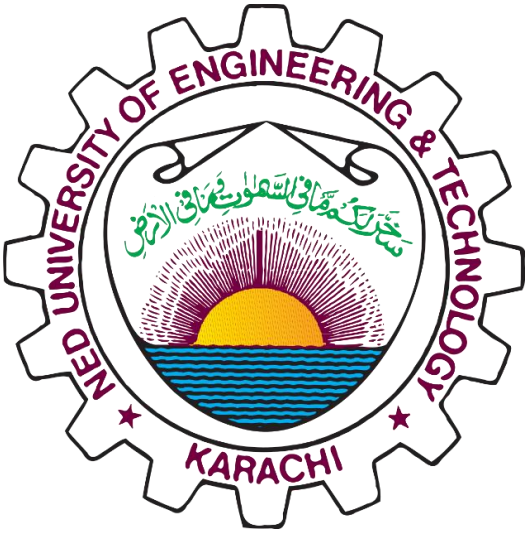


Department of Computer Science and Information Technology

Project Report

Course: Artificial Intelligence & Expert System (CT-361)
Submitted To: Sir. Abdullah Siddique

Group Members:

1. **Saim Fayyaz (CT-22031)**
2. **Umair Lafir (CT-22050)**
3. **Sherial Hassan (CT-22046)**

Contents

AI-Powered Meeting Transcription and Summarization System

Introduction

Meetings are essential in today's collaborative work culture but often lead to unorganized takeaways, missed action points, and wasted time reviewing lengthy recordings. The **AI-Powered Meeting Transcription and Summarization System** addresses this by using speech-to-text and generative AI technologies to transcribe and summarize meetings in real-time. This mobile/web application enables asynchronous meeting access, improves productivity, and ensures that all key information is retained and shareable.

Objective

To build an AI-powered productivity tool that allows teams to:

- Get real-time transcription of meeting conversations.
- Automatically generate structured summaries including agendas, discussion points, and conclusions.
- Allow remote participants to access meeting content post-session.
- Eliminate manual note-taking and improve knowledge transfer across distributed teams.

Overview

This system is designed to:

- Record and transcribe meetings using speech-to-text AI.
- Generate real-time or post-session summaries via GPT-based models.
- Allow meeting hosts to create sessions and share public links for remote access.
- Enable secure storage and asynchronous access to meeting content.

The application leverages **FastAPI** for backend processing, **React Native or Next.js** for frontend development, **Whisper by OpenAI** for transcription, and **GPT** for summarization. Data is stored securely using **MongoDB**, ensuring scalability and performance.

Key Features:

- Google Sign-In Authentication (via Clerk)
- Create and Host Meeting Rooms with shareable public links
- Real-time Transcription using Whisper or DeepSpeech
- Automatic Summary Generation via GPT with sections:
    - Agenda
    - Discussion Points
    - Conclusion
- Transcript Storage & Display post-meeting
- Recorded Sessions Archive for viewing and exporting summaries
- Cross-Platform Access (mobile and web)

Functional Specifications

User Authentication

- Implemented using Clerk for secure login.
- Supports OAuth via Google for easy access.

Meeting Creation & Participation

- Host can create a meeting room and share public links.
- Participants can join via web or mobile interface.

Audio Recording & Processing

- Meeting audio is captured and sent to the backend.
- FFmpeg converts the stream into MP3 format.
- Whisper transcribes audio in real-time or batch mode.

Summary Generation

1) The transcript is processed by OpenAI's GPT-3.5 Turbo.

2) Generates:
    a) Agenda
    b) Discussion Points
    c) Conclusion

Storage & Access

- All recorded meetings and summaries are accessible via a dashboard and are retrieved from Stream.io Library.

User Interface Design

- Developed using **TailwindCSS** for sleek and responsive design.
- **Radix UI** ensures accessibility in dropdowns/dialogs.
- Icons from **FontAwesome**.

Key Screens:

- **Home/Dashboard**: Start or join a meeting.
- **Meeting Room**: Host and transcribe meetings.
- **Recordings Page**: View and manage previous sessions.
- **Summary View**: Display structured outputs from OpenAI.

Technology Stack Used

System Architecture

The system follows a modular architecture with separation of concerns:

- **Frontend Layer:** Built using Next.js or React Native, handling all UI/UX interactions.
- **Backend Layer:** Developed with FastAPI, managing APIs, data flow, and integration.
- **AI Processing Layer:** Speech-to-text transcription (Whisper) and summarization (GPT-3.5) modules.

- **Media Processing Layer:** FFmpeg handles audio conversion for transcription.

Frontend:

- Next.js 14
- Clerk for authentication
- Stream-io Library for Enabling Meetings

Backend:

- **FastAPI** for API handling
- **Whisper** for transcription
- **OpenAI GPT-3.5** for summarization
- **FFmpeg** for audio processing

Workflow of the Meeting Summarization System

- User logs in via Google OAuth.
- Host creates a meeting room.
- Participants join the meeting via link.
- Audio stream is recorded and sent to the FastAPI backend.
- FFmpeg converts audio to MP3.
- Whisper transcribes the audio.
- Transcript is sent to GPT-3.5 for summarization.
- Users can view summaries and transcripts in the recording tab.

Future Scope and Enhancements

- Cloud Storage Integration (e.g., AWS S3 or Firebase)
- Action Item Extraction alongside summary points
- Searchable Transcripts across all sessions
- User Roles & Permissions for Hosts, Viewers, Admins
- Automated Email Reports after each meeting
- Multi-language Support for global organizations
- Native Mobile App Release for Android & iOS

Industrialization Aspect

The app has immense potential for enterprise adoption. It can be packaged as a SaaS solution targeting:

- Corporate teams for meeting automation
- Educational institutions for lecture archiving
- Media firms for interview transcriptions
- Customer support centers for QA purposes

Monetization can be achieved via subscription models or integrations with tools like Zoom, Google Meet, and Microsoft Teams.

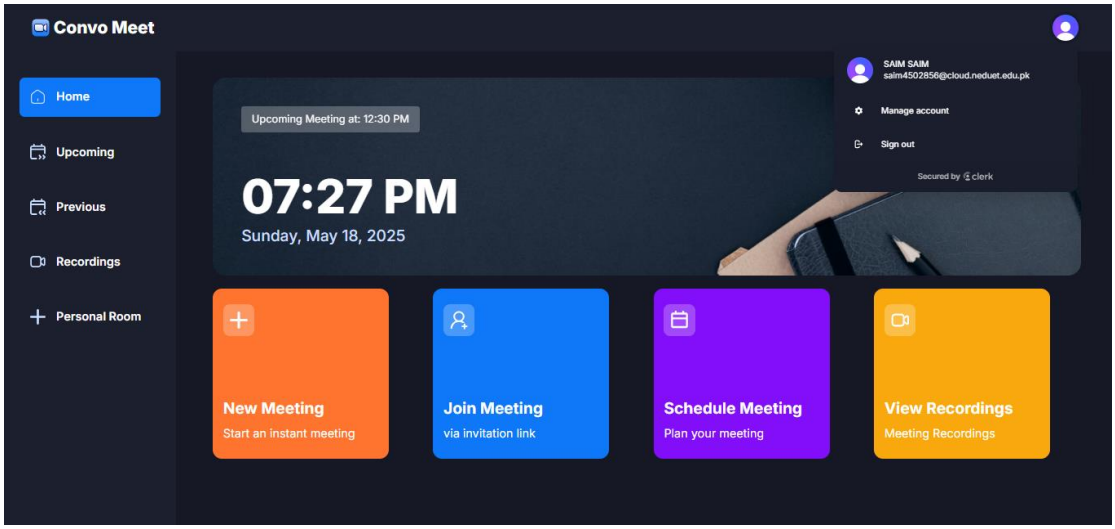
Limitations

- Accuracy may vary with noisy environments.
- Summarization depends on model limitations.
- Real-time transcription may lag slightly on low-performance devices.

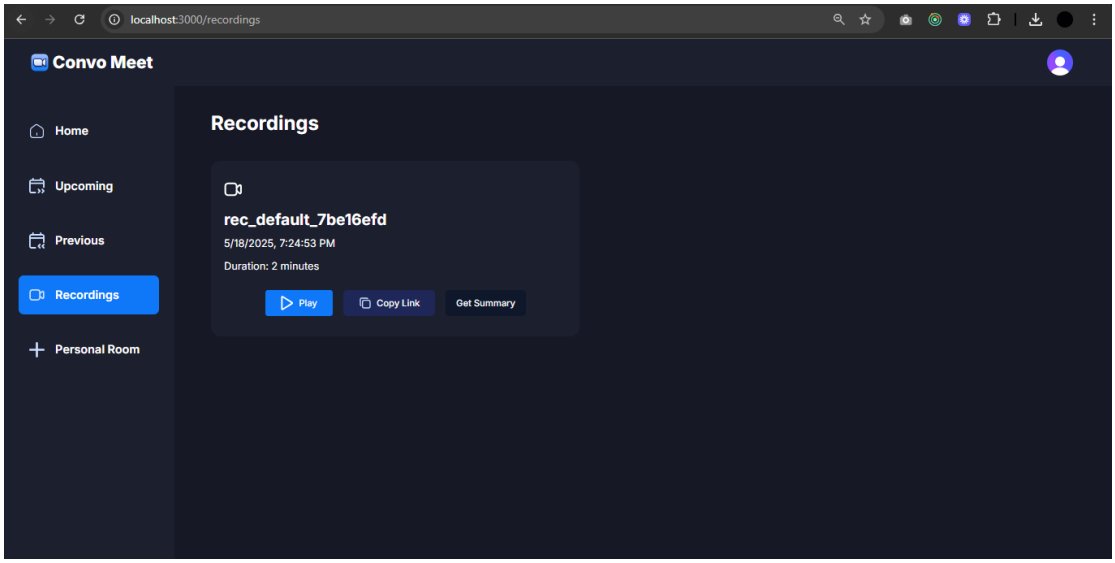
Recommendations

- Improve transcription through custom noise-filtering layers.
- Fine-tune GPT model on domain-specific meeting data.
- Add offline meeting support with delayed sync.

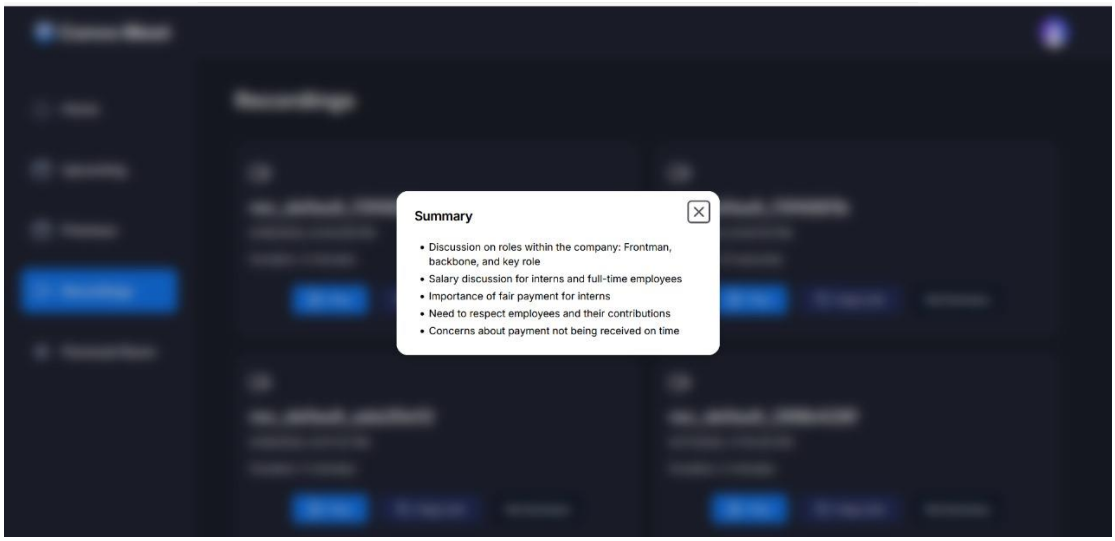
Project Visuals



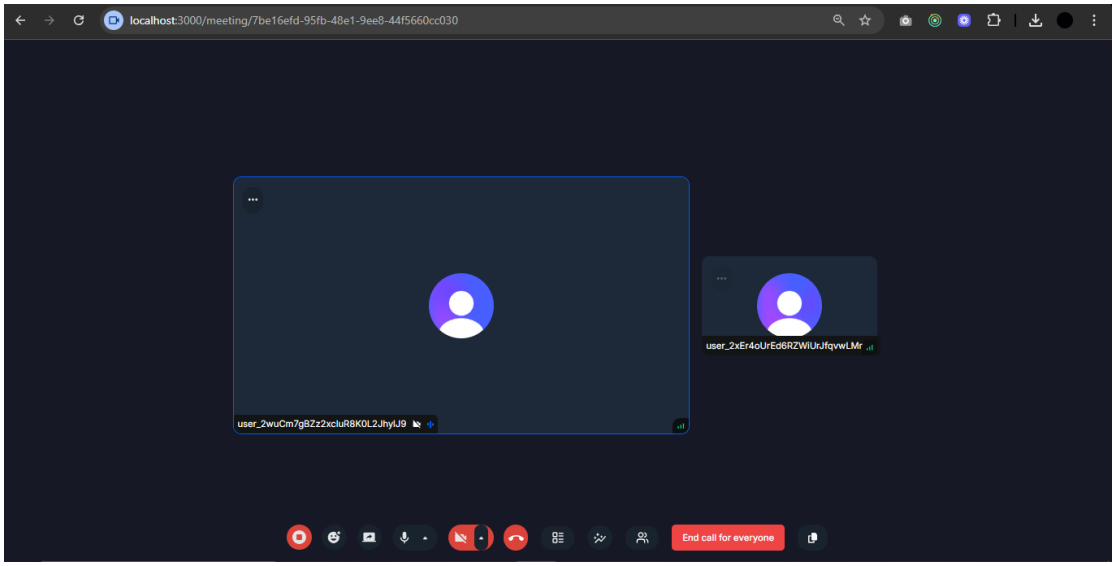
This is the main Home Page of our Product which enables users, different options for meetings, the user can make an instant meeting, join any meeting via link, etc.



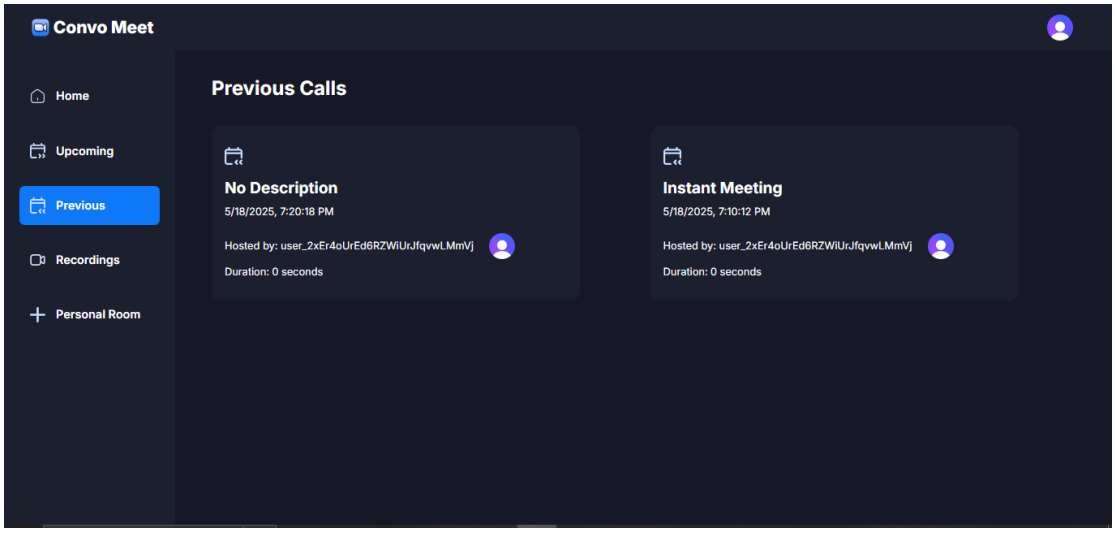
The recordings page shows recordings done during the meeting. These will be shown to the host only and he can then get the Summary from it, download the video or share via link.



The summary will be shown to the host like this generated by AI. The summary includes some points discusses in to meeting.



This is a sample meeting page which is joined by a member via the link shared to him. It includes controls on the meeting like recording(which is shown in recordings page to host).



Previous meetings page fetches the meetings hosted before. Likewise, the page Upcoming shows any scheduled meeting.

Project Links

Backend: https://github.com/DevoUmair/Automatic-Meeting-Minutes-Generator

Frontend: https://github.com/saim04/Meetsumfrontend

Conclusion

This AI-powered application redefines how teams interact with meeting content. By combining speech-to-text transcription and GPT-based summarization in one seamless workflow, the tool improves productivity, ensures better knowledge transfer, and reduces the overhead of manual documentation. It’s scalable, efficient, and tailored for modern collaboration needs.