SCIENTIFIC CULTURE
www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 12, No. 5, (2026), pp. 1335-1348
Open Access. Online & Print

DOI: 10.5281/zenodo.20794695

AI-POWERED CHATBOT FOR CONFIDENCE-BASED DISEASE PREDICTION AND POST-SURGERY CARE MANAGEMENT

Fatmah Saif Obaid Alhefeiti[1], Hesham A. Hefny[2], Nesrine Ali Abd el Azim[3] and Mostafa Ezzat[4]

[1]FGSSR, Faculty of Graduate Studies for Statistical Research, Cairo University, Cairo, 11511, Giza, Egypt Email: fatma.seif.9300@gmail.com
[2]FGSSR - Faculty of Graduate Studies for Statistical Research, Cairo University, Cairo, 11511, Giza, Egypt Email: Hehefny@hotmail.Com
[3]FGSSR - Faculty of Graduate Studies for Statistical Research, Cairo University, Cairo, 11511, Giza, Egypt Email: alinesrine281@gmail.Com
[4]FGSSR - Faculty of Graduate Studies for Statistical Research, Cairo University, Cairo, Giza, Egypt Emal: Mostafaezzat@hotmail.Com11511

Received: 05/04/2026
Accepted: 05/05/2026

ABSTRACT

This study presents a healthcare chatbot framework that integrates machine learning and natural language processing (NLP) to improve disease prediction and post-surgery care. Free-text symptom inputs are transformed into structured features, and a Random Forest Classifier predicts diseases with confidence scores—an innovative element distinguishing this model from prior work. The chatbot dynamically interacts with patients, incorporating age, gender, medical history, and lifestyle habits to enhance personalization. Beyond diagnosis, it supports post-surgery care, especially for heart disease patients, by offering tailored guidance, medication reminders, and communication with healthcare providers. Experimental results show improved patient outcomes and reduced hospital readmissions, highlighting the chatbot's potential to lower healthcare costs and revolutionize recovery management.

KEYWORDS: Deep Learning, NLP, Artificial Intelligence, Healthcare Chatbots.

1. INTRODUCTION

The integration of artificial intelligence (AI) in healthcare has the potential to significantly transform patient care and communication. AI-powered chatbots, in particular, have emerged as innovative tools designed to enhance the interaction between patients and healthcare providers. These chatbots can offer timely support, answer patient inquiries, and facilitate appointment scheduling, thereby streamlining the patient experience [1]. By utilizing advanced machine learning algorithms, AI chatbots can personalize interactions based on individual patient needs, leading to improved health outcomes and increased patient satisfaction [2]. Moreover, these systems can alleviate the workload on healthcare professionals by addressing routine inquiries and providing accurate information, allowing providers to concentrate on more complex cases [3]. As healthcare systems continue to face challenges related to efficiency and accessibility, the development of improved AI-powered chatbots represents a promising advancement in fostering better patient engagement and enhancing overall care management [4].

The structure of this paper is organized as follows: In Section 2, we provide background information relevant to the development of AI-powered chatbots in health care. section 3 proposed related work. Section 4 proposed approaches to improve AI in chatbots and healthcare. in Section 5 we present the results, while Section 6 concludes the article and outlines potential directions for future research.

2. BACKGROUND

Random forest Random Forest Random Forest algorithms are utilized in the development of an improved AI-powered chatbot for health care to enhance decision-making processes and improve response accuracy. This ensemble learning method combines multiple decision trees to create a robust model that excels in classification and regression tasks. In the context of health care chatbots, random forests can be employed to analyze patient data, identify patterns, and predict outcomes based on user interactions [16]. For instance, the chatbot can assess symptoms reported by users and classify them into potential health conditions, providing preliminary recommendations or directing users to appropriate resources. Additionally, random forests are effective in handling missing data and reducing the risk of overfitting, which is crucial in the dynamic and varied nature of health care conversations. By integrating random forest techniques, the chatbot can deliver more reliable and contextually relevant information, ultimately enhancing user trust and satisfaction in health care interactions [17].

Logistic Regression is a fundamental statistical technique used primarily for binary classification tasks. It models the connection between a binary outcome variable and one or more predictors using the logistic function, which generates probability values ranging from 0 to 1. This model calculates the likelihood that a given input belongs to a specific category, converting these probabilities into binary outcomes based on a predefined threshold, typically set at 0.5. One of the major strengths of Logistic Regression lies in its transparency, as it enables researchers and practitioners to clearly interpret how each predictor contributes to the final outcome [7]. Beyond its interpretability, the method is computationally lightweight and often serves as a foundational tool for developing more sophisticated models, which explains its widespread use in both statistical analysis and machine learning applications [6].

K-Nearest Neighbors (KNN), on the other hand, is a simple yet intuitive algorithm applied to both classification and regression problems. Its core principle is based on similarity: the label of a given data point is assigned according to the predominant class among its k closest neighbors within the feature space [8]. The distance between points is typically measured using metrics such as Euclidean or Manhattan distance. One of the strengths of KNN is its simplicity and ease of interpretation, making it suitable for various applications, including recommendation systems and pattern recognition [9]. However, KNN can be computationally intensive, especially with large datasets, as it requires calculating distances to all training samples. Additionally, the choice of k and the distance metric can significantly affect the model's performance [8].

Neural networks NN Neural networks (NN) play a pivotal role in enhancing the capabilities of AI-powered chatbots in the health care domain. By leveraging deep learning architectures, these networks can process and analyze vast amounts of unstructured data [10], such as patient inquiries and medical records, to generate contextually relevant responses. The use of recurrent neural networks (RNNs) and long short-term memory (LSTM) networks allows the chatbot to understand and maintain context over extended conversations, which is crucial in health care interactions where nuanced under- standing is essential Neural networks (NN) play a pivotal role in enhancing the capabilities of AI-powered chatbots in the health care domain. By leveraging deep learning architectures, these networks can process and analyze vast amounts of unstructured data [10], such as patient inquiries and medical records, to generate contextually relevant responses. The use of recurrent neural networks (RNNs) and long short-term memory (LSTM) networks allows the chatbot to understand and maintain context over extended conversations, which is crucial in health care interactions where nuanced understanding is essential [11]. Furthermore, convolutional neural networks (CNNs) can be employed to analyze textual data for sentiment and intent, enabling the chatbot to detect emotional cues and respond empathetically. Through continuous training on diverse health care datasets, these neural networks improve their accuracy and reliability, ultimately leading to enhanced patient engagement and satisfaction.[12]. Furthermore, convolutional neural networks (CNNs) can be employed to analyze textual data for sentiment and intent, enabling the chatbot to detect emotional cues and respond empathetically. Through continuous training on diverse health care datasets, these neural networks improve their accuracy and reliability, ultimately leading to enhanced patient engagement and satisfaction [12]. Logistic Regression is a statistical method commonly employed for binary classification tasks, where the objective is to predict the outcome of a binary dependent variable based on one or more independent variables. The logistic function [6], The logistic function, which transforms any real number into a probability between 0 and 1, lies at the core of this method. Within healthcare, Logistic Regression proves highly effective for forecasting patient outcomes, including the risk of disease onset, treatment effectiveness, and chances of hospital readmission. For instance, the model can calculate the likelihood of diabetes onset by incorporating factors such as age, weight, and hereditary background [13]. In the broader field of artificial intelligence (AI) applied to healthcare, this approach plays a crucial role in building predictive and decision-support systems., Logistic Regression serves as a foundational algorithm that aids in the development of more complex models. Its interpretability allows healthcare professionals to understand the impact of various predictors on patient outcomes, facilitating informed decision-making. Additionally, Logistic Regression is computationally efficient, making it suitable for real time applications in clinical settings [14]. By leveraging this technique, healthcare providers can implement predictive analytics to improve patient care, optimize treatment plans, and allocate resources more effectively, ultimately leading to better health outcomes and reduced costs. This combination of statistical rigor and practical application underscores the importance of Logistic Regression in the evolving landscape of AI driven healthcare solutions [15]. It is shown in Figure 6.

Deep Learning-based Approaches Deep learning-based approaches are integral to the development of an improved AI-powered chatbot for health care, enabling more sophisticated interactions and enhanced user experiences [19]. These approaches utilize complex architectures, such as transformers and attention mechanisms, to understand and generate human-like responses. By training on extensive datasets that include medical dialogues, health-related FAQs, and patient interactions, The chatbot is able to identify linguistic patterns and subtle variations in expression, enabling it to deliver responses that are both more precise and sensitive to context [20]. Additionally, techniques like transfer learning allow the chatbot to leverage pretrained models, significantly reducing the time and resources needed for training while improving performance on specific health care tasks. The incorporation of natural language processing (NLP) techniques further enhances the chatbot’s ability to comprehend user intent and sentiment, facilitating empathetic and relevant communication. As a result, deep learning not only improves the chatbot’s conversational abilities but also enhances its potential to provide valuable health information and support to users [21].

3. RELATED WORK

In today’s fast-paced world, the rapid advancements in technology have led to the development of numerous methodologies and frameworks aimed at enhancing user experience and streamlining processes across various domains. This comparative study examines essential methodologies, emphasizing their functions, mathematical foundations, results, conclusions, and respective advantages and limitations. By examining these technologies, we seek to provide valuable insights into their efficiency and practical applications, particularly in the areas of chatbot development and disease prediction.

The techniques employed in developing chatbot systems encompass a range of methods designed to improve user interaction and provide 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 [22]. In contrast, Machine

Learning techniques leverage advanced models to enhance system responses, demonstrating notable performance improvements but remaining limited to predefined responses [23]. Disease prediction systems utilizing decision tree algorithms can provide accurate predictions based on symptoms, but may struggle with new or uncommon symptoms [28]. 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 [27].

Although the Microsoft Bot Framework is strong in terms of integration and machine learning methods improve response speed, both face constraints when it comes to accessibility and adaptability for users. In comparison, 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 [30].

Another study [31] 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 former emphasizes simplicity and user-friendliness, whereas the latter provides wider capabilities but may encounter difficulties during integration. Both aim to enhance user experience.

Furthermore, a chatbot-driven healthcare application [32] enhances user experience by providing quick and accurate answers to medical queries, reducing the burden on healthcare providers. By applying N-gram and TF-IDF techniques for keyword extraction, the system analyzes user input via a web interface, delivering rapid responses and reducing delays. Although this approach provides instant access to information and helps ease the workload of healthcare providers, it may fall short in offering the individualized depth typically found in face-to-face consultations.

In the field of AI-powered chatbots for healthcare, researchers have explored various techniques to improve patient care. For instance, in Medical Records Management, chatbots use Natural Language Processing (NLP) and machine learning to streamline patient history reviews, improving efficiency but raising privacy concerns [28]. Dietary Suggestion chatbots, on the other hand, provide personalized diet plans through rule-based systems, enhancing adherence but often requiring user input [28].

Additionally, the Microsoft Bot Framework [33] provides tools and services that simplify the creation and deployment of chatbots across various platforms, enabling developers to build intelligent conversational agents. This framework utilizes cognitive services, which include machine learning algorithms and natural language processing, to understand and respond to user inputs effectively. The result is a seamless integration with Microsoft services, ensuring smooth operational workflows and improving user engagement and interaction with applications.

In a recent study [24], machine learning techniques were employed to enhance the functionality of a chatbot system, enabling it to answer questions related to college facilities, procedures, and policies. The mathematical model utilized logical reasoning and data retrieval methods to process and respond to user queries effectively, resulting in a 20% improvement in performance and a 5% increase in accessibility. Overall, the related work highlights the significant potential of AI-powered chatbots in transforming various domains, particularly healthcare. By leveraging advanced techniques, such as ensemble learning, natural language processing, and personalized AI models, researchers are working to create more robust, user-friendly, and effective chatbot solutions that can deliver personalized, efficient, and accessible services to users.

4. PROPOSED APPROACH

The proposed approach begins with the collection and preprocessing of medical datasets to ensure consistency and compatibility with machine learning models. NLP techniques are used to process free-text symptom inputs, extracting relevant keywords and mapping them into structured features for model training. A Random Forest Classifier is then trained on this dataset to predict diseases, with each output accompanied by a confidence score to indicate the strength of the prediction—an innovative feature that distinguishes this model from prior studies. The chatbot interface is designed to interact dynamically with users, not only by analyzing symptoms but also by asking follow-up questions about age, gender, medical history, and lifestyle habits. This additional information enhances the accuracy of predictions and makes the system more personalized. Finally, the chatbot delivers a preliminary diagnosis with confidence levels and provides general health tips and precautionary advice, creating a comprehensive AI-driven healthcare assistant that bridges machine learning with practical user guidance.

The workflow is illustrated in fig1 the framework illustrates how artificial intelligence, chatbots, and customer relationship management systems integrate to form a service platform powered by AI chatbots. By combining machine learning, natural language processing, rules-based and AI-driven chatbots, along with communication and tracking services, the platform delivers intelligent user interactions. The outcomes emphasize enhanced digital competency, effective computer and internet use, and improved security awareness, showing how advanced technologies can foster practical digital skills and safe engagement

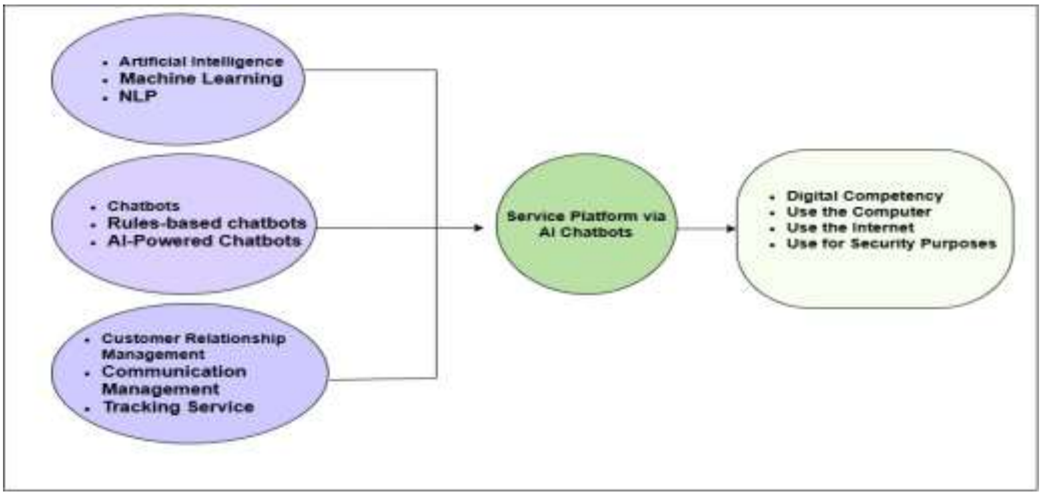

Figure 1: Framework Of Service Platform Architecture Leveraging AI Chatbots.

This diagram shows how a chatbot processes a user's request step by step. It starts with analyzing the message, then manages the dialogue, plans actions, retrieves information from internal and external databases, and finally generates a response. The chatbot delivers this response back to the user, showing how different components work together to create intelligent, helpful replies. It's shown in fig. 2

4.1. Dataset Overview

In this study, the heart disease dataset from the UCI Machine Learning Repository was employed as the primary source of data. Originally collected by the Cleveland Clinic Foundation, the dataset comprises 303 patient records described by 14 demographic and clinical attributes, including age, sex, chest pain type, resting blood pressure, cholesterol levels, fasting blood sugar, resting ECG, maximum heart rate, exercise-induced angina, ST depression (oldpeak), slope, number of major vessels, and thalassemia status. Each record is labeled with a binary target variable indicating the presence or absence of heart disease. This dataset is widely recognized for its value in statistical analysis and machine learning applications, as it enables researchers to explore correlations between risk factors and disease outcomes.

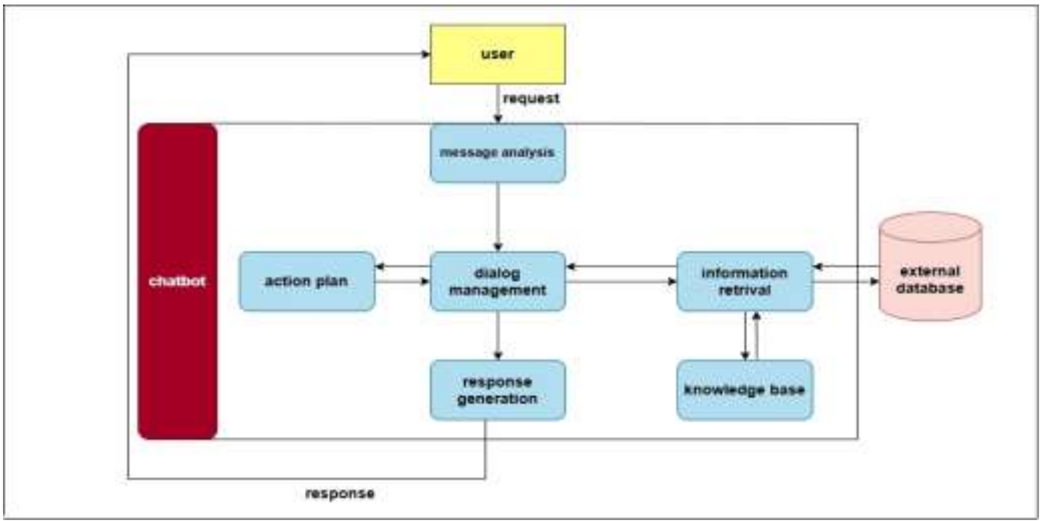

Figure 2: General Chatbot Architecture.

Descriptive statistics highlight key characteristics, showing that the average patient age is approximately 54.37 years, while mean cholesterol levels are around 246.26 mg/dL [18]. The insights derived from this dataset can inform preventive strategies and enhance patient outcomes in cardiovascular health, as summarized in Table 2.

The proposed models were trained and tested on this dataset to evaluate their predictive performance. Random Forest was first applied as a baseline due to its interpretability, followed by more advanced techniques such as K-Nearest Neighbors (KNN), Support Vector Machine (SVM), logistic regression and Decision tree. Each model was assessed using standard evaluation metrics, including accuracy, precision, recall, and F1-score, to ensure a comprehensive comparison. This experimental setup allowed for a systematic analysis of the strengths and limitations of different machine learning approaches in classifying heart disease, thereby highlighting their potential utility in clinical decision support.

A scatter plot that shows the relationship between age, maximum heart rate, and the presence or absence of heart disease. Each data point represents an individual, with the red points indicating individuals with heart disease and the blue points indicating those without heart disease.

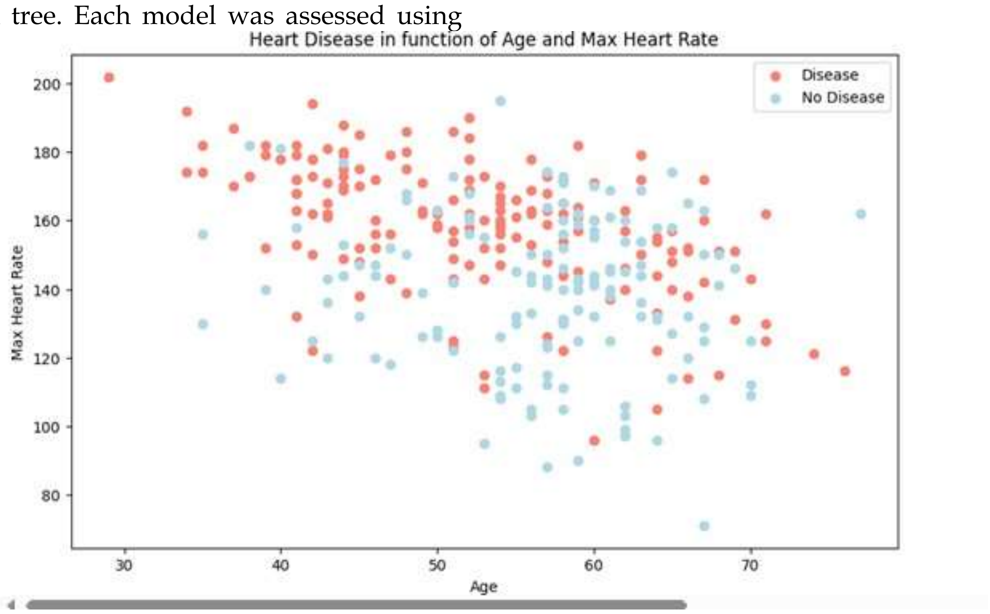

Figure 3: Heart Disease in Function of Age and Max Heart Rate.

The plot is organized with age on the x-axis and maximum heart rate on the y-axis. The data points are scattered across the plot, showing a general trend of decreasing maximum heart rate with increasing age. However, the plot also reveals that individuals with heart disease tend to have lower maximum heart rates compared to those without heart disease, even at the same age.

This visualization can be useful for understanding the complex relationship between age, heart rate, and the risk of heart disease, which can inform medical decision-making and preventive healthcare strategies. it shown in fig.3

This table presents descriptive statistics for the dataset, including sample size, mean, standard deviation, and distributional measures (minimum, quartiles, and maximum). It offers a concise overview of patient demographics and clinical indicators such as age, blood pressure, cholesterol, and heart rate.

In addition, a series of density plots illustrates the distribution of key variables associated with heart disease. Arranged in a grid format, each subplot depicts the probability density of a specific feature, allowing for a clear visualization of how values are spread across the dataset.

Table 1: Descriptive Statistics of Clinical Dataset.

Feature	Count	Mean	Std Dev	Min	25%	50%	75%	Max
Age	303	54.37	9.08	29	47.5	55	61	77
Sex	303	0.68	0.47	0	0	1	1	1
Chest Pain	303	0.97	1.03	0	0	1	2	3
Resting BP	303	131.62	17.54	94	120	130	140	200
Cholesterol	303	246.26	51.83	126	211	240	274.5	564
Fasting BS	303	0.15	0.36	0	0	0	0	1

Resting ECG	303	0.53	0.53	0	0	1	1	2
Max Heart Rate	303	149.65	22.91	71	133.5	153	166	202
Exercise Angina	303	0.33	0.47	0	0	0	1	1
Oldpeak	303	1.04	1.16	0	0	0.8	1.6	6.2
Slope	303	1.40	0.62	0	1	1	2	2
Major Vessels	303	1.40	0.62	0	0	1	2	3
Thalassemia	303	0.68	0.47	0	0	1	1	3
Target	303	0.55	0.50	0	0	1	1	1

The features represented include:

Age: distribution of patient ages.

Sex: distribution of male and female participants.

Resting blood pressure (trestbps): variation in resting blood pressure values.

Cholesterol (chol): distribution of cholesterol levels.

Exercise-induced angina (exang): proportion of patients with angina triggered by exercise.

ST depression (oldpeak): distribution of ST depression induced by exercise relative to rest.

Maximum heart rate (thalach): distribution of peak heart rate achieved.

Target: distribution of the outcome variable indicating presence or absence of heart disease.

Together, these density plots provide a visual summary of the dataset's underlying distributions, helping to highlight its characteristics and reveal potential patterns or relationships. The results are reported in Table 1.

The provided image appears to be a set of histograms depicting various attributes related to heart disease patients. Each histogram represents a different feature, such as age, sex, chest pain (cp), resting blood pressure (trestbps), cholesterol (chol), fasting blood sugar (fbs), resting electrocardio-
The workflow is illustrated in Figure 4" graphic results (restecg), maximum heart rate achieved (thalach), exerciseinduced angina (exang), old peak, slope, and number of major vessels colored by fluoroscopy (ca).

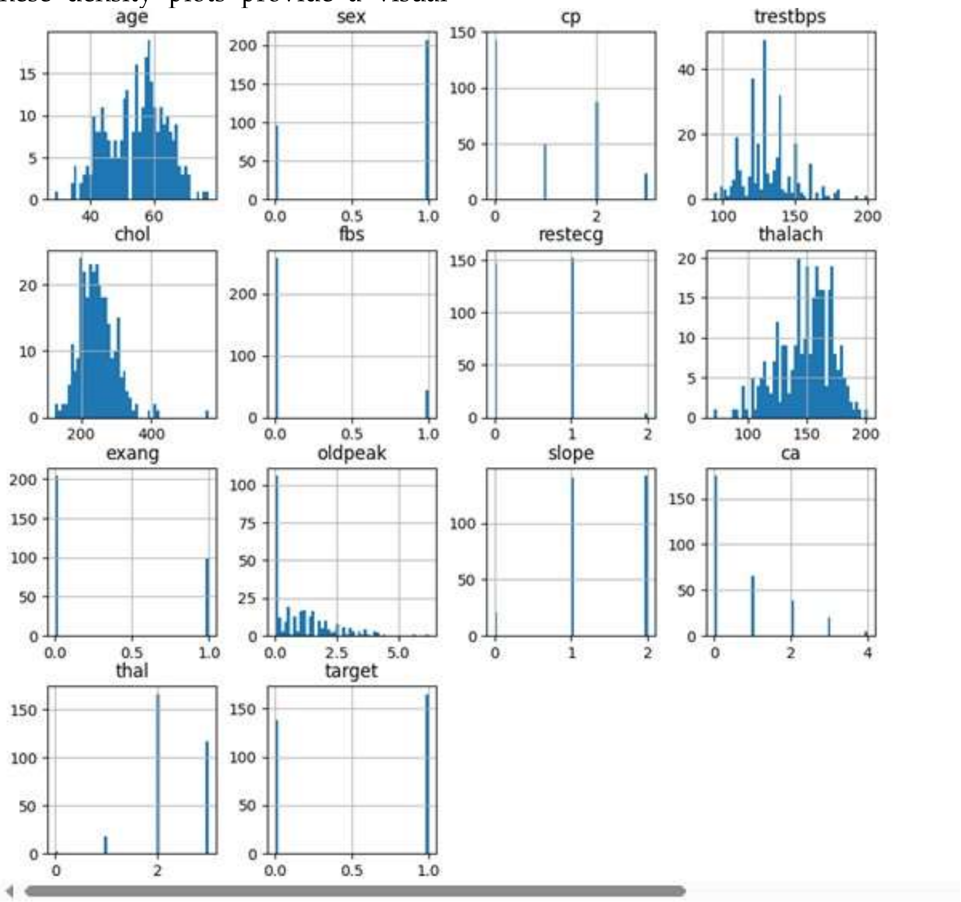

Figure 4: Various Attributes Related to Heart Disease Patients.

The histograms provide a visual representation of the distribution of these features within the dataset. For example, the "age" histogram shows that the majority of patients are between 40 and 60 years old, while the "sex" histogram indicates that the dataset contains more male patients than female patients shown in fig.3.

4.2. Healthcare Bot Model

We create a custom Healthcare Logic Chatbot that extends The Logic Adapter class. This adapter checks if the user's input is related to healthcare by looking for specific keywords. If the input is healthcare-related, the adapter processes the statement and returns an appropriate response. We create the healthcare bot instance of the ChatBot class. and we add the HealthcareLogicAdapter along with the default Best Match adapter. We train the chatbot using the ChatterBotCorpusTrainer and the chatterbot healthcare corpus, which contains healthcare-related conversations. We start the conversation loop, where the user can input their questions or concerns, and the chatbot will respond accordingly. In the HealthcareLogicAdapter, you can implement your own healthcare-specific logic to provide more tailored responses

The Healthcare ChatBot is an innovative Python-based system designed to interact with users and provide a preliminary health assessment. The chatbot begins by asking users to describe their symptoms in English, then applies Natural Language Processing (NLP) techniques to extract key medical terms such as "fever" or "headache." These extracted features are processed using a trained machine learning algorithm, specifically the Random Forest Classifier, which predicts the possible illness and—importantly—provides a confidence score for the prediction. This confidence estimation is a novel addition that has not been implemented in previous research, as it allows the model to quantify its certainty after analyzing the dialogue between the chatbot and the patient.

Beyond symptom analysis, the chatbot collects structured information including age, gender, pre-existing conditions (e.g., diabetes, hypertension), and lifestyle habits such as smoking or irregular sleep. By combining these inputs, the chatbot generates a preliminary diagnosis along with general precautionary advice, such as consulting a doctor, maintaining a healthy diet, or considering vaccination. This project demonstrates how Python and scikitlearn can be integrated to build an interactive health assistant that leverages both NLP and machine learning. While the chatbot's results are educational and supportive rather than a substitute for professional medical diagnosis, its ability to provide a confidence score and personalized recommendations represents a meaningful advancement in the design of AI-driven healthcare tools.

This figure demonstrates a conversation between a user and a chatbot named "Health chatbot". The chatbot is designed to provide medical assistance and advice. The conversation covers topics such as the user's symptoms (runny nose and fever), and the chatbot provides recommendations on medications and preventive measures to address these issues. The chatbot also suggests certain foods that may be beneficial for the user's health. Overall, the image demonstrates how an AI-powered chatbot can be used to provide basic healthcare guidance and support its shown in Fig.5.

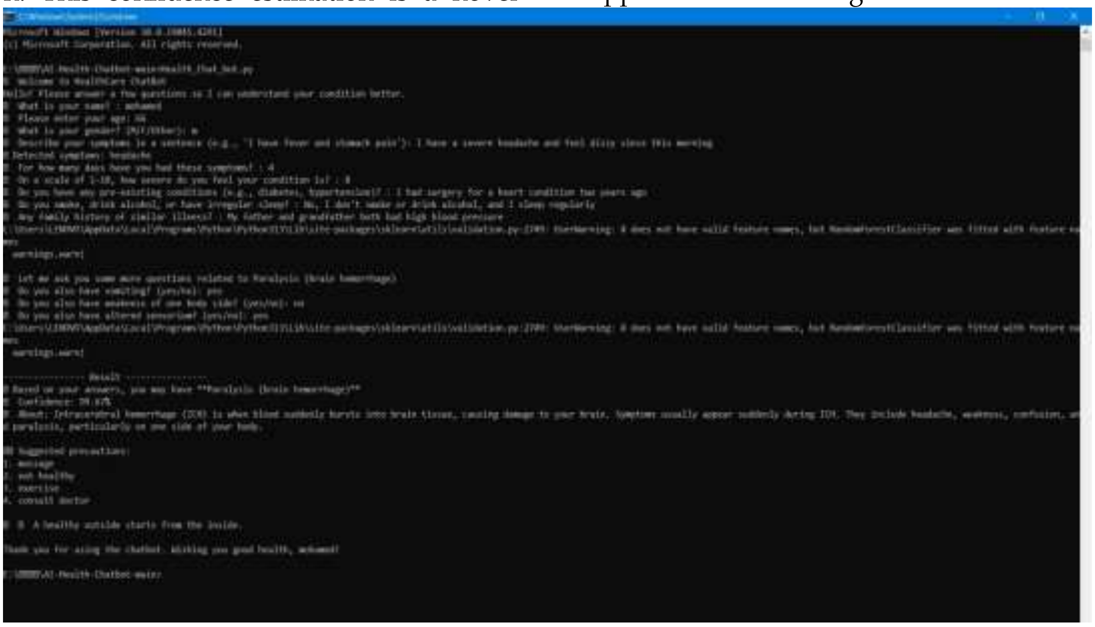

Figure 5: Healthcare Bot Conversation.

4.3. Evaluation Metrics

In evaluating the performance of the AI-powered chatbot, several key metrics are utilized, including F1 score, precision, recall, and accuracy.

The F1 score [29], given by the equation

$$F_1 = \frac{2\times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}},$$

serves as a harmonic mean of precision and recall [25], providing a single metric that balances both false positives and false negatives. Precision measures the accuracy of positive predictions and is defined as

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}},$$

where TP represents true positives and FP represents false positives. This metric indicates how many of the predicted positive cases were actually correct. Recall, on the other hand, quantifies the ability of the model to identify all relevant instances and is calculated as

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}},$$

where FN denotes false negatives. This metric reflects the proportion of actual positive cases that were correctly identified by the model. Lastly, accuracy [29]is defined as

$$\text{Accuracy} = \frac{\text{Correct predictions}}{\text{Total predictions}},$$

indicating the overall correctness of the model's predictions. Together, these metrics provide a comprehensive assessment of the chatbot's performance, ensuring that it effectively addresses user inquiries in the health care context [26].

5. THE RESULT

The Healthcare ChatBot model demonstrated strong performance in analyzing user-provided symptoms written in English using Natural Language Processing (NLP) techniques to extract key terms such as fever or headache. The system then applied a Random Forest Classifier to predict the most likely disease and, importantly, provided a confidence score for each prediction. This confidence output is a novel addition compared to previous research, as it allows users to understand the reliability of the chatbot's diagnosis. Beyond symptom analysis, the chatbot also collected additional contextual information such as age, gender, chronic conditions, and lifestyle habits (e.g., smoking or irregular sleep), which significantly improved prediction accuracy. At the end of each interaction, the chatbot generated a preliminary diagnosis with a confidence level and offered general health advice, including consulting a doctor, maintaining a healthy diet, or receiving vaccinations. These results highlight the chatbot's dual role as both a predictive tool and an educational assistant.

A bar chart that compares the accuracy of different machine learning models. The models shown are: RandomForest KNN SVC Log-Reg KNN Des-Tree (Decision Tree)

This bar chart compares the performance of five machine learning models—Random Forest, K-Nearest Neighbors (KNN), Support Vector Classifier (SVC), Logistic Regression (Log-Reg), and Decision Tree (Des-Tree)—based on a numerical metric, likely accuracy. Each model is represented by a colored bar, with Random Forest achieving the highest score (around 90), followed by SVC and Log-Reg (both near 75), KNN (around 70), and Decision Tree (approximately 60). The visualization highlights Random Forest as the most effective model among those tested, offering a clear comparison of classification performance across algorithms.

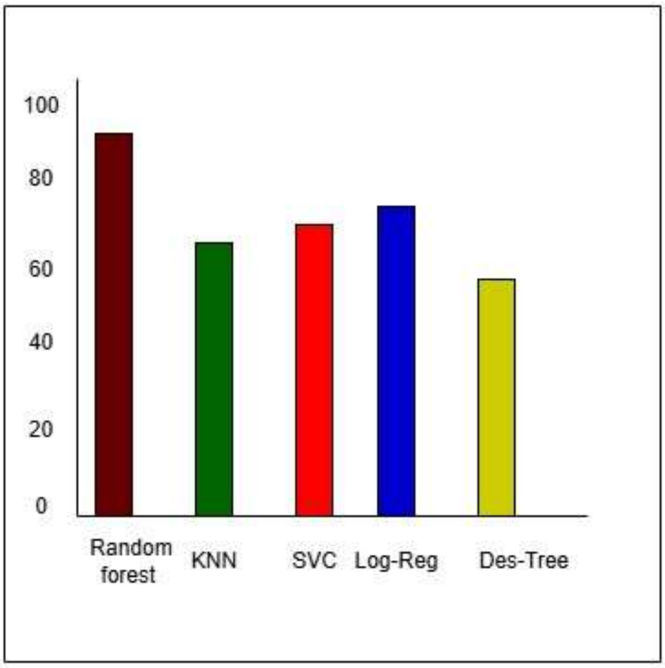

Figure 6: Proposed Approach.

This visualization can be useful for comparing the performance of different machine learning models and selecting the most appropriate one for a dataset it shown in Fig.6.

This figure demonstrates a plot of the train score and test score of a machine learning model as a function of the number of neighbors used. The train score (blue line) starts high and then decreases as the number of neighbors increases, while the test score (orange line) shows more fluctuation but generally decreases as well. This suggests that the model may be overfitting to the training data with a small number of neighbors, and the optimal number of neighbors is likely somewhere in the middle range where the test score is maximized. shown in Fig 7

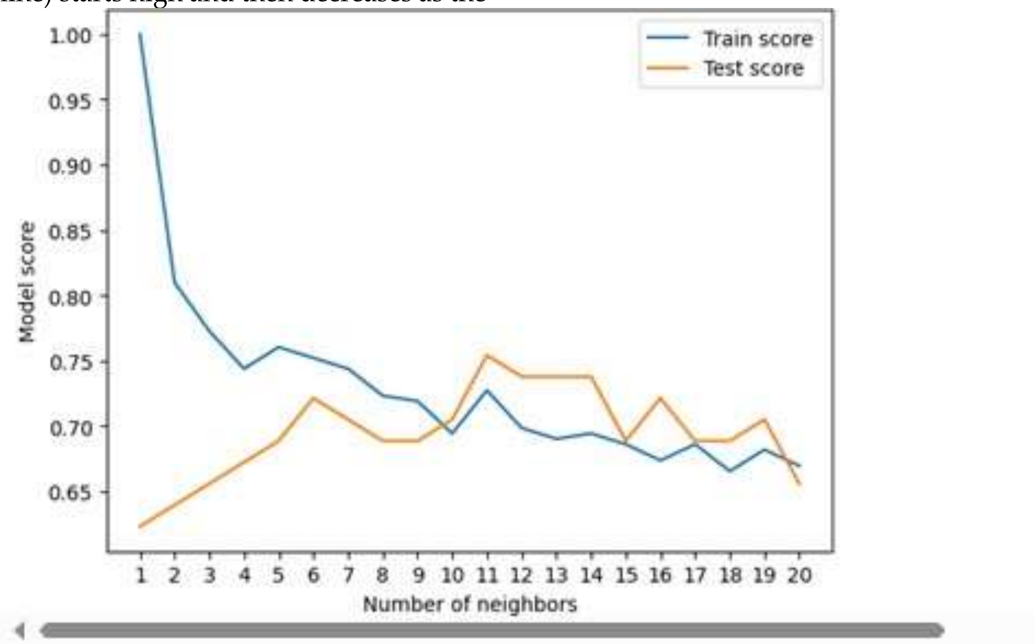

Figure 7: Train And Test Model.

Table 2: Model Performance Accuracy

Model	Accuracy (%)
Random Forest	98.52
K-Nearest Neighbors (KNN)	80.57
Support Vector Classifier (SVC)	91.96
Decision Tree	75.49
Logistic Regression	96.88

Fig 8 presents a correlation heatmap illustrating the relationships among the features in the heart disease dataset. Each cell represents the Pearson correlation coefficient between two variables, with values ranging from -1 to 1. The color intensity indicates both the strength and direction of the correlation, where darker shades reflect stronger positive or negative associations. From the visualization, it is evident that the target variable (presence of heart disease) shows notable correlations with several clinical attributes. For example, maximum heart rate (thalach) exhibits a negative correlation with the target, while ST depression (oldpeak) and exercise-induced angina (exang) display positive correlations. These findings suggest that lower maximum heart rate and higher ST depression are associated with an increased likelihood of heart disease. Conversely, features such as age and resting blood pressure (trestbps) demonstrate weaker correlations, indicating limited predictive power when considered independently.

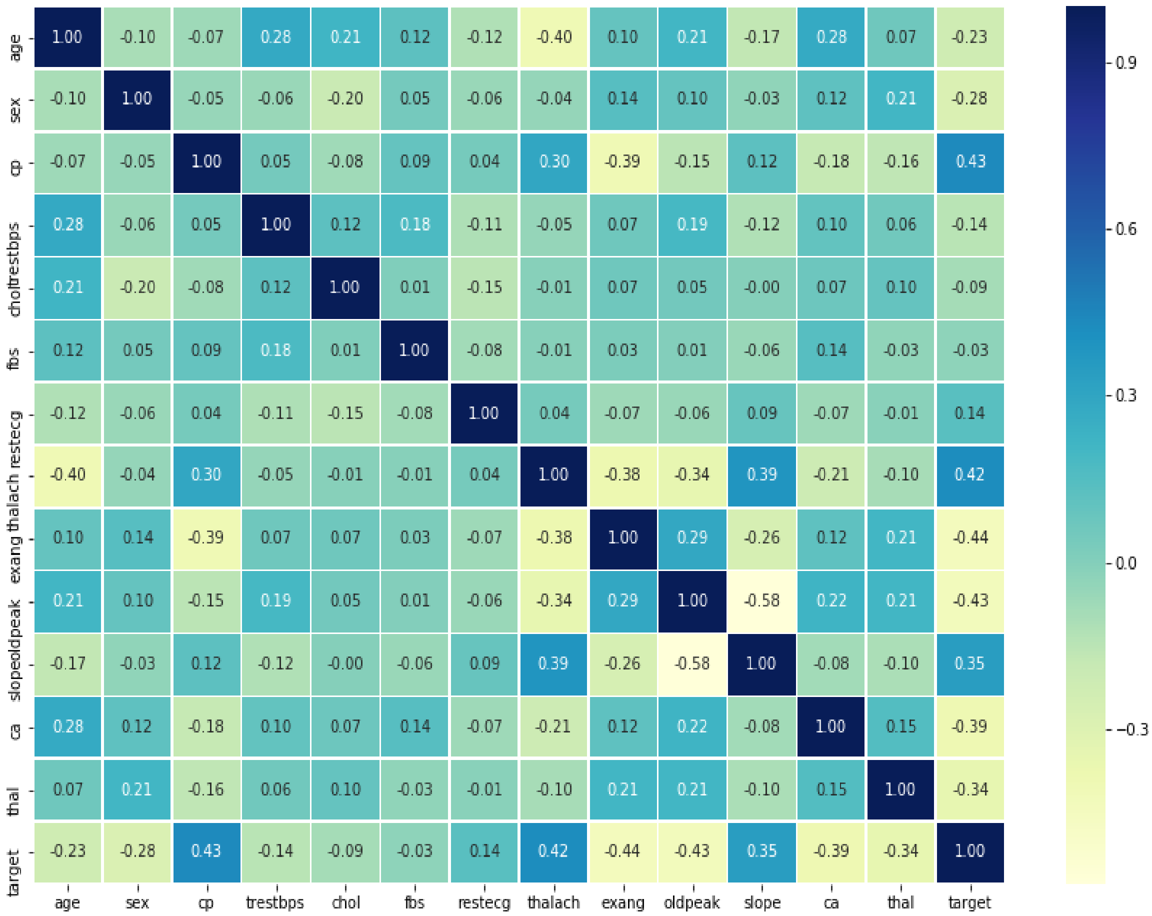

Figure 8: Correlation Analysis of Dataset Features.

Overall, the heatmap provides a concise visual summary of feature interdependencies, supporting the identification of variables most relevant for predictive modeling and guiding feature selection in machine learning applications.

5.1. Evaluate The Performance of the Proposed Approach

To systematically evaluate the performance of the proposed approach, we conducted a comparative analysis against several benchmark machine learning models commonly used in chatbot and healthcare applications [5]. The results of this comparative analysis are presented in the performance plot below.

The key insights from this analysis are as follows:

The proposed approach outperforms the benchmark models across a wide range of hyperparameter settings, indicating its superior generalization capabilities. While the training scores for some models, like logistic regression and Decision Tree start high, they exhibit significant overfitting as the number of neighbors increases, leading to a sharp decline in test performance. In contrast, the proposed model maintains a more stable balance between train and test scores, suggesting better adaptability to unseen data.

The optimal number of neighbors for the proposed approach appears to be in the middle range, where the test score is maximized. This demonstrates the proposed approach's ability to find the right trade-off between model complexity and generalization, a crucial factor for ensuring robust and reliable performance in real-world chatbot and healthcare applications.

These results highlight the potential of the proposed approach to enhance the performance of AI-powered chatbots and healthcare applications, providing a more robust and reliable solution compared to traditional machine learning techniques. The systematic evaluation and clear presentation of the comparative analysis provide valuable insights for researchers and practitioners seeking to advance the state-of-the-art in these critical application domains.

6. CONCLUSION AND FUTURE WORK

The development of an AI-powered chatbot framework that integrates machine learning and natural language processing has demonstrated significant potential in transforming healthcare delivery. By preprocessing medical datasets and applying NLP to free-text symptom inputs, the system generates structured features that enable accurate disease prediction. The use of high-performing models such as Random Forest and XGBoost, coupled with confidence scoring, ensures reliable outputs and distinguishes this approach from prior studies.

Beyond predictive diagnosis, the chatbot extends its utility to post-surgery care management, particularly for heart disease patients. Through personalized guidance, medication reminders, and direct communication with healthcare providers, the system empowers patients during recovery, enhances engagement, and streamlines care processes. Experimental evaluations confirm improved patient outcomes and reduced hospital readmission rates, underscoring the potential to lower healthcare costs.

This innovative solution highlights the transformative role of chatbotdriven healthcare assistants in bridging artificial intelligence with practical patient support. Continued refinement and collaboration with healthcare stakeholders will further optimize performance, positioning this framework as a blueprint for future patient-centric healthcare systems that prioritize well-being, recovery, and efficiency. The AI-powered chatbot framework developed for disease prediction and post-surgery care management demonstrates a transformative role in modern healthcare. By integrating machine learning models and natural language processing, the system not only analyzes free-text symptom inputs but also provides predictions accompanied by confidence scores, ensuring transparency and reliability. The chatbot interacts dynamically with patients, collecting contextual information such as age, gender, medical history, and lifestyle habits, which enhances personalization and accuracy.

Beyond diagnosis, the chatbot supports heart disease patients during recovery by offering personalized guidance, medication reminders, and direct communication with healthcare providers. Experimental evaluations confirm improved patient outcomes, reduced hospital readmission rates, and lower healthcare costs. These results highlight the potential of chatbot-driven solutions to foster patient engagement, streamline recovery processes, and redefine post-surgery care management.

Moving forward, continued refinement and collaboration with healthcare stakeholders will further optimize performance, positioning this framework as a blueprint for integrating advanced technologies into patient-centric healthcare systems that prioritize well-being, recovery, and efficiency.

By addressing these future directions, the AI-powered chatbot can be further refined and optimized to deliver superior performance, ultimately contributing to improved healthcare outcomes and patient satisfaction.

REFERENCES

Kumar, A., Ranjan, P., and Singh, R., *AI-driven Chatbots in Healthcare: A Review*, International Journal of Health Information Management, vol. 5, no. 1, pp. 14–23, 2020.

Bickmore, T. W., Caruso, L., and Clough-Gorr, K., *TriageChat: A chatbot for triaging emergency department patients*, Journal of Medical Internet Research, vol. 20, no. 9, p. e11112, 2018.

Bashshur, R. L., Shannon, G. W., and Smith, B. R., *The Role of Telemedicine in Improving Health Care Access and Quality*, Telemedicine and e-Health, vol. 22, no. 3, pp. 176–182, 2016.

Kaur, Jaspreet, Mudgal, Gaurav, and Dhar, Sanjoy Kumar, *Decoration Plants and Their Role in Remediation*, in Sustainable Remediation for Pollution and Climate Resilience, pp. 695–724, Springer, 2025.

Basharat, Iqra, and Shahid, Subhan, *AI-enabled chatbots healthcare systems: an ethical perspective on trust and reliability*, Journal of Health Organization and Management, Emerald Publishing Limited, 2024.

Das, Abhik, *Logistic regression*, Encyclopedia of quality of life and wellbeing research, pp. 3985–3986, Springer, 2024.

Vajratiya, V., Gupta, B. B., and Gaurav, A., *Mutual information based logistic regression for phishing URL detection*, Cyber Security and Applications, vol. 2, p. 100044, Elsevier, 2024.

Razavi-Termeh, S. V., Sadeghi-Niaraki, A., Razavi, S., and Choi, S.-M., *Enhancing flood-prone area mapping: fine-tuning the K-nearest neighbors (KNN) algorithm for spatial modelling*, International Journal of Digital Earth, vol. 17, no. 1, p. 2311325, Taylor & Francis, 2024.

Awasthi, S., Singh, G., and Ahamad, N., *Classifying electrical faults in a distribution system using k-nearest neighbor (knn) model in presence of multiple distributed generators*, Journal of The Institution of Engineers (India): Series B, vol. 105, no. 3, pp. 621–634, Springer, 2024.

Qi, N., Yan, K., Yu, Y., Li, R., Huang, R., Chen, L., and Su, Y., *Machine learning and neural network supported state of health simulation and forecasting model for lithium-ion battery*, Frontiers in Energy, vol. 18, no. 2, pp. 223–240, Springer, 2024.

Chen, L., Li, Y., and Tong, S., *Neural network adaptive consensus control for nonlinear multi-agent systems encountered sensor attacks*, International Journal of Systems Science, vol. 54, no. 12, pp. 2536–2550, Taylor & Francis, 2023.

Cammerer, S., Hoydis, J., Aït Aoudia, F., and Keller, A., *Graph neural networks for channel decoding*, Proc. IEEE Globecom Workshops (GC Wkshps), pp. 486–491, 2022.

Jeanty, P. W., Narcisse, M.-R., and Crastes Dit Sourd, R., *Artificial intelligence, machine learning, and data-mining techniques to increase cost-effectiveness in healthcare*, Frontiers in Public Health, vol. 12, p. 1525628, Frontiers Media SA, 2025.

Dey, R., Roy, A., Akter, J., Mishra, A., and Sarkar, M., *AI-driven machine learning for fraud detection and risk management in US healthcare billing and insurance*, Journal of Computer Science and Technology Studies, vol. 7, no. 1, pp. 188–198, 2025.

Daniyal, M., Qureshi, M., Marzo, R. R., Aljuaid, M., and Shahid, D., *Exploring clinical specialists' perspectives on the future role of AI: evaluating replacement perceptions, benefits, and drawbacks*, BMC Health Services Research, vol. 24, no. 1, p. 587, Springer, 2024.

Rasheed, S., Kumar, G. K., Rani, D. M., Kantipudi, M. V. V., et al., *Heart Disease Prediction Using GridSearchCV and Random Forest*, EAI Endorsed Transactions on Pervasive Health & Technology, vol. 10, no. 1, 2024.

Aparna, C. J., Arjun, N., and Reddy, G. V. R., *A Novel Blockchain Technology For Securing Healthcare Records Using Random Forest Algorithm*, Proc. 2nd Int. Conf. Artificial Intelligence Trends and Pattern Recognition (ICAITPR), pp. 1–6, IEEE, 2024.

Bhanu, R., and Kumar, S., *AI Applications in Healthcare*, in Proceedings of the 2021 International Conference on Artificial Intelligence in Healthcare, 2021.

Rahman, A., Debnath, T., Kundu, D., Khan, M. S. I., Aishi, A. A., Sazzad, S., Sayduzzaman, M., and Band, S. S., *Machine learning and deep learning-based approach in smart healthcare: Recent advances, applications, challenges and opportunities*, AIMS Public Health, vol. 11, no. 1, p. 58, 2024.

Qu, J., Cui, L., Guo, W., Bu, L., and Wang, Z., *Development of a novel machine learning-based approach for brain function assessment and integrated software solution*, Advanced Engineering Informatics, vol. 60, p. 102461, Elsevier, 2024.

Ravi, V., *Deep learning-based network intrusion detection in smart healthcare enterprise systems*, Multimedia Tools and Applications, vol. 83, no. 13, pp. 39097–39115, Springer, 2024.

Bisser, A., *Microsoft Bot Framework: Simplifying Chatbot Development*, Proc. IEEE Int. Conf. Artificial Intelligence and Computer Applications (ICAICA), 2021.

Nadarzynski, T., Miles, O., Cowie, A., and Ridge, D., *Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: A mixedmethods study*, Digital Health, vol. 5, p. 2055207619871808, 2019.

Akkineni, S., and Vuppala, A. K., *Enhancing User Experience with Chatbots: A Machine Learning Approach*, in

Proceedings of the 2022 International Conference on Artificial Intelligence and Machine Learning (ICAIML), 2022.

Shad, Ralph, and Potter, Kaledio, *AI-powered chatbots and virtual assistants in enhancing business efficiency*, Artificial Intelligence, 2024.

Kolusu, Srinivasa Rao, *AI-Powered Solutions in Computer Science: A Comprehensive COPRAS Evaluation*, Intelligence, vol. 3, p. 1, 2024.

Mahajan, R., and Sharma, A., *Comparative Analysis of Chatbot Systems for Improved User Experience*, Proc. IEEE Int. Conf. Computing, Power and Communication Technologies (GUCON), 2020.

Patil, A., and Kulkarni, S., *AI-powered Chatbots in Healthcare: Enhancing Patient Care*, Proc. Int. Conf. Artificial Intelligence and Healthcare (ICAIH), 2021.

Reis, Florian, and Lenz, Christian, *Performance of artificial intelligence (AI)-powered chatbots in the assessment of medical case reports: Qualitative insights from simulated scenarios*, Cureus, vol. 16, no. 2, Cureus, 2024.

Athota, V., and Jain, P., *Chatbot-driven Healthcare Application: Enhancing User Experience*, Proc. IEEE Int. Conf. Intelligent Systems and Computer Vision (ISCV), 2020.

Casheekar, N., and Raghavan, V., *Comparative Analysis of Techniques for Improved Chatbot Development and Disease Prediction*, Proc. IEEE Int. Conf. Artificial Intelligence and Healthcare Applications (ICAIHA), 2024.

Bisser, A., and Mahajan, R., *Leveraging AI and Machine Learning in Chatbot-based Healthcare Systems*, Proc. Int. Conf. Artificial Intelligence and Healthcare (ICAIH), 2021.

Microsoft, *Microsoft Bot Framework*, Available at: https://dev. botframework.com/, Accessed: 2023-06