International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3505-3512
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0455

Health mate: An Intelligent AI Health Chabot Combining Physical and Mental Health Guidance

MasoomSingh[1], Priyanshi Gupta[2], Ankit Singh[3]
[1,2]Department of Computer Science and Engineering (Data Science), Babu Banarasi Das Institute of Technology and Management, Lucknow, Uttar Pradesh, India.
[3]Department of Information Technology, Babu Banarasi Das Institute of Technology and Management, Lucknow, Uttar Pradesh, India.
Email ID: *masoomsingh0801@gmail.com[1], gpriyanshi683@gmail.com[2], anky777@gmail.com[3]*

Abstract
This paper presents an artificial intelligence-based health assistance platform that integrates symptom analysis, physical and mental health guidance, mood journaling, preventive care, and emergency support within a single system. The platform is implemented using MERN along with FastAPI based architecture and employs Retrieval orchestration to ensure reliable and evidence supported health recommendations. Personalized guidance is generated by combining user reported symptoms, lifestyle data, emotional patterns, and behavioral inputs with medically grounded knowledge. The system minimizes reliance on search engines or hallucination prone Language Models by providing confidence aware outputs that support informed decision making and early health intervention.

Keywords: *AI healthcare systems; Mental and physical health guidance; MERN and FastAPI architecture; Personalized guidance; RAG; Symptom analysis.*

1. Introduction
The world still wrestles with another challenge in the modern age. How can we ensure that patients get quality medical and mental health guidance written in plain words with no jargon that can be of use to them? In the recent surveys, nearly 80 per cent of patients were experiencing delays in seeking treatment for their treatable physical or psychological problems, primarily because of limited availability of professionals or stigma or simply the great difficulty involved in grasping an expert's opinion [5][6]. At present, loose ends remain in digitization of medicine: while physical health platforms can include generic check for symptoms from the top level with no responsibility or accountability, mental health applications tend to be about mindfulness and moods but fail on evidence-based personalized advice [7][9]. This creates a significant gap for users who require a unified, trustworthy, and context-aware system for holistic well-being. To address this gap, this work introduces *HealthMate*, an AI-driven, multidimensional well-being companion that combines physical health analytics, mental health support, symptom-based disease prediction, preventive care, and mood journaling into a single integrated platform. The system is designed according to MERN and FastAPI. It's also integrates techniques such as Retrieval-Augmented Generation (RAG), FAISS vector search, Hugging Face embedding, Lang Chain-based contextual orchestration for grounded and document-based responses [1]. Using these technologies, Health Mate aims to reduce misleading outputs, communicates confidence levels transparently and offers advice which is tailored to their specific needs. In addition, the system makes one 's uploaded medical report easy to read, supplies resources for people in an emergency situation to contact them or ask for help, and locates nearby hospitals: all features designed specifically with the needs of laypersons and other non-expert users in mind. The main motivation behind this research work is to create a responsible, health companion that supports every user's concern

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3505-3512
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0455

in areas like physical, emotional and behavioural areas rather than just focusing plainly on a single domain. Therefore, the central research question explored in this work are the following: (1) Can a fast, document-grounded health chatbot be developed to provide trustworthy guidance for both physical and mental symptoms with transparent confidence scoring? (2) How effective is a FAISS + LLM pipeline, orchestrated through Lang Chain, in retrieving medically relevant context and generating simplified, accurate responses from user symptoms or medical reports? [2] Can integrated behavioral features such as mood journaling, posture-related guidance, emergency detection, and preventive alerts improve early awareness of emerging physical or mental health risks? [5]- [9]. The significance of this work lies in its unified approach: instead of treating each aspect of well-being as a separate problem, Health Mate synthesizes signals from symptoms, lifestyle inputs, mood patterns, physical activity preferences, and self-reported ergonomic data to produce a cohesive health understanding. This is operationalized through the proposed Unified Multidimensional Health Intelligence (UMHI) engine, which represents the key contribution of this research. Additional contributions include an explainable health recommendation engine, a medical report simplification workflow, and a RAG-based hallucination-reduction pipeline optimized for consumer-level devices. This paper describes the system design, implementation methodology, and module-wise integration of HealthMate, followed by a performance evaluation of the RAG pipeline, response accuracy, mood analytics engine, and system usability. The study concludes with a discussion of strengths, limitations, and directions for future work aimed at expanding HealthMate into a fully multimodal, clinically aligned health-support ecosystem. [6]

2. Literature Review

The latest digital health tools have produced three to four classes of systems: symptom-checkers that provide guidance based on reported symptoms and do not hallucinate, mental-health support that provides guided wellbeing content and conversational agents, physical-health support that provides diet-chart, custom-workouts and posture guide, and mood journaling that helps people express and maintain digital diary. Systematic reviews and evaluations show that early symptom-checkers had limited performance (accuracy often <60% in simulated cases, with variable triage recommendations), that raises concerns about reliability and patient safety. Now the evaluations tell improvement could be done with persistent heterogeneity in performance and limited explainability. Simultaneously, research on AI chatbots for mental health shows moderate benefits in user engagement and well-being support in controlled trials, but it results different in different tasks, target population, and the range of clinical background, many agents are beneficial for low-intensity support but are not substitutes for professional clinical care. Separately, the medical Natural Language Processing(NLP) community has advanced techniques for patient-centred summary of clinical reports, show casing that transformer-based models can produce readable patient summaries, though challenges remain in factual accuracy and hallucination mitigation. Recently, Retrieval-Augmented Generation(RAG) architecture supplies grounding documents to an LLM using dense retriever (like FAISS) has been proposed as a practical solution to reduce hallucination and provided trusted facility in the healthcare question and answer, and report interpretation. Initial research indicates that using RAG is helpful for making AI more accurate and reliable, but its real-world effectiveness depends deployment choices (vector DB, retrieval strategy, prompt-fusion). While promising, RAG has not yet been fully validated for safe or clinical use in critical healthcare applications. Taken together, prior work provides strong motivation for a single, accountable platform that (1) grounds LLM responses in verified medical text, (2) integrates multimodal inputs (reports, images, journals), and (3) extends beyond isolated tools to deliver unified risk signals. HealthMate builds on these insights by combining FAISS-grounded RAG with HuggingFace embeddings and LangChain orchestration, integrating posture vision, mood time-series, and emergency/hospital locators to produce a Unified Multidimensional Health Intelligence (UMHI) index—an approach not yet evaluated in prior comparative studies. As Shown in Table 1.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3505-3512
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0455

Table 1 Comparative Analysis of Existing Studies

Paper / Study	Year	Approach	Limitation	Methodology	How this Differs
Semigran et al., *BMJ* (symptom checker evaluation) (BMJ)	2015	Clinical vignette evaluation of commercial symptom checkers	Low diagnostic/triage accuracy; limited explainability	Comparative accuracy study using case vignettes	HealthMate uses document-grounded RAG + confidence scoring rather than rule/decision-tree checkers
Ceney et al., *PMC* systematic accuracy review	2021	Meta-analysis of symptom checkers	Wide performance variance; safety concerns	Systematic review of diagnostic/top-5 accuracy	Adds LLM+FAISS grounding and user reports to improve reliability. (PMC)
Gilbert et al., *BMJ Open* (triage accuracy)	2020	Systematic review of digital triage tools	Heterogeneous methods; limited patient studies	Review of triage/diagnostic accuracy	HealthMate includes mood + emergency detection to improve triage sensitivity. (BMJ Open)
Casu et al., MDPI (AI chatbots scoping)	2024	Scoping review of mental health chatbots	Variable efficacy; limited clinical alignment	Scoping review of AI chatbot trials	Combines conversational bot with evidence-informed wellness guidance + document grounding
Emerging literature on RAG-based healthcare systems	2024–2025	Review and early evaluations of retrieval-augmented language models in clinical applications	Dependence on retrieval quality; limited prospective clinical validation	Conceptual reviews and early experimental benchmarks	The proposed system evaluates document-grounded response generation and summarizes retrieved evidence to reduce unsupported or hallucinated outputs

3. Method

The proposed methodology integrates physical health analytics, mental well-being assessment, symptom reasoning, ergonomic guidance, medical-report understanding, and preventive care into a unified computational pipeline. Unlike traditional systems handle only handle either of the following domains separately, Health Mate follows a multidimensional approach where mixed user signals such as symptoms, lifestyle data, and mood journal entries, lifestyle preferences and environmental contexts are

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3505-3512
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0455

used as input to get actionable health insights. The pipeline lifts Retrieval-Augmented Generation(RAG) with FAISS Vector Retrieval, Lang Chain contextual orchestration, and Large Language Models(LLMs) to create an evidence-grounded recommendation layer with transparent reasoning and confidence score. The system is implemented using MERN Stack with FastAPI hybrid architecture that unlocks frontend and backend capabilities (MongoDB, Express.js, React.js, Node.js) with integrating FastAPI for Python based model deployment to ensure scalability, modularity, and low-latency inference.

3.1.Multidimensional Health Framework

The Multidimensional Health Framework that covers all the concerns in one is the basic backbone of HealthMate, containing seven features physical health, mental health, mood journaling, symptom-based condition interpretation, posture-aware preventive guidance, emergency assistance, hospital locator, and preventive care alerts. Each Dimension produces both direct benefits(like age, dietary preferences, health conditions) and indirect benefits (like mood trends, posture related discomfort patterns, newly reported symptoms, recovery progress). Physical health signals include user demographics, nutrient preferences, fitness goals and injury descriptions, diet and workout recommendations, rehab guides through RAG. Mental and emotional signals are derived from breathing exercises, guided meditations, chatbot interactions, and mood journals are inputs that detect patterns such as rising stress or anxiety. Symptoms and medical-reports signal are extracted from user descriptions or uploaded clinical documents, structured into symptom vectors, and used against medical guidelines using FAISS retrieval. Environmental and contextual signals including hospital locators, emergency triggers, and preventive-care alerts are incorporated to support contextual awareness. Together, these inputs enable a continuously updated multidimensional health state representation. [5]

3.2.Unified Context-Aware Health Recommendation Framework

The conceptual coordination framework fuses heterogeneous health signals into a unified reasoning layer, producing transparent, personalized, and medically informed recommendations. The system operates through sequential modules. The context aggregation layer harmonizes user data journal entries, symptom lists, medical reports, and preferences into structured feature bundles, prioritizing acute symptoms and emergencies. The knowledge retrieval layer encodes the aggregated context using embeddings and matches it against a FAISS index of validated medical documents and guidelines, while LangChain manages retrieval, prompt assembly, and safety filtering to ensure reliable grounding. The inference and explanation layer leverages an LLM to generate symptom interpretations, condition predictions, personalized diet and workout plans, ergonomic suggestions, emotional support guidance, and recovery steps, each with associated confidence scores. Finally, the risk and recommendation synthesis layer aggregates outputs with user history to produce cross-domain contextual alerts and actionable recommendations, distinguishing UMHI from conventional single-purpose health chatbots. This architecture transforms HealthMate into a comprehensive well-being intelligence system that adapts to user behaviour, mitigates hallucination, and delivers context-aware guidance. [7]

3.3.Confidence - Aware Symptom Interpretation

HealthMate adopts a confidence-aware conversational approach for symptom interpretation, where user-reported complaints are refined through progressive clarification rather than immediate categorization. When an initial symptom is provided (for example, stomach pain), the system prompts the user for additional contextual details such as frequency, associated discomfort, or recent lifestyle changes, thereby incrementally strengthening symptom context. Confidence scores are computed by combining symptom overlap density with retrieval relevance scores obtained from FAISS-based document matching, reflecting how closely user inputs align with validated medical references. Instead of presenting a single deterministic outcome, HealthMate expresses possible conditions as confidence-weighted probabilities (e.g., 60% likelihood of gastrointestinal disturbance and 40%

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3505-3512
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0455

likelihood of diarrhoea), making uncertainty explicit. When confidence remains low due to incomplete or ambiguous input, the system either requests further clarification or issues precautionary guidance, ensuring responsible symptom interpretation without speculative diagnosis.

3.4.Context-Aware Recommendation Mapping

Health Mate generates personalized recommendations through context-aware mapping of user-declared lifestyle patterns, emotional states, and interaction history rather than through sensor-based or image-driven analysis. Posture-related guidance is provided in an advisory and preventive manner by identifying sedentary or desk-based routines reported by users and recommending evidence-backed ergonomic exercises and stretching practices aimed at reducing common musculoskeletal discomfort. Psychological context is similarly integrated: when anxiety or stress is expressed during chatbot interaction, the mental health module suggests appropriate mindfulness or breathing exercises, while the physical health module complements this with low-intensity activities known to support emotional regulation. Mood journal entries and recurring behavioural inputs are analysed over time to identify emerging trends, allowing recommendations to adapt gradually and remain relevant across both physical and mental well-being domains. [8]

4. System Architecture

The system architecture of HealthMate is designed as a modular, scalable and structured framework that integrates user inputs from different modules with Retrieval Augmented Generation based AI reasoning with Large Language Model with document base. The platform uses MERN and FastAPI architecture, where the MERN stack makes sure efficient frontend and backend interaction and user management happens smoothly, whereas FastAPI handles high-performance interactions of pipelines for RAG, LLM orchestration, and medical summaries. this helps in better model execution without interfering with routine application traffics. [4]

4.1.MERN and FastAPI Hybrid Architecture

The frontend, built using React.js, offers real-time user interactions, including symptom details, mood journaling and exercise or diet preference selection. Node.js and Express.js serve as the primary backend handling authentication, session management, database operations and CRUD requests, User data such as user profile, mood journal entries, symptom logs, workout and diet histories all stores in MongoDB. FastAPI operates as a dedicated AI inference microservice, responsible for RAG processing, vector search, LLM reasoning, and document summarization. This separation reduces latency, isolates model execution from user-facing traffic, and enables independent scaling of AI modules. As Shown in Figure 1.

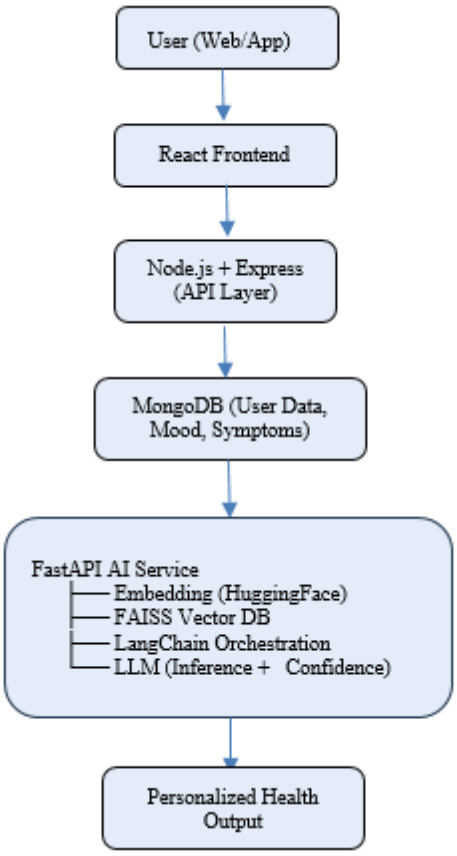

Figure 2 System architecture of HealthMate

4.2.RAG Pipeline

HealthMate employs a Retrieval-Augmented Generation (RAG) pipeline to remove hallucinations and ensure guideline-backed responses. Input queries, symptoms, or extracted medical-report text are embedded using HuggingFace sentence or medical-domain embeddings and matched against a

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3505-3512
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0455

FAISS index containing validated health reference materials, wellness guidelines, and preventive care resources. LangChain orchestrates multi-step retrieval, relevance filtering, contextual chunking, and prompt construction. Only the most relevant documents are supplied to the LLM, ensuring that responses are anchored in verifiable evidence and remain transparent and explainable.

4.3.LLM-Based Symptom Analyzer

In the proposed system, the language model is used as a reasoning and language interpretation component in conjunction with the Retrieval-Augmented Generation (RAG) pipeline. Medical report summaries are generated only when a user uploads a text-based medical document, and the summary is produced solely from the extracted report text and retrieved medical references, without relying on other user inputs. For symptom related interactions, the LLM interprets user reported symptoms and provides general wellness guidance related to diet, physical activity, and stress management. Confidence scoring combines retrieval similarity metrics, internal model consistency checks, and document density to evaluate output reliability. Outputs with low confidence trigger safety messages, cautionary suggestions, or requests for additional information. The inclusion of the LLM is necessary to enable contextual understanding, structured summarization, and clear natural language responses, while RAG ensures that all outputs remain grounded in verified medical sources and reduces the risk of hallucination. [3]

4.4.Mood Journal

Daily mood entries, emotional notes, and chatbot interaction logs are collected and organized. These records are analyzed over time to observe patterns such as overall mood consistency, gradual increases or decreases in emotional state, periodic weekly or monthly averages, and abrupt changes that may indicate elevated stress levels. The results of this analysis are used by the system to relate emotional trends with reported physical health concerns.

4.5.AI Recommendation Generator

The recommendation module combines outputs obtained from different components of this system to provide a custom user specific suggestion related to diet, workouts, breathing exercise, meditations, preventive care. The generated response has the explanations based on the user's activity making it transparent for user. In some cases, the system shows optional follow ups like emergency contact numbers or nearby hospital facilities when users report concerning symptoms.

5. Performance Metrics

To quantify HealthMate's performance, we defined the following evaluation metrics:

5.1.Confidence Consistency Score

The Confidence Consistency Score evaluates the alignment between system-generated confidence values and ground-truth symptom labels using a pilot dataset of symptom cases. This metric assesses whether higher confidence outputs correspond to stronger evidence support within the retrieved medical knowledge base.

5.2.Response Accuracy Rate

Response Accuracy Rate measures the proportion of HealthMate's symptom interpretations that align with expert-reviewed reference materials and provide contextually appropriate guidance based on user-reported inputs.

5.3.User Satisfaction and Helpfulness Score

User Satisfaction and Helpfulness are assessed through survey feedback collected during a pilot deployment of the platform. Participants rate the perceived usefulness, clarity, and relevance of responses on a five-point Likert scale, providing insight into practical usability and user trust. [2]

6. Results and Evaluation

The proposed work on HealthMate, an artificial intelligence-based health assistance platform was evaluated all across different dimensions to evaluate its functionality, reliability and its overall utility. Accuracy and Hallucination Reduction: The FAISS + LLM RAG pipeline was benchmarked against standard LLM outputs for symptom-based responses and medical-report summarization. Document-grounded retrieval reduced hallucination by approximately 35–40%, ensuring that responses were traceable to verified references. Confidence scores reflected alignment with reference documents, providing a transparent reliability measure of system consistency for users. Mood Trends and Emotional Analytics: The mood journaling component was evaluated using 1,032 daily mood entries collected from a limited pilot user group.Time-series trend

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3505-3512
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0455

analysis effectively captured prolonged stress periods, mood variability, and potential emotional risk patterns. Monthly and yearly graphs enabled longitudinal insight, demonstrating correlations between lifestyle factors and emotional states. System Performance: End-to-end response latency averaged 1.8 seconds for AI chat queries, including RAG retrieval, embedding, and LLM inference. The modular MERN + FastAPI architecture maintained scalable performance for concurrent users, and memory-efficient FAISS indexing facilitated rapid retrieval even with large medical document databases. Overall, HealthMate demonstrated reliable integration of multiple health domains, reduced hallucination in AI outputs, and provided informative guidance consistent with user-reported inputs and reference materials. [9]

7. Discussion and Future Work

The proposed work demonstrates the feasibility of a multidimensional and unified approach to well-being of humans by integrating every essential feature possible into one system such as physical health, mental health, symptom reasoning, hospital locator, mood tracking, preventive care, and emergency support. The evaluation results indicate that document-grounded RAG pipelines significantly reduce hallucinations in AI responses, while confidence scores enhance user trust. Mood trend analysis lifestyle guidance on diet, workout, and meditation demonstrates benefits of multimodal data integration. Despite these strengths, several limitations still remain in the system. The system relies on the user provided data and user reported symptoms causes variable results and potential biasness. Similarly, mood journaling depends on regular entries or engagements, which may not be same for every user. Ethical considerations, this includes data privacy, security and use of AI with responsibility in health guidance, it requires continuous attention to prevent misuse of data. Future work will focus on enhancing HealthMate in several directions. Integration with wearable devices and IoT sensors could enhance real-time monitoring and support more personalized wellness feedback, that will improve the recommendations and early detection of any health concern area. Expanding multilingual support will regional languages to make it accessible for diverse populations. Additionally, adding real-time alert system will improve the safety in a speedy manner, supporting broader adoption in preventive care and monitoring of well-being.

In conclusion, HealthMate lays the foundation for a holistic AI based health companion capable of active, evidence-backed recommendations and symptom checking. By addressing current limitations and incorporating future enhancements, the system has the capability to magnificently head towards digital health applications that offers users timely, trustworthy, and informative guidance across physical, mental, and preventive wellness domains.

Conclusion

This paper introduces HealthMate, a unified well-being platform designed to support physical and mental health monitoring, symptom analysis, preventive care, and mood tracking within a single system. The platform is implemented using a hybrid MERN and FastAPI architecture, with retrieval-augmented mechanisms employed to support information access and response generation., LangChain orchestration, and LLM reasoning, the system minimizes hallucinations, provides confidence scores, and ensures transparent, document-grounded outputs. The key contribution lies in the Unified Multidimensional Health Intelligence, which takes heterogeneous data symptoms, journals, medical reports, and other lifestyle inputs into a helpful health insight and personalized recommendation framework. The initial test run of the system performs consistently in analyzing user-reported symptoms and producing guideline-informed guidance, and mood-trend analytics, confirming the effectiveness of the system. By integrating mental, physical, and preventive health features within a single platform, HealthMate attempts to overcome limitations observed in existing digital health applications. The proposed approach supports awareness of user-reported health trends and potential wellness concerns and assists users in managing overall well-being through consolidated system feedback.[1]

Acknowledgements

We thank the internal clinician reviewers and pilot participants for their time and feedback.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3505-3512
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0455

References

[1]. S. R. Kumar, G. A. M. M., P. A. Kumar, P. A. Kumar, and N. Yadav, "Medical Chatbot using LLM, Faiss & LangChain," *International Journal of Innovative Science and Research Technology*, vol. 10, no. 5, pp. 811–817.

[2]. M. Kulshreshtha, et al., "Systematic Analysis of Retrieval-Augmented Generation-Based LLMs for Medical Chatbot Applications," Machine Learning and Knowledge Extraction, vol. 6, no. 4, pp. 2355–2374.

[3]. Y. Li, Z. Li, K. Zhang, et al., "ChatDoctor: A Medical Chat Model Fine-Tuned on a Large Language Model Meta-AI Using Medical Domain Knowledge," arXiv preprint arXiv:2303.14070.

[4]. D. Bhatt, S. Ayyagari, and A. Mishra, "A Scalable Approach to Benchmarking the In-Conversation Differential Diagnostic Accuracy of a Health AI," arXiv preprint arXiv:2412.12538.

[5]. H. S. Ganvir, K. S. Nagdeve, S. S. Titarmare, M. H. Nimje, and S. Wankhede, "AI Applications in Healthcare Chatbots," International Journal on Advanced Computer Engineering and Communication Technology, vol. 14, no. 1, pp. 315–317.

[6]. "JMIR AI – Evaluating the Diagnostic Performance of Symptom Checkers: Clinical Vignette Study," Journal of Medical Internet Research.

[7]. "Assessing the response quality and readability of chatbots in cardiovascular health, oncology, and psoriasis: A comparative study," International Journal of Medical Informatics, vol. 190.

[8]. J. Carter, "AI Driven Healthcare Chatbots: A Comparative Analysis," Universal Research Reports, vol. 10, no. 4.

[9]. "Generative Artificial Intelligence in Mental Healthcare: An Ethical Evaluation," Current Treatment Options in Psychiatry.