IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429 Volume 6, Issue 5, April 2026 Impact Factor: 8.2

AI-Driven Public Healthcare Chatbot for Disease Awareness

Rehan Inamdar , Akshay Adhav, Vishal Fulari, Ritesh Gaikwad, Vishvesh Arote, Tushar Sonawane

Department of AIML Engineering

Assistant Professor, Department of AIML Engineering

Sanjivani University, Kopargaon, India

Abstract: *The rapid expansion of digital health platforms has created a pressing opportunity to deliver reliable, accessible disease awareness to the general public. This paper presents the design and development of an AI-driven public healthcare chatbot capable of providing symptom awareness, preventive health guidance, and disease-related information through natural language interaction. The proposed system integrates state-of-the-art Natural Language Processing (NLP) techniques, including intent classification and named entity recognition, with a curated medical knowledge base to facilitate meaningful, context-sensitive conversations. A hybrid architecture combining rule-based logic and machine learning-based response generation ensures both accuracy and conversational fluency. The system supports multilingual input, making it broadly accessible across diverse linguistic communities. Evaluation using standard metrics — Precision, Recall, F1-Score, and Mean Reciprocal Rank (MRR) — demonstrates strong performance across a range of healthcare-related queries. The chatbot was validated through user testing with a representative sample of general public users, yielding favorable usability scores. This work contributes a scalable, cost-effective solution to bridge the gap between healthcare professionals and the public at large.*

Keywords: Artificial Intelligence, Healthcare Chatbot, Natural Language Processing, Disease Awareness, Multilingual Support, Symptom Guidance, Deep Learning

I. INTRODUCTION

The intersection of artificial intelligence and public health has given rise to transformative tools that promise to democratize access to medical information. Despite significant advances in healthcare infrastructure, a considerable portion of the global population still lacks timely access to basic disease-related guidance. According to the World Health Organization, preventable diseases continue to claim millions of lives annually, often due to late detection and inadequate awareness (WHO, 2022). Chatbots powered by AI offer a scalable, always-available alternative that can serve individuals regardless of their geographic location or socioeconomic status.

Existing healthcare information systems are frequently fragmented, difficult to navigate, and not designed with the general public in mind. Most medical websites present information in technical language that is inaccessible to individuals without formal health education (Bickmore et al., 2018). There is therefore a growing need for an intelligent, conversational agent that can translate complex medical knowledge into accessible, personalized guidance.

This paper proposes an AI-driven healthcare chatbot specifically designed for disease awareness and basic health guidance. The system leverages Natural Language Processing (NLP) to understand user queries and respond with contextually appropriate information. Key features include a symptom awareness module, preventive health tips, a frequently asked questions (FAQ) engine, and multilingual support to cater to a linguistically diverse user base.

The remainder of this paper is organized as follows: Section II reviews related work in AI-based healthcare chatbots. Section III describes the proposed system architecture. Section IV details the implementation methodology. Section V presents the evaluation results. Section VI discusses the findings, and Section VII concludes the paper with directions for future work.

II. LITERATURE REVIEW

The domain of conversational agents in healthcare has received considerable scholarly attention over the past decade. Early systems relied primarily on decision-tree logic and pattern-matching algorithms that limited their ability to handle open-ended queries (Woebot Health, 2017). While such rule-based approaches offered predictable, structured responses, they lacked the adaptability required for real-world health conversations.

The emergence of neural language models has substantially improved the quality of chatbot interactions. Devlin et al. (2019) introduced Bidirectional Encoder Representations from Transformers (BERT), which enabled machines to understand context in a more human-like manner. Subsequent healthcare-specific adaptations, such as BioBERT (Lee et al., 2020), demonstrated that domain-specific pre-training yields significant performance improvements on medical NLP tasks including named entity recognition and question answering.

Palanica et al. (2019) conducted a systematic survey of existing healthcare chatbots and noted that most tools focused narrowly on mental health or medication reminders, leaving disease awareness — particularly in low-resource language settings — underserved. Similarly, Laranjo et al. (2018) highlighted that multilingual functionality remains a critical gap, since the majority of deployed health chatbots operate exclusively in English.

More recent work by Abd-Alrazaq et al. (2020) evaluated the effectiveness of AI chatbots in managing chronic conditions and found that consistent, timely information delivery improved patient adherence and awareness. However, these systems were largely confined to clinical settings with structured datasets rather than open-domain public health contexts.

The proposed system addresses these shortcomings by combining transformer-based NLP with a curated health knowledge base and multilingual translation support, thereby serving a broader demographic without sacrificing accuracy or conversational quality.

III. PROPOSED SYSTEM ARCHITECTURE

The architecture of the proposed healthcare chatbot is organized into four principal layers: the User Interface Layer, the NLP Processing Layer, the Knowledge Base Layer, and the Response Generation Layer. Together these layers form a coherent pipeline that transforms a raw user query into a contextually appropriate health-related response.

A. User Interface Layer

The front-end of the system is implemented as a responsive web and mobile application built using the React.js framework. Users interact with the chatbot through a text-based messaging interface that supports both typed input and voice-to-text conversion. The interface is designed to be intuitive and accessible, following Web Content Accessibility Guidelines (WCAG 2.1) to accommodate users with disabilities. A language selector allows users to choose their preferred language, which dynamically adjusts both input processing and output delivery.

B. NLP Processing Layer

Upon receiving a user message, the system passes it through a preprocessing pipeline that includes tokenization, stopword removal, and lemmatization. The cleaned tokens are then fed into a fine-tuned BERT-based intent classification model that categorizes the query into one of several predefined intents: symptom inquiry, disease information request, preventive advice, FAQ, and general greeting. For each identified intent, a named entity recognizer (NER) extracts key medical entities such as disease names, body parts, and symptoms using a BioBERT-based tagging model trained on the MedMentions dataset (Mohan and Li, 2019).

Multilingual support is implemented through a language detection module that identifies the user's input language and routes it through the mBART-50 multilingual translation model (Tang et al., 2020) before NLP processing. The translated query is processed in English, and the response is translated back to the user's original language before delivery.

C. Knowledge Base Layer

The system's knowledge base consists of two components: a structured disease ontology and an unstructured document corpus. The ontology encodes relationships between diseases, symptoms, risk factors, and preventive measures using

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 5, April 2026
Impact Factor: 8.2

the Resource Description Framework (RDF) standard, informed by established medical taxonomies such as ICD-11 (World Health Organization, 2019) and SNOMED CT. The unstructured corpus contains curated articles from peer-reviewed health journals and reputable public health organizations, which are indexed using Elasticsearch for efficient retrieval. A retrieval-augmented generation (RAG) strategy is employed to dynamically fetch relevant passages, enhancing the contextual richness of responses (Lewis et al., 2020).

D. Response Generation Layer

Response generation combines template-based and generative approaches. For structured queries such as FAQs and symptom checklists, the system retrieves pre-authored responses from the knowledge base and populates them with entity-specific details. For open-ended queries, a fine-tuned GPT-based generative model produces fluent, contextually grounded responses, conditioned on the retrieved knowledge passages. A safety filter, implemented as a binary classifier trained on a dataset of harmful medical advice examples, screens all generated responses before delivery to ensure that no dangerous or misleading information is presented to the user.

IV. METHODOLOGY

A. Dataset Collection and Preparation

Training data for the intent classifier was assembled from three publicly available sources: the MSMACRO health question dataset (Bajaj et al., 2016), the HealthQA dataset (Zhu et al., 2019), and a custom dataset of 5,200 domain-specific question-answer pairs curated by the research team in collaboration with a certified medical practitioner. Data augmentation techniques including synonym substitution and back-translation were applied to address class imbalance across intent categories.

For the NER module, the MedMentions corpus was combined with a custom annotation dataset of 1,800 sentences describing disease symptoms, which were labeled using the BRAT annotation tool. The multilingual translation component was validated against a bilingual health glossary covering English, Hindi, Marathi, and Tamil.

B. Model Training

The intent classification model was initialized with BERT-base-uncased weights and fine-tuned for five epochs on the combined training corpus using the AdamW optimizer with a learning rate of 2×10^{-5} and a batch size of 32. The NER model was fine-tuned from BioBERT-base-cased weights under similar hyperparameter settings. All models were trained on an NVIDIA Tesla T4 GPU instance using PyTorch 2.0 and the HuggingFace Transformers library.

The generative response model was adapted from the GPT-2 medium architecture using a retrieval-augmented training procedure. During training, retrieved passages were prepended to the input context to encourage the model to ground its outputs in factual health information. Early stopping with a patience of three epochs was applied based on validation perplexity.

C. System Integration

The back-end services are containerized using Docker and orchestrated with Kubernetes, enabling horizontal scalability. A RESTful API built on Flask serves as the communication bridge between the front-end interface and the NLP processing pipeline. Session management is handled using JWT-based authentication, and all user query logs are anonymized before storage in compliance with data privacy regulations. The complete system was deployed on a cloud infrastructure supporting up to 500 concurrent users without performance degradation.

V. RESULTS AND EVALUATION

The system was evaluated across three dimensions: intent classification accuracy, information retrieval quality, and overall user satisfaction. Table II summarizes the performance metrics obtained on the held-out test set comprising 1,040 samples.

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429 Volume 6, Issue 5, April 2026 Impact Factor: 8.2

TABLE II: PERFORMANCE METRICS OF THE PROPOSED CHATBOT

Module	Precision (%)	Recall (%)	F1-Score (%)	MRR
Intent Classification	92.4	91.7	92.0	—
Named Entity Recognition	89.1	87.6	88.3	—
Information Retrieval (RAG)	—	—	—	0.84
FAQ Response Matching	94.3	93.1	93.7	0.91
Multilingual Translation	91.0	90.2	90.6	—

The intent classifier achieved an F1-Score of 92.0%, which represents a notable improvement over the baseline TF-IDF with SVM approach (F1: 81.3%) and is competitive with more resource-intensive Transformer architectures. The NER module attained an F1-Score of 88.3%, demonstrating reliable extraction of medical entities across diverse query formulations.

The retrieval-augmented generation pipeline yielded a Mean Reciprocal Rank of 0.84 on the information retrieval benchmark, indicating that the most relevant health passage was consistently ranked among the top responses. FAQ matching, benefiting from the structured knowledge base, achieved the highest F1-Score of 93.7%, confirming that template-based retrieval remains highly effective for common health inquiries.

A user study was conducted with 80 participants recruited from a general public sample, none of whom possessed formal medical training. Participants interacted with the chatbot over a 20-minute session involving a standardized set of ten health-related scenarios. Post-session usability was measured using the System Usability Scale (SUS), yielding a mean score of 82.4 out of 100, which falls in the 'excellent' category (Bangor et al., 2009). Qualitative feedback highlighted the chatbot's conversational tone, clarity of explanations, and the convenience of multilingual interaction as primary strengths.

VI. DISCUSSION

The results collectively affirm that the proposed architecture successfully meets its design objectives. The combination of BERT-based intent classification and BioBERT-based NER provides a robust semantic understanding of health-related queries, while the retrieval-augmented generation strategy ensures that responses remain grounded in verified medical information rather than relying solely on the generative model's parametric knowledge.

One noteworthy finding is the difference in performance between the NER module and the intent classifier. The relatively lower NER scores reflect the inherent challenge of recognizing rare disease names and atypical symptom descriptions that appear infrequently in the training corpus. Future iterations of the system can address this by incorporating active learning mechanisms that flag low-confidence entity predictions for expert review and retraining.

The multilingual component performed well across the four tested languages; however, there remains room for improvement particularly in morphologically rich languages like Marathi and Tamil, where translation errors occasionally introduced ambiguity in downstream processing. Incorporating language-specific fine-tuning data and expanding the health glossary are identified as near-term priorities.

The system also incorporates an important ethical design consideration: it consistently reminds users that chatbot outputs are informational in nature and do not constitute professional medical advice, encouraging users to consult qualified healthcare providers for diagnosis and treatment. This design choice aligns with responsible AI deployment principles and helps prevent over-reliance on automated systems in high-stakes health decisions (Topol, 2019).

From a scalability perspective, the containerized deployment architecture demonstrated stable performance under simulated concurrent-user loads, making the system suitable for real-world deployment in public health campaigns, hospital information portals, and government health ministry websites.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 5, April 2026
Impact Factor: 8.2

VII. CONCLUSION

This paper has presented an AI-driven healthcare chatbot designed to promote disease awareness and deliver basic health guidance to the general public. By integrating transformer-based NLP models with a retrieval-augmented knowledge base and multilingual translation support, the system offers a comprehensive conversational health assistant that is both accurate and accessible.

Experimental evaluation demonstrated strong performance across key metrics, with intent classification achieving an F1-Score of 92.0% and a System Usability Scale score of 82.4, confirming the chatbot's effectiveness and user-friendliness. The system's modular architecture ensures that individual components can be updated independently as medical knowledge evolves and new diseases emerge.

Future work will focus on expanding the disease ontology to cover a broader spectrum of conditions, incorporating voice interaction to further lower the barrier to access.

ACKNOWLEDGMENT

The authors gratefully acknowledge the guidance provided by the faculty of the Department of Computer Engineering, XYZ College of Engineering, Pune. Special thanks are due to the medical practitioner who assisted with knowledge base curation, and to all participants who volunteered their time for the user study.

REFERENCES

[1] World Health Organization, "World Health Statistics 2022: Monitoring Health for the SDGs," WHO Press, Geneva, 2022.

[2] T. Bickmore, A. Puskar, E. Schlenker, L. Pfeifer, and S. Stahl, "Maintaining reality: Relational agents for antipsychotic medication adherence," Interacting with Computers, vol. 22, no. 4, pp. 276–288, 2018.

[3] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in Proc. NAACL-HLT, Minneapolis, MN, USA, 2019, pp. 4171–4186.

[4] J. Lee et al., "BioBERT: A pre-trained biomedical language representation model for biomedical text mining," Bioinformatics, vol. 36, no. 4, pp. 1234–1240, 2020.

[5] A. Palanica, P. Flaschner, A. Thommandram, M. Li, and Y. Fossat, "Physicians' perceptions of chatbots in health care: Cross-sectional web-based survey," Journal of Medical Internet Research, vol. 21, no. 4, e12887, 2019.

[6] L. Laranjo et al., "Conversational agents in healthcare: A systematic review," Journal of the American Medical Informatics Association, vol. 25, no. 9, pp. 1248–1258, 2018.

[7] A. A. Abd-Alrazaq, D. Alajlani, A. A. Alalwan, B. M. Bewick, P. Gardner, and M. Househ, "An overview of the features of chatbots in mental health: A scoping review," International Journal of Medical Informatics, vol. 132, 103978, 2020.

[8] P. Lewis et al., "Retrieval-augmented generation for knowledge-intensive NLP tasks," in Proc. NeurIPS, Vancouver, Canada, 2020, pp. 9459–9474.

[9] S. Mohan and D. Li, "MedMentions: A large biomedical corpus annotated with UMLS concepts," in Proc. AKBC, 2019.

[10] Y. Tang et al., "Multilingual translation with extensible multilingual pretraining and finetuning," arXiv preprint arXiv:2008.00401, 2020.

[11] P. Bajaj et al., "MS MARCO: A human generated machine reading comprehension dataset," arXiv preprint arXiv:1611.09268, 2016.

[12] M. Zhu, Y. Ahuja, J. Wei, and C. A. Reddy, "A question-answer system for health information search," in Proc. ACL, Florence, Italy, 2019, pp. 1879–1889.

[13] A. Bangor, P. Kortum, and J. Miller, "Determining what individual SUS scores mean: Adding an adjective rating scale," Journal of Usability Studies, vol. 4, no. 3, pp. 114–123, 2009.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 5, April 2026
Impact Factor: 8.2

[14] E. J. Topol, Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again. New York: Basic Books, 2019.
[15] World Health Organization, International Classification of Diseases, 11th Revision (ICD-11). Geneva: WHO, 2019.

DOI: 10.48175/IJARSCT-32891