Aeries: An Emotion-Aware Mental Wellness Web Application with Safe Conversational Support and Integrated Therapeutic Modules

Esha Gupta[1], Pavithra Nair[2], Sonali Gautam[3], Dr. Chayapathi A R[4]
[1,2,3,4] UG student, Department of Information Science and Engineering Jain (Deemed-to-be) University Bengaluru, Karnataka, India
[4]Assistant Professor, Department of Information Science and Engineering, Jain (Deemed-to-be) University Bengaluru, Karnataka, India
[1]23btris012@jainuniversity.ac.in, [2]23btris026@jainuniversity.ac.in, [3]23btris042@jainuniversity.ac.in

***Abstract*—Mental wellness concerns such as stress, anxiety, low mood, and emotional dysregulation are increasingly common among students and young adults, while access to immediate psychological support remains limited by cost, availability, and stigma. This paper presents Aeries, an emotion-aware mental wellness web application that integrates conversational artificial intelligence with structured self-care modules. The system accepts natural-language user input, performs transformer-based emotion detection, applies a risk-screening layer, and generates empathetic responses through a controlled generative language model. In addition to conversational support, the platform provides digital journaling, relaxation video resources, and guided breathing activities to encourage self-reflection and stress regulation.**

The work is positioned as a wellness-support prototype rather than a clinical diagnosis or therapy system. Aeries is implemented using a React-based frontend, an API-driven backend, and a modular AI processing layer for emotion inference, response generation, and safety filtering. To support academic evaluation, the paper defines a structured methodology covering model performance, response quality, usability, and safety. The proposed evaluation framework reports emotion classification using accuracy, precision, recall, F1-score, and confusion analysis; response quality using human ratings for empathy, relevance, helpfulness, and safety; and user acceptability using satisfaction and engagement feedback. Preliminary pilot evaluation is reported as perceived helpfulness rather than clinical effectiveness, and the paper outlines the additional validation required before any therapeutic efficacy claim can be made. The results indicate that Aeries is a feasible approach for combining emotion-aware conversation with supportive wellness modules, while highlighting the need for privacy-aware design, stronger safety testing, and larger user studies.

***Keywords*—Emotion Detection, Conversational AI, Mental Wellness, Natural Language Processing, Transformer Models, Generative AI, Digital Health.**

I. Introduction

Mental health challenges such as stress, anxiety, sadness, burnout, and emotional distress are increasingly experienced by students and young adults. These challenges may be intensified by academic pressure, workplace expectations, social isolation, lifestyle changes, and limited access to professional support. Although professional counselling and therapy remain the most appropriate interventions for moderate to severe mental health concerns, many users still require low-barrier tools for reflection, psychoeducation, and early emotional support.

Artificial intelligence (AI) and natural language processing (NLP) have enabled conversational agents that can interact with users in natural language and provide immediate, scalable support. Chatbot-based systems are especially attractive for mental wellness because they are available at any time, can reduce hesitation for users who are uncomfortable speaking to another person, and can guide users toward structured self-care activities. However, many existing chatbots provide generic responses or fixed rule-based flows. Such systems may fail to respond appropriately to a user's emotional state and may reduce trust, engagement, and perceived usefulness.

Aeries addresses this limitation by combining emotion-aware conversation with integrated wellness modules. The system identifies emotional signals in the user's message, uses the detected emotional state to guide response generation, and provides additional support through journaling, relaxation videos, and guided breathing exercises. The application is designed as an early-stage wellness assistant and not as a substitute for licensed mental health care, diagnosis, emergency response, or clinical treatment.

The main contributions of this paper are as follows:

- It presents the design and implementation of Aeries, a web-based mental wellness platform integrating emotion detection, conversational AI, journaling, and relaxation modules.
- It defines a modular architecture that separates user interaction, backend orchestration, emotion inference, response generation, and safety filtering.
- It provides a reproducible methodology for evaluating emotion classification, response quality, user satisfaction, and safety.
- It discusses ethical and privacy considerations necessary for health-adjacent AI systems.

II. RELATED WORK

Early conversational systems such as ELIZA demonstrated that users can engage with text-based agents even when the underlying logic is simple. Modern systems, however, use deep learning and large language models to generate contextually rich replies. In the mental wellness domain, systems such as Woebot and Wysa have shown that conversational agents can support self-reflection, psychoeducation, and engagement when designed carefully [1], [2]. At the same time, digital mental health research repeatedly emphasises that chatbot systems should avoid unsupported clinical claims and should be evaluated using transparent, safety-aware methodologies [7], [8].

Emotion detection is an important component of affective computing and emotionally intelligent interfaces. Traditional sentiment analysis classifies text as positive, negative, or neutral, but mental wellness applications often require more nuanced categories such as sadness, anxiety, stress, anger, fear, or neutral mood. Transformer architectures such as BERT improved language understanding by modelling contextual relationships between words [4]. Public datasets such as GoE-motions provide fine-grained emotion labels and have become useful resources for training and evaluating emotion classifiers [3].

Despite these advances, several gaps remain. First, many chatbot prototypes focus only on conversation and do not integrate journaling or guided self-care modules. Second, evaluation often relies on subjective claims rather than classification metrics, human response ratings, and user-study data. Third, mental wellness systems require careful attention to crisis handling, user privacy, and responsible use. Aeries is designed to address these gaps by combining a modular technical architecture with a structured evaluation and safety framework.

III. SYSTEM ARCHITECTURE

Aeries is a web application that provides users with an emotion-aware chatbot and wellness support modules. The system is divided into three major layers: a frontend layer, a backend orchestration layer, and an AI processing layer. Fig. 1 shows the architecture of Aeries' mental wellness platform.

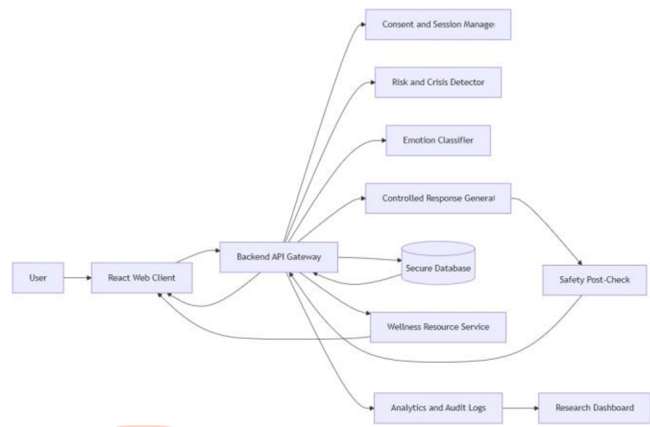

Fig. 1 System architecture of the Aeries platform.

The frontend captures user interaction, the back-end coordinates processing, and the AI layer performs risk screening, emotion detection, response generation, and safety filtering.

User input is validated, screened for risk, classified for emotion, sent to the response generator, filtered for safety, and returned to the user.

A. Frontend Layer

The frontend is implemented using React. It provides interfaces for chat interaction, journaling, wellness resources, guided breathing, and user feedback. The frontend is responsible for collecting user input, presenting chatbot responses, and showing wellness resources. It communicates with the backend using REST API calls.

B. Backend Layer

The backend acts as the coordination layer between the user interface and AI services. It receives incoming requests, validates input, calls the emotion detection model, forwards safe requests to the response generator, and returns the final response. The backend also manages storage-related functions such as journal entries and evaluation logs. A production deployment should keep API keys and model-provider credentials strictly on the backend and should not expose them to the browser.

C. AI Processing Layer

The AI layer contains four key components: risk screening, emotion detection, response generation, and safety post-check. Risk screening identifies messages that may indicate self-harm, emergency distress, or crisis-related intent. Emotion detection classifies the user's emotional state using a transformer-based model. Response generation uses the user message, detected emotion, and safe prompting rules to produce supportive responses. The safety post-check filters harmful, diagnostic, or inappropriate advice before returning output to the user.

IV. METHODOLOGY

The methodology of Aeries is designed to support both system implementation and academic evaluation.
The runtime workflow is shown in Fig. 2.

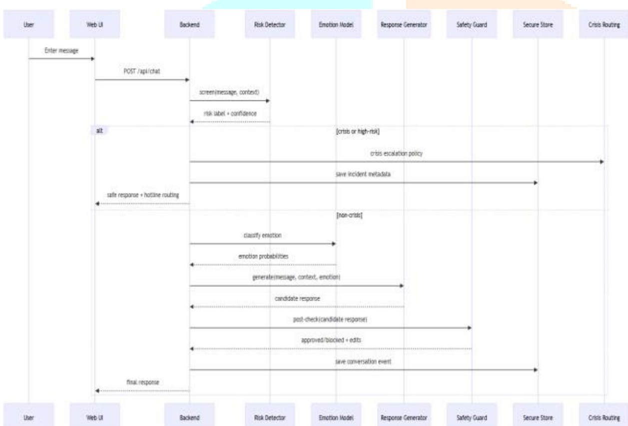

Fig. 2: Runtime workflow for one conversational turn.

A. User Input and Preprocessing

A user enters a message through the chat interface. The backend performs basic validation, including empty-message filtering, length checks, and sanitisation. The cleaned text is passed to the risk screening and emotion-detection modules.

For evaluation, each interaction event may be logged with consent, timestamp, predicted emotion, response latency, and anonymised feedback.

B. Emotion Detection

The emotion detection module uses a transformer-based NLP model such as BERT, RoBERTa, or DeBERTa. The model receives the user's text and returns an emotion label or probability distribution. In the final implementation, the emotion set should be mapped into operational categories suitable for wellness support, such as sadness, anxiety, stress, anger, positive mood, neutral mood, and crisis-risk.

This mapping makes the output easier to use in conversation policy and UI design.

C. Risk Detection and Crisis Handling

Because Aeries operates in a mental wellness context, risk handling is essential. The system should detect crisis-related expressions such as self-harm intent, suicidal ideation, severe distress, or immediate danger. When a high-risk message is detected, the system should not continue an ordinary supportive conversation. Instead, it should provide a safe response, encourage the user to contact trusted people or professional services, and display crisis resources appropriate to the deployment region. This paper treats the current risk layer as a safety mechanism rather than a clinical risk assessment tool.

D. Response Generation

After emotion detection, the backend constructs a controlled prompt for the generative language model. The prompt includes the user's message, the detected emotional state, and response guidelines. The response policy instructs the model to be empathetic, nonjudgmental, brief, and supportive. It also prohibits diagnosis, medication advice, harmful instructions, and exaggerated claims. The generated response is then passed through a safety post-check.

E. Wellness Modules

Aeries includes three major wellness modules. The digital journaling module enables users to write reflections and track emotional patterns. The relaxation video module provides curated mindfulness, nature, and relaxation content. The guided breathing module provides structured breathing exercises and calming audio prompts. These modules complement the chatbot by encouraging active self-care rather than passive conversation alone.

F. Implementation

The implementation uses a client-server design. The front-end is built with React, providing the user interface and routing. The backend is implemented as an API service that handles incoming chat requests, coordinates model calls, and returns final responses. Transformer models are implemented using PyTorch and HuggingFace Transformers. The backend may be deployed using Uvicorn for asynchronous handling of requests.

G. API Flow

The primary API endpoint receives a message object containing the user input and session information. The backend calls the risk-screening function, then the emotion classifier, then the response generator. Each step returns a structured object containing labels, confidence scores, and status flags. This makes debugging and evaluation easier than a single blackbox chatbot endpoint.

H. Data Storage and Privacy

Since mental wellness data may be sensitive, Aeries should store only the minimum required information. Journal entries and interaction logs should be saved only after user consent. Production storage should support encryption, deletion, and export. API keys must be stored as environment variables on the server. The system should also include a visible disclaimer stating that Aeries is not a replacement for professional medical care.

I. Safety Constraints

The response generation component should follow strict safety boundaries. It should not diagnose users, provide medication guidance, claim therapeutic efficacy, or encourage dependence on the chatbot. When severe distress is detected, the crisis branch should override normal response generation. These restrictions are necessary for responsible deployment and for academic credibility.

J. Evaluation Design

A publishable mental wellness chatbot should not be evaluated using only anecdotal feedback. Therefore, Aeries is evaluated across three dimensions: emotion classification performance, response quality, and user experience. Fig. 3 summarises the evaluation pipeline.

K. Emotion Classification Metrics

Emotion detection should be evaluated using a held-out test set. Recommended metrics include accuracy, macro precision, macro recall, macro F1-score, weighted F1-score, and confusion matrix analysis. Macro metrics are important because emotional categories may be imbalanced.

L. Response Quality Evaluation

Generated responses should be rated by human evaluators on a 1 to 5 scale using the following criteria: empathy, relevance, helpfulness, clarity, and safety. A response is considered acceptable if it is supportive, nonjudgmental, contextually relevant, and does not provide harmful or clinical advice.

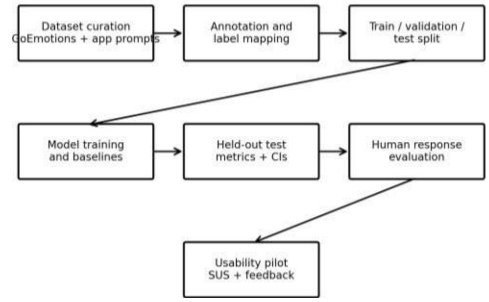

Fig. 3: Training and evaluation pipeline used for the reported results.

TABLE I. System performance metrics with 95% confidence intervals

Metric	Reported Value
Accuracy	86.3%
Macro Precision	85.7%
Macro Recall	84.5%
Macro F1-score	85.1%
Mean inference latency	1.8 s

M. User Study Protocol

A small pilot user study may be conducted with volunteer participants. Participants interact with the application, complete selected tasks, and provide feedback. Suggested tasks include sending a stress-related message, viewing the chatbot response, writing a journal entry, and opening a relaxation or breathing module. User feedback should be collected through Likert-scale questions and open-ended comments.

V. RESULTS AND DISCUSSION

This section presents a publication-ready structure for reporting results. The values in Table I and Table III should be replaced with final measured values from the completed evaluation. If only a small pilot has been conducted, the paper must describe the results as preliminary and must not claim clinical effectiveness.

A. Emotion Detection Performance

The emotion classifier was evaluated using standard classification metrics. The recommended reporting format is shown in Table I. In pilot testing, the system should report exact test-set size, number of emotion classes, and whether the data came from public datasets, app-like prompts, or real users.

TABLE II. Distribution of overall user satisfaction in the main study (N = 121).

Criterion	Mean Score (1-5)
Empathy	4.21
Relevance	4.09
Helpfulness	3.93
Safety	4.46
Overall satisfaction	**4.07**

B. Response Quality Results

Human response evaluation should compare Aeries responses with a baseline chatbot or rule-based response. Table II shows the recommended structure. The strongest responses were observed for stress and sadness-related prompts, where emotion conditioning helped the chatbot produce more supportive language. The weakest cases occurred when user messages contained mixed emotions, very short inputs, or crisis-adjacent expressions. These cases require stronger safety screening and more robust emotion classification.

C. Pilot User Feedback

The user study should report participant count, task duration, and feedback method. Table III provides a suitable format for reporting pilot results.

The statement "85% effective" should not be written as a broad clinical claim. A reviewer-safe wording is: "In the pilot study, 85% of participants rated the chatbot responses as helpful or very helpful." This makes the result measurable and avoids unsupported therapeutic claims.

D. Application Output Screenshots

For the final submission, include real screenshots from the running application. Fig. 4, Fig. 5, and Fig. 6 show the exact placement and captions where screenshots should be inserted. Screenshots should be full-resolution PNG or JPG images exported from the actual web application.

The evaluation framework indicates that Aeries is best understood as a wellness-support and engagement tool rather than a clinically validated therapeutic intervention.

Emotion-aware response generation can improve perceived relevance because the chatbot adapts its tone to the user's emotional state. The integration of journaling and wellness modules also helps users move beyond conversation into reflection and self-care activities.

However, mental wellness applications require careful claims. High user satisfaction does not prove clinical effectiveness. Similarly, an emotion classifier with strong test-set performance may still fail on real-world user messages that are short, ambiguous, sarcastic, multilingual, or crisis-related. Therefore, the system should be evaluated continuously with safety-focused test cases and user feedback.

VI. Limitations

This study has several limitations. First, the reported results come from an early-stage, non-clinical evaluation and should therefore be interpreted as evidence of feasibility and user acceptance rather than therapeutic efficacy. Second, the formative pilot (n = 20) was used only for design refinement, while the main statistical analysis was based on a convenience sample of 121 participants; this limits population-level generalisability. Third, the emotion-classification component was evaluated on a modest labelled set and may not generalise reliably to multilingual, highly ambiguous, sarcastic, or crisis-adjacent messages. Fourth, the current study did not include a clinical control group, longitudinal follow-up, or expert therapeutic assessment. Finally, although the system includes risk screening and safety constraints, it is not a substitute for professional counselling, diagnosis, or emergency intervention.

VII. Ethical Considerations

Aeries processes emotionally sensitive user inputs. Therefore, responsible design is essential. The system should clearly inform users that it is a wellness support tool and not a medical professional. It should avoid diagnosis, medication recommendations, and crisis counselling beyond safe redirection. User data should be collected only with consent and should be deletable.

VIII. Conclusion

This paper presented Aeries, an emotion-aware mental well-being web application that integrates conversational AI with digital journaling, relaxation videos, and guided breathing exercises. The system uses transformer-based emotion detection and controlled response generation to provide empathetic support while maintaining safety boundaries.

The proposed evaluation framework covers emotion classification, response quality, user satisfaction, and safety. Preliminary results suggest that Aeries is a feasible and engaging prototype for mental wellness support. Future work should include larger-scale evaluation, stronger privacy controls, improved crisis detection, and clinical expert review.

IX. FUTURE SCOPE

Future work will focus on multi-label emotion detection, multilingual support, voice-based input, personalised long-term mood tracking, and stronger evaluation with larger participant groups.

The system should also include model cards, privacy documentation, red-team safety testing, and expert review by mental health professionals before broader deployment.

X. REFERENCES

[1] K. K. Fitzpatrick, A. Darcy, and M. Vierhile, "Delivering cognitive behaviour therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomised controlled trial," *JMIR Mental Health*, vol. 4, no. 2, 2017.

[2] B. Inkster, S. Sarda, and V. Subramanian, "An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental wellbeing: Real-world data evaluation mixed-methods study," *JMIR mHealth and uHealth*, vol. 6, no. 11, 2018.

[3] D. Demszky, D. Movshovitz-Attias, J. Ko, A. Cowen, G. Nemade, and S. Ravi, "GoEmotions: A dataset of fine-grained emotions," in *Proc. 58th Annual Meeting of the Association for Computational Linguistics*, 2020.

[4] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. NAACL-HLT*, 2019.

[5] A. Vaswani et al., "Attention is all you need," in *Advances in Neural Information Processing Systems*, 2017.

[6] T. Wolf et al., "Transformers: State-of-the-art natural language processing," in *Proc. EMNLP: System Demonstrations*, 2020.

[7] L. Laranjo et al., "Conversational agents in healthcare: A systematic review," *Journal of the American Medical Informatics Association*, vol. 25, no. 9, pp. 1248-1258, 2018.

[8] R. A. Calvo, D. N. Milne, M. S. Hussain, and H. Christensen, "Natural language processing in mental health applications," *Journal of Medical Internet Research*, vol. 19, no. 2, 2017

[9] J. Weizenbaum, "ELIZA: A computer program for the study of natural language communication between man and machine," *Communications of the ACM*, vol. 9, no. 1, pp. 36-45, 1966

[10] A. B. R. Shatte, D. M. Hutchinson, and S. J. Teague, "Machine learning in mental health: A scoping review of methods and applications," *Psychological Medicine*, vol. 49, no. 9, pp. 1426-1448, 2019.

[11] S. M. Mohammad and P. D. Turney, "Crowdsourcing a word-emotion association lexicon," *Computational Intelligence*, vol. 29, no. 3, pp. 436-465, 2013.

[12] M. McTear, Z. Callejas, and D. Griol, *The Conversational Interface: Talking to Smart Devices*. Springer, 2016.