ÇANKAYA UNIVERSITY FACULTY OF ENGINEERING

COMPUTER ENGINEERING DEPARTMENT

CENG 407

Innovative System Design and Development I Project Report

Team ID: 202421

SenseAI

Ece Heval Ünal - 202111014
Zeliha Aybüke Baştürk - 202028034
Özge Alkan - 202128406
Mustafa Tozman – 202128014
Sadrettin Anıl Karaçay - 202011046

Advisor:
Dr. Atila BOSTAN
Dr. İsmail Bora ÇELİKKALE

Contents

Project Plan

Week		1	2	3	5	6	7	8	9	10	11	12	13	14	15	16
CENG407 Deadlines	Progress Status	11 October 2024	18 October 2024	25 October 2024	1 November 2024	8 November 2024	15 November 2024	22 November 2024	29 November 2024	6 December 2024	13 December 2024	20 December 2024	27 December 2024	3 January 2025	10 January 2025	17 January 2024
Team Setup	Completed															
Project Proposal Form	Completed															
Project Selection Form	Completed															
GitHub Repository	Completed															
Project Work Plan	Completed															
Literature Review	Completed															
Software Requirements Specification	Completed															
Project Webpage	Completed															
Software Design Description	Completed															
Project Report/Project Tracking Form	Completed															
Presentation	In Progress															

Literature Review

Abstract

This paper reviews the literature on multimodal, AI-driven psychology chatbots that aim to assess and support emotional well-being by integrating inputs from speech, facial expressions, text, and biometric data. Leveraging advances in Artificial Intelligence (AI), Natural Language Processing (NLP), computer vision, and signal processing, such systems enable real-time emotional analysis and personalized
feedback. This literature review covers key research and technologies underpinning the project, including sentiment analysis, computer vision, wearable devices, and multimodal data fusion, and examines their applications in psychological support tools. The review highlights advances in sentiment analysis with transformers, facial recognition with convolutional neural networks (CNNs), and biometric analysis through machine learning models, providing a roadmap for developing comprehensive, real-time emotional tracking systems.

Introduction

Recent years have seen significant advances in digital mental health tools, particularly AI-driven systems that analyze and interpret emotional states for improved psychological support. The growth in wearable technologies and machine learning models has created new opportunities to integrate multimodal data sources, such as text, speech, facial expressions, and biometric signals, allowing for a holistic understanding of emotional well-being. This project aims to build a psychology chatbot that leverages multimodal data to analyze emotional states, offer personalized feedback, and track users' emotional patterns over time. This literature review outlines relevant studies on multimodal emotional analysis, describes technologies such as NLP and computer vision, and discusses the ethical considerations involved in mental health chatbot development.

1. Background and Motivation

1.1 The Role of Chatbots in Mental Health Support

In recent years, digital tools have increasingly played a significant role in supporting mental health, with AI-powered chatbot systems emerging as key tools for addressing emotional needs. Mental health support chatbots provide users with a safe, non-judgmental environment to express their emotions and thoughts, offering a valuable space for self-reflection, stress management, and basic psychological support.

Research has shown that even text-based chatbots have a positive impact on users' mental well-being. For instance, applications like Wysa and Woebot employ Cognitive Behavioral Therapy (CBT) techniques to help users manage stress and anxiety effectively. However, these text-based systems have inherent limitations, as they rely solely on verbal input to interpret users' emotions and thoughts. This project, SenseAI, aims to overcome these limitations by incorporating multimodal

data—such as voice tone, facial expressions, and biometric signals—to enable a more comprehensive emotional assessment, thus fostering more empathetic and accurate interactions.

The potential for chatbots to offer mental health support is immense, particularly as they can provide immediate assistance in times of need, help track emotional changes over time, and offer coping strategies that may reduce the burden on traditional mental health services. By incorporating multimodal emotional analysis, SenseAI seeks to deliver a more human-like, empathetic interaction that surpasses the capabilities of standard text-based chatbot solutions.

1.2 Advancements in Multimodal Emotional Analysis

The field of emotion recognition has progressed significantly, moving beyond traditional text-based sentiment analysis to sophisticated multimodal approaches that integrate visual and physiological data. Multimodal emotional analysis enables more accurate predictions by combining inputs from various sources, such as voice tone, facial expressions, and heart rate variability (HRV).

Multimodal emotion recognition systems leverage physical and behavioral indicators—such as speech tone, facial cues, and biometric signals—to construct a nuanced understanding of users' emotional states. For instance, HRV is a valuable biometric indicator often associated with stress levels, while facial expression analysis can reveal emotions like happiness, sadness, or anger. By combining these modalities, multimodal systems achieve higher accuracy and allow for a deeper understanding of the user's emotional context, ultimately enabling more empathetic responses in chatbot interactions. This project leverages these advancements to build a mental health chatbot that provides robust emotional support by integrating multimodal data sources to improve both the relevance and empathy of its responses.

2. Key Technologies and Methodologies

2.1 Natural Language Processing (NLP) for Sentiment Analysis

Natural Language Processing (NLP) plays a fundamental role in understanding the emotional tone conveyed in a user's text or speech input. In recent years, advancements in transformer-based models like BERT (Bidirectional Encoder Representations from Transformers) and GPT (Generative Pre-trained Transformer) have significantly improved the ability to capture complex linguistic patterns, making them ideal for sentiment analysis tasks. Transformers use self-attention mechanisms that allow models to consider the context of each word within a sentence, capturing subtle nuances in language that can reflect emotional states.

In this project, transformer models will be fine-tuned specifically for detecting emotional cues in conversational text. Research suggests that when sentiment analysis is combined with other emotional indicators, such as voice and facial expressions, the overall accuracy and reliability of emotion detection improve. This approach provides a solid foundation for a multimodal chatbot system like SenseAI, enabling it to interpret and respond to user inputs with greater emotional sensitivity and accuracy.

2.2 Computer Vision for Facial Expression Analysis

Computer vision, particularly using Convolutional Neural Networks (CNNs), is essential for analyzing facial expressions that convey emotions. CNNs have proven highly effective in identifying facial features and understanding emotions such as happiness, sadness, or anger through the recognition of subtle changes in expressions. Studies have shown that combining facial expression data with other modalities, like text and biometric signals, enhances the accuracy of emotion recognition.

For the SenseAI project, pre-trained CNN models, such as those trained on the FER+ or AffectNet datasets, can be adapted to recognize real-time facial expressions. This functionality enables the chatbot to respond in a manner that reflects empathy and aligns with the user's emotional state. Facial recognition techniques are crucial for providing non-verbal cues, which add depth to the chatbot's responses and allow for a more human-like interaction.

2.3 Biometric Data Analysis from Wearable Devices

Biometric data, such as heart rate variability (HRV) and skin conductance, provide valuable insights into a user's emotional arousal and stress levels. Wearable devices can capture this data in real time, offering an additional layer of information that complements text and visual inputs. Machine learning algorithms can process these biometric signals to detect emotional states, such as anxiety or calmness, by analyzing patterns over time.

The literature indicates that combining wearable-derived data with other modalities improves the accuracy of emotional assessments. For instance, HRV is a commonly used metric in stress detection, as fluctuations in heart rate are associated with changes in emotional arousal. In the SenseAI project, biometric data will be processed using time-series machine learning models to identify stress patterns, allowing the chatbot to offer timely interventions and personalized feedback based on the user's physiological state.

3. Multimodal Data Fusion

3.1 The Importance of Multimodal Fusion in Emotion Recognition

Multimodal fusion combines data from various sources—text, visual, and biometric—to create a comprehensive emotional profile. Research has demonstrated that multimodal fusion techniques, such as feature-level and decision-level fusion, yield significantly higher accuracy in emotion detection by integrating complementary data streams. Each modality brings unique insights: while text may reveal explicit thoughts or concerns, facial expressions and biometric data provide non-verbal cues that deepen the understanding of a user's emotional state.

For psychological applications like SenseAI, multimodal fusion is crucial to improving both response relevance and empathic accuracy. By integrating various emotional indicators, the chatbot can respond in a manner that aligns more closely with the user's overall emotional profile. This feature is essential in mental health contexts, where nuanced and context-aware responses foster a supportive and safe environment for the user.

3.2 Data Fusion in Multimodal Emotion Recognition

Data fusion in multimodal emotion recognition involves combining insights from various data sources to enhance the accuracy, reliability, and robustness of emotional assessments. In SenseAI, fusion techniques are applied across different levels to create a nuanced understanding of emotional states. Below are the relevant fusion levels utilized within the project:

- **Low-Level Fusion (Raw Data Combination):** At this level, raw data from each modality—such as speech tone, facial expressions, and heart rate—is collected and processed in parallel. This approach provides a holistic view of the user's physiological and behavioral states, forming a foundation for more detailed feature extraction and emotional analysis.
- **Feature-Level Fusion:** In feature-level fusion, specific features indicative of emotions are extracted from each modality and then combined to create a unified representation of the user's emotional state. For example, speech features may include pitch and tone, facial expressions may focus on key landmarks like eyebrow movements or mouth curvature, and biometric data may capture HRV or skin conductance. By merging these features, the system leverages the complementary strengths of each modality to provide a more accurate and dynamic emotional profile.
- **Decision-Level Fusion:** In decision-level fusion, separate predictions are generated from each modality and then combined to form a final emotional assessment. This level relies on algorithms, such as voting schemes or weighted decision models, that merge outputs from each modality, allowing for increased accuracy by mitigating the limitations of individual data sources. This method enhances robustness in emotional recognition, as it synthesizes

information from text, visual, and biometric sources to deliver consistent emotional assessments.

4. Data Sources and Datasets

A variety of datasets are essential for training multimodal emotion recognition models, as they provide a foundation for developing accurate and reliable AI-driven psychology chatbots. SenseAI will utilize datasets that span textual, visual, and biometric data to ensure comprehensive emotional understanding across different modalities.

- **Text and Speech Datasets:** For sentiment and emotional tone detection, datasets such as the **IEMOCAP** (Interactive Emotional Dyadic Motion Capture) and **EmoReact** are instrumental. These datasets contain annotated conversational text and speech samples labeled with emotions like happiness, sadness, and anger, providing a robust foundation for training Natural Language Processing (NLP) models to detect emotional cues. Multilingual resources, such as **EmoReact**, help the model adapt to users from diverse linguistic backgrounds, enhancing the inclusivity and applicability of the chatbot.
- **Facial Expression Datasets:** Datasets like **FER+ (Facial Expression Recognition)** and **AffectNet** provide large collections of labeled facial images that are crucial for facial emotion recognition models. These datasets are used to train Convolutional Neural Networks (CNNs) to analyze facial expressions in real time, supporting emotion recognition based on visual input. By incorporating these datasets, SenseAI can assess non-verbal cues like smiles, frowns, and other subtle facial movements, making its responses more empathetic and accurate.
- **Biometric Data Datasets:** For emotional detection through biometric signals, the **WESAD (Wearable Stress and Affect Detection)** dataset is a valuable resource. WESAD includes sensor data collected from wearable devices, capturing physiological indicators such as heart rate, skin conductance, and respiratory rate. This dataset allows machine learning models to understand and interpret stress and arousal levels in real-time, enabling SenseAI to respond to the physiological states of users dynamically and with heightened emotional sensitivity.

By combining these diverse datasets, SenseAI will achieve a balanced understanding of user emotions across modalities, allowing for holistic emotional assessments. The diversity and quality of these datasets play a critical role in building a chatbot that adapts accurately to various emotional cues and supports a broad range of users.

5. Model Training

Training the emotion recognition model in SenseAI involves processing multimodal data—including text, facial expressions, and biometric signals—to accurately assess and track emotional states. Each data modality is processed individually to harness its unique features, followed by integration to create a cohesive emotional profile.
Below is an outline of the training approach for each modality:

- **Text Data Processing (NLP):** Text data, derived from user interactions, is tokenized and fed into a transformer-based model, such as BERT or GPT. These models are fine-tuned to detect subtle emotional cues in conversational text, leveraging attention mechanisms to capture the context and sentiment behind user statements. Transformer models are particularly effective in understanding nuances, making them an essential component for analyzing language that conveys emotional depth.
- **Facial Expression Processing (Computer Vision):** Facial expression data is processed using Convolutional Neural Networks (CNNs), which are pre-trained on extensive facial emotion datasets like FER+ and AffectNet. These CNN models are then fine-tuned to detect relevant facial expressions in real-time video feeds, enabling SenseAI to capture emotional changes in users' facial expressions during conversations. By recognizing expressions such as smiles, frowns, and surprise, the chatbot can adjust its responses to better align with the user's current emotional state.
- **Biometric Signal Processing (Time-Series Analysis):** Biometric signals, including heart rate and skin conductance, are analyzed using time-series machine learning models that can detect patterns associated with emotional arousal and stress. For instance, fluctuations in heart rate variability (HRV) are closely associated with stress levels, making HRV a key metric for emotional analysis. By training on the WESAD dataset, these models learn to interpret biometric signals in real-time, adding an additional dimension of emotional awareness to the chatbot's response generation.

Performance Metrics and Model Evaluation: The performance of each model is evaluated using accuracy, precision, and recall metrics, aiming for a balanced trade-off between true positive and false positive rates. Cross-modal consistency is also monitored to ensure that emotional predictions remain coherent across different data types, enhancing the model's reliability. This cross-checking mechanism helps refine the chatbot's responses, making them more aligned with the user's overall emotional state.

Cross-Modality Training: A unique feature of SenseAI's model training is the integration of cross-modality learning, where the predictions from one modality (e.g.,

text sentiment) are cross-validated with data from another (e.g., facial expressions or biometric signals). This approach helps improve the accuracy and consistency of the emotional assessment, providing a holistic view that draws from multiple data points. This training approach enables SenseAI to deliver nuanced, data-driven emotional insights, making it a reliable tool for mental health support.

6. Challenges

The development of a multimodal, AI-driven chatbot for mental health support presents various challenges that must be addressed to ensure effectiveness, reliability, and user trust. Below are some of the primary technical, ethical, and practical challenges encountered in this project, along with proposed solutions:

6.1 Data Privacy and Security

Handling multimodal data—particularly sensitive information such as biometric signals and facial expressions—raises significant privacy concerns. Users may be reluctant to share personal data if they are unsure of its security or potential misuse. Ensuring the confidentiality of this information is paramount in mental health applications.

- **Solution:** Data anonymization and secure data storage techniques, such as end-to-end encryption, can mitigate privacy risks. Additionally, implementing data minimization practices ensures that only essential information is collected. Informed consent and transparency about data use further build trust. Techniques such as federated learning, where data is processed locally on the user's device, can be explored to enhance privacy while minimizing the transfer of sensitive data.

6.2 Computational Efficiency and Real-Time Processing

Multimodal data fusion and real-time emotion recognition require substantial computational resources, especially when integrating multiple complex models, such as CNNs for facial recognition and transformers for NLP. For the system to deliver timely and responsive feedback, computational efficiency is critical.

- **Solution:** Optimization strategies such as model pruning and quantization can reduce the computational load, making real-time processing more feasible. Additionally, exploring edge computing—where data processing occurs locally on the device rather than in the cloud—can reduce latency and enhance user experience. By using lightweight, pre-trained models and adapting them for on-device processing, SenseAI can improve responsiveness without sacrificing accuracy.

6.3 Emotional Sensitivity and User Safety

In mental health applications, incorrect or insensitive responses can lead to emotional distress, undermining the effectiveness of the chatbot. If the chatbot misinterprets or overreacts to certain emotions, it could inadvertently exacerbate a user's emotional state rather than providing support.

- **Solution:** SenseAI incorporates a feedback mechanism where users can rate responses, helping the system improve over time based on user feedback. Additionally, the chatbot's responses are calibrated to avoid overly assertive or diagnostic language, maintaining a supportive tone that prioritizes empathy and encouragement. Safeguards such as a crisis detection module can alert users to seek human assistance in cases where severe distress is detected.

6.4 Ethical Concerns in Emotional Analysis

Automated emotional analysis introduces ethical considerations, especially regarding potential biases in the models. For instance, emotional recognition models trained on specific demographic datasets may not generalize well to diverse populations, potentially leading to inaccurate assessments.

- **Solution:** To address this, SenseAI utilizes diverse, inclusive datasets for model training to minimize biases and enhance the generalizability of emotional assessments across different demographic groups. Regular audits and model evaluations are conducted to detect any unintended biases. Furthermore, the project aligns with established ethical frameworks, such as the IEEE's Ethically Aligned Design guidelines, to ensure responsible development and deployment.

6.5 Model Interpretability

The complex nature of multimodal systems—especially those involving neural networks—can make it challenging to interpret model decisions. This lack of transparency can be problematic in mental health applications, where understanding the reasoning behind certain assessments is valuable for user trust.

- **Solution:** Explainable AI (XAI) techniques can be applied to enhance the interpretability of the model's outputs. For example, attention visualization for transformer models in NLP or heatmaps for CNN-based facial recognition can provide insights into which inputs influenced the emotional prediction. By offering transparent explanations for its assessments, SenseAI builds user confidence and allows for greater accountability.

7. Ethical and Practical Considerations

The development of a multimodal psychology chatbot like SenseAI raises critical ethical and practical concerns, especially given the sensitive nature of emotional and

biometric data. This section outlines the primary ethical guidelines followed and practical strategies adopted to ensure that SenseAI remains a responsible, trustworthy, and supportive mental health tool.

7.1 Privacy and Data Security

Ensuring the privacy and security of user data is fundamental, as SenseAI collects sensitive multimodal data, including text inputs, facial expressions, and biometric signals. Unauthorized access or misuse of this information could lead to privacy violations and erode user trust.

- **Privacy-by-Design:** SenseAI incorporates privacy from the outset by embedding security features directly into the system architecture. Data encryption, secure data storage, and anonymization techniques help protect user information. For instance, all user data is encrypted both in transit and at rest, reducing the risk of unauthorized access.
- **User Control and Consent:** Transparency and user control are prioritized, allowing users to choose which data modalities they wish to share (e.g., opting out of facial expression analysis). Clear consent forms inform users of how their data will be used, processed, and stored, allowing them to make informed choices.
- **Federated Learning for Enhanced Privacy:** To minimize the need for transferring sensitive data, federated learning approaches are explored, where models are trained on-device, and only anonymized updates are shared with central servers. This method significantly reduces data transfer, enhancing privacy without compromising model effectiveness.

7.2 Emotional Sensitivity

Chatbots designed for mental health must be particularly cautious in handling user emotions. Misinterpreting or inadequately responding to emotions can result in user distress or harm. Emotional sensitivity is critical, especially when addressing users experiencing negative emotional states.

- **Empathetic Response Design:** SenseAI's response system is calibrated to prioritize empathy, validation, and encouragement, rather than providing definitive diagnostic statements. The chatbot's responses are designed to be supportive, reassuring, and reflective of the user's emotional state, ensuring that users feel heard and validated.
- **Crisis Detection and Escalation:** Recognizing that some emotional states may require professional intervention, SenseAI includes a crisis detection module. This module identifies severe emotional distress (e.g., high levels of anxiety or depressive indicators) and provides resources, such as helpline

information or the option to speak with a mental health professional, if available.

7.3 Bias and Fairness

Bias in AI systems can lead to inaccuracies, especially if the training data lacks diversity. For a mental health chatbot, biased emotion recognition could result in unintentional misunderstandings, particularly for users from different cultural backgrounds.

- **Inclusive and Diverse Training Data:** SenseAI's emotion recognition models are trained on datasets that represent diverse populations across age, gender, ethnicity, and cultural backgrounds. Regular audits and evaluations are performed to ensure that the chatbot remains fair and accurate for all users, regardless of demographic differences.
- **Continuous Bias Monitoring:** The models are continually monitored and adjusted based on user feedback to identify and address any emerging biases. By integrating explainable AI methods, SenseAI can make model predictions more transparent, allowing developers to understand and mitigate biases effectively.

7.4 Ethical Framework Compliance

SenseAI aligns with established ethical frameworks to ensure that its development and deployment adhere to accepted standards in AI ethics, especially those concerning user welfare and accountability.

- **IEEE Ethically Aligned Design and WHO Guidelines:** SenseAI follows ethical guidelines set forth by reputable organizations like the IEEE and the World Health Organization (WHO). These frameworks provide principles for ensuring user welfare, fairness, and accountability in AI systems, particularly within healthcare and mental health applications.
- **Informed Consent and Transparency:** Ethical AI deployment requires that users are fully informed about the data collected, how it will be used, and their rights to access or delete their data. SenseAI ensures that all users are aware of these aspects and provides them with access to their personal data history, ensuring transparency and fostering trust.

8. Related Articles

Research in multimodal emotion recognition and AI-driven mental health support has expanded in recent years. This section reviews notable studies and advancements that provide a foundation for SenseAI's approach to multimodal emotional analysis.

- **Multimodal Emotion Recognition Based on Facial Expressions, Speech, and EEG**
  Wang et al. (2022) propose a deep learning-based multimodal emotion recognition system called “Deep-Emotion,” which combines facial expressions, voice, and electroencephalogram (EEG) data for high accuracy in emotion recognition. This study highlights the potential of integrating physiological signals for robust emotional analysis, which informs SenseAI’s approach in utilizing biometric data alongside text and visual inputs to capture user emotions comprehensively.[1]
- **A Survey of Deep Learning-Based Multimodal Emotion Recognition: Speech, Text, and Face**
  Lian et al. (2023) examine various emotion recognition techniques across modalities like text, audio, and facial expressions. This survey categorizes deep learning models based on their performance in emotion recognition and offers a comparison of popular datasets. The findings emphasize the importance of choosing appropriate datasets and techniques for each modality, a principle applied in SenseAI’s model selection and data fusion methods.[2]
- **Multimodal Emotion Recognition Based on Facial Expressions, Speech, and Body Gestures**
  Yan et al. (2024) explore a three-modality system for emotion recognition, integrating facial expressions, speech, and body gestures to enhance recognition accuracy. This study demonstrates the benefits of analyzing
  non-verbal cues like gestures in emotional detection, underscoring the value of multimodal fusion. Although SenseAI does not use body gestures, this study supports the efficacy of combining visual and auditory data to achieve a more nuanced emotional understanding.[3]
- **Depression Detection Using Multimodal Analysis with Chatbot Support**
  Sharma et al. (2023) investigate an AI system for early depression detection through multimodal analysis, using text, audio, and image data. This approach aligns closely with SenseAI’s objectives, as both systems aim to provide
  real-time emotional support using AI-driven insights. The study emphasizes the role of multimodal data in detecting nuanced emotional states, reinforcing SenseAI’s focus on combining text, facial expressions, and biometric data for accurate emotional tracking.[4]

9. Similar Projects

Several existing projects have explored the use of AI-driven chatbots in mental health support, each with unique features and limitations. Reviewing these projects provides context for SenseAI's unique approach to multimodal integration.

9.1 Only Text-Based Chatbots

Most psychology-focused chatbots currently rely solely on text input to gauge user emotions and provide support. These include:

- **Woebot**: Uses Cognitive Behavioral Therapy (CBT) techniques to help users manage emotions. It focuses on emotional support but lacks multimodal analysis capabilities, relying only on text-based interactions.[5]
- **Wysa**: Offers tools for stress and anxiety management, including CBT and mindfulness exercises. The premium version provides access to therapists, though the chatbot's emotional understanding is limited to text.[6]
- **Youper**: Combines mood tracking with CBT techniques to enhance emotional awareness. It directs users to therapists when needed, although its capabilities are restricted to analyzing textual inputs.[7]
- **Headspace**: Primarily focused on stress and anxiety management, allowing users to connect with mental health professionals. While effective for guidance, it does not leverage multimodal data to improve emotional understanding.[8]

9.2 Ellie (SimSensei)

Developed at the University of Southern California and supported by DARPA, **Ellie** is an AI-driven system designed to assess mental health, primarily for military veterans.
It combines facial expression analysis and voice tone detection, offering a closer parallel to SenseAI's multimodal approach. Ellie leverages tools like Stanford NLP, OpenFace, and OpenCV to analyze user emotions, with techniques like Pitch Tracking and Vocal Timbre Analysis to improve emotional accuracy. While Ellie focuses on veterans' mental health assessment, its multimodal architecture validates the use of both facial and vocal data in emotion recognition.[9]

9.3 Replika AI

Replika AI provides a virtual companion for users, allowing communication through text and voice, though it does not utilize image or biometric data for emotion recognition. While Replika aims to fulfill companionship needs, it does not offer advanced emotional support features based on real-time multimodal inputs. Replika's focus on companionship rather than psychological support sets it apart from SenseAI's goal of delivering empathetic and clinically supportive interactions.[10]

10. Conclusion

The development of a multimodal, AI-driven psychology chatbot such as SenseAI represents a significant step forward in digital mental health support. By integrating advanced technologies in Natural Language Processing (NLP), computer vision, and biometric data analysis, SenseAI aims to provide users with personalized, context-aware emotional assistance that goes beyond traditional, text-only chatbots. Leveraging multimodal data fusion enables SenseAI to interpret users' emotional states with greater depth and accuracy, enhancing the empathy and relevance of its responses.

This literature review has explored the key technologies, datasets, fusion techniques, and ethical considerations essential for creating a chatbot capable of real-time emotional tracking. The use of transformer models for sentiment analysis, CNNs for facial expression recognition, and machine learning for biometric analysis provides a robust framework that addresses the complexities of emotional detection across multiple data sources. Additionally, the ethical guidelines and privacy measures adopted by SenseAI ensure that user data is handled with utmost care, fostering trust and transparency in a highly sensitive domain.

SenseAI's approach stands out in its commitment to creating a truly empathetic and adaptive mental health tool. By focusing on both technological innovation and ethical integrity, SenseAI aims to set a new standard in AI-driven emotional support, providing users with a safe, responsive, and supportive virtual companion that can positively impact mental well-being.

Software Requirements Specification

1 Introduction

1.1 Purpose

The purpose of this Software Requirements Specification (SRS) is to provide a comprehensive overview of the functional and non-functional requirements for the **SenseAI** project. This document serves as a foundational guide for the design, development, and deployment of the application. It ensures all stakeholders have a shared understanding of the system's objectives and specifications while reducing ambiguity and promoting consistency during the project's lifecycle.

The primary audience for this SRS includes:

- **Development Team:** To implement the features and functionality outlined in the requirements.
- **Project Stakeholders:** To ensure the application aligns with business objectives and user needs.
- **Quality Assurance Team:** To validate that the system meets the specified requirements.
- **End-User Representatives:** To provide feedback and confirm the application addresses their expectations.
- **Future Developers/Maintainers:** To understand the original design and purpose for updates or maintenance.

This document is designed to be accessible and useful for both technical and non-technical stakeholders.

1.2 Scope

The software product to be developed is **SenseAI**, a mobile psychology chatbot application designed to analyze emotional states and assist users in managing their mental well-being.

What SenseAI will do:

- **Camera Integration:** Analyze facial expressions to interpret the user's emotional state.
- **Speech-to-Text (STT):** Convert user speech to text for processing by the AI.
- **Heart Rate Monitoring:** Retrieve BPM data from connected smartwatches for additional mood analysis. (Google Fit will be utilised to achieve this)
- **Chatbot Support:** Provide AI-driven conversational support to guide users through mental wellness strategies.
- **Recommendations:** Offer personalized activities or content based on user mood and inputs.

What SenseAI will not do:

- SenseAI will not replace professional therapy sessions.

Application Benefits and Goals:

SenseAI is targeted at individuals seeking a convenient, AI-powered companion for mental health support. It aims to:

- Improve accessibility to basic mental wellness tools.
- Provide emotional feedback and personalized suggestions.
- Leverage non-invasive biometric data to enhance user experience.
- Deliver a user-friendly interface for seamless interaction.
- Ensure data security and user privacy.

1.3 Definitions, acronyms, and abbreviations

Term	Description
CBT	Cognitive Behavioral Therapy
AI	Artificial Intelligence
NLP	Natural Language Processing
CNN	Convolutional Neural Network
DBMS	Database Management System
SRS	Software Requirements Specification
HA	High Availability

1.4 Overview

This SRS contains a comprehensive explanation of the system's functional and non-functional requirements and interfaces. It describes the features and capabilities SenseAI will provide, outlines system design constraints, assumptions, and dependencies to guide development.

This SRS is separated into different sections to increase readability, these sections include:

Overall description (Section 2)

A section for general description of the product and product perspective.

Specific Requirements (Section 3)

A section for describing both functional and non-functional requirements, design constraints, and software interfaces.

References (Section 4)

The section containing the articles in which we determined our requirements and the documents that helped us create our report.

2 Overall description

2.1 Product perspective

This project aims to develop a multi-modal psychology chatbot that analyzes users' emotional states and contexts and helps people with a psychological guideline. This chatbot will perform emotional analysis using biometric data obtained from voice, facial expressions and wearable devices (e.g. smart watches) and provide personalized feedback. Users can use this application when they need instant support, check their old data and perform emotional state analysis. The application can be authenticated with the authentication API. Users' previous conversations will be deleted after a certain period of time in the database.

2.1.1 System interfaces

User Interface: The interface through which users interact with the chatbot. This interface receives input from the user, such as voice, facial expressions, biometric data, and text input.

Voice Recognition Interface: Analyzes the user's voice commands and emotional tone.

Image Processing Interface: A system interface that analyzes the user's facial expressions.

Biometric Data Interface: An interface that analyzes the biometric data the user receives from wearable devices (e.g. smart watches).

API Interfaces: An interface that analyzes both historical and real time data from the user and provides instant feedback and suggestions by enabling the chatbot to integrate with external services (emotion analysis, biometric data collection, etc.). It is required for the client and server to communicate, allowing real-time data exchange and feedback. It ensures that requests sent by the client (user's device) are properly handled and processed by the server, which then returns appropriate responses to enrich the user experience.

Database Interface: An interface where user data is stored and queried.

2.1.2 User interfaces

- The interfaces in the system should be designed to be user-friendly, guiding and easy to use.
- The login screen should be large enough to enter the username and password and should be adaptable to different screens. There should be no elements that will make it difficult for the user to click on a field or impair the readability of the interface.

- The buttons that the user will click on to select the desired data type on the homepage should be large enough and the fonts should be readable.
- The home screen should provide users with easy access to the new chat, previous chats and the user page.
- The menu structure should be intuitive so that users can easily access the information they need. The main menu should provide users with quick access to the desired section.
- Warning messages should provide clear and direct information to the user.
- The interface should allow users to perform functions quickly and effectively.
- The steps required for any operation should be minimized. Unnecessary complexity should be avoided and a user should be able to easily use this application with an average of 5 minutes of training.
- The interfaces should be designed in a way that the user can easily remember how to use the application.
- The application shall provide support for Dark Mode on both Android and iOS platforms. Dark Mode should be available as an optional interface theme for users who prefer a darker, more visually comfortable interface. This feature will automatically respect the device's system-wide theme settings (if enabled) and allow the user to toggle between Light and Dark modes from within the app settings.
- The application shall ensure that the user interface (UI) remains consistent and responsive (e.g. enabling screen rotation) across different screen sizes, resolutions, and platforms (Android and iOS). The design will be adaptive, ensuring that the app provides an optimal and consistent user experience regardless of whether the user is on a phone, tablet, or other device.

2.1.3 Hardware interfaces

The system will interact with several hardware components (cameras, microphones, biometric sensors, etc.) to collect data necessary for emotional analysis. The interfaces between the system (application) and these hardware components will be defined as follows:

1. **Camera Interface (for Facial Expression Analysis)**

○ **Functionality:** Captures users' facial expressions for emotion detection.

○ **Connection:** Uses the Android **Camera2 API** to access the camera hardware and process image data. The app will interface with image-processing libraries like OpenCV to analyze facial expressions.

○ **Interaction:** The system will request camera permissions from the user and activate the camera only with the user's consent.

2. **Microphone Interface (for Voice Emotion Detection)**

○ **Functionality:** Captures voice input for analyzing emotional tone and speech commands.

○ **Connection:** Utilizes Android's **AudioRecord** class or **MediaRecorder API** to access microphone data. The app will process the recorded audio using speech-to-text or emotion detection algorithms.

○ **Interaction:** The microphone will be activated only upon user permission, respecting privacy.

1. **Biometric Sensor Interface (for Wearable Device Data)**

○ **Functionality:** Collects biometric data such as heart rate and stress levels from wearable devices (e.g., smartwatches).

○ **Connection:** Communicates with wearable devices via Bluetooth or wireless protocols using **Google Fit API** or third-party SDKs like **Samsung Health SDK** to retrieve data.

○ **Interaction:** Data will be fetched from the wearable device, with user consent for access to health-related data.

2.1.4 Software interfaces

Operating System (OS): The application is built using the Flutter framework and relies on Android and iOS as the primary operating systems. The minimum required versions are Android 10 (API level 29) for Android devices and iOS 13.0 for iOS devices. These versions ensure that the application can access modern system APIs needed for biometric data, voice recognition, image processing, and other critical features. These OS versions also support improved privacy controls, security enhancements, and compatibility with Flutter plugins necessary for device functionalities.

User Interface: This is the interface where users interact with the chatbot. This interface receives various inputs from the user, such as voice, facial expressions, biometric data, and text input. It will be developed to have a user-friendly design, so that users can easily communicate with the chatbot.

Voice Recognition Interface: This is an interface that analyzes the user's voice commands and emotional tone. This interface ensures that voice commands are correctly perceived and that user emotions are extracted from the tone of voice. Real-time processing of the user's voice is a critical component for instant responses.

Image Processing Interface: This is a system interface that analyzes the user's facial expressions. This interface processes the necessary image data to determine the user's emotional state using face recognition and emotion analysis technologies. Real-time face recognition and emotion detection play an important role in the chatbot providing personalized feedback to the user.

Biometric Data Interface(Google Fit): An interface that analyzes biometric data received from wearable devices (e.g. smartwatches). This interface collects and interprets information such as heart rate, stress levels, and other health data. Integration of biometric data allows for more accurate feedback by providing a deeper understanding of the user's emotional state.

API Interfaces: An interface that analyzes data from users and provides feedback and recommendations. This interface is used to provide sentiment analysis, biometric data collection, and integration with other external services. API interfaces enhance the functionality of the chatbot, enriching the user experience. It is required for the client and server communication.

Database Interface: An interface where user data is stored and queried. This interface securely stores users' past sentiments, interactions, and other important information, and provides quick access when needed. The database interface is critical for managing and analyzing user data.

2.1.5 Communication interfaces

The system will use secure and efficient methods for communication between its components. Firebase will handle real-time data updates and ensure everything is synced quickly across devices. Google Authentication API will securely handle user sign-ins and Google Fit API will help capture biometric data like heart rate. The system will also use RESTful APIs to make connecting the app to the backend simple and scalable.

The device should have access to a stable internet connection, Wi-Fi, or mobile data, to use all the functionalities.

2.1.6 Memory constraints

- The SenseAI app should require a minimum of 2 GB of RAM for optimal performance.
- Facial recognition and speech-to-text functionality require real-time processing, which can consume significant memory resources. While modern mobile devices have 4 GB or more of RAM, setting 2 GB as the minimum ensures compatibility with a broad range of devices.
- Users might run the app alongside other applications. Allocating enough memory ensures smooth operation without crashes or significant lag.
- The app should require approximately 100 MB of available storage space for installation and operation.
- Previous queries will be kept in the database and will not be deleted unless the user is deleted. It will be used to analyze the query before and after.

2.1.7 Operations

Regular Users:

1. **Registration and Login:**
- Users can register and log in to the system.
- The system integrates with an external authentication API to secure user accounts.
2. **Data Input and Interaction:**

- Users can choose the type of query to send: text, image, audio, or biometric data.
- Users can initiate new queries at any time.
- Before starting the emotional analysis process, users select the query type.

3. **Profile and Data Management:**
- Users can update their registration information.
- Users can update their mood in their profile section.
- Users can access, search, rename, and delete previous queries or chats.

4. **Error Handling and Feedback:**
- Users will be notified about issues such as exceeding the character limit for text inputs or uploading unsupported file types.

5. **Access to System-Generated Insights:**
- The AI system analyzes input (text, image, audio, biometric data) and provides emotional feedback to users.
- Users can view historical emotional insights and feedback from the system.

6. **Query History:**
- Users can view their last query's access time and title.

7. **Data Privacy:**
- The application ensures compliance with privacy laws, securely storing and processing user data.

AI System Responsibilities

1. Authenticate users via the external authentication API.
2. Analyze user inputs (text, image, audio, biometric data) to generate emotional insights.
3. Provide real-time emotional feedback based on user inputs.
4. Generate historical insights for users to access.
5. Notify users about errors, such as invalid data formats or network issues.
6. Support multilingual functionality in future versions.

Administration:

1. **User Management:**
- Admins can manage user accounts, including registration approval, deactivation, or deletion.

2. **Access to All Data:**
- Admins can access all users' historical data, chats, and other sensitive information (compliant with privacy laws).

3. **System Monitoring and Management:**
- Admins can monitor server health, query load, and user activity.
- Admins are responsible for system-level troubleshooting and resolving issues.
- Admins can back up and restore system data.
- Admins can add or remove users, reset user passwords, and modify user credentials.
- Admins can add new admin users or change admin credentials.

5. **Data Privacy and Compliance:**
- Admins ensure secure data management in compliance with regulations.

2.1.8 Site adaptation requirements

Platform Support: The application will be developed for both iOS and Android platforms, with Android as the primary target.

User Customization Requirements

1. Theme Selection: Users can toggle between Light and Dark Mode.
2. Profile Personalization: Users can update their name, picture
3. Personalized Feedback: Tailored emotional feedback based on input type.
4. Language Preference: Users can choose their preferred language.
5. Input Mode Preference: Users select their preferred input (text, voice, image, biometric).
6. Saved Query Names: Users can rename saved queries for easier tracking.

2.2 Product functions

- The system will analyze the users' emotional state and provide real-time feedback.
- The system should comply with ethical standards while providing advice on mental health and should be sensitive to the needs of the users.[11][12]
- Users can log in to the system by registering.
- Users can send text, image, audio and biometric query.
- The user should be able to choose what type of data to send from the homepage.
- Users can start new query whenever they want and access and delete old query from the user page.
- With the user-friendly interface, the user's interface will be easy to learn and remember with the use of the application.
- Wearable devices or existing health applications will be used for biometric data integration.
- The application will perform emotional analysis through multi-modal inputs such as voice, facial expressions, and biometric data collected from wearable devices like smartwatches.
- The system will analyze user responses using quantifiable metrics, such as numerical or statistical analysis, to derive insights about the mental state of the user. These insights will not be shared directly with the user but will inform the system's logic to provide personalized and contextually relevant responses back to the user.

2.3 User characteristics

The intended users of SenseAI are primarily individuals seeking therapeutic assistance through a digital platform. Users are expected to have a general understanding of using mobile applications. No specific educational background is required; however, the app will be designed to cater to users with diverse educational levels, from high school students to professionals.

Given the nature of SenseAI as a mental support app using AI, users may have varying

levels of comfort with AI interactions. Some users may feel hesitant or cautious about AI's ability to understand emotional context. Consequently, the app will prioritize transparent communication about data handling, privacy, and the limitations of AI responses to build user trust and enhance user comfort.

2.4 Constraints

Hardware Constraints (HW)

● 5 users supported simultaneously in the first stage: The system will support up to 5 concurrent users in the initial phase, and this number can be scaled as the hardware capacity improves.

● App size under 150 MB: The total size of the application should not exceed 150 MB to ensure it can be efficiently installed and run on various devices.

Software Constraints (SW)

● Text input character limit: The application will enforce a character limit on text inputs, with the exact limit determined during development. A warning will be shown for users who exceed the character limit.

● Video size and image resolution limits: The size and resolution of videos and images that users send will have specific limits, to be defined by the development team. Unsupported formats or sizes will trigger a warning message.

● 95% response time under 1 minute: The system is expected to process and return 95% of the responses to users within 1 minute.

● Previous query titles cannot exceed 100 characters: Users will not be able to set query titles exceeding 100 characters when renaming them.

Operating System Constraints (OS)

● Data privacy and compliance: User data must be stored and processed in compliance with data privacy laws, including the regulations specified for the target OS (e.g., Android and iOS).

User Constraints

● Data storage duration for previous chats: Users' previous AI chats will be stored in the database for a period to be determined by the development team, and users will be informed about this period.

● Unsupported features warning for users: Users who try to upload unsupported data (e.g., large video files or exceeding text input limits) will receive an appropriate warning message.

2.5 Assumptions and dependencies

- **User permissions**: Permissions such as camera and microphone access are assumed to be allowed by the user.
- **User Device Specifications:** The characteristics of the devices that users use to access the software can affect the performance of the application. For example, optimization requirements may arise for devices with lower hardware specifications. The mobile devices are assumed to have working camera microphone and touch screen. Also users are assumed to be familiar with basic smart device interactions.
- **Network Conditions:** Users' internet connection speeds and network conditions can affect the performance of the software. Conditions such as longer response times or data loss for low-bandwidth connections may require reconsideration of the requirements.
- **Audio and Visual Quality:** The audio and video quality delivered to users can significantly impact the user experience. Inadequate audio or video quality can make communication difficult to understand and can reduce the effectiveness of the application. Therefore, the requirements should be updated to ensure high-quality audio and video streaming.

2.6 Apportioning of requirements

- **Additional Emotion Recognition Modes:** Additional emotion recognition modes (e.g. stress management or relaxation techniques) added to better understand the user's mood.
- **More Integration Options:** A wider range of integrations with different wearable devices. For example, connection with more brands and models of smartwatches or fitness trackers.
- **Enhanced UI Features:** Added more customization options in the UI design to enhance the user experience.

3 Specific Requirements

3.1 External Interface Requirements

3.1.1 User interfaces

This section will describe the logical characteristics of each interface in the system.
The section will include wireframes to better describe the interface.

3.1.2.1 Sign Up

On the Sign Up page, users will register to the system with an e-mail address and a password they will specify. E-mail and password are mandatory fields on this page. With ”Google” sign users can login with their google account. Users must click the "Sign Up" button to complete their registration to the system. (Figure 2)

3.1.2.2 Login/Sign In

The Login page is for users to access the system through their existing accounts. The required fields for this page are e-mail and password. With ”Google” sign users can login with their google account. After the fields are provided, users can click the “Log In” button to access the system. With “Remember Me” button users can access their account without sign in process. (Figure 1)

3.1.2.3 Chat History

In this interface, users can access, delete and update their previous queries. Each query is kept with its date and title. Users can search for old queries with the search icon. They can add new queries with the "+" button. (Figure 5)

3.1.2.4 Profile

Users can update their mood in the profile section, see the last query access time and the title of the last query they talked about. They can start a new query with the "Start conversation" button. They can also access the settings page, chat section and profile page via the corresponding buttons. (Figure 6)

3.1.2.5 Settings

On this page, users can go to account settings, share the application URL, manage notifications, access privacy policy, terms C conditions and log out. They can go back to the previous page with the button on the top left. Users can access the chat section and profile page with the buttons below. (Figure 9)

3.1.1.6 Admin

When the admin clicks on the "Review User Record" button, they can search for a user by ID or Username with the search button in the window that opens. Users are listed according to the search. The admin selects the user to see the record. They can see ID, username, e- mail, last access date, query contents and log in/log out activities. They can see the queries of the users with the "query contents" button, the "log in/log activity history screen opens. They select a user for controlling from the list that appears. Admin can make the user an admin,

block them or manage access. Admin can go to the settings page with the button on the top right. (Figure 7)

3.1.1.7 Home Page

Users can go to the previous query page with the "Chat history" button, write their own thoughts in the "My Inbox" section, access the settings page with the icon at the top left, and access the chat section and profile page with the buttons at the bottom. (Figure 3)

3.1.1.8 Chat Page

On this page, users send queries to the system and receive answers. Users can select the query type they will send with the paperclip icon at the top, and switch to dark mode with the icon next to it. They can go to the previous page. (Figure 4)

3.1.1.8 Forgot Password

If the user clicks on the "forgot password" button, the system directs them to a page where they will write the e-mail that will reset the password. If they enter a valid e-mail, the user should be directed to the page where they will determine the new password. They should click on the "Submit Password" button to save the new password. (Figure 8)

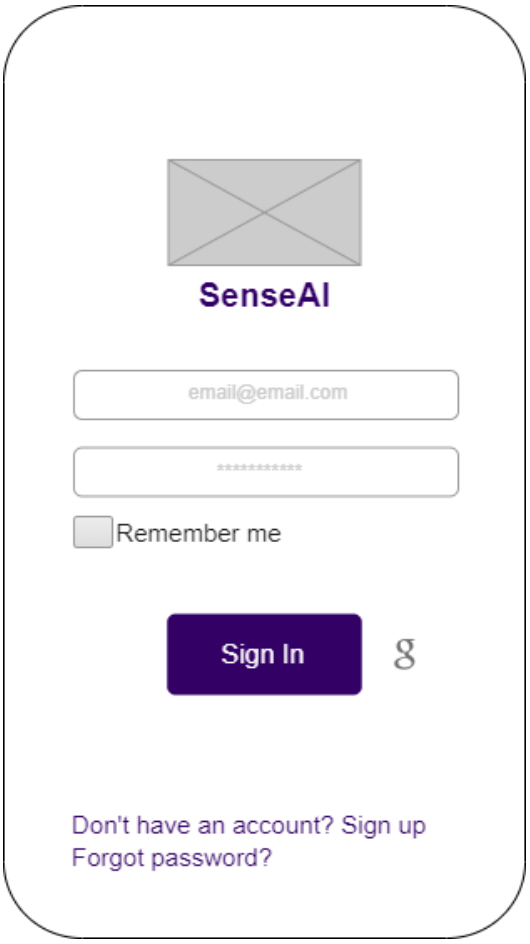


Figure 1 Login page

Figure 2 Sign up page

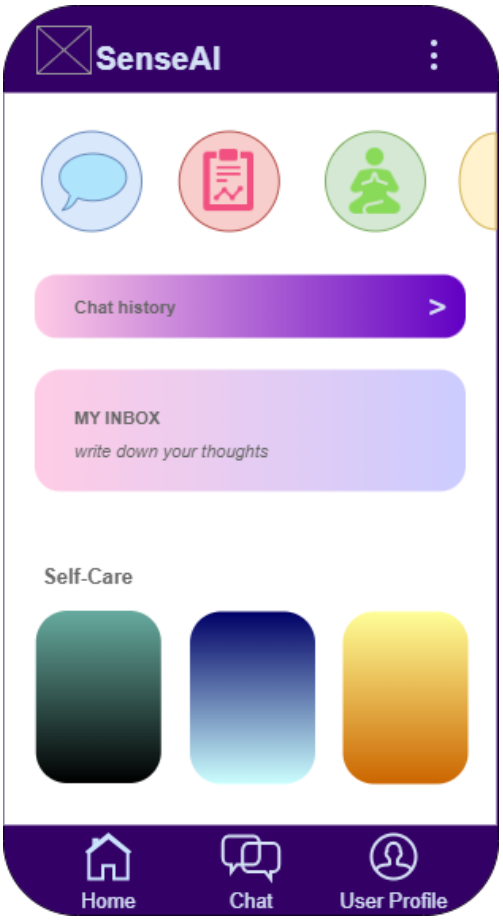


Figure 3 Main page

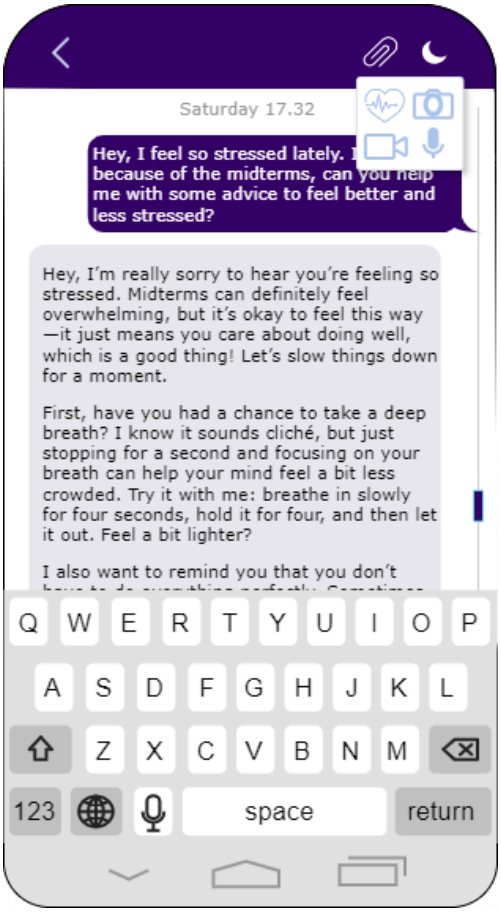


Figure 4 AI chat page

Figure 5 Chat history page

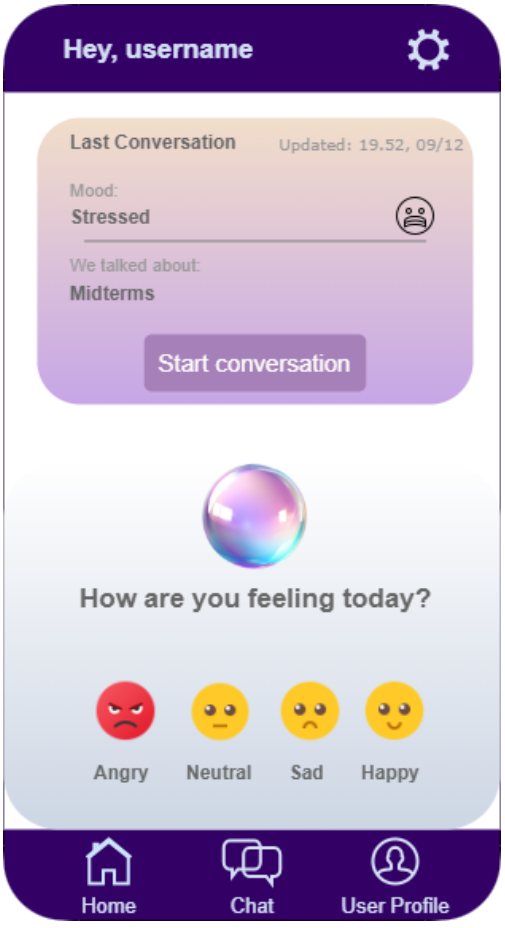
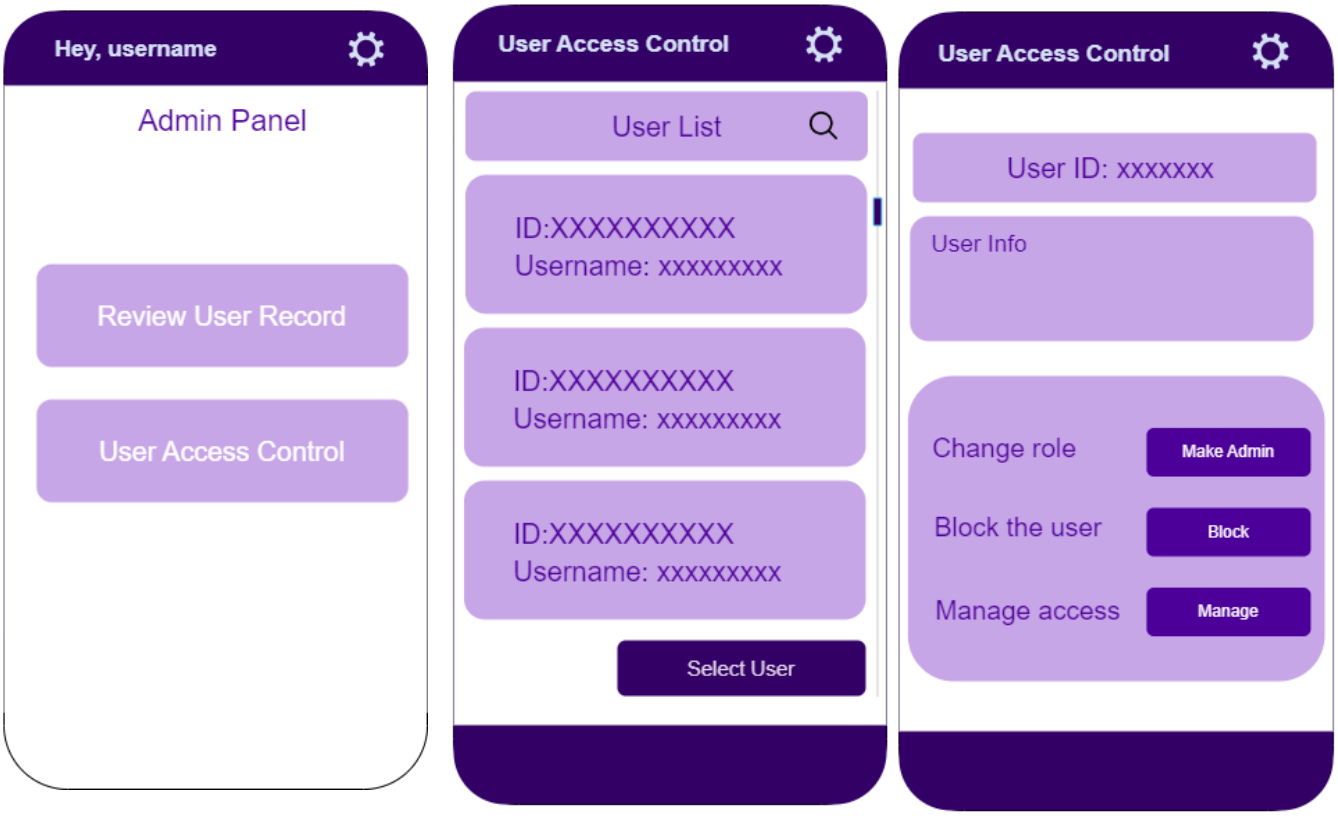


Figure c User profile page



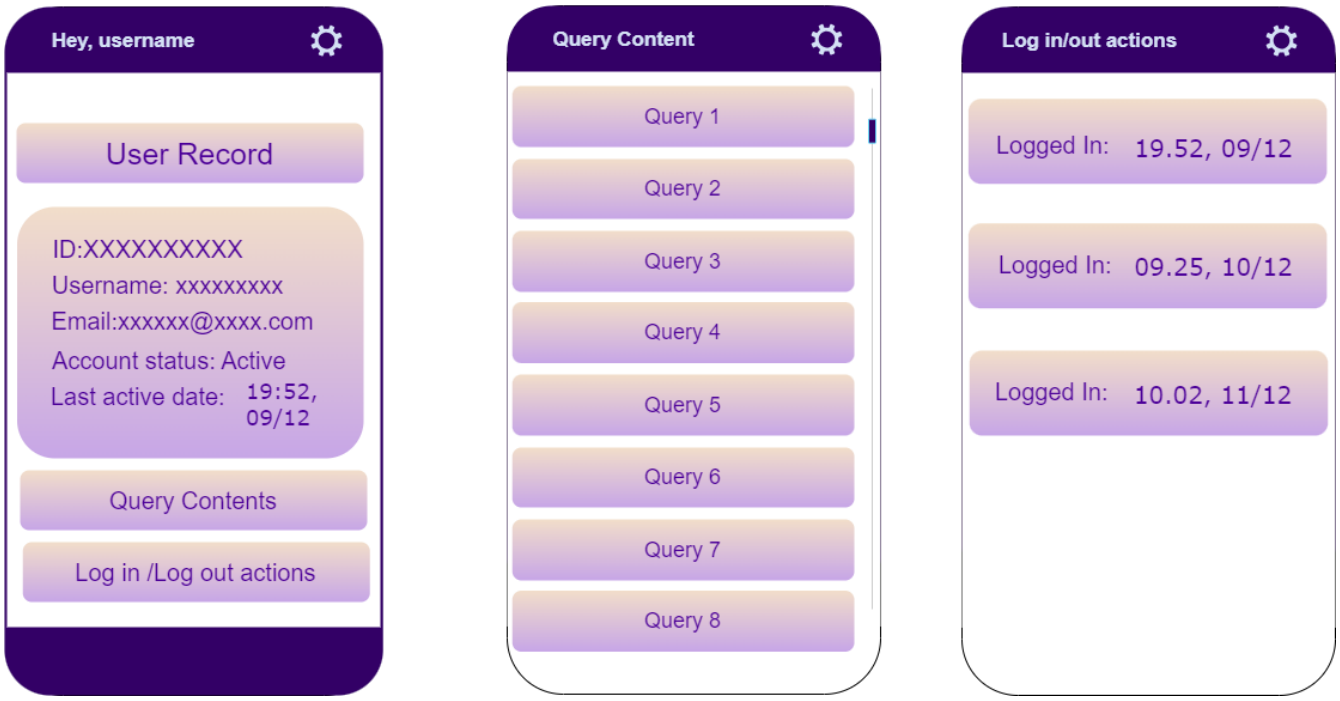


Figure 7 Admin pages

Figure 5 Chat history page

Figure c User profile page

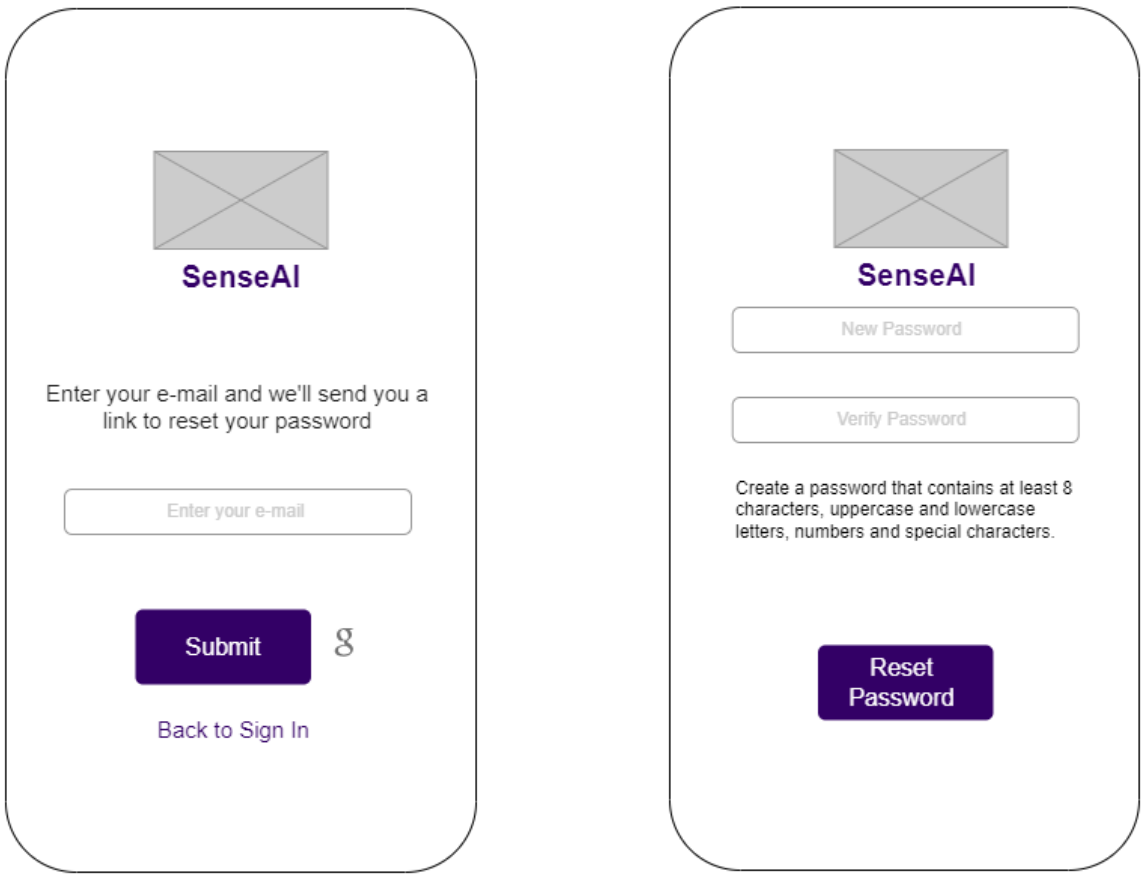


Figure 8 Forgot password pages

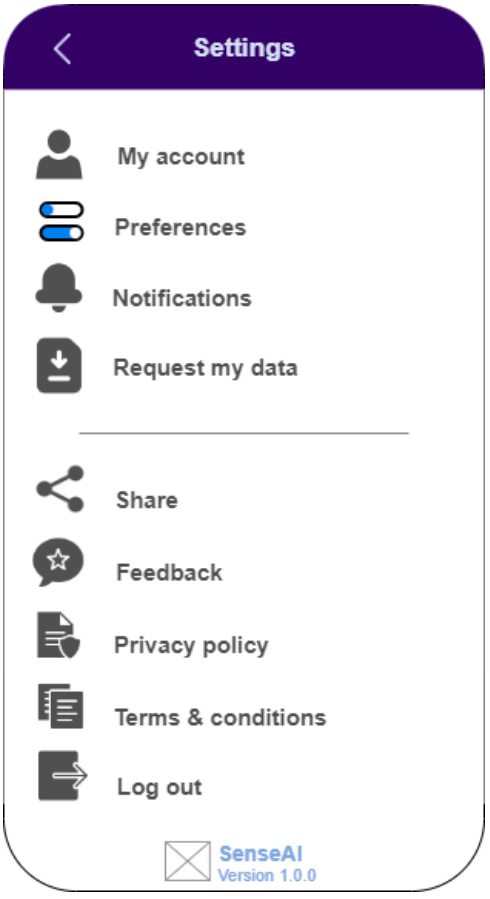


Figure 8 Settings page

3.1.2 Hardware interfaces

3.1.2.6 Input Devices

- Touchscreen: SenseAI should support touch input on the mobile device's touchscreen for user interactions such as tapping and swiping.
- Microphone: The devices that users use should support microphone input for inputs sent via voice.
- Camera: Camera should be supported for inputs received from users' facial expressions.
- Sensor: Smartwatch input should be supported for inputs received from users' watches.

3.1.2.7 Output Devices

Display: SenseAI should be able to render the user interface on the mobile device's screen, utilizing the available screen resolution and aspect ratio.

3.1.2.8 Network Connectivity

SenseAI will require access to the device's internet connection to communicate with external services.

3.1.3 Software interfaces

Software Interfaces will have the features mentioned in the 2.1.4 Software Interfaces section.

3.1 4 Communications interfaces

Communication Interface will have the features mentioned in the 2.1.5 Communication Interfaces section.

3.2 Functional Requirements

Users *User Functional Requirements (User-FR)*

- User-FR-01: The system shall allow users to register and log in to the application.
- User-FR-02: Users shall be able to update and delete the information entered during registration.
- User-FR-03: The system shall enable users to send data via text, images, audio, and biometric data.
- User-FR-04: Users shall be able to choose the type of query (text, image, audio, or biometric data) from the homepage.
- User-FR-05: Users shall be able to start new AI chats and access or delete old chats from their history page.
- User-FR-06: Users shall be able to search past queries by title and change the query name from history.
- User-FR-07: Users shall be able to choose between dark mode and light mode.
- User-FR-08: Users shall be able to subscribe to premium features or make in-app purchases for advanced functionalities.

- User-FR-09: Users shall be guided by the system to ensure images are taken in sufficient lighting and at appropriate angles for accurate analysis.
- User-FR-10: The system shall warn users when videos/photos exceed the size limit or when texts exceed the character limit.

Administrators *Admin Functional Requirements (Admin-FR)*

- Admin-FR-01: Administrators shall be able to monitor the system's compliance with ethical standards.
- Admin-FR-02: Administrators shall manage premium feature access and subscription statuses for users.
- Admin-FR-03: Administrators shall be able to track data quality issues and provide resolution guidelines to users.
- Admin-FR-04: Administrators shall oversee system updates to support additional languages and ensure accurate localization.

System ***System Functional Requirements (System-FR)***

- System-FR-01: The system shall analyze the user’s emotional state and provide feedback.
- System-FR-02: The system shall comply with ethical standards when providing mental health advice, ensuring sensitivity to users' needs.
- System-FR-03: Wearable devices or existing health applications shall be integrated for biometric data.
- System-FR-04: The system shall provide a user-friendly interface that is easy to learn, remember, and visually appealing.
- System-FR-05: The system shall clearly convey problems to users when errors occur and allow recovery.
- System-FR-06: The system shall support user authentication via an external authentication API to secure access to personal data and interactions.
- System-FR-07: The system shall initially support the English language, with future multilingual expansion potential.

Non-Functional Requirements (NFR)

- NFR-01: The system shall ensure high performance, with minimal response time for user queries.
- NFR-02: The system shall adhere to security standards to protect user data and privacy.
- NFR-03: The interface shall maintain visual consistency and responsiveness across multiple devices.
- NFR-04: The system shall be scalable to handle increased user traffic and data volume in the future.
- NFR-05: The system shall log errors and system performance metrics for analysis and troubleshooting.
- NFR-06: The system shall maintain a 99.9% uptime reliability rate.
- NFR-07: The application shall comply with applicable data protection regulations (e.g., GDPR, HIPAA).

Figure 10 Use Case Diagram

3.2.1 SenseAI users

3.2.2.1 Logging in

3.2.2.1.1 Introduction/Purpose of feature

The Logging In feature allows authorized users to securely access the SenseAI application and their personalized data. This feature serves as an entry point to ensure data privacy, personalized experience, and user accountability. The system should let users stay logged in once login is successfully completed.

3.2.2.1.2 Stimulus/Response sequence

1. System: Check if user logged in before
2. System: Direct to landing page if user logged in
3. System: Else, show login screen with text fields for e-mail and password
4. User: Fills in account info
5. User: Tap login
6. System: Check account credentials and direct user to landing page if correct
7. System: Else show error

3.2.2.1.3 Associated functional requirements

REQ-1:Users should be able to login with their google accounts.
REQ-2: Users should be able to sign up by inputting their email and a password.
REQ-3: Users should be able to log out whenever they want
REQ-4: Logged in users should stay logged in unless they logged out (closing the app should not log out the users). The login screen should be skipped for already logged in users.

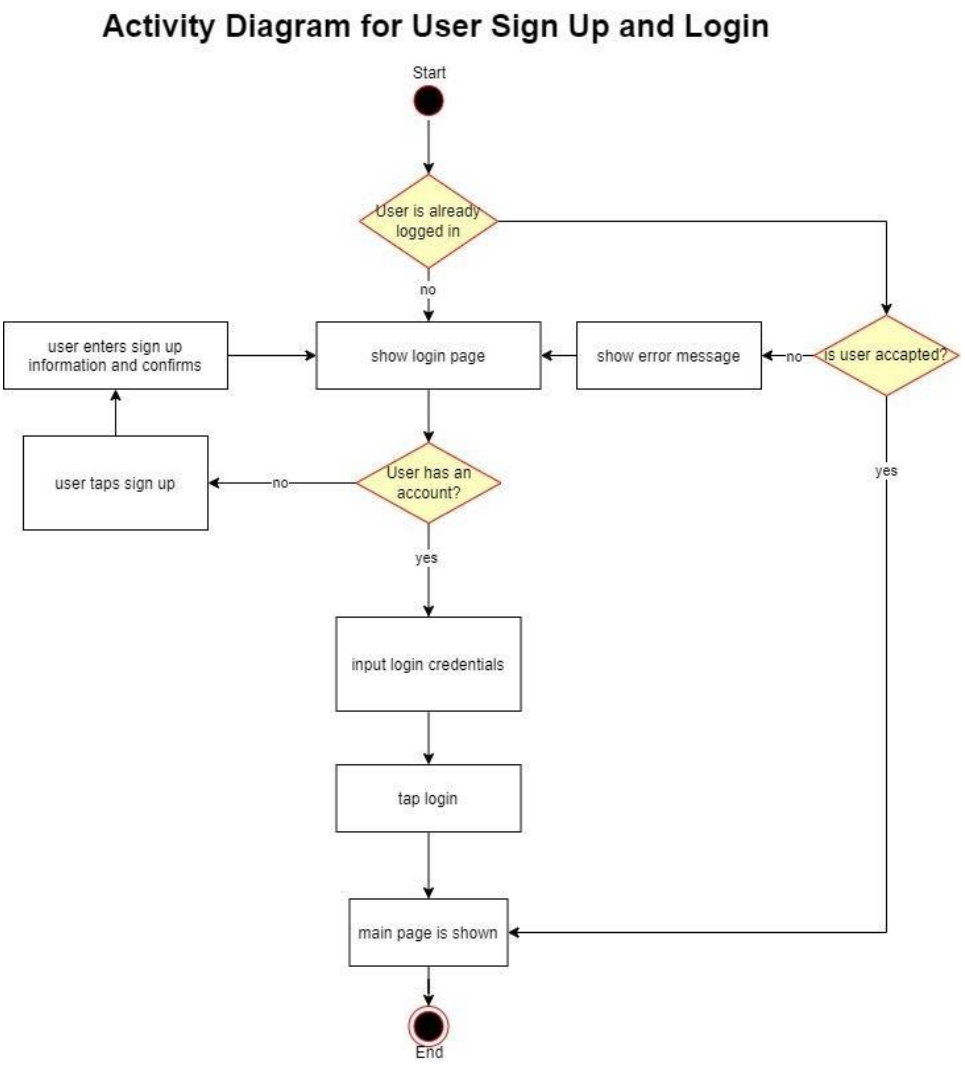


Figure 11 Login/Sign Up Activity Diagram

3.2.2.2 Forgot password

3.2.2.2.1 Introduction/Purpose of feature

The system shall allow users to securely update their passwords to maintain account security and protect their personal information. This feature ensures that users can regularly update their credentials or recover access in case of a potential breach.

3.2.2.2.2 Stimulus/Response sequence

1. System: Display a "Change Password" option in the account settings.
2. User: Taps on the "Change Password" option.
3. System: Prompt the user to input their current password and the new password twice for confirmation.
4. User: Enters the current password and the new password.
5. System: Validates the current password and checks if the new password meets the required security criteria (e.g., length, character mix).
6. System: If validation passes, update the password and confirm the change to the user.
7. System: If validation fails, display an error message and prompt the user to try again.

3.2.2.2.3 Associated functional requirements

REQ-1: The system must verify the user's current password before allowing the password to be updated.

REQ-2: The new password must meet defined security standards, including minimum length, inclusion of special characters, and exclusion of easily guessable patterns.
REQ-3: The system should send a confirmation email or notification upon successful password change.
REQ-4: The feature must block repeated use of the last X passwords for additional security.
REQ-5: If the password change attempt fails multiple times, the system must temporarily lock the account to prevent unauthorized access.

3.2.2.3 Sign up

3.2.2.3.1 Introduction/Purpose of feature

The system shall provide new users with a secure and user-friendly way to create an account. This feature ensures the onboarding process is seamless while verifying the user's identity and maintaining system integrity.

3.2.2.3.2 Stimulus/Response sequence

1. System: Display a "Sign Up" form with fields for required information (e.g., email, password, name).
2. User: Fills out the required information and submits the form.
3. System: Validate the input data for accuracy and completeness (e.g., valid email format, strong password).
4. System: Check if the email is already associated with an existing account.
5. System: If validation passes, create a new user account and send a confirmation email or message to the provided address.
6. System: If validation fails, display an error message and prompt the user to correct the input.

3.2.2.3.3 Associated functional requirements

REQ-1: The system must ensure that email addresses are unique and not already in use.
REQ-2: Passwords must meet security criteria, including minimum length, inclusion of special characters, and resistance to dictionary attacks.
REQ-3: The system should require users to verify their email address by sending a confirmation link or code.
REQ-4: The system must encrypt sensitive user data, such as passwords, during storage and transmission.
REQ-5: The system should provide error messages that guide users in correcting any issues with their input (e.g., “Email already in use” or “Password is too weak”). REQ-6: The feature must comply with data privacy regulations by securely handling user-provided information.
REQ-7: Optional fields for user profile information (e.g., phone number, date of birth) should be clearly marked and not mandatory for account creation.

3.2.2.4 Sending text for analysis

3.2.2.4.1 Introduction/Purpose of feature

The system shall allow users to input text for emotion and sentiment analysis, enabling SenseAI to interpret emotional cues within user responses. By processing the input, SenseAI's analysis module will detect underlying emotions (e.g., happiness, sadness, anger) and overall sentiment (positive, negative, neutral).

3.2.2.4.2 Stimulus/Response sequence

1. System: Show textbox for input
2. User: Enter input.
3. User: Taps "Send" button.
4. System: Validates text input and sends it to the model for analysis with other input types (if exists).
5. System: Displays loading indicator.
6. Model: Processes the input.
7. System: Receives analysis results, saves it to the database and the device.
8. System: Displays results to the user.

3.2.2.4.3 Associated functional requirements

REQ-1: Chatting with ai should look and feel like the users are chatting with a friends on an app. e.g. both input and output text should be visible as text bubbles and the screen should be vertically scrollable.

REQ-2: The device's keyboard should not obstruct the previous text while writing new inputs.

3.2.2.5 Using voice as input

3.2.2.5.1 Introduction/Purpose of feature

The system shall allow users to input audio by processing the audio through text-to-speech conversion. It shall also take audio input and analyze it to return a response to the user.

3.2.2.5.2 Stimulus/Response sequence

1. System: Show microphone icon for input
2. User: Taps the icon
3. User:Speaks to the microphone.
4. System: Detects speech and uses TTS (text to speech) and converts it to text input.
5. System: Validates input and sends it to the model for analysis with other input types (if exists).
6. System: Displays loading indicator.
7. Model: Processes the input.
8. System: Receives analysis results, saves it to the database and the device.
9. System: Displays results to the user.

3.2.2.5.3 Associated functional requirements

REQ-1: Transcript of the user's input should be added to the input textbox, ready to be sent.

3.2.1.6 Sending image for analysis

3.2.1.6.1 Introduction/Purpose of feature

The system shall allow users to input photos using their smart devices for analysis, enabling SenseAI to process the images and extract actionable insights, such as object recognition, emotional tone assessment, or visual patterns, to enhance the overall analysis process.

3.2.1.6.2 Stimulus/Response sequence

1. System: Show image icon for input
2. User: Taps the icon.
3. System: Enables the front camera for taking a photo
4. User: Takes a photo.
5. System: Validates input and sends it to the model for analysis with other input types (if exists).
6. System: Displays loading indicator.
7. Model: Processes the input.
8. System: Receives analysis results, saves it to the database and the device.
9. System: Displays results to the user.

3.2.1.6.3 Associated functional requirements

REQ-1: Users should be able to take photos without exiting the app.

3.2.1.7 Sending heartbeat info for analysis

3.2.1.7.1 Introduction/Purpose of feature

The system shall enable SenseAI to receive and analyze users' heartbeat information to assess emotional or physical states. This analysis will take into account the user's age, as age-related differences can impact heart rate norms and interpretations. By factoring in age, SenseAI can more accurately identify deviations from typical heartbeat patterns, helping to gauge stress, anxiety, or relaxation levels.

3.2.1.7.2 Stimulus/Response sequence

1. System: Display option for biometric data input.
2. User: Either check or uncheck the option
3. System**:** Check if the user has checked or unchecked the option for biometric data input. If checked, proceed; if unchecked, skip the biometric data retrieval step.
4. System**:** Check if Google Fit API is successfully connected to the user's account (is the user authenticated on Google Fit?).
5. System**:** Check if the required biometric data (BPM) is available in Google Fit. If data is unavailable or incomplete, prompt the user to provide manual input or skip this step.
6. System: Access today's google fit information (BPM) of the user and send it to the model along with other input type (if exists)
7. System: Displays loading indicator.
8. Model: Processes the input.
9. System: Receives analysis results, saves it to the database and the device.
10. System: Displays results to the user.

3.2.1.7.3 Associated functional requirements

REQ-1: The system should be able to access the latest heartbeat info from the user's google fit account.

REQ-2: The users should be only able to use this input option if they have a google fit account linked. If no account is linked the option should be grayed out. And tapping it should show a text bubble telling users they need to link their google fit accounts.

REQ-3: If an error occurs while retrieving info or if no data exists, the app should inform the user about the error.

3.2.1.8 Viewing previous analysis (history)

3.2.1.8.1 Introduction/Purpose of feature

The system shall provide users with access to a history feature, allowing them to view previous analyses and insights generated by SenseAI. This history will include past sentiment, emotion analyses from text, heartbeat, and image inputs, enabling users to track trends and changes over time. By offering a historical view, SenseAI aims to support users in understanding their emotional patterns and progress

3.2.1.8.2 Stimulus/Response sequence

1. System**:** Display option to view past analysis results
2. User**:** Selects the option to view previous analysis.
3. System**:** Checks if the user has any stored analysis results in the database or device.
   - If results are found, proceed.
   - If no results are found, display a message such as "No previous analysis available."
4. System**:** Retrieves and displays a list of previous analysis results.
5. User**:** Selects a specific past analysis entry to view detailed results.
6. System**:** Displays the selected analysis details, including insights, graphs, and any other relevant data.
7. User**:** Optionally, chooses to delete or update past analysis results (if allowed by the system).
8. System**:** If the user deletes or updates, confirms action and updates the database accordingly.
9. System**:** displays the updated history of deletion.

3.2.1.8.3 Associated functional requirements

REQ-1: History of previous interactions should be displayed with dates. REQ-2: Users should be able to view the history chronologically ordered. REQ-3: Users should be able to delete previous talks.
REQ-4: The system should provide a search or filter option, allowing users to find specific past interactions based on keywords or date ranges.

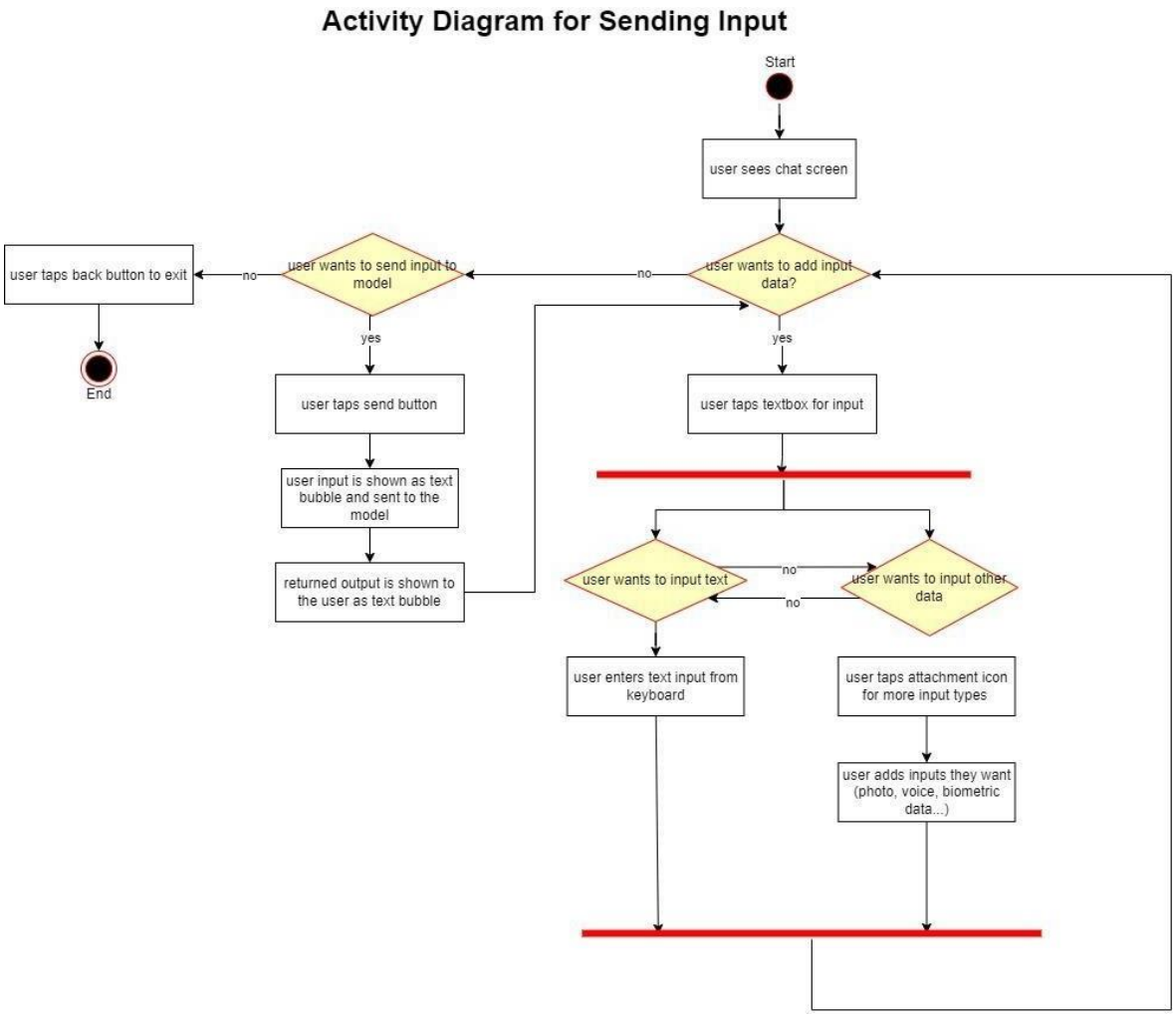


Figure 12 Sending Input Activity Diagram

3.2.2 Admin

3.2.2.6 Logging in

The Logging In feature allows admins to securely access the SenseAI application and their personalized data. This feature serves as an entry point to ensure data privacy, personalized experience, and user accountability. The system should let users stay logged in once login is successfully completed.

3.2.2.6.1 Stimulus/Response sequence

1. System: Check if user logged in before
2. System: Direct to landing page if user logged in
3. System: Else, show login screen with text fields for email and password
4. User: Fills in account info
5. User: Tap login
6. System: Check account credentials and direct user to landing page if the user is admin
7. System: Else show error

3.2.2.6.2 Associated functional requirements

REQ-1: The system must securely authenticate the admin's credentials using industry-standard encryption methods.

REQ-2: The system should store a secure session token to keep the admin logged in until the session expires or the admin logs out manually.

REQ-3: If the user has previously logged in and their session is active, the system should skip the login screen and redirect them directly to the landing page.

REQ-4: Invalid login attempts should trigger appropriate error messages (e.g., "Invalid email or password") without exposing sensitive information.

REQ-5: The system must enforce a retry limit to prevent brute-force login attempts and temporarily lock the account after exceeding the limit.

REQ-6: The login feature must validate that the user has an admin role before granting access to the admin dashboard or personalized data.

REQ-7: A "Forgot Password" option should be available to help admins recover access in case of credential loss.

3.2.2.7 Forgot password

3.2.2.7.1 Introduction/Purpose of feature

The system shall allow users to securely update their passwords to maintain account security and protect their personal information. This feature ensures that users can regularly update their credentials or recover access in case of a potential breach.

3.2.2.7.2 Stimulus/Response sequence

1. System: Display a "Change Password" option in the account settings.
2. User: Taps on the "Change Password" option.
3. System: Prompt the user to input their current password and the new password twice for confirmation.
4. User: Enters the current password and the new password.
5. System: Validates the current password and checks if the new password meets the required security criteria (e.g., length, character mix).
6. System: If validation passes, update the password and confirm the change to the user.
7. System: If validation fails, display an error message and prompt the user to try again.

3.2.2.7.3 Associated functional requirements

REQ-1: The system must verify the user's current password before allowing the password to be updated.

REQ-2: The new password must meet defined security standards, including minimum length, inclusion of special characters, and exclusion of easily guessable patterns.

REQ-3: The system should send a confirmation email or notification upon successful password change.

REQ-4: The feature must block repeated use of the last X passwords for additional security.
REQ-5: If the password change attempt fails multiple times, the system must temporarily lock the account to prevent unauthorized access.

3.2.2.8 Sign up

3.2.2.8.1 Introduction/Purpose of feature

The system shall provide new users with a secure and user-friendly way to create an account. This feature ensures the onboarding process is seamless while verifying the user's identity and maintaining system integrity.

3.2.2.8.2 Stimulus/Response sequence

1. System: Display a "Sign Up" form with fields for required information (e.g., email, password, name).
2. User: Fills out the required information and submits the form.
3. System: Validate the input data for accuracy and completeness (e.g., valid email format, strong password).
4. System: Check if the email is already associated with an existing account.
5. System: If validation passes, create a new user account and send a confirmation email or message to the provided address.
6. System: If validation fails, display an error message and prompt the user to correct the input.

3.2.2.8.3 Associated functional requirements

REQ-1: The system must ensure that email addresses are unique and not already in use.
REQ-2: Passwords must meet security criteria, including minimum length, inclusion of special characters, and resistance to dictionary attacks.
REQ-3: The system should require users to verify their email address by sending a confirmation link or code.
REQ-4: The system must encrypt sensitive user data, such as passwords, during storage and transmission.
REQ-5: The system should provide error messages that guide users in correcting any issues with their input (e.g., “Email already in use” or “Password is too weak”). REQ-6: The feature must comply with data privacy regulations by securely handling user-provided information.
REQ-7: Optional fields for user profile information (e.g., phone number, date of birth) should be clearly marked and not mandatory for account creation.

3.2.2.9 Review user record

3.2.2.9.1 Introduction/Purpose of feature

This feature enables admins to access and manage detailed records of individual users, including their activities, queries, and biometric or manual inputs. It supports efficient monitoring and accountability while ensuring user data privacy and system transparency.

3.2.2.9.2 Stimulus/Response sequence

1. System: Display a list of user records in the admin panel, with options to search and filter users.
2. Admin: Search for or filter the user list to locate a specific user ID,username and select the user

record.

3. System: Retrieve and display the selected user's record with the following sections:
    - Personal Information:
        - Username, email address, user ID, account status (active/inactive).
    - System Interaction Logs:
        - User's login/logout actions, connection status.
        - The last active date and time on the application.
    - Query History:
        - A chronological list of all queries submitted by the user:
            - Query content.
            - Submission date and time.
            - Results or error messages returned for the query.
            - Query status (completed, failed, pending).
    - Biometric Data:
        - Latest biometric data (e.g., BPM) if the user is connected to Google Fit.
        - Connection status: "Connected" or "Not Connected."
    - Linked Services:
        - Status of connections with external services like Google Fit.
4. System: Display a confirmation dialog for each selected action.
5. Admin: Confirm or cancel the action.
6. System: Execute the action, display a record.

3.2.2.9.3 Associated functional requirements

REQ-1: The system must provide an interface for searching, filtering, and viewing user records.

REQ-2: The user's query history must be displayed in detail, including query content, results, timestamps, and statuses.

REQ-3: Biometric data should only be displayed if the user is connected to an external service like Google Fit.

REQ-4: All admin actions must be logged with action type, timestamp, and admin identity for auditing purposes.

REQ-5: User details can only be edited or deleted by authorized admins.

REQ-6: The query history must support search and filter functionality for efficient review.

REQ-7: The system must maintain separate logs for user and admin activities.

3.2.2.10 User access control

3.2.2.10.1 Introduction/Purpose of feature

This feature allows admins to manage and regulate user access within the system. It enables the assignment of roles, permission settings, and account status management, ensuring that users can only access the resources and actions appropriate to their role. This is critical for maintaining security, privacy, and operational efficiency within the application.

3.2.2.10.2 Stimulus/Response sequence

1. System: Search users with options to search, filter, and sort by roles, permissions, or activity status (active/inactive).

Admin: Select a specific user from the list to review or modify their access permissions.
System: Retrieve and display the selected user's access details.

2. Admin: Closes an account, changes access permissions, or changes the role of another account.
3. System: Display a confirmation dialog summarizing the changes.
4. Admin: Confirm or cancel the action.

System: Execute the changes, update the user's record in the database, and display a success or error message.

3.2.2.10.3 Associated functional requirements

REQ-1 The system must provide a unified dashboard for searching, filtering, and managing user access.

REQ-2 Changes to roles, permissions, or account statuses must show their impact clearly before confirmation.

REQ-3 The system must maintain an audit log of all actions performed, including timestamps, changes, and admin details.

REQ-4 Account status changes should have safeguards (e.g., warnings) to prevent unintentional lockouts.

REQ-5 Permissions and roles should be manageable independently to ensure granular control over access.

3.3 Performance requirements

- The application shall be available 24/7 to ensure accessibility at any time.
- 95% of system responses to user requests shall be completed within 1 minute to ensure timely feedback.
- The system shall initially support up to 5 simultaneous users, with scalability to support more users as the application matures and a more advanced hardware is provided.
- The system shall handle multimodal data types, including text, images, audio, and biometric data.
- Users who engage with audio and video functionalities shall use compatible devices to ensure smooth data transmission and processing.

3.4 Logical database requirements

The system must ensure 24/7 availability of the database without any interruption or service interruption. The database will be backed up daily to ensure uninterrupted accessibility of the database in the event of hardware or software failure, to ensure data integrity and to

support disaster recovery procedures. Users' past queries will be stored in the database indefinitely until the user deletes them.

3.5 Design constraints

The server's processor power, RAM and storage capacity should not negatively affect the performance of the Chatbot as the number of users increases. The device should support microphone, camera, smart watch, touchscreen features.

3.5.1 Standards compliance

3.5.2.1 Material Design Guidelines

Sense AI's user interface will be designed in accordance with Google's Material Design Principles[13] to provide a consistent, modern, and user-friendly experience. Material Design provides a comprehensive guide to visual design, motion, and interaction principles that enhance usability and add aesthetic appeal. This design approach makes it easy for users to interact naturally with the chatbot, while providing a professional and reliable interface standard.

3.6 Software system attributes

3.6.1 Reliability

A chatbot system that provides psychological support is critical in terms of reliability. The system must provide a sustainable service without losing users' past conversations and analyses. Therefore, regular data backup mechanisms will be implemented. In addition, continuous tests will be conducted to minimize system errors and provide uninterrupted service.

3.6.2 Availability

The system should be available 24/7 and include recovery and restart mechanisms for disaster recovery situations. High availability (HA) infrastructure will be used to minimize user waiting time.

3.6.3 Security

Users' personal data and chat content must be protected against both accidental and malicious access. Security measures will include:

- Use of reliable cryptographic algorithms (e.g. AES-256) to encrypt data,

- Two-factor authentication for user logins,
- Role-based access control to ensure only authorized individuals can access sensitive data,
- Regular verification processes to ensure data integrity.

3.6.4 Maintainability

The chatbot will be developed in a modular structure. Each function will be in a separate module and these modules can be updated or changed independently when necessary. Thanks to the object-oriented design, it will be easy to expand the chatbot by adding new features or improving existing features. The system should allow automatic updates to be installed remotely via the Google Play Store and Apple App Store without requiring user intervention. Updates should not disrupt the user experience and should ensure the integrity and security of the data. The system will apply Over-the-Air (OTA) updates to keep the system up to date via the Google Play Store and Apple App Store.

3.6.5 Portability

The mobile application that offers psychological support will be developed in a way that it can work seamlessly on different mobile devices and operating systems (iOS, Android). In this direction:

- Platform-dependent codes will be minimized,
- Tools that provide cross-platform compatibility will be used in the development process,
- It will be optimized to meet the requirements of mobile application stores,
- Its compatibility with various screen sizes and resolutions will be tested.
- This approach will ensure that the application offers a seamless user experience on different devices.

Software Design Description

Introduction

Purpose

The purpose of this Software Design Document (SDD) is to provide a comprehensive and detailed description of the architecture, design, and implementation approach for the SenseAI project. The SDD serves as a blueprint to ensure that all stakeholders have a clear understanding of the system's structure, its components, and their interactions.
This document is intended for:

- **Development Team:** To implement the features and functionality outlined in the requirements.
- **Project Stakeholders:** To review and validate the proposed design.
- **Quality Assurance Team:** To understand the system design for testing purposes.
- **End-User Representatives:** To provide feedback and confirm the application addresses their expectations.
- **Project Managers**: To ensure the design aligns with project objectives.
- **Future Developers/Maintainers:** To understand the original design and purpose for updates or maintenance.

This document is designed to be accessible and useful for both technical and non-technical stakeholders.

Scope

This SDD outlines the design of the SenseAI system, which aims to facilitate mood tracking, data analysis, and session-based communication.
The document will cover:

- The overall system architecture.
- Detailed descriptions of system components and their interactions.
- Data models and storage mechanisms.
- Interfaces and dependencies.
- Non-functional requirements, such as performance and scalability.

The SDD will not include:

- Detailed implementation code.
- End-user documentation or training materials.

The intended audience includes developers, testers, project managers, and stakeholders who are involved in the project lifecycle.

Definitions, acronyms, and abbreviations

Term	Description
AI Model	Artificial Intelligence Model
IOS	iPhone Operating System

IP	Internet Protocol Address
OS	Operating System
RAM	Random Access Memory
SSD	Solid State Disk

Overview

This document details the design and development process of a software or system. Below are summaries of the topics covered in each section:

Architecture (Section 2): This section outlines the overall architecture of the system. It describes the architectural components, structures on the client and server sides, and how these components interact with each other. The main components of the system and their interconnections are detailed.

User Interface (Section 3): The user interface design is explained in this section. The design of different screens and the functionalities they provide are discussed. Interface components and the interaction methods used to improve user experience are described.

High Level Design (Section 4): This section provides a high-level design of the system. The structures and functions of the main modules are discussed in broad terms. The basic roles of each module, their functions, and the relationships between the modules are summarized.

Low Level Design (Section 5): This section delves into the detailed design of each module. It explains how each module operates, including the algorithms used and the data flows. Technical details of each module are provided.

Database Design (Section 6): This section explains the database design. It provides information on the database structure, relationships, tables, and how the database will be organized. Entity-relationship diagrams, tables, and their relationships are detailed.

References (Section 7): The final section lists all the sources and references cited throughout the document. It includes technical documents, libraries, and other relevant materials used.

Architecture

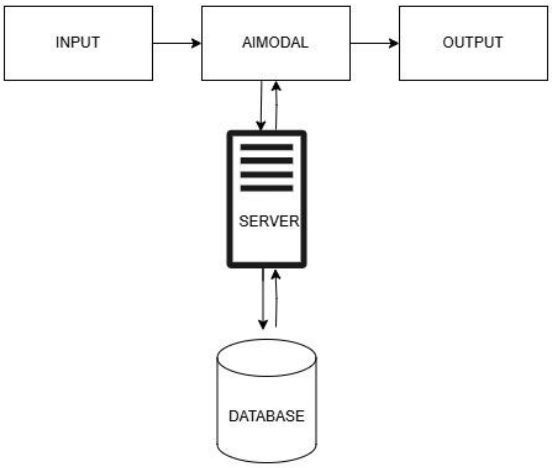


Figure 1 Hardware

Clients

Mobile Phones / Tablets

Minimum Requirements:

- **OS: Android 10.0 or iOS 13.0 and above**
- **RAM: 4 GB minimum (6 GB or more recommended for better performance)**
- **Storage: 200 MB of available space (additional space may be required for user data and media)**
- **Network: Active internet connection (Wi-Fi or 4G/5G recommended for reliable performance)**

Application Server

Minimum Requirements:

- OS: Ubuntu 20.04 LTS or Windows 10/11 (Desktop versions)
- CPU: Dual-core processor (Intel i3/Ryzen 3 or higher)
- RAM: 8 GB minimum (16 GB recommended for smooth operation)
- Storage: 50 GB available space (SSD for better performance)
- Network: Standard internet connection (at least 20-50 Mbps)

Database Server

Minimum Requirements:

- OS: Ubuntu 20.04 LTS or Windows 10/11
- CPU: Dual-core processor (Intel i3/Ryzen 3 or higher)
- RAM: 8 GB minimum (16 GB recommended)
- Storage: 256 GB SSD (for fast read/write operations)
- Network: Standard internet connection (at least 20-50 Mbps).

User Interface

Login Page

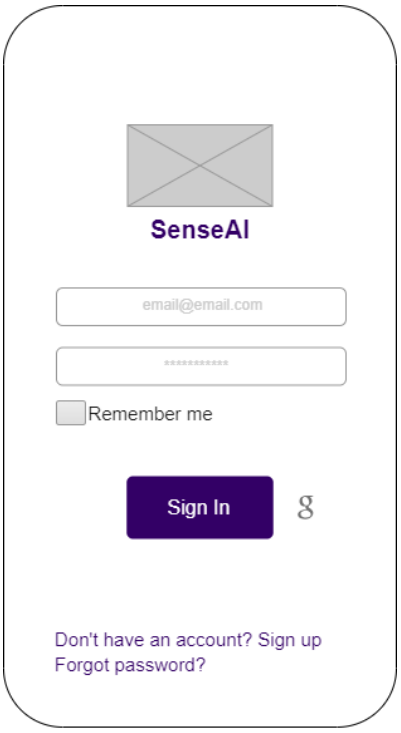


Figure 2 Login Page

Title	SenseAI
E-mail Field	Field where the user will enter his/her e-mail address. Accepts data only in e-mail format.
Password Field	The field where the user will enter the password. The characters are displayed hiddenly.
Remember Me Checkbox	Provides the option for the user to remember their login information.
Sign In Button	The main action button that allows the user to log in.
Google Login Icon	Icon that allows you to log in with a Google account.
Sign Up Link	Referral link for user to create account
Forgot Password Link	Allows the user to go to the reset screen when they forget their password.

Sign up Page

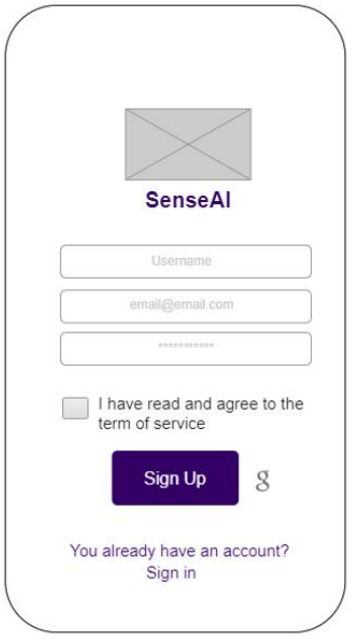


Figure 3 Sign up

Title	SenseAI
Username Field	Textbox for entering the user's username.
E-mail Field	Field where the user will enter his/her e-mail address. Accepts data only in e-mail format.
Password Field	The field where the user will enter the password. The characters are displayed hiddenly.
Checkbox	Checkbox labeled *"I have read and agree to the term of service"*. Users must check this box before signing up.
Sign Up Button	Main button for account creation, highlighted in purple.
Google Sign-Up Button	Secondary button with Google icon for signing up via Google authentication.
Sign In Link	Text link for existing users to navigate to the login page. Label: *"You already have an account? Sign in"*

Main Page

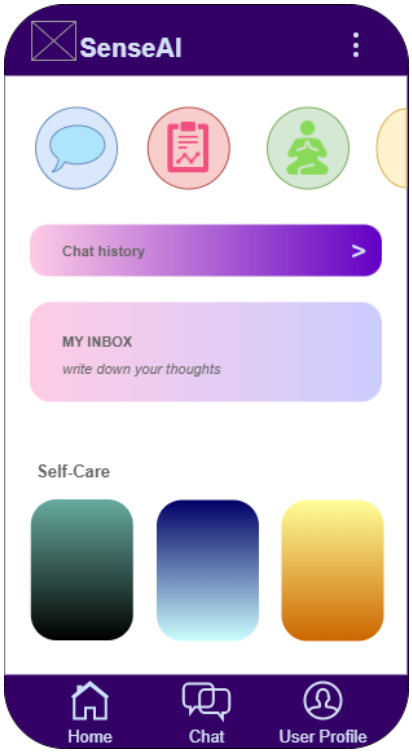


Figure 4 Main Page

Title	SenseAI
Chat History Section	Button with text and a right arrow for accessing previous chat sessions. Label: *"Chat history"*.
My Inbox Section	Button with text for accessing a note-taking feature. Placeholder: *"Write down your thoughts"*.
Self-Care	A section displaying gradient-colored cards/buttons representing various self-care activities or tools.
Home Button	Tab with a home icon for navigating to the main page.
Chat Button	Tab with a chat bubble icon for accessing chat page.
User Profile Button	Tab with a user icon for see account details.
3 Dots Button	Tab with 3 dots icon for managing account.

AI Chat Page

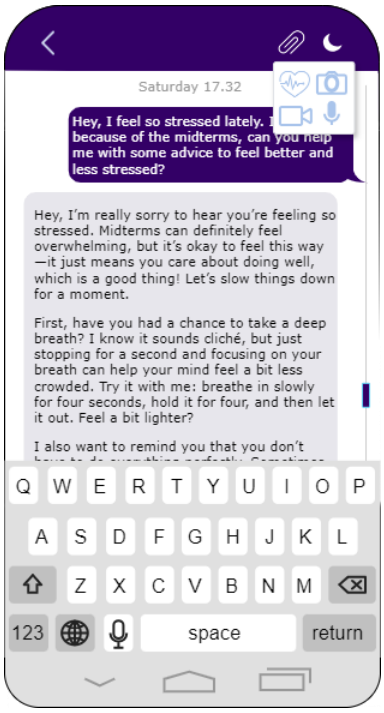


Figure 5 AI Chat

Date and time	Shows last conversation time.
Paperclip Button	Shows the options for the type of data to be sent.
Moon Button	Button that changes dark mode/light mode.
Heart Rhythm Button	Chatbot allows biometric data collection.
Camera Button	Chatbot allows to get photo data.
Video Button	Chatbot allows to get video data.
Microphone Button	Receives voice data from the user for the text-to-speech feature.
Purple Speech Bubble	Shows the query sent by the user to the chatbot.
Gray Speech Bubble	Shows the answers the chatbot returns to the user.
Keyboard	Used to receive text data from the user.
Back Button	Redirects to the previous page.

Chat History Page

Figure 6 Chat History

Title	SenseAI
Sub-Title	Recent chats
Title (in box)	Shows the title of the chat.
Date	Shows the chat's date.
Pen Button	Edit the chat.
Bin Button	Delete the chat.
Search Button	Search the chat from the chats.
Sub-Title	Previous 7 days
Plus Button	Redirects to the chat page to add a new chat.
Back Button	Redirects to the previous page.

User Profile Page

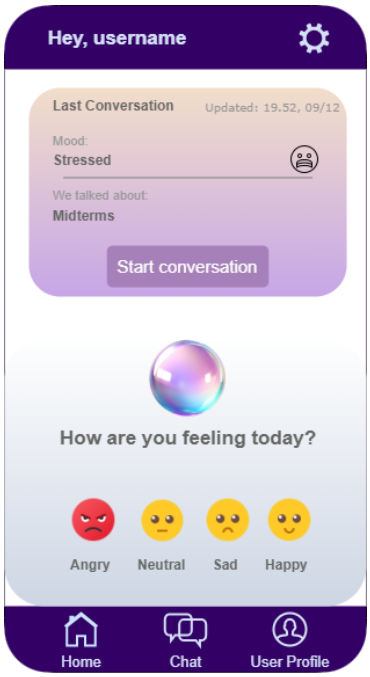


Figure 7 User Profile Page

Title	Hey, [username]
Sub-Title	Last Conversation
Date and time	Shows the last conversation date and time.
Sub-Title	Mood
Mood Field	Shows mood type.
Mood Icon	Shows mood.
Sub-Title	We talked about. (Shows the latest chat topic)
Chat name field	Shows the latest chat title.
Start Conversation Title	Tab for start new chat.
Sub-Title	How are you feeling today?
Mood Icons	Users select mood for the day.
Home Button	Tab with a home icon for navigating to the main page.
Chat Button	Tab with a chat bubble icon for accessing chat page.
User Profile Button	Tab with a user icon for see account details.
Settings Button	Tab for managing account.

Admin Pages

Admin Panel

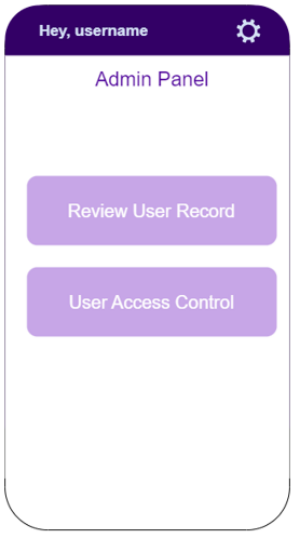


Figure 8 Admin Panel

Title	Hey, [username]
Sub-Title	Admin Panel
Settings Button	Tab for managing account.
Review User Record	Shows users records to admin.
User Access Control Button	Shows user access control page to admin.

User Record Page

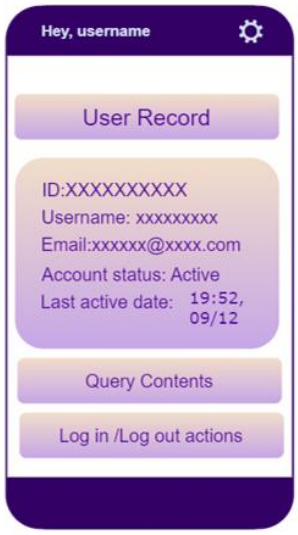


Figure 9 User Record Page

Title	Hey, [username]
Sub-Title	User Record
Record Field	Shows the recorded information of user to admin.

Query Contents	Redirected query content page.
Log in/Log out actions	Redirected log in/out actions page.
Settings Button	Tab for managing account.

Query Page

Figure 10 Query

Title	Query Content
Query Botton	Shows the queries to admin. Tab to see information about query.
Settings Button	Tab for managing account.

Log Information Page

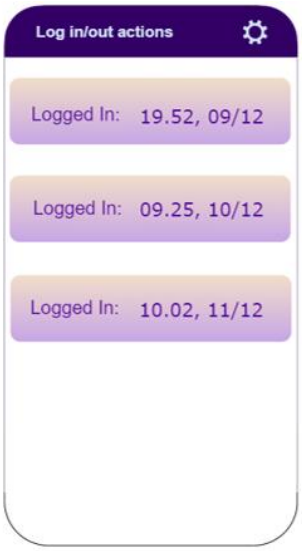


Figure 11 Log Information Page

Title	Log in/out actions
Logged in/out Fields	Shows the login/out information of user to admin.

Date and time	Shows users log in/out information to admin.
Settings Button	Tab for managing account.

User Access Control Page

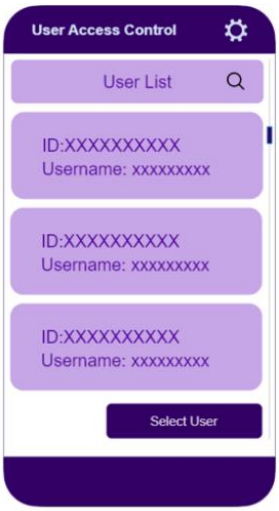


Figure 12 User List Page

Title	User List
User Fields	Shows users as a list.
Search Button	Admin searches users with ID or username.
Select User Button	Tab the one of user then tab to this button for select the user.
Settings Button	Tab for managing account.

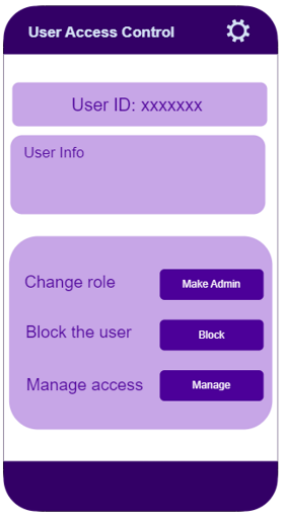


Figure 13 Access Control Page

Title	User Access Control
User ID Field	Shows user's ID information.
User Info Field	Shows the information of user.
Sub-Title	Change Role
Make Admin Button	Tab to button form changing user role.
Sub-Title	Block the user.

Block Button	Tab to block the user.
Sub-Title	Manage Access
Manage Button	Tab to manage accessing of user.
Settings Button	Tab for managing account.

Configuration Page

Figure 14 Admin Configuration

Title	Configuration
Search Bar	Allows searching configurations.
Add Button	Allows adding new configurations.
Configuration Box	Opens configuration detail popup.
Edit Button	Allows editing the configuration.
Settings Button	Tab for managing account.

Forgot Password Pages

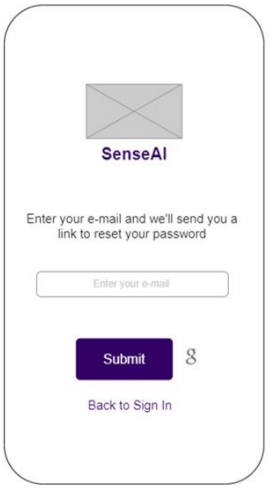


Figure 15 Entering E-mail Address Page

Title	SenseAI
Explanation Part	Explanation for directing user.
E-mail Field	Field where the user will enter his/her e-mail address. Accepts data only in e-mail format.
Submit Button	Submit for e-mail address.
Back to Sign in Link	Redirected user Sign in Page.
Google Sign-Up Button	Secondary button with Google icon for signing up via Google authentication.

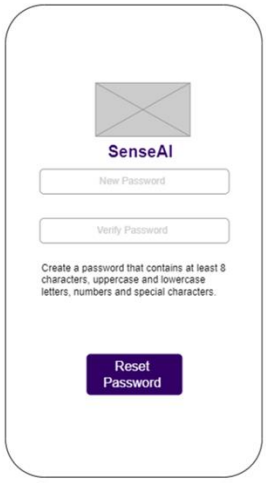


Figure 16 New Password Page

Title	SenseAI
New Password Field	Field where the user will enter new password.
Verify Password Field	Field where the user will verify new password.
Reset Password Button	Button for resetting password.
Explanation Part	Explain the new password constraints.

Settings Page

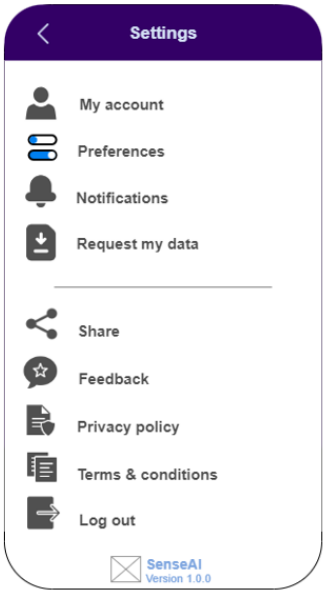


Figure 17 Settings

Title	Settings
My Account Button	With this button user can see their account.
Preferences Button	With this button users can manage preferences.
Notifications Button	With this button users can manage notification preferences.
Request my data Button	Used to obtain user information.
Share Button	Copies the link to share the app.
Feedback Button	With this button users can rate the system.
Privacy Policy Button	Contains information regarding the use of personal data.
Terms & conditions Button	Shows the terms of use of the service.
Log out Button	With this button users can log out the system.

High Level Design

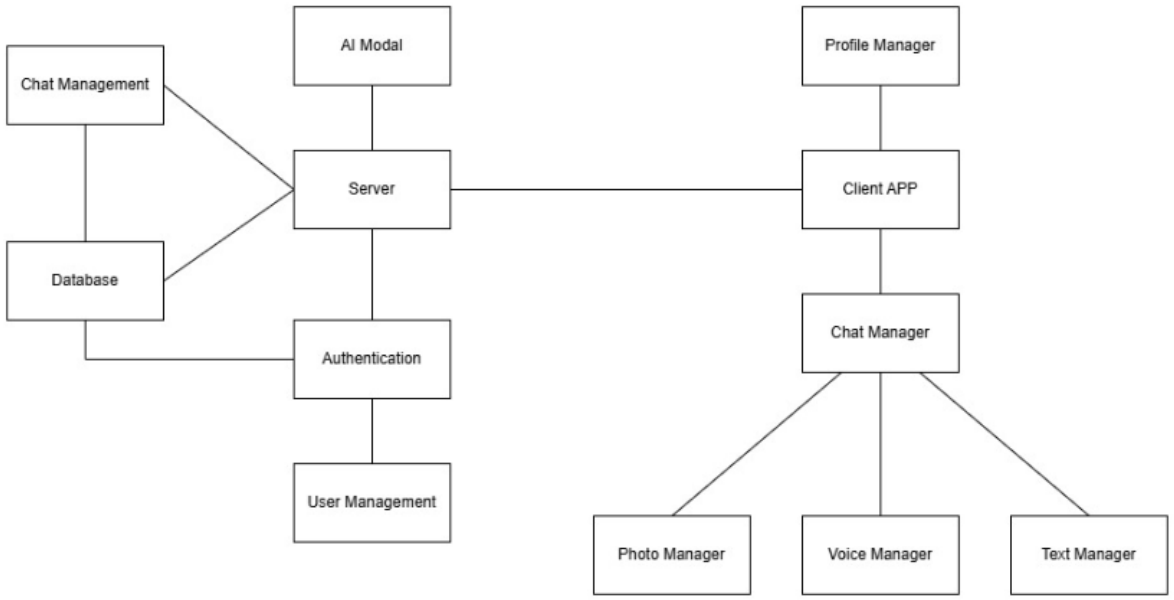


Figure 18 Module Diagram

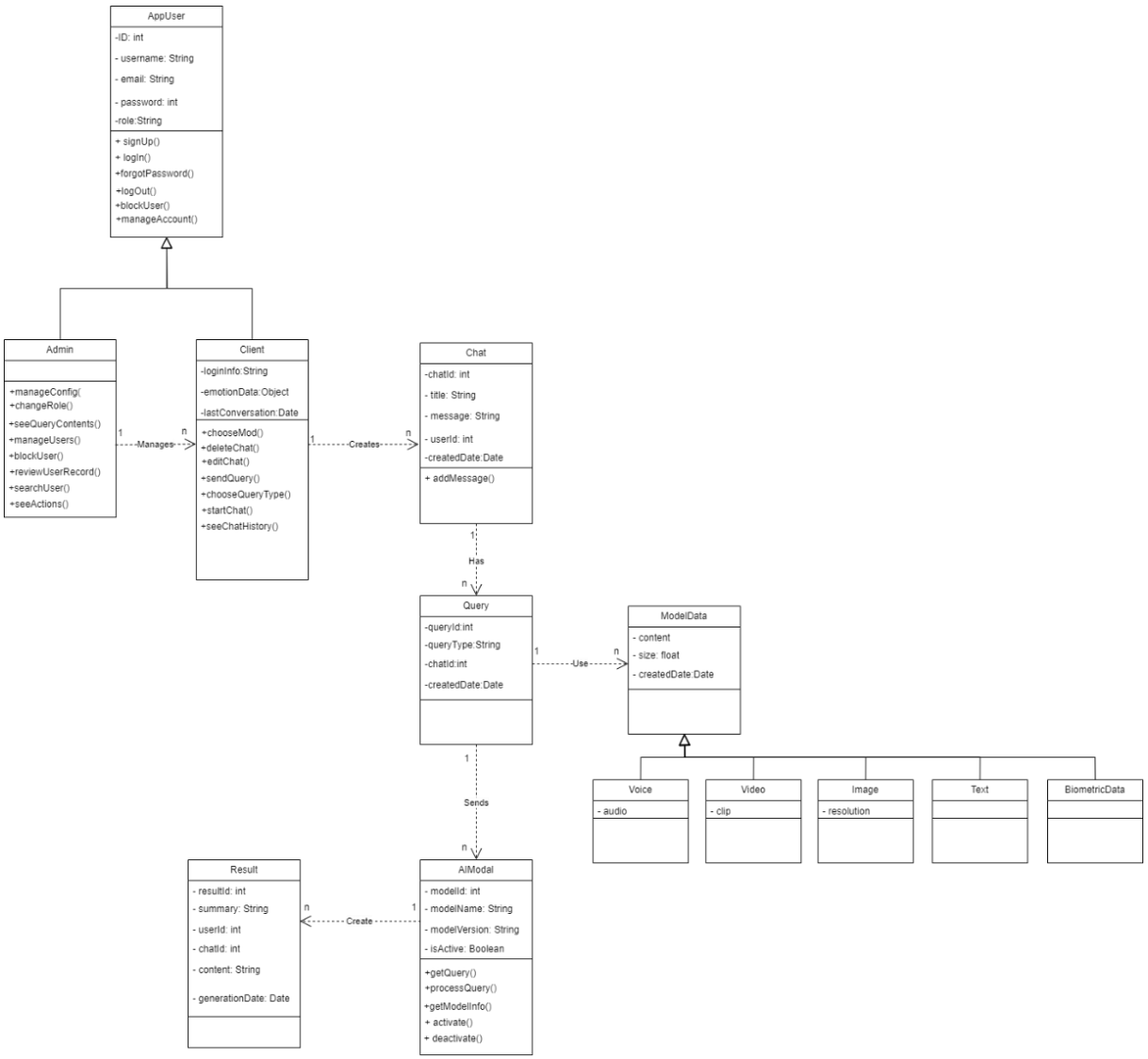


Figure 19 Class Diagram

Server Module

The Server module provides the core backend infrastructure that runs the entire system. It handles HTTP requests and manages data communication between the client and server. Additionally, it provides real-time messaging services using technologies like socket.io. The server coordinates the overall operation of the application and works in integration with other modules to fulfill user requests. Data from users is processed through the server and appropriate responses are sent back to the user.

AIModel Module:

The AIModel module is an artificial intelligence model that generates responses to user messages. This model works with a specific AI algorithm to analyze each incoming query and generate the most accurate response. AIModel can be customized based on the AI model used in the project and can continuously evolve based on user interactions. This module plays a critical role in ensuring that the conversation occurs in a meaningful and natural way.

Chat Management Module

The Chat Management module is responsible for managing the chats that users have. This module ensures that chat histories are stored, messages are organized and sorted. Each message is stored with details like which user it belongs to and which chat session it is associated with. This module is critical for providing users with access to previous chats, ensuring a more comprehensive and consistent conversation history.

Profile Management Module

The Profile Management module manages the editing and updating of user profiles. User information is stored and can be edited here. This module allows users to manage all information related to their accounts.

Authentication Module

The Authentication module handles the authentication process of users. This module ensures a secure authentication process when users log in or create a new account. Typically, token-based authentication is used to manage each user's access level. The Authentication module is crucial for maintaining the security of the system.

Client App Module

The Client App module is the interface where the user interacts with the chatbot. This module typically works as a web or mobile application, enabling users to chat using text, voice, images, and other media types. The Client App captures the messages typed by users, sends them to the server, and displays the responses from the server. The user interface (UI) and user experience (UX) are optimized to provide a friendly and engaging conversation experience. In this module, users receive responses generated by the AIModel and continue their interactions with the chatbot.

Chat Manager Module

The Chat Manager module is responsible for handling the type of queries sent during a chat session. It allows users to select the type of query (such as text, photo, video, etc.) they want to send. This module ensures that the correct processing is applied based on the selected query type, enabling flexible interaction with the chatbot. This module is also responsible for displaying the output to the user.

Low Level Design

Authentication Module

This diagram explains the user's registration and login to the system.

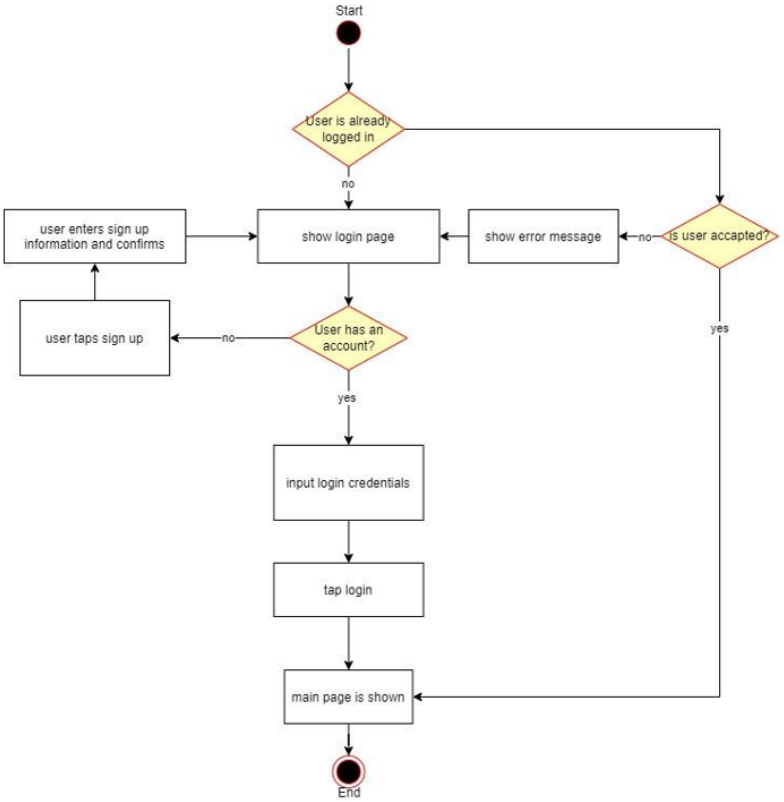


Figure 20 Authentication Activity Diagram

Chat Manager Module

It is a module where the type of query that the user will send is determined. The modules and steps selected by the user are shown in the diagrams below.

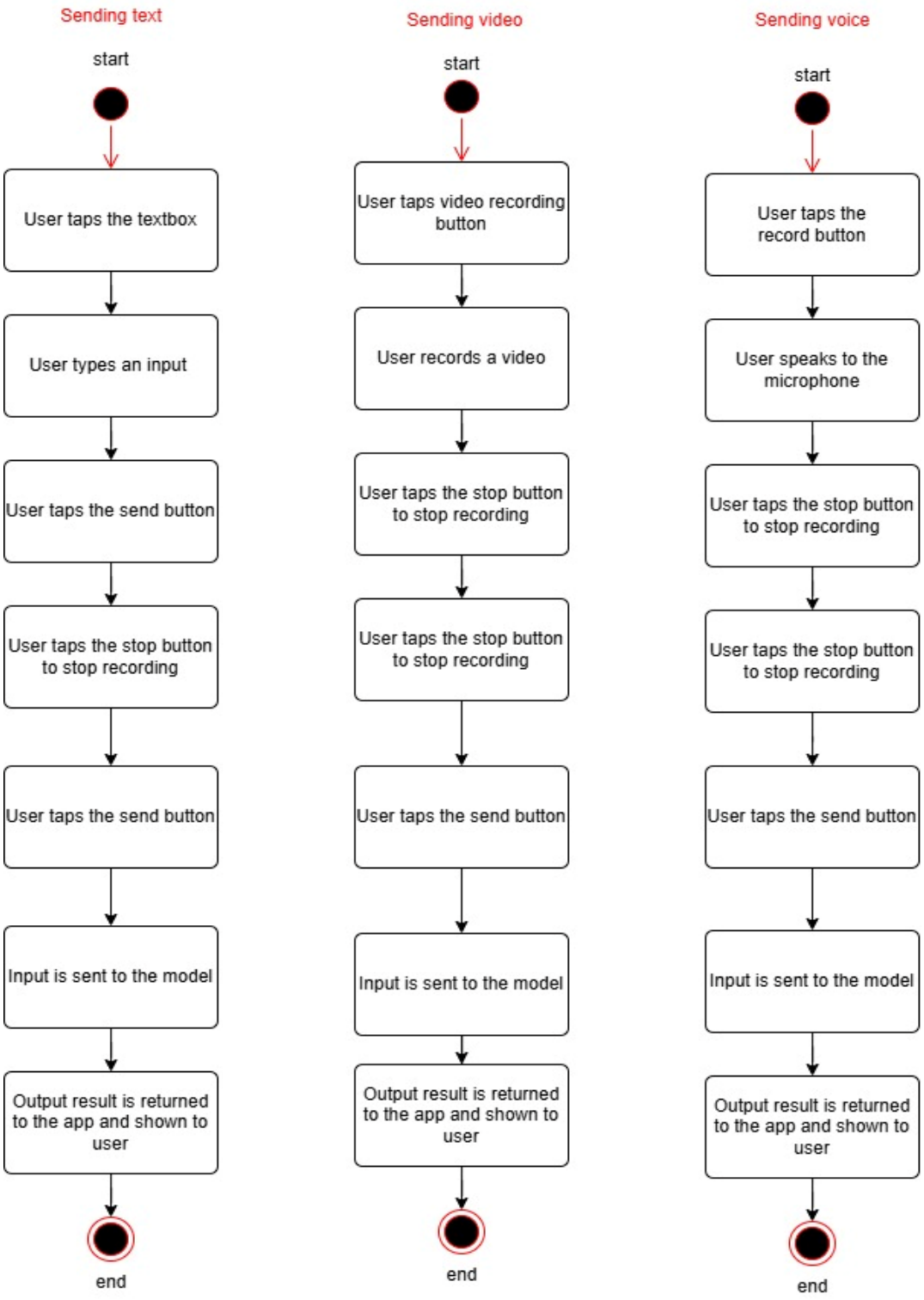


Figure 21 Chat Activity Diagram

Profile Management Module

The steps that users take to change their account properties are shown in the diagram below.

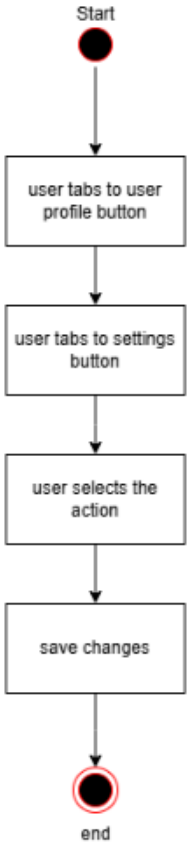


Figure 22 Settings Activity Diagram

Database Design

E-R Diagram

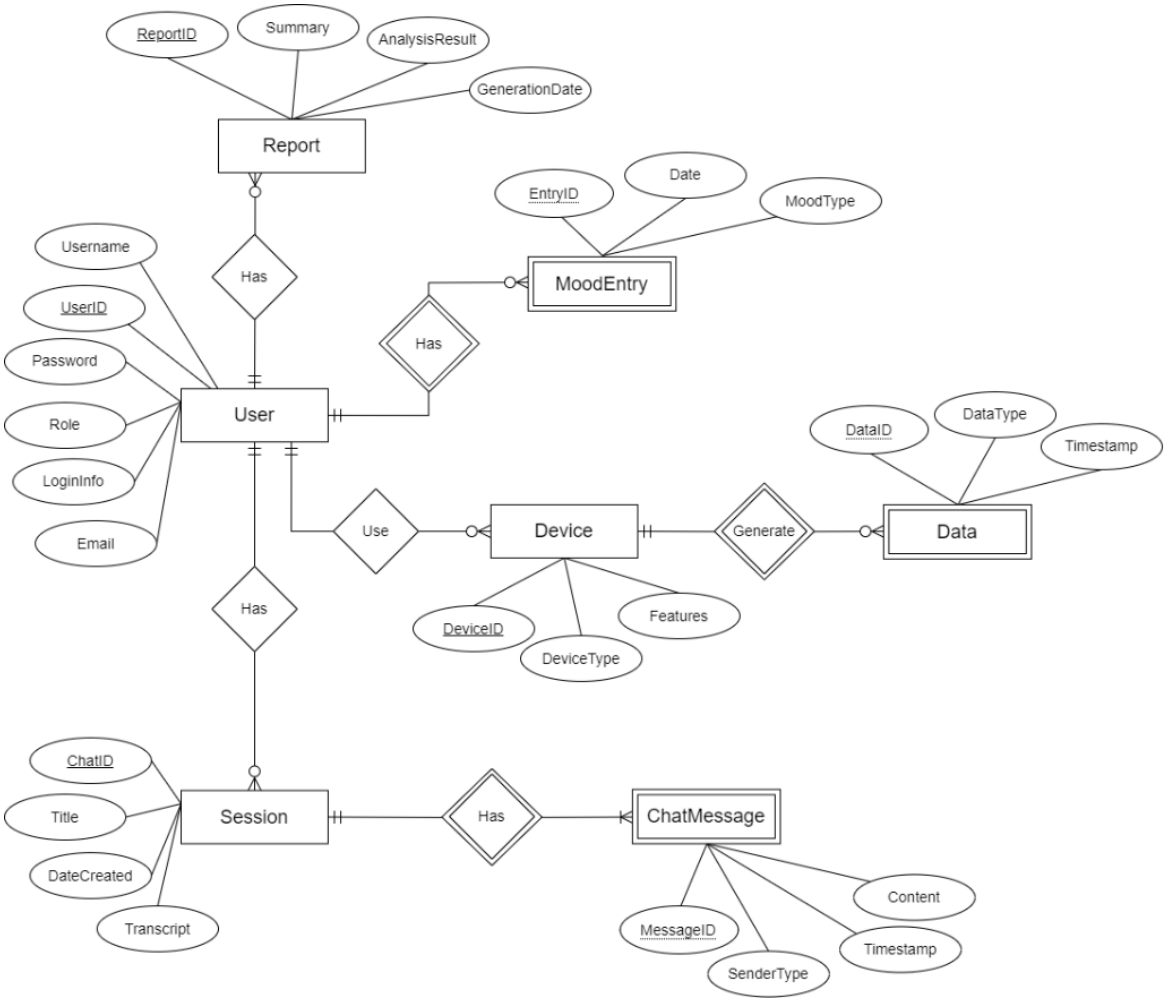


Figure 23 ERD (Entity-Relationship)

Tables

User

Attributes:

1. **UserID**: Primary Key, Integer, Unique identifier for each user.
2. **Username**: String, the name chosen by the user.
3. **Password**: String, Encrypted password for the user.
4. **Role**: String, The role of the user (e.g., Admin, User).
5. **LoginInfo**: String, Information about the user's login records.
6. **Email**: String, User's email address.

Report

Attributes:

1. **ReportID**: Primary Key, Integer, Unique identifier for each report.
2. **Summary**: String, A summary of the report.
3. **AnalysisResult**: String, the analysis result included in the report.
4. **GenerationDate**: Date, The date the report was generated.
5. **UserID**: Foreign Key, Integer, References the User table.

MoodEntry

Attributes:

1. **EntryID**: Partial Key, Integer, Unique identifier for each mood entry.
2. **Date**: Date, the date the mood entry was recorded.
3. **MoodType**: String, The type of mood (e.g., Happy, Sad).
4. **UserID**: Foreign Key, Integer, References the User table.

Session

Attributes:

1. **ChatID**: Primary Key, Integer, Unique identifier for each session.
2. **Title**: String, The title of the session.
3. **DateCreated**: Date, the date the session was created.
4. **Transcript**: Text, The transcript of the session.
5. **UserID**: Foreign Key, Integer, References the User table.

ChatMessage

Attributes:

1. **MessageID**: Partial Key, Integer, Unique identifier for each chat message.
2. **Content**: Text, The content of the message.
3. **Timestamp**: DateTime, the time the message was sent.
4. **SenderType**: String, Type of sender (e.g., User, System).
5. **ChatID**: Foreign Key, Integer, References the Session table.

Device

Attributes:

1. **DeviceID**: Primary Key, Integer, Unique identifier for each device.
2. **DeviceType**: String, The type of device (e.g., Wearable, Smartphone).
3. **Features**: String, Features of the device.
4. **UserID**: Foreign Key, Integer, References the User table.

Data

Attributes:

1. **DataID**: Partial Key, Integer, Unique identifier for each data entry.
2. **DataType**: String, The type of data (e.g., Sensor Data, Analysis Data).
3. **Timestamp**: DateTime, the time the data was generated.
4. **DeviceID**: Foreign Key, Integer, References the Device table.

Conclusion

This document outlines the AI-based Psychology Chatbot project, including a Literature Review, Software Requirements Specification (SRS), and Software Design Document (SDD). The Literature Review explores multimodal emotional analysis, technologies like NLP and computer vision, and ethical considerations in mental health chatbot development. The SRS defines the project's objectives, scope, and specifications, while the SDD details its architecture and features, supported by UML diagrams to visualize system interactions and workflows.

This project leverages multimodal data—text, speech, facial expressions, and biometrics—to analyze emotional states, provide personalized feedback, and track users' emotional patterns over time. By integrating advanced AI technologies, the chatbot aims to enhance psychological support and emotional well-being responsibly.

In conclusion, the AI-based Psychology Chatbot project represents a comprehensive effort to create a meaningful digital mental health tool. Through the integration of the literature review, clearly stated requirements, and detailed design with UML diagrams, this document provides a solid foundation for the project. As students, we aimed to address the complexities of emotional analysis while ensuring the chatbot's functionality is robust and user-focused. The project reflects our understanding and application of modern AI techniques to a critical area of need.

References

1. Wang, S., Zhang, Z., Chen, B., Zhang, Z., Fang, W., and Pan, J. (2022). "Multimodal Emotion Recognition Based on Facial Expressions, Speech, and EEG." *IEEE Transactions on Affective Computing*, vol. 13, no. 1, pp. 142-154. doi: 10.1109/TAFFC.2021.10026861.
2. Lian, H., Lu, C., Li, S., Zhao, Y., Tang, C., and Zong, Y. (2023). "A Survey of Deep Learning-Based Multimodal Emotion Recognition: Speech, Text, and Face." *Entropy*, vol. 25, no. 10, pp. 1440. doi: 10.3390/e25101440.
3. Yan, J., Li, P., Du, C., Zhu, K., Zhou, X., Liu, Y., and Wei, J. (2024). "Multimodal Emotion Recognition Based on Facial Expressions, Speech, and Body Gestures." *Electronics*, vol. 13, no. 18, p. 3756. doi: 10.3390/electronics13183756.
4. Sharma, A., Saxena, A., Kumar, A., and Singh, D. (2023). "Depression Detection Using Multimodal Analysis with Chatbot Support." *2023 IEEE International Conference on Computational Intelligence and Data Science (ICCIDS)*, Chennai, India, pp. 1-6. doi: 10.1109/ICCIDS56901.2023.10489080.
5. Woebot Health. "Woebot Health - Mental Health for the Digital Age." Available: https://woebothealth.com/. [Accessed: 24-November-2024].
6. Wysa. "Reports." Available: https://www.wysa.com/reports#6. [Accessed: 24-November-2024].
7. Youper. "AI Therapy." Available: https://www.youper.ai/. [Accessed: 24-November-2024].
8. Headspace for Work. "Headspace for Work." Available: https://organizations.headspace.com/. [Accessed: 24-November-2024].
9. News.com.au. "Meet Ellie, the robot therapist treating soldiers with PTSD." Available: https://www.news.com.au/technology/innovation/meet-ellie-the-robot-therapist- treating-soldiers-with-ptsd/news-story/0201fa7cf336c609182cffd637deef00. [Accessed: 24-November-2024].
10. Replika. "Replika - My AI Friend." Available: https://replika.ai/. [Accessed: 24-Oct-2024].
11. American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5). Arlington, VA: American Psychiatric Publishing, 2013.
12. American Psychological Association, Guidelines for the Practice of Telepsychology. [Online]. Available: https://www.apa.org/practice/guidelines/telepsychology, 2013.
13. Google. (2024). Material Design Guidelines. Retrieved from https://material.io/design