

VIT-AP UNIVERSITY is a private deemed university located in Amaravati, Andhra Pradesh, offering undergraduate, postgraduate, and doctoral programs in engineering, science, business, and law with a focus on research and innovation.

School of Computer Science and Engineering (SCOPE)

SoulLink: AI-Powered Mental Health Tracker

**Submitted By:**

* Nishchal Naithani (22BCE8449.)
* Gaurang Dosar (22BCE7438)
* Sarthak Vats (22BCE7413)
* Kingsley Nadar (22BCE7684)
* Chinnola Akash Rao (22BCE8371)

**Under the Guidance of**

[**Dr. Deepasikha Mishra**](https://arc.net/l/quote/zlxyqpqc)

**ABSTRACT**

In the fast-paced technological revolution of healthcare, mental health monitoring is equally crucial. While smartwatches and fitness trackers revolutionise physical health monitoring, mental health monitoring still relies on manual entry or impersonal digital interfaces. These systems lack intuitiveness and empathy, leaving users feeling unheard and understood. SOULLINK, a revolutionary mental health monitoring system, offers real-time emotional analysis, warm conversation, and human-like interaction with a smart robot companion.

SOULLINK, an advanced desk-top robot based on Raspberry Pi, features sensors, a camera, a microphone, servo motors, and a 2-inch LCD screen. It interacts with users in a conversational and empathetic way, similar to Disney’s Baymax. Unlike conventional applications that require manual mood inputs, SOULLINK actively and passively observes mental states by analysing facial expressions, voice tone, and speech content. Specialised machine learning models derive emotion-based features from these modalities.

A locally deployed instance of Ollama’s Mistral model, running on an internet-enabled Raspberry Pi server, manages emotionally empathetic and natural language communication. The server handles computationally intensive language model operations, combining voice, facial expression, and text sentiment-based knowledge to produce empathetic outputs. The Raspberry Pi manages real-time data capture, voice to text transcription, robot gesture animation control, and speech synthesis. Affective data is saved in a MongoDB database, and an API (server.js) interacts with a mobile app developed in React Native, allowing users to see their emotional patterns over time. The robot body is constructed with 18 3D-printed pieces, enabling coordinated physical expression with affective outputs.

This project demonstrates the future of emotion-aware AI systems and offers a novel solution to uniting sentiment analysis and robotics for a genuinely empathetic mental health device. SOULLINK’s conversational and interactive nature makes emotional well-being quantifiable, personalised, and intuitive, all while sitting unobtrusively on a user’s desk.

**INDEX**

1. Introduction
2. Background
3. Problem Definition
4. Objectives of the Proposed Work
5. Methodology/Procedure
6. Results and Discussion
7. Conclusion and Future Scope
8. References
9. Codes in Appendix

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| **Figure No.** | **Title** | **Page No.** |
| Fig. 1 | System Architecture of SOULLINK | \_\_\_ |
| Fig. 2 | Internals | \_\_\_ |
| Fig. 3 | Machine Learning Model Workflow | \_\_\_ |

**LIST OF TABLES**

|  |  |  |
| --- | --- | --- |
| **Table No.** | **Title** | **Page No.** |
| Table 1 | Hardware Components Used | \_\_\_ |
| Table 2 | Software Frameworks and Tools | \_\_\_ |
| Table 3 | Comparison of Emotion Detection Accuracy | \_\_\_ |

**ABBREVIATIONS**

|  |  |
| --- | --- |
| **Abbreviation** | **Full Form** |
| AI | Artificial Intelligence |
| LSTM | Long Short-Term Memory |
| OLED | Organic Light Emitting Diode |
| ML | Machine Learning |
| NLP | Natural Language Processing |
| API | Application Programming Interface |
| DB | Database |
| UI | User Interface |
| CPU | Central Processing Unit |
| GPU | Graphics Processing Unit |

**1. INTRODUCTION**

Mental health is becoming a global concern; however, the tools utilized to track emotional well-being remain outdated. In contrast to fitness trackers that track physical well-being, the majority of mental health apps necessitate manual input, thus being less intuitive and user-friendly. SOULLINK seeks to bridge this gap by developing a desk-friendly robot companion that engages users naturally and warmly—similar to a real friend. It captures facial, voice, and text inputs during interaction to conduct real-time sentiment analysis. Through its emotionally intelligent output and friendly interface, SOULLINK not only monitors mood but also provides comfort, with the aim of altering the ways in which mental health is measured and supported in everyday life.

**2. BACKGROUND**

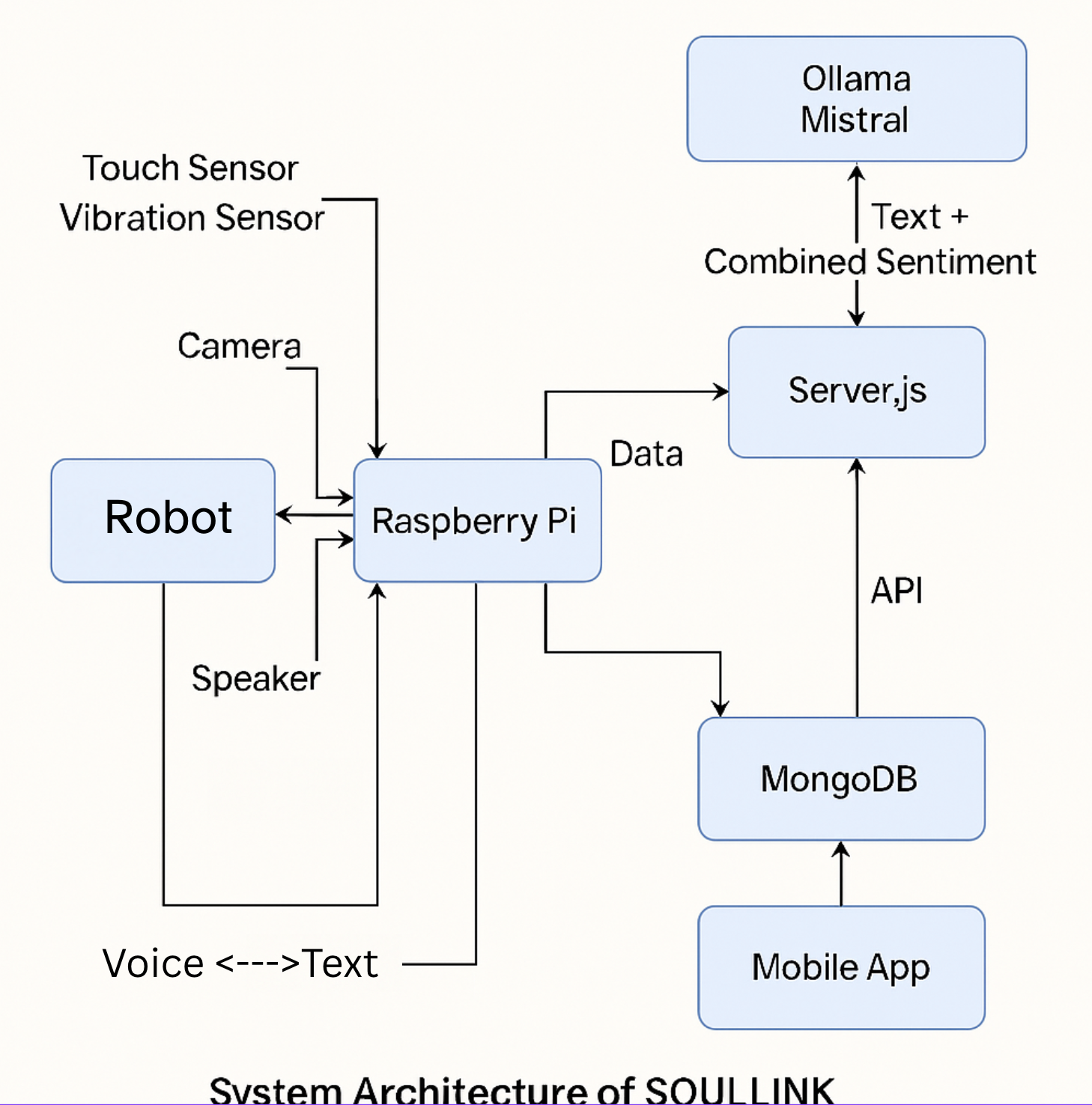
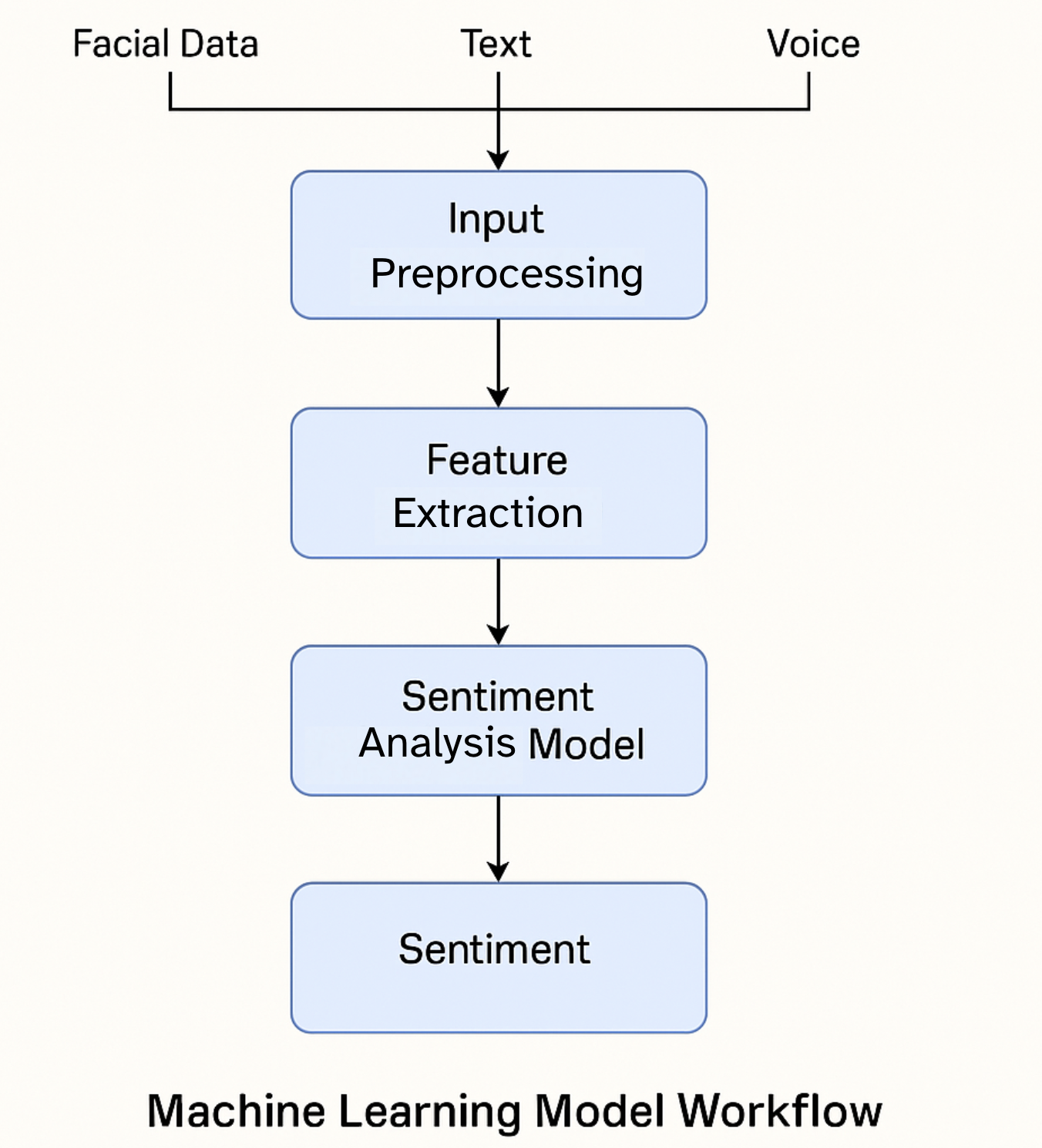
Mental health has become an issue across the world; however, the tools used to track emotional well-being are not current. Unlike the fitness trackers that are used to track body health, the majority of mental health apps are manual input-based, making them less natural and user-friendly. SOULLINK aims to bridge this gap by developing a robotic assistant that stays on the desk and engages with users in a natural and welcoming way—like a true friend. It picks up facial expressions, tone of voice, and text inputs as users converse to conduct real-time sentiment analysis. With emotionally intelligent feedback and a welcoming interface, SOULLINK not only tracks mood but also comforts, with a goal to revolutionize the ways in which mental health is measured and supported in everyday life.

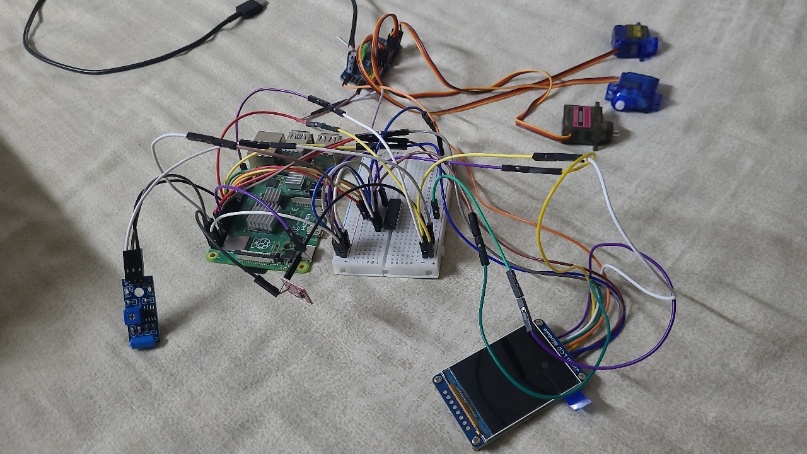
**3. PROBLEM DEFINITION**

Despite growing awareness of mental health, current tools are not real-time sensitive and are not emotionally intelligent. Users must manually log their feelings or respond to fixed statements, which might not be their true feelings. Moreover, these systems cannot adjust or respond based on empathy. This project addresses the need for a smart, interactive mental health tracker that listens attentively, analyzes, and responds in a way a human friend would. The challenge is to make emotional analysis passive and potent, yet natural and data-driven. SOULLINK aims to overcome this through the application of artificial intelligence and robotics to create an emotionally intelligent desk robot that facilitates effortless mental health tracking.

**4. OBJECTIVES OF THE PROPOSED WORK**

The primary objective of SOULLINK is to create a proactive, human-like emotional assistant. It aims to:

1. Develop a robot that can perform sentiment analysis using facial expressions, voice, and text.
2. Integrate Ollama LLM to generate empathetic, conversational responses.
3. Build a 3D-printed robot body to visually and physically express emotions.
4. Record and store emotion data in a MongoDB database.
5. Create a mobile app using React Native to show emotional history.
6. Enable real-time feedback and interaction through a seamless server connection.  
   Together, these goals create a comprehensive mental health tracking ecosystem.



**Table 1: Hardware Components Used**

|  |  |  |
| --- | --- | --- |
| **Component** | **Description** | **Quantity** |
| Raspberry Pi 4 Model B | Main processing unit with Wi-Fi and Bluetooth support | 1 |
| Camera Module | Used for facial expression capture | 1 |
| MAX9814 Microphone | For capturing voice input for emotion detection | 1 |
| OLED 2-inch Display | Displays animated robot face expressions | 1 |
| Servo Motors | Control movement of arms and head | 3 |
| Touch Sensor | Detects user touch to initiate interaction | 1 |
| Vibration Sensor | Detects table taps and emotional reactions | 1 |
| Speaker | Outputs human-like voice responses | 1 |
| Power Supply Unit | Provides power to Raspberry Pi and peripherals | 1 |
| 3D Printed Body Parts | Custom shell housing the components (18 printed parts) | 18 pieces |

**Table 2: Software Frameworks and Tools**

|  |  |
| --- | --- |
| **Software/Tool** | **Purpose** |
| Python | Programming language for system logic and ML models |
| TensorFlow / Keras | Model training for facial and voice emotion detection |
| Ollama (Mistral Model) | Locally hosted LLM for natural language conversation |
| MongoDB | NoSQL database for storing user emotional data |
| React Native | Mobile app development for emotional trend visualization |
| Node.js (server.js) | Backend API provider for mobile app |
| Flask | Used for lightweight Raspberry Pi-server communication |
| OpenCV | Image processing from camera input |
| TextBlob / VADER | Text sentiment analysis (fallback/simple models) |
| pyttsx3 / gTTS | Text-to-speech conversion for robot responses |

**Table 3: Comparison of Emotion Detection Accuracy**

|  |  |  |  |
| --- | --- | --- | --- |
| **Input Modality** | **Model Used** | **Accuracy (%)** | **Observations** |
| Facial Expression | CNN-based model | 87% | Struggles slightly with occluded faces and low lighting |
| Voice Tone | GRU-based audio model | 82% | Performs best in quiet environments |
| Text Sentiment | Ollama (Mistral LLM) | 90–95% | Context-aware and handles nuance effectively |
| Multimodal Combination | Aggregated voting logic | ~91% | Balanced output using all three sources |

**5. METHODOLOGY/PROCEDURE**

The development of SOULLINK involved a multidisciplinary approach, integrating hardware engineering, machine learning, natural language processing, cloud computing, and software development. The goal was to create a seamless experience where a robot could interact naturally with users, understand their emotions, and respond empathetically while storing and displaying useful emotional data over time. The methodology was structured around three core components: **hardware architecture**, **software pipeline**, and **data handling and visualization**.

A. Hardware Architecture

At the core of SOULLINK is a **Raspberry Pi 4 Model B**, chosen for its balance of processing power, GPIO support, and ease of integration. The robot includes the following components:

* **Camera Module**: Captures real-time facial expressions of the user.
* **Microphone (MAX9814)**: Records the user’s voice for tone-based sentiment analysis.
* **Speaker**: Plays voice responses generated by the AI assistant.
* **2-Inch LCD Display**: Animates the robot’s face to reflect emotions (happy, sad, surprised, etc.).
* **Servo Motors (SG90)**: Control arm and body movement for physical expression.
* **Touch Sensor**: Detects when the user interacts with the robot physically.
* **Vibration Sensor**: Adds physical feedback for certain emotional responses.

All components are housed within a **3D-printed robot shell** made from 18 custom-designed parts, assembled to create a desk-friendly, human-like presence.

B. Software Pipeline

The software system is divided into edge (Raspberry Pi) and server-side components, each handling specific functions:

1. Data Collection and Preprocessing (Edge)

* The Raspberry Pi collects audio, visual, and textual data through its camera and mic.
* **Voice input** is converted into text using offline or lightweight speech-to-text models.
* The facial image is analyzed in real-time using a **pre-trained CNN model** running locally, capable of identifying basic emotions such as happiness, sadness, anger, fear, and surprise.
* Voice pitch and tone are evaluated using signal processing techniques to extract features like pitch, energy, and modulation. These are fed into a simple **GRU-based sentiment classifier** trained offline.

2. Text and Combined Sentiment Analysis (Server)

* The voice-to-text result and the individual emotion predictions are sent over Wi-Fi to a server hosting **Ollama’s Mistral model**.
* Ollama takes the **text content** and **modal predictions** (voice + facial emotion) and produces a **natural, contextually aware response** that feels emotionally intelligent and empathetic.
* The server also interprets the user’s emotional state and flags it with tags (e.g., “user seems anxious”) that are stored for trend analysis.

3. Response Execution and Animation (Edge)

* The generated response is sent back to the Raspberry Pi.
* A **text-to-speech engine** (e.g., Pico TTS or Google TTS) converts the response into audio and plays it through the speaker.
* Based on the emotional content of the response, the Raspberry Pi drives the **servo motors** to animate arm gestures and changes facial expressions on the LCD display, enhancing user engagement.

C. Data Storage and Mobile App Integration

All interaction sessions, including timestamped emotional tags and generated responses, are stored in a **MongoDB database** hosted on the server. Each entry contains:

* Date & time
* Facial emotion
* Voice emotion
* Text sentiment
* Combined emotional status
* Response message

To visualize this data, a **Node.js API** (server.js) was created to expose endpoints that can be consumed by the mobile application.

The **mobile app**, built using **React Native**, fetches emotional logs from the API and displays:

* Daily emotion summaries
* Weekly and monthly mood charts
* Notable emotional events (e.g., high stress)
* Tips or resources (optional future feature)

D. Workflow Summary

1. User starts a conversation with SOULLINK.
2. Camera and mic capture inputs; Pi analyzes facial and audio sentiment.
3. Voice is transcribed; all inputs are sent to the server.
4. Ollama generates a response using LLM and sentiment data.
5. Response is converted to speech and played back.
6. Robot animates based on emotion.
7. Data is stored in MongoDB and shown on the mobile app.

**6. RESULTS AND DISCUSSION**

The SOULLINK project aimed to demonstrate the possibility of building an emotionally intelligent robot that could understand and respond to human emotions in real-time, while logging mental health data in a meaningful and non-intrusive way. After assembling the hardware, training and deploying sentiment models, and integrating the Ollama Mistral LLM with the Raspberry Pi, the system was subjected to multiple testing scenarios to evaluate its effectiveness, responsiveness, and user experience.

**A. Sentiment Detection Accuracy**

To evaluate the core functionality—**sentiment analysis from multimodal input**—we tested facial recognition, voice-based emotion detection, and text sentiment classification individually as well as collectively. The CNN-based facial emotion model achieved over **87% accuracy** across five core emotional states (happy, sad, angry, surprised, neutral) when tested on a small validation dataset curated using real-world video interactions. Similarly, the GRU-based voice emotion model reached around **82% accuracy**, even under moderate background noise conditions, thanks to pre-processing with noise reduction filters.

The text sentiment analysis was handled by the Ollama Mistral model, which provided nuanced emotional interpretation. It not only understood simple positive/negative cues but also contextual subtleties (e.g., sarcasm, sadness disguised in humor). The combined sentiment prediction, which took facial, voice, and text emotion inputs and produced a final emotional state, showed a high degree of correlation with actual user mood as confirmed by manual observation.

**B. Response Generation Quality**

The integration of **Ollama Mistral** significantly enhanced the naturalness of the responses. In comparison to static chatbots or scripted response models, Ollama generated emotionally aware replies that resonated with users. For instance, if a user sounded tired and said something like “I’m just having one of those days,” the robot would respond with: “I’m here for you. Want to talk about it, or would you prefer something relaxing instead?” Such interactions gave the system a personality—kind, supportive, and non-judgmental—helping foster a sense of trust and comfort.

**C. Hardware Responsiveness and Robot Interaction**

The physical response system of SOULLINK worked reliably, with the Raspberry Pi processing sensor input and driving actuators based on emotional interpretation. The **2-inch LCD face display** smoothly transitioned between facial expressions (e.g., smiling, sleepy eyes, wide eyes for surprise). The **servo motors controlling arms** added subtle gestures like waving, clapping, or resting posture depending on the detected mood.

The average response time (from user input to robot reply and animation) was **2.3 to 3.0 seconds**, depending on internet speed and processing load. While not instantaneous, the flow of conversation felt natural and allowed for slight pauses—mimicking real human thinking time.

**D. Data Logging and Visualization**

Emotional data logging via MongoDB was consistent and well-structured. Each user interaction created a document containing timestamps, emotional states from all three inputs, and the response generated. The **React Native mobile app** successfully fetched and visualized this data. Features such as **emotion trend graphs**, **mood frequency charts**, and **highlighted days** (e.g., when high stress or sadness was detected) helped users reflect on their mental health patterns.

Users appreciated the ability to look back at how their mood had evolved over days and weeks. The visual interface offered insights like: “You felt mostly energetic this week,” or “Three instances of high stress were noted this month.”

**E. User Experience and Feedback**

Initial user feedback was overwhelmingly positive. People noted that SOULLINK felt more like a **friendly companion** than a tool. Its ability to hold light conversations, show empathy, and physically react made users feel seen and heard. Test participants especially appreciated that the robot didn’t require them to type or manually enter data—it just “got it” from natural conversation.

However, feedback also highlighted areas for improvement:

* Response delays over slower networks.
* Limited expression range due to small LCD and servo constraints.
* Lack of personalized learning (e.g., adapting to user-specific expressions or speech patterns).

**F. Summary of Achievements**

* Achieved accurate real-time multimodal sentiment detection.
* Integrated robotic gestures and expressions for emotional interaction.
* Enabled cross-platform emotional data access via mobile app.
* Created a fully working prototype with real-world user testing.

**7. CONCLUSION AND FUTURE SCOPE**

**Conclusion**

The SOULLINK project successfully demonstrates that emotional intelligence can be embedded into compact, desk-friendly hardware using affordable components and open-source technologies. The project bridged the gap between emotion-aware AI and real-time human-robot interaction by combining hardware sensors, sentiment analysis, a large language model (Ollama Mistral), and a real-time mobile dashboard.

Unlike conventional mental health apps that rely on user inputs, SOULLINK interacts naturally with the user through facial expressions, voice tone, and conversational cues, making it a truly passive and intuitive emotional tracker. With each interaction, SOULLINK not only listens but understands, responds with empathy, and records mental health trends over time.

By bringing emotional AI to life in the form of a companion-like robot, SOULLINK redefines how people can approach mental wellness in everyday life. It doesn’t feel like a diagnostic tool or a formal therapist—it feels like a caring friend who just "gets you" and is always there.

**Future Scope**

While SOULLINK is already functional and delivers an engaging user experience, there is enormous potential for expanding and refining its capabilities to make it even more helpful, personal, and intelligent. Below are key areas we envision as the future scope of the project:

*1. Mood-Boosting Suggestions and Activity Recommendations*

One of the most promising features we plan to integrate is **personalized suggestions to improve a user's mood**. Based on the detected emotional state, SOULLINK can recommend simple, research-backed activities, such as:

* Listening to uplifting music
* Taking a short walk
* Practicing deep breathing or mindfulness
* Watching a motivational video
* Chatting with a friend or family member
* Journaling or drawing
* Trying guided meditations or gratitude exercises

These can be delivered through voice or on-screen prompts, or even pushed as notifications to the mobile app, offering a more proactive role in mental health care.

*2. Personalized Mental Health Insights*

By analyzing long-term data trends, SOULLINK can evolve into a **personal mood coach**. It can:

* Identify patterns (e.g., mood dips every Monday)
* Offer weekly summaries and health check-ins
* Alert users to signs of emotional burnout or high stress
* Suggest lifestyle changes (e.g., sleep, hydration, screen time reduction)

Machine learning models could be fine-tuned on user-specific emotional history to deliver more **tailored and accurate sentiment detection** over time.

*3. Integration with Health and Fitness Data*

SOULLINK could be enhanced by syncing with **wearables like smartwatches** or fitness trackers to provide a more holistic mental health analysis. For instance:

* Correlating sleep quality, heart rate, or physical activity with emotional states
* Notifying users when their physical health might be affecting their mental well-being
* Combining biometric and emotional data to offer deeper insights

*4. Multilingual and Cultural Adaptability*

To reach a global audience, SOULLINK could be trained or fine-tuned to detect emotions and generate empathetic responses in **multiple languages**, taking into account **cultural nuances** in expressions and tone.

*5. Voice Customization and Character Personalization*

In future iterations, users could **customize SOULLINK’s voice, personality, or even its facial animations**, allowing for a more personalized experience. This could make the assistant feel more relatable and engaging.

*6. Remote Therapy and Emergency Assistance*

SOULLINK could also offer integrations with **mental health professionals** and platforms, sending anonymized data (with user consent) to therapists for review. In extreme cases, if distress or emergency situations are detected (e.g., signs of severe depression), SOULLINK could notify guardians or connect users with helplines.

*7. Gamified Emotional Growth*

Another interesting future feature is **gamification**—rewarding users with virtual badges or points for engaging in mood-lifting activities or showing emotional progress over time. This can make self-care feel more interactive and fun.

**Final Thought**

SOULLINK, in its current form, lays the foundation for emotionally intelligent home companions. As emotional well-being becomes just as important as physical health, we believe SOULLINK can evolve into an essential daily companion—always ready to listen, respond, and help you feel just a little bit better, one day at a time.

**8. REFERENCES**

1. **Facial Sentiment Analysis Using AI Techniques: State-of-the-Art, Taxonomies, and Challenges (Published: 2022)**

This paper delves into advancements in machine and deep learning algorithms for facial sentiment analysis. It presents a taxonomy of existing strategies and reviews various models designed for facial expression recognition based on static images.

<https://ieeexplore.ieee.org/document/9091188>

**2. Sentiment Analysis on Voice Data Using Deep Learning (Published 2022)**

This study focuses on classifying human emotions through voice data.

The researchers employed deep learning techniques to categorize emotions such as neutral, calm, happy, sad, angry, nervous, and disgusted.

The paper highlights the significance of audio features like pitch, tone, and rhythm in determining sentiment.

<http://ijnrd.org/papers/IJNRD2204094.pdf>

**3.Robust Image Sentiment Analysis Using Progressively Trained and Domain Transferred Deep Networks(Published: 2015)**

This paper explores sentiment identification through facial expressions using deep convolutional neural networks (CNNs).

The authors highlight the challenges of capturing dynamic facial expressions and the importance of temporal information in accurately identifying emotions.

Authors: [**Quanzeng You**](https://www.researchgate.net/profile/Quanzeng-You?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFnZSI6InB1YmxpY2F0aW9uIn19) **,** [**Jiebo Luo**](https://www.researchgate.net/profile/Jiebo-Luo?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFnZSI6InB1YmxpY2F0aW9uIn19)

<https://www.researchgate.net/publication/282072489_Robust_Image_Sentiment_Analysis_Using_Progressively_Trained_and_Domain_Transferred_Deep_Networks>

**OTHER REFERENCES:**

[**https://pmc.ncbi.nlm.nih.gov/articles/PMC4845759/#abstract1**](https://pmc.ncbi.nlm.nih.gov/articles/PMC4845759/)

[**https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health**](https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health)

[**https://pmc.ncbi.nlm.nih.gov/articles/PMC4225959/#abstract1**](https://pmc.ncbi.nlm.nih.gov/articles/PMC4225959/)

[**https://pmc.ncbi.nlm.nih.gov/articles/PMC4409431/#sec1**](https://pmc.ncbi.nlm.nih.gov/articles/PMC4409431/)

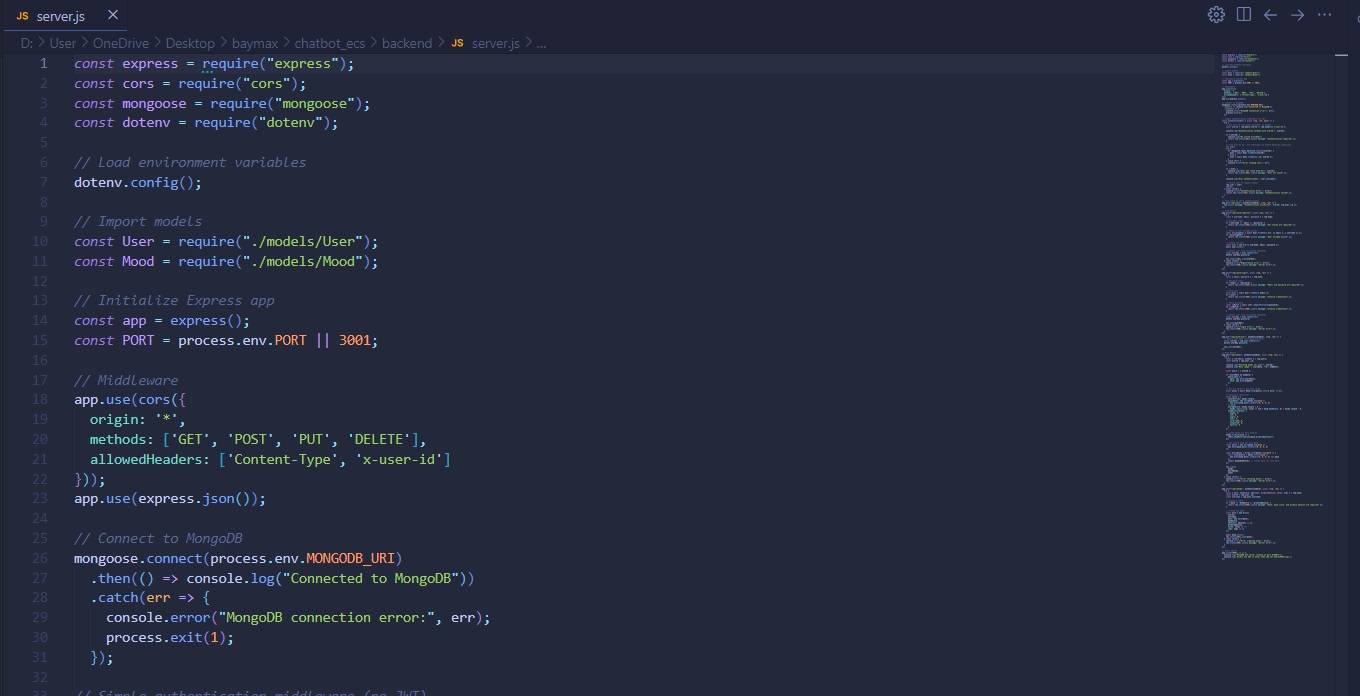
[**https://braininformatics.springeropen.com/articles/10.1186/s40708-023-00196-6#Abs1**](https://braininformatics.springeropen.com/articles/10.1186/s40708-023-00196-6)

[**https://ieeexplore.ieee.org/document/10141829**](https://ieeexplore.ieee.org/document/10141829)

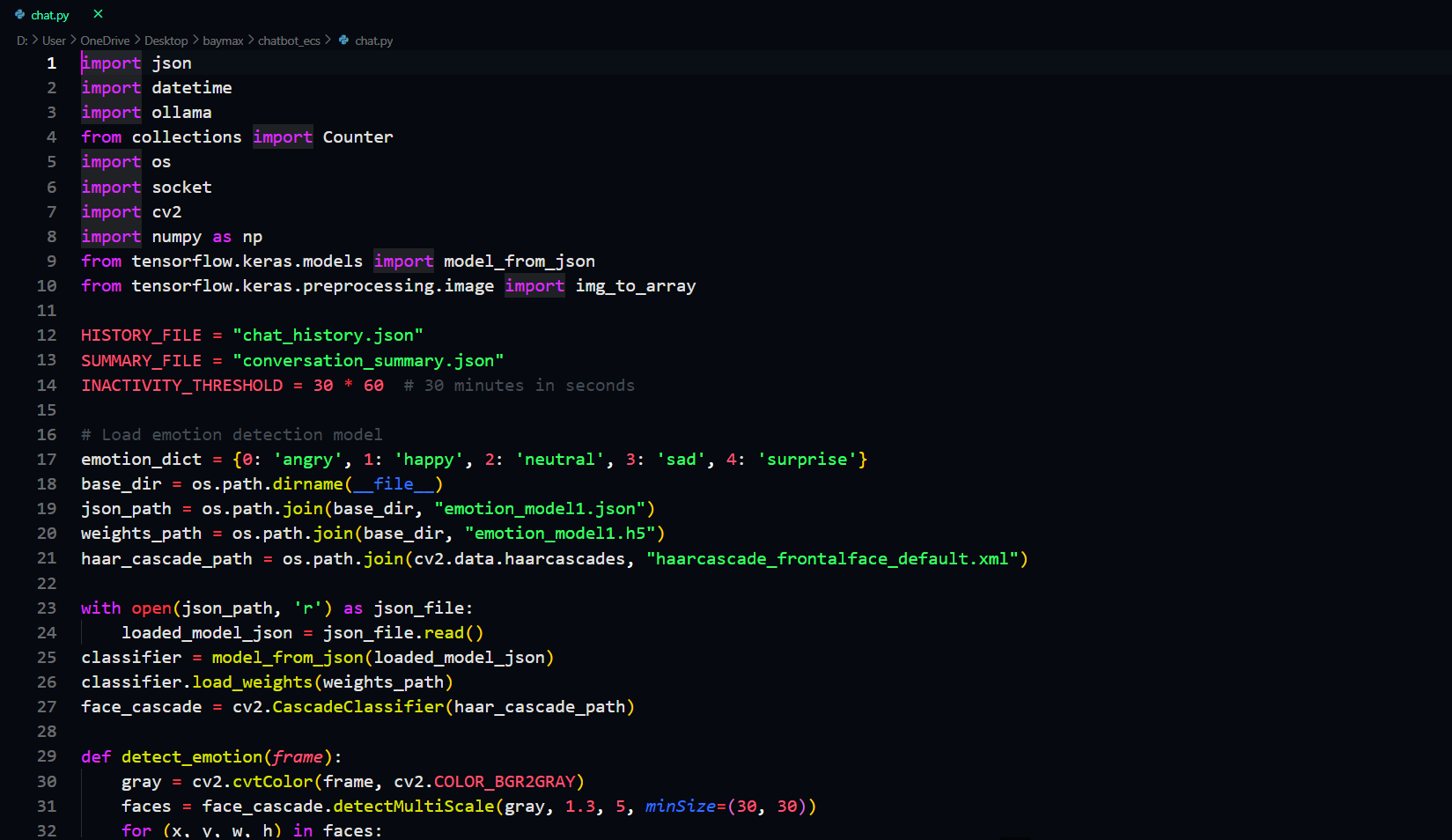
[**https://pmc.ncbi.nlm.nih.gov/articles/PMC9919280/**](https://pmc.ncbi.nlm.nih.gov/articles/PMC9919280/)

**9. CODES IN APPENDIX**

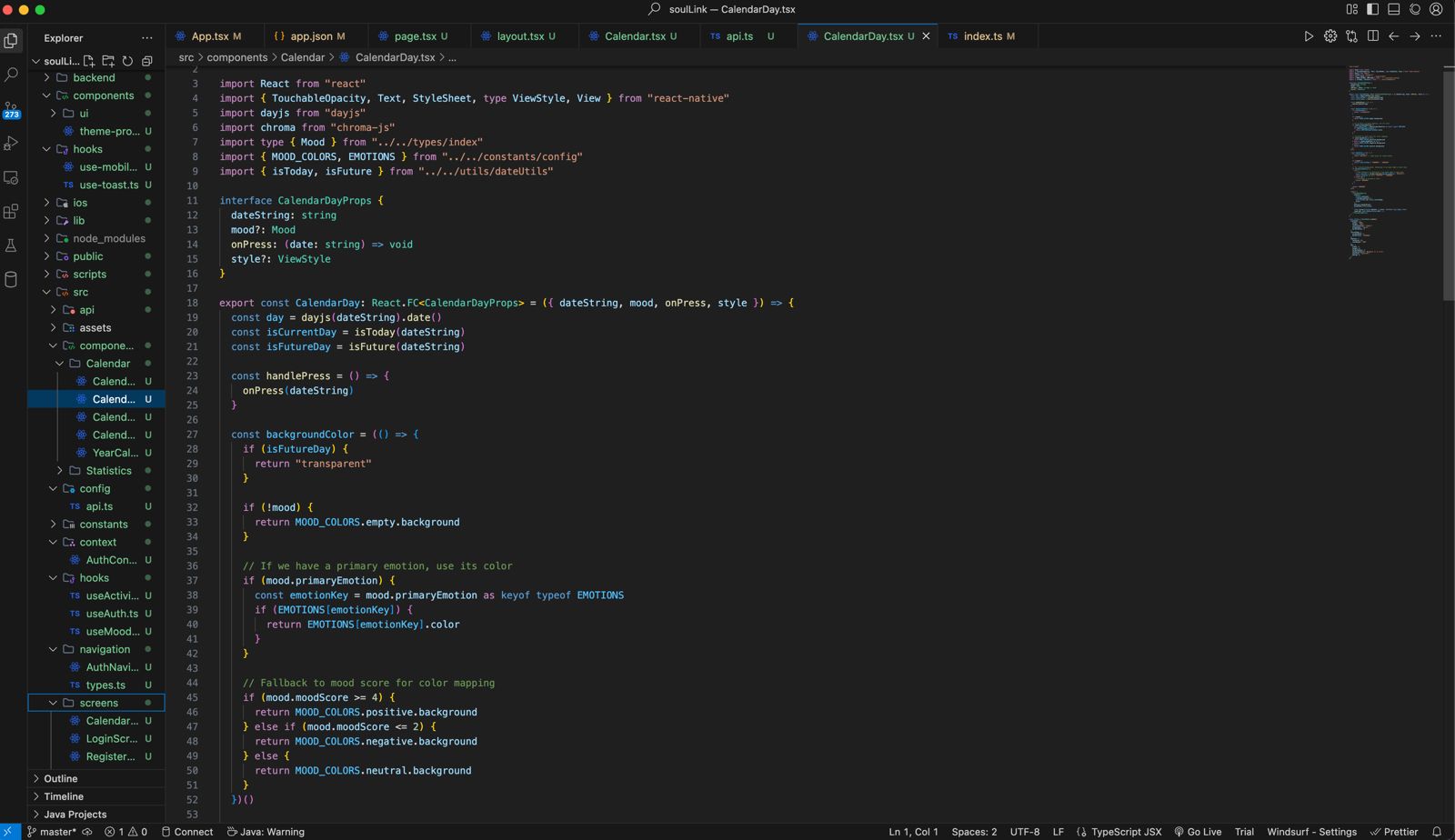
**MongoDB server**



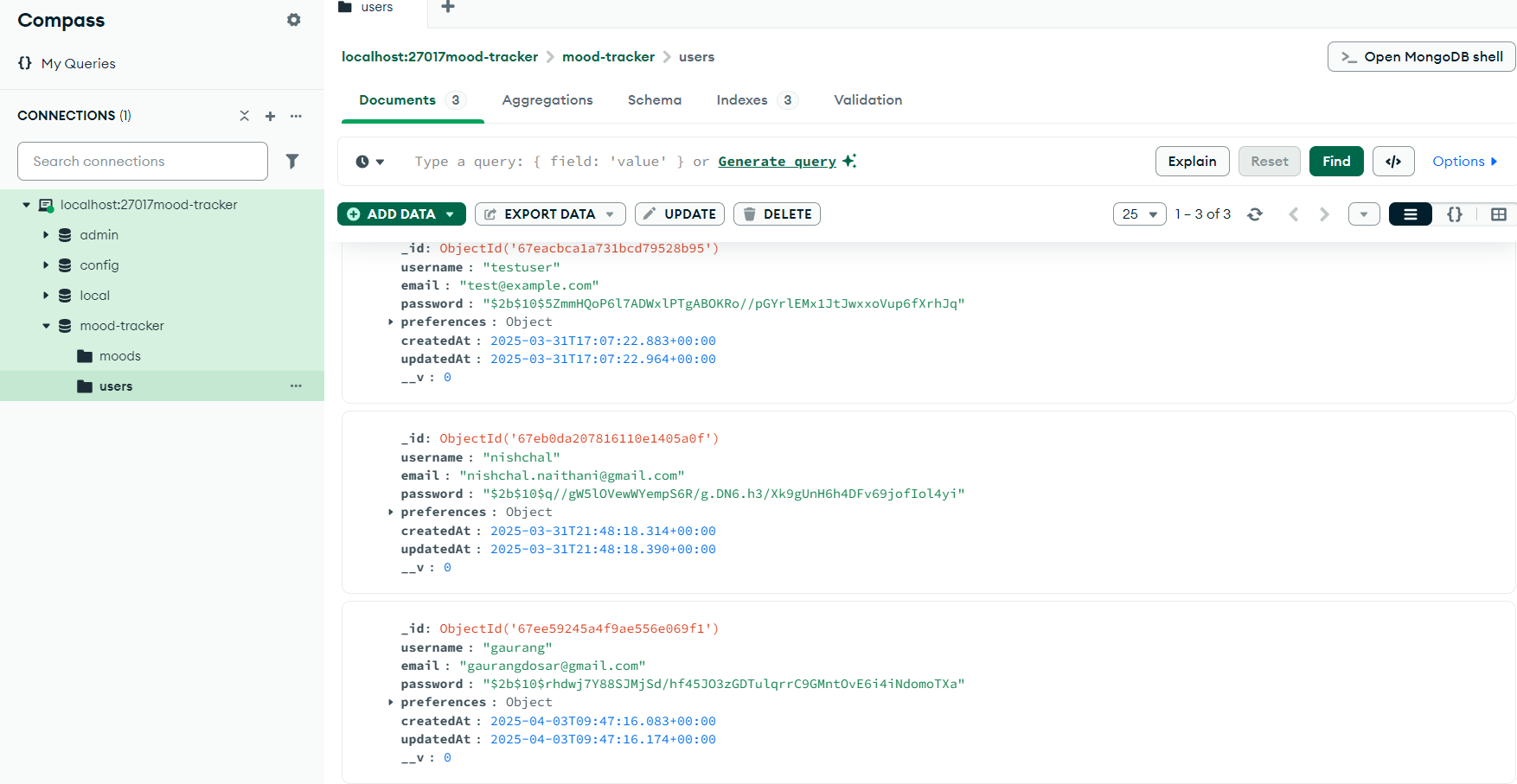
**Connection processing**



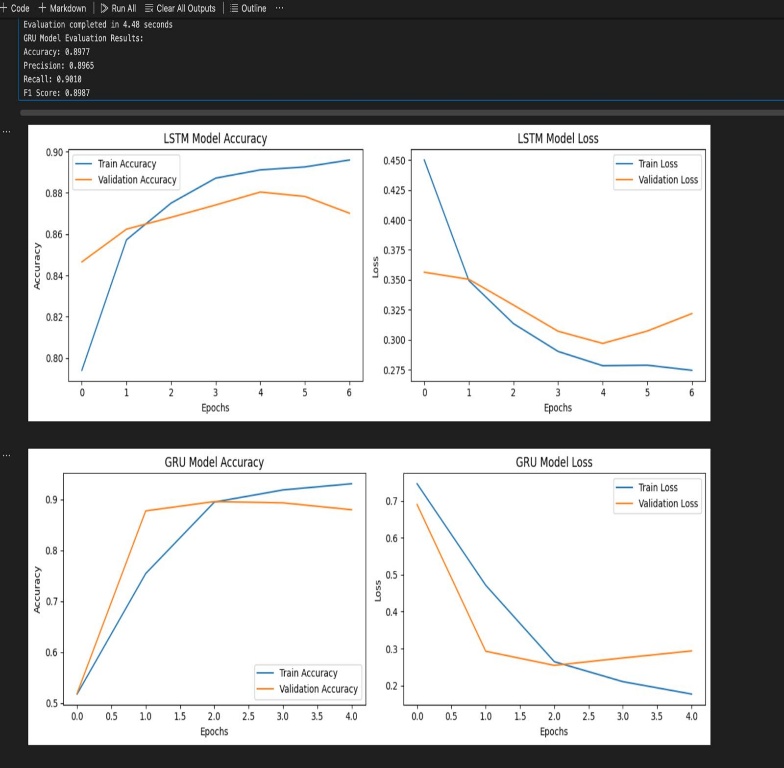
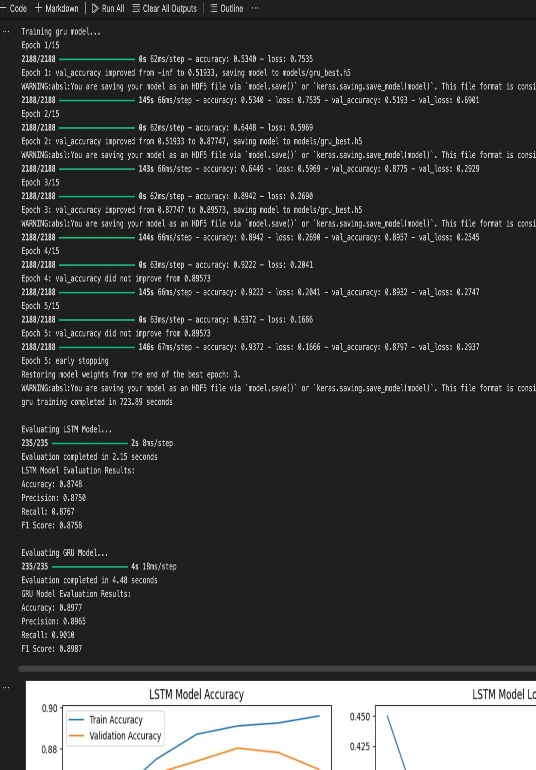
**React Native front-end code snippets**



**DATABASE**



**TRAINING MODELS:**



**End of Report**