Multilingual Meeting Summarizer

Presented by Team No 41

Motivation for Summarization

Enhancing Collaboration Across Language Barriers

In each and every meeting, prescription, or sincere conversation, something is worth saying but not necessarily recalled. Decisions are lost in prolonged recordings. Directions get fuzzy with the passage of time. And what should have been a moment of insight turns into a lost opportunity.
This project was conceived out of a mere but strong idea: "Let spoken words turn into organized memory."
By creating a speech-to-text summarizer that pulls actionable information out of meetings and doctors orders, we intend to close the gap between what happens to be said and what really needs to be done. It's not transcription–it's comprehension, structuring, and enabling.
From boardrooms to hospital rooms, this tool listens with accuracy and replies with clarity so that all voices lead to action

Challenges in Summarizing

Addressing Multilingual Meeting Complexities

Language Diversity

The wide range of languages spoken can lead to misunderstandings and omissions in key information, making it difficult to create coherent summaries for all participants.

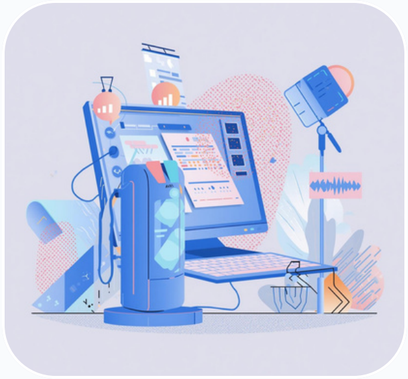
Accurate Summarization Challenges

Large meetings and discussions may have long audio summarization models struggle to capture all key points. Ensuring summaries are concise but not missing critical info is difficult.

Speaker identification Difficulties and unstructured Noisy input

Accurately detecting who said what in real-time meetings is hard when voices sound similar.Overlapping conversations make diarization less reliable.Raw transcriptions contain filler words, grammar errors, repetitions, and broken sentences.

Related Works

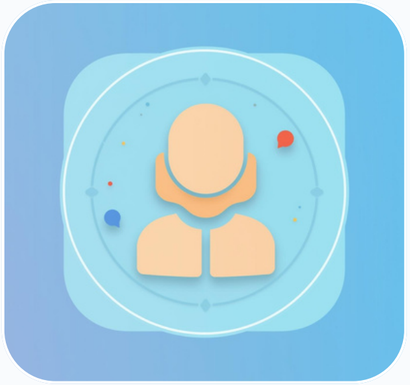


Transcription Focus

Traditional ASR systems only handle transcription tasks.

Limited Multilingual Support

Current tools struggle with true multilingual conversation support.



End-to-End Multilingual Assistant

There's a need for an integrated multilingual solution.

English Datasets Focus

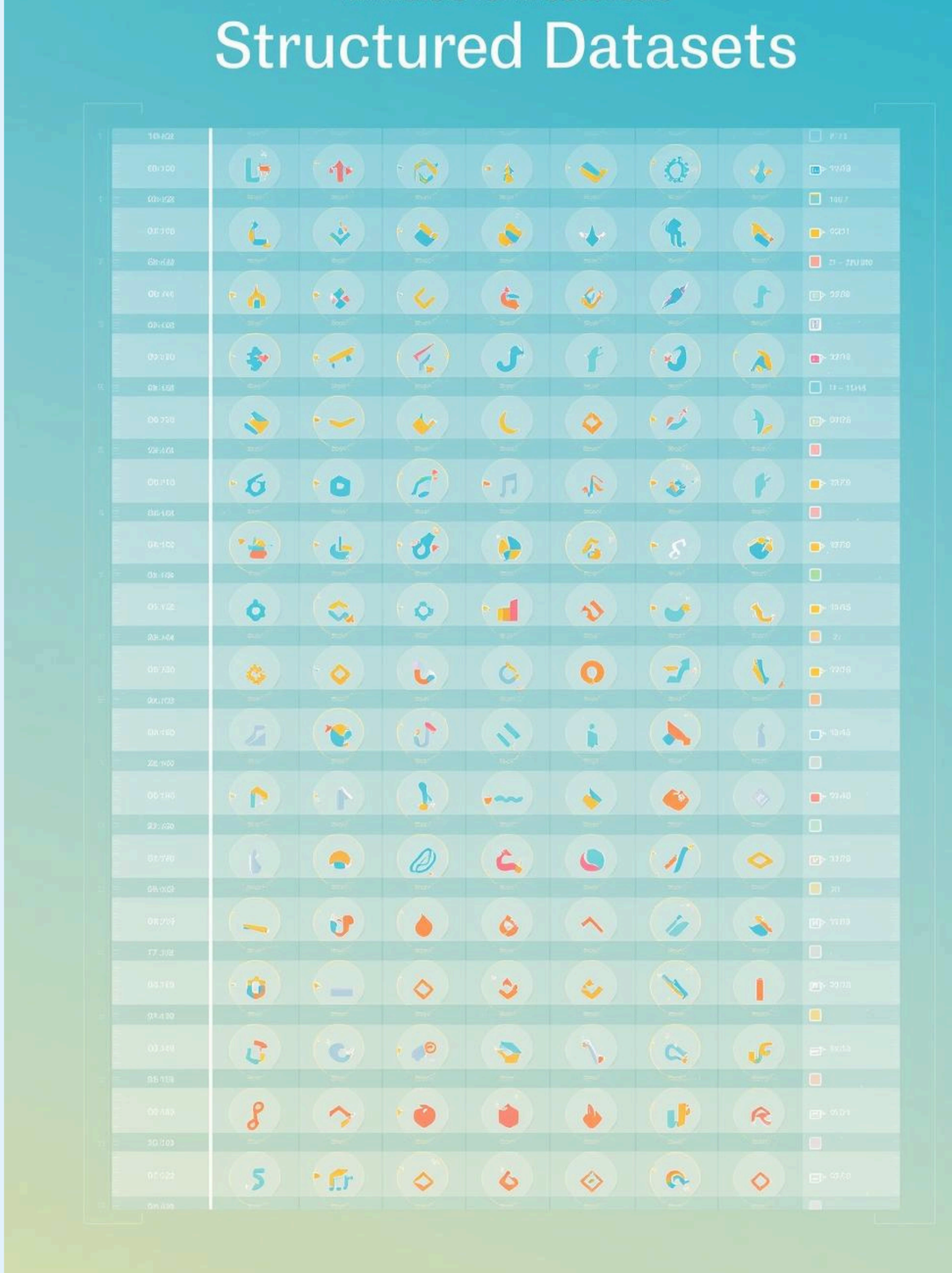
Most meeting summarization research relies on English datasets.

Raw Text Output Only

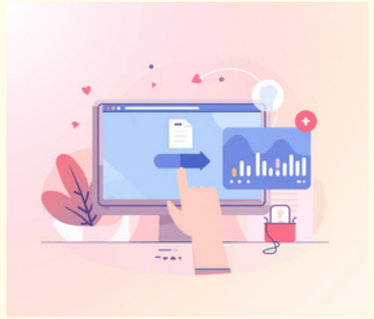
Existing tools provide only unformatted raw text output.

Datasets Used

- AMI Data Corpus
- Microsoft Hinglish Speech - OpenSLR 94
- IndicVoices



Methodology



Audio Handling

Upload and convert audio files efficiently.

Speech Recognition

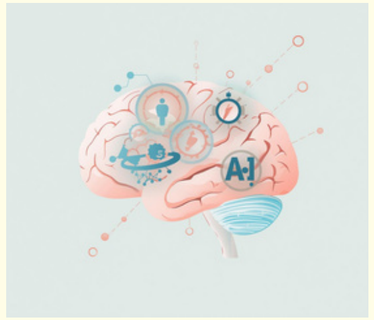
Convert speech to text seamlessly.

Language Detection

Identify language with high accuracy.

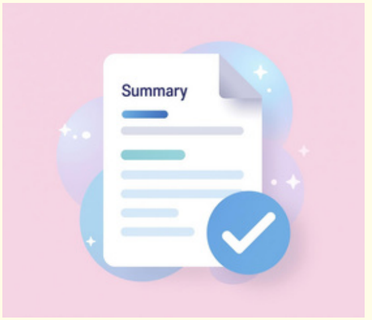
Translation Process ASR Whisper

Translate content conditionally to English.



Summarization

Generate concise summaries using AI.



Output File

Produce a summary text file and Actionable Tasks

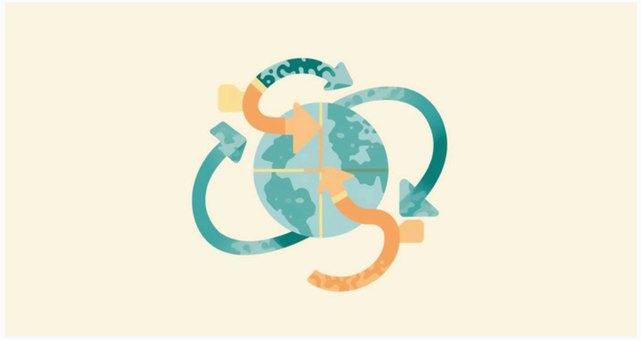
Experiments

Experimental Setup

The pipeline transforms audio into summarized text effectively.

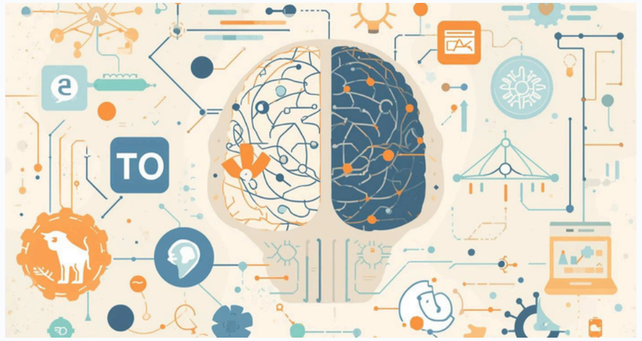
Tools & Libraries

Utilizes various libraries for audio and language processing.



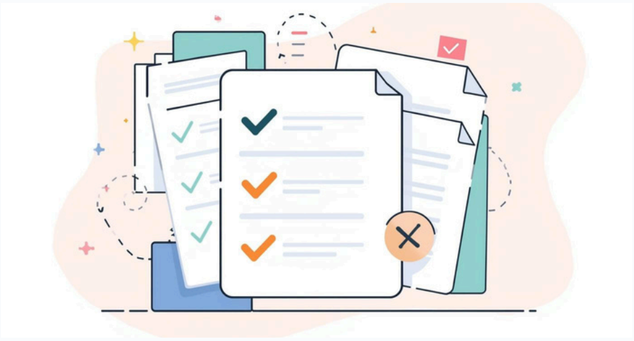
Challenges & Solutions

Addressed issues through effective tools and configuration fixes.



Key Observations

Notable successes and challenges identified during the experiments.



Future Work

Focus on enhancing ASR models and features for improvement.

Additional Improvements

Addressing future challenges for better overall performance.

Results

Summary Text

The final summarized meeting text captures key points, ensuring clarity and coherence while providing a concise representation of the discussions held during the multilingual meeting.

Language Analysis

Comprehensive analysis of speech-to-text output revealed effective language detection across multiple languages, including English, Indonesian, Portuguese, and Norwegian, enhancing overall understanding of meeting content.

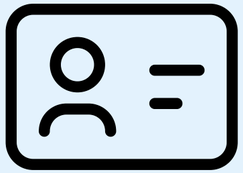
Translation Performance

The translation process integrated non-English segments seamlessly, facilitating effective communication and maintaining the integrity of the original content, contributing to improved summary quality.

Quality Evaluation

Evaluation of summary quality highlighted strengths in clarity, coverage, and coherence, indicating that the multilingual summarization pipeline effectively conveyed essential meeting details to participants.

Identified Limitations



Notable limitations include the absence of speaker IDs and timestamps, which could impact the referencing of specific contributions and the overall organization of the summarized content.

Visual Enhancements

Suggested visuals for enhancement include a language chart and a word cloud, which would provide additional insights and engage the audience with compelling graphical representations of the data.

Insights Analysis

Understanding Summaries' Impact and Relevance

Key Findings

The analysis reveals significant insights into **multilingual communication**, highlighting how effective summarization can bridge language barriers and improve collaboration among global teams.

User Feedback

Collecting user feedback has shown that **summarization accuracy** directly influences user satisfaction, underscoring the importance of refining our model to meet diverse linguistic needs.

Contributions

Aditya

- Collected, tested, and organized multilingual meeting audio datasets.
- Assisted with text preprocessing and documented the workflow.
- Helped test outputs, refine summaries, and provide feedback.

Meghana

- Built text-processing, action extraction, and NER-based task modules.
- Integrated summarization models into the pipeline.
- Planned the system's multilingual expansion.

Akshay

- Built the Whisper-based Speech-to-Text and English summarization modules.
- Helped integrate preprocessing steps and debug the NLP pipeline.
- Supported dataset testing and refining the final workflow

Project Timeline