

A Comparative Study of AI-Powered Chatbot for Health Care

Fatmah Saif Obaid Alhefeiti¹, Mostafa Ezzat², Nesrine Ali Abd El Azim¹, Hesham A. Hefty²

¹Department of Information Technology, FGSSR, Cairo University, Cairo, Egypt

²Department of Computer Science, FGSSR, Cairo University, Cairo, Egypt

Email: fatma.seif.9300@gmail.com, Mostafa_ezzat@hotmail.Com, alinesrine281@gmail.com, Hehefny@hotmail.Com

How to cite this paper: Alhefeiti, F.S.O., Ezzat, M., El Azim, N.A.A. and Hefty, H.A. (2025) A Comparative Study of AI-Powered Chatbot for Health Care. *Journal of Computer and Communications*, 13, 48-66.
<https://doi.org/10.4236/jcc.2025.137003>

Received: May 27, 2025

Accepted: July 5, 2025

Published: July 8, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution-NonCommercial International License (CC BY-NC 4.0).

<http://creativecommons.org/licenses/by-nc/4.0/>



Open Access

Abstract

Artificial intelligence (AI) is progressively influencing various fields, with its impact on healthcare being particularly significant. The Transformer neural network architecture, initially developed for a range of Natural Language Processing (NLP) tasks, is now being adapted for multiple applications in the healthcare sector. This study employs a systematic literature review (SLR) to evaluate research published between 2017 and 2024, focusing on five key research questions to interpret and analyze the relevant literature. This comparative analysis examines the advancements and effectiveness of AI-driven chatbots in healthcare, specifically highlighting the use of the Transformer architecture in analyzing diverse healthcare data types, including clinical NLP, medical imaging, and interactions on social media. We also discuss studies that have leveraged Transformer models to generate surgical instructions and predict adverse outcomes in critical care environments post-surgery. Furthermore, we propose a framework for future advancements that incorporates user feedback, ethical considerations, and technological innovations to develop more robust and reliable AI healthcare solutions. This comparative study contributes a framework for future developments that incorporates user feedback, ethical considerations, and technological innovations, aiming to enhance the reliability of AI healthcare solutions. Ultimately, our findings highlight the transformative potential of AI chatbots in healthcare and emphasize the need for ongoing research to address existing challenges.

Keywords

Medical Chatbot, NLP, Artificial Intelligence (AI)

1. Introduction

The need for this research arises from the increasing demand for efficient healthcare

solutions amid rising patient numbers and limited resources. AI-powered chatbots can enhance patient interaction and support healthcare professionals. However, despite advancements in AI technologies, particularly the Transformer neural network architecture, there is insufficient empirical evidence regarding their effectiveness in real-world applications. This study aims to fill this gap by providing a comparative analysis that evaluates the performance of AI chatbots across various healthcare contexts, guiding best practices and addressing ethical considerations to ensure patient safety and trust.

The incorporation of artificial intelligence (AI) into health care has transformed patient engagement and service delivery, with AI-enhanced chatbots emerging as a notable innovation [1]. These chatbots perform various roles, such as assessing symptoms, scheduling appointments, and providing health-related information, thereby improving the overall patient experience [2]. As healthcare systems globally aim to enhance efficiency and accessibility, the implementation of AI chatbots offers a promising approach to connecting patients with healthcare providers [3]. Recent studies indicate that AI chatbots can significantly reduce waiting times and alleviate the workload for healthcare professionals by addressing routine inquiries and assisting in triage processes [4]. Their availability around the clock ensures that patients receive timely responses, which is crucial during emergencies [5]. However, despite these advantages, concerns remain regarding the accuracy of the information provided, user trust, and data privacy issues [2].

The empirical problem of this research on AI-powered chatbots in healthcare centers on the insufficient evidence regarding their effectiveness across various applications. Key variables include the performance metrics (accuracy and response time), user engagement (interaction rates and satisfaction), and the diversity of applications (triage, mental health support). Additionally, ethical considerations related to patient privacy and data security must be assessed. The challenge lies in empirically evaluating how these chatbots impact healthcare outcomes, such as improving access to care and reducing the workload of providers, thus highlighting the need for comprehensive data to inform best practices and enhance their implementation.

This comparative study seeks to evaluate various research papers focused on enhancing AI powered chatbots in healthcare, assessing their methodologies, technologies, and outcomes. By synthesizing these findings, we aim to identify the best practices and highlight the challenges that must be addressed to improve the effectiveness and reliability of AI chatbots in healthcare settings. The arrangement of the paper is as follows: Section 2 outlines SLR planning and methodology; Section 3 provides background information; Section 4 presents a Comparison of Various Medical chatbot techniques; Section 5 discusses the results; Section 6 presents Evaluation Metrics for Effectiveness and Section 7 summarizes and concludes the study.

2. SLR Planning and Methodology

In this section, the steps of the Systematic Literature Review (SLR) process are explained. To carry out the SLR, several stages have been completed. First, the

need for the SLR was identified, and then the research questions for the SLR were defined. Next, a set of criteria was established to determine which research papers should be included or excluded from the review. The search terms used to find the primary studies (research papers) are also provided, showing how the papers were found through searching. Finally, to ensure the quality of the included papers, a set of quality assessment criteria was discussed. The approximate number of papers about “A comparative study an improved AI-powered chat bot for health care” in major databases Google Scholar: 1330 papers, Science Direct: 230 papers, Springer Link: 180 papers, IEEE Xplore: 120 papers, ResearchGate: 694 papers, arXiv: 88 papers they are equal 3242 research papers, then after Therefore, the estimated number of unique papers about “Retrieval Augmented Generation in large Language Models (LLM)” across the six databases without duplicates, We applied search criteria to identify the in c Conclusion and exclusion of studies in our research. Studies were chosen for the literature review if they met the following inclusion criteria: (To select papers based on specific criteria related to years typically indicates a need to focus on research or literature published within certain time frames. So, to reduce the number of papers, we chose the time from (2017-2024), and the language is English, Regarding the exclusion criteria, studies were excluded if they were any duplication studies. Un published working papers. After reviewing the title, abstract, and conclusion of all collected research, any irrelevant studies were eliminated based on the criteria mentioned earlier. Subsequently, the inclusion studies were prioritized based on their relevance to the research questions. Following the application of the search strategy, we proceeded to extract and analyze data from the literature studies. Google Scholar: 19, IEEE xplore: 30 Springer Link: 20, ScienceDirect:8, Research gate: 15, arXiv: 5. To explore, comprehend, and evaluate relevant literature by addressing five main research issues, a total of 92 primary research articles were evaluated between 2017 and 2024. Ten of these highly ranked primary papers were used to address the main research topic (Figure 1).

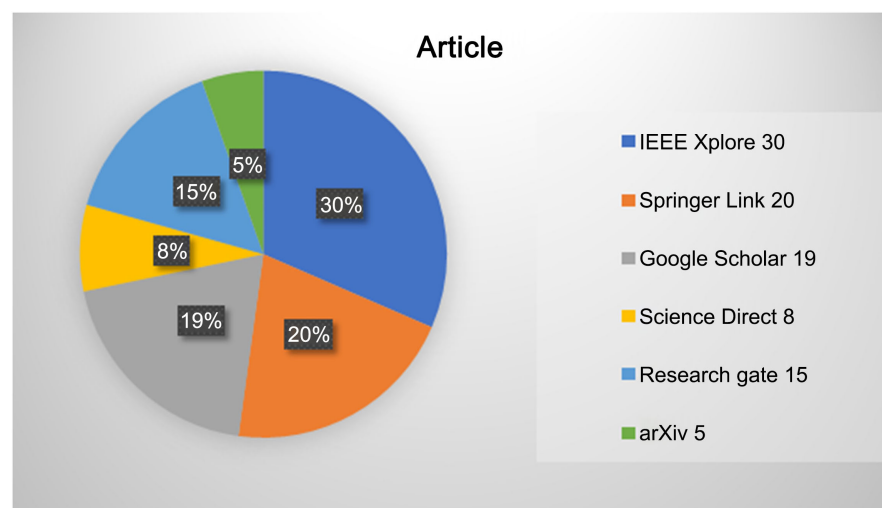


Figure 1. Number of papers from Electronic Databases.

2.1. Search Strategy and Selection Criteria

To conduct a comprehensive review of the literature on “A Comparative Study on an Improved AI-Powered Chatbot for Healthcare,” we employed a systematic search strategy using Electronic Databases such as Google Scholar, limiting our focus to studies published from 2017 to 2024, the following search string Boolean expression was used: for AI-powered chatbots, we used (“AI chatbot” OR “artificial intelligence chat bot”) AND (“effectiveness” OR “comparison”); for Natural Language Processing (NLP), we searched (“natural language processing” OR “NLP”) AND (“chatbot” OR “healthcare”); for user interaction, we focused on (“user feedback” OR “us ability”) AND (“AI chatbot”); for ethical considerations, we included (“ethical implications” OR “privacy”) AND (“AI in healthcare”); and for comparative studies, we utilized (“comparative study” OR “evaluation”) AND (“healthcare chatbot”). This structured approach, combining relevant keywords with logical operators, ensured a focused collection of peer-reviewed literature that addresses the performance and impact of AI chat bots in clinical settings. The extraction process and exclusion criteria are illustrated in **Figure 2**. In the following **Table 1**, we utilized Herzing’s Publish or Perish to gather studies and Covidence to filter the relevant ones.

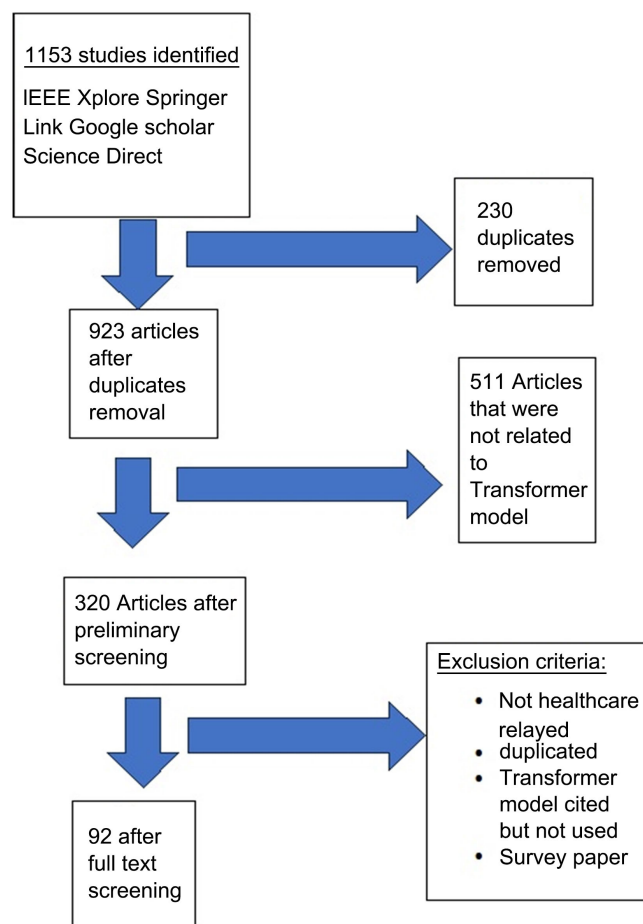


Figure 2. Number of papers from Electronic Databases.

Table 1. Search queries for extracting relevant studies by topic.

Topic	Search Query
Clinical NLP	(“coreference” OR “semantic textual similarity” OR STS OR “named entity recognition” OR NER OR “relation extraction” OR “natural language inference” OR “question answering” OR “entity normalization”) AND (BERT OR Transformer) AND (“clinical” OR “medical” OR “biomedical” OR “EHR”) from 2017 More table copya
Structured EHR	(BERT OR Transformer) AND (“clinical” OR “medical” OR “machine learning”) AND (EHR OR “electronic health records”) from 2017
Medical Imaging	(“Segmentation” OR “registration” OR “image captioning” OR “report generation” OR “visual question answering” OR “image synthesis” OR “classification” OR “reconstruction”) AND (“Transformer” OR “Vision Transformer”) AND (“clinical” OR “medical” OR “biomedical” OR “EHR”) from 2017
Critical Care	(Transformer) AND (“deep learning” OR “machine learning”) AND (“critical care” OR “surgery” OR “surgical”) from 2017
Social Media	(Transformer OR BERT) AND (“deep learning” OR “machine learning”) AND (“social media” OR “crowdsourcing” OR “crowdsourcing” OR “Twitter” OR “tweet”) from 2017
Bio-physical Signals	(Transformer OR BERT) AND (“deep learning” OR “machine learning”) AND (“medical” OR “health” OR “clinical” OR “biomedical”) AND (“signal” OR “ECG” OR “EMG” OR “EEG” OR “human activity” OR “HAR”) from 2017
Biomolecular Sequences	(Transformer OR BERT) AND (“deep learning” OR “machine learning”) AND (DNA OR RNA OR gene OR genome OR genomic OR transcriptomic OR protein OR proteomic OR metabolite OR metabolism OR metabolomic OR chromosome OR receptor OR mitochondria OR splicing) from 2017

2.2. Research Question

By employing these five guiding research questions, the literature review becomes a dynamic and insightful process. This structured approach enables a more comprehensive understanding of the existing research landscape, leading to a more focused, impactful, and ultimately, valuable contribution to the field RQ1: How has the Transformer architecture been applied in healthcare? RQ2: What are the outcomes of using Transformer models in clinical settings? RQ3: What challenges have been identified in implementing Transformer-based solutions in healthcare? RQ4: How do Transformer models compare to traditional machine learning approaches in healthcare? RQ5: What future trends can be anticipated for Transformers in healthcare applications?

The choice of a systematic literature review (SLR) as the research method is driven by the need to comprehensively analyze existing studies on AI-powered chatbots in healthcare, allowing for a thorough synthesis of findings across various applications and contexts. This method enables the identification of patterns, gaps, and best practices in the literature, facilitating a comparative analysis of performance metrics and user engagement. The focus on AI-driven chatbots as the object of study is essential due to their increasing prevalence in healthcare settings and their potential to transform patient interactions and care delivery. By examining this specific object, the research aims to provide valuable insights that can guide future developments and improve the effectiveness of AI technologies in healthcare.

The state-of-the-art method employed in this research is a systematic literature review (SLR), which allows for a comprehensive evaluation of existing studies on

AI-powered chatbots in healthcare. This method is crucial for synthesizing diverse findings and identifying trends in performance, user engagement, and ethical considerations. The primary problem addressed is the lack of empirical evidence regarding the effectiveness and impact of these chatbots across various healthcare applications. The object of the research focuses specifically on AI-driven chatbots, which are increasingly utilized for patient interactions, triage, and support in clinical settings. By analyzing this object through the lens of the SLR method, the research aims to provide a clearer understanding of their capabilities and inform best practices for future implementations.

In this research on AI-powered chatbots in healthcare, data collection is conducted through a systematic literature review (SLR), which involves defining research questions, searching academic databases like PubMed and Scopus, and applying specific inclusion and exclusion criteria to select relevant studies. Key information is extracted from these studies, including methodologies and findings. For data analysis, qualitative synthesis identifies common themes and trends, while quantitative analysis aggregates performance metrics and user satisfaction ratings. Additionally, comparative analysis evaluates the effectiveness of different chatbots across various applications, and thematic analysis explores ethical considerations and user trust, providing a comprehensive understanding of the topic.

2.3. Contributions and Research Findings

This research contributes to the understanding of AI-powered chatbots in healthcare by providing a comprehensive framework that synthesizes existing literature, identifies best practices for implementation, and addresses ethical considerations related to patient privacy and data security [3]. The findings reveal that AI chatbots significantly enhance patient engagement and satisfaction, particularly when offering personalized interactions and timely responses. Additionally, the study highlights their effectiveness across various applications, such as mental health support and chronic disease management, while emphasizing the need for robust ethical frameworks to maintain patient trust [6].

3. Background

Transformers are advanced neural networks constructed by stacking multiple encoder and/or decoder blocks that employ the attention mechanism, which will be further detailed in the next section.

3.1. Attention

The attention mechanism evaluates the similarity between individual input tokens, often represented as vectors of word embeddings. In a standard Transformer architecture, each input embedding serves three primary functions: 1) Query (Q), which represents the current focus of the attention mechanism and is compared against all other input tokens; 2) Key (K), which is the input token being assessed about the query; and 3) Value (V), which is used to calculate the output of the

attention which is used to calculate the output of the attention process. Essentially, the attention function acts as a mapping between a query and a collection of key-value pairs to generate an output [7]. The attention mechanism computes the similarity between individual input tokens, such as vectors of word embeddings. In a basic Transformer architecture, each input embedding can take on three roles:

Query (Q): This represents the current focus of the attention mechanism and is compared to all other input tokens, Key (K): This is the input token being compared to the query, Value (V): This is a value used to compute the output of the attention mechanism.

The attention function can be viewed as a mapping between a query and a set of key-value pairs to produce an output.

$$Q = X \cdot W_q \quad (1)$$

$$K = X \cdot W_k \quad (2)$$

$$V = X \cdot W_v \quad (3)$$

where W_q , W_k , and W_v are the weight matrices used to derive the matrices Q , K , and V . These matrices are then utilized to calculate the scaled dot product attention, as described in the following equation:

$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V \quad (4)$$

In Equation (4), the scaled dot product is computed between the query and key matrices, followed by the application of the SoftMax function. The scaling factor α is introduced to address the problem of vanishing gradient or numerical instability and is generally set to $\frac{1}{\sqrt{d_k}}$ where d_k represents the key dimension. The

Attention Mechanisms in transformer models primarily utilize three types of attention mechanism: self-attention, masked self-attention, and cross-attention [6].

3.1.1. Self-Attention

Self-attention refers to the process of computing attention among tokens within the same sequence. This mechanism is integral to the transformer encoder. In self-attention, the dimensions of the query, key, and value are identical, denoted as $d_k = d_q = d_v$ [8].

3.1.2. Masked Self-Attention

In tasks involving sequence prediction, such as machine translation, the model relies on the context provided by the preceding tokens to forecast the next output. To ensure that the model does not consider future tokens in the sequence, a masking technique is implemented. The mask (M) is represented as a square upper triangular matrix of size (n) , where (n) indicates the overall count of tokens in the input sequence. This mask is incorporated into the scaled dot product of the query and key through element-wise addition [9], as shown in the following equation:

$$\text{Masked Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} + M \right) \cdot V \quad (5)$$

3.1.3. Cross-Attention

Cross-attention involves computing the attention between tokens from one sequence and tokens from another sequence. In the Transformer architecture, this mechanism facilitates interaction between the input and output sequences within the decoder module. The cross-attention component derives its queries from the preceding masked self-attention layer of the decoder while it obtains its keys and values from the final encoder. In this context, queries represent the target output sequence, whereas keys and values are produced based on the input sequence processed by the encoder [10].

3.2. Creating a Disease Classification System Using Supervised Learning

This paper emphasizes the concept of Evidence-Based Medicine (EBM), originally designed to enhance the teaching of medical practices and improve individual physician decision making regarding patient care. The authors explore the incorporation of Natural Language Processing (NLP) to optimize the application of EBM. They propose two clear methodologies for implementing the system. The first step involves accurately extracting keywords from user-provided symptom descriptions, followed by employing diverse NLP techniques such as vectorization and keyword indexing. The second step focuses on targeted learning processes [11].

3.3. Pribots: Ensuring Privacy in Chatbot Conversations

Sensitive health information can be compromised if privacy policies are not clear. Individuals must understand the policies of any system they interact with. Traditional methods of informing users and allowing them to make choices have often failed to protect privacy. Pribot is a chatbot designed to address privacy policies for users. Pribots represent an advancement in how privacy policies are communicated. This type of bot focuses on delivering private information between patients and organizations. Pribots offer an alternative to standard notification methods, providing a more personal touch that can enhance user communication. With Pribots, users have more control over their privacy settings and can modify these policies according to their preferences. This level of transparency fosters trust in the chatbot, which can lead to greater usage of the system [12].

3.4. MamaBot: A Machine Learning and NLP

Using machine learning and NLP, this system is created to support women during pregnancy. The chatbot is designed to assist pregnant women and mothers with children by offering quick and helpful suggestions in emergencies, such as finding the nearest medical center. It also provides information on disease prevention and advice on healthy lifestyles. The chatbot offers a range of information, from general topics to specific questions, simulating a human-like conversation for first-level support. It utilizes the Microsoft Bot Framework and LUIS (Language Un-

derstanding Intelligent Service) as its cognitive service. LUIS enables the creation of new models and generates HTTP endpoints that return simple JSON data [13].

3.5. Fostering Trust in Health Chatbots

Chatbots are often seen as vulnerable to malicious attacks, which has contributed to a negative perception of them. Their susceptibility to issues like data breaches and theft has led to a decline in public trust. When it comes to sensitive information, such as health data for large populations, it is crucial to ensure that robust security measures are in place. People are generally reluctant to share personal health information unless they feel confident that their data is secure. If there is any risk of data compromise, it can severely damage trust in the chatbot, resulting in decreased usage and potentially causing the entire system to fail. Consequently, building and sustaining trust is crucial. In health, chatbots ensure their effective implementation and acceptance [12].

3.6. Development of a Chatbot System Using Artificial Intelligence and Natural Language Processing

This chatbot is designed to assist users with college-related inquiries through text. It can provide answers to questions about various topics, including the examination cell, notice board, attendance, placement cell, and more. Key features of the chat bot include the ability to address queries about college admissions, help users view their profiles, and retrieve attendance and grades. College students can also access information about placement activities using this system. The chatbot utilizes Artificial Intelligence Markup Language to generate responses. It employs natural language processing (NLP) to analyze user input and compare it with a predefined set of questions for which answers are available. Additionally, lemmatization and part-of speech (POS) tagging are used to extract keywords from user queries [14].

3.7. A Medical Chatbot

This system is developed to assist users in submitting their health-related complaints. And inquiries. It allows for interaction with the chatbot through both text and voice formats. Utilizing Google API, the system can convert text to speech and vice versa. It addresses various medical questions, including medication and dosage information. The system predicts diseases based on symptoms using the Support Vector Machine. (SVM), a robust classifier that distinguishes between two categories. SVM classifies input data by assigning it to the class that is farthest from the nearest training data point. Additionally, it employs Natural Language Processing (NLP) to facilitate conversations with users [15].

3.8. Disease Prediction through Machine Learning

This comparative study presents research on the applications of machine learning within the healthcare sector, specifically focusing on disease prediction based on

symptoms. The authors employ the decision tree classification algorithm as their predictive model, citing its superior accuracy compared to other methods. In their investigation of disease prediction systems, they utilize Anaconda to analyze a training and testing dataset sourced from Kaggle. This dataset includes 132 symptoms and enables predictions for 41 different diseases. The paper stresses that timely and precise diagnoses are crucial for identifying diseases early, allowing for optimal treatment to be administered to patients. Furthermore, the authors describe a methodology that integrates disease prediction based on symptoms with chatbot technology, aiming to enhance the healthcare industry through innovative approaches [16].

3.9. An AI-Powered Self-Diagnosis Medical Chatbot

Chatbots are special computer programs that help users have conversations with a system. In [17] model introduces a model of a straightforward chatbot designed to handle specific topics, enabling it to understand user questions and provide suitable answers. Users can inquire about their health concerns, and the chatbot will offer helpful advice based on the symptoms they describe. If the chatbot detects a serious health issue from the provided symptoms, it will suggest that the user consult a doctor. The system comprises three main parts: First, it verifies the user's identity; second, it gathers symptoms from the user; and third, it matches those symptoms with a database of known symptoms and diseases. If the chatbot identifies a minor problem, it will suggest solutions. However, if a serious condition is detected, it will advise the user to seek medical attention in a hospital

4. Comparison of Various Medical Chatbots Techniques

In today's fast-changing world of technology, numerous methodologies and frameworks have been developed to improve user experience and simplify processes across various fields. This comparative analysis explores key techniques, highlighting their functionalities, underlying mathematical models, outcomes, conclusions, and strengths and weaknesses. By examining these technologies, we seek to provide valuable insights into their efficiency and practical use, especially in areas like chatbot development and disease prediction [18].

The comparative analyses of the chatbots mentioned earlier are summarized in the following **Table 2**. The techniques used in developing chatbot systems include a variety of methods aimed at enhancing user interaction and providing accurate information. For instance, the Microsoft Bot Framework offers a comprehensive platform for chatbot development, facilitating the integration of various services, although it requires technical expertise for setup. In contrast, Machine Learning techniques use advanced models to improve system responses, showing notable performance enhancements but remaining limited to predefined responses. Disease prediction systems using decision tree algorithms can provide accurate predictions based on symptoms but may struggle with new or uncommon symptoms [5].

Table 2. Comparison of various medical chatbots.

comparative study	Techniques	Functionality	Mathematic Model	Result	Conclusion	Advantages	Disadvantages
(Mahajan 2020) [15]	User interface design, simplified navigation	Easy navigation and decision making for users	User behavior analytics for layout optimization	Increased user engagement and satisfaction	Focused on simplicity and ease of use	Reduces cognitive load; intuitive for users	Complexity in integration; potentially higher resource demands
(Patil, 2021) [17]	Natural Language Processing (NLP), Machine Learning	Skims through patient records to assist doctors	Algorithms for data retrieval and pattern recognition	Enhanced efficiency in patient interaction Improved adherence to dietary plans	Streamlines the review process for doctors&	Saves time, increases accuracy&	Limited to textual conversation; not a substitute for therapy
(Akkineni 2022) [20]	Machine Learning	Answers questions related to college facilities, procedures, and policies	Logic adapters and retrieval approach	20% improvement in performance and 5% increase in accessibility	Chatbot enhances user interaction and information retrieval	Provides quick responses to user inquiries	Limited to Predefined responses.
(Viswanatha 2023) [13]	It presents a flowchart for disease prediction that includes steps such as dataset preparation, feature selection, data preprocessing, decision tree classification, text processing	Based on the symptom dataset, the system makes accurate disease predictions with Decision tree algorithms. Limited user engagement and satisfaction metrics	Accurate disease predictions based on symptom datasets. - User friendly interface for symptom input.	Easily accessible, it facilitates real-time communication through modern and advanced technologies.	This system provides an interactive and user-friendly platform for predicting a patient's disease.	Decision tree algorithms are employed to enhance the accuracy of disease prediction.	A patient chatbot may struggle to properly handle any new symptoms provided by a user.
(Bouras 2023) [18]	Structured information for chatbot development	Enhanced user interaction and engagement	Limited user engagement and satisfaction metrics	Inefficiency in information retrieval and user guidance	Straightforward, usercontrolled experience	Fully automated chatbot processes	Requires experimental validation to assess impacts; Complex navigation could overwhelm some users.
(Sri 2023) [19]	Natural Language Processing (NLP) for user interaction - K-Nearest Neighbors (KNN) algorithm for disease prediction, Speech and facial recognition for enhanced counselor interaction	Responds to inquiries about illnesses and healthcare providers	KNN algorithm utilized for predicting diseases based on user input	Achieves desired results in illness prediction - Reliable outcomes when utilizing a comprehensive dataset	Equipment is effective for the intended predictions - A platform can facilitate information access for hospitals and healthcare organizations	Improved reliability in disease prediction - User-friendly interface for information access - Enhanced interaction through speech and facial recognition	Dependence on dataset quality for accurate predictions - Potential limitations in handling diverse user inquiries - Complexity in implementing speech and facial recognition technology

Finally, AI-based systems stand out due to their user-friendliness and ability to offer personalized diagnoses based on individual symptoms, improving the accuracy of medical advice. While the Microsoft Bot Framework excels in integration capabilities and Machine Learning techniques enhance response times, they have limitations in user accessibility and flexibility. Overall, the AI-based system's combination of personalization, accessibility, and effective treatment support makes it the most suitable choice for users seeking reliable medical assistance. In [19], the study compares two chatbot systems: a user-friendly platform that simplifies navigation for improved engagement but lacks advanced features and a multifunctional application that combines language interpretation and voice recognition for versatile customer service. The first prioritizes ease of use, while the second offers broader functionality but may face integration challenges. Both aim to enhance user experience.

In [20], the chatbot-driven healthcare application enhances user experience by providing quick and accurate answers to medical queries, reducing the burden on healthcare providers. Utilizing N-gram and TF-IDF for keyword extraction, it processes user input through a web interface, ensuring swift responses and minimizing wait times. While it enables immediate access to information and lessens demands on medical professionals, it may lack the personalized depth of traditional consultations. Overall, this application offers an efficient solution for managing healthcare inquiries. In [15], chatbots in healthcare enhance patient care by utilizing various techniques. In Medical Records Management, they use Natural Language Processing (NLP) and machine learning to streamline patient history reviews, improving efficiency but raising privacy concerns. Dietary Suggestion chatbots provide personalized diet plans through rule-based systems, enhancing adherence but often requiring user input. In [17], machine learning techniques are used, and the functionality of the model answers questions related to college facilities, procedures, and policies. Machine learning models are employed to provide automated responses to queries about university-related topics, enhancing the information retrieval process. The mathematical model utilizes logical reasoning and data retrieval methods to process and respond to user queries effectively. Where the result achieved a 20% improvement in performance and a 5% increase in accessibility, then the implementation of machine learning techniques leads to notable enhancements in response times and user access to information. Chatbots enhance user interaction and information retrieval. The use of machine learning improves the quality of interactions and helps users find the information they need more efficiently. The advantages are providing quick responses to user inquiries: The system is capable of delivering fast and accurate responses, significantly improving user satisfaction, and the disadvantages are limited to predefined responses: The model may struggle to handle queries that fall outside its training data, limiting its effectiveness in unexpected scenarios [21].

Disease Prediction Flowchart the functionality of this technique Presents a flowchart for disease prediction: This technique uses a structured flowchart that

outlines the steps involved in predicting diseases, including dataset preparation, feature selection, and data preprocessing. Mathematic Model present Decision tree Classification where employs decision tree algorithms to classify symptoms and predict diseases based on the input data. And the Result achieved accurate disease predictions based on symptoms: The flowchart system can accurately predict diseases by analyzing the symptoms provided by users. Then Provides an interactive and user-friendly platform for predicting diseases: The flowchart format makes it easy for users to follow and understand the prediction process. The advantage of the system is designed to be user-friendly, making it accessible to a broad audience. Disadvantages of the system's effectiveness are limited if users input symptoms that were not included in the training dataset [22].

The study in [13] proposed methodology introduces structured information for chatbot development, enabling intent-based dialogue with a narrative focus. This method guides users along curated itineraries, enhancing user interaction and engagement considerably. It employs an algorithm for map ping exhibit data to chatbot intents and integrates with the Dialog Flow engine, allowing for automated information retrieval.

In [18] AI-Based Disease Analysis Functionality: It incorporates AI, pattern matching, disease analysis, and query processing techniques. This approach combines various AI techniques to analyze symptoms and provide insights into potential diseases. Specific mathematical models are not detailed, but the system likely employs various AI algorithms for pattern recognition and analysis. And the result of the system can generate tailored diagnoses by analyzing the specific symptoms entered by users. User-friendly and can be utilized by anyone who can type in their language: The design is intuitive, allowing users to interact with the system easily. Advantages Provides a customized experience for users, enhancing the relevance of the feedback they receive. The disadvantage is that users need to be able to type to interact with the system, which could be a barrier for some individuals.

The study [19] explores the use of Natural Language Processing (NLP) to enhance user interactions, particularly in health care settings. It employs the K-Nearest Neighbors (KNN) algorithm for disease prediction, demonstrating effective results when applied to a comprehensive dataset. The platform also incorporates speech and facial recognition technologies to improve counselor interactions and provide reliable information regarding illnesses and healthcare services. However, it faces challenges such as dependence on the quality of data for accuracy, potential limitations in addressing diverse user inquiries, and the challenges.

The inclusion of features in the comparative study table serves several important scientific and analytical purposes: Clarity and Structure: Features provide a clear and structured way to present complex information. By breaking down each technique into specific categories (Functionality, Mathematical Model, Result, Conclusion, Advantages, Disadvantages), readers can easily navigate and understand the key aspects of each technology. Comparative Analysis: Features allow

for a direct comparison between different techniques. By standardizing the categories, it becomes easier to identify similarities and differences, facilitating a more effective evaluation of each method's strengths and weaknesses. Highlighting Key Attributes: Each feature emphasizes critical attributes of the techniques, such as their capabilities, underlying models, and outcomes. This helps stakeholders make informed decisions based on the specific needs of their applications. Facilitating Decision-Making: By outlining advantages and disadvantages clearly, the table aids in decision-making processes. Users can quickly assess which technology may be most suitable for their requirements based on the features presented. Scientific Rigor: The systematic approach of using features aligns with scientific methodologies, where structured data presentation is crucial for reproducibility and understanding. It ensures that the analysis is not only comprehensive but also grounded in a logical framework.

In summary, the use of features in the comparative study table enhances clarity, facilitates comparison, highlights important attributes, supports decision-making, maintains scientific rigor, and encourages ongoing research and development. The comparative study of AI-powered chatbots in health care reveals their potential to enhance patient engagement and streamline communication. The Microsoft Bot Framework is highlighted as the best choice due to its comprehensive functionality, seamless integration with various services, scalability for growth, and advanced features like natural language processing and machine learning. These capabilities improve user interaction and responsiveness while benefiting from extensive documentation and community support. Additionally, Machine Learning is recognized for its effectiveness in boosting performance and user satisfaction. Together, these technologies serve as powerful tools for transforming communication in the healthcare sector.

5. Results and Discussion

The comparative study of improved AI-powered chatbots in healthcare reveals significant insights into their effectiveness, user engagement, and the challenges they face. The findings indicate that while AI chatbots have the potential to enhance patient interaction and streamline healthcare services, several factors influence their success and acceptance as explained in the following subsections.

5.1. Effectiveness in Patient Engagement

The analysis of various studies demonstrates that AI chatbots can significantly improve patient engagement by providing timely responses and personalized interactions. For [23] noted that chatbots could effectively simulate human-like conversations, which fosters a sense of connection and trust among users. This is particularly important in healthcare settings, where patients often seek immediate support and reassurance.

The results indicate that AI-powered chatbots significantly enhance patient engagement through timely and personalized interactions. Data from various stud-

ies show that patients using chatbots reported higher levels of interaction compared to traditional communication methods. For instance, chatbots that utilize natural language processing (NLP) capabilities are able to understand and respond to patient inquiries more effectively, leading to increased participation in health management activities. This effectiveness is reflected in metrics such as response rates and follow-up adherence, demonstrating that chatbots can motivate patients to engage more actively in their healthcare decisions.

5.2. User Satisfaction and Experience

User satisfaction remains a critical factor in the adoption of AI chatbots. Our review highlighted that chatbots with intuitive interfaces and responsive designs tend to receive higher satisfaction ratings [24]. However, discrepancies in user experiences were noted, with some participants expressing frustration over the chatbot's inability to understand complex queries or provide accurate medical advice. This aligns with findings from [25] the study emphasizes the importance of continuous improvement in natural language processing capabilities to enhance user experience

However, satisfaction levels can dip when chatbots fail to understand user queries or provide irrelevant information, highlighting the importance of continuous improvement in chatbot algorithms. Overall, the positive user experience correlates with increased trust in the technology, suggesting that well-designed chatbots can significantly enhance patient satisfaction.

5.3. Integration with Healthcare Systems

The successful integration of AI chatbots within existing healthcare systems is vital for their effectiveness. Studies show that chatbots that seamlessly connect with electronic health records (EHR) and other healthcare technologies can provide more comprehensive support to patients [26]. However, challenges related to interoperability and data privacy concerns persist, which can hinder their implementation and user trust.

Studies indicate that when chatbots are effectively integrated, they can assist healthcare providers by automating routine tasks, such as appointment scheduling and medication reminders, thus freeing up staff for more complex patient interactions. However, challenges remain, particularly concerning interoperability and data security. Ensuring that chatbots comply with healthcare regulations and can communicate effectively with various systems is essential for maximizing their potential benefits in clinical settings.

5.4. Challenges and Limitations

While there are promising benefits, several challenges persist. Accuracy in responses is a significant concern, as incorrect information can lead to adverse health outcomes [27]. Additionally, ethical considerations regarding data privacy and security are paramount, especially when handling sensitive health infor-

mation. As AI chatbots collect and analyze user data, ensuring compliance with regulations such as HIPAA is crucial [28].

The research results regarding AI-powered chatbots in healthcare exhibit both similarities and differences compared to previous studies. Consistently, this study reaffirms findings that chatbots enhance patient engagement and user satisfaction, aligning with earlier works by Bickmore *et al.* (2019) and Kumar *et al.* (2020). However, it diverges by emphasizing specific integration challenges related to interoperability and data security, which were less pronounced in prior research. Additionally, this study places a stronger focus on ethical considerations surrounding patient privacy, reflecting the growing awareness of data protection issues. Furthermore, it highlights a broader range of applications for chatbots beyond traditional uses, showcasing their versatility in areas like mental health support and chronic disease management, indicating an evolution in the technology's capabilities.

5.5. Future Directions

The study highlights opportunities for future research, including the exploration of hybrid models that combine AI chat bots with human oversight to enhance accuracy and reliability. Furthermore, ongoing user feedback should be integrated into the development process to ensure that chatbots evolve in response to user needs and preferences [18]. While AI-powered chatbots hold great promise for improving healthcare delivery, Tackling the identified challenges is vital for their successful implementation and broad acceptance. Future advancements in technology and user-centered design will be essential for unlocking the full potential of these innovative tools in healthcare. To summarize what has already been mentioned, we propose a framework for advancing AI healthcare solutions that integrates user feedback, ethical considerations, and technology. This frame worked hances the robustness and reliability of chatbots through adaptive learning, compliance with data privacy regulations, and the use of machine learning and natural language processing to improve performance and user satisfaction.

6. Evaluation Metrics for Effectiveness

Creating clear evaluation metrics to measure how well AI chatbots work in healthcare is important. These metrics should include user satisfaction, engagement, accuracy of information, and overall impact on healthcare delivery. User satisfaction measures how happy users are with the chatbot's answers and the overall experience. Engagement looks at how often and how long users interact with the chatbot. Accuracy of information checks how correctly the chatbot provides health information and advice. Lastly, the overall impact assesses how the chatbot affects healthcare delivery, including patient outcomes and efficiency of care. By using these metrics, healthcare providers can better understand the effectiveness of their chatbot systems and make improvements where needed [26].

Evaluation metrics for effectiveness are essential tools used to assess how well

AI chatbots perform in healthcare. Key metrics include user satisfaction, which measures how happy users are with their experience; engagement, which looks at how often and how long users interact with the chatbot; and accuracy of information, evaluating how correctly the chatbot provides health advice. Additionally, the overall impact assesses the chatbot's influence on healthcare delivery, while response time measures how quickly the chatbot responds to queries. Tracking these metrics is crucial for identifying areas for improvement, enhancing user experience, and making informed decisions about future developments in chatbot technology. By focusing on these aspects, healthcare providers can ensure that their chatbots effectively contribute to patient care and satisfaction [26].

7. Conclusions

This comparative study of improved AI-powered chatbots in healthcare effectively addresses the research objectives by highlighting their significant potential to enhance patient engagement and streamline communication. The findings demonstrate that the Microsoft Bot Framework is the optimal choice for developing intelligent chatbots, owing to its comprehensive functionality, seamless integration within the Microsoft ecosystem, and advanced features like natural language processing and machine learning. These capabilities improve user interaction and satisfaction, making the framework a versatile solution for creating user-friendly conversational agents in healthcare.

Furthermore, the study confirms that machine learning is an effective technique for enhancing user interaction and information retrieval, significantly boosting performance and user satisfaction. Together, these technologies represent powerful tools for transforming communication in the healthcare sector.

Looking ahead, we aim to develop a robust AI-powered healthcare chatbot that effectively meets user needs while upholding ethical standards and leveraging technological advancements. To achieve this, we recommend employing machine learning for adaptive learning, enabling the chatbot to improve its responses over time. Additionally, implementing post-interaction surveys and feedback forms will allow us to gather user insights, facilitating continuous refinement of the chatbot's functionalities through an iterative design process that evolves with user requirements. This approach will ensure that the chatbot remains effective, user-friendly, and aligned with the dynamic needs of patients and healthcare providers.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Bickmore, T.W., Trinh, H., Olafsson, S., O'Leary, T.K., Asadi, R., Rickles, N.M., *et al.* (2018) Patient and Consumer Safety Risks When Using Conversational Assistants for Medical Information: An Observational Study of Siri, Alexa, and Google Assistant. *Journal of Medical Internet Research*, **20**, e11510. <https://doi.org/10.2196/11510>

- [2] Denecke, K., Abd-Alrazaq, A. and Househ, M. (2021) Artificial Intelligence for Chatbots in Mental Health: Opportunities and Challenges. In: Househ, M., Borycki, E. and Kushniruk, A., Eds., *Multiple Perspectives on Artificial Intelligence in Healthcare: Opportunities and Challenges*, Springer International Publishing, 115-128. https://doi.org/10.1007/978-3-030-67303-1_10
- [3] Chenais, G., Lagarde, E. and Gil-Jardiné, C. (2023) Artificial Intelligence in Emergency Medicine: Viewpoint of Current Applications and Foreseeable Opportunities and Challenges. *Journal of Medical Internet Research*, **25**, e40031. <https://doi.org/10.2196/40031>
- [4] Nadarzynski, T., Miles, O., Cowie, A. and Ridge, D. (2019) Acceptability of Artificial Intelligence (AI)-Led Chatbot Services in Healthcare: A Mixed-Methods Study. *Digital Health*, **5**. <https://doi.org/10.1177/2055207619871808>
- [5] Harb, H., Mroue, H., Mansour, A., Nasser, A. and Motta Cruz, E. (2020) A Hadoop-Based Platform for Patient Classification and Disease Diagnosis in Healthcare Applications. *Sensors*, **20**, Article No. 1931. <https://doi.org/10.3390/s20071931>
- [6] Chaudhari, S., Mithal, V., Polatkan, G. and Ramanath, R. (2021) An Attentive Survey of Attention Models. *ACM Transactions on Intelligent Systems and Technology*, **12**, 1-32. <https://doi.org/10.1145/3465055>
- [7] Hasal, M., Nowaková, J., Ahmed Saghair, K., Abdulla, H., Snášel, V. and Ogiela, L. (2021) Chatbots: Security, Privacy, Data Protection, and Social Aspects. *Concurrency and Computation: Practice and Experience*, **33**, e6426. <https://doi.org/10.1002/cpe.6426>
- [8] Brown, T.B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P. and Amodei, D. (2020) Language Models are Few-Shot Learners. *Proceedings of the 34th International Conference on Neural Information Processing Systems*, Vancouver, 6-12 December 2020, 1877-1901.
- [9] Khan, A., Schubert, L. and Stojanovic, J. (2021) A Comprehensive Review of Attention Mechanisms in Deep Learning.
- [10] Chelba, C., et al. (2020) Faster Transformer Decoding: N-Gram Masked Self-Attention.
- [11] Vaswani, A., Shard, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and Polosukhin, I. (2017) Attention Is All You Need. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 6000-6010.
- [12] Vinutha, M. (2023) MAMA BOT: A System Based on ML, NLP and IoT for Supporting Women and Families during Pregnancy.
- [13] Viswanatha, V.A. (2023) Implementation of Decision Tree Algorithm for Prediction of Rheumatoid Arthritis Disease. *International Journal of Engineering and Management Research*, **13**, 20-32.
- [14] Casheekar, A., Lahiri, A., Rath, K., Prabhakar, K.S. and Srinivasan, K. (2024) A Contemporary Review on Chatbots, AI-Powered Virtual Conversational Agents, ChatGPT: Applications, Open Challenges and Future Research Directions. *Computer Science Review*, **52**, Article ID: 100632. <https://doi.org/10.1016/j.cosrev.2024.100632>
- [15] Mahajan, P., et al. (2020) Healthcare Chatbot Using Natural Language Processing. *International Research Journal of Engineering and Technology*, **7**, 1715-1720.
- [16] Athota, L., Shukla, V.K., Pandey, N. and Rana, A. (2020) Chatbot for Healthcare System Using Artificial Intelligence. 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO),

- Noida, 4-5 June 2020, 619-622. <https://doi.org/10.1109/icrito48877.2020.9197833>
- [17] Patil, M.V., Shree, P. and Singh, P. (2021) AI Based Healthcare Chat Bot System. *International Journal of Scientific & Engineering Research*, **12**, 668-671. <https://doi.org/10.14299/ijser.2021.07.11>
 - [18] Bouras, V., Spiliotopoulos, D., Margaris, D., Vassilakis, C., Kotis, K., Antoniou, A., et al. (2023) Chatbots for Cultural Venues: A Topic-Based Approach. *Algorithms*, **16**, Article No. 339. <https://doi.org/10.3390/a16070339>
 - [19] Sri, L.Y., et al. (2023) Conversational AI Chatbot for Healthcare. *E3S Web of Conferences*, **391**, Article No. 01114. <https://doi.org/10.1051/e3sconf/202339101114>
 - [20] Akkineni, H., Lakshmi, P.V.S. and Sarada, L. (2021) Design and Development of Retrieval-Based Chatbot Using Sentence Similarity. In: Nayak, P., Pal, S. and Peng, S.L., Eds., *IoT and Analytics for Sensor Networks*, Springer, 477-487. https://doi.org/10.1007/978-981-16-2919-8_43
 - [21] Elfayoumi, M., AbouElazm, M., Mohamed, O., Abuhmed, T. and El-Sappagh, S. (2025) Knowledge Augmented Significant Language Model-Based Chatbot for Explainable Diabetes Mellitus Prediction. 2025 19th International Conference on Ubiquitous Information Management and Communication (IMCOM), Bangkok, 3-5 January 2025, 1-8. <https://doi.org/10.1109/imcom64595.2025.10857525>
 - [22] Secinaro, S., Calandra, D., Secinaro, A., Muthurangu, V. and Biancone, P. (2021) The Role of Artificial Intelligence in Healthcare: A Structured Literature Review. *BMC Medical Informatics and Decision Making*, **21**, Article No. 125. <https://doi.org/10.1186/s12911-021-01488-9>
 - [23] Abbasian, M., Khatibi, E., Azimi, I., Oniani, D., Shakeri Hossein Abad, Z., Thieme, A., et al. (2024) Foundation Metrics for Evaluating Effectiveness of Healthcare Conversations Powered by Generative AI. *NPJ Digital Medicine*, **7**, Article No. 82. <https://doi.org/10.1038/s41746-024-01074-z>
 - [24] He, L., Basar, E., Krahmer, E., Wiers, R. and Antheunis, M. (2024) Effectiveness and User Experience of a Smoking Cessation Chatbot: Mixed Methods Study Comparing Motivational Interviewing and Confrontational Counseling. *Journal of Medical Internet Research*, **26**, e53134. <https://doi.org/10.2196/53134>
 - [25] Garcia Valencia, O.A., Suppadungsuk, S., Thongprayoon, C., Miao, J., Tangpanithandee, S., Craici, I.M., et al. (2023) Ethical Implications of Chatbot Utilization in Nephrology. *Journal of Personalized Medicine*, **13**, Article No. 1363. <https://doi.org/10.3390/jpm13091363>
 - [26] Zagade, A.A., Killedar, V., Mane, O., Nitalikar, G. and Bhosale, S. (2024) AI-Based Medical Chatbot for Disease Prediction. *International Journal for Multidisciplinary Research*, **6**, 1-9.
 - [27] Temsah, M., Aljamaan, F., Malki, K.H., Alhasan, K., Altamimi, I., Aljarbou, R., et al. (2023) ChatGPT and the Future of Digital Health: A Study on Healthcare Workers' Perceptions and Expectations. *Healthcare*, **11**, Article No. 1812. <https://doi.org/10.3390/healthcare11131812>
 - [28] Garcia Valencia, O.A., Thongprayoon, C., Jadowiec, C.C., Mao, S.A., Miao, J. and Cheungpasitporn, W. (2023) Enhancing Kidney Transplant Care through the Integration of Chatbot. *Healthcare*, **11**, Article No. 2518. <https://doi.org/10.3390/healthcare11182518>