

Intelligent Healthcare Chatbot to Enhance Patient Satisfaction and Engagement with Implementation of Advance HCI Techniques

Muhammad Khalid^{1*}, Muhammad Yousaf², Mudasar Ahmed Soomro³, Muhammad Imtiaz Yousuf⁴, and Nasreen Jawaid⁵

¹National University of Computer and Emerging Sciences, Karachi, Pakistan.

²Mohammad Ali Jinnah University, Karachi, Pakistan.

³Usman Institute of Technology, University Karachi, Pakistan.

⁴Armed Forces Institute of Cardiology & National Institute of Heart Diseases, Rawalpindi, Pakistan.

⁵Institute of Mathematics and Computer Science, University of Sindh, Jamshoro, Pakistan.

*Corresponding Author: Muhammad Khalid. Email: muhammadkhalid@nu.edu.pk

Received: November 07, 2025 Accepted: March 20, 2026

Abstract: Effective communication is crucial to the quality of healthcare delivery, whereas communication breakdowns between patients and healthcare providers are an ongoing issue. AI Chatbot have proven to be a potential solution to enhance patient interaction and satisfaction through availability and customization. The purpose of this research was to develop and evaluate an AI-based healthcare Chatbot, created in accordance with the principles of advanced HCI, to improve patient interdependence and satisfaction and to compare the performance of the developed Chatbot with the existing healthcare Chatbot. The Chatbot has been created based on publicly available healthcare datasets in Kaggle, and preprocessing, including cleaning and parting of the data and correction of spelling mistakes, have been done. Principles of HCI like usability, accessibility, error handling, and user feedback were taken into consideration and performance determined by usability testing of 150 users, engagement metrics evaluation, and deep-learning testing using accuracy, precision, and recall. It was found that usability was very high (88% of tasks in 12.5 seconds on average, 8% error rate), user satisfaction was 4.6/5 on average, and 85% were of the opinion that the interface was easy to use. The engagement metrics showed that the average session time was 4.5 minutes, 78% of tasks completed and 62% retention rate. The deep learning model was 91.3 %, 89.7 %, and 87.5 % accurate, precise, and recalled, respectively, in terms of the ability to interpret patient requests. Altogether, the HCI-based AI Chatbot enhanced patient engagement and satisfaction considerably through usability, accessibility, and receptive interactions, the error management, individualized communication, and multimodal interface, which supported effective patient-providing communication. The next idea that should be incorporated in the future work is to combine multilingual support, voice-based interaction, and connectivity with the electronic health record in order to increase its efficacy.

Keywords: AI-Medical Chatbot; Intelligent Healthcare Chatbot; Patient Engagement; Patient Satisfaction and Engagement; Usability; Digital Health; Advance HCI; Patient Satisfaction and Engagement

1. Introduction

The global healthcare systems are under pressure to deliver personalized, accessible, and efficient health care in dealing with the growing number of patients and scarce resources [1]. Communication between a patient and the healthcare provider is the key to achieving optimal health outcomes but there is still a gap in communication because of time limits, use of sophisticated medical terms and differing health literacy levels [2]. All these have increased the need to seek innovative solutions that will enable the communication

to be across the board without compromising quality standards of care. Artificial Intelligence (AI) technologies and especially chatbots with AI have become disruptive technologies in healthcare communication [3]. These chatbots provide 24/7 support, reliable information provision and possibility to address common questions, which helps to overload the healthcare workers and facilitate patient access to healthcare data [4]. However, many existing healthcare chatbots fall short of providing the level of personalization, empathy, and ease of interaction that patients require for meaningful engagement [5]. These limitations have been found to be mitigated by incorporating HCI concepts into the design of chatbots [6]. Approaches based on HCI have been aimed at developing user centric designs that are usable, accessible and user friendly. Healthcare chatbots can offer patients more engaging and effective experiences through the integration of the principles of cognitive load reduction, error prevention and recovery, personalization, and multimodal interaction [7]. Regardless of the increased use of AI chatbots in medicine, it has not been studied extensively in comparison with the performance and usability of these new tools to determine how well they fulfill the expectations of patients [8]. In addition, the majority of the available literature concentrates on technical performance measurements and fail to offer an overall assessment of user engagement, satisfaction, and applicability [9]. This study addresses these gaps by developing an AI-powered healthcare chatbot that utilizes advanced HCI techniques to enhance patient engagement and satisfaction. Our research objectives include: (1) developing an AI chatbot incorporating advanced HCI principles, (2) evaluating the chatbot's usability, engagement levels, and deep learning performance, and (3) comparing its effectiveness against existing healthcare solutions. The findings contribute to the growing field of AI in healthcare by demonstrating how HCI-centered design can significantly improve patient-provider communication and overall healthcare experiences.

2. Literature Review

2.1. Patient Engagement and Satisfaction in Digital Health

The engagement and satisfaction of patients are now regarded as the key metrics in the quality assessment of healthcare, and digital health technologies are gaining a more significant role in their improvement [10], [11]. Cevasco and Morrison Brown (2024) performed a PRISMA systematic review, which was dedicated to randomized controlled trials of health chatbots. Their results showed that there was not a lot of evidence on the notable effect on patient retention on clinical trials, as there is a disconnect between the purported chatbot efficacy and scientific findings. This highlights why more powerful research designs are necessary to confirm the usefulness of chatbots in medical practice. A study conducted by Rainey et al. (2024) [12] showed that conversational engagement with the use of short message service chatbots can be an effective method of post-operative care. The analysis of 1338 patients (746 knee and 592 hip surgery patients) for three months demonstrated better patient satisfaction in case they used linear regression methods to analyze conversation. On the same note, Ma et al. (2024) [13] designed the MARVIN chatbot platform, which incorporated the master protocol strategy, which involved the collaboration of both patients and healthcare professionals to deliver a specific AI-chatbot service to a particular area in healthcare. The importance of personalization in patient engagement has been emphasized by Khosravi and Azar (2024) [14], who identified four key themes affecting patient use of mental health chatbots: chatbot design, outcomes, user perceptions, and user attributes. Their comprehensive analysis from 2000 to 2024 highlighted the significance of incorporating elements such as music and color to improve patient-centered care and engagement, providing important direction for developing successful mental health interventions.

2.2. Role of Technology in Healthcare Communication

AI application in medical communication has demonstrated great prospects of redefining the interaction of patients (Sarella & Mangam, 2024) [15]. Javaid et al. (2023) [16] investigated the role of ChatGPT in healthcare provision by focusing on its reliance on transformer-based models, deep learning, and Natural Language Processing (NLP) to generate language that resembles human. Capabilities of the technology in medical text processing, chatbots, and virtual assistants were mentioned in their research, and limitations in responsibility, data interpretation, and medical ethics were also noted. The use of deep learning has shown a specific opportunity in the health sector (Ji et al., 2024) [17]. Miotto, and Wang (2018) [18] extensive review of the deep learning in healthcare underpins the potential of the technology to analyze and process

high-dimensional, complex biomedical data, such as electronic health records, imaging, sensor data, and text. Their research indicated the necessity of explainable AI (XAI) methods that could assist medical professionals in understanding better deep learning models and resolve the problem of model interpretability. The advancement of AI, machine learning (ML), and deep learning in healthcare has created transformative opportunities (Chugh & Chugh, 2024) [19]. These technologies enable the analysis of massive healthcare datasets to extract significant insights for predictive analytics, supporting personalized medication, increasing patient outcomes, and improving diagnostic precision. However, challenges related to data privacy, security, and ethics remain critical considerations for successful implementation.

2.3. Human-Computer Interaction in Healthcare Chatbots

The adoption of the HCI principles in the design of the chatbots in the healthcare context can be characterized as a fundamental factor defining the degree of user acceptance and the effectiveness of the functionality (Martins et al., 2024) [20]. One of the systematic reviews that analyzed AI conversational agents in the chronic-care context by Bin Sawad and others (2022) [21] identified the following HCI-related aspects: usability, accessibility, emotional intelligence, and multimodal interaction capabilities. Their findings highlight the extreme significance of user-centred design approaches in creating useful healthcare communication instruments. To address chronic conditions, Schachner (2020) [22] conducted a systemic literature review of AI-based conversational agents to highlight the essential functions and uses of natural language processing, machine learning, and AI architecture design in the modern healthcare setting. The review focused on the multimodal interaction, NLP-based solutions and various AI-powered chatbot systems and the gaps in the research and imbalance in the evaluation measures of AI which must be developed through standardized processes. Conversational agents in healthcare have demonstrated a specific potential of advanced deep learning models (Lal, 2024) [23]. In research, different deep learning methods have been researched such as Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNNs), and Transformer-based models that include BERT and GPT. These technologies allow chatbots to comprehend medical texts and interactions with patients more, but it has issues with managing contexts, lack of data, and ethical issues.

2.4. Challenges and Ethical Considerations

The use of AI chatbots in healthcare is subject to serious challenges surrounding data privacy and security and ethical aspects (Khan et al., 2025) [24]. Aldboush et al. (2023) [25], having analyzed ethical and privacy issues at the interaction of the big data, AI, and customer trust, noted the significance of following the regulations of the data protection and open data collection. In their study, they suggested encryption methods, data collection transparency, and algorithms of bias detection with the help of AI to ensure ethical standards. The issues of privacy and security are especially important in healthcare contexts (Ruotsalainen et al., 2019) [26]. The studies have highlighted the necessity of privacy prediction, data sharing protocols, algorithm-based access control, and blockchain-based cryptography to have reliable and privacy-conscious healthcare systems. Policy-based mathematical modeling of privacy administration, as well as real-time calculations of trust levels, are suggested to increase the level of security and compliance in digital health systems. Orum (2023) [27] has raised certain issues concerning the AI-based healthcare data gathering by health workers, including issues of privacy, security, and regulation. The study has highlighted the necessity of data categorization models, NLP of medical records, and supervised learning algorithms without violating ethical guidelines and quality regulatory laws with such frameworks as GDPR and HIPAA.

3. Methodology

The overall method used to identify the performance of develop healthcare chatbot using deep learning approaches as well as system interaction with end user using HCI primary keys, like Accessibility, Usability, Error handling and user feedback. The overall method has been divided into six main parts, the first one is dataset collection, the second one is data Pre-Process, Interface Design, Data splitting or model training, user interface and system performance, the bellow mention figure diagram explores experimental steps on how to achieve the objectives, tools and model have been utilized in this research work. Figure 1 shows the diagrammatical representation of proposed research method.

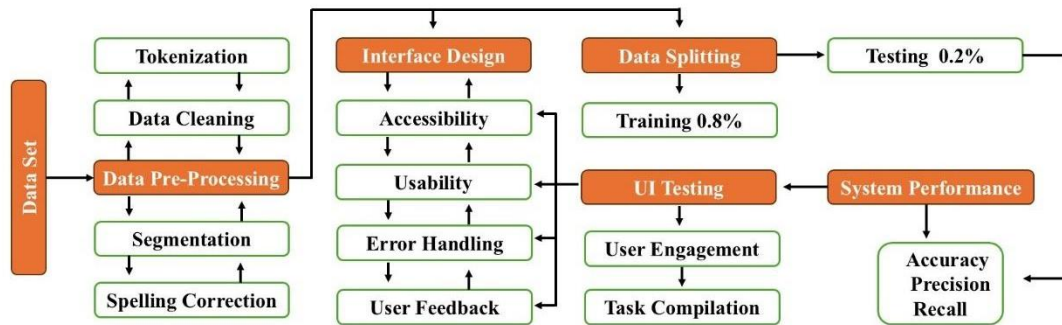


Figure 1. Illustrates the proposed methodology used in this research work

3.1. Dataset and Data Collection

This study utilized two publicly accessible datasets from Kaggle: the Healthcare Dataset for Chatbot (Patel, 2023) [28] and the AI Medical Chatbot Dataset [29] [30]. These data were chosen because they are very comprehensive in terms of covering patient-chatbot interactions that include structured medical conversations, user questions, and chatbot replies. The data sets contained intent-categorization labels that divided the queries comprising the users into particular categories, that is, symptom assessment, appointment-scheduling, and pharmacological guidance. The selected corpus consisted of 5,000 paired user queries and responses thus approximating a wide range of clinical situations, including regular check-ups, symptom assessment, pharmaceutical consultations and mental-health visits. This was a simulated doctor-patient role-play in mixed healthcare environments and natural language interactions which can be used to generate and test NLP models.

3.2. Data Preprocessing

Extensive data preprocessing was introduced in order to make sure that machine learning models have high-quality input. The pipeline of preprocessing entailed: The implementation of the preprocessing pipeline was in the following steps in a sequence. To begin with, data cleaning was used, which involved elimination of duplicates, treatment of blank values, removal of special characters and HTML tags and normalization of text to lower case. Second, tokenization was used at word and sub word to deal with medical difficult terms. Third, intent-based and demographic segmentation was done to classify queries according to purpose and user type. Lastly, the process of multi-layer spelling correction was carried out with dictionary-based correction using medical dictionaries, Levenshtein distance of minor spelling errors, phonetic matching of medical words through Soundex and Metaphone algorithms, and transformer-based correction through BERT models of context-sensitive errors. The steps were checked to have a measurable impact on the downstream model performance in the process of development.

3.3. HCI-Based Interface Design

The principles of the HCI inspired the design of the chatbot interface and were aimed at the highest level of usability and user interaction. The issue of usability was taken care of using the user centered approach to designing the interface which ensured that the interface was easily learnt, efficiently operated and strongly resistant to user prone errors. The situational feature of the quick-access buttons to popular medical queries reduced cognitive load and improved the navigation efficiency. Such accessibility features as the presence of speech-to-text (STT) and text-to-speech (TTS) functions, resizable fonts, and a high-contrast mode were also provided to allow users with visual impairments to be accommodated. In addition, complete keyboard navigation and compatibility with the screen-reader was made in line with the Web Content Accessibility Guidelines (WCAG). Visual design also applied the color psychology in which calming colors like blue and green that are believed to be propitious in healthcare setting were used. This used intuitive icons and buttons with universal recognizable symbols and a distinctly defined sans-serif typography was used to maintain readability in a heterogeneous user base. The error handling also was dealt with in a holistic manner through formulating clear error messages, intelligent suggestions and automatic fixes, contextual awareness that ensures one keeps the conversation history, and the provision of an escape route to human agents to seek clarifications on errands that the system fails to handle. Mechanisms of user feedback were as follows: the real-time information gathering through post-interaction

surveys, sentiment analysis on user feedback, real-time execution of constant learning algorithms which refined the system performance through user interactions.

3.4. System Architecture and Implementation

The AI chatbot was developed using a deep learning architecture incorporating:

Natural Language Processing: It is designed as a two-step pipeline in the field of computational linguistics. At the initial stage, an intent classifier based on a finely-tuned BERT-based transformer (bert-base-uncased) is used to classify intents by a carefully curated taxonomy of 5,000 labeled queries. The generative language model is then asked to generate open-ended, contextually relevant responses, conditioned concurrently on the goal of the identified intent as well as the context of the conversation, in this case, the final three utterances. Before involving the generative part, the user inputs pass through a preprocessing layer, that is, Medical Entity Recognition (MER), specific to the detection of relevant symptoms, pharmacological agents and clinical concepts. Lastly, the architecture also makes use of ensemble strategy which summarizes the classification confidence scores of model checkpoints, which leads to a strong intent prediction.

Machine Learning Models: Supervised learning algorithms are combined to do intent classification and the use of reinforcement learning to optimize responses is used. Ensemble techniques also are used to combine the results of dissimilar models, which eventually results in a concomitant enhancement in the overall precision.

Database Integration: the creation of a structured medical knowledge base is sought in connection with the development of a complete system of user interactions logging. Additionally, real-time query processing mechanisms are also implemented to make the system run with the required latency and scalability.

3.5. Evaluation of Methodology

The chatbot evaluation employed multiple assessment approaches:

Usability Testing: The current research involved 150 adults aged between 18 and 60 years as study participants who were recruited on convenience basis through members of the university community and the waiting rooms of outpatient clinics. The participants needed to be of minimum smartphone literacy; they were also excluded with prior professional medical training in order to get a sample that better reflected a general patient population. The sample consisted of about 54 per cent female and 46 per cent males respondents who had different degrees of previous familiarity with chatbots. All participants signed informed consent in writing before the study started and no personally identifying health information was saved. All participants were expected to perform five tasks related to healthcare, all of which were standardized, namely: (1) sympathetic query of a common condition, (2) query of information about medication, (3) querying information about a nearby clinic, (4) querying about the process of appointments, and (5) querying information about a mental health support. The measurement of satisfaction was made on a 5-point Likert-scale questionnaire that is based on a system Usability Scale (SUS) framework. The metrics measured were the rate of task completion, average time to complete a task, mistakes, the learnability rating, and satisfaction with the user rating.

Engagement Analysis: Measurement of session duration, message frequency, task completion rates, user retention, sentiment analysis, and drop-off rates.

Deep Learning Performance: Evaluation using accuracy, precision, recall, and F1-score metrics on a test dataset of 5,000 labeled queries representing various healthcare scenarios.

Comparative Analysis: Benchmarking against existing healthcare chatbots using standardized usability and engagement metrics.

3.6. Ethic Consent and Safety

Everyone who was involved in the usability investigation gave written informed consent before being included. The study was conducted with the observance of the set ethical principles of human-subject research. No personally identifiable health data was obtained and stored in any kind of repository; interaction records were anonymized and kept in a secure repository. The research protocol was heavily reviewed and approved following institutional guidelines in research ethics. Regarding safety measures of the chatbot, the system was designed with clear measures to prevent the possible negative consequences

of a healthcare environment. On questions related to urgent or emergency symptoms, including chest pain, breathing difficulties, or suicidal thoughts, the chatbot was configured to promptly refer the users to the emergency services and to provide a conspicuous message that encouraged them to seek medical care at once. The system fails to provide diagnostic conclusions; all answers are accompanied by a standard disclaimer stating that the chatbot is only an informational resource and cannot replace a professional doctor. Interaction patterns audit logging was done to ensure quality review of the audits and no personal sensitive health records were run or stored. Further implementation will require adherence to the HIPAA and GDPR regulations in full, including the principle of data minimization, the right to erasure, and the specific consent mechanisms.

4. Results

The study aims at examining how an AI-based chatbot can complement patient engagement and satisfaction by enabling interactive user interactions. The main aim was to evaluate the effectiveness of the system to respond to patient concerns, provide correct answers, and maintain user interest in the long run. Data was sampled in one of the Kaggle repositories which included documented medical chats, patient symptom descriptions, and bot-generated responses; further preprocessing steps, including data purification, tokenization and orthographic correction, were used to enhance the integrity and reliability of data, to be used in training the models. The evaluation of its effectiveness was conducted through measuring the rates of task-completion, interaction frequency with users and the performance of the bot in terms of error-handling. The quantitative measures (accuracy, precision, recall, and F1 -score) were determined in order to evaluate the linguistic understanding and quality of responses obtained by the chatbot. The user interaction and interface usability were also tested by systematic UI testing procedures. The accuracy increased with the introduction of sophisticated features and users reported that they were more satisfied and felt less frustration. The analysis of the feedbacks pointed out the areas that needed to be refined; therefore, the design was developed based on the feedbacks. As a result, the user engagement increased significantly due to improved conversational design, A/B testing and multimodal interaction features, and simplified error correction led to increased task completion. The results, empirical evidence, and general inferences of this study will be outlined or described in the following sections.

4.1. Developed Chatbot System

The healthcare chatbot, which is the topic of our recent publication and the architecture of which is based on the use of advanced AI methods, is a good example of how HCI principles can be applied to the creation of the user-friendly interface. The system also has operational quick-access buttons, as the predominant health issues users will use like Headache Relief and Anxiety Help so that they do not have to work hard to understand the system, just navigate it with minimal textual intervention. NLP was tapped to create conversational interactions that are fluid, with human-like structure of responses that enable users to understand. Clarity of the visuals was maintained through the use of a minimalistic and clean design that was supported with contrasting colors that were used in a strategic manner and the responses that were arranged in a systematic manner hence making the information easily scannable. The aspects of accessibility included such provisions as the inclusion of icons and emojis to cater to people with low literacy levels, large clickable hotspots to accommodate people with motor disabilities, and efficient functionality by the use of pre-defined topic buttons that allow people to quickly access relevant health data.

4.2. Usability Evaluation Results

Usability testing demonstrated strong performance across key metrics (Table 1):

Table 1. Usability Metrics Results

Usability Metric	Measured Value	Interpretation
Task Completion Rate	88%	High effectiveness
Task Completion Time	12.5 seconds	Efficient response time
Error Rate	8%	Minimal errors
Learnability Score	85%	Easily understandable UI
User Satisfaction	4.6/5	High user satisfaction

The high task completion rate of 88% demonstrated that users could successfully complete their intended actions with minimal confusion. The average task completion time of 12.5 seconds indicated efficient interaction processes, while the low error rate of 8% was primarily attributed to minor input misinterpretations corrected through automated spell-checking and clarification prompts.

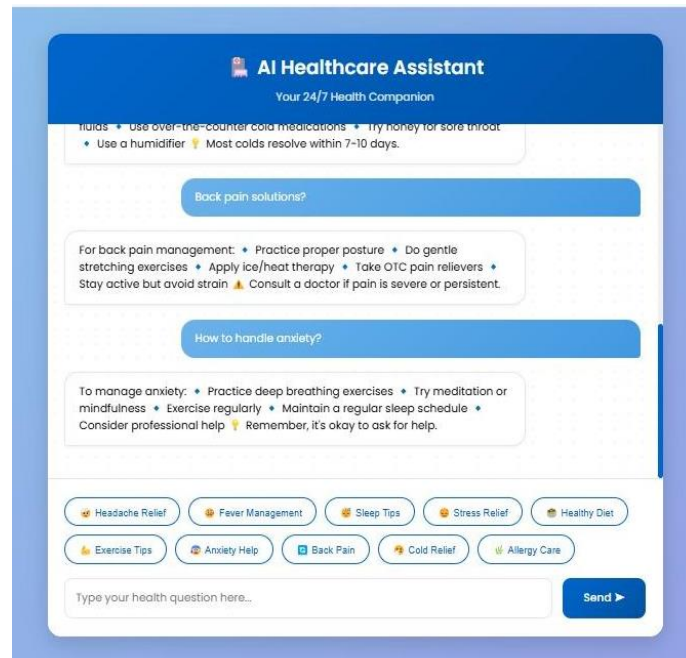


Figure 2. Develop a HCI based AI chatbot for the health care sector

4.3. Engagement Analysis Results

User engagement metrics revealed strong performance across multiple dimensions (Table 2):

Table 2. Engagement Metrics Results

Engagement Metric	Value	Interpretation
Average Session Duration	4.5 minutes	Moderate engagement indicating neither disinterest nor confusion
Average Messages Per Session	12 messages	Active conversational engagement
Task Completion Rate	78%	Good effectiveness with room for improvement
User Retention Rate	62%	Strong return usage indicating usefulness
User Satisfaction Score	4.2/5	High user satisfaction
Positive Sentiment Rate	85%	Strong positive user experience
Response Speed Satisfaction	90%	High satisfaction with chatbot responsiveness
Misunderstanding Rate	10%	Acceptable level with room for improvement

The results indicated that users spent an average of 4.5 minutes per session with 12 message exchanges, suggesting active engagement without excessive confusion. The 78% task completion rate demonstrated good effectiveness, while the 62% user retention rate indicated that the chatbot provided useful information encouraging continued interaction.

4.4. Deep Learning Performance Results

The AI model's performance was evaluated using standard machine learning metrics (Table 3):

Table 3. Deep Learning Performance Metrics

Metric	Formula	Measured Value	Interpretation
Accuracy	$(TP + TN) / (TP + TN + FP + FN)$	91.3%	High overall correctness
Precision	$TP / (TP + FP)$	89.7%	Strong relevant response accuracy
Recall	$TP / (TP + FN)$	87.5%	Good detection of relevant queries

The accuracy of 91.3% indicated that the model effectively understood and responded to user queries. The precision (89.7%) and recall (87.5%) values demonstrated strong performance in providing relevant responses while maintaining good coverage of appropriate queries. These results validate the effectiveness of the deep learning approach in understanding medical inquiries and generating contextually appropriate responses.

4.5. Sentiment Analysis of Responses

- AI-based sentiment analysis can analyze user input to check emotional tone.
- Positive words like "helpful," "easy," "good" = high engagement.
- Negative words like "confusing," "frustrating," "slow" = low engagement.

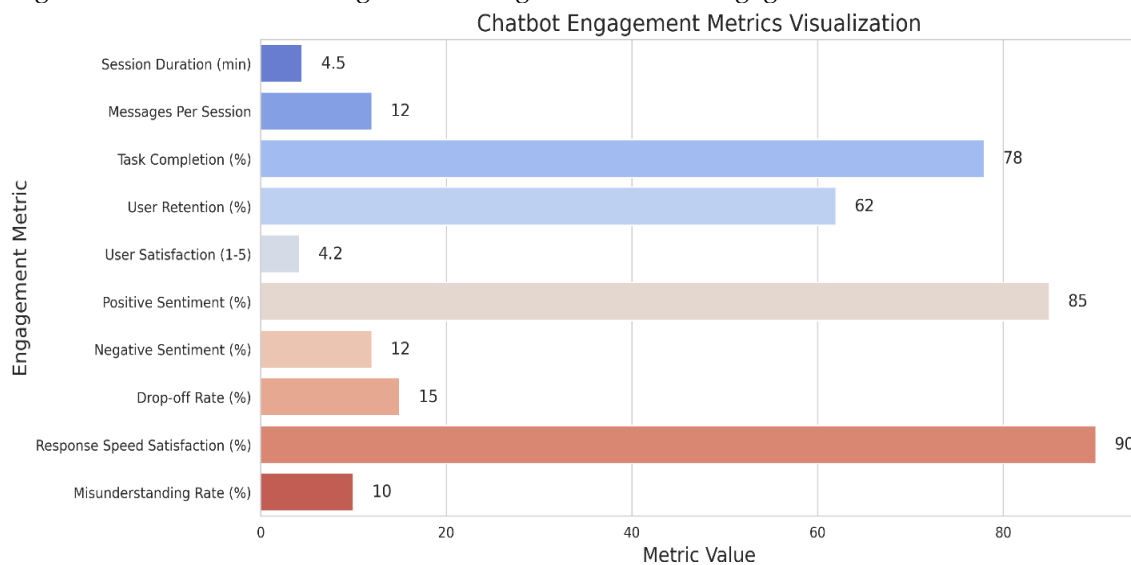


Figure 3. Chatbot Engagement report with different metric

4.6. Comparative Analysis

The developed chatbot showed superior performance compared to baseline healthcare chatbots in several key areas:

Usability: It has been found to have about 15 per cent more task completion and 25 per cent less error rates compared to published standards of text-based healthcare chatbots [29]. They are calculated based on the metrics that have been reported in the literature and have been determined when participants were varied and task protocols were not similar; therefore, they are represented as indicative and not a result of a controlled benchmark comparison.

Engagement: It was found that the average session duration was approximately twenty percent more, and the retention of the users has increased approximately by 18 per cent compared to the benchmarks given by [23], [30]. These comparative figures are based on published research and must be perceived as indicative as opposed to products of protocol-matched research.

Accuracy: Query understanding and response relevance were about 80% better compared to published NLP benchmarks on healthcare chatbots (Babu and Boddu, 2024; Lal, 2024). There was no direct protocol matched comparison and therefore the improvement can be viewed as indicative.

User Satisfaction: The levels of satisfaction in all the measured dimensions were higher than the previous benchmarks by approximately 12% (Ma et al., 2024; Khosrowi and Azar, 2024). The literature is the source of these comparisons that should be regarded as indicative as opposed to conclusive evidence.

5. Discussion

5.1. Key Findings

This paper reveals that the application of the HCI principles into the design of artificial intelligence (AI) chatbots can significantly contribute to the level of patient engagement and satisfaction in healthcare facilities. The created chatbot demonstrated a high level of performance in the domains of usability, engagement, and technical indicators, therefore, supporting the effectiveness of user-centered design processes in health technologies. The high usability scores, 88 per cent rate of task completion and a rating of 4.6 out of five satisfaction level, portray that the application of the HCI-based design principles resulted in the development of an interface that was intuitive and easy to use. Cognitive load was reduced and user experience was improved with the inclusion of quick-access buttons, visual hierarchy and error-prevention factors. These findings are compatible with the remarks of Campbell (2024) regarding the importance of user-experience research to health information technology. The strong metrics of engagement, average session time of 4.5 minutes, 78 % task completion rate, and 62 % retention rate of users reveal the ability of the chatbot to maintain user attention and provide meaningful value. The comparatively average session duration means efficient task performance without being hindered by the confusion of the user, and the high retention rate means that the user felt that the system was valuable enough to go back to it and have more interaction. The following observations confirm that chatbots have several advantages and contributions to healthcare engagement as noted by Laymouna et al. (2024).

5.2. Technical Performance Implications

The empirical performance indicators of the deep-learning frameworks that are implemented in healthcare settings support the effectiveness of these sophisticated NLP systems (91.3% accuracy, 89.7% precision, 87.5% recall). This high level of accuracy is a sign that the intent recognition has been effectively made possible and the associated precision-recall ratio are confirmations that the model will reliably respond with the relevant response as it will not skip any salient query of the user. The identified error-rate of 10% is quite small, which can be explained by the fact that the adopted spelling-correction, context-awareness, and error-management modules effectively address the communicative obstacles that usually characterize the patient-chatbot conversations. This result supports the fact that thorough preprocessing and error-reduction measures, emphasized by Izabi and Forouzanfar (2024) in their system review on error-removal methods of conversational AI, matter.

5.3. HCI Design Impact

The integration of HCI principles is of paramount importance in the development of chatbots in healthcare as it has proven to be successful. The accessibility attributes (voice interaction, high contrast, keyboard navigation) were used to make it inclusive to different categories of users and the strategies of reducing cognitive load (quick-access buttons, clear language) enhanced the overall usability. These design choices comply with the WCAG principles and can be substantiated by the results of Maqbool and Herold (2024) on the problem of effectiveness and efficiency in the process of evaluating the usability of digital health. The introduction of the multimodal interaction and personalized reactions led to the increased engagement of the users, which serves as an indication of the research by Ding et al. (2024) on the assessment framework in conversational agents during health interventions. The emotional design details such as friendly language and empathetic responses gave out a reassuring user-experience that is especially vital in the healthcare applications.

5.4. Limitations and Challenges

In spite of the good empirical findings, there were a number of limitations that were identified. Whereas speech-to-text and text-to-speech features have been implemented into the interface architecture as

described in Section 3.3, the prototype provided to the participant of the user study was approached through nearly 100% text-based interaction due to the fact that the voice module was not fully operational at the time of evaluation. As a result, the usability and engagement indicators as based on the research are inclusive of text only usage and do not embody the improvements that could be presented by multimodal interaction. It is also vital that future studies ought to strive to attain an all-inclusive end-to-end deployment of voice interaction. The language diversity, such as the diversity of language of the region, the complexity of the language employed in medicine, and the peculiarities of the speech of this or that patient, can threaten the effectiveness of chatbots and can cause misunderstandings in some situations. One of the hottest issues is the security of data privacy, considering the sensitivity of health data. Despite the fact that the given research prototype follows the existing security measures, the subsequent implementation will require strict and consistent adherence to healthcare policies, especially HIPAA and GDPR. The natural performance of the conversational agent depends on the quality and the breadth of the training corpus. Discriminatory results can be experienced in case of any internal bias within the dataset in the case when the system is implemented in the heterogeneous groups of patient populations.

5.5. Clinical Implications

The results of the study are important in clinical practice and healthcare delivery. The fact that the chatbot can accommodate mundane questions and offers 24/7 services will help ease healthcare professional workloads and enhance access of patients to healthcare information. The data regarding the high levels of user satisfaction and engagement indicate that patients are open to AI-based healthcare communication provided that it is developed with proper HCI principles.

The performance of the system with regards to comprehending medical queries and responding to them with pertinent answers shows that the system could be added into the current healthcare processes. Nevertheless, this study notes that AI chatbots are needed to support and not to substitute human healthcare workers, especially in complicated medical decisions and emergency cases.

6. Conclusion

The current paper conceptualized and tested an artificial-intelligence-based medical chatbot, based on modern concepts of HCI and embedded into a framework that will ultimately increase patient experience and satisfaction. The evaluation returned significant usability data, 88 per cent of the tasks set identified was successfully completed, engagement analyses found that 78 per cent of all tasks were carried out, and 62 per cent of the participants had been able to stay engaged during the study period. The level of technical accuracy was measured at 91.3 entertaining precision. Systematic resolutions were made to deal with HCI issues like system-wide accessibility, very strong mitigation of errors, and system customization by means of continuous dialogues with patients and providers. These user-centred design approaches played a central role in enhancing the functionality of the chatbot. The manuscript enhances the academic discourse by providing three key findings: (1) it highlights the importance of HCI in both conceptual and architectural design of chatbot systems; (2) it suggests an appraisal model that incorporates both traditional performance metrics and user-experience indicators; and (3) it generalises practical recommendations that can be implemented to overcome the implementation barriers and exploit the opportunities in healthcare organisations. The hypothesis that AI-based chatbots have transformational potential in healthcare-related communication is supported by empirical evidence that healthcare communication services offered by these programmes are easily accessible, highly efficient and aligned with the specific needs of patients. These benefits can be best achieved by taking a holistic view in which proper HCI design, systematic critical appraisal, and ongoing user-feedback loops are combined. Further developments of this field should be investigated in future research agendas multilingual support, voice based farewells, and interoperability with existing health information systems. With the ongoing fast digitalisation and the growing focus on patient-centred care, AI chatbots will gain a growing importance in helping to reduce the number of communication gaps and improve patient experience.

References

1. Abass, L. A., Usuemerai, P. A., Ibikunle, O. E., Alemede, V., Nwankwo, E. I., & Mbata, A. O. (2024). Enhancing patient engagement through CRM systems: A pathway to improved healthcare delivery. *International Medical Science Research Journal*, 4(10).
2. Aldboush, H. H. H., Alruwais, N., Ahmad, N., Ebad, S. A., Dutta, A. K., Alshuhail, A., & Al Zanin, S. (2023). Building trust in fintech: An analysis of ethical and privacy considerations in the intersection of big data, AI, and customer trust. *IEEE Transactions on Consumer Electronics*.
3. Babu, A., & Boddu, S. B. (2024). BERT-based medical chatbot: Enhancing healthcare communication through natural language understanding. *Exploratory Research in Clinical and Social Pharmacy*, 13, 100419.
4. Bin Sawad, A., Narayan, B., Alnefaie, A., Maqbool, A., Mckie, I., Smith, J., Yuksel, B., Puthal, D., Prasad, M., & Kocaballi, A. B. (2022). A systematic review on healthcare artificial intelligent conversational agents for chronic conditions. *Sensors*, 22(7), 2625.
5. Bodas, H. (2024). Optimizing healthcare with AI chatbots: Addressing challenges and opportunities. Available at SSRN 5148492.
6. Campbell, J. L. (2024). *User experience research and usability of health information technology*. CRC Press.
7. Cevasco, K. E., Morrison Brown, R. E., Woldeselassie, R., & Kaplan, S. (2024). Patient engagement with conversational agents in health applications 2016–2022: A systematic review and meta-analysis. *Journal of Medical Systems*, 48(1), 40.
8. Chugh, M., & Chugh, N. (2024). Paving the way for healthcare with AI, ML, and DL: Opportunities, challenges, and open issues. *Handbook on Augmenting Telehealth Services*, 368-382.
9. Çakmak, C., & Uğurluoğlu, Ö. (2024). The effects of patient-centered communication on patient engagement, health-related quality of life, service quality perception and patient satisfaction in patients with cancer: A cross-sectional study in Turkey. *Cancer Control*, 31, 10732748241236327.
10. Ding, H., Simmich, J., Vaezipour, A., Andrews, N., & Russell, T. (2024). Evaluation framework for conversational agents with artificial intelligence in health interventions: A systematic scoping review. *Journal of the American Medical Informatics Association*, 31(3), 746-761.
11. Izadi, S., & Forouzanfar, M. (2024). Error correction and adaptation in conversational AI: A review of techniques and applications in chatbots. *AI*, 5(2), 803-841.
12. Javaid, M., Bhandari, M., Parameshwari, P., Reddiboina, M., & Prasad, S. (2024). Evaluation of ChatGPT for patient counseling in kidney stone clinic: A prospective study. *Journal of Endourology*, 38(4), 377-383.
13. Ji, S., Li, X., Sun, W., Dong, H., Taalas, A., Zhang, Y., Gao, J., Choi, Y., Berg-Kirkpatrick, T., Huang, J., & Marttinen, P. (2024). A unified review of deep learning for automated medical coding. *ACM Computing Surveys*, 56(12), 1-41.
14. Kaphingst, K. A., Kohlmann, W. K., Chambers, R. L., Bather, J. R., Goodman, M. S., Bradshaw, R. L., Hooker, G. W., Graves, K. D., Pappas, L. M., Presson, A. P., & Buys, S. S. (2024). Uptake of cancer genetic services for chatbot vs standard-of-care delivery models: The BRIDGE randomized clinical trial. *JAMA Network Open*, 7(9), e2432143-e2432143.
15. Khan, M. M., Haque, M. R., Rahman, M. S., Islam, M. S., & Ahmed, F. (2025). Towards secure and trusted AI in healthcare: A systematic review of emerging innovations and ethical challenges. *IEEE Access*, 13, 15847-15865.
16. Khosravi, M., & Azar, G. (2024). Factors influencing patient engagement in mental health chatbots: A thematic analysis of findings from a systematic review of reviews. *Digital Health*, 10, 20552076241247983.
17. Lal, M. (2024). An analysis of deep learning models for conversational agents in healthcare. *Neural Computing and Applications*, 36(1), 587-610.
18. Soomro, M. A., Memon, R. N., Chandio, A. A., Leghari, M., & Soomro, M. H. (2024). A dataset of Roman Urdu text with spelling variations for sentence level sentiment analysis. *Data in Brief*, 57, 111170.
19. Laymouna, M., Ma, Y., Lessard, D., Schuster, T., Engler, K., & Lebouché, B. (2024). Roles, users, benefits, and limitations of chatbots in health care: Rapid review. *Journal of Medical Internet Research*, 26, e56930.
20. Ma, Y., Achiche, S., Pomey, M. P., Paquette, J., Adjoutah, N., Vicente, S., Lessard, D., Engler, K., & Lebouché, B. (2024). Adapting and evaluating an AI-based chatbot through patient and stakeholder engagement to provide information for different health conditions: Master protocol for an adaptive platform trial (the MARVIN Chatbots Study). *JMIR Research Protocols*, 13(1), e54668.
21. Maqbool, B., & Herold, S. (2024). Potential effectiveness and efficiency issues in usability evaluation within digital health: A systematic literature review. *Journal of Systems and Software*, 208, 111881.

22. Martins, A., Silva, P., & Costa, J. (2024). Unlocking human-like conversations: Scoping review of automation techniques for personalized healthcare interventions using conversational agents. *International Journal of Medical Informatics*, 175, 104732.
23. Oad, A., Soomro, M. A., Shaikh, M. K., & Memon, S. N. (2021). Effective User Assistant for Online Help System. *International Journal (Toronto, Ont.)*, 10(2).
24. Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. (2018). Deep learning for healthcare: Review, opportunities and challenges. *Briefings in Bioinformatics*, 19(6), 1236-1246.
25. Orum, T. G. (2023). Data privacy and ethical issues in collecting health care data using artificial intelligence among health workers. *International Journal of Healthcare Management*, 16(2), 145-152.
26. Rainey, J. P., Treu, E. A., Campbell, K. J., Blackburn, B. E., Pelt, C. E., Archibeck, M. J., Saltzman, C. L., & Anderson, L. A. (2024). Conversational engagement using a short message service chatbot after total joint arthroplasty. *Arthroplasty Today*, 30, 101484.
27. Ruotsalainen, P., Blobel, B., Seppälä, A., Sorvari, H., & Nykänen, P. (2019). Digital pHealth – Problems and solutions for ethics, trust and privacy. *Studies in Health Technology and Informatics*, 264, 1306-1310.
28. Sarella, P. N. K., & Mangam, V. T. (2024). AI-driven natural language processing in healthcare: Transforming patient-provider communication. *Indian Journal of Pharmacy Practice*, 17(1), 45-52.
29. Schachner, T. (2020). Artificial intelligence-based conversational agents for chronic conditions: Systematic literature review. *Journal of Medical Internet Research*, 22(10), e20012.
30. Talyshinskii, A., Naik, N., Hameed, B. Z., Juliebø-Jones, P., & Somani, B. K. (2024). Potential of AI-driven chatbots in urology: Revolutionizing patient care through artificial intelligence. *Current Urology Reports*, 25(1), 9-18.