

DEPARTMENT OF COMPUTER APPLICATION

MSE2 PROJECT – Speech to text converter

Submitted by :

Riya Gaur [202410116100171]

Priyanshi[202410116100153]

Riya [202410116100170]

Pallavi [202410116100139]

Introduction

- Speech recognition technology allows a computer to understand and convert spoken language into text.
- Our project is a basic speech-to-text converter built using Python, which:
 - Takes an audio file as input
 - Processes it using Python libraries
 - Returns readable text output
- This improves human-computer interaction and helps with accessibility and fast transcription.



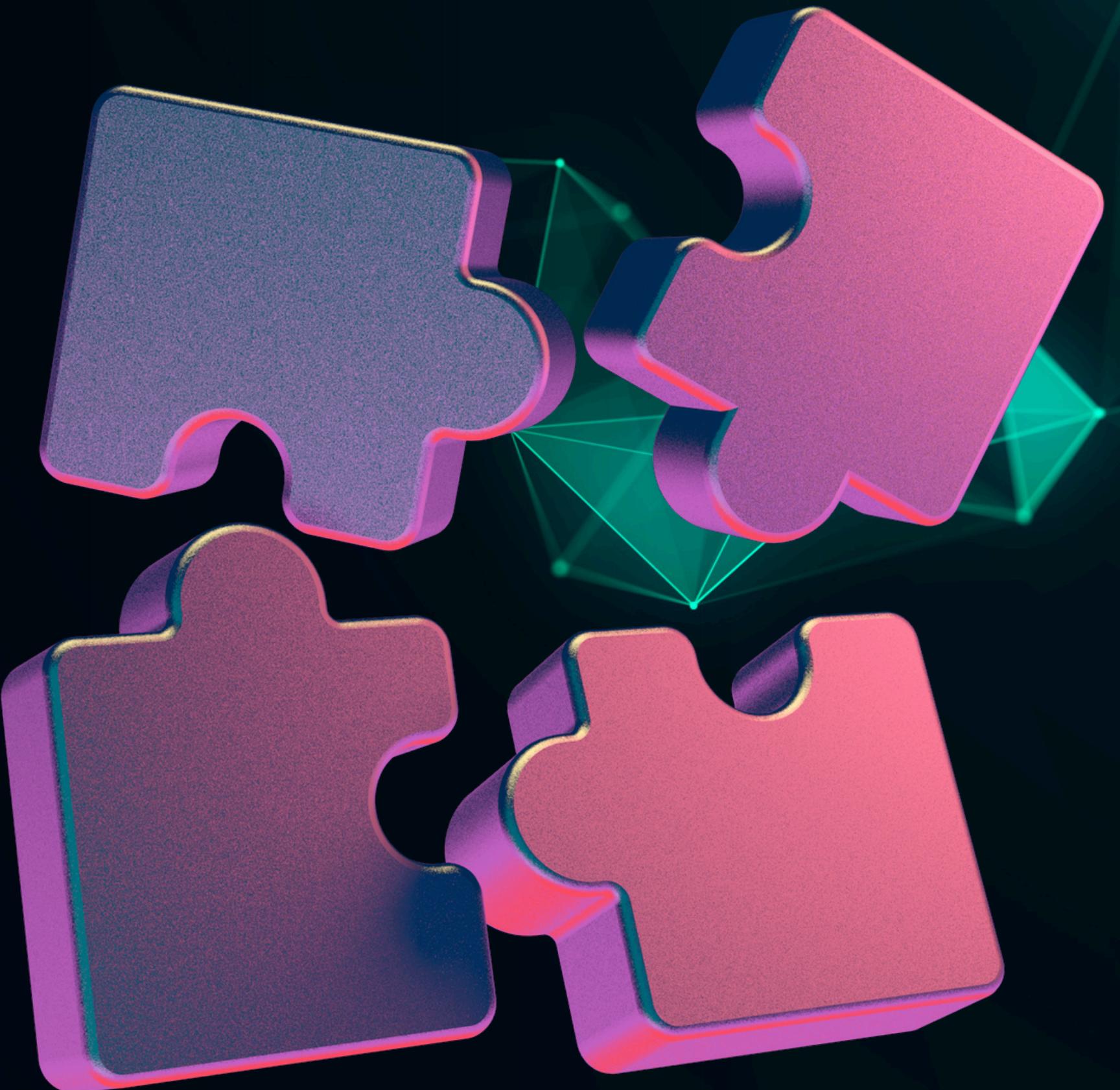


Problem Statement

- Manual transcription is time-consuming, especially:
 - For long lectures, meetings, or interviews
 - For people with disabilities
 - When automation is required
- A Python-based tool can automate this task and improve efficiency

Objectives

- Convert audio speech to text automatically.
- Use Python and open-source libraries.
- Demonstrate API integration and real-world AI concepts.
- Educational tool for voice tech learner.



Technology Stack / Tools Used



- Python – Core programming language .
- SpeechRecognition – Speech to text using Google API .
- pydub – Convert MP3 to WAV.
- IPython.display – Play audio in Colab.
- Google Colab – Cloud-based platform.
- Google Speech API – Converts speech to text .

System Architecture / Workflow

1. Upload Audio File (.mp3 or .wav)
2. Playback using IPython.display.Audio
3. Convert MP3 to WAV using pydub (if needed)
4. Run SpeechRecognition with Google API
5. Display recognized text
6. Handle errors: unclear audio or API/internet issues.



Implementation (Without Code)

- Upload .mp3/.wav file in Colab
 - Play audio for verification
 - Convert MP3 to WAV (if required)
- Run SpeechRecognition to convert speech to text
 - Display clean, readable transcription
- Handle unclear audio or no internet error

Features

- Accepts .mp3 and .wav formats
 - Audio playback before transcription
- Uses Google API for speech recognition
- Error handling for unclear audio or no internet
- Runs on Google Colab – no installation needed
 - Beginner-friendly and customizable



Applications

- Transcription Tools – lectures, meetings
- Accessibility – help for disabled users
- AI Voice Assistants – base logic of Alexa/Siri
- Smart Homes – voice-controlled devices •

Machine Learning – train NLP models



Limitations & Future Scope

- **Limitations:**

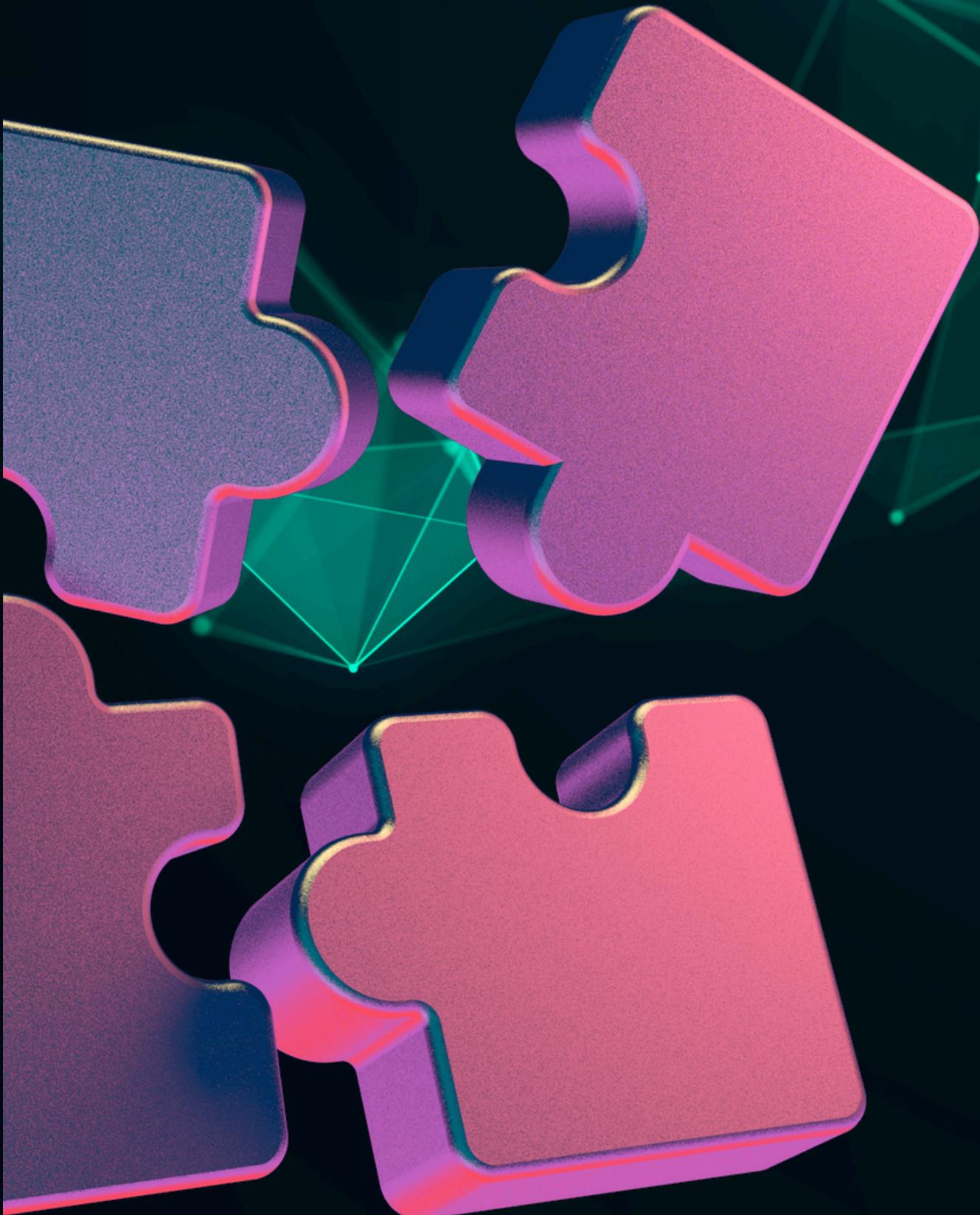
- Needs internet (Google API)
- No live mic input
- Limited language support

- **Future Scope:**

- Add real-time (mic) input
- Support multilingual transcription
- Add GUI/voice commands

Conclusion

- Demonstrates real-world AI use with Python
- Simplifies speech technology for students
- Teaches API integration, NLP basics
- Strong foundation for voice-based applications



Thank You

- Thank you for your attention!
- We'll be happy to answer any questions

