Bilkent University
Department of Computer Engineering

Senior Design Project
T2315
SummarEyes

Analysis and Requirements Report

21902103, Kürşad Güzelkaya,
kursad.guzelkaya@ug.bilkent.edu.tr
21903435, Hazal Buluş,
hazal.bulus@ug.bilkent.edu.tr
21902512, Süleyman Gökhan Tekin,
gokhan.tekin@ug.bilkent.edu.tr
21902270, Burcu Kaplan,
burcu.kaplan@ug.bilkent.edu.tr
22002850, Deniz Gökçen,
deniz.gokcen@ug.bilkent.edu.tr

Supervisor
Uğur Güdükbay

Course Instructors
Atakan Erdem
Mert Bıçakçı

25.11.23

This report is submitted to the Department of Computer Engineering of Bilkent University in partial fulfillment of the requirements of the Senior Design Project course CS491/2.

Contents

Analysis and Requirements Report

SummarEyes

1   Introduction

In the wake of the pandemic, remote work and online learning have become increasingly popular due to their time-saving benefits, comfortable work environments, and cost-effectiveness. A pivotal aspect of these remote setups is online meetings, which enable effective collaboration without the need for physical presence. However, one significant challenge in online meetings is maintaining focus, as extended screen time can lead to distractions.

Recently, there has been a surge in interest regarding the development of summary systems tailored for virtual meetings. However, many of these systems fall short in meeting the diverse needs of users. Some end up consuming more time than they save, particularly for those who miss only brief portions of meetings. Another critical issue is accuracy, especially when dealing with languages other than English. This language barrier diminishes the reliability of such tools, especially for users seeking thorough and precise meeting summaries in different languages.

Introducing *SummarEyes*, an innovative solution designed to enhance online meetings and lectures. By harnessing the power of machine learning and computer vision, SummarEyes identifies moments when participants become distracted, ensuring that crucial information is not overlooked. The system's pretrained language model generates comprehensive summaries, complete with clear topic titles and timestamps for key discussions during these distracted intervals. These summaries are invaluable for attendees looking to catch up and provide speakers with valuable insights into audience engagement. SummarEyes empowers speakers to deliver more engaging presentations by offering feedback on the most common moments of distraction. This innovation promises to revolutionize the way companies and educational institutions approach online communication, making meetings and lectures more productive and engaging for all participants.

2   Current System

In the realm of virtual meeting platforms, we find various meeting summarization systems in use today. Notable among them are platforms like Zoom Meeting and Microsoft Teams, which have integrated meeting summarization features to enhance user experiences. However, these existing systems exhibit certain limitations that create an opening for innovative alternatives like our project [1][2].

The current meeting summarization systems primarily focus on the English language, excelling in summarizing English-language meetings with a certain degree of accuracy. However, their effectiveness takes a hit when it comes to processing meetings conducted in languages other than English. This language barrier narrows their accessibility and effectiveness for non-English speakers, limiting their global reach.

Furthermore, these existing systems tend to offer a one-size-fits-all approach, generating generic summaries that lack personalization. Such summaries do not cater to the specific needs of individual meeting participants. Consequently, attendees receive summaries that may not effectively capture the key points relevant to them, particularly if they have only missed specific portions of the meeting.

In contrast, our project introduces several advantages that address these limitations. Firstly, it is designed to operate proficiently in both English and Turkish, breaking down language barriers and making it accessible to a broader audience. Additionally, our system focuses on personalized summaries that target moments when a participant's attention wanes, enhancing the utility of the summaries. Unlike real-time systems, our approach involves post-meeting video analysis, ensuring comprehensive and accurate results. Furthermore, our project offers a feedback mechanism that empowers meeting hosts and speakers to improve their presentations based on insights into when distractions occur most frequently.

3   Proposed System

3.1   Overview

Our project introduces a revolutionary system that synergistically integrates emotion recognition and eye tracking in meeting summarization technology to revolutionize how attention levels are gauged during online meetings. This cutting-edge approach surpasses traditional single-score attention systems by offering a nuanced analysis that correlates emotional responses and eye movement data with attention spans.

The methodology involves capturing video footage of the entire meeting while meticulously tracking each participant's eye movements. Following the meeting, a detailed analysis of each individual's video ensues. This phase is critical, as it focuses on identifying periods of diminished attention, as evidenced by specific eye movement patterns and emotional indicators. These moments form the cornerstone of our unique meeting summaries, ensuring no crucial information slips through the cracks.

What sets our system apart are these tailored summaries, which emphasize the portions of the meeting where a participant's focus waned, thereby providing highly targeted and relevant content. While standard, comprehensive meeting summaries are also available, the personalized summaries are where our technology truly shines, catering to the distinct attention dynamics of each participant, informed by both emotional and eye movement data.

Moreover, our system enriches the overall meeting experience by offering invaluable feedback to hosts and speakers. This feedback, derived from the amalgamated data on emotions and eye movements, provides insights into the instances of highest distraction, the emotions linked to attention lapses, and actionable advice to refine presentation skills in future meetings.

3.2 Functional Requirements

3.2.1 User

- Registration: Users can sign up for the platform by providing required personal details and setting up login credentials.
- Login/Logout: Users have the ability to securely log in to access the platform and log out to maintain data security.
- Data Usage Permission: Upon registration or initial login, users will be prompted to consent to the platform's use of their personal data, including camera and microphone feeds, to comply with privacy regulations.
- Profile Management: Users can view and edit their profile information.
- Create a Meeting: Any user can schedule new online meetings, set the date, time etc.
- Join a Meeting: Any user can join a meeting using meeting ID's or secure links.

3.2.1.1 Participant

- Audio/Video Settings: Participants can manage their audio and video settings during meetings.
- Leave Meeting: Participants can leave meetings as needed.
- Meeting Interaction: Participants have the capability to view other participants during the meeting.
- View Summaries: Participants can download both default or customized summaries of meetings they have attended.

3.2.1.2 Host

- End a Meeting: Hosts have the authority to end the meeting for all participants, ensuring an orderly conclusion to the session.
- Manage Participants: Hosts can view participants and have the ability to invite or remove participants from the meeting.
- View Summary/ Feedback: Hosts can access and download default summaries and feedback reports post-meeting to review participant engagement and their own performance.

3.2.2 Organization

- Subscribe to Plan: Organizations can choose and manage subscription plans according to their needs.
- Organizational User Management: Organizations have the authority to add users to or remove users from their organizational structure within the platform.

3.2.3 Admin

- Manage User Accounts: Admins are responsible for the management of user accounts within the system, including General Users, Participants, and Hosts.

- Access Reports: Admins can access detailed reports related to user activities, meeting analytics, and system usage.

3.3 Non-functional Requirements

3.3.1 Usability

The system prioritizes a user-centric design, providing an intuitive interface that accommodates users of varying expertise. The process for account creation, meeting management, and feature utilization should be streamlined and clear. The system's layout will adopt a minimalist approach, reducing cognitive load and potential user errors. Accessibility to instructions and support documentation will be a click away, enhancing user comfort with data usage and system functionalities.

3.3.2 Reliability

The system's reliability remains paramount, ensuring consistent delivery of precise meeting summaries and robust performance under diverse operational conditions. It will support a wide range of languages and dialects accurately, and maintain high-quality processing regardless of audiovisual variations. A commitment to continuous operation will be upheld, with the system demonstrating a high availability quotient to meet user scheduling and participation needs.

3.3.3 Performance

Performance optimization is crucial, with the system capable of swiftly processing and analyzing data to provide post-meeting summaries efficiently. The analysis algorithms will operate discreetly in the background, preserving meeting integrity and participant experience. The system will be fine-tuned for varying internet bandwidths, providing a stable experience for all users, irrespective of their network conditions.

3.3.4 Supportability

The system's life cycle will be supported by regular updates and a dedicated support team, addressing technical inquiries and user feedback proactively. The infrastructure will be designed for straightforward maintenance and updates, minimizing service interruptions and ensuring the system evolves in line with technological advancements and user feedback.

3.3.5 Scalability

Scalability is designed into the system's architecture, ready to accommodate growth in user numbers, meeting frequency, and data throughput. The infrastructure will be capable of scaling up to meet increased demand while preserving performance levels. The system's roadmap includes provisions for future feature enhancements and capacity expansions to remain adaptable and forward-facing in the dynamic tech environment.

3.3.6 Robustness

The system will remain stable under various stress conditions, whether due to high user load, large data processing demands, or network variability. Also the system will include robust recovery mechanisms that enable quick restoration of service in the event of failure, with minimal impact on the user experience and data loss.

3.4   Pseudo Requirements

3.4.1 Implementation Constraints

- Google Colab will be utilized to leverage high-performance CPU and GPU resources for intensive machine learning computations, ensuring efficient processing of data analysis tasks.
- The MVC architectural pattern and OOP paradigm will govern the software development process.
- Support for English and Turkish languages will be included.
- Only participants with an active camera will be included in attention and emotion analysis.
- To follow the development process of the project, Jira and Github technologies will be used.
- The project will be a Web application that uses Agora SDK to create the online meeting platform of the project.
- PostgreSQL will be used as the database management system.
- Django, a Python-based framework, will be used to build the website of the application.
- ReactJS library will be used for the front-end development of the project.
- To transcript the online meeting, a pre-trained AI model Whisper (by OpenAI) will be used as well as some other helper machine learning libraries.
- Python's FER library will be used to perform emotion detection of the users to help with distraction detection [3].
- Gensim, a machine learning library, will be used to generate the summaries of the transcripted meeting record.
- Llama2, a pre-trained language model, will be used to generate personalized summaries and topic predictions.

3.4.2 Economic Constraints

- Domain service is free of charge.
- Open source libraries are free of charge.
- Development and testing tools are free of charge.
- Cloud servers will be free of charge.
- Github repositories are free of charge for student accounts.
- To be able to provide online meeting features, we need to use Agora SDK, which gives 10,000 minutes of video call free each month and charges additional time with 3.99$/1000 minutes [4].

- We plan to use a pre-trained model of Whisper AI to convert speech-to-text, which is open-source and free. When the user count increases, using the premium version, which provides optimization with GPUs, will be an option with 6$ for 1000 minutes [5].
- Language model Llama2 is free [6].
- The libraries and frameworks we will use to implement our application are free.
- We plan to use Google Colab, which provides free GPUs for training the DL model.

3.4.3 Safety Constraints

- Comprehensive security protocols will be in place to protect user data and ensure the privacy of stored meeting videos.
- Clear and transparent consent mechanisms will be implemented for participants, detailing the storage and analysis of video data.

3.4.4 Ethical Constraints

- Necessary consent will be taken from the user to record their video and analyze it to produce personalized summaries. After analyzing video data which will be stored privately, it will be deleted.
- We will inform users about the precision of the technology's evaluations and offer clarity on the methods used to identify distractions and produce summaries. Users will be aware of the system's limitations and possible mistakes.
- The gathered data —including video feeds and summaries— will be protected from abuse, breaches, and unwanted access. It is planned to use strong encryption techniques and safe storage mechanisms.

3.4.5 Environmental Constraints

- Since our project has no connection with environmental factors, there will be no environmental constraint.

3.4.6 Manufacturability Constraints

- Since our project has no connection with manufacturing, there will be no manufacturability constraint.

3.4.7 Technological constraints

- Users will need devices with internet connectivity and camera functionality to participate in meetings and to have their videos recorded for subsequent analysis.
- The program will be suitable for different platforms like Chrome, Firefox, etc.
- Demo will only support 2 participants (excluding the presenter) per meeting due to the limitation of processing power at the current stage.

3.5   System Models

3.5.1   Scenarios

Scenario1: Sign Up

Actors: Users who don't have an account in the SummarEyes app.

Entry Conditions: User doesn't have an account clicks on the sign up button.

Exit Conditions: User successful sign up.

User signing up fails.

Flow of Events: 1. User enters the personal information (name, surname, email, password etc.) and clicks on the "submit" button.

2. A verification code is sent to the user's email account.

3. User enters the received code to the related field of the authentication page.

4.1. If the entered code is correct, the user is directed to the home page.

4.2. If the entered code is wrong, the user tries again.

Scenario2: Log in

Actors: Users who have an account in the SummarEyes app.

Entry Conditions: User isn't signed in to the app.

Exit Conditions: User successful log in.

User logging in fails.

Flow of Events: 1. User fills in the "email" and "password" boxes.

2.1. If the entered email and password is correct, the user is directed to the home page.

2.2 If the entered email or password is wrong, the user tries again.

Scenario3: Host a meeting

Actors: User.

Entry Conditions: User is signed in and clicks on the "Host a Meeting" button in the home page.

Exit Conditions: User clicks on the "cancel" button.

Flow of Events: 1. User enters a display name.

2. User clicks on the "Start the Meeting" button.

3. A new meeting id is assigned to the created meeting.

4. User is directed to the meeting room with the assigned id.

Scenario4: Invite a Participant

Actors: Host

Entry Conditions: Host is hosting an on-going meeting and clicks on the "Invite Participants"

button.

Exit Conditions: Host clicks on the "Cancel" button.

Host invites a participant.

Flow of Events: 1. Host clicks on the "Invite Participants" button.

2. A pop up which has the meeting id of the current meeting shows up.

3. Host copies the meeting id and shares it to the other users to join.

Scenario5: Join a Meeting

Actors: User.

Entry Conditions: User is signed in and received a valid meeting id.

Exit Conditions: User clicks on the "Cancel" button.

Meeting id the user enters is not valid.

Flow of Events: 1. User clicks on the "Join a Meeting" button in the home page.

2. A pop up occurs for users to enter the meeting id they wish to join.

3. User enters the meeting id and clicks on the "Enter" button.

4.1. If the entered meeting id is valid and the meeting is still continuous, the

user will be directed to the meeting with the entered meeting id.

4.2 If the entered meeting id is not valid or the meeting has ended, the user will
be directed to the home page.

5. The user gains access to the online meeting and can control the camera and
microphone preferences.

Scenario6: End a Meeting

Actors: Host, User.

Entry Conditions: Host is hosting an on-going meeting and clicks on the "End the Meeting"

button.

Exit Conditions: Host clicks on the "Cancel" button.

Host ends the meeting.

Flow of Events: 1. Host clicks on the "End the Meeting" button.

2. All users in the meeting are directed to the home page.

3. Host receives the feedback of the meeting which includes the time ranges

when the most users are distracted.

Scenario7: Leave a Meeting

Actors: User

Entry Conditions: User is in an on-going meeting and clicks on the "Leave the Meeting"

button.

Exit Conditions: User clicks on the "Cancel" button.

User leaves the meeting.

Flow of Events: 1. User clicks on the "Leave a Meeting" button.

2. User is directed to the home page.

Scenario8: View Profile

Actors: User.

Entry Conditions: User is signed in and clicks on the "My Profile" button.

Exit Conditions: User is directed to his/her profile.

Flow of Events: 1. User clicks on the "submit" button in the sidebar.

2. User is directed to his/her profile.

Scenario9: View Meeting Summaries

Actors: User.

Entry Conditions: User is signed in and views his/her profile.

Exit Conditions: User is directed to the Summary History page.

Flow of Events: 1. User views his/her profile and clicks on the "Meeting Summaries" button.

2. User is directed to the Summary History page.

3. User views all the summaries he/she received with descending order by the meeting date.

4. User clicks on the summary he/she wishes to view.

5. A pop up which includes the summary occurs.

Scenario10: View Meeting Feedbacks

Actors: User.

Entry Conditions: User is signed in and views his/her profile.

Exit Conditions: User is directed to the "Feedback History" page.

Flow of Events: 1. User views his/her profile and clicks on the "Feedbacks" button.

2. User is directed to the Feedback History page.

3. User views all the feedback reports of the meetings he/she hosted before with descending order by the meeting date.

4. User clicks on the feedback he/she wishes to view.

5. A pop up which includes the feedback occurs.

Scenario11: View Subscription Plan

Actors: User.

Entry Conditions: User is signed in and views the home page.

Exit Conditions: User is directed to the "Subscription Plan" page.

Flow of Events: 1. User views the home page and clicks on the "Subscription Plan" button.

2.User is directed to the Subscription Plan page.

3. User views the current subscription plan he/she has and the other

subscription plans to pass on.

Scenario12: Change Subscription Plan

Actors: User.

Entry Conditions: User is signed in and views the "Subscription Plan" page.

Exit Conditions: User is directed to the "Payment" page or User clicks on cancel button. .

Flow of Events: 1. User views his/her "Subscription Plan" page and clicks on the "Upgrade

Plan" button on one of the plan options.

2.User is directed to the "Payment" page.

3. User enters the credit card information and clicks on the "Pay Now" button .

3.1. If the entered credit card information is correct, the subscription plan of the

user will be upgraded.

3.2. If the entered credit card information is incorrect, the user will continue

using the application with the previous plan.

Scenario13: Add Member to My Organization

Actors: User.

Entry Conditions: User is signed in and views the "Profile" page.

Exit Conditions: User is directed to the "My Organization" page.

Flow of Events: 1. User views his/her profile page.

2. User clicks on the "My Organization" button.

3.User enters an username and clicks on the "Add Member" button.

4.1. If the entered username exists, the user with the entered username is

added to the organization.

4.2. If the entered username doesn't exist, a pop up occurs in the page saying

"The username you entered is invalid!".

Scenario14: Remove Member From My Organization

Actors: User.

Entry Conditions: User is signed in and views the "Profile" page.

Exit Conditions: User is directed to the "My Organization" page.

Flow of Events: 1. User views his/her profile page.

2. User clicks on the "My Organization" button.

3.User clicks on the "Remove" button next to the user he/she wishes to

remove.

4. Selected user is removed from the organization.

Scenario15: Change Password

Actors: User.

Entry Conditions: User is signed in and views the "Profile" page.

Exit Conditions: User's password is changed OR User clicks on cancel.

Flow of Events: 1. User views his/her profile page.

2. User clicks on the "Change Password" button.

3. User enters the current password and the new password (twice).

4.1. If the entered current password is valid, the new password is being set.

4.2. If the entered current password is invalid, the password remains the

same.

Scenario16: Contact Summareyes

Actors: User.

Entry Conditions: User is signed in and views the "Home" page.

Exit Conditions: User sends a message to Summareyes.

Flow of Events: 1. User views Home page.

2. User clicks on the "Contact Us" button.

3. User is directed to the "Contact" page.

4. User enters his/her name, email and message.

5. User clicks on the "Send Message" button.

3.5.2   Use-Case Model
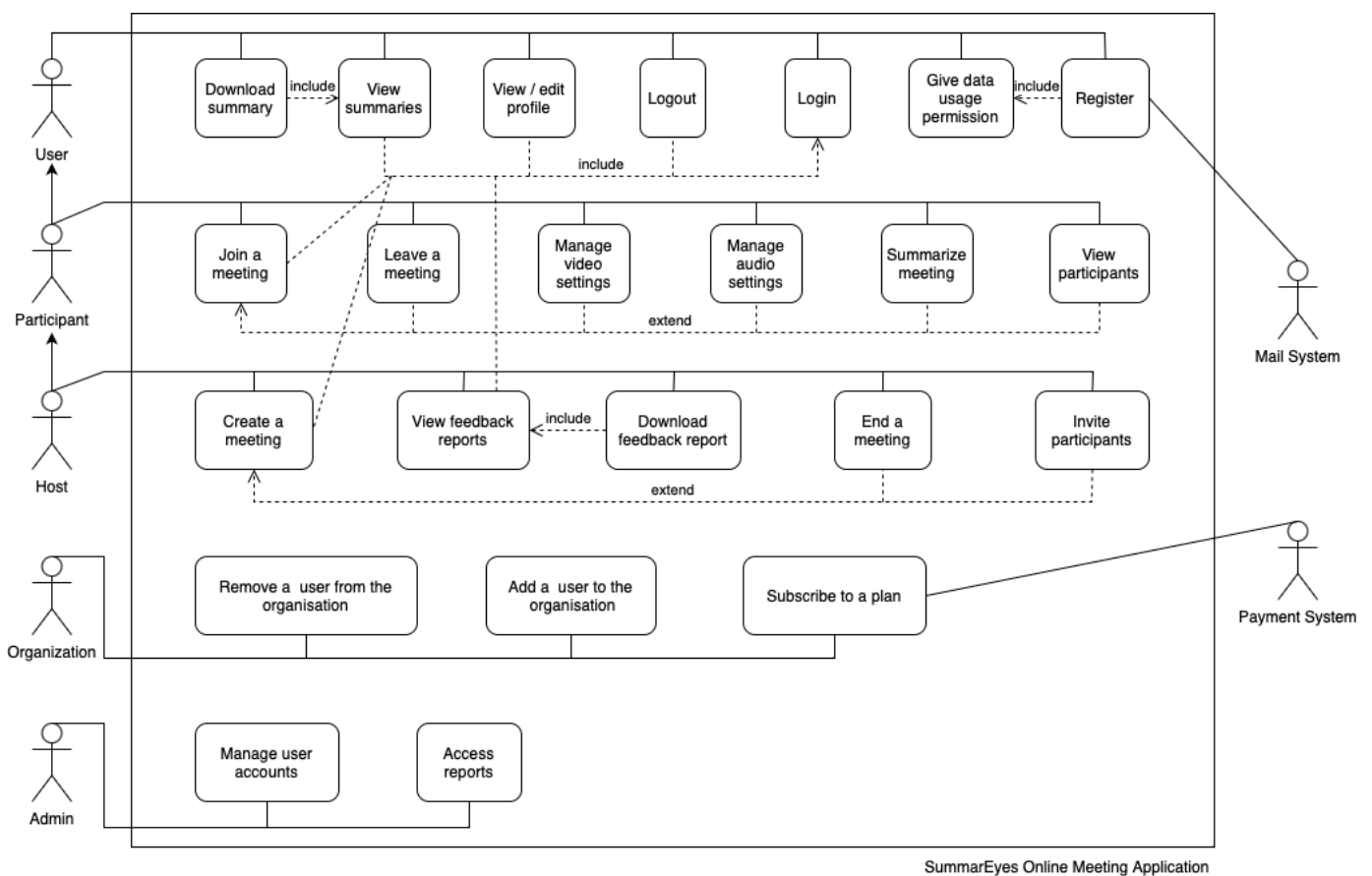
Figure 1: Use Case Diagram

3.5.3 Object and Class Model
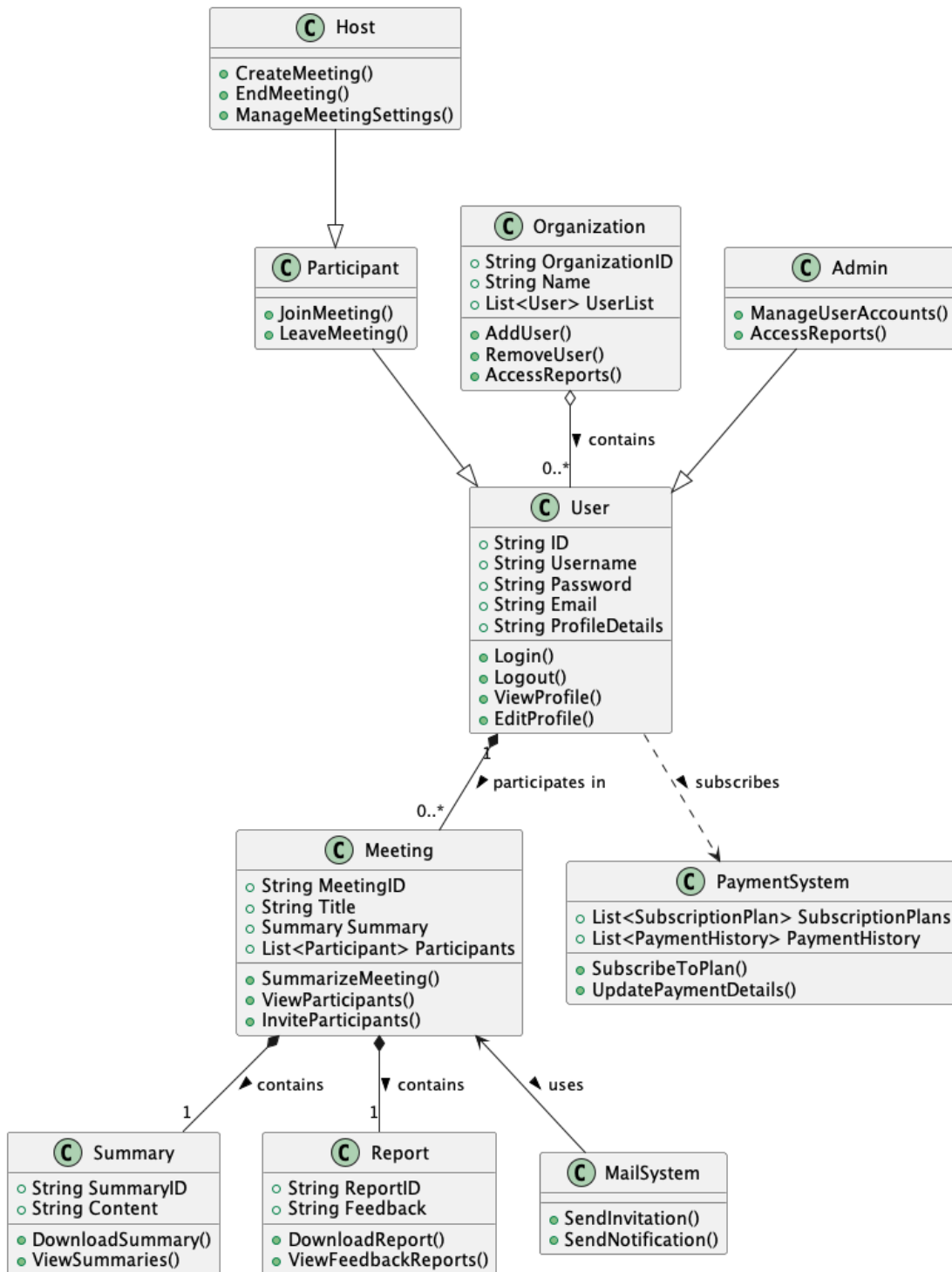
Figure 2: Object and Class Model
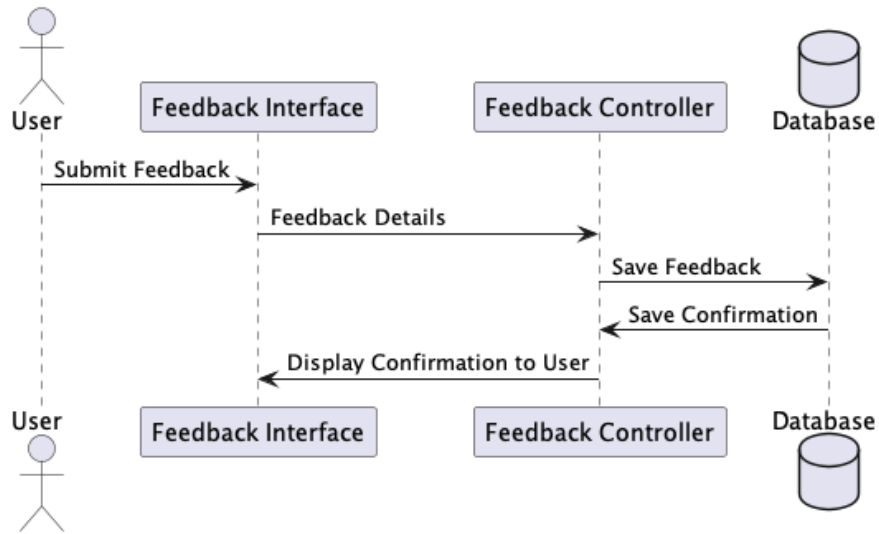
3.5.4 Dynamic Models
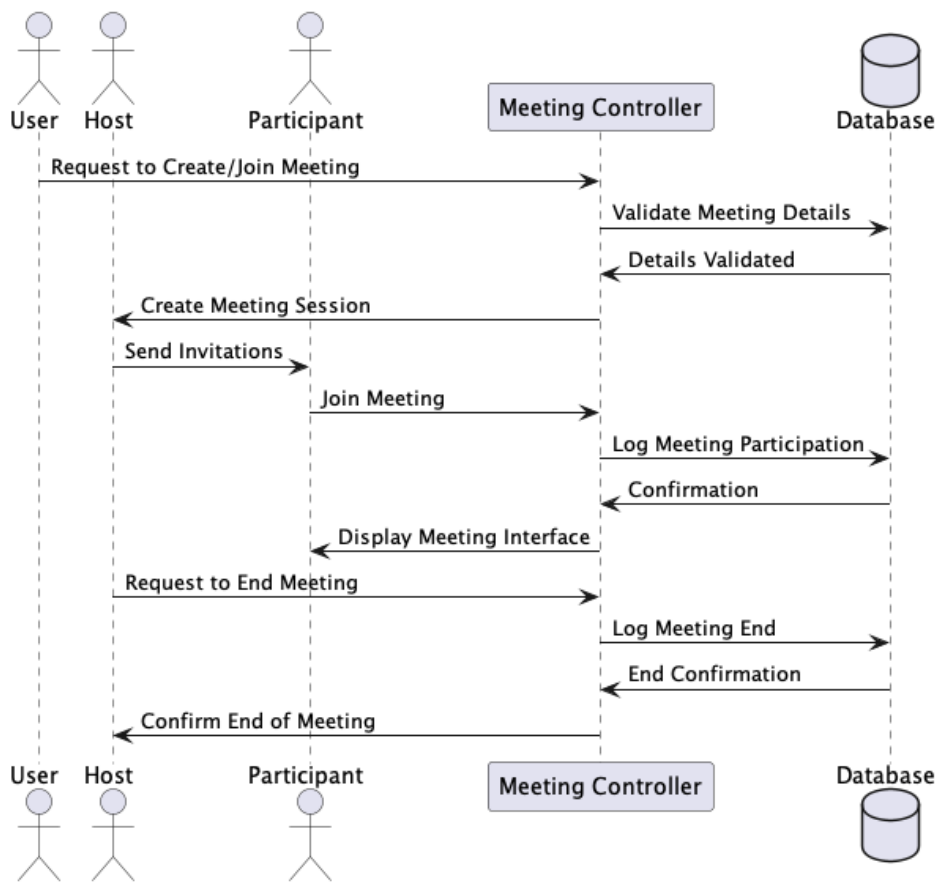

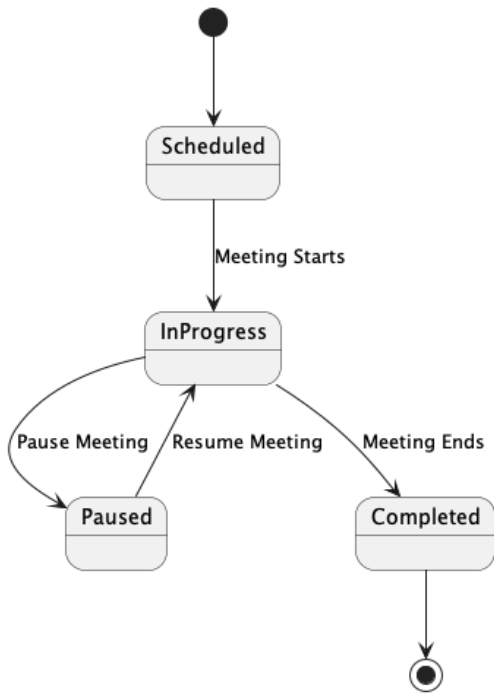Figure 3: Feedback System Sequence Diagram

Figure 4: Meeting Management Sequence Diagram

Figure 5: Meeting State Diagram
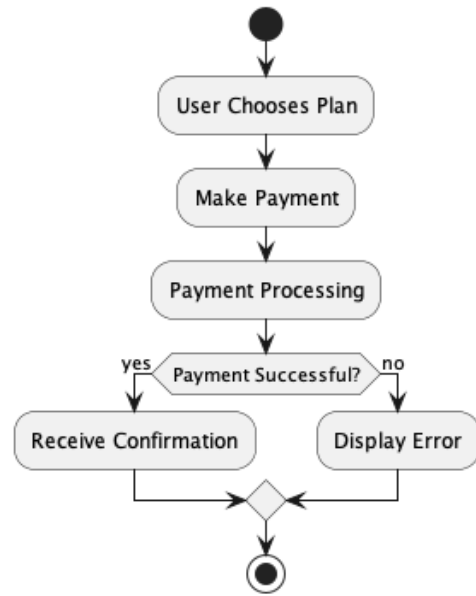
Figure 6: Subscription Activity Diagram
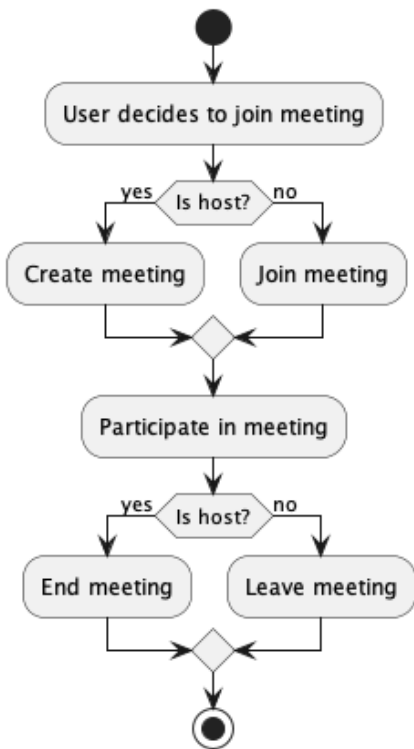
Figure 7: Meeting Workflow Activity Diagram
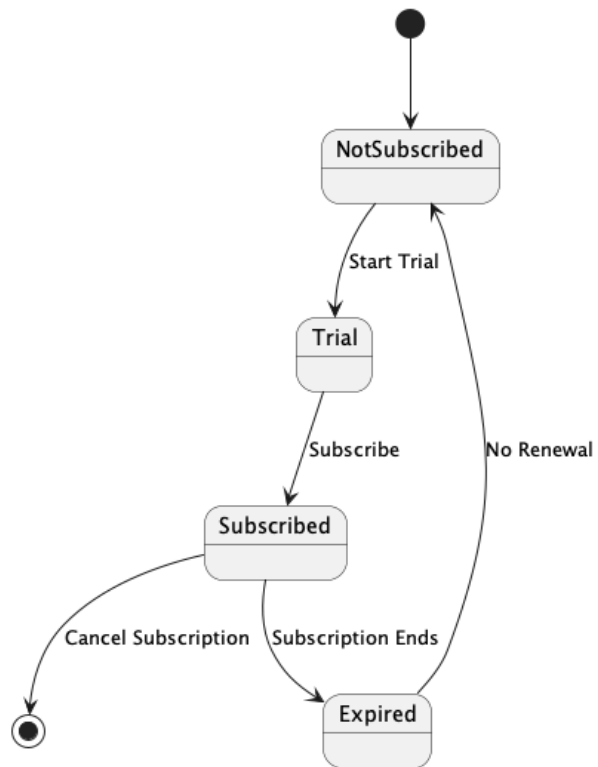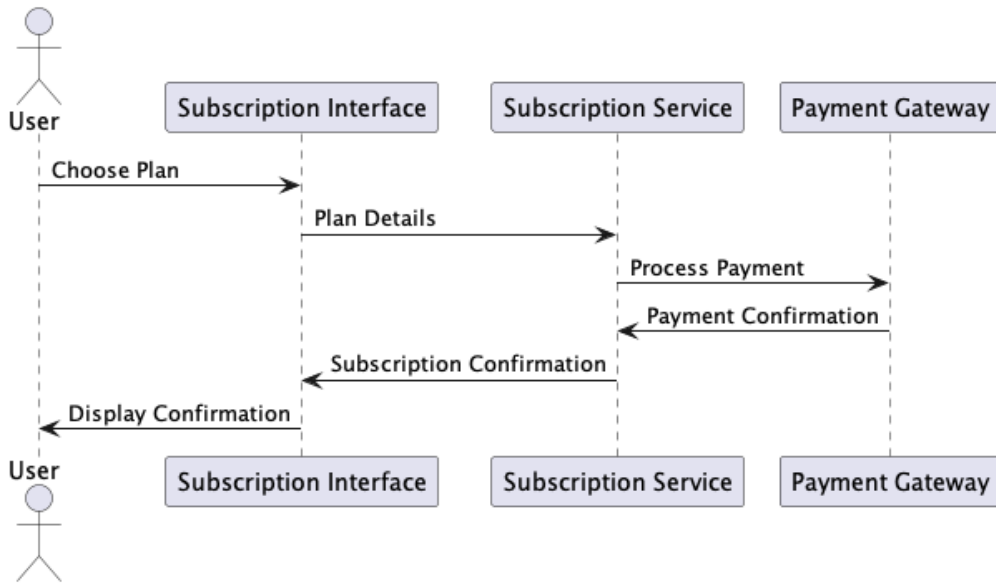

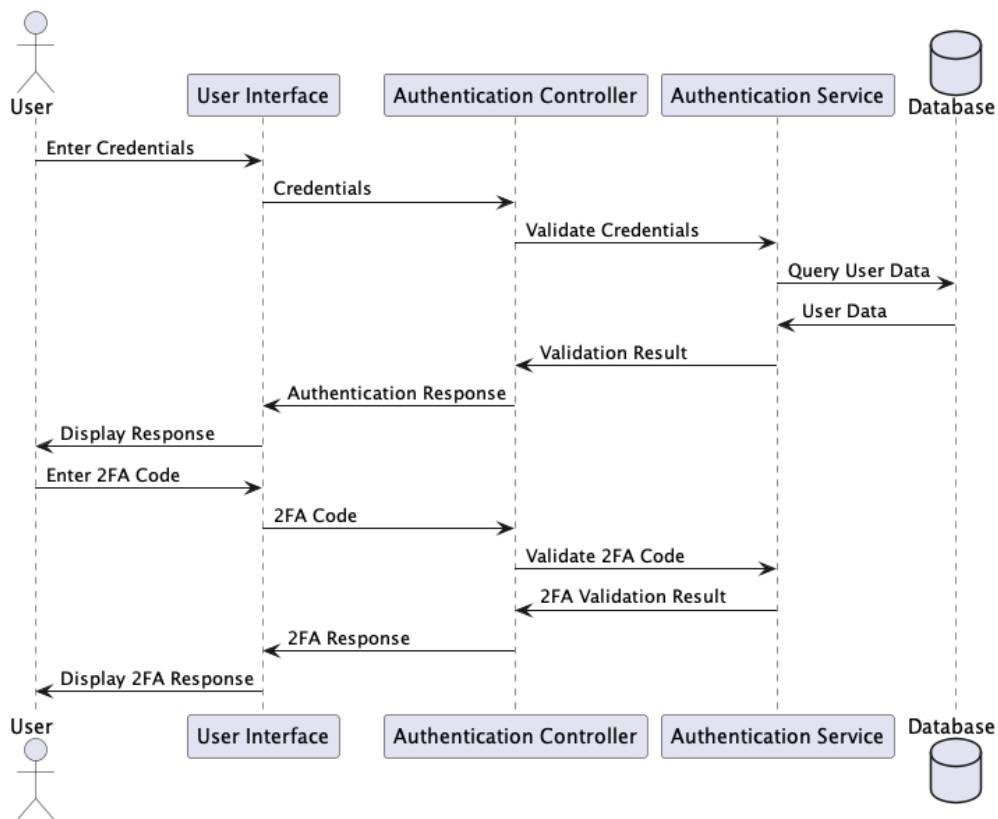Figure 8: Subscription State Diagram

Figure 9: Subscription Sequence Diagram
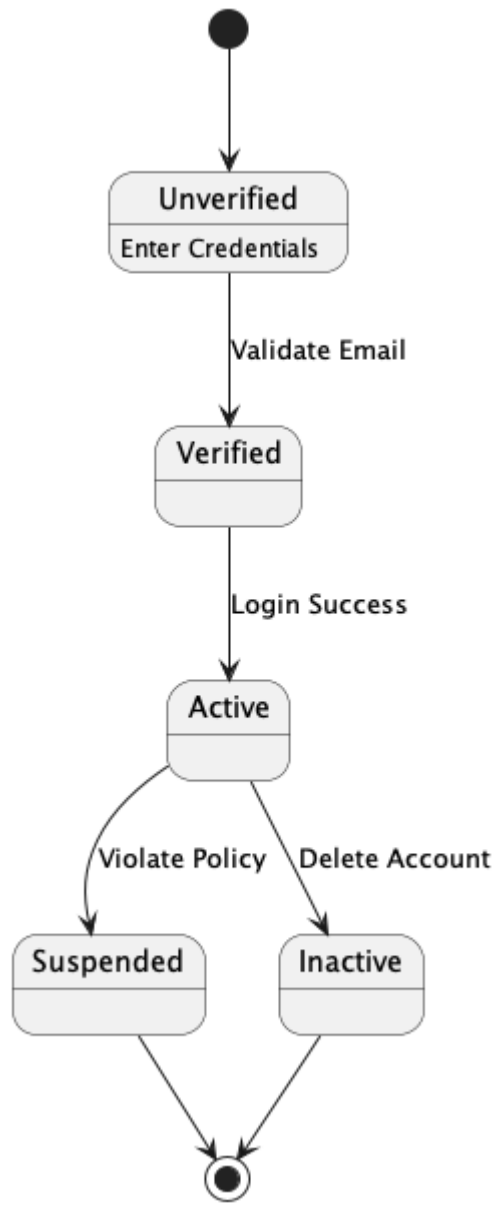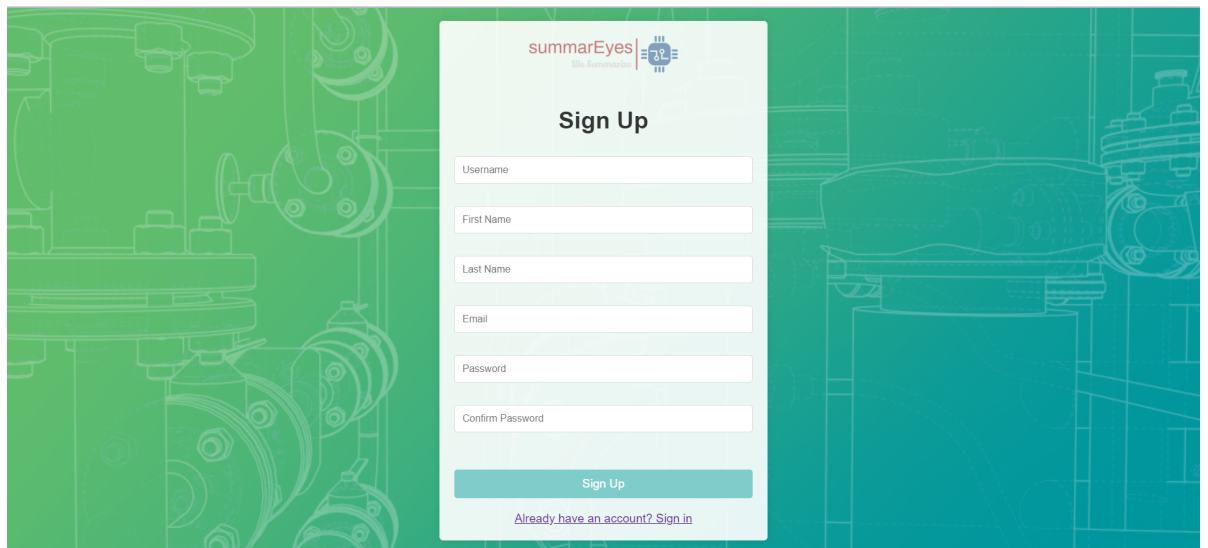

Figure 10: User Authentication Sequence Diagram

Figure 11: User Account State Diagram

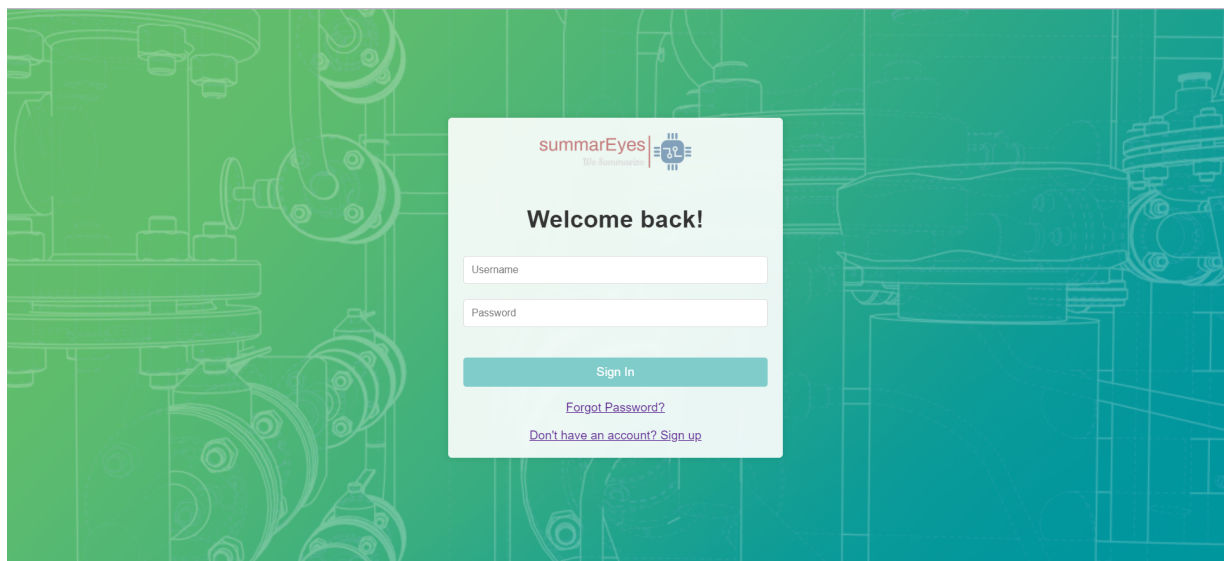### 3.5.5　User Interface- Navigational Paths and Screen Mock-ups

3.5.5.1 Sign Up Page

Figure 12: Sign Up Page

3.5.5.2 Sign In Page

Figure 13: Sign In Page

3.5.5.3 Homepage
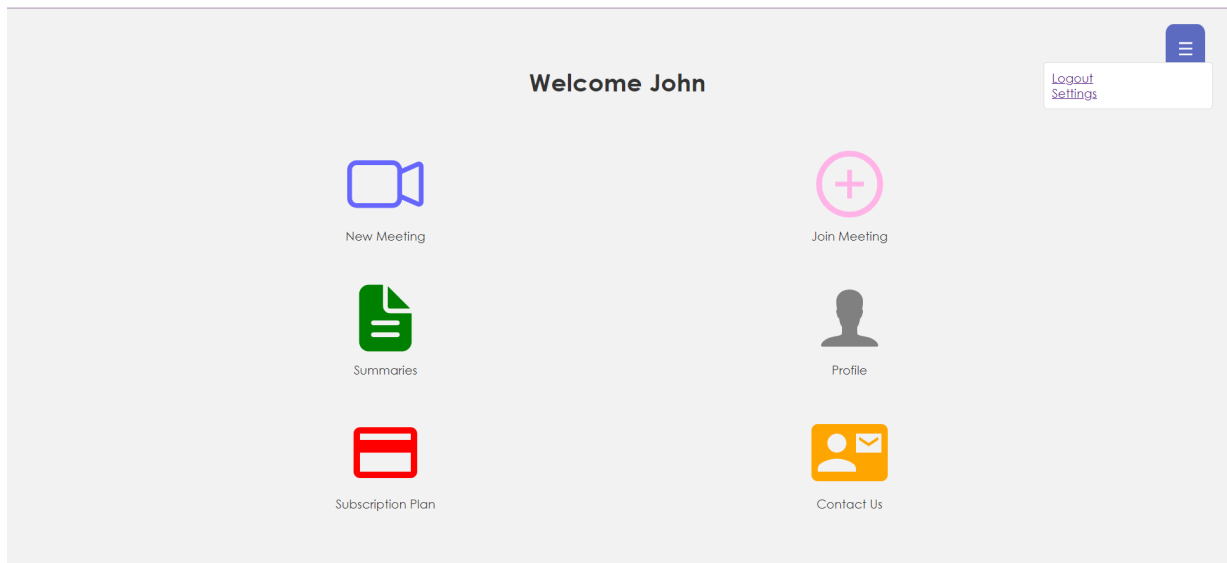
Figure 14: Home Page

3.5.5.4 Homepage Join Meeting Pop-up
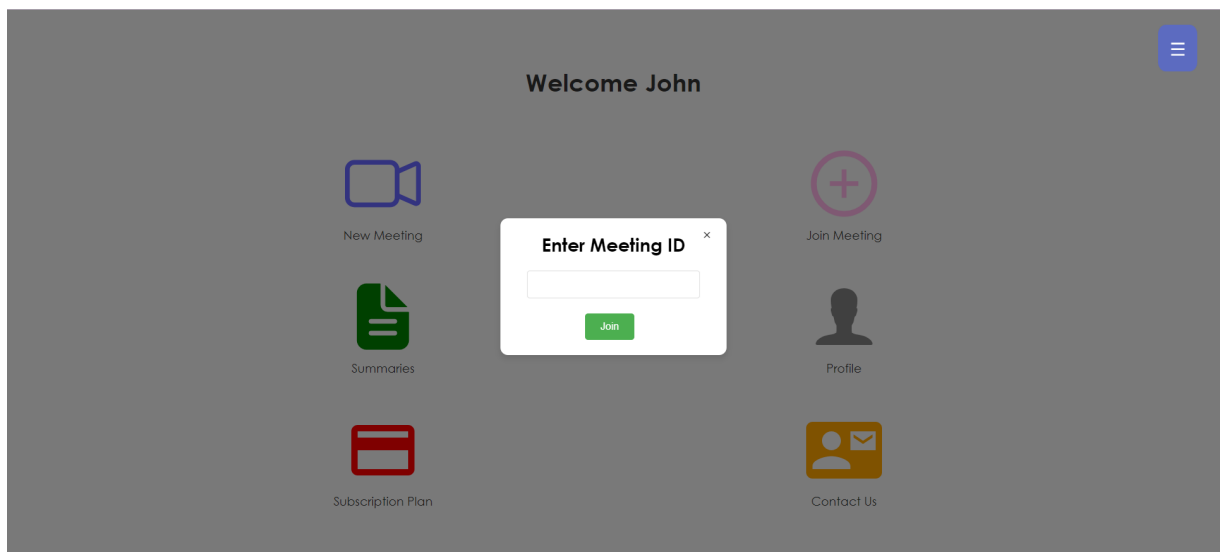
Figure 15: Join Meeting Pop-up

3.5.5.5 Meeting Page
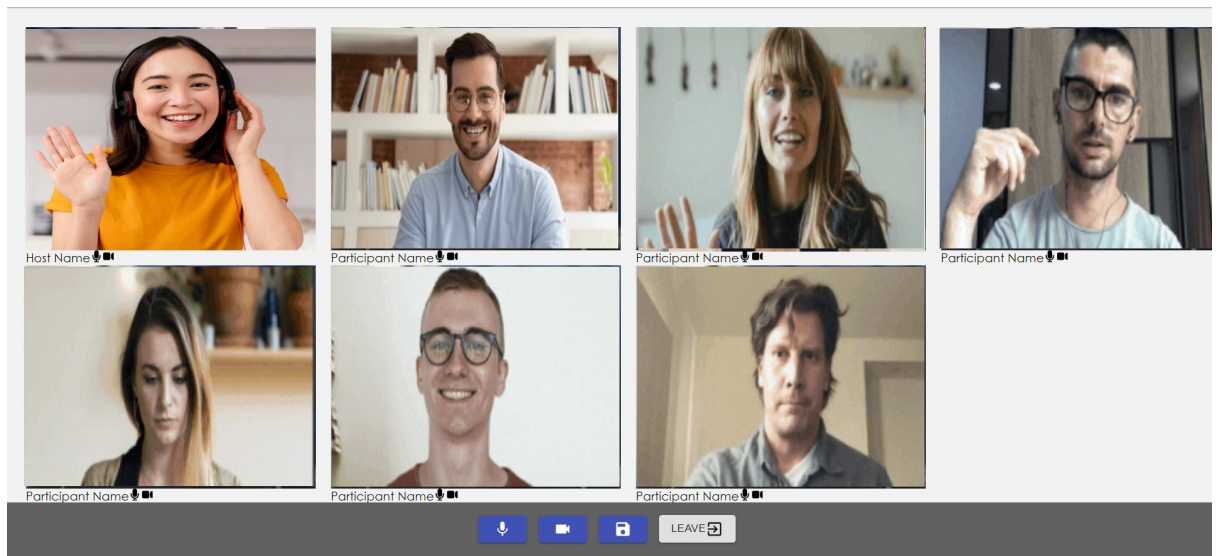
Figure 16: Meeting Page

3.5.5.6 Profile Page
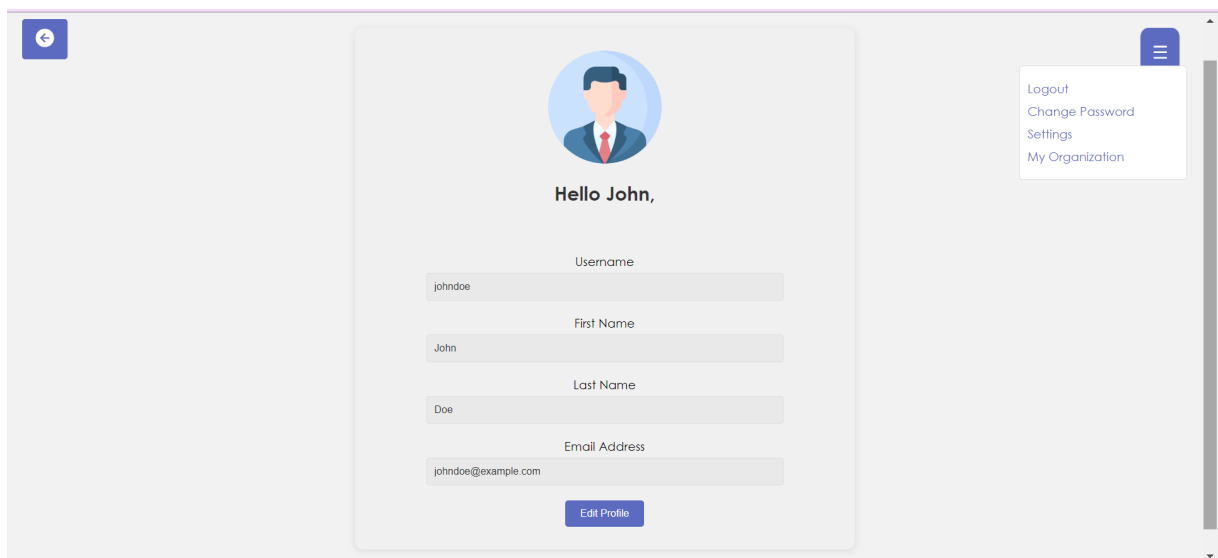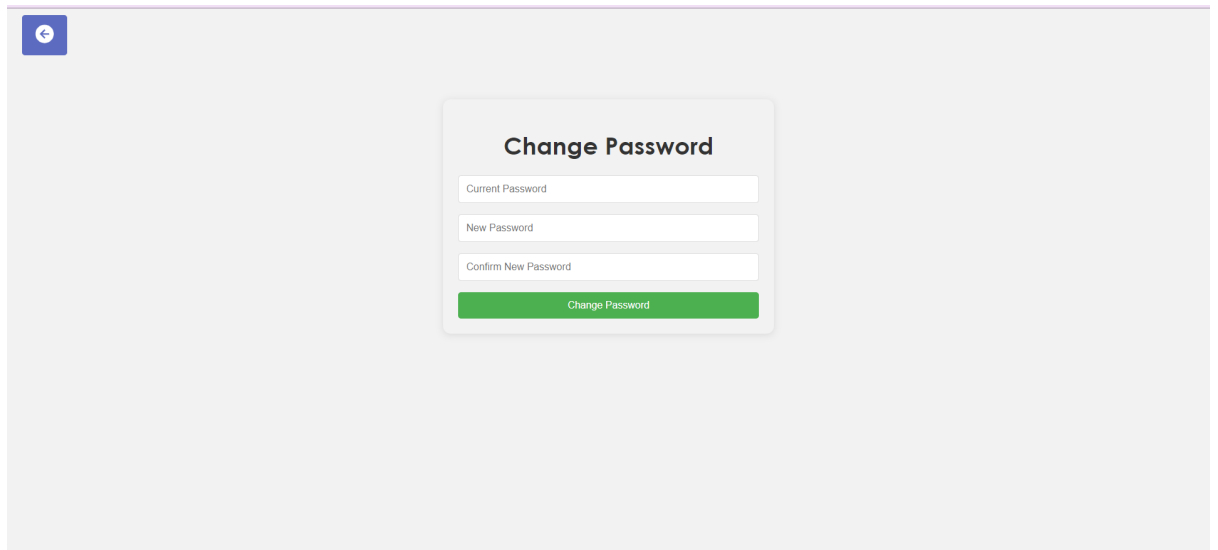
Figure 17: Profile Page

3.5.5.7 Change Password Page
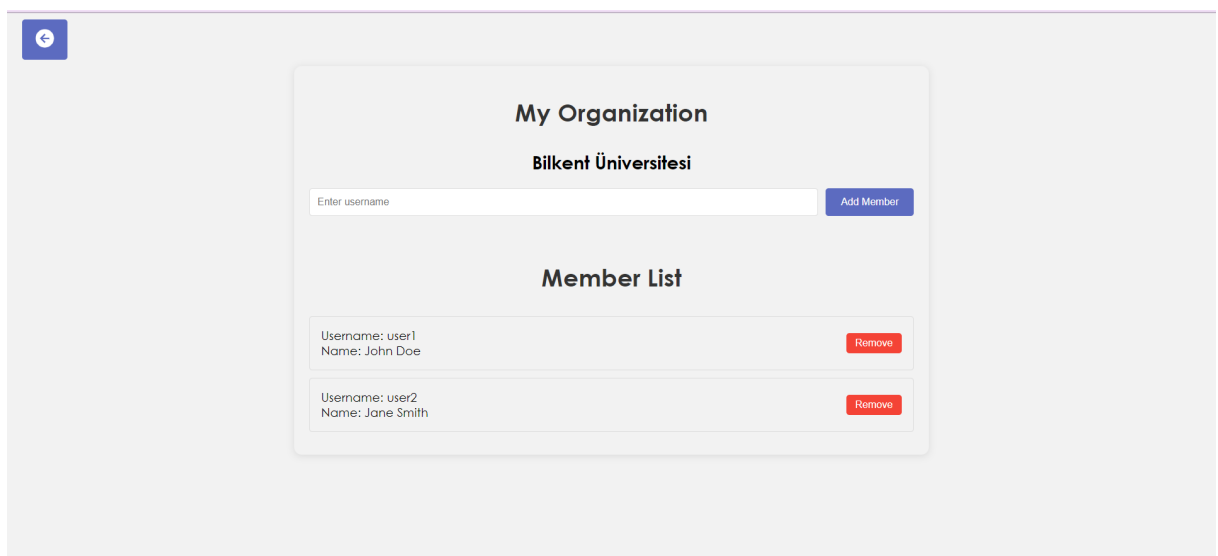
Figure 18: Change Password Page

3.5.5.8 My Organization Page

Figure 19: My Organization Page

3.5.5.9 Previous Summaries/Feedbacks Page
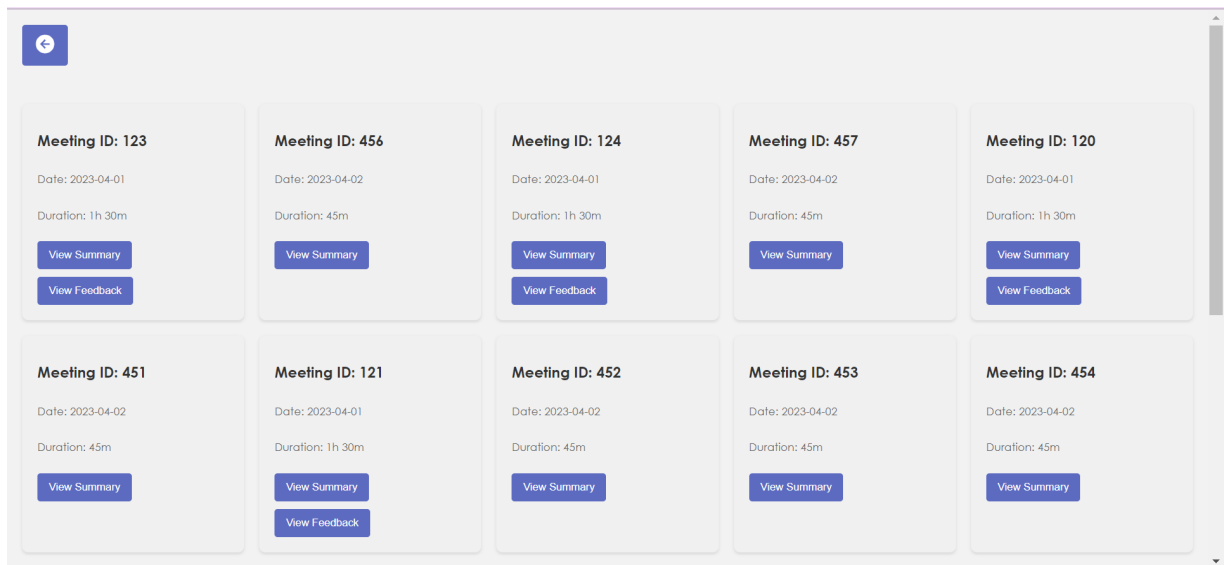
Figure 20: Summaries Page

3.5.5.10 View Summary
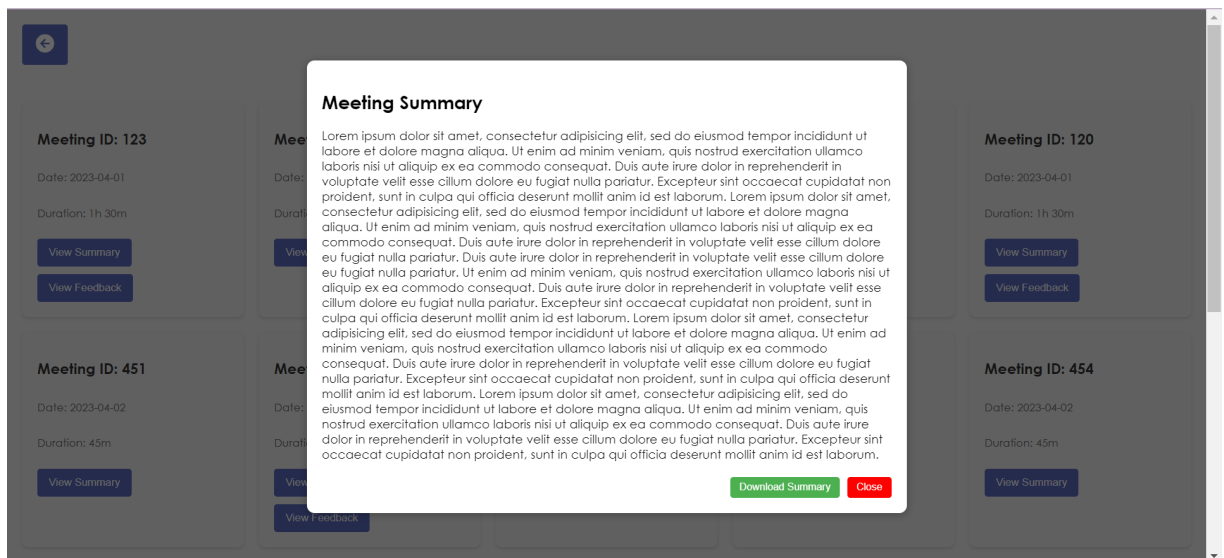
Figure 21: View Summary Page

3.5.5.11 View Feedback
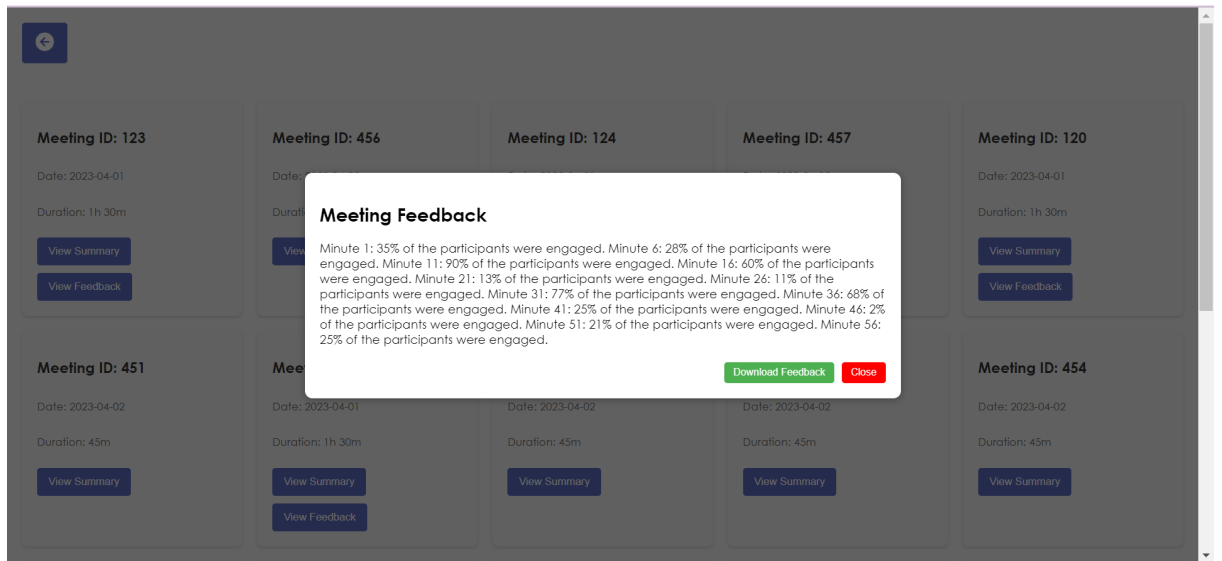
Figure 22: View Feedback Page

3.5.5.12 Subscription Plan Page

Figure 23: Subscription Plan Page
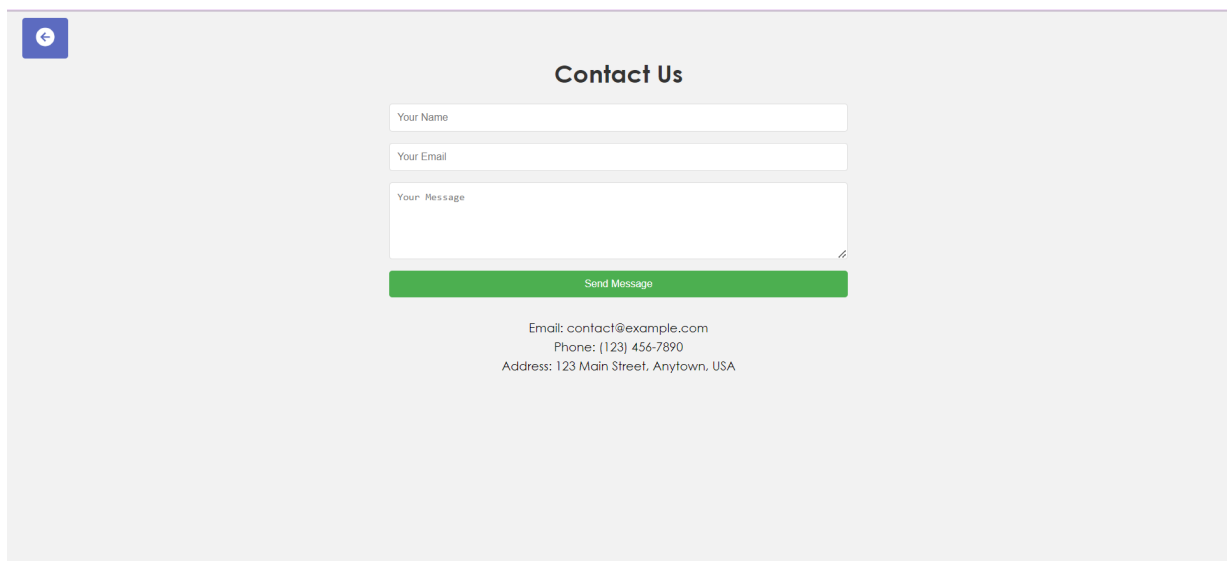
3.5.5.13 Payment Page

Figure 24: Payment Page

3.5.5.14 Contact Page

Figure 25: Contact Page

4 Other Analysis Elements

4.1 Consideration of Various Factors in Engineering Design

4.1.1 Constraints

The SummarEyes project faces several implementation constraints, pivotal for guiding the development process. For project management and version control, important tools like Jira and GitHub have been selected, guaranteeing streamlined workflow and cooperation. Because it is a web-based platform, the application uses the Agora SDK to facilitate online meetings. The database management system of choice is PostgreSQL since it offers dependable and sturdy data storage. In terms of web development, frontend development will make use of the React library, and backend development will employ the Python-based Django framework. For online meeting transcription, Whisper—OpenAI's pre-trained AI model—must be included. Additionally, Python's FER library will be used for emotion detection, and Gensim for summarizing transcripts. The use of Llama2, a pre-trained language model, will further enhance our ability to generate personalized meeting summaries and topic predictions.

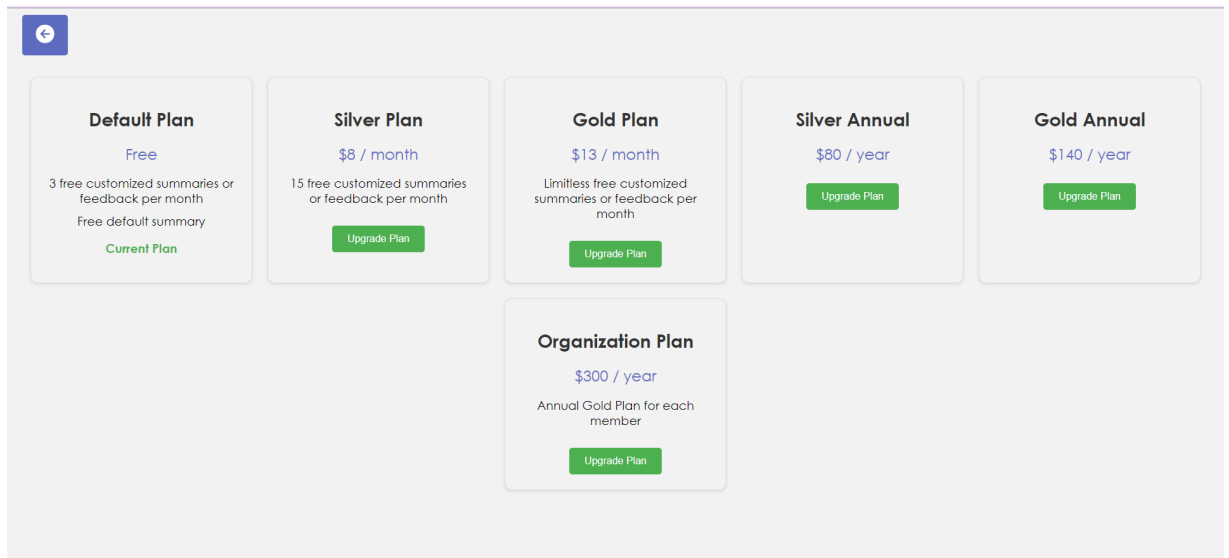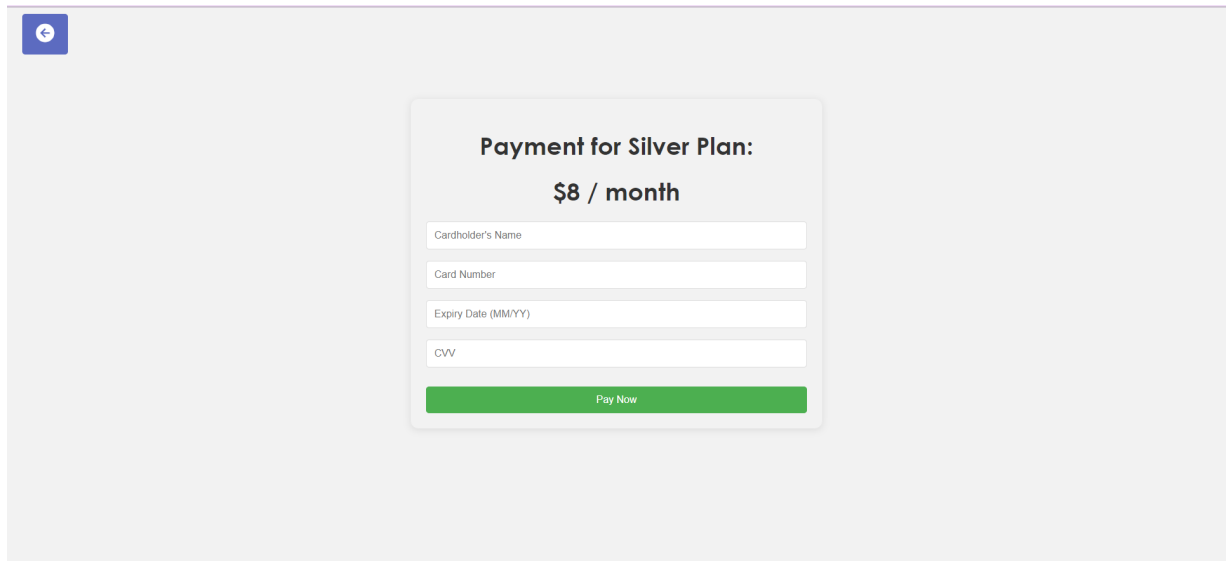Economically, the project is constrained by the costs associated with using certain technologies. For video call functionality, the Agora SDK is required. It provides 10,000 minutes of free usage each month, after which additional usage is charged at $3.99 per 1000 minutes. The project also intends to employ Whisper AI for speech-to-text conversion. Whisper AI is open-source and free in its basic version, but when the number of users increases, it could be necessary to upgrade to a premium version for optimized GPU support, which will cost $6 every 1000 minutes. Thankfully, there is some financial relief available in the form of free development libraries and frameworks, such as Google Colab's GPU support for deep learning model training, and the language model Llama2.

Concerns of ethics are central to the SummarEyes project. Users' informed consent will be strictly sought before any video data is recorded and analyzed. The data will be kept confidential and removed after analysis. Users will be informed in a transparent manner about the methods and accuracy of our technology, especially when it comes to summarizing and identifying distractions. This entails admitting to the shortcomings and possible mistakes of the system. Since data security is a top priority, robust encryption methods and secure storage systems will be used to guard against misuse, security breaches, and unauthorized access to all collected data, including video feeds and meeting summaries.

4.1.2 Standards

The standards used are UML 2.5.1 for modeling and IEEE 830 for documentation.

4.1.3 Considerations about Public Health

Growing worries about "Zoom fatigue," or physical and psychological exhaustion, have arisen from the increased use of virtual meetings. This phenomena, which was thoroughly investigated during the COVID-19 epidemic, emphasizes the mental strain that results from using digital interfaces for extended periods of time as well as the difficulty of deciphering

nonverbal signs in a virtual setting [7]. By offering brief meeting summaries based on participant attention and emotion detection, SummarEyes seeks to reduce this stress. This feature enables users to participate in meetings with less intensity, which may lessen exhaustion. SummarEyes might also include user interface modifications, including the ability to see just the presenter's video, which would simplify the visual field and reduce mental stress.

4.1.4 Considerations about Public Safety

In the realm of video conferencing, safeguarding sensitive information is paramount. SummarEyes will protect privacy by not keeping meeting recordings on file permanently. Only feedback and summaries will be kept instead, and users will have the option to download and then remove this data. In order to enhance safety even more, hosts will be able to monitor participant interactions, such as adding or removing attendees, in order to preserve meeting integrity and security. Access to meetings will also be restricted through authentication procedures.

4.1.5 Considerations about Public Welfare

Since SummarEyes is a web program that people will use for personal purpose, there is no impact on public welfare.

4.1.6 Considerations about Global Factors

While no immediate global factors impact the design of SummarEyes, lessons learned from the COVID-19 pandemic underscore the importance of adaptability in communication tools. Therefore, SummarEyes will be designed with scalability and flexibility in mind, ready to adapt to changing global conditions and emerging user needs.

4.1.7 Considerations about Cultural Factors

Diversity in communication styles between cultures is an important factor. These differences will be taken into consideration by SummarEyes's mood and attention analysis algorithms. For example, gesture expressiveness differs between communication cultures with high and low context. The model for emotion detection and the dataset for eye tracking will comprise a wide variety of cultural expressions to enable correct analysis and improve the app's cross-cultural usability.

4.1.8 Considerations about Social Factors

Reducing social biases is an essential part of developing AI. SummarEyes will take extra care to make sure that biases pertaining to age, gender, or race are not present in its attention and emotion analysis algorithms. In order to do this, inclusive datasets will be used, and universal measures of attention and emotion, like eye contact and facial expressions, that are free from bias will be the main focus.

4.1.9 Considerations about Economic Factors

A wide range of users, including companies and academic organizations, are the focus of SummarEyes. The software will be available for free at first to promote user input and wide adoption. Later, tier-based subscription models will be implemented. These models will serve a variety of user needs, including business customers with access to more sophisticated features and support available in premium packages, and educators with easily accessible solutions. This approach seeks to strike a compromise between wide accessibility and financial viability.

Factor	Effect Level	Effect
Public Health	7	Stress during the online meetings should be reduced
Public Safety	8	User's privacy must be secured.
Public Welfare	0	None
Global Factors	2	Scalability and flexibility must be a concern in our development
Cultural Factors	5	Datasets and models should be inclusive
Social Factors	6	No biases should be included
Economic Factors	2	Initially free, then tier-based subscription

Figure 26: Factor and Effect Table

4.2   Risks and Alternatives

In the deployment of our project, we recognize several risks and have developed alternative strategies to mitigate them effectively.

Performance and Response Time: Due to the computing needs of image processing techniques, a major worry is the possible delay in processing and assessing the emotional state and attention level of presentation participants. As data is transmitted to our systems for processing, this delay might become more pronounced. When we improve the speed and efficiency of our models and algorithms, this restriction will be reviewed and perhaps increased.

Technical Difficulties with Agora SDK: The Agora SDK is a technology that none of our team has ever used before, thus successfully implementing it is critical to the success of our project. Although initial investigation indicates that Agora SDK should fulfill our needs, we are

conscious of the dangers involved in utilizing new technologies. We are ready to investigate the possibility of creating these features from scratch if Agora SDK constraints prevent us from implementing crucial video-conferencing features; however, this may require more time and resources.

Accuracy of Models and Algorithms: We are currently using the Facial Expression Recognition (FER) library and Pretrained Language Models for emotion detection, along with well-known eye-tracking datasets like MPIIGaze and GazeCapture for attention assessment, in an effort to address the accuracy of models and algorithms for our SummarEyes application. Understanding that these sources might not meet our requirements for precision, we are ready to investigate other or additional datasets that might provide better results. In the event that this strategy proves ineffective, we have a backup plan that entails creating a bespoke dataset that is especially suited to the particular needs of our application, with an emphasis on enhancing the precision of attention level assessments. This preemptive approach is essential to guaranteeing SummarEyes' dependability and efficiency in providing excellent meeting summaries.

Security Problems: In the development of SummarEyes, a significant security risk pertains to the unauthorized access of meetings, a concern that has become increasingly relevant in the era of widespread digital communication. The potential for "meeting bombing," in which uninvited attendees join a meeting through a shared or publicly accessible link, endangers the confidentiality of the conversations as well as the integrity of the meeting space. This intrusion may cause discomfort, the disclosure of private information to uninvited parties, and a general feeling of unease among individuals who are authorized.To mitigate this risk, SummarEyes could implement a multi-layered security protocol. A viable solution is to introduce a two-step verification process for joining meetings. This could involve a combination of a meeting link and a unique, time-sensitive passcode or a waiting room feature where the host has to manually admit participants. Additionally, hosts could be given the ability to lock meetings once all expected participants have joined, preventing any further access. These measures would significantly reduce the likelihood of unauthorized access, thereby enhancing the overall security and trustworthiness of the SummarEyes platform.

These proactive measures are designed to mitigate risks while offering viable alternatives, ensuring the project's adaptability and resilience in the face of unforeseen challenges.

Risks	Likelihood	Effect	Solution
Performance and Response Time	8/10	Analysis of emotion of participants, or eye tracking can take some time which can lead to delay of generating summary or feedback of the meeting, reducing the functionality.	Limited number of users will be serviced.
Impracticability of Agora SDK	3/10	Applications for video conferences may not offer the capabilities that users want, which reduces their usability.	An alternative SDK or a custom video conferencing program can be utilized.
Accuracy of Models	6/10	The system can draw an incorrect conclusion about the attention level of the participant, resulting in decreased functionality.	A new dataset can be used or created.
Security Problem	5/10	The potential for meeting bombing	Two-step verification or waiting room in the meetings

Figure 27: Risks and Solutions Table

4.3   Project Plan

Below you can see various tables that will be made use of. The project plan can be reported by a list of work packages and their content. For better readability, a Gant chart based on work packages can also be added.

WP#	Title	Leader	Members involved
WP1	Reports	Kürşad Güzelkaya	Hazal Buluş, Deniz Gökçen, Burcu Kaplan, Gökhan Tekin
WP2	Frontend Development	Hazal Buluş	Gökhan Tekin
WP3	Backend Development	Burcu Kaplan	Deniz Gökçen
WP4	Emotion Detection	Kürşad Güzelkaya	Gökhan Tekin

WP5	Speech-to-Text	Burcu Kaplan	Kürşad Güzelkaya
WP6	Summarization	Kürşad Güzelkaya	Hazal Buluş
WP7	Eye Tracking	Kürşad Güzelkaya	Gökhan Tekin
WP8	Demo Development 1	Deniz Gökçen	Hazal Buluş, Kürşad Güzelkaya, Burcu Kaplan, Gökhan Tekin
WP9	Demo Development 2	Gökhan Tekin	Hazal Buluş, Kürşad Güzelkaya, Burcu Kaplan, Deniz Gökçen
WP10	Testing	Burcu Kaplan	Hazal Buluş, Kürşad Güzelkaya, Gökhan Tekin, Deniz Gökçen

Figure 28: Work Packages Table

WP1: Reports			
Start Date: 03.10.23 **End Date:** 15.05.24			
Leader:	Kürşad Güzelkaya	**Members involved:**	Hazal Buluş, Deniz Gökçen, Burcu Kaplan, Gökhan Tekin

Objectives: The reports' primary objective is to set some limitations and make decisions regarding the project in advance, allowing it to be constructed more quickly and clearly. The project's requirements, their analysis, and a comprehensive design are among the goals.

Tasks:

Task 1.1 Brainstorm Ideas: Try to understand the needs of the project and decide on specific functionalities for the project.

Task 1.2 Requirements Elicitation: Research and decide on functional and non-functional

requirements in order to meet the expectations of the application.

Task 1.3 Specification Report: The report is a non ambiguous documentation of requirements of the project. It can be used as a guideline by the developers while building the application.

Task 1.4 Analysis Report: The report includes the UI design, scenarios, use cases, models such as object, class, and dynamic. It also includes the project plan which will be helpful through the project.

Task 1.5 Detailed Design Report: The report is to show class packages and interfaces.

Task 1.6 Final Report: The report includes the packages, tools and technologies, frameworks that have been used while developing the project. Overall design will be included as well.

Deliverables:

D1.1: Project Website

D1.2: Project Specification Report

D1.3: Analysis Report

D1.4: Detailed Design Report

D1.5: Final Report

Figure 29: Work Package 1

WP2: Frontend Development			
Start Date: 15.11.23 **End Date:** 15.05.24			
Leader:	Hazal Buluş	**Members involved:**	Gökhan Tekin
Objectives: Making a user interface that satisfies all functional, non-functional, and UI requirements is the goal of frontend development. The frontend development goals include the application's visual components as well as user interaction tools.			
Tasks:			

Task 2.1 Welcome Screen: Brief introduction of the mobile application with the options to view frequently asked questions, terms of service, and privacy policy.

Task 2.2 Sign Up and Sign In Screens: The interactive pages allows users to sign up and sign in by working with the backend logic.

Task 2.3 Home Page: Personal home page of a user which will offer options to create/attend meetings, view summaries and feedback.

Task 2.4 Profile Page: Personal information and settings with change options.

Task 2.5 Meeting Page: Meeting view which shows cameras of the users, participant list and options to video/audio settings, summarization and leaving meeting.

Task 2.6 Summaries/Feedback Page: Summary, personalized summary and feedback report with the related meeting information.

Task 2.7 Pricing Page: Subscription packages with their price and contents.

Task 2.8 Organization Page: User adding to and removing from subscribed organization.

Task 2.9 Admin Page: User/Organization adding to and removing from application.

Deliverables:

D2.1: User Interface

Figure 30: Work Package 2

WP3: Backend Development			
Start Date: 15.11.23 **End Date:** 15.05.24			
Leader:	Burcu Kaplan	**Members involved:**	Deniz Gökçen
Objectives: Backend development involves architecting the server-side components of applications, encompassing tasks such as database modeling, business logic implementation, and data management. It's focused on creating the infrastructure that enables data storage, retrieval, and manipulation while ensuring security and scalability.			
Tasks:			

Task 3.1: Initializing Django project with correct dependencies

Task 3.2: Implementation of basic classes according to class diagram

Task 3.3: Implementing service layer

Task 3.4: Connecting the external services to service classes

Task 3.5: Implementing the controller classes

Task 3.6: Testing controller endpoints via postman

Task 3.7: Connecting the back-end with front-end

Task 3.8: Deployment of the app

Deliverables:

D3.1: The backend of the application

D3.2: API collection

Figure 31: Work Package 3

WP4: Emotion Detection			
Start Date: 20.11.23 **End Date:** 10.02.24			
Leader:	Kürşad Güzelkaya	**Members:**	Gökhan Tekin
Objectives: It involves detecting emotion from facial expressions in the given video file and converting the result to concentration measurements by using a point base system for emotions.			
Tasks: **Task 4.1:** Implementation of FER library in Python **Task 4.2:** Test library with images **Task 4.3:** Test library with videos **Task 4.4:** Process emotion results to find concentration results			

Task 4.5: Connect emotion detection to the application
Deliverables: D4.1: Emotion detection Python script

Figure 32: Work Package 4

WP5: Text-to-speech			
Start Date: 27.11.23 End Date: 15.12.23			
Leader:	Burcu Kaplan	Members:	Kürşad Güzelkaya
Objectives: Takes an audio file and produce a transcript by converting speech to text			
Tasks: Task 5.1: Implementation of Whisper AI in Python Task 5.2: Test the tool with audio files Task 5.3: Connect transcription to the application			
Deliverables: D5.1: Speech-to-text Python script			

Figure 33: Work Package 5

WP6: Summarization			
Start Date: 23.11.23 End Date: 15.05.24			
Leader:	Kürşad Güzelkaya	Members:	Hazal Buluş

Objectives: An LLM that takes emotion detection result, eye tracking result and transcript of meeting to produce a summary that contains parts mostly missed by user

Tasks:

Task 6.1: Initialize a pre-trained LLM (Llama2)

Task 6.2: Test summary feature

Task 6.3: Implement langchain with a proper prompt to produce summary

Task 6.4: Test langchain

Task 6.5: Connect summarization to the application

Deliverables:

D6.1: Langchain Python script

Figure 34: Work Package 6

WP7: Eye Tracking			
Start Date: 10.02.24 **End Date:** 15.05.24			
Leader:	Kürşad Güzelkaya	**Members:**	Gökhan Tekin

Objectives: It performs gaze estimation to detect whether user looking the screen or not

Tasks:

Task 7.1: Find a proper dataset for eye tracking

Task 7.2: Develop an ML model for eye tracking

Task 7.3: Research existing ML models and tools

Task 7.4: Compare and test ML models

Task 7.5: Connect eye tracking to the application

Deliverables:

D7.1: An ML model for eye tracking

Figure 35: Work Package 7

WP8: Demo Development 1			
Start Date: 05.12.23 **End Date:** 15.12.24			
Leader:	Deniz Gökçen	**Members:**	Hazal Buluş, Kürşad Güzelkaya, Burcu Kaplan, Gökhan Tekin

Objectives: A demo will be given to highlight the project and the advancements made in development. There will be a live demo in addition to a presentation. The goal is to ensure that the demo is professionally done and presents the work in an engaging way.

Tasks:

Task 8.1 Demo Presentation : Create a powerpoint presentation that covers the description, completed features and the future work.

Task 8.2 Demo Application: Have an application that has the features of online meeting. Show examples of emotion detection and summarization. Create sample users to present in the demo and show all the completed features up to that point.

Deliverables:

D8.1: Live demo of the working web application

D8.2: Demo Presentation

Figure 36: Work Package 8

WP9: Demo Development 2
Start Date: 01.05.24 **End Date:** 15.05.24

Leader:	Gökhan Tekin	**Members:**	Hazal Buluş, Kürşad Güzelkaya, Burcu Kaplan, Deniz Gökçen

Objectives: A demo will be given to highlight all features and developments of the project. There will be a live demo in addition to a presentation. The goal is to ensure that the demo is professionally done and presents the work in an engaging way.

Tasks:

Task 9.1 Demo Presentation: A powerpoint presentation that covers the description, completed features and the future work.

Task 9.2 Demo Application: Have an application that has the features of online meeting, emotion detection, eye detector, summary and feedback. Create sample users to present in a demo and show all the features of the application.

Deliverables:

D9.1: Live demo of the working web application

D9.2: Demo Presentation

Figure 37: Work Package 9

WP10: Testing			
Start Date: 01.12.23 **End Date:** 15.05.24			
Leader:	Burcu Kaplan	**Members:**	Hazal Buluş, Kürşad Güzelkaya, Deniz Gökçen, Gökhan Tekin

Objectives: Test the emotion detection, speech-to-text transcription and summarization. Furthermore, test the frontend and backend of the application to observe any possible errors and issues.

Tasks:

Task 10.1 Test Emotion Detection: The emotion detection will be tested through application to see if it is giving accurate results.

Task 10.2 Test Transcription: The speech-to-text features will be tested through web application to see if it is giving accurate results.

Task 10.3 Test Summarization: The summarization features will be tested through web application to see if it is giving accurate results.

Task 10.4 Test WebApplication: The web application will be tested to see if there are any bugs or issues within the frontend or the backend.

Deliverables:

D10.1: Test results

D10.2: Bug fixes

Figure 38: Work Package 10
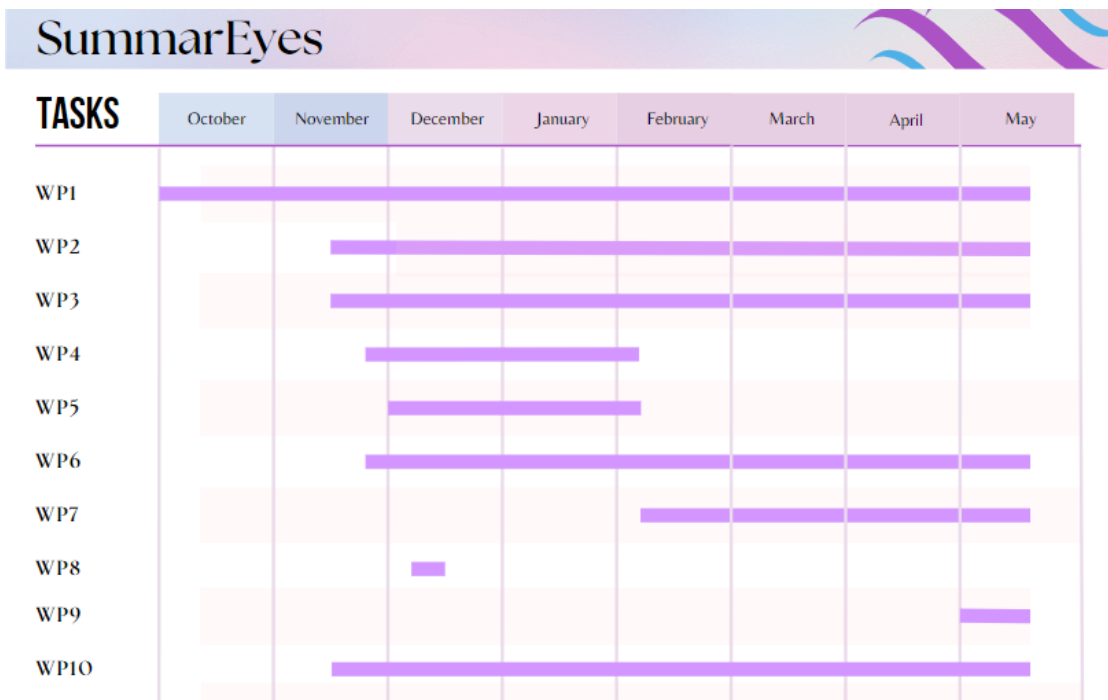
Figure 39: Gant chart displaying work packages and time period

4.4 Ensuring Proper Teamwork

We have five team members in our team, so teamwork is significant for us to develop our project in time. We divided the total work into three parts which are frontend development, backend development and ML development. They are assigned to team members according to their background knowledge and workload.

- Frontend Development: Hazal Buluş, Süleyman Gökhan Tekin

- Backend Development: Deniz Gökçen, Burcu Kaplan
- ML Development: Kürşad Güzelkaya

Frontend developers are working on the development of the user interface of the website while backend developers are managing databases, Agora SDK for online meeting tools and server side logic. The ML developer is working on ML technologies that are used in our project like Whisper AI, Pretrained Language Models, Emotion Detection with FER library and eye tracking system.

We are meeting every week to keep in touch, let each other know about the developments that are performed and share works to do. To monitor and follow developments and responsibilities we are using Jira Software.

4.5   Ethics and Professional Responsibilities

In the development of our project, we place utmost importance on ethical practices and professional responsibilities, particularly in relation to user data privacy and compliance with legal standards. Our system makes sure that no personally identifiable information, including images or videos, is kept in order to protect user privacy. We closely adhere to data reduction standards, retaining just the minimum data needed for emotion analysis. This method protects against the misuse of sensitive information by complying with data protection laws and ethical norms.

We're devoted to maintaining openness regarding our data usage guidelines. User data is never shared with outside parties for profit; instead, it is utilized only to improve user experience and service quality. Additionally, we carefully consider licenses that protect privacy and do not require the disclosure of enhancements when using open-source software, so we can maintain user confidentiality. Our procedures are continuously reviewed to make sure they still meet changing user expectations and standards for ethics.

By following these guidelines, we hope to maintain the greatest standards of moral purity and professional accountability, guaranteeing our users' trust and security constantly.

4.6   Planning for New Knowledge and Learning Strategies

Creating a thorough plan for improving knowledge and learning techniques necessitates focusing on particular topics. In order to improve knowledge of Machine Learning (ML), specifically Convolutional Neural Networks (CNNs) for eye tracking, the emphasis will be on participating in specialized online tutorials and courses that are accessible on sites like Coursera and Fast.ai. Theoretical knowledge will be strengthened through practical application through projects utilizing publicly available eye-tracking datasets such as MPIIGaze and GazeCapture. Acquaintance with scholarly papers and publications will also guarantee that you are in line with the most recent developments in CNNs for eye tracking.

Parallel to this, learning how to deploy Django applications on platforms such as Python Anywhere will need you to delve deeply into the official Django documentation and then put it into practice by following online tutorials and guidelines. Deployment techniques inside Django

will be solidified through the creation of example projects and their subsequent deployment across many platforms.

In addition, exploring Language Chains (LangChain) and Language Models (LLM) will need learning about their principles and uses from already-existing materials and playing with related APIs. A small-scale website project including LangChain or LLM will show how useful these technologies are for improving text processing or language-related features.

All things considered, a systematic and successful learning journey throughout these targeted disciplines will be ensured by upholding regular learning methods, involvement with practical projects, active participation in communities linked to ML, Django, and NLP, along with careful documentation and note-taking. Achieving the stated objectives will require modifying timetables and goals in accordance with the changing interests and pace of learning.

5   Glossary

LLM: Large Language Model is a deep learning algorithm that can perform a variety of natural language processing (NLP) tasks. Ex: GPT-4 and Llama2.

Langchain: A framework for developing applications powered by language models.

6   References

[1] "Enabling Meeting Summary with AI Companion", Zoom, Jul. 11. 2023. Available: https://support.zoom.com/hc/en/article?id=zm_kb&sysparm_article=KB0057960.   [Accessed: 29-Nov-2023].

[2] "Generate a Meeting Summary," Microsoft, Sept. 1, 2023. Available: https://learn.microsoft.com/en-us/microsoft-sales-copilot/generate-meeting-summary. [Accessed: 29-Nov-2023].

[3] "fer," PyPI, Jun. 08, 2023. Available: https://pypi.org/project/fer/. [Accessed: 6-Dec-2023]

[4] "Pricing". Available: https://www.agora.io/en/pricing/. [Accessed: 7-Dec-2023].

[5] "Pricing". Available: https://openai.com/pricing. [Accessed: 6-Dec-2023].

[6] "Introducing Llama 2". Available: https://ai.meta.com/llama/. [Accessed: 6-Dec-2023].

[7] "Zoom Fatigue – Why video calls are so exhausting," allianzcare.com. [Online]. Available: https://www.allianzcare.com/en/about-us/blog/2021/03/zoom-fatigue.html.   [Accessed: 6-Dec-2023].

[8] Object-Oriented Software Engineering, Using UML, Patterns, and Java, 2nd Edition, by Bernd Bruegge and Allen H. Dutoit, Prentice-Hall, 2004, ISBN: 0-13-047110-0.