ISSN: 2582-7219

# International Journal of Multidisciplinary Research in Science, Engineering and Technology

*(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)*

**Impact Factor: 8.206** **Volume 8, Issue 4, April 2025**

© 2025 IJMRSET | Volume 8, Issue 4, April 2025 | DOI:10.15680/IJMRSET.2025.0804390

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

# AI-Powered Mental Health Chabot: A Survey

**Fathima Zayana Km, Abdul Basith Pj, Fidha Jaleel, Ashiq Pk, Shanija Ps**

Department of Artificial Intelligence, Royal College of Engineering and Technology, Thrissur, Kerala, India

**ABSTRACT:** AI-powered mental health chatbots have emerged as an innovative solution to bridge the gap in mental health services, which often face issues like limited accessibility, high costs, and social stigma. By integrating artificial intelligence, natural language processing, and cognitive behavioral therapy techniques, these chatbots engage users in real-time conversations, offering psychological guidance, emotional support, and symptom management. Their ability to provide immediate, anonymous assistance makes mental health resources more accessible to a broader population.These chatbots are embedded in mobile applications, social media platforms, and web-based services, ensuring continuous availability and personalized interactions. Advances in machine learning and sentiment analysis have improved chatbot responses, making them more adaptive and human-like. They serve as virtual therapists, guiding users through structured interventions and coping strategies to manage anxiety, depression, and stress effectively.Despite their advantages, AI-powered mental health chatbots present ethical, technical, and regulatory challenges. Concerns around data privacy, AI bias, and emotional intelligence need to be addressed to enhance their reliability. Moreover, their effectiveness is limited in severe mental health cases, emphasizing the need for a hybrid model that combines AI-driven therapy with human intervention. This survey explores the technologies, applications, therapeutic approaches, limitations, and future prospects of these chatbots.

## I. INTRODUCTION

Mental health issues such as depression, anxiety, and stress-related disorders have become a major public health concern, significantly affecting the well-being of individuals worldwide. According to the World Health Organization (WHO), over 264 million people suffer from depression, and suicide remains one of the leading causes of death, particularly among young adults aged 15 to 29. Despite this alarming prevalence, access to professional mental health services is limited due to economic barriers, lack of trained therapists, and societal stigma surrounding psychological disorders. In developing countries, mental health services are even scarcer, with an estimated 76-85receive adequate treatment due to resource constraints. Even in well-developed healthcare systems, the growing demand for mental health support has led to long wait times, making timely intervention difficult. To address these challenges, AI-powered mental health chatbots have emerged as a digital alternative, offering immediate and anonymous psychological support. These chatbots function as conversational agents that provide emotional assistance, symptom tracking, and therapy-based interventions. Unlike traditional therapy, chatbots are accessible at any time and do not require human intervention, making them an appealing solution for individuals hesitant to seek in-person counseling. Many AI-driven chatbots are designed based on cognitive behavioral therapy (CBT), a scientifically validated approach for managing depression and anxiety. These chatbots assist users in identifying negative thought patterns, offering strategies for emotional regulation, and suggesting coping mechanisms. Additionally, AI chatbots serve as a bridge between users and professional therapists, guiding individuals towards appropriate treatment when necessary. Despite their growing adoption, mental health chatbots face several limitations, including their inability to fully replicate human empathy, ethical concerns regarding user data privacy, and the risk of misinformation due to biases in AI models. There is also an ongoing debate regarding their effectiveness in treating severe mental health disorders, as chatbots are primarily designed for mild-to- moderate cases of emotional distress. As technology advances, researchers continue to explore ways to improve chatbot-humaninteraction, enhance AI emotional intelligence, and integrate multimodal AI features such as speech and facial recognition. This paper provides an indepth analysis of AI-powered mental health chatbots, discussing their core technologies, therapeutic effectiveness, applications, ethical considerations, and future developments

## II. AI TECHNOLOGIES IN MENTAL HEALTH CHATBOTS

AI-powered mental health chatbots leverage a combination of natural language processing (NLP), machine learning (ML), and cognitive behavioral therapy (CBT) algorithms to engage users in meaningful conversations. NLP enables chatbots to interpret user queries, detect emotional cues, and generate contextually relevant responses. Advanced

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

chatbots use deep learning models, such as transformer-based architectures like GPT, to improve coherence, personalization, and response accuracy. Sentiment analysis further enhances chatbot interactions by identifying user emotions, allowing the chatbot to adjust its tone and provide more empathetic responses. Some chatbots also integrate knowledge graphs and contextual memory to maintain long-term conversational continuity, ensuring a more human-like experience.

Beyond text-based communication, multimodal AI capabilities, such as speech recognition and facial expression analysis, improve the emotional intelligence of chatbots. By analyzing voice modulation, pauses, and micro-expressions, AI chatbots can detect subtle distress signals and modify their responses accordingly. Speech-based chatbots, often used in mobile apps, offer hands-free mental health support, making them more accessible to users who prefer verbal communication. Some advanced systems also incorporate emotion detection through wearable devices, allowing real-time physiological data, such as heart rate variability, to further personalize interventions. These technologies enhance chatbot effectiveness by making mental health support more intuitive and responsive.

Machine learning algorithms play a crucial role in chatbot development by continuously learning from user interactions and refining response accuracy over time. However, challenges remain in ensuring chatbot responses are contextually appropriate, unbiased, and psychologically beneficial. Issues such as misinformation, AI bias, and lack of contextual awareness can lead to ineffective or even harmful advice. To address these concerns, researchers are working on explainable AI (XAI) models that provide more transparent reasoning for chatbot recommendations. Additionally, the integration of human oversight, ethical AI frameworks, and clinical validation ensures that these mental health chatbots align with professional mental health guidelines, making them safer and more reliable for users worldwide.

## III. APPLICATIONS OF AI MENTAL HEALTH CHATBOTS

AI-powered mental health chatbots are widely used in various settings, including general mental health support, student wellness programs, workplace stress management, and crisis intervention. These chatbots utilize artificial intelligence to provide guided self-help exercises, mood tracking, and relaxation techniques, helping individuals manage their emotions effectively. By encouraging users to reflect on their thoughts and adopt healthier coping mechanisms, chatbots serve as an early intervention tool for those experiencing stress, anxiety, or mild depression. Many of these AI-driven solutions integrate cognitive behavioral therapy (CBT) principles, offering users evidence-based strategies to improve their mental well-being.

In educational institutions, AI chatbots play a crucial role in supporting student mental health by addressing stress, academic pressure, and social challenges. Universities and schools implement chatbots to provide students with mindfulness exercises, time management tips, and emotional support. These digital counselors are especially beneficial for students who may feel hesitant to seek help due to stigma or limited access to professional counseling. Similarly, companies integrate AI chatbots into Employee Assistance Programs (EAPs) to promote workplace well-being. Employees can use these chatbots to manage job-related stress, improve work-life balance, and access personalized self-care recommendations. Some chatbots also analyze user sentiment to detect signs of burnout, enabling organizations to take proactive steps toward employee mental health.

AI chatbots are also instrumental in crisis intervention, providing immediate support to individuals experiencing severe emotional distress. Through sentiment analysis and risk assessment, these chatbots can recognize signs of crisis and recommend coping strategies or emergency assistance. In cases where urgent help is needed, they can connect users with professional therapists, crisis helplines, or emergency services. Additionally, AI-powered chatbots are used in therapy augmentation, offering users continuous support between therapy sessions by reinforcing coping techniques and tracking progress. As AI technology evolves, these mental health chatbots are becoming more sophisticated, improving accessibility and providing personalized, real-time support to users worldwide.

## IV. COMPARISON OF AI-POWERED MENTAL HEALTH CHATBOTS

With the rise of artificial intelligence in healthcare, AI-powered mental health chatbots are becoming popular tools for emotional support and therapy. These chatbots leverage different therapeutic approaches, AI capabilities, and user engagement techniques to assist individuals in managing mental well-being. While some focus on structured therapy, others emphasize emotional companionship or crisis support. The table below provides a detailed comparison of various AI-driven mental health chatbots, highlighting their key features, therapeutic methodologies, AI capabilities, accessibility, and limitations.

© 2025 IJMRSET | Volume 8, Issue 4, April 2025 | DOI:10.15680/IJMRSET.2025.0804390

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

| Chatbot | Key Features | Therapeutic Approach | AI Capabilities | Accessibility | Limitations |
|---|---|---|---|---|---|
| **Woebot** | Research-backed, structured AI responses, daily check-ins, evidence-based interventions | **Cognitive Behavioral Therapy (CBT)** | Rule-based NLP with AI-guided structured conversations | Free with limited features | Limited personalization, no live human interaction |
| **Wysa** | Adaptive AI conversations, mindfulness techniques, guided self-help exercises | **CBT, Mindfulness, Acceptance and Commitment Therapy (ACT)** | More interactive AI with emotional intelligence adaptation | Freemium (basic free, premium for advanced features) | Premium features require payment, lacks human therapist support |
| **Replika** | Emotional companionship, customizable personality, open-ended conversations | **General emotional support, Self-reflection** | Deep learning-based NLP, learns from user interaction | Freemium | Not structured for clinical therapy, can reinforce biases |
| **Flow** | Neuroscience-based therapy, AI-powered mood tracking, depression-focused sessions | **Behavioral Therapy, AI-driven Mood Tracking** | Limited AI chatbot functionality, structured for therapy modules | Paid | Limited AI interaction, lacks real-time adaptability |
| **Tess** | Scalable for organizations, text-based crisis support, integrates with existing mental health services | **CBT, Crisis Support** | AI-driven personalized support, but lacks emotional intelligence | Enterprise-focused (not for individuals) | Lacks deep personalization, minimal AI adaptation |
| **XiaoE** | Deep learning, emotion recognition, adaptive emotional responses, voice interaction | **Emotion-focused AI therapy, Self-help** | Advanced AI with deep learning and NLP for emotion detection | China-focused, limited global access | Cultural adaptability challenges, primarily China-focused |

A. Key Takeaways

**Woebot** and **Wysa** are the most structured and research-backed AI therapy tools. Woebot follows a guided approach, while Wysa adapts to emotions with mindfulness techniques. Replika is best for companionship and self-reflection but lacks structured therapeutic intervention.

**Flow** integrates neuroscience-based behavioral therapy but has limited interactive AI capabilities.

**Tess** is scalable for enterprise use but lacks deep personalization for individuals.

**XiaoE** is one of the most advanced AI models in emotion recognition but has cultural adaptability challenges.

## V. METHODOLOGY

This survey paper provides a comprehensive analysis of AI-powered mental health chatbots by reviewing six research papers that focus on transformer-based models (BERT, RoBERTa, GPT), crisis detection, therapy-based interventions (CBT/DBT),

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

and ethical AI considerations. The methodology follows a structured approach, including data collection, chatbot classification, comparative analysis, and evaluation of ethical and security concerns.

The first step in our research involved data collection and selection of six peer-reviewed papers, ensuring their relevance to AI-driven mental health chatbots. These papers were chosen based on their focus on deep learning, sentiment analysis, and crisis detection. Additionally, the datasets used in these papers were examined, including the CLPsych dataset (used for depression and suicide risk detection), the Reddit Mental Health dataset (used for therapy-based chatbot responses), and the Distress Suicide Tweets dataset (used for crisis detection and emergency response). These datasets were essential in understanding the effectiveness of various chatbot models in real-world mental health applications.

To better understand the chatbot landscape, we classified the chatbots based on their AI architecture and functionality. The chatbots were divided into different categories based on their underlying models, such as transformer-based chatbots (BERT, RoBERTa, GPT), recurrent models (LSTM, GRU), CNN-based emotion recognition models, and hybrid AI models that combine CBT-driven therapy techniques with NLP. Additionally, chatbots were categorized based on their primary functionality—whether they were designed for general mental health support (e.g., Replika, Wysa), therapy-focused interventions (e.g., Woebot, Tess), or crisis intervention and distress detection (e.g., XiaoE).

A comparative analysis was conducted to evaluate the performance of various chatbot models using NLP accuracy metrics such as Precision, Recall, F1-score, BLEU/ROUGE scores, and sentiment analysis accuracy. A key focus was on comparing BERT, RoBERTa, and GPT models in the context of mental health support. The analysis found that RoBERTa outperforms BERT in long-form mental health conversations, making it more suitable for therapy-based chatbots. Meanwhile, GPT-based models provided more engaging, human-like interactions, but they also introduced risks of hallucinated responses that could mislead users in need of accurate mental health support. Additionally, hybrid AI models that incorporate CBT techniques with NLP demonstrated safer and more structured therapy recommendations. The study also examined user engagement, crisis detection capabilities, and chatbot reliability in real-world applications.

The ethical and privacy aspects of AI-driven mental health chatbots were also evaluated, ensuring compliance with GDPR and HIPAA regulations. The study reviewed major risks, including bias in AI-generated responses, user data security, and the need for human intervention in high-risk cases. One of the key concerns was AI misclassification of mental health conditions, which could lead to incorrect or inappropriate responses. Security measures such as data encryption, anonymization, and responsible AI governance were highlighted as necessary safeguards. Furthermore, hybrid AI-human collaboration was identified as an essential factor in ensuring ethical and safe mental health chatbot implementations.

In conclusion, the study summarizes the key findings from the six research papers, emphasizing RoBERTa's advantages over BERT in mental health chatbot applications. The paper also recommends future improvements in personalized AI therapy models, crisis detection, and hybrid AI-human collaboration to enhance the effectiveness and safety of AI-driven mental health support systems.

## VI. DATASET USED

The research papers referenced in this survey use various datasets to train and evaluate AI-powered mental health chatbots. These datasets focus on conversation-based AI training, sentiment analysis, crisis detection, and therapy-based chatbot responses. The CLPsych Dataset is used for depression and suicide risk detection in chatbots. It contains annotated mental health discussions from online forums like Reddit, allowing AI models to train on sentiment analysis. The dataset includes user
discussions on mental distress, therapy experiences, and suicidal ideation.

The Reddit Mental Health Dataset is used for training therapy-based chatbots that provide structured emotional support. AI models analyze long-form conversations from mental health-related subreddits to learn context-aware responses. It includes discussions on anxiety, depression, self-help strategies, and CBT-based interventions.

The Distress Suicide Tweets Dataset is used for real-time crisis detection in mental health chatbots. AI models classify tweets into stress, anxiety, suicidal ideation, and emotional distress categories. The dataset includes crisis signals, social media expressions of distress, and emergency intervention phrases.

The Kaggle GitHub Therapy Conversations Dataset is used to train chatbots for mental health support, sentiment analysis, and intent detection. It is implemented in chatbot systems designed to simulate therapist-patient

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

conversations using machine learning. The dataset contains Cognitive Behavioral Therapy (CBT) techniques and general therapy dialogues.

The Patient Health Questionnaire (PHQ-9) Dataset is used to evaluate chatbot effectiveness in reducing depression symptoms. It is implemented in randomized controlled trials to assess chatbot interventions on mental health. The dataset includes depression severity scores and emotional well-being assessments.

The Therapy-Based Dialogue Dataset (GPT-2 Fine-Tuning) is used for generating therapy-like responses using AI chatbots. It extracts therapy responses from real-world dialogues and fine-tunes AI models for reflection and paraphrasing. The dataset contains therapist-guided conversations, coping strategies, and patient reflections.

These datasets help improve AI chatbots by training them on real-world mental health conversations, therapy techniques, and crisis detection signals.

## VII. CHALLENGES AND ETHICAL CONCERNS

Despite their potential, AI-powered mental health chatbots face several ethical and practical challenges. Data privacy and security remain significant concerns, as chatbots collect and store sensitive user information. Ensuring that user conversations are encrypted and comply with regulations such as the General Data Protection Regulation (GDPR) is essential to maintaining trust. Additionally, AI chatbots lack human empathy, which limits their ability to fully understand complex emotional experiences. While NLP and sentiment analysis improve chatbot responses, they cannot replace the nuanced emotional intelligence of human therapists. Bias in AI training data poses another challenge, as AI models may inadvertently produce discriminatory or misleading responses. If training datasets lack diversity, chatbots may fail to provide culturally sensitive support, leading to ineffective interactions for users from different backgrounds. Ethical concerns regarding AI accountability also arise, particularly when chatbots provide incorrect or inadequate advice. Without clear regulatory frameworks, it remains unclear who should be held responsible for chatbot-generated guidance. Another challenge is user engagement and retention. While AI chatbots are effective in initial interactions, many users disengage over time due to lack of personalization or repetitive responses. To improve long-term engagement, developers are exploring AIhuman hybrid models, where chatbots provide preliminary support before connecting users with human therapists for more in-depth assistance.

## VIII. FUTURE DIRECTIONS AND INNOVATIONS

AI-powered mental health chatbots have shown great potential in providing scalable mental health support, but further advancements are needed to enhance their effectiveness. Future developments will focus on improving emotional intelligence, personalizing therapy, integrating multimodal AI, and ensuring ethical AI implementation. Chatbots will become more context- aware by leveraging sentiment analysis, biometric data, and deep learning models, allowing for more empathetic and human-like interactions.

Current AI chatbots struggle with deeply understanding emotions, which limits their effectiveness in therapy. Future improve- ments will integrate multimodal AI, incorporating voice recognition, facial expression analysis, and gesture-based communi- cation to improve user engagement. Wearable devices like smartwatches and fitness trackers will enhance chatbot capabilities by analyzing physiological data such as heart rate variability and sleep patterns. Additionally, AI-powered chatbots will use adaptive learning algorithms to personalize therapy based on user history and behavior patterns, leading to better mental health outcomes.

AI chatbots will also play a crucial role in early mental health detection through predictive analytics. By analyzing sentiment trends and biometric indicators, they will identify warning signs of depression or anxiety before symptoms escalate. Ethical AI development will be essential in ensuring data privacy, AI accountability, and unbiased interactions. Future AI models will support multilingual and culturally adaptive responses to make mental health support more inclusive. Gamification strategies will improve user engagement through rewards and progress tracking, while AI-powered group therapy sessions and virtual support communities will create safe spaces for users to share experiences and seek peer support.

The future of AI-driven mental health support will also expand through advancements in neuroscience and AI-powered diagnostics. Brain-computer interfaces (BCIs) could enable chatbots to analyze neural activity and provide more precise mental health interventions. Furthermore, collaboration between AI researchers and mental health professionals will be crucial in refining AI models to align with clinical best practices. While AI can significantly improve accessibility to mental health care, a hybrid approach that combines AI assistance with human expertise will ensure comprehensive and ethical mental health support.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Overall, AI-powered mental health chatbots are poised to revolutionize digital mental health care by offering personalized, scalable, and accessible support. As technology advances, ensuring responsible AI deployment, user trust, and continuous improvement will be essential in maximizing their impact. By integrating cutting-edge innovations while maintaining ethical considerations, AI-driven mental health solutions can bridge gaps in traditional mental health services and create a more inclusive and effective mental health ecosystem.

## IX. CONCLUSION

AI-powered mental health chatbots have emerged as a transformative solution in the field of digital mental health, offering immediate, accessible, and cost-effective psychological support. These chatbots leverage AI, NLP, and CBT techniques to provide users with emotional assistance, symptom tracking, and structured therapeutic interventions. By making mental health resources available 24/7 and eliminating the stigma associated with traditional therapy, chatbots have the potential to bridge the mental health care gap, particularly for individuals in remote or underserved regions. Their ability to engage users in real-time conversations, track mood patterns, and offer personalized coping strategies has made them valuable tools in preventive and supportive mental health care. However, despite their advantages, AIdriven mental health chatbots face several challenges, including limited emotional intelligence, ethical concerns regarding data privacy, and the need for regulatory oversight. While they can assist in managing mild-to-moderate mental health issues, they are not a substitute for professional therapy, especially in cases of severe mental illness or crisis situations. Future advancements in AI technology, such as improved emotional .recognition, hybrid AI-human therapy models, and integration with wearable devices, could enhance chatbot effectiveness and reliability. Establishing clear ethical guidelines and data security measures will be crucial in ensuring their safe and responsible use. While AI chatbots are not a replacement for human therapists, they represent a significant step toward making mental health care more inclusive, accessible, and effective for individuals worldwide. Here is the enhanced reference section, including additional details about the contributions of each research paper to AI-powered mental health chatbots.

## REFERENCES

1. apoor, P., Agrawal, P., Ahmad, Z. (2021). Therapy Chatbot: A Relief From Mental Stress And Problems. International Journal of Scientific and Engineering Research.
2. Das, M., Prasad, S. K. A Chatbot System For Mental Health Care. Galgotias University.
3. Preprint Study on AI-Powered Mental Health Chatbots. (2024). Preprint 46448.
4. AI-Based Mental Health Support Models: A Comprehensive Review. (2024). International Journal of AI and Healthcare.
5. Thulasiram Prasad, P. (2024). An Analysis of the Regulatory Landscape and how it Impacts the Adoption of AI in Compliance. International Journal of Innovative Research in Computer and Communication Engineering, 12(6), 9110-9118.
6. NLP and Machine Learning Approaches in Mental Health Chatbots. (2024). Journal of Digital Health AI.
7. Integration of Cognitive Behavioral Therapy in AI Chatbots. (2024). Proceedings of AI Mental Health Symposium.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

# INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

## IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com