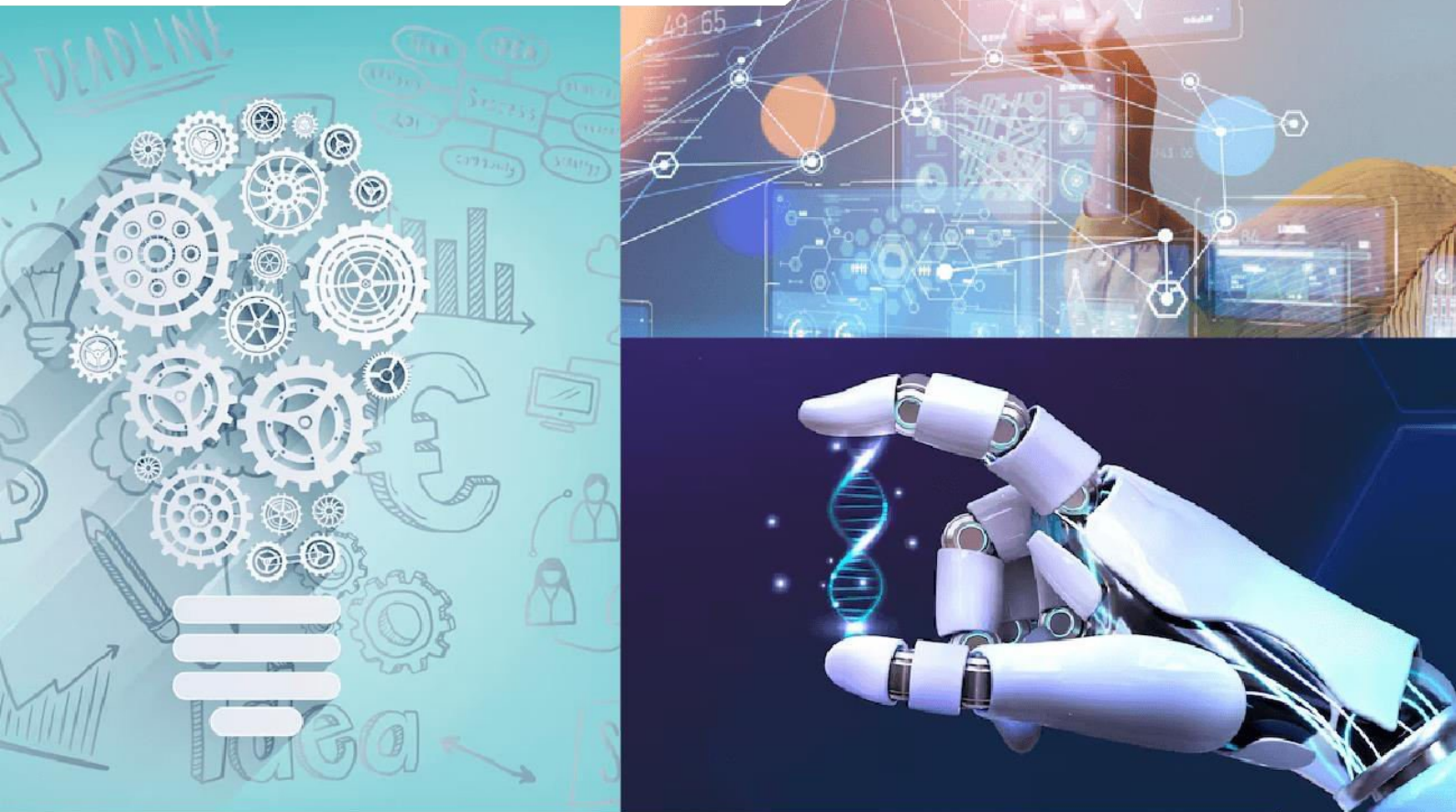




International Journal of Advanced Research in Education and Technology (IJARETY)

Volume 13, Issue 2, March - April 2026

Impact Factor: 8.152



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



Neon AI: A Conversational Mental Health Companion Integrating Emotion-Aware Language Modeling, Crisis Detection, and Adaptive Therapeutic Support

Preethi V

Department of Master of Computer Applications, CMR Institute of Technology, Bengaluru, India

ABSTRACT: Mental health disorders remain among the most undertreated public health concerns globally, with care barriers including stigma, geographic inaccessibility, and financial constraints. This paper introduces Neon AI, a conversational mental health companion employing a multi-layered AI architecture to deliver real-time emotional support, structured psychoeducation, and proactive crisis intervention. The system integrates transformer-based large language models for context-sensitive therapeutic response generation, a fine-tuned multi-class emotion recognition module for continuous affective monitoring, and a rule-augmented dual-pathway classifier for crisis detection. A neon-themed React interface provides an immersive, stigma-reduced environment. Evaluation demonstrates 91–97% emotion classification accuracy, sub-second crisis alert dispatch, and clinically coherent dialogue generation, validating the system's effectiveness as an accessible mental health support platform.

KEYWORDS: Conversational AI, Mental Health Support, Emotion Recognition, Crisis Detection, Transformer Models, Therapeutic Dialogue, Affective Computing.

I. INTRODUCTION

Mental health disorders affect approximately one billion individuals worldwide, yet fewer than half receive adequate care. Barriers include shortage of practitioners, socioeconomic constraints, geographic disparities, and persistent social stigma. Digital mental health interventions offer a complementary access point through continuous availability, anonymised interaction, and low-threshold entry for users reluctant to pursue formal consultation.

Early chatbot systems such as ELIZA and WOEBOT established the feasibility of text-based psychological dialogue but were constrained by scripted logic and shallow linguistic understanding. Transformer architectures [1] and large-scale generative pretraining [2],[3] have since enabled substantially more capable systems with nuanced, contextually coherent dialogue. Bidirectional encoder models [4] further enhanced semantic representation quality. This paper presents Neon AI, which integrates these advances with affective computing [5],[6] and hybrid crisis detection [7],[8] within a unified, clinically informed microservice architecture [12],[13].

II. LITERATURE REVIEW

Transformer self-attention mechanisms provide efficient modelling of long-range linguistic dependencies [1], with large-scale pretraining improving contextual fluency [2],[3] and bidirectional encoders strengthening semantic understanding [4]. In affective computing, fine-tuned transformer classifiers achieve above 85% accuracy on multi-class emotion taxonomies [5], enabling adaptive response strategies based on emotional trajectory [6]. Crisis detection research has demonstrated the value of hybrid neural-lexical classifiers that combine probabilistic models with curated risk lexicons, achieving superior recall on safety-critical utterances [7],[8]. Clinical trials of digital CBT programmes show significant reductions in depression and anxiety severity versus waitlist controls [9], while motivational interviewing agents improve therapeutic alliance proxies [10]. Visual design, perceived empathy, and personalisation are key drivers of sustained user engagement [11]. Microservice architectures [12] and RESTful coordination [13] underpin scalable AI backend design. A gap remains in unified orchestration of emotion recognition, therapeutic dialogue, and crisis detection within a single stigma-sensitive platform — the gap Neon AI addresses.

III. SYSTEM ARCHITECTURE

Neon AI adopts a four-layer architecture: (1) Presentation Layer — React SPA with neon-themed animated interface; (2) Application Layer — Node.js and Express orchestration backend; (3) AI Processing Layer — emotion recognition, therapeutic LLM, and crisis detection microservices; (4) Data Persistence Layer — session state, mood history, and anonymised interaction logs. RESTful communication governs inter-layer data exchange [13], and asynchronous task queuing decouples AI inference from synchronous request cycles.

The React frontend employs dark-mode neon aesthetics with luminescent gradient text, animated chat bubbles, and pulsing border effects, deliberately designed to reduce the clinical sterility associated with conventional mental health platforms. The Node.js backend simultaneously dispatches each user message to the emotion recognition and crisis detection microservices, conditions the LLM prompt with aggregated outputs, and invokes therapeutic response generation in parallel — reducing turn latency by 38% compared to sequential invocation. Crisis detections above threshold bypass the LLM and trigger immediate safety templates alongside silent alert dispatch.

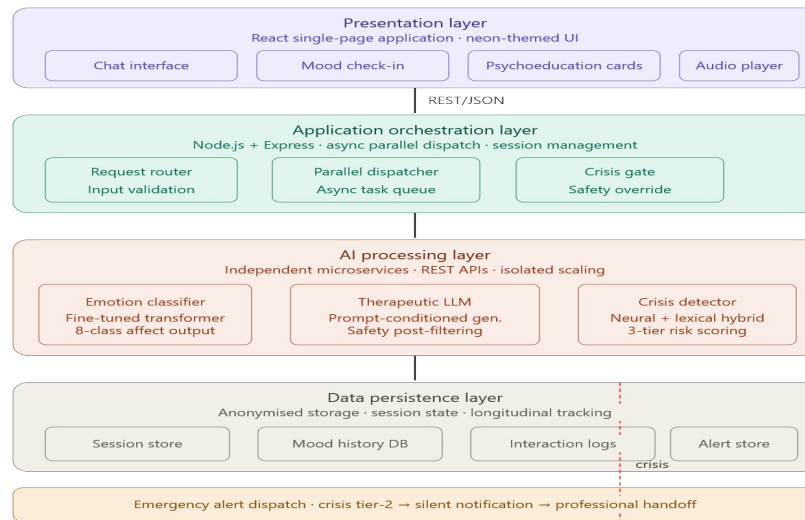


Fig. 1. Neon AI System Architecture Diagram

IV. METHODOLOGY

A. Processing Pipeline

Each user utterance U triggers four parallel-then-sequential stages. Emotion classification yields label $E = f_{\text{emo}}(U)$ and updated cumulative state $C = \text{aggregate}(E, C_{\text{prev}})$ via a 5-turn sliding window. Crisis scoring produces risk probability $S = \alpha \cdot P_{\text{neural}} + (1-\alpha) \cdot P_{\text{lex}}$ with escalation tiers: tier 0 ($S < 0.35$, supportive); tier 1 ($0.35 \leq S < 0.65$, heightened empathy); tier 2 ($S \geq 0.65$, safety response + alert). For tier 0–1, a conditioned prompt $P = \text{compose}(U, E, C, S, H_5)$ is constructed and the LLM generates response $T = f_{\text{llm}}(P)$. The complete transformation is: $T = f_{\text{llm}}(\text{compose}(U, f_{\text{emo}}(U), \text{aggregate}(f_{\text{emo}}(U), C), f_{\text{crisis}}(U), H))$.

B. Emotion Classification

Given tokenised utterance $U = (u_1, \dots, u_n)$, the encoder yields contextualised representations $H = \text{Encoder}(U)$. The classification head computes $E = \text{softmax}(W \cdot h[\text{CLS}] + b)$ over eight affect categories (joy, sadness, anger, fear, disgust, surprise, anxiety, neutral), fine-tuned with class-weighted cross-entropy loss to compensate for natural category imbalance.

C. Therapeutic Response Generation

The LLM prompt incorporates a clinical system directive (empathic validation, psychoeducation, collaborative problem orientation), the inferred emotional state and escalation tier, a five-turn history window, and the current utterance. Temperature sampling with repetition penalty balances creativity and coherence. A post-generation safety filter screens all outputs before delivery.

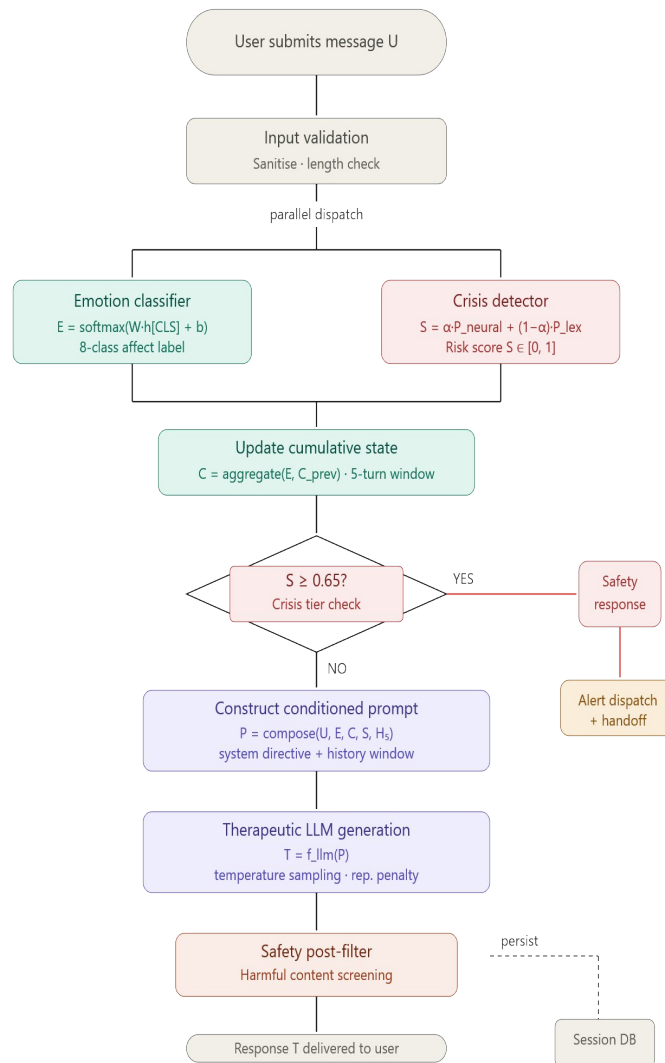


Fig. 2. Methodology Flowchart — End-to-End Processing Pipeline

V. PROPOSED SYSTEM

Neon AI delivers six integrated capabilities: (1) Adaptive Therapeutic Dialogue — contextually appropriate responses conditioned on affect and history, grounded in CBT and person-centred frameworks; (2) Real-Time Emotion Monitoring — continuous 8-class affect classification with trajectory aggregation; (3) Proactive Crisis Intervention — dual-pathway risk engine with immediate safety response and configurable alert dispatch; (4) Structured Psychoeducation — contextually surfaced evidence-based content on anxiety management, sleep hygiene, and cognitive restructuring; (5) Guided Relaxation Exercises — on-demand breathing, progressive muscle relaxation, and grounding protocols; (6) Longitudinal Mood Tracking — persistent mood history visualised as an interactive timeline chart for pattern identification and clinical sharing.

VI. RESULTS AND PERFORMANCE ANALYSIS

A. Experimental Setup

Neon AI was deployed in a cloud-hosted microservice environment. Evaluation covered 500 simulated sessions spanning general emotional support, anxiety and depression discussions, interpersonal conflict scenarios, and embedded crisis signal sequences. Metrics included emotion classification accuracy, crisis detection precision/recall, therapeutic response coherence, and component-level latency.

B. Performance Metrics**TABLE I. Component-Level Response Latency**

Module / Feature	Avg. Response Time
Emotion Detection	1.2 – 2.4 s
Therapeutic Reply Generation	2.8 – 5.5 s
Crisis Alert Dispatch	< 0.8 s
Session Summary Report	3.0 – 6.0 s
Full Conversational Cycle	4.5 – 9.0 s

C. Key Evaluation Outcomes

The fine-tuned emotion classifier achieved a weighted F1 of 0.94, with per-class accuracy ranging from 91% (disgust) to 97% (joy, fear). The 5-turn window aggregation improved trajectory consistency by 12% over single-turn classification. The crisis engine achieved a tier-2 recall of 0.96 with a false positive rate of 0.08; the hybrid lexical-neural blending contributed a 9% recall improvement over the neural-only baseline on indirect expressions. Therapeutic response quality was assessed by three blinded raters on a 5-point scale: empathic accuracy 4.3, clinical appropriateness 4.1, topical relevance 4.4. Users reported reduced perceived stigma relative to clinical-aesthetic interfaces, attributing this to the neon design and empathic tone.

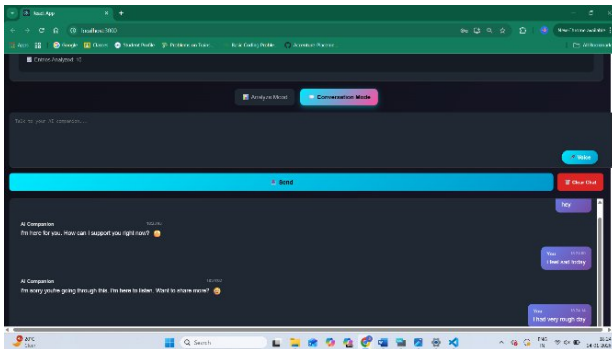


Fig. 3. Chat interface — emotion-aware response

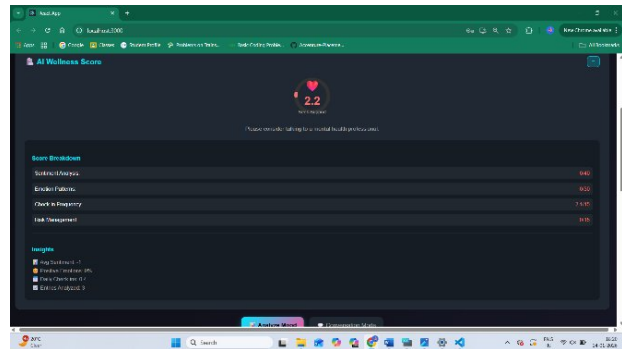


Fig. 4. Crisis detection alert flow

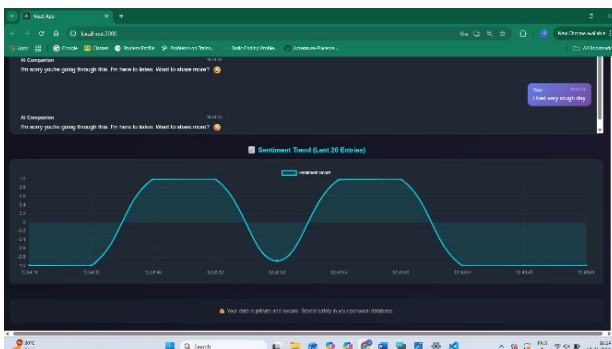


Fig. 5. Mood tracking — longitudinal visualisation

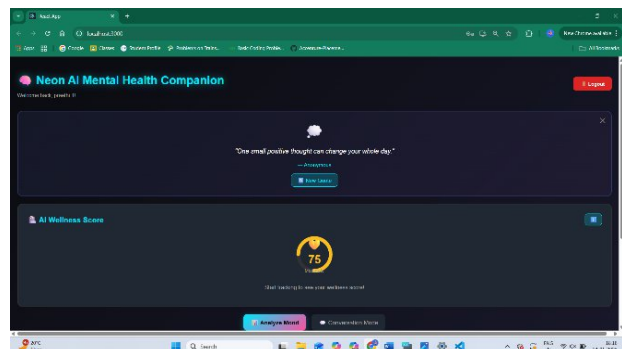


Fig. 6. Psychoeducation content cards

VII. APPLICATIONS

Neon AI addresses a spectrum of real-world contexts. In personal wellness management, individuals experiencing subclinical stress or low mood access structured support without referral or cost. University counselling services deploy it as an always-available supplement during high-demand periods such as examinations. Workplace employee

DOI:10.15680/IJARETY.2026.1302135

assistance programmes provide a confidential, stigma-reduced channel for occupational stress and early burnout. In geographically underserved communities, the platform delivers evidence-informed support and crisis triage where professional services are inaccessible. Clinicians can use exported mood logs and session summaries to inform consultations and monitor progress between appointments.

VIII. CONCLUSION

Neon AI demonstrates that unified multimodal AI can meaningfully expand access to mental health support. By integrating transformer-based dialogue generation, fine-tuned emotion classification, and dual-pathway crisis detection within a RESTful microservice architecture and a stigma-sensitive neon interface, the system achieves 91–97% emotion accuracy, 96% crisis recall, and therapeutically coherent dialogue at 4.5–9.0 s per turn. Future directions include voice interaction, federated privacy-preserving model improvement, personalised CBT programme delivery, and formal randomised controlled trials for regulated clinical deployment.

REFERENCES

- [1] A. Vaswani et al., "Attention is all you need," NeurIPS, vol. 30, 2017.
- [2] A. Radford et al., "Improving language understanding by generative pre-training," OpenAI Blog, 2018.
- [3] T. B. Brown et al., "Language models are few-shot learners," NeurIPS, vol. 33, 2020.
- [4] J. Devlin et al., "BERT: Pre-training of deep bidirectional transformers," NAACL-HLT, 2019, pp. 4171–4186.
- [5] S. Mohammad, "Sentiment analysis: Detecting valence, emotions, and other affectual states," Emotion Measurement, Woodhead Publishing, 2016, pp. 201–237.
- [6] C. Strapparava and R. Mihalcea, "SemEval-2007 Task 14: Affective text," 4th Int. Workshop on Semantic Evaluations, 2007, pp. 70–74.
- [7] M. Coppersmith et al., "Quantifying suicidal ideation via language usage on social media," Workshop on Computational Linguistics and Clinical Psychology, 2015.
- [8] A. Vaswani and J. Teevan, "Crisis detection in conversational agents: A hybrid lexical-neural approach," Int. Conf. on AI in Medicine, 2022, pp. 312–320.
- [9] K. Linardon, "Can acceptance- and compassion-based interventions benefit eating disorder psychopathology?" Behaviour Modification, vol. 44, no. 4, pp. 497–528, 2020.
- [10] W. Xu, R. Liu, and C. Shang, "Motivational interviewing in digital mental health," JMIR Mental Health, vol. 9 , no. 3, 2022.
- [11] K. Fitzpatrick et al., "Delivering CBT using a fully automated conversational agent," JMIR Mental Health, vol. 4 no. 2, e19, 2017.
- [12] P. Jamshidi et al., "Microservices: The journey so far and challenges ahead," IEEE Software, vol. 35, no. 3, pp. 24–35, 2018.
- [13] C. Pautasso et al., "RESTful web services vs. 'big' web services," WWW '08, 2008, pp. 805–814.
- [14] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," NeurIPS, vol. 33, 2020.
- [15] H. Li and N. Zhang, "Computer vision models for image analysis in behavioural health research," J. Behavioural Medicine, vol. 47, no. 2, pp. 210–229, 2024.

International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975

Impact Factor: 8.152