LUMINA

Your beacon of hope and support

Kaleemunnisa[1], Snehanjani Pati[2], Pravallika Challa[3], Pavan Ganeshreddy Yeruva[4], Sravya Vemireddy[5], and Sireesha Gangara
Pace University, Seidenberg School of CSIS, One Pace Plaza, New York, NY 10038, USA
{klnu, sp16045n, pc94436n, py40786n, sv65573n, sg48113n}@pace.edu

***Abstract*—Mental health disorders such as anxiety, depression, and ADHD are rising globally due to increasing social isolation, academic and work-related stress, and limited access to timely care. Traditional therapy remains inaccessible for many due to high costs and long waiting periods. This paper presents *LUMINA*, a real-time, AI-assisted mental health support platform that connects users to trained volunteers and peers through voice or chat interfaces. LUMINA integrates WebRTC for real-time communication, Firebase for backend services, and an ensemble of BiLSTM and GPT-2 models for mental health classification using NLP-based emotional tone analysis. The system escalates high-risk cases to human moderators or emergency services when necessary. A comprehensive literature review explores existing AI-driven mental health prediction models, chatbot-based interventions, ICF-based community support systems, and robotics in ADHD therapy. This study highlights the potential of hybrid AI-human models in improving the accessibility, scalability, and responsiveness of digital mental health care while addressing challenges related to privacy, ethics, and long-term engagement.**

***Index Terms*—Mental Health Support, Artificial Intelligence (AI), Natural Language Processing(NLP), Machine Learning(ML), Ensemble Learning, BiLSTM, GPT-2, Real-Time Communication, Digital Health Platforms.**

I. Introduction

Mental health disorders, including stress, anxiety, and depression, are becoming increasingly prevalent worldwide due to work pressures, personal struggles, and social isolation. Despite the growing awareness of mental health issues, traditional therapy remains expensive and inaccessible to many individuals, with long waiting times for appointments. People often require immediate emotional support, but existing mental health hotlines and professional services may not always be available when needed. To address this challenge, the Live Mental Health Support Network is designed as a real-time, accessible, and empathetic mental health platform that enables individuals to connect instantly with trained volunteers and peers via text chat or voice calls, eliminating delays in accessing support. The platform is built to be free or low-cost, ensuring accessibility regardless of financial background. Additionally, AI-driven monitoring is incorporated to detect harmful content or high-risk situations, triggering escalation to human moderators or emergency services when necessary.

Identify applicable funding agency here. If none, delete this.

A. Key Technological Aspects of the Solution

The platform utilizes cutting-edge technologies to deliver real-time support:

- Front-end: React.js and React Native for seamless user interaction.
- Backend: Firebase for real-time database and authentication.
- Real-Time Communication: WebRTC for voice and video calls, Socket.IO for instant messaging.
- AI Moderation: NLP models including BiLSTM and GPT-2 , trained using PyTorch and Transformers.

B. Impact of the Solution

- Immediate, low-cost mental health support that is available anytime.
- Anonymous interactions to reduce stigma and encourage more users to seek help.
- AI-driven crisis prevention, ensuring timely intervention for high-risk users.
- A supportive community that fosters long-term mental well-being.

C. Scope of the Literature Review

This literature review examines existing AI-driven mental health models and community-based intervention systems to evaluate the feasibility of this approach. It explores: 1. AI-Driven Mental Health Prediction Models – Utilizing machine learning, sentiment analysis, and mobile interventions for identifying and predicting mental health conditions. 2. ICF-Based Community Mental Health Systems – Implementing digital frameworks for community rehabilitation, peer support, and crisis management. 3. Challenges in Scalability, Security, and Ethical Considerations – Addressing data privacy, AI bias, and sustainable user engagement in digital mental health platforms. By analyzing prior research, this study aims to enhance digital mental health support by integrating AI-driven intervention and real-time community engagement, ensuring that mental health resources are accessible, efficient, and secure for all individuals in need.

II. Literature Review

Mental health disorders, including depression, anxiety, and ADHD, have seen a sharp rise in recent years, necessitating

technological interventions that provide early detection, real-time support, and AI-driven mental health solutions. Existing research explores the role of machine learning, AI-based predictive models, community-driven support systems, and real-time interactive technologies in addressing mental health challenges.

This literature review examines four major approaches:

1) AI-Driven Mental Health Prediction Models – Utilizing machine learning (ML), natural language processing (NLP), and AI models for early detection and predictive analysis of mental health conditions.
2) Chatbot and AI-Assisted Mental Health Platforms – Investigating AI-based conversational agents for mental health support, including their effectiveness in delivering Cognitive Behavioral Therapy (CBT) and crisis intervention.
3) ICF-Based Community Mental Health Systems – Implementing digital frameworks for information sharing among healthcare providers to improve rehabilitation and crisis prevention.
4) AI and Robotics in ADHD Therapy – Evaluating robotic assistants powered by Large Language Models (LLMs) for personalized ADHD treatment and cognitive support.

A. AI-Based Mental Health Prediction Models (Revised & Enhanced)

A notable contribution by Choudhary et al. introduced **SentiSense**, an AI-driven application that uses Natural Language Processing (NLP) to assess user sentiment in real-time chats and conversations [1]. The system classifies emotional tone and provides timely feedback to users based on emotional fluctuation, promoting self-awareness and early intervention. Though the results were promising, the authors emphasized the need for validation with broader datasets and long-term user engagement.

In a separate study, Smith et al. developed machine learning models using Support Vector Machines (SVMs) and Random Forest classifiers to predict symptoms of depression and anxiety [2]. These models achieved up to 80% accuracy in classification tasks, but their generalizability was limited by small dataset size and lack of real-world deployment.

Further research by Johnson et al. employed **Long Short-Term Memory (LSTM)** and **Recurrent Neural Networks (RNNs)** to analyze user-generated content on social media platforms [3]. Their model reached an 82% accuracy rate in identifying patterns associated with mental health distress, such as sleep disruption, mood swings, or withdrawal from social activity. However, the study highlighted privacy risks and algorithmic bias as key challenges when using social media data for mental health analysis.

Beyond user-generated content, Electronic Health Records (EHRs) have also been explored as a valuable data source for AI-driven mental health prediction. Several UK-based studies have implemented **CogStack**, an open-source data pipeline for processing unstructured clinical text in EHRs, in combination with NLP tools like **MedCAT** and **BioYODIE**, to extract relevant psychiatric information from hospital records [4]. These systems enabled the training of models for identifying mental health deterioration, suicide risk, and crisis prediction. The use of CogStack has shown promise in hospital settings, particularly for creating real-time alert systems. However, challenges remain regarding data standardization, patient consent, and the integration of these insights into routine clinical workflows.

While Random Forest and SVMs were effective in structured data classification, LSTM networks outperformed them in sequential data processing — making them particularly suitable for applications involving user text input or chat logs. Recent research also highlights the effectiveness of ensemble approaches that combine RNN-based architectures with transformer models to achieve greater robustness across varying class distributions.

Collectively, these AI models showcase the potential for early mental health intervention at scale. However, challenges like data diversity, ethical AI implementation, and integration with clinical practices remain vital areas for future research.

B. Chatbot and AI-Assisted Mental Health Platforms

Chatbot-based mental health support systems have gained traction as cost-effective and accessible digital therapy solutions. The Saarthi mental health chatbot, introduced at the 2023 International Conference on Disruptive Technologies (ICDT), employs AI-driven NLP algorithms to deliver Cognitive Behavioral Therapy (CBT), peer support, and real-time intervention [4]. This study highlighted that chatbots can effectively assist individuals suffering from depression and anxiety, offering structured CBT-based interventions through virtual assistants. However, it also noted that chatbots cannot fully replace human interaction and are best used as complementary support systems.

Furthermore, research suggests that chatbots can track user mental health over time using machine learning algorithms to analyze user input and classify distress levels. These models segment mental health conditions into different severity levels and can escalate high-risk cases to professional intervention [5].

While chatbots provide scalable mental health solutions, studies indicate limitations in deep emotional intelligence, ethical considerations, and privacy concerns when using AI for psychological therapy [6].

C. ICF-Based Community Mental Health Systems

Beyond AI-based chatbots and predictive models, ICF-based community health systems play a crucial role in mental health rehabilitation and crisis prevention. A study published in Healthcom 2004 (IEEE) introduced an ICF-based electronic health-sharing platform that enables psychiatric hospitals, community support centers, and rehabilitation facilities to share patient information in real-time [7]. This system integrates electronic bulletin boards and clinical data repositories to enhance collaboration among mental health professionals.

The study emphasized that standardizing mental health records across different care providers can improve long-term patient monitoring and reduce hospital readmissions. However, data security and privacy measures remain a challenge, as many digital community health systems lack strong encryption protocols [7].

Additionally, Assertive Community Treatment (ACT) models have been explored for long-term mental health support. ACT integrates networked healthcare teams to provide 24/7 mental health assistance. Studies suggest that integrating AI-driven analytics into ACT models could improve early detection and personalized care [8].

D. AI and Robotics in ADHD Therapy

Recent advancements in robot-assisted mental health therapy have introduced Large Language Models (LLMs) and Socially Assistive Robots (SARs) for treating neurodevelopmental disorders like ADHD. A study presented at the 2024 International Conference on Intelligent Environments (IEEE) explored the integration of ChatGPT-4 Turbo and Claude-3 Opus in robotic assistants for ADHD therapy [9].

The study compared two AI models in robot-assisted therapy:

- **ChatGPT-4 Turbo** excelled in performance and responsiveness, making it suitable for time-sensitive interactions.
- **Claude-3 Opus** demonstrated higher coherence and ethical considerations, prioritizing safety and engaging dialogue.

The results indicated that robotic assistants powered by LLMs could provide personalized cognitive support for children with ADHD, enhancing engagement and therapeutic outcomes. However, further research is needed to address ethical concerns, AI bias, and long-term integration into clinical practice [9].

Similarly, Tamdjidi et al. investigated the role of AI in ADHD reading comprehension support and found that individuals with prior experience using ChatGPT showed improved comprehension abilities [10]. These findings suggest that AI-powered cognitive assistants could serve as complementary tools in ADHD therapy.

E. Challenges in Security, Privacy, and Ethical Considerations

While AI and digital mental health platforms improve accessibility and efficiency, they pose significant security and ethical concerns.

- EHR-based predictive models raise concerns about data ownership and patient confidentiality, especially when utilizing unstructured clinical notes and NLP algorithms [1], [7].
- AI-driven chatbots may introduce bias and misinformation, potentially leading to inaccurate self-diagnoses [4], [5].
- Robotic assistants for ADHD therapy need stricter regulatory frameworks to ensure ethical AI use and child safety [9].

Studies recommend implementing blockchain-based security models for AI-driven mental health platforms and GDPR/HIPAA-compliant encryption for community-based digital health networks [8].

III. Methodology

This study employs a dual-model architecture combining a traditional BiLSTM network with a transformer-based GPT-2 model to classify mental health-related text. The objective is to leverage the complementary strengths of both models to improve classification accuracy, especially on underrepresented classes.

A. Dataset and Preprocessing

A custom-labeled dataset comprising 823 text samples categorized into three mental health classes (LABEL_0, LABEL_1, and LABEL_2) was used. The data was split into 80% training and 20% validation sets. Preprocessing included token normalization, padding, truncation (to a maximum length of 128 tokens), and vocabulary building for the BiLSTM model. The GPT-2 model used pre-trained tokenizers from the Hugging Face Transformers library.

B. BiLSTM Model Architecture

The BiLSTM model was developed using PyTorch and consists of:

- An embedding layer (dimension = 128)
- Two bidirectional LSTM layers (hidden size = 256)
- A dropout layer (rate = 0.3)
- A fully connected classification layer

The model was trained using dynamic batching, padded sequences, and a cross-entropy loss function. Validation was conducted at each epoch to monitor overfitting and performance degradation.

C. GPT-2 Fine-Tuning

The GPT-2 model (125M parameters) was fine-tuned using the Hugging Face pipeline with the following configurations:

- A classification head added to the transformer output
- Max sequence length of 128
- AdamW optimizer with a learning rate of 2e-5
- Training batch size of 8 over 4 epochs

The fine-tuning aimed to adapt GPT-2's semantic context capabilities to domain-specific classification tasks.

D. Ensemble Strategy

An ensemble approach is being implemented to combine the predictions from both models. The method involves:

1) Generating independent predictions and confidence scores from BiLSTM and GPT-2
2) Normalizing the scores to ensure compatibility

3) Combining predictions through either majority voting or a weighted averaging scheme based on class-wise F1 performance
4) Producing the final label from the ensemble decision

The ensemble is expected to enhance generalization and particularly boost performance on the minority class (LABEL_2).

E. Implementation Details

The system is modularized into:

- Data loading and preprocessing scripts
- Model training modules for BiLSTM and GPT-2
- Evaluation tools using accuracy, F1-score, and confusion matrix
- An inference API for deploying the models in real-time applications

All components are built using PyTorch and the Hugging Face Transformers library for reproducibility and scalability.

IV. Results

V. Future Scope

1) Research Gaps and Future Directions: Despite advancements in AI-driven mental health solutions, several critical research gaps remain:

1) Hybrid AI-Human Mental Health Models – Most platforms rely on either AI-driven chatbots or human therapists. Research should focus on integrating AI with human intervention models for better crisis detection and response [2], [4].
2) Scalability of Real-Time Mental Health Support – WebRTC-based video support and AI chatbots have been explored, but scalability for high-volume, real-time mental health consultations remains a challenge [3], [5].
3) Enhanced Privacy Mechanisms for AI-Driven Therapy – Digital mental health platforms require stronger encryption, ethical AI governance, and improved user anonymity protocols [7], [9].

References

[1] Jayant Choudhary; Devashish Awasthi; Bhaskar Bhatt; Harshit Narang; Satvik Vats; Vikrant Sharma, "SentiSense: Pioneering Mental Health Application Through Advanced Technology" IEE int.2024 International Conference on Electronics, Computing, Communication and Control Technology (ICECCC)

[2] S. Manabe; N. Ashida; T. Takemura; K. Murase; N. Nishiura "ICF-based community mental health care" iEEE int. Proceedings. 6th International Workshop on Enterprise Networking and Computing in Healthcare Industry - Healthcom 2004 (IEEE Cat. No.04EX842)

[3] Yamiko Joseph Msosa; Arturas Grauslys; Yifan Zhou; Tao Wang; Iain Buchan; Paul Langan, "Trustworthy Data and AI Environments for Clinical Prediction: Application to Crisis-Risk in People With Depression" IEEE Journal of Biomedical and Health Informatics (Volume: 27, Issue: 11, November 2023)

[4] Yanjuan Shen; "Analysis on Mental Health Education Strategy of Higher Vocational College Students Based on Information Technology" IEEE 2021 2nd International Conference on Education, Knowledge and Information Management (ICEKIM)

[5] Komal Rani; Harshit Vishnoi; Manas Mishra; "A Mental Health Chatbot Delivering Cognitive Behavior Therapy and Remote Health Monitoring Using NLP And AI" 2023 International Conference on Disruptive Technologies (ICDT)

[6] Reza Tamdjidi et al., "Enhancing ADHD Reading Comprehension with AI: A Comparative Study of ChatGPT in Education," 2023.

[7] OpenAI, "GPT-4 Technical Report," arXiv preprint arXiv:2303.08774, 2023.

[8] Santiago Berrezueta-Guzman; Mohanad Kandil; María-Luisa Martín-Ruiz; Iván Pau de la Cruz; Stephan Krusche "Exploring the Efficacy of Robotic Assistants with ChatGPT and Claude in Enhancing ADHD Therapy: Innovating Treatment Paradigms" 2024 International Conference on Intelligent Environments (IE)

[9] I. Bean, L. Teo, R. Dobson, R. Stewart, and Z. Jackson, "CogStack–Experiences of deploying integrated information retrieval and extraction services in a large NHS Foundation Trust hospital," *BMC Medical Informatics and Decision Making*, vol. 19, no. 1, pp. 1–12, 2019.