

SPEECH RECOGNITION

NAME -
SUDHANSU KAKKAR
STUDENT ID -
20036779

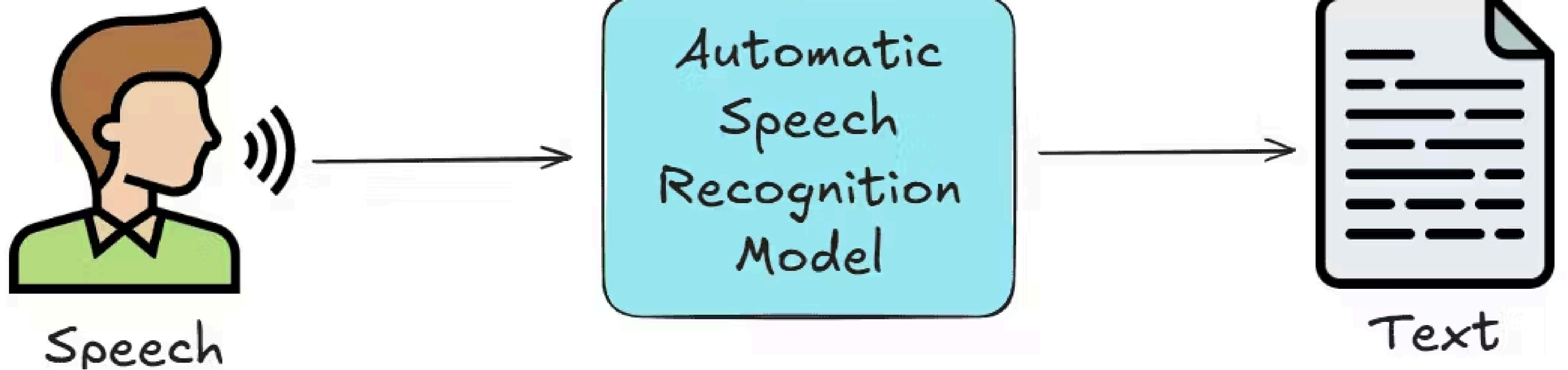
ENABLING MACHINES TO UNDERSTAND AND PROCESS HUMAN SPEECH

INTRODUCTION

Speech recognition is the process of converting spoken language into text using AI-powered systems.

APPLICATIONS

- Virtual assistants (Siri, Alexa, Google Assistant)
- Healthcare (medical dictation, voice-driven diagnostics)
- Customer service (call center automation, chatbots)
- Accessibility (voice control for differently-abled users)
- Transcription (meetings, lectures, media production)



MACHINE LEARNING (ML):

Role: Learns from speech datasets to improve recognition of accents, dialects, and context.

Example: Logistic regression or SVM models used for phoneme classification

DEEP LEARNING (DL)

Role: Neural networks (e.g., RNNs, Transformers) model sequences of speech for high accuracy.

Example: End-to-end systems like DeepSpeech achieving near-human transcription performance.

ARTIFICIAL INTELLIGENCE (AI):

Role: Provides the foundation for speech-to-text pipelines by combining language understanding and signal processing.

Example: AI systems that filter background noise to improve recognition accuracy.

CHATGPT

Role: Processes recognized text into meaningful dialogue, generating natural responses.

Example: Integrated with voice assistants for conversational AI.

LARGE LANGUAGE MODELS (LLMs)

Role: Add contextual understanding to speech recognition, correcting errors and predicting intent.

Example: GPT-4 used to enhance accuracy in domain-specific voice transcription.

