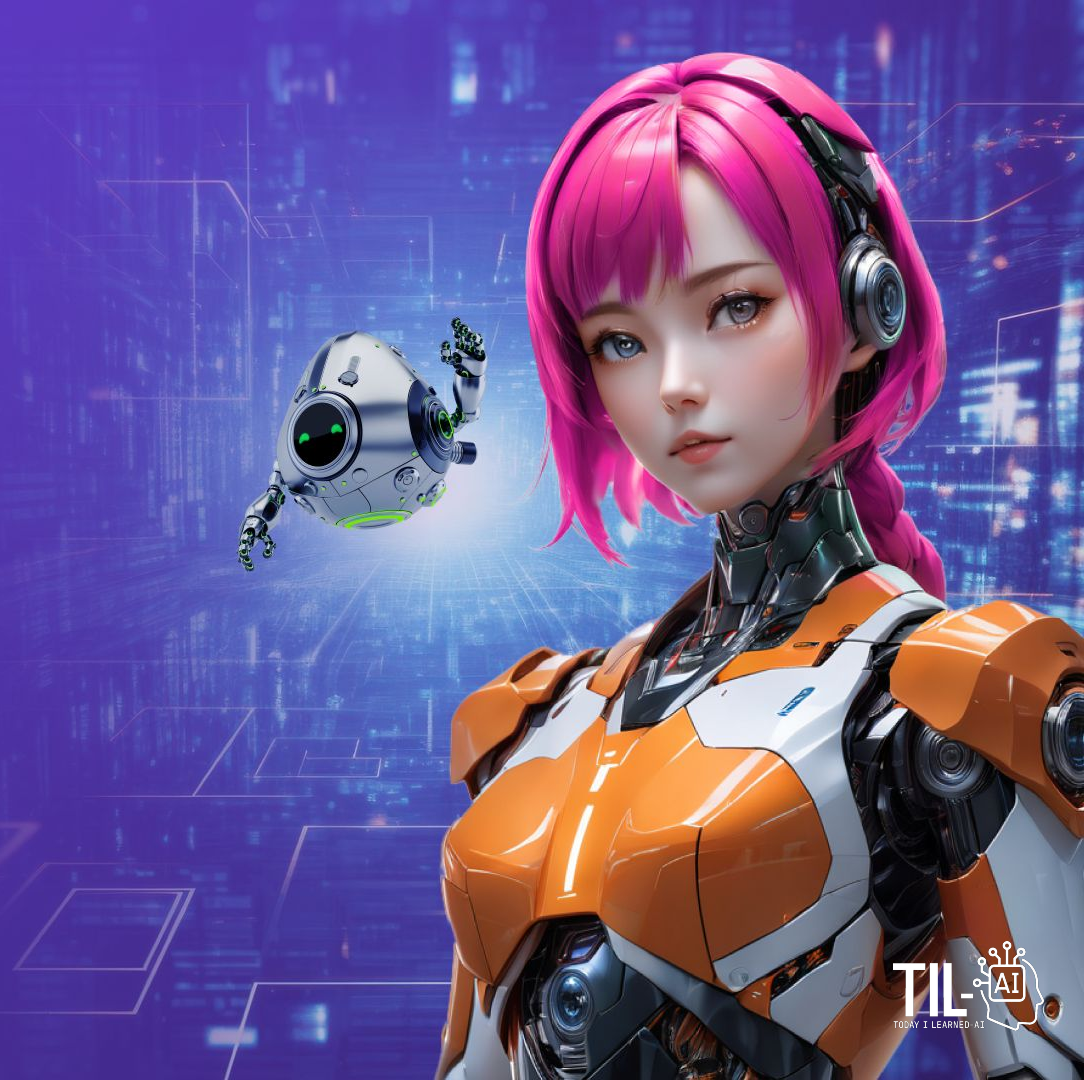
NLP/ASR

Unit 5: Advanced Topics in ASR/NLP



5.2.1

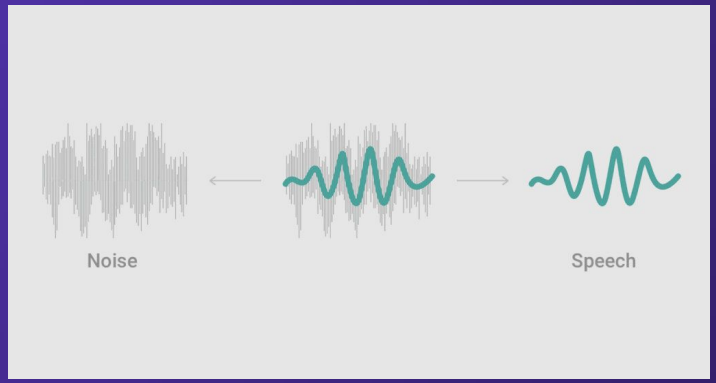
Advanced Audio Processing

Noise reduction and Audio enhancement techniques

TIL-AI
TODAY I LEARNED-AI

Understanding the Problem of Noise

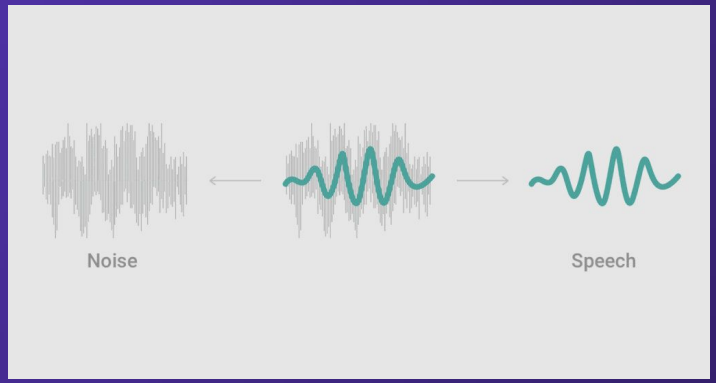
- Noise can come from various sources: background conversations, traffic, machinery, wind, etc
- Noise masks important speech features and disrupts the acoustic signal
- This noise makes it harder for ASR systems to identify words and phrases correctly
- Noise reduction helps minimize background noise without distorting speech



TIL-AI
TODAY I LEARNED-AI

Types of Noise

- Broadband Noise: Spread across a wide range of frequencies (e.g., white noise, fan noise)
- Narrowband Noise: Concentrated in a narrow frequency band (e.g., hum from electrical equipment)
- Impulsive Noise: Short bursts of noise (e.g., clicks, pops)



TIL-AI
TODAY I LEARNED AI

Noise Reduction Techniques

- Spectral Subtraction
    - Estimates the noise spectrum during non-speech segments (silent periods)
    - Subtracts the estimated noise spectrum from the speech spectrum
    - Effective for broadband noise, but it can introduce artifacts if the noise is not stationary or if it overlaps with speech frequencies

Noise Reduction Techniques

- Wiener Filtering
    - More sophisticated approach than spectral subtraction
    - Improves speech clarity by calculating a filter that minimizes the mean squared error between the estimated clean speech and the original noisy speech
    - Adapts to varying noise conditions including non-stationary noise but requires accurate noise estimation

Noise Reduction Techniques

- Deep Neural Networks-Based
    - Neural Networks can be trained to distinguish between speech and noise, enhancing speech recognition accuracy
    - Can learn complex relationships between speech and noise, allowing for highly adaptive noise reduction in various scenarios
    - Example: Deep Denoising Convolutional Neural Network (DnCNN)

Audio Enhancement

- Audio enhancement techniques improve the clarity, intelligibility, and overall quality of speech

- Beamforming:
  - Uses multiple microphones to enhance the audio signal from a specific direction while suppressing others

- Equalization:
  - Adjusts the balance of different frequency bands to improve speech clarity