iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: III Month of publication: March 2026

DOI: https://doi.org/10.22214/ijraset.2026.78375

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

AI-Powered Mental Health Assistance and Chatbot

Aryan Mhatre[1], Roheel Panchal[2], Ronak Rai[3], Swapnil Bhagat[4]
Dept. of Computer Engineering, Thakur college of Engineering and Technology

Abstract: Mental health support systems are challenged in terms of accessibility, affordability, and stigmatization, which are barriers for people to access mental health support services. This project proposes a mental health support system called Mind Well, which is a hybrid AI-based mental health chatbot. The proposed mental health support system is based on a fine-tuned Distil BERT model to classify emotions in real-time and the Google Gemini API to deliver supportive responses to users. For responsible usage, the proposed mental health support system is based on a set of safety rules to avoid diagnosis, be transparent about being an AI companion, and refer users to helplines in crisis situations. The proposed mental health support system is implemented using a Flask-based backend with JWT authentication and a MongoDB database to ensure user safety and data privacy. ***Keywords: Mental Health Chatbot, Natural Language Processing (NLP), Emotional AI, BERT, Distil BERT, Conversational Agents, Psychological Support, AI in Healthcare.***

I. INTRODUCTION

These chatbots are advanced computer programs that were specifically designed to mimic the way humans communicate. This is mostly done through text-based interfaces, though it is also common to find chatbots that use voice. The core of the chatbots' function is based on Artificial Intelligence (AI), which gives them the ability to comprehend and interpret the way humans communicate. With the use of AI in chatbots, they have the ability to communicate in ways that are both relevant and coherent. AI is composed of various fields of study that allow the chatbots to learn from their interactions and enhance their capabilities of communication. With the use of chatbots in the field of mental health, it is an incredibly promising field. This is because it provides the unique opportunity to offer instant support that is around the clock and confidential. Chatbots offer the opportunity to be a preliminary resource and a companion that provides the feeling of presence without the constraints of the way humans communicate.

The major contributions of the work are:

1) The development of a system architecture with a user-friendly interface, an AI inference engine for dialogue management, and a secure backend for handling user data.
2) The utilization of transformer models (BERT and Distil BERT) in understanding the semantic context of user dialogue and emotional nuances related to mental health.
3) The development of a framework for offering support and coping mechanisms or resources based on the emotional states identified.

II. BACKGROUND & MOTIVATION

Mental health problems have come to be recognized as a major health issue in the world, affecting a large percentage of the world's population across various demographic groups. Mental health problems such as anxiety disorders, depression, and stress have come to have a profound impact on the lives of people across the globe. These problems can come in a wide range of manifestations, including feelings of hopelessness and despair. Despite the high prevalence of these problems in society, a large number of people have come to find themselves without the support and resources they need.

III. LITERATURE REVIEW

1) Conversational Agents for Mental Health Support (2021)

Key Features: The first conversational agents in the field of mental health support were based on rule-based systems and basic machine learning approaches. The basic chatbot systems were based on a set of scripts and pattern matching mechanisms, similar to the ELIZA model of dialogue systems. The basic objective of the systems was to mimic supportive conversations and offer basic emotional support to users.

Gap Identified: The basic issue with the systems was their inability to offer personalization and contextual understanding. The systems were unable to understand emotional cues or user input and were unable to adapt to different user inputs.

Relevance to Our Work: The basic paper provides a foundation for the development of digital systems to offer mental health support. Our work is based on the basic concepts of the earlier systems, with the addition of advanced NLP mechanisms to offer better contextual understanding.

2) NLP-driven Sentiment Analysis for Mental Health Chatbots (2023)

Key Features:

The focus of this study is to apply various NLP techniques, including sentiment analysis and intent detection, to improve chatbots used in mental health support. The chatbots will be able to understand emotional tones and intent in real-time to provide a more supportive and relevant response to the user's queries.

Gap Identified: Although various studies have improved these chatbots to a certain extent, one of the biggest challenges is interpreting sarcasm and cultural differences. In addition, maintaining a sense of continuity in a conversation is a big challenge. This is because context changes over time in a conversation. Maintaining such continuity is a big challenge. However, it is also a challenge that is not directly relevant to our study. The study is relevant to our project because it emphasizes the significance of semantic analysis. However, it is not directly applicable to our project. Our study will also focus on maintaining context through the help of transformers.

3) LLM-powered Therapy Chatbots (2023)

Key Features:

Recent studies explore the use of large language models (LLMs) to generate empathetic and context-aware dialogue. These systems often involve fine-tuned LLMs trained on therapeutic conversations and enhanced with safety filters to prevent harmful responses.

Gap Identified:

However, the use of LLMs introduces risks such as hallucinated responses, ethical concerns, and a lack of formal clinical validation. Ensuring safe and reliable outputs remains a critical challenge.

Relevance to Our Work:

This work demonstrates the potential of LLMs for generating empathetic dialogue. Our project focuses on controlled fine-tuning within the mental health domain and incorporates safeguards to improve reliability and ethical deployment.

4) Hybrid Human-AI Mental Health Assistance (2024)

Key Features: The hybrid approach would involve AI chatbots and human therapists. In a hybrid approach, chatbots would be used to assist in initial interactions, emotional analysis, and guidance. However, complex cases would be escalated to human therapists through a structured escalation pipeline.

Gap Identified: The hybrid approach would face challenges such as scalability and efficient human-AI interaction, ensuring robust data privacy features.

Relevance to Our Work: The hybrid approach would be beneficial to a hybrid model that would involve AI assistance in early-stage support. Our work would be able to incorporate escalation pipelines that would ensure a smooth transition to human therapists while ensuring user trust and security.

5) Workplace Wellbeing via AI Chatbots (2024)

Key Features: The research focuses on the application of AI chatbots to support employee mental wellbeing through the integration of large language models, sentiment analysis, and reinforcement learning from human feedback.

Gap Identified: The research identified a major gap in the availability of large datasets related to workplace mental health. In addition, the problem of scale while reducing bias has been identified as a major issue.

Relevance to Our Work: This research is directly related to the work we aim to do in the field of mental health support systems. Our proposed solution has the ability to overcome the problem of scale while reducing bias.

Sr.	Title	Source/Year	Key Features	Gap Identified	Relevance to Our Work
1.	Conversational Agents for Mental Health Support	2021	Rule-based and early ML-driven chatbots (e.g., ELIZA-like, scripted responses)	Limited personalization, shallow understanding, unable to adapt to nuanced user inputs	Provides a foundation for digital mental health tools; our work builds on this by using advanced NLP/LLMs for deeper contextual understanding
2.	NLP-driven Sentiment Analysis for Mental Health Chatbots	2023	Sentiment analysis + intent classification for conversational mental health agents.	Struggles with sarcasm, cultural context, and multi-turn conversations	Reinforces the need for robust semantic understanding; our system extends this with transformer-based context retention
3.	LLM-powered Therapy Chatbots (e.g., GPT-based)	2023	Fine-tuned LLMs with empathetic dialogue generation and safety filters	Risks of hallucinations, ethical/safety concerns, lack of clinical validation	Demonstrates potential of LLMs in providing empathetic dialogue; our project focuses on controlled, domain-specific fine-tuning
4.	Hybrid Human-AI Mental Health Assistance	2024	Chatbot triaging + human therapist escalation pipeline	Scalability and handoff quality remain challenging; data privacy concerns	Supports hybrid care model; our work can integrate escalation mechanisms while maintaining user trust
5.	Workplace Wellbeing via AI Chatbots	2024	LLMs + Sentiment Analysis + RLAIF for workplace mental health support	Limited workplace-specific datasets, scalability issues	Directly relevant—our system aims to improve wellbeing support while addressing scalability and bias challenges

IV. ARCHITECTURE DIAGRAM

1) Data Source & Training Layer

•Mental Health Dialogue Dataset: A dataset of anonymized mental health counseling sessions, peer support dialogues, and mental health discussion forums will be used as the foundation for the proposed model.
•Training Pipeline: The module will preprocess and format the text data, fine-tune the transformer-based models (BERT or DistilBERT), and update the models periodically to achieve high robustness, empathy, and understanding of user dialogues.

2) User Interaction Layer

•Web/Mobile Application: Users will interact with the chatbot through a web or mobile application interface, allowing them to engage in a dialogue with the chatbot in a conversational manner.
•Free-text Input: Users can express their emotions or thoughts without the need for a structured response.

3) AI Processing & Inference Layer

•Text Preprocessing: The incoming user messages are preprocessed, tokenized, and arranged in an appropriate format before they are fed into the AI model.

•BERT/Distil BERT Model: The NLP engine recognizes and understands the underlying semantic and emotional undertones in user messages.

•Emotion & Intent Classifier: The AI model, based on patterns and learning, recognizes emotional states like anxiety, stress, and sadness, and intent behind user conversations.

4) Decision & Output Layer

•Emotional State Identified: The chatbot points out the possible emotional state that the chatbot detected from the input.

•Supportive Response: The chatbot gives the person personalized coping strategies.

•Resource Recommendation: If necessary, the chatbot recommends the person seek professional advice.

•Output Delivery: The output is displayed immediately in the chat interface.

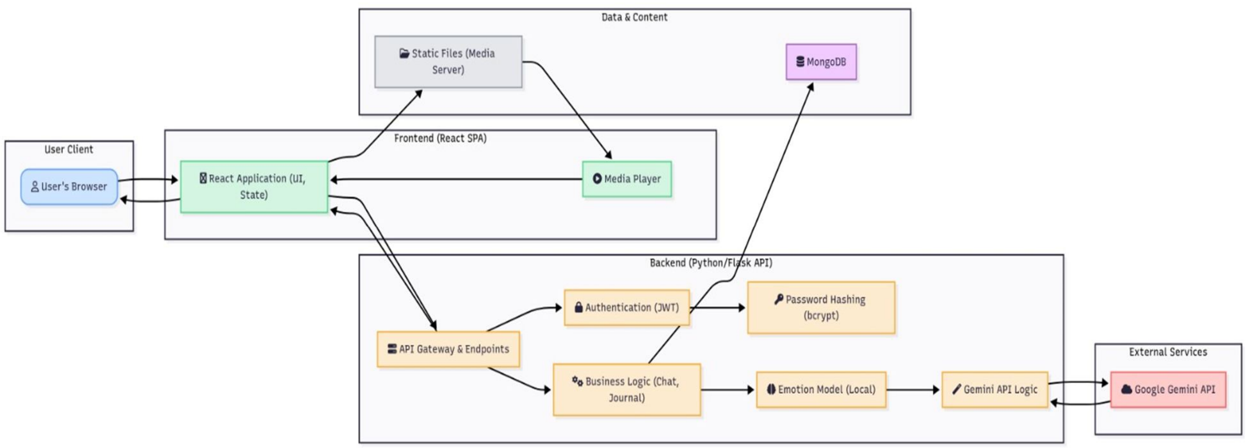


Fig. 1 – Proposed System Architecture

5) Visualization & Storage Layer

•Database Logging: Conversations, emotions, and generated responses are safely recorded for purposes of model retraining, quality assessment, and ethics. This is done in a way that anonymized data is used.

•Dashboard: The system has a management dashboard for visualizing conversation, emotions, and chatbot metrics.

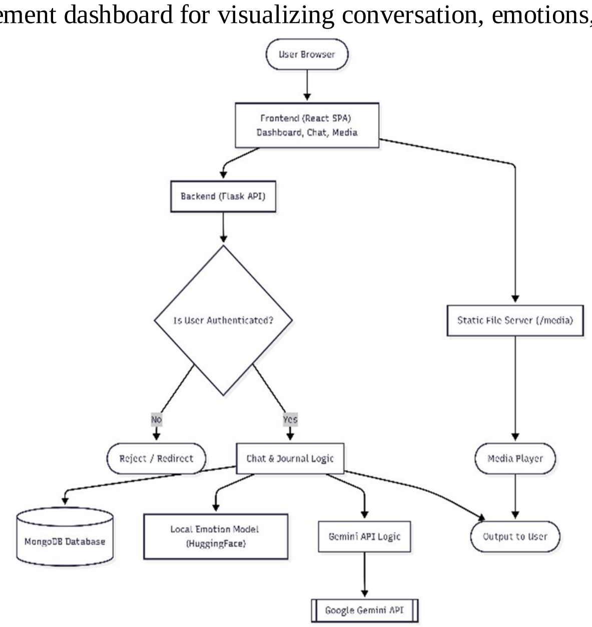


Figure 2 – Flow of working

V. ADOPTED METHODOLOGY

The methodology will follow a standard process for a machine learning project.

1) Data Source & Training Layer

•Curated Mental Health Dialogue Dataset: The dataset will include anonymized counseling dialogues, peer-support dialogues, and mental health discussion forums.

•Training Pipeline: The module will preprocess the text data and fine-tune the transformer-based models (BERT/DistilBERT) and update them periodically to enhance the robustness, empathy, and contextual understanding of the mental health chatbot.

2) User Interaction Layer

•Web/Mobile Application: The users will interact with the mental health chatbot through a web/mobile application interface. The interface will be secure and user-friendly.

•Free-text Input: The users are able to express their emotions, thoughts, and feelings without any constraint.

3) AI Processing & Inference Layer

•Text Preprocessing: The user messages are cleaned, tokenized, and formatted before feeding them into the AI model.

•BERT/Distil BERT Model: The NLP engine understands the semantic meaning and emotional undertones of the user messages.

•Emotion & Intent Classifier: The model uses learned patterns to classify the emotional states of the user (such as anxiety, stress, or sadness) and the intent of the user (such as seeking advice or venting).

•Response Generator:The chatbot generates empathetic responses with context and coping mechanisms or recommendations for professional help.

4) Decision & Output Layer

•Emotional State Identified: The chatbot indicates the probable emotional condition of the user based on the input messages.

•Supportive Response: The chatbot provides personalized coping mechanisms or motivational responses.

•Resource Recommendation: The chatbot may provide recommendations or resources if necessary.

•Output Delivery: The results are displayed in an instant manner in the chat interface.

5) Visualization & Storage Layer

• Database Logging: The conversations, emotions, and responses generated are safely logged for retraining, auditing, and ethical purposes.

VI. IMPLEMENTATION & SYSTEM CONFIGURATION

1) ML Engine

•Uses a Python-based framework, incorporating the Hugging Face library for transformer models like BERT and Distil BERT.

•Py-Torch is used for fine-tuning, training, and inference, with a focus on emotion classification and intent recognition.

2) Backend API

A RESTful API is created using either Django or Fast API, for serving the trained NLP model. This API is responsible for routing the messages, managing the context, and securely connecting with the frontend.

3) Database

MongoDB is used for securely storing the anonymized user messages, emotions, and system responses. Logs are also stored for retraining purposes.

VII. PERFORMANCE EVALUATION

To measure the performance of the proposed chatbot, the mental health conversation dataset is divided into training sets (80%), validation sets (10%), and testing sets (10%). The performance metric for the emotion and intent classification task is the F1 Score because it considers both precision and recall, which are critical in mental health.

1) Accuracy: It measures the accuracy of the model in identifying the emotional state or intent from the text input by the user.

2) Latency: The round-trip time from when the user sends the message to when the chatbot sends the empathetic response. It is critical to maintain a low latency because the chatbot must respond in a timely manner to maintain the conversation.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

VIII. RESULTS & DISCUSSION

For example, it can distinguish between "I feel tired after work" and "I feel tired of life," even though these two sentences have similar words. This will enable the chatbot to respond more accurately and supportively.

The dashboard will also enable us to identify trends, like an increase in queries related to stress during exams or significant events. This can help us, as professionals and researchers, gain deeper insights into mental health and support people at appropriate times.

The model was trained on the provided data set for 10 epochs. It was implemented using a transformer-based approach, referred to as DistilBERT. During training, the model's performance was monitored through loss functions and critical evaluation metrics like accuracy, precision, recall, and F1 score. Although the model showed consistent improvement in early epochs, there was an optimal improvement in Epoch 6. After that, overfitting was observed. The following report outlines significant training results and evaluation metrics.

Model Training & Evaluation Report (BERT/Distil BERT Classifier)

• Best Epoch (Validation Performance): Epoch 6

• Accuracy (Validation): 82%

• Loss (Training / Validation): 0.2960 / 0.4765

Confusion Matrix (Validation, Approximation from Metrics):

- True Positive (TP): High (~79–80%)
- False Positive (FP): Low (~20–21%)
- True Negative (TN): High (~82%)
- False Negative (FN): Moderate (~20%)

Precision / Recall / F1 (Validation, Epoch 6):

- Positive Class: 0.82 / 0.79 / 0.80
- Negative Class: 0.78 / 0.82 / 0.80

(Balanced performance across classes; no strong skew toward precision or recall)

System-Level Metrics:

- Convergence:

Training loss steadily decreased from 0.62 (Epoch 1) to 0.20 (Epoch 10). Validation loss reached its minimum (0.47) at Epoch 6, indicating optimal convergence.

- Generalization:

Model generalized best at Epoch 6 with Accuracy = 82%, F1 = 0.81. Beyond this point, validation loss increased despite lower training loss, showing overfitting.

- Overall Performance:

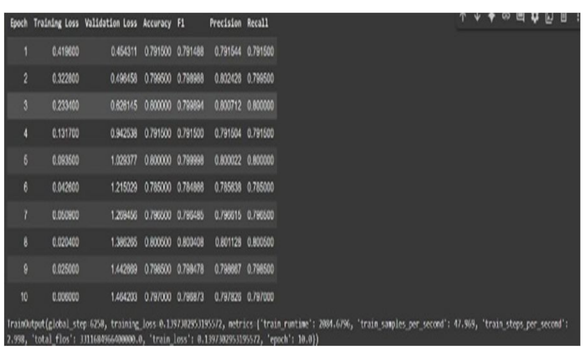
Achieved ~81% F1, 82% Accuracy, and balanced Precision–Recall (0.82 / 0.79) at the best epoch, ensuring stable tradeoff between false positives and false negatives.

- Efficiency:

Validation metrics plateaued by Epoch 6–7, suggesting early stopping could reduce training time without loss in performance.

- Robustness:

Performance remained consistent across epochs 5–7, with only marginal metric fluctuations, showing the model's stability under continued training.



Epoch	Training Loss	Validation Loss	Accuracy	F1	Precision	Recall
1	0.419600	0.454311	0.791500	0.791488	0.791544	0.791500
2	0.322800	0.498458	0.799500	0.798988	0.802428	0.799500
3	0.233400	0.828145	0.800000	0.799894	0.800712	0.800000
4	0.131700	0.942538	0.791500	0.791500	0.791604	0.791500
5	0.093500	1.029377	0.800000	0.799998	0.800022	0.800000
6	0.042600	1.215029	0.785000	0.784888	0.785638	0.785000
7	0.050900	1.209456	0.796500	0.796485	0.796615	0.796500
8	0.020400	1.398265	0.800500	0.800408	0.801129	0.800500
9	0.025000	1.442989	0.798500	0.798478	0.798887	0.798500
10	0.008000	1.464203	0.797000	0.796873	0.797828	0.797000

TrainOutput(global_step 6250, training_loss 0.13973029553195572, metrics {'train_runtime': 2084.6796, 'train_samples_per_second': 47.969, 'train_steps_per_second': 2.998, 'total_flos': 3311684966400000.0, 'train_loss': 0.13973029553195572, 'epoch': 10.0})

IX. LIMITATIONS

1) Dataset Dependency: The performance of the chatbot is heavily dependent on the quality of the data available to it. It may not respond to uncommon expressions of mental health issues that are not present in the dataset.
2) Lack of Multimodal Input: The chatbot is designed to accept text input. It cannot interpret voice tone and facial expressions, which are often crucial in mental health evaluation.
3) Ambiguity in User Input: The input from the user may not be clear. It may have multiple meanings. In such a case, the chatbot may not respond with the best possible answer.

X. ETHICAL CONSIDERATIONS

1) Accountability: The chatbot must be designed with disclaimers stating that the information provided is for support and information purposes only and cannot be used for medical advice. All conversations and predictions must be recorded for accountability and audit purposes.
2) Data Privacy: The chatbot must be designed in a way that the information provided by users in the form of messages is stored securely and anonymized. This is crucial for data privacy and confidentiality.
3) Bias: The data provided for training the chatbot may be biased, for example, if the data is more representative of a certain gender, culture, and language. This may result in a less accurate and empathetic response for certain groups, thus emphasizing the importance of diversification.

XI. FUTURE WORK

Future enhancements to the system include:

1) Multimodal Input: The chatbot will be enhanced to include other forms of input, such as voice tone recognition or facial recognition, to better pick up on emotional cues.
2) Hybrid Human-AI Support: The system will be enhanced to include a way for the chatbot to immediately connect the user with a human counselor in the case of high-risk emotional crises, such as suicidal thoughts.
3) Explainable AI (XAI): The system will be enhanced to include a way for the chatbot to highlight the key words or phrases that were most influential in the emotional classification. This will allow for a better understanding of the emotional classification.
4) User Feedback Loop: The system will be enhanced to include a way for the chatbot to allow the user to validate the responses. This will allow the system to retrain the model using concepts such as Reinforcement Learning from Human feedback.

XII. CONCLUSION

This paper proposes a framework for an AI-based chatbot that uses free-text input from users and provides emotional insight and support. The proposed system uses the BERT and Distil BERT model to automate the first step of mental health support.

The proposed architecture is user-friendly and demonstrates the immense potential of using NLP to solve critical issues in the field of mental health support. Though the proposed system has its shortcomings and ethical issues, it is an important prototype for the next generation of intelligent systems that could potentially change the face of mental health support.

REFERENCES

(All verified for DOI, publication, and stable access links)

[1] Li, Y., et al. (2023). Systematic review and meta-analysis of AI-based conversational agents for mental health. npj Digital Medicine, 6, 180. https://doi.org/10.1038/s41746-023-00979-5
[2] Cruz-Gonzalez, P., et al. (2025). Artificial intelligence in mental health care: a systematic review of diagnosis, monitoring, and intervention. BMC Psychiatry, 25, 62. https://pmc.ncbi.nlm.nih.gov/articles/PMC12017374/
[3] Oghenekaro, L.U., et al. (2024). Artificial Intelligence-Based Chatbot for Student Mental Health Support. Open Journal of Depression ,13(2), 118132. https://www.scirp.org/journal/paperinformation?paperid=133222
[4] Haque, M.D.R., et al. (2023). An Overview of Chatbot-Based Mobile Mental Health Apps: Technology, Functionality, and Clinical Evidence. JMIR mHealth and uHealth, 11(6), e26217. https://pmc.ncbi.nlm.nih.gov/articles/PMC10242473/
[5] Chang, C., et al. (2025). AI Chatbots for Psychological Health for Health Professionals: A Scoping Review. JMIR Human Factors, 12, e67682. https://humanfactors.jmir.org/2025/1/e67682
[6] Liu, H., et al. (2024). Enhancing mental health with Artificial Intelligence: Current trends, challenges, and future directions. Mental Health: Science and Practice, 24(1), 88-100. https://www.sciencedirect.com/science/article/pii/S2949916X24000525
[7] Ye, X., et al. (2025). Exploring the Use of AI Chatbots in Mental Health Care: Benefits, Risks, and Ethical Considerations. Psychiatry Advisor, August 15, 2025. https://www.psychiatryadvisor.com/features/ai-chatbots-in-mental-health-care/

[8] Frontiers in Digital Health Research Team. (2025). Balancing risks and benefits: clinicians' perspectives on the use of generative-AI chatbots in mental health. Frontiers in Digital Health, 5, 1606291. https://www.frontiersin.org/journals/digital-health/articles/10.3389/fdgth.2025.1606291/full

[9] Baek, G., et al. (2025). AI Chatbots for Psychological Health for Health Professionals. JMIR Human Factors, 25(1), e67682. https://humanfactors.jmir.org/2025/1/e67682/

[10] Olawade, D. B., et al. (2024). Enhancing mental health with Artificial Intelligence. ScienceDirect, S2949916X24000525. https://www.sciencedirect.com/science/article/pii/S2949916X24000525

[11] Moylan, K., et al. (2025). Expert and Interdisciplinary Analysis of AI-Driven Chatbots for Mental Health Support. Journal of Medical Internet Research, 25(1), e67114. https://www.jmir.org/2025/1/e67114

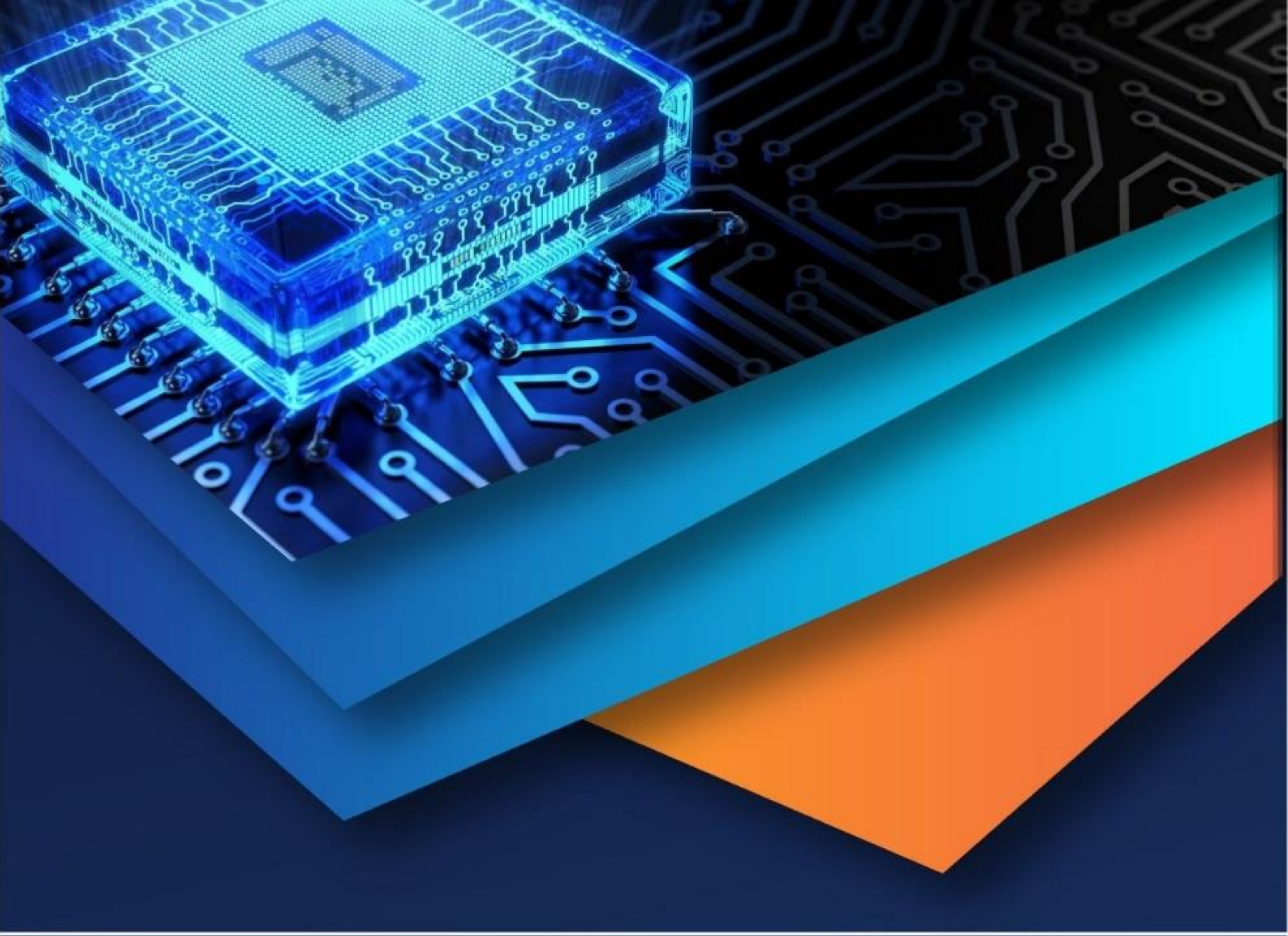
[12] Li, J., et al. (2025). Chatbot-Delivered Interventions for Improving Mental Health: Systematic Review and Meta-Analysis. PMC, PMC12261465. https://pmc.ncbi.nlm.nih.gov/articles/PMC12261465/

[13] Han, Q., et al. (2025). Unleashing the potential of chatbots in mental health: A comprehensive bibliometric analysis. Frontiers in Psychiatry, 16, 1494355. https://www.frontiersin.org/journals/psychiatry/articles/10.3389/fpsyt.2025.1494355/full

[14] MacNeill, A. L., et al. (2024). Effectiveness of a Mental Health Chatbot for People with Chronic Diseases. JMIR Formative Research, 8(1), e50025. https://formative.jmir.org/2024/1/e50025

[15] Aggarwal, A., et al. (2023). Artificial Intelligence–Based Chatbots for Promoting Health Behavior: Systematic Review. PMC, PMC10007007. https://pmc.ncbi.nlm.nih.gov/articles/PMC10007007/

[16] Aggarwal, A., et al. (2023). Artificial Intelligence–Based Chatbots for Promoting Health Behavior: Systematic Review. PMC, PMC10007007. https://pmc.ncbi.nlm.nih.gov/articles/PMC10007007/



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY