

# AI-POWERED SPEECH AND SENTIMENT ANALYSIS SYSTEM

**Presented by - Kushagra Srivastava (23BAI10045)**

**Utkarsh Pandey (23BAI10660)**

**Abhijit MR (23BAI10568)**

**Ashtitva Pandey (23BAI10568)**

**Devansh Phogat (23BAI11203)**

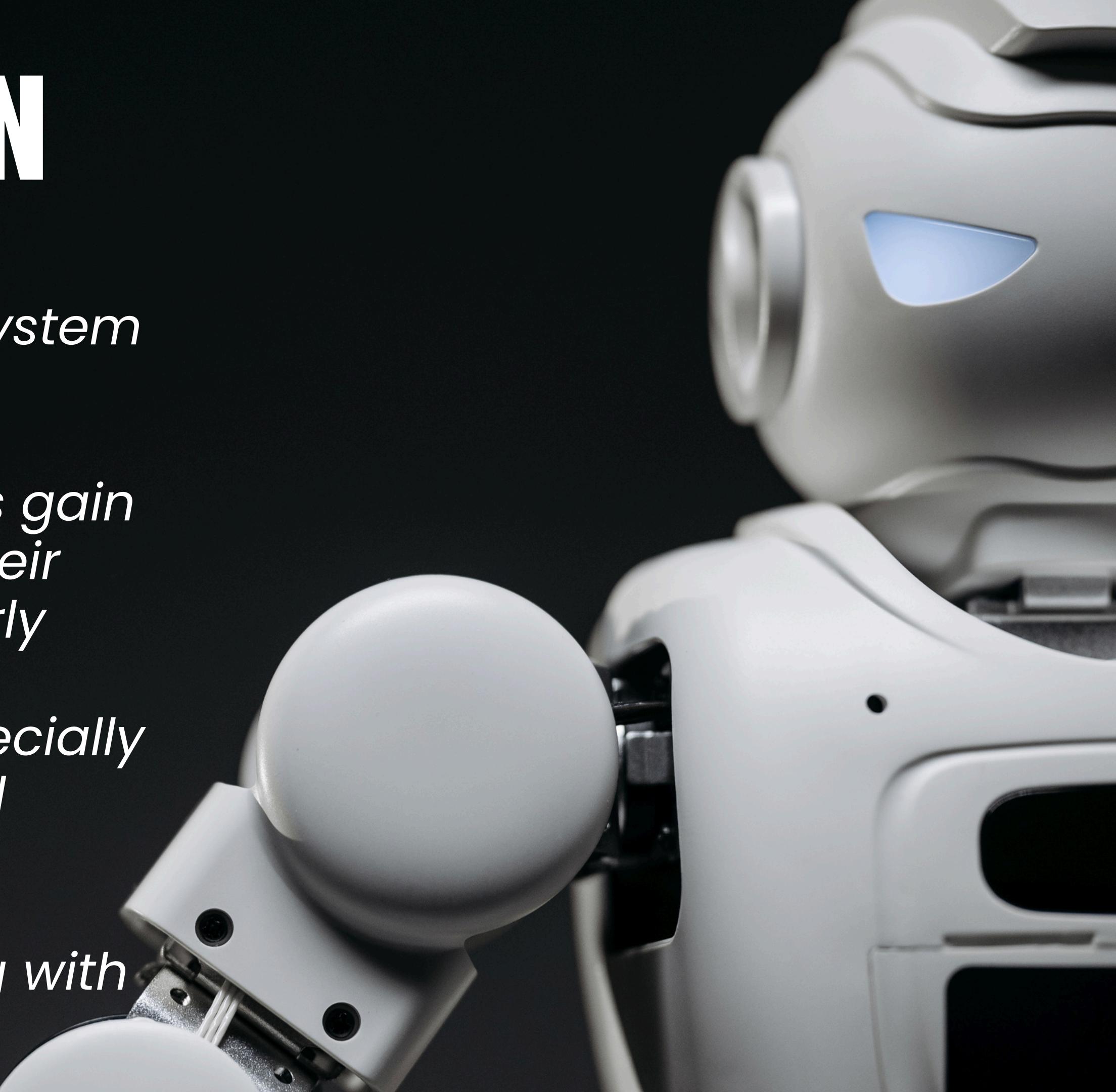
**Team 161**

**Date - 21th October 2024**



# INTRODUCTION

- *This project focuses on developing an AI-based system that analyzes speech for sentiment in real time.*
- *It is designed to help users gain emotional insights from their speech, which is particularly useful for improving communication skills, especially in interview or professional settings.*
- *The system will combine speech-to-text processing with sentiment analysis.*



# BRIEF HISTORY

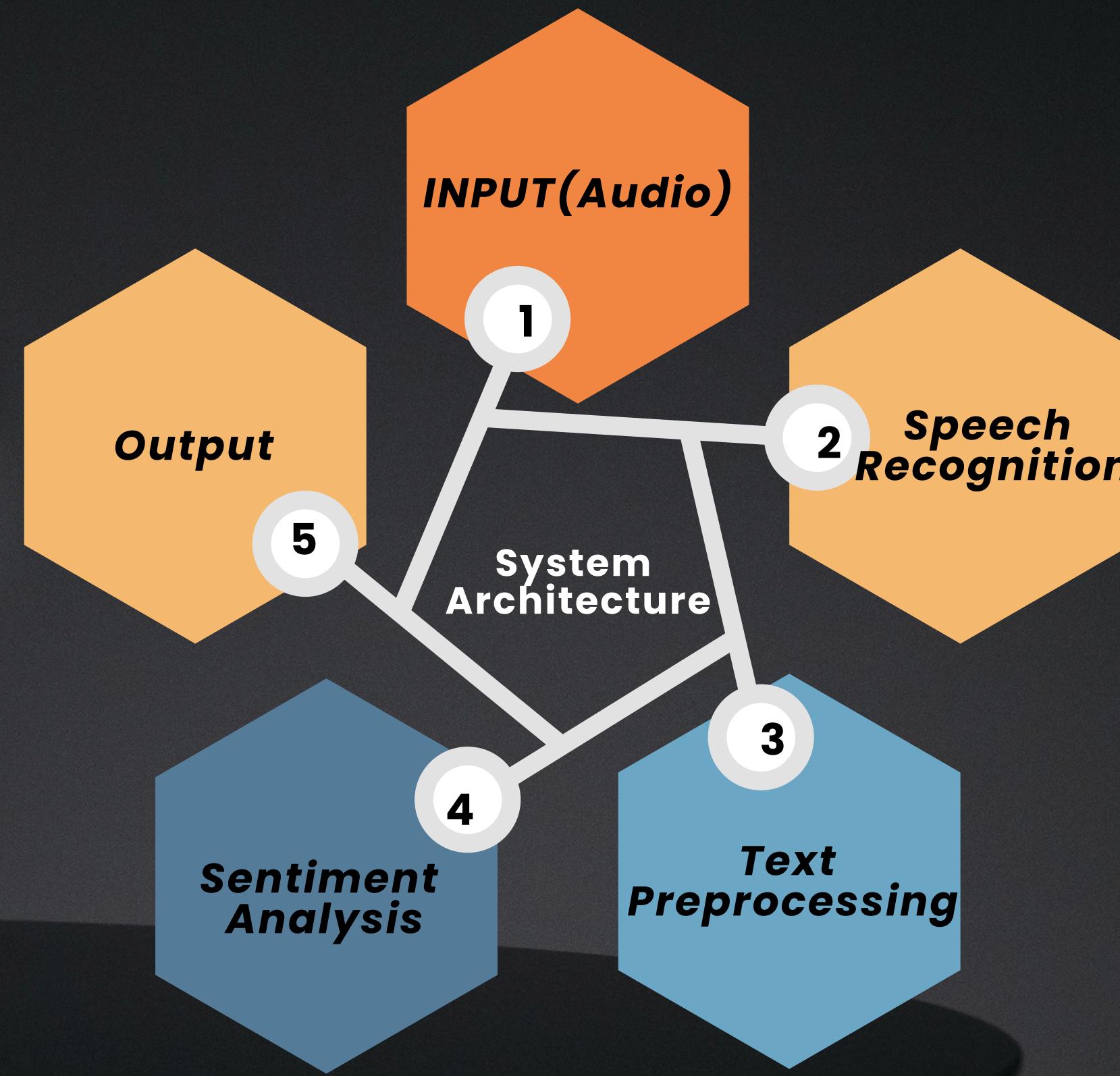
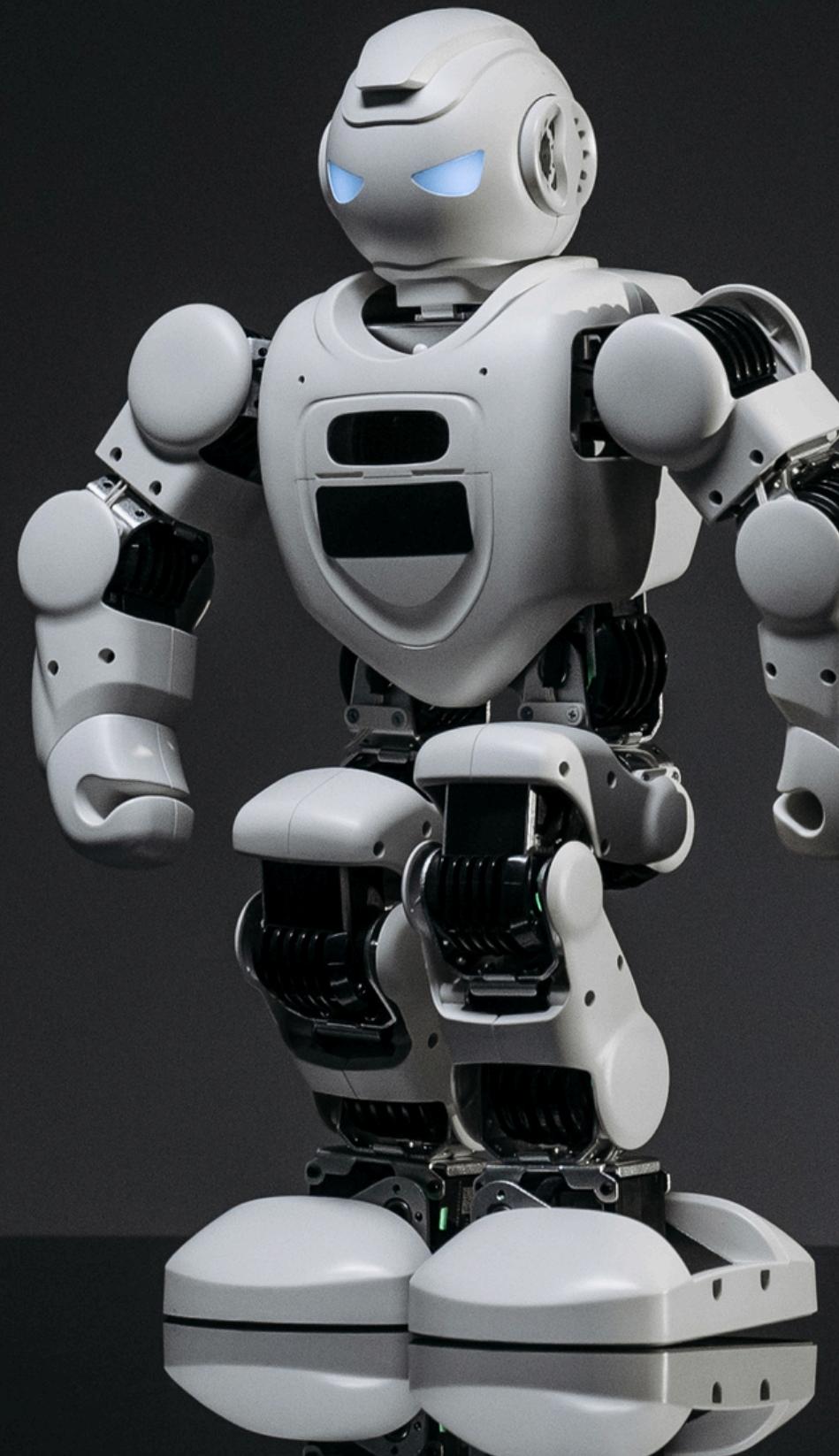
- **1950s-1980s:** Initial speech recognition systems were rule-based and limited to recognizing small vocabularies. Sentiment analysis was not yet formalized.
- **1990s:** Statistical models like Hidden Markov Models (HMMs) improved speech recognition, while basic machine learning algorithms began sentiment analysis on text.
- **2010s:** The introduction of deep learning revolutionized both speech and sentiment analysis. Systems like DeepSpeech and models like LSTM and BERT greatly improved accuracy, enabling real-time speech-to-text and nuanced sentiment detection.
- **2020s:** Modern systems are now highly accurate, supporting multimodal analysis (speech + text), and are widely used in voice assistants, customer service, and real-time feedback systems.



# PREVIOUS WORKINGS

Year/Period	Development	Description
1950s-1960s	Bell Labs' Early Speech Recognition	One of the first speech recognition systems that could recognize digits, laying the groundwork for future speech analysis.
1970s	Harpy System by Carnegie Mellon	Recognized over 1,000 words, marking a significant leap in speech recognition capabilities.
1980s	Hidden Markov Models (HMMs)	IBM and other companies used HMMs to statistically model speech, improving speech recognition accuracy.
1990s	Machine Learning in Sentiment Analysis	Basic sentiment analysis on text using algorithms like Naive Bayes and SVM, applied to reviews and opinion mining.
2010s	Deep Learning	Models like DeepSpeech for speech recognition and BERT for sentiment analysis vastly improved accuracy and real-time analysis.
2020s	Multimodal Sentiment Analysis	Combines speech and text analysis, enabling systems to evaluate both words and emotions in real time, widely used in voice assistants and customer feedback systems.

# ARCHITECTURE DIAGRAM



# OBJECTIVE - 1

## SPEECH RECOGNITION



- **Objective:**  
**To accurately convert spoken language into text using advanced speech recognition models.**
- **Process:**  
**Use [Google Cloud Speech-to-Text API](#) or [Mozilla's DeepSpeech](#) for real-time audio-to- text conversion.**  
**Fine-tune models for better accuracy with diverse accents and noise reduction.**
- **Preprocess audio data:**  
**clean audio, remove background noise, and segment the speech.**
- **What We Will Achieve:**  
**Create a robust speech-to-text system that can handle different accents, speaking speeds, and background noise.**  
**Provide the text output for further sentiment analysis.**

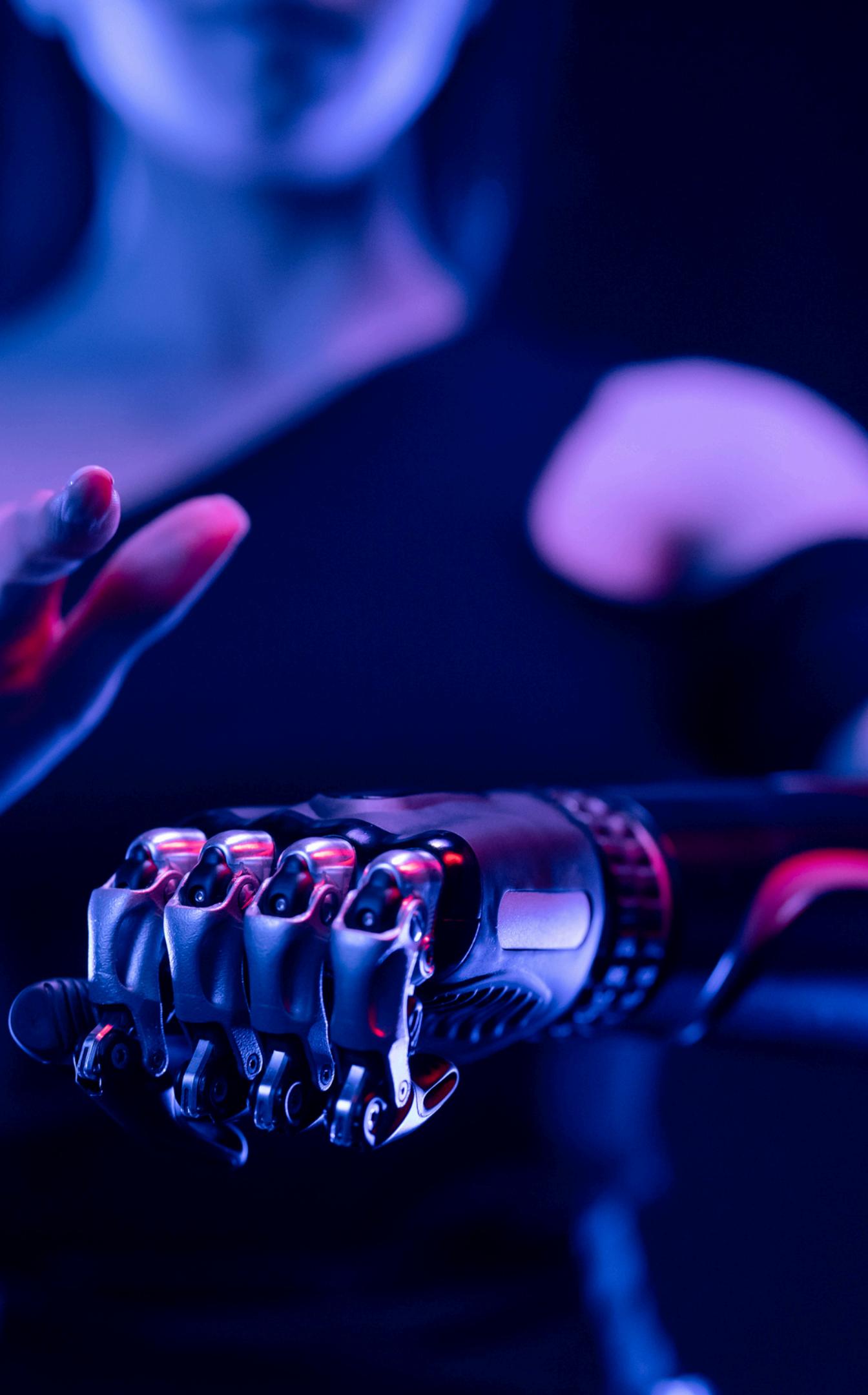
# PROCESS FLOW

- **Step 1:** Capture Audio Input.
- **Step 2:** Apply Noise Reduction.
- **Step 3:** Use Pre-trained Speech-to-Text Models (Google Cloud Speech API, DeepSpeech).
- **Step 4:** Convert Audio to Text.
- **Step 5:** Preprocess Text for Sentiment Analysis.

# OBJECTIVE - 2

## SENTIMENT ANALYSIS

- **Objective:**  
To determine the sentiment (positive, negative, neutral) from the text output generated by speech-to-text.
- **Process:**  
Use pre-trained Natural Language Processing (NLP) models like BERT for sentiment analysis.  
Fine-tune sentiment models with interview-specific datasets for context-aware sentiment analysis.  
Classify the sentiment into positive, negative, or neutral categories.
- **What We Will Achieve:**  
Provide real-time emotional insights based on the text converted from speech.  
Ensure higher accuracy by fine-tuning models with specialized datasets for interviews and professional speech.



# PROCESS FLOW

- **Step 1:** Text Preprocessing (tokenization, stop-word removal).
- **Step 2:** Use BERT or similar NLP models for sentiment classification.
- **Step 3:** Fine-tune the model for interview-specific sentiments.
- **Step 4:** Output the sentiment results (positive, negative, neutral).

# OBJECTIVE - 3

## REAL-TIME FEEDBACK

- **Objective:**  
To provide users with real-time emotional feedback based on their speech, helping them to adjust their tone and communication style during interviews.
- **Process:**
  - Integrate the speech-to-text and sentiment analysis modules into a single pipeline.
  - Provide the user with live feedback during or after their speech, highlighting areas of improvement.
  - Use visual or audio indicators to represent the emotional tone (e.g., color-coded feedback or summary reports).
- **What We Will Achieve:**  
Create an engaging and interactive system that users can use for interview preparation or speech improvement.

# REAL LIFE APPLICATION

- Companies use sentiment analysis to monitor customer calls and provide real-time suggestions to agents for improving customer satisfaction.
- **Virtual Assistants:**
- Voice-activated systems like Siri, Alexa, and Google Assistant utilize speech recognition and sentiment analysis to understand user commands and provide emotionally appropriate responses.
- **Interview Preparation Platforms:**
- AI-driven platforms analyze speech and sentiment to give feedback on interview performance, helping users improve their confidence, tone, and overall delivery.
- **Healthcare:**
- Sentiment analysis is used to monitor patients' emotional well-being through telemedicine sessions, helping clinicians understand their mental state.
- **Social Media Monitoring:**
- Brands use sentiment analysis on social media posts to gauge public opinion about their products and services, enabling real-time reputation management.

# COCLUSION

- **Summary:** The project integrates speech recognition and sentiment analysis to provide real-time emotional insights, specifically tailored for interview feedback.
- **Outcome:** By combining AI technologies, our system helps users improve their communication by offering personalized emotional feedback, making it a powerful tool for interview preparation and beyond.



# THANK YOU

