scientific reports

OPEN

Responsible CVD screening with a blockchain assisted chatbot powered by explainable AI

Salman Muneer[1], Sagheer Abbas[2], Asghar Ali Shah[3]✉, Meshal Alharbi[4], Haya Aldossary[5], Areej Fatima[6], Taher M. Ghazal[7] & Khan Muhammad Adnan[8]✉

Cardiovascular disease (CVD) is rising as a significant concern for the healthcare sector around the world. Researchers have applied multiple traditional approaches to making healthcare systems find new solutions for the CVD concern. Artificial Intelligence (AI) and blockchain are emerging approaches that may be integrated into the healthcare sector to help responsible and secure decision-making in dealing with CVD concerns. Secure CVD information is needed while dealing with confidential patient healthcare data, especially with a decentralized blockchain technology (BCT) system that requires strong encryption. However, AI and blockchain-empowered approaches could make people trust the healthcare sector, mainly in diagnosing areas like cardiovascular care. This research proposed an explainable AI (XAI) approach entangled with BCT that enhances healthcare interpretability and responsibility to cardiovascular health medical experts. XAI is significant in addressing cardiovascular prediction issues and offers potential solutions for complex communication and decision-making in cardiovascular care. The proposed approach performs better, with the highest accuracy of 97.12% compared to earlier methods. This achievement shows its ability to tackle complex issues, accessible during healthcare sector communication and decision processes.

Keywords Blockchain, Chatbot, CVD screening, Explainable AI

Chatbots[1] are AI programs that understand human language. They answer questions and help people. Chatbots are very useful for customer service. They can quickly solve problems and make the user experience better. But chatbots also help in other areas like healthcare and education. They can do repetitive tasks automatically. It makes processes more efficient. Chatbots use machine learning (ML)[2] to keep improving. They adapt to how users communicate and learn new language patterns. Chatbots increase productivity and make information more accessible. People can use technology and get services more efficiently with chatbots.

Chatbots work using complex AI systems. They employ natural language processing (NLP)[3] to grasp users' words. NLP helps chatbots comprehend context, tone, and intent. This allows more natural conversations. Also, chatbots leverage ML algorithms[4,5]. ML enables chatbots to learn from interactions. They can then modify responses based on feedback and trends. The ability to learn dynamically improves chatbot skills. It ensures responses stay relevant and personalized. Chatbots combine cutting-edge tech to make information accessible. They transform how people engage digitally. Healthcare widely adopts chatbots nowadays.

Healthcare[6] keeps people healthy. It combines science, caring, and technology. Healthcare aims to help people and communities. It prevents sickness, identifies illness, gives treatment, and offers ongoing support. Today, Healthcare faces challenges. More people are born[7]. People live longer[8]. Medical knowledge overgrows[9]. So, Healthcare must change to improve. It must offer better care. Care must be easier to access. Care must be more efficient[10]. Healthcare isn't just about curing illness. It helps society progress. It helps people thrive. New technology promises better Healthcare. Telemedicine connects patients and doctors remotely. Smart medical equipment makes care more personalized[11], effective, and accessible. Healthcare is undergoing a shift, with patient-centered

[1]Department of Computer Science, University of Central Punjab, Lahore, Pakistan. [2]Department of Computer Science, Prince Mohammad Bin Fahd University, Alkhobar, Kingdom of Saudi Arabia. [3]Department of Computer Science, Kateb University, Kabul, Afghanistan. [4]Department of Computer Science, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, 11942 Alkharj, Saudi Arabia. [5]Computer Science Department, College of Science and Humanities, Imam Abdulrahman Bin Faisal University, 31961 Jubail, Saudi Arabia. [6]Department of Computer Science, Lahore Garrison University, Lahore, Pakistan. [7]Department of Networks and Cybersecurity, Hourani Center for Applied Scientific Research, Al-Ahliyya Amman University, Amman, Jordan. [8]Department of Software, Faculty of Artificial Intelligence and Software, Gachon University, Seongnam-si 13557, Republic of Korea. ✉email: asghar.ali.shah@kateb.edu.af; adnan@gachon.ac.kr

care and equal access to resources taking center stage. This field strives to push boundaries, merging advanced tech and empathy to foster well-being for everyone. Chatbots are renovating patient engagement, offering real-time assistance, appointment booking, and more services. The sector's evolving mindset prioritizes redefining what's achievable by blending cutting-edge innovations and compassion for optimal health outcomes.

Healthcare chatbots[12–14] have become important in the medical industry. They utilize AI and NLP to talk with patients in real time. They offer easy access to healthcare information, appointments, and answers. They enhance patient experience by providing quick help, reducing wait times, and promoting smooth communication. They also boost efficiency in admin work. It allows healthcare professionals[15] to focus more on patient care. The chatbots help create an efficient, patient-focused healthcare system. They improve patient involvement, encourage proactive health management, and leverage technological advancements. However, they still face some security issues and BCT can resolve these issues.

BCT's distributed, secure platform[16,17] is a game-changer in the healthcare industry. Patient data protection is crucial in digital health, and BC offers a clear ledger. Since BC is also distributed, medical chatbots built on this technology have a distributed network or network of networks. This minimizes the risk of a single point of failure and boosts security against cyber attacks. Transactions become more visible and traceable, further strengthening private health information's privacy. Cryptographic principles[18] nurture the trust of patients and healthcare chatbots in each other. In medicine, where the field is ever-changing and emerging, healthcare chatbots become the platform for ensuring the confidentiality of information, faithfulness and general security, thus contributing to a more reliable and effective medical environment.

Integrating XAI into healthcare chatbots marks a significant step toward secure and transparent healthcare systems. This study introduces a novel approach by combining XAI's interpretability with BC technology to enhance data security and trust in medical diagnostics. Unlike previous models, this framework not only provides explainable diagnostic results but also ensures tamper-proof storage of medical records[19–21]. This unique combination improves transparency, reliability, and patient trust, addressing key challenges in AI-driven healthcare applications.

The combination of BC and XAI in medical chatbots introduces a novel approach to healthcare delivery. This study uniquely integrates BC for secure data storage with XAI to enhance the transparency of AI-based diagnostic decisions. Unlike previous studies, this framework addresses both data security and interpretability, ensuring that medical information remains protected while being understandable to all stakeholders. This innovation fosters greater trust in AI-powered healthcare systems, promoting wider acceptance among users.

This paper explores the integration of blockchain-assisted chatbots with XAI to enable responsible CVD screening. While existing studies have separately addressed AI-based diagnostics and blockchain for secure data management, limited research combines both technologies to enhance transparency and trust in healthcare applications. This work bridges the gap by proposing a novel framework that ensures secure data storage while offering interpretable diagnostic outcomes. The study aims to guide future advancements in ethical and transparent AI-driven healthcare systems.

Related work

Numerous researchers have earlier applied BCT and XAI approaches to develop chatbots in the healthcare sector. Some of their study are mentioned in this portion.

The researchers[22] explained that AI chatbots provide multiple advantages to patients in Healthcare. One main usage is for the primary screening and suggestion of interventions. In this scenario, a disease-affected person can engage with a chatbot that inquires about their indications and hearing history, subsequently offering sanctions for self-management, further assessment, or treatment built on the patient's input. This proves especially valuable when individuals are uncertain about experiencing hearing loss, reluctant to seek medical attention, or facing profound hearing impairments hindering communication with a clinician. Chatbots can also serve as teaching tools, supporting self-management and identifying circumstances related to social demands. Patients are also provided with health information, safety tips, and advice on how best to manage medical problems, while chatbots also offer advice on how to use management tools and suggest strategies for addressing common problems. Conversely, the risk of bad advice from chatbots poses problems, such as providing incorrect guidance and prompting patients to treat themselves inappropriately, delaying cure, or leading to more harmful outcomes.

This research[23] demonstrated how a chatbot quickly assists patients, medical staff, and the hospitalization of individuals with critically ill renal disease. When powered with innovative AI algorithms, chatbots can support patients and physicians 24/7, as each common and frequently raised question/issue is worked out instantly, which validates that the before- and after-hours patients with the condition receive correct information and help, thus significantly enhancing the recipients' utility of Healthcare[24]. Chatbots may become vital tools for kidney health-related individuals due to quick advice and relieving anxiety.

In this research[25], the authors highlighted that typical medical chatbots rely on AI and NLP to understand a patient's words, predict what they mean, and produce an answer that fits. Then, the chatbot's programmers modify it every so often based on the patients' real questions, their satisfaction, and some language performance metrics. These systems are governed by a developer for the chatbot, who cares for a human-legible, human-moderated database. In contrast, ChatGPT's advanced AI technology deviates from fetching internet-based data, which may cause readers to wonder just how accurate and current medical data it access could be. Filling its database by this method may be easy, but ChatGPT will require slow teaching from all of the people its clients, a process it could only allow a medical team to do, as any client can affect the AI's learning, and hence its ensuring errors. Because ChatGPT's responses can be unexpected and can vary depending on the corpus it's trained on, it was paramount to rigorously test and evaluate its performance. A robust quality assurance framework will have to be established, with systematic tracking of database changes and preservation, to ensure that what ChatGPT delivers online is accurate. Creating a custom dataset and working with a team of healthcare experts to review and validate the training dataset significantly heightens the accuracy and importance of ChatGPT's healthcare data. As a healthcare chatbot, ChatGPT must be constantly updated and improved to maintain current relevance and consistently offer new and accurate data.

The authors of this study argued that medical chatbots advanced by dedicated healthcare professionals may not have the same degree of accuracy as ChatGPT. Over time, healthcare chatbots, enriched with AI functionalities like NLP, have evolved to concentrate on specific functions, such as addressing user queries across diverse medical subjects[26,27]. Developers of these chatbots can construct the foundational software, integrate it with a well-maintained database, and readily adjust both the conversational structure and data as needed.

In this research[28], the authors emphasize elevating awareness about cyber threats and stimulating organizations' cybersecurity by honing in on the frailest connection, the human factor layer. They also propose the deployment of an AI-driven conversational bot, functioning as a personalized support to augment awareness of cyber threats and disseminate the latest data and training to company employees. Designed explicitly for communication via WhatsApp, the bot can maintain individual records for each employee, assess their progress, and recommend training measures to mitigate vulnerabilities. Implementation of this bot has demonstrated significant positive effects on employees, enabling the system to update its database in the event of a security.

breach and suggest appropriate actions during an attack. However, the article underscores cybersecurity's dynamic nature, emphasizing the need to incorporate new features into the bot to stay abreast of emerging threats. Proposed enhancements include a feature for validating procedural applications, immediate notification to the IT team in case of a severe attack, the integration of voice generation for employee focus, and linking the bot to the latest security webpages and databases to promptly inform employees and the IT department about new threats.

The authors[29] presented that the Honest Chain system utilizes both BCT and chatbot functionalities to facilitate secure, expedited, and standards-compliant sharing of health information. Employing a consortium blockchain approach, Honest Chain ensures efficient data sharing by incorporating reputation value calculations for both Requesters and Providers. Furthermore, it uses risk valuation for each operation through computerization, ensuring auto-assurance and auto-audibility. The effectiveness of the Honest Chain process hinges on the computerization of distributed trust. The serviceability of the chatbot, depending upon requester supervision, may either enhance or impede the process of reducing Loss of Value and Loss of Chance challenges. Additionally, our method grants access to secure data sets, but their investigation requires the corporation of several systematic tools and visualizations, for example, Jupyter notebooks, by users.

Numerous previous studies have tackled the issue of trust deficits in the sharing of secured healthcare data. For instance, in[30], they propose a brokering architecture focused on building trust and fostering disease-affected person-centric cloud medical facilities. This approach actively pursues patient feedback and introduces auditability by monitoring communications via the BC solution. Brokering processes incorporating BCT have demonstrated the potential to enhance the quality of patient care and reduce healthcare costs through targeted and secure data sharing, as evidenced in[31]. They are addressing the shortcomings of centralized architectures in medical data exchange, which include high reliance on web connection and vulnerability to one point of disaster. In[32], ML is primarily categorized into supervised learning and unsupervised learning. In supervised learning, algorithms are trained on labeled data by comparing predicted outcomes with actual results to improve accuracy for future predictions. In another study[33], the researchers highlighted that Electronic Health Records (EHRs) have revolutionized CVD prediction by storing comprehensive patient information in digital form. These records contain vital data such as medical history, demographics, and laboratory results, playing a crucial role in forecasting disease progression. The rich dataset in EHRs enhances the accuracy of predicting patient outcomes and supports early diagnosis.

The authors explained that the utilization of BCT is undergoing a conceptual evolution in Healthcare, delivering significant value to information management functions through enhancements in efficiency, access control, technical innovation, privacy protection, and security. Research findings indicate that existing limitations primarily revolve around Approach performance, implementation constraints, and associated costs[34]. The term "blockchain" is derived from its methodology of maintaining transaction data in sequentially connected "blocks." These blocks, forming a continuous "chain," grow in length alongside the increasing volume of transactions. Each interaction is logged in a personal ledger, with entries stored as blocks on the chain. The fundamental components of a block include the data or information segment, the hash, and the preceding hash. Blockchain encompasses features such as a peer-to-peer network, cascaded encryption, a distributed database, transparency with pseudonymity, and irreversible records. Significant applications of BC in healthcare span drug development, clinical trials, medical data management, and security[35].

The authors presented that three distinctive features of BCT: immutability, cybersecurity, and interoperability can effectively support comprehensive data secrecy, loading, and management at the lowest cost and hazard[36]. A method was implemented to remotely detect and treat cancer tumours for selected disease-affected persons, utilizing a BC Approach for telemonitoring medical and dermatologic challenges[37]. It is emphasized that rules should be established to employ conventions containing BCs, which may validate information generated at medical services and by distinct inhabitants. BCT has found application in gerontology, chronic disease management, and Healthcare and pharmacological firms for study and medical practice[38].

The authors[39] highlighted that AI chatbots, also known as conversational agents, utilize dialog systems to engage in natural language conversations with users through speech, text, or a combination of both. In terms of conceptualization, the fundamental technical capability of AI chatbots differs from that of personified virtual conversational representatives, which focus on making multimodal motions to pretend face-to-face human talk. This study concentrates on the emerging main feature of natural language discussion in AI chatbots, aiming to help more adaptable data sharing among people and the chatbot. The conversational capacity may vary, ranging from constrained to unconstrained discussion (where users can respond naturally by inputting their conversational lines).

The authors presented that AI chatbots may be implemented in the shape of mobile applications on smartphones, ensuring their availability around the clock. The rapid evolution of AI chatbots has led to significant transformations in various sectors, encompassing business[40], governance[41], education[42], and healthcare[43]. Amazon Alexa boasted over 100,000 programs as a prominent platform for chatbot progress. Facebook Messenger had over 300,000 active chatbots as of 2019, with a substantial portion dedicated to Healthcare and well-being. An illustrative example is the WHO's launch of a chatbot on Facebook Messenger in April 2020, aimed at combating misinformation and providing immediate and correct data related to COVID-19[44].

In[45], the authors describe that Telehealth and telemedicine systems aim to provide remote healthcare services to alleviate the transmission of COVID-19. These systems are crucial in efficiently managing limited healthcare resources and addressing the overwhelming load of COVID-19 disease-affected people in hospitals. Nevertheless, many current telehealth and telemedicine processes exhibit a unified structure, lacking essential features such as information security, privacy, decision-making power, operational transparency, health records immutability, and traceability. These shortcomings pose challenges in detecting and preventing fraudulent activities related to disease-affected persons' insurance privileges and surgeon credentials.

In[46], the authors explain that the evolution of electronic information technology has led to the widespread adoption of electronic medical records (EMRs) as a conventional method for storing patient data in hospital settings. Patient records are dispersed across various hospital databases, even on the same individual. Consequently, constructing a merged and summarized EMR for a single patient from multiple hospital databases is challenging due to concerns related to security and privacy. They also highlight that existing EMR systems lack a standardized data management and sharing policy. This absence of a general policy poses difficulties for pharmaceutical scientists striving to develop precise medicines, as they must contend with data obtained under different policies. In response to these challenges, they have introduced MedBlock, a blockchain-assisted information management system designed to address patient information issues; this system has no decision-making power, transparency, or accountability. However, a blockchain-assisted AI chatbot is an intelligent information system that uses BC for secure data storage and AI for delivering transparent and reliable healthcare recommendations.

In[47], researchers propose implementing a secure BCT system to protect electronic healthcare records (EHR). This framework integrates sensors, the Internet of Things (IoT), databases, and other computing resources. By employing this framework to secure EHR, the authors anticipate an overall security and privacy enhancement compared to traditional healthcare systems. However, the study does not explicitly address data purity, transparency, and accountability concerns.

The authors[49] developed and assessed a novel evidence-based health information tool called PROSCA, a chatbot designed for the field of Prostate Cancer (PC). This tool shows great promise in raising awareness, assisting patients with knowledge, and providing support. Its primary goal is to give targeted help for doctor-patient communication. The study discovered that a medical chatbot with an early PC detection emphasis helps patients by providing them with an extra educational resource. Nevertheless, it is noteworthy that authors avoided discussing responsibility and transparency in their work.

The authors' research[50] emphasizes certain aspects of the question-answering system (QAS) that are now utilized in the healthcare industry. According to their research, people view conversational bots as practical and easy to use. These agents show promise regarding time and resource savings but also have issues with data integrity, secure communication, accountability, and transparency.

Despite significant advancements, previous studies have not sufficiently addressed the simultaneous integration of model interpretability, scalability, and data privacy in AI-based healthcare systems. This gap limits the ability of these models to provide transparent and secure solutions. The proposed approach leverages blockchain-assisted AI to overcome these limitations and deliver more reliable healthcare diagnostics.

Existing studies have explored the integration of BCT in healthcare, primarily focusing on secure data storage, patient data management, and decentralized access control. These implementations provided enhanced data security and privacy but often lacked the capability to offer transparent and interpretable AI-based decision-making. Moreover, many of these studies struggled with scalability issues and failed to ensure seamless integration with AI diagnostic systems. The proposed blockchain-assisted AI chatbot addresses these limitations by not only securing patient data through decentralized ledgers but also offering explainable predictions and scalable solutions, fostering greater trust and reliability in healthcare diagnostics.

Limitations of previous work

Several limitations have been observed in the previous research regarding healthcare chatbots. Previous healthcare chatbots face challenges such as insecure communication, lack of decision-making authority, and poor transparency. Sensitive patient data may be exposed to breaches without robust security measures. Limited decision-making capacity can lead to suboptimal healthcare support. Blockchain-assisted AI chatbots address these issues by providing secure, transparent, and decentralized data storage, enhancing both trust and decision-making reliability. Some previous publications' work with the proposed system is shown in Table 1.

Table 1 highlights various limitations from the literature review, including a lack of performance regarding secure communication systems, decision-making power, and transparency and accountability. This proposed research work addresses these critical issues by incorporating innovative technologies. The integration of BCT ensures secure communication and protects patient data, while machine learning algorithms enhance decision-making with accurate, personalized responses derived from comprehensive healthcare data analysis. XAI raises transparency and accountability through clear, interpretable explanations for chatbot recommendations,

Ref.	Authors	Preprocessing layer	Secure communication	Decision making	Transparency and accountability
[45]	Ahmad RW et al.	✗	✓	✗	✗
[46]	Fan K et al.	✓	✓	✗	✗
[47]	Quasim MT et al.	✗	✓	✓	✗
[48]	Zarour M et al.	✗	✓	✓	✗
[49]	Görtz M et al.	✓	✓	✓	✗
[50]	Budler LC et al.	✗	✗	✓	✗
[51]	Proposed responsible healthcare chatbot approach	✓	✓	✓	✓

Table 1. Comparison of previously published works with the proposed approach.

Attribute	Description	Data type
Id	Unique identifier for each record	Integer
Age	Age of the patient (in years)	Integer
Gender	Gender of the patient (1 = male, 0 = female)	Integer (binary)
Height	Height of the patient (in cm)	Float
Weight	weight of the patient (in kg)	Float
ap_hi	Systolic blood pressure (in mm Hg)	Integer
ap_lo	Diastolic blood pressure (in mm Hg)	Integer
Cholesterol	Cholesterol level (1: Normal, 2: Above Normal, 3: High)	Integer
Gluc	Glucose level (1: Normal, 2: Above Normal, 3: High)	Integer
Smoke	Whether the patient smokes (1 = yes, 0 = No)	Integer (binary)
Alco	Whether the patient consumes alcohol (1 = yes, 0 = No)	Integer (binary)
Active	Whether the patient is physically active (1 = yes, 0 = No)	Integer (binary)
Cardio	Target attribute indicating CVD (1 = yes, 0 = No)	Integer (binary)

Table 2. Dataset attributes with its datatype.

fostering trust and understanding in user interactions and overcoming prior limitations to enhance secure, intelligent, and accountable healthcare chatbot systems.

Dataset structure

The dataset used for CVD prediction consists of several attributes, each representing a different health-related feature. These attributes are numerical and categorical, and they provide essential information about the patient's physical condition. The structure of the dataset is as follows: it includes attributes such as "id," "age," "gender," "height," "weight," "ap_hi" (systolic blood pressure), "ap_lo" (diastolic blood pressure), "cholesterol" (cholesterol level), "gluc" (glucose level), "smoke" (whether the patient smokes), "alco" (whether the patient consumes alcohol), "active" (whether the patient is physically active), and "cardio" (target attribute, indicating whether the patient has cardiovascular disease or not). The "cardio" attribute is binary (0 or 1), where 1 indicates the presence of cardiovascular disease, and 0 indicates the absence. The dataset structure is shown below in Table 2.

Table 2 represents that the dataset for CVD prediction is a comprehensive collection of clinical and lifestyle attributes that provide critical insights into a patient's cardiovascular health. Key features such as age, blood pressure, cholesterol levels, and lifestyle habits like smoking and alcohol consumption offer a holistic view of potential risk factors. The target variable, "cardio," serves as an indicator of whether CVD is present, enabling predictive modeling and analysis. This structured data allows researchers and clinicians to identify patterns and make informed decisions to improve patient outcomes.

System block diagram

This research proposes a system that leverages historical data for predictive analysis, as depicted in Fig. 1 It incorporates Exploratory Data Analysis (EDA) to determine the necessity of preprocessing and to detect any outliers in the data. The preprocessing phase addresses issues such as missing values, duplicate records, outliers, and class imbalance to ensure data quality. Subsequently, the dataset is partitioned into training and testing subsets, with 70% allocated for training and 30% for testing. The model is trained and evaluated on these subsets, employing metrics such as accuracy, precision, recall, F1-score, and a confusion matrix to identify the optimal model. Finally, the selected model is utilized for accurate outcome prediction, with interpretability achieved through applying SHAP & LIME.

Figure 1 depicts the workflow of a healthcare prediction model integrating ML and XAI techniques. The process starts with data collection, followed by EDA and preprocessing to prepare the dataset. An XGBoost model is trained on the processed data, and its predictions are explained using SHAP and LIME to enhance interpretability. Positive predictions are flagged for further medical review, ensuring transparency and efficiency in healthcare delivery.

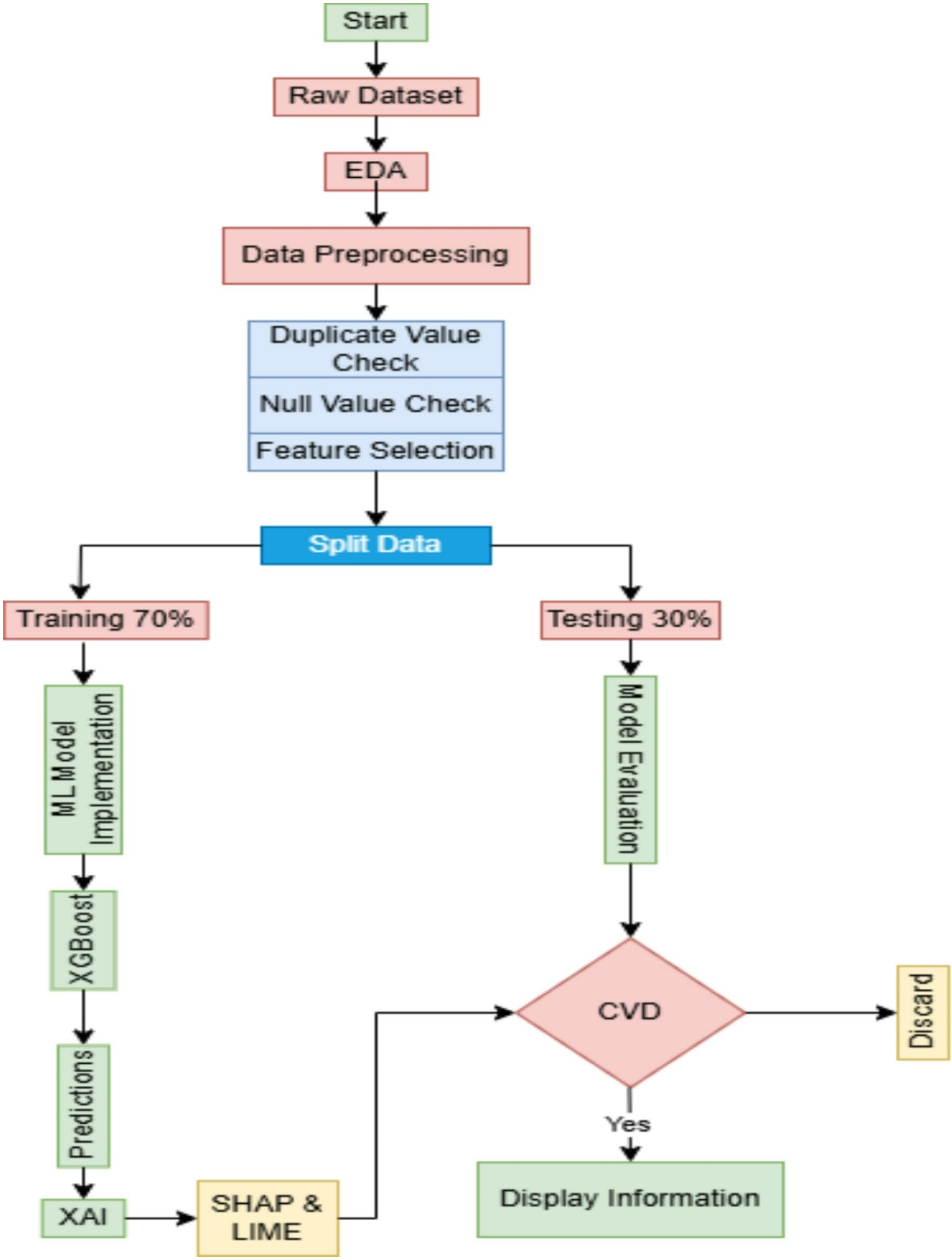

Fig. 1. System block diagram.

Pseudo code

The step-by-step pseudo code for the prediction of HD is shown below in Table 3.

Table 3 presents a structured pseudo-code for processing healthcare data and predicting CVD. It begins with loading the dataset, performing EDA, handling missing values, and using SMOTE to balance classes. The data is split for training an XGBoost model, and XAI techniques are applied to enhance prediction transparency, aiding informed healthcare decisions.

START
// Step 1: Data Loading
Load dataset from the source
// Step 2: Data Preprocessing
Perform EDA
Remove duplicates and check for null values
Handle missing values if any
Feature selection based on relevance
Apply SMOTE for handling class imbalance (if needed)
Normalize/Standardize numerical features
// Step 3: Splitting Data
Split data into Training set (70%) and Testing set (30%)
// Step 4: Model Implementation
Select an ML model (e.g., XGBoost)
Train the model on the Training set
Evaluate the model using the Testing set
// Step 5: Model Evaluation
Calculate performance metrics (e.g., Accuracy, Precision, Recall)
Check if the model detects 'CVD' or 'No CVD
IF model performance is satisfactory THEN
Proceed to XAI implementation
ELSE
Modify model or preprocessing steps and retrain
// Step 6: XAI Implementation (e.g., SHAP & LIME)
Generate explanations for model predictions using XAI techniques
Display or store the explanations
// Step 7: Make Predictions
Input new data into the trained model
Predict whether the patient has 'CVD' or 'No CVD'
Display predictions and explanations
// Step 8: Decision and Action
IF 'Health Disease' is predicted THEN
Recommend further medical consultation
ELSE
No further action required
END

Table 3. Pseudo code.

Proposed framework

With technological advancement, various fields have adopted autonomous systems, with chatbots emerging as a prominent application. Chatbots have gained much popularity, especially in the healthcare sector, as vital guides to patients and health-related inquiries. Nonetheless, these chatbots face several problems, including patient medical records issues, insecure communication, and failure to give clear and accurate answers.

Knowing the need to address these challenges has led to an increasing demand for developing intelligent approaches that can yield better results. Therefore, this research aims to design an intelligent approach for chatbot development using BCT and XAI. It sets out to mitigate prevailing challenges around chatbots in general with special emphasis on the healthcare industry to ensure that secure communications are implemented while providing clear responses and retrieving accurate information from patient records. The proposed responsible healthcare chatbot using machine learning is shown in Fig. 2. Figure 1 represents the proposed approach, which comprises the training and validation phases. During the training phase, initially, patient data is acquired from the patient through a chatbot and undergoes a BC for secure and tamper-proof transactions, thereby enhancing data integrity and transparency within healthcare systems. The tamper-proof transactions are then forwarded to the preprocessing layer, involving normalization, handling missing values, and moving averages, and the processed data is then divided into training and testing sets, with respective ratios of 70% and 30%. Subsequently, the approach is trained on 70% of the data for predictive analysis using a ML algorithm (XGBoost). The predictions are directed to XAI for comprehensive output explanations. If the output aligns with the predefined learning criteria, the results are stored in the cloud; otherwise, they are returned to the approach if the learning rate is not achieved.

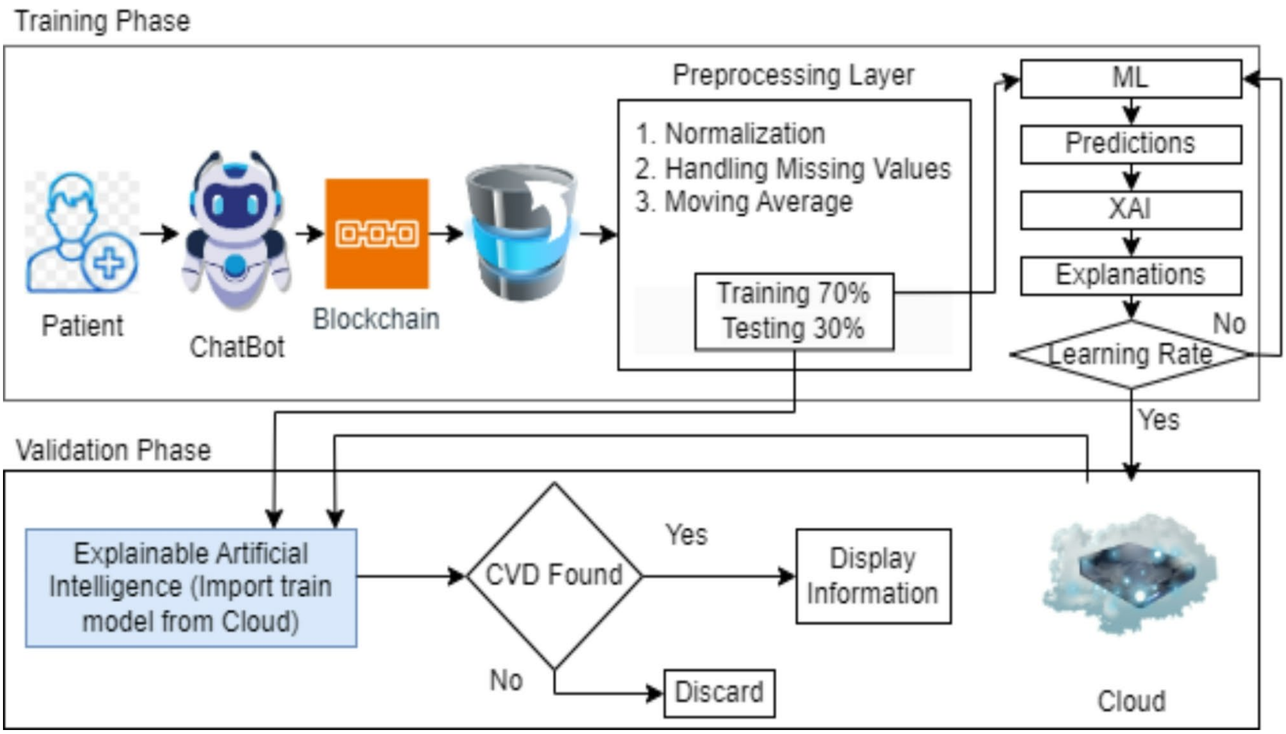

Fig. 2. Proposed responsible healthcare chatbot approach.

In the validation phase, patient data is directly compared with the imported data stored in the cloud. If the criteria are met, indicating the presence of CVD, the system displays as "yes". On the other hand, if the criteria are not met, signifying the absence of CVD, the system is discarded.

Simulation results

This research proposed a responsible healthcare chatbot using ML approach and implemented it on the dataset[51] containing 5391 samples. The data were distributed into 70% training (3774 samples) and 30% validation (1617 samples). As the equations state, this approach finds the result using multiple statistical measures as shown below.

$$\text{Accuracy (Acc)} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

$$\text{Sensitivity (TPR)} = \text{TP}/(\text{TP} + \text{FN})$$

$$\text{Specificity (TNR)} = \text{TN}/(\text{TN} + \text{FP})$$

$$\text{Miss Rate (FNR)} = 1 - \text{Acc}$$

$$\text{Fall out (FPR)} = \text{FP}/(\text{FP} + \text{TN})$$

$$\text{LR} + \text{ive (LR+)} = \text{TPR}/\text{FPR}$$

$$\text{LR} - \text{ive (LR}-) = \text{FNR}/\text{TNR}$$

$$\text{Precision or Positive Predictive Value (PPV)} = \text{TP}/(\text{TP} + \text{FP})$$

$$\text{Negative Predictive Value (NPV)} = \text{TN}/(\text{TN} + \text{FN})$$

It is shown in Table 4 that the proposed responsible healthcare chatbot approach predicts the CVD during the training period using XGBoost. During training, 3774 samples are divided into 1535, 2239 positive, and negative samples. 1505 true positives are successfully forecasted, and no CVD is recognized, but 30 records are mistakenly predicted as negatives, indicating the CVD is recognized. Likewise, 2239 samples are obtained, with negative showing CVD is identified and positive indicating no CVD. With 2200 samples correctly identified as negative, showing the CVD is recognized, and 39 samples inaccurately foreseen as positive, representing no CVD is identified despite the presence of the CVD.

Figure 3 shows that the charts compare the model's performance in identifying CVD in the training and validation datasets. The left chart shows the counts of true positives (CVD correctly identified) and false negatives (CVD missed) during training, while the right chart displays the same for validation. High true positive counts

	Total number of samples (3774)	Result (output)	
Input	Expected output	Predicted positive	Predicted negative
		True positive (TP)	False positive (FP)
	1535 positive	1505	30
		False negative (FN)	True negative (TN)
	2239 negative	39	2200

Table 4. Proposed responsible healthcare chatbot training using XGBoost.

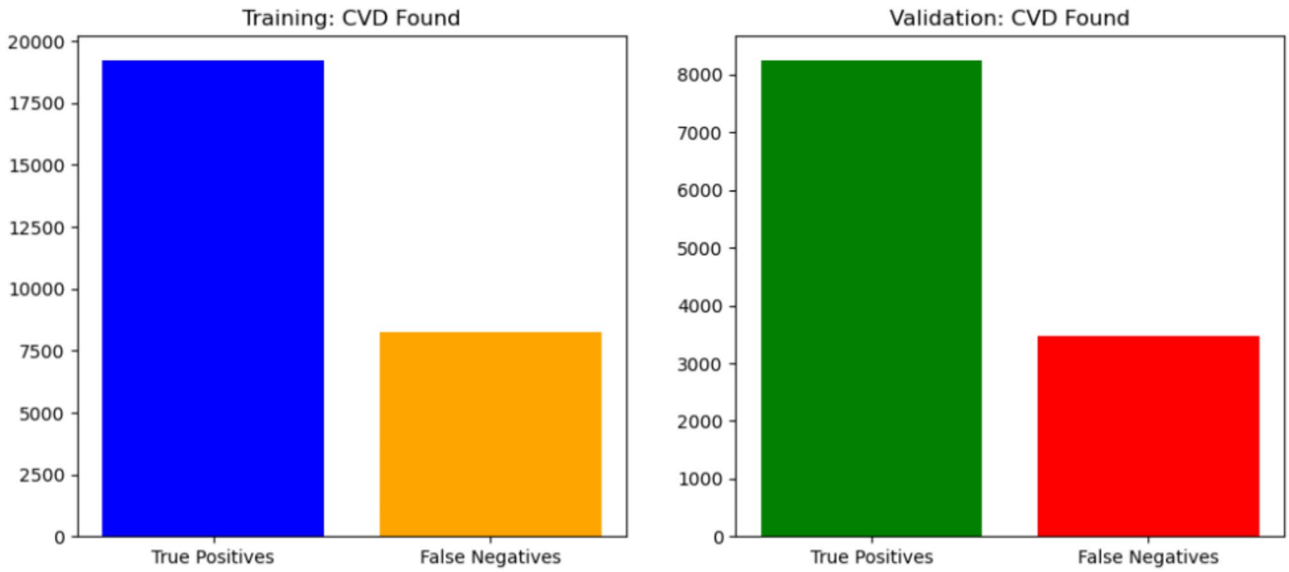

Fig. 3. 'CVD found' in the training and validation phase.

	Samples (1617)	Result	
Input	Expected output	Predicted positive	Predicted negative
		True positive (TP)	False positive (FP)
	922 Positive	900	22
		False negative (FN)	True negative (TN)
	695 Negative	20	675

Table 5. Proposed responsible healthcare chatbot validation using XGBoost.

indicate good detection, but false negatives suggest some cases of CVD were missed, highlighting areas for improvement.

It is shown in Table 5 that the proposed responsible healthcare chatbot approach predicts the CVD during the training period using XGBoost. During validation, 1617 samples are divided into 922, 695 positive, and negative samples. 900 true positives are successfully forecasted, and no CVD is recognized, but 22 records are mistakenly predicted as negatives, indicating the CVD is recognized. Likewise, 695 samples are obtained, with negative showing CVD is identified and positive indicating no CVD. With 675 samples correctly identified as negative, showing the CVD is recognized, and 20 samples inaccurately foreseen as positive, representing no CVD is identified despite the presence of the CVD.

Figure 4 shows that the charts illustrate the model's ability to identify cases where CVD is not found. The left chart (training data) shows true negatives (correctly identified as no CVD) and false positives (incorrectly predicted as CVD). Similarly, the right chart represents these metrics for validation data. The high true negative counts indicate strong performance, but false positives highlight instances where CVD was incorrectly flagged, requiring attention for further refinement.

It is shown in Fig. 5 that the correlation matrix illustrates the relationships between various features in the dataset. Attributes like age, cholesterol, and weight show moderate positive correlations with the target variable cardio (presence of CVD). Features such as smoke, alco, and active exhibit minimal correlation with cardio,

Fig. 4. 'CVD Not found' in the training and validation phase.

indicating limited direct influence. This matrix helps identify key predictors and their interactions, aiding in feature selection for model building.

Figure 6 is showing that the correlation between ap_hi (systolic blood pressure) and cholesterol is minimal, as indicated by a near-zero value in the heatmap. This suggests that changes in blood pressure levels have little to no direct association with cholesterol levels in this dataset. Both features may independently contribute to cardiovascular risk but do not show a strong relationship with each other.

Figure 7 shows that the ROC curve is a graphical representation of the model's ability to predict CVD. It plots the true positive rate (sensitivity) against the false positive rate (1-specificity) at various threshold levels. A higher area under the curve (AUC) indicates better model performance, with an AUC close to 1 signifying excellent discrimination between patients with and without CVD. It provides a clear visual of the trade-off between correctly identifying CVD cases and minimizing false alarms.

Figure 8 presents the distribution of predicted probabilities for CVD prediction, representing the spread of likelihoods that individuals in the dataset will develop CVD based on model predictions. Typically, these probabilities range from 0 (no risk) to 1 (high risk). Visualizing this distribution helps in understanding the overall risk profile of the population, identifying areas with higher or lower probabilities, and assessing the model's ability to distinguish between high-risk and low-risk individuals effectively. A balanced distribution often indicates good model calibration.

Figure 9 shows that SHAP interaction values in CVD prediction explain how pairs of features work together to influence the model's predictions. They highlight not just the individual contribution of a feature but also how its effect changes in the presence of another feature. For example, the interaction between age and cholesterol levels might reveal nuanced insights into cardiovascular risk. These values provide deeper interpretability by uncovering complex relationships within the data.

Figure 10 describes how the SHAP values measure the impact of each feature on the model's prediction for an individual's CVD risk. They quantify how much a specific feature, such as age or blood pressure, pushes the prediction towards higher or lower cardiovascular risk. This ensures transparency by breaking down the contribution of each feature to the model's output.

Figure 11 highlights that a decision model in CVD prediction explains how the model combines different features to arrive at a prediction. It shows the sequence and importance of features (like age, blood pressure, and cholesterol) influencing the final output, helping to visualize the reasoning process behind high or low cardiovascular risk predictions.

Figure 12 illustrates that the feature importance in CVD prediction highlights which factors, such as age, cholesterol levels, or blood pressure, have the greatest influence on the model's predictions. By ranking these features, it provides insights into the key drivers of cardiovascular risk, aiding in model interpretation and aligning predictions with medical knowledge.

Figure 13 explains how SHAP values for the "age" feature in CVD prediction quantify how an individual's age impacts the model's output. Age often shows high feature importance, as older individuals typically have a higher cardiovascular risk. This insight underscores age's critical role in shaping predictions and aligns with its established relevance in medical risk assessment.

Figure 14 presents that the SHAP (SHapley Additive exPlanations) summary plot for the training phase in CVD prediction highlights the importance and impact of each feature on the model's predictions. It visually represents how individual features (e.g., age, blood pressure) contribute to increasing or decreasing the likelihood of CVD across the dataset. The plot helps identify the most influential factors, aiding in model interpretability and ensuring key predictors align with domain knowledge about cardiovascular health.

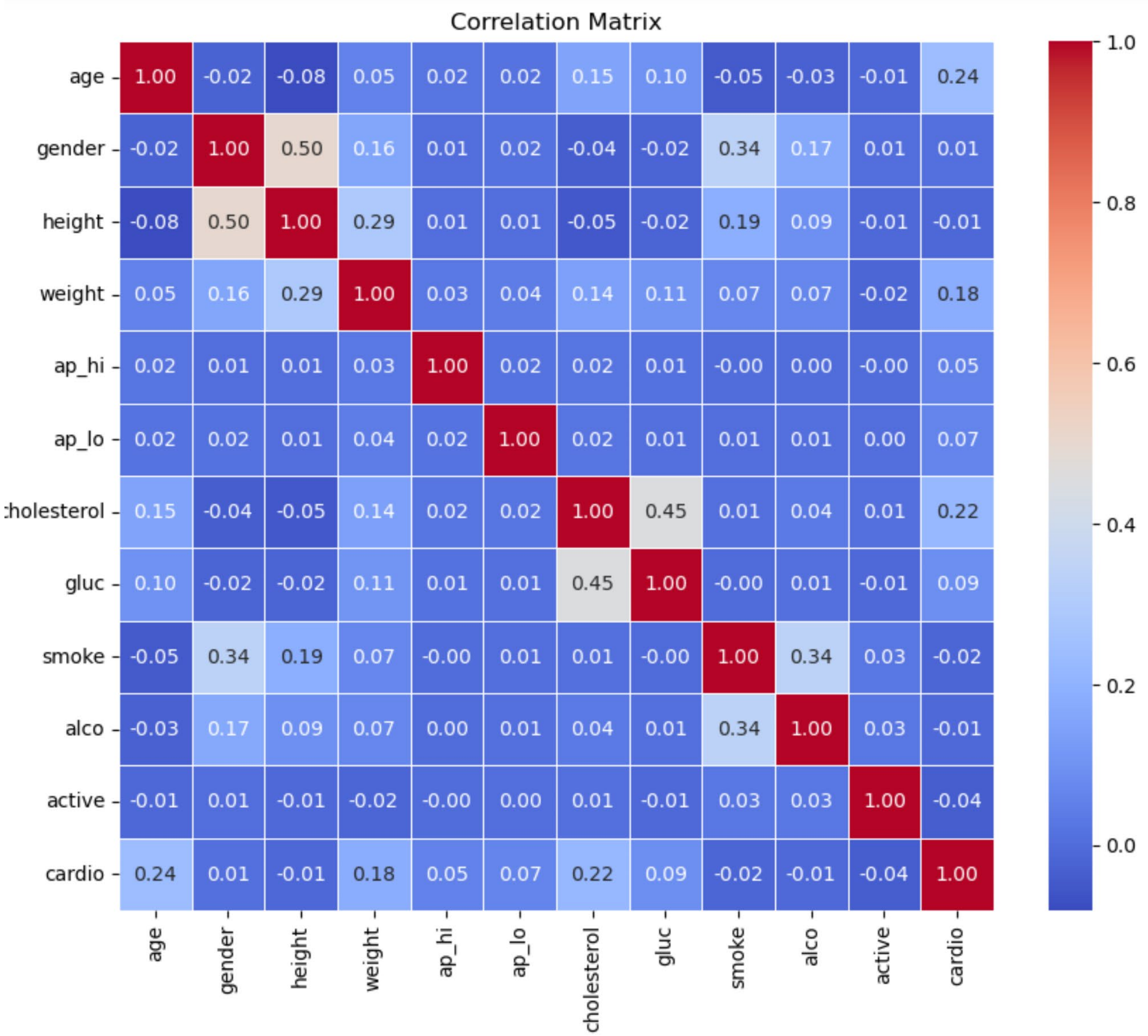

Fig. 5. Correlation matrix.

It is shown in Fig. 15 that the SHAP summary plot for the validation phase in CVD prediction shows how well the model generalizes to unseen data by illustrating the influence of features on predictions for the validation set. It highlights whether the most important features identified during training remain consistent in the validation phase. This helps verify the stability and reliability of the model's feature importance, ensuring its applicability to new data and alignment with expected CVD risk prediction patterns.

Table 6 shows that the proposed responsible healthcare chatbot using ML approach performance in terms of accuracy sensitivity, specificity, miss rate, and precision during training using XGBoost provides 98.17, 97.47, 98.65, 1.83, and 98.05, respectively. The suggested approach yields 97.40, 97.83, 96.84, 2.60, and 97.61 during the validation phase's accuracy, sensitivity, specificity, miss rate, and precision. Furthermore, the proposed responsible healthcare chatbot using ML approach yields 1.35, 72.2, 0.019, and 98.26 in terms of fall-out likelihood positive ratio, likelihood negative ratio, and negative predictive value during training and 3.16, 30.96, 0.027, 97.12 in terms of validation.

According to Fig. 16, the Proposed responsible healthcare chatbot using ML approach with the XAI shows a high level of confidence, about 97.12%, in predicting the forecasting of normal cardiovascular conditions. This confidence is influenced by several factors, including the presence of id, age, education, sex, is_smoking, cigsPerDay, BPMeds, prevalent stroke, prevalent Hyp, diabetes, totChol, sysBP, diaBP, BMI, heart Rate, glucose, exng, caa, Triglyceride, hdl_cholestrol, ldl_cholestrol, CPK_MB_Percentage and TenYearCHD. These factors contribute to a higher likelihood of classifying the cardiovascular as normal.

According to Fig. 17, the proposed responsible healthcare chatbot using ML approach with the XAI shows high confidence, about 97.12%, in predicting abnormal cardiovascular conditions. This confidence is influenced

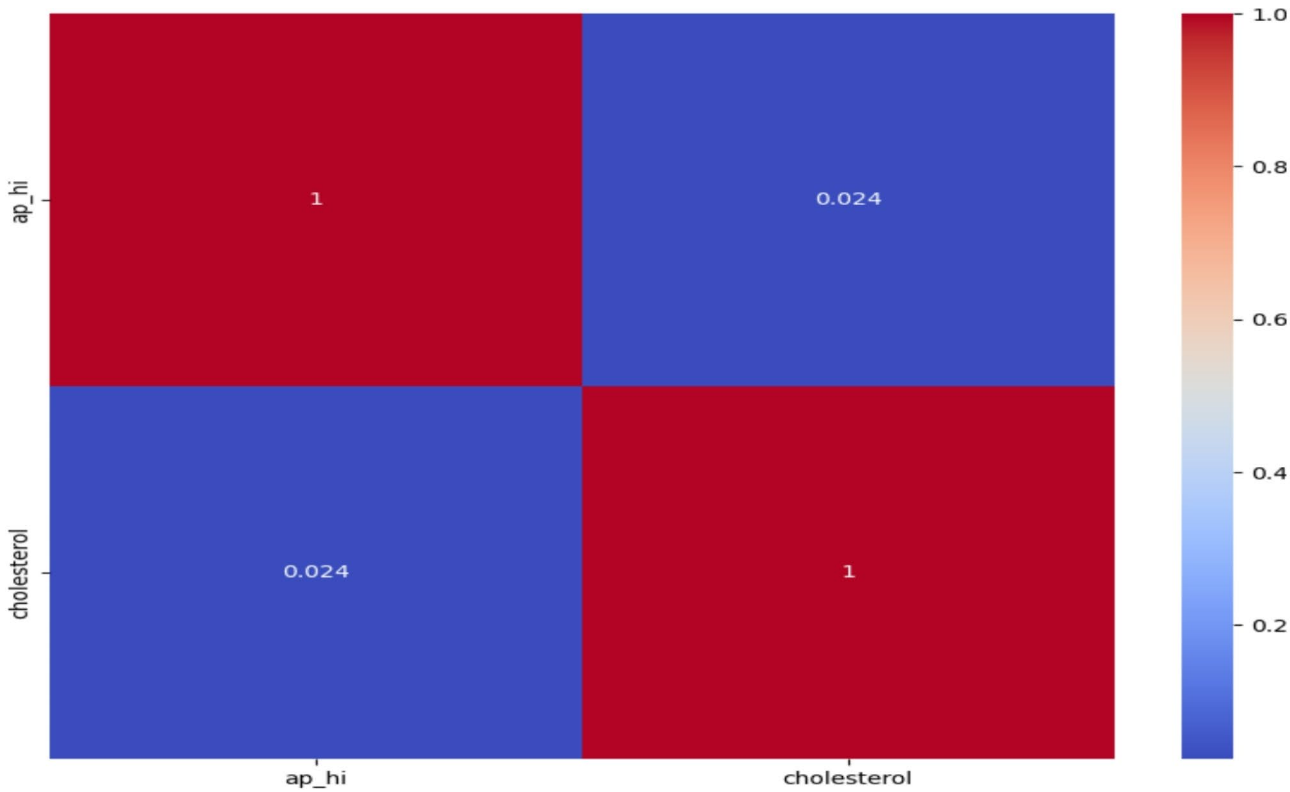

Fig. 6. Correlation heatmap of selected features.

by several factors, including the presence of id, age, education, sex, is_smoking, cigsPerDay, BPMeds, prevalent Stroke, prevalent Hyp, diabetes, totChol, sysBP, diaBP, BMI, heart rate, glucose, exng, caa, Triglyceride, hdl_cholestrol, ldl_cholestrol, CPK_MB_Percentage and TenYearCHD. These factors contribute to a higher likelihood of classifying the cardiovascular as normal.

Table 7; Fig. 18 compare the performance of the proposed responsible healthcare chatbot approach with previous ML approaches to predict CVD. It is clearly shown that this approach is better than the previous results in terms of accuracy and miss rate.

Discussion

The proposed blockchain-assisted chatbot powered by XAI demonstrates significant strengths in ensuring secure data storage, transparency, and reliable CVD screening[57]. The integration of BC enhances data privacy and trust, while XAI improves the interpretability of diagnostic decisions. However, the approach may face limitations in terms of computational complexity and response time, particularly in real-time medical consultations. Future improvements could focus on optimizing the system's efficiency, scalability, and seamless integration with existing healthcare infrastructures to enhance overall performance and user experience.

Conclusion

Healthcare systems increasingly incorporate AI into their systems, but it is not a solution to all difficulties. Healthcare data, complete with sensitive patient information, demands strong protections against breaches and unauthorized access. Artificial intelligence (AI) has been readily adopted to solve all healthcare industry issues better. This is one of the vital areas to bother about data security, including health data, which is very sensitive since it is patients' private information. Patient data privacy is now exposed by data breaches and unauthorized access, which are now the biggest threat to healthcare, putting both security and privacy at the forefront. The inherent decentralization of the BC[58] entails the use of powerful encryption methods and access controls, which lead to data integrity and privacy preservation, but become a key issue. The vast potential of AI is constrained by a trust issue that rises because of AI work like a black box, that is, can't be explained. Integrating XAI demonstrates the deployability of the Chatbot to the users where they will be comprehending the rationale behind the answers provided and the advice given. It may cause the responsible and ethical behaviors of the patients so that they can make the informed decision and this in turn may help in the utilization of health services. It is obvious, XAI is likely to become an integral part of patient-territorialization and de-monopolizing of healthcare information, making clinical decision support systems more sufficient for better patient outcomes and the overall provision efficiency improvement in general.

The development of a responsible healthcare chatbot framework integrated with XAI components represents a transformative advancement in healthcare technology. The proposed approach demonstrates notable

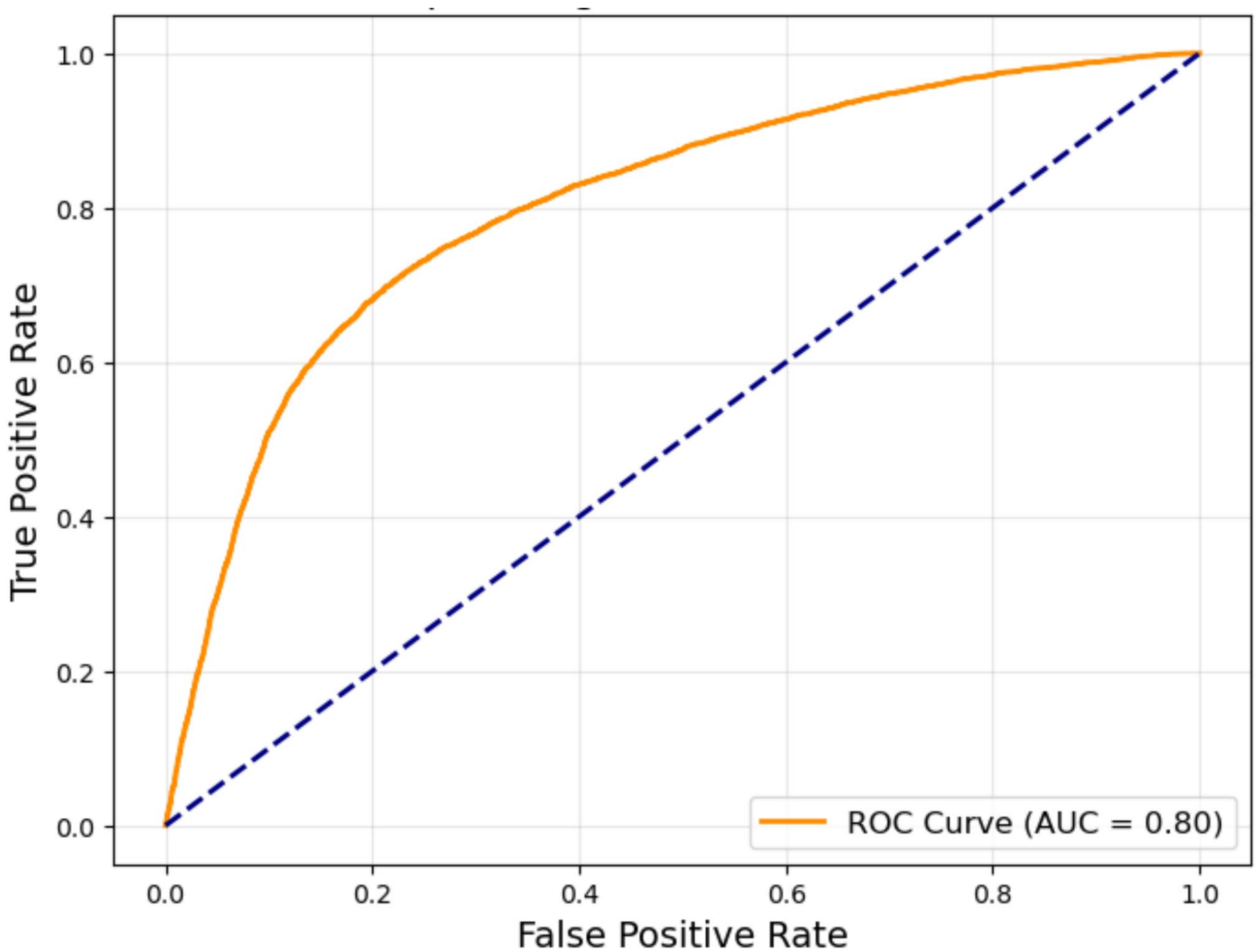

Fig. 7. ROC curve.

performance metrics, achieving 97.40% accuracy, 97.83% sensitivity, 96.84% specificity, and a 2.60% miss rate. The method outperforms previous strategies by attaining the highest precision of 97.12%, underscoring its efficacy in addressing complex health communication challenges and facilitating informed decision-making. Future research directions may involve scaling the proposed solution or integrating it with emerging technologies to optimize care delivery efficiency and enhance patient outcomes.

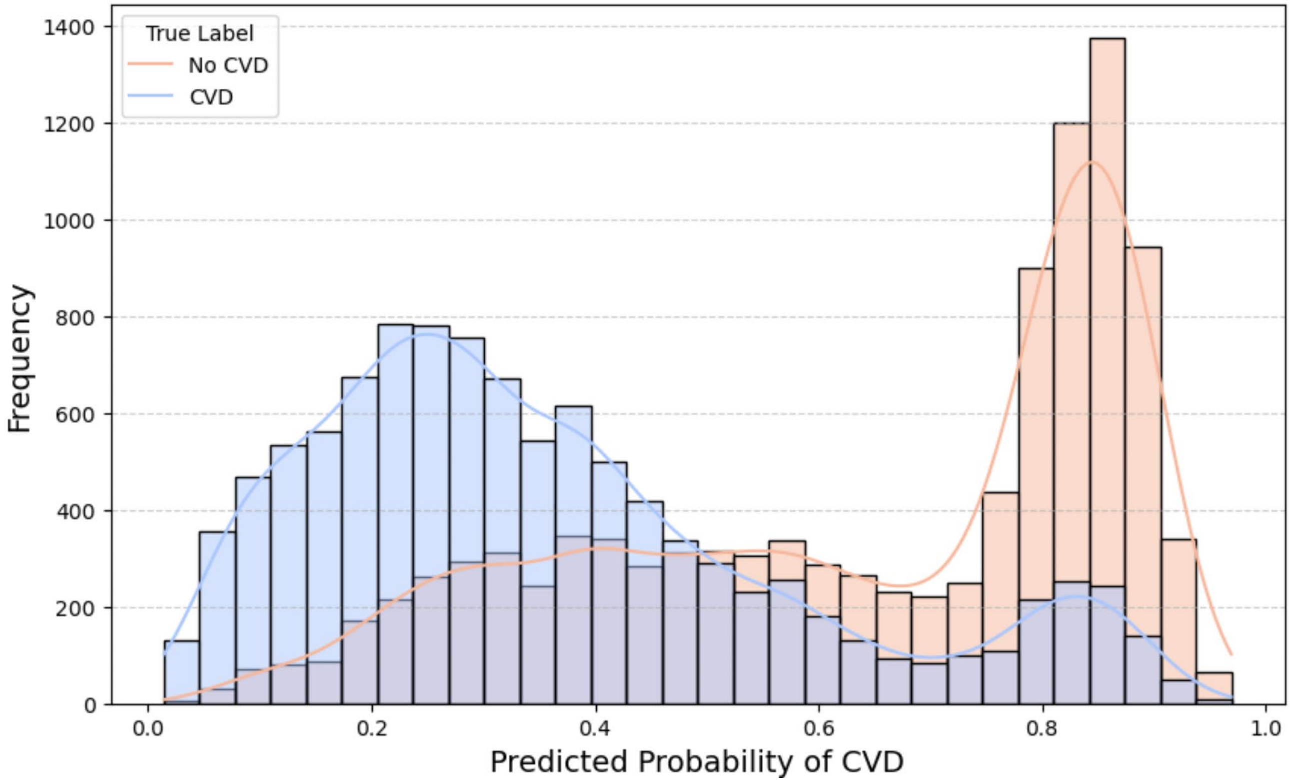

Fig. 8. Distribution of predicted probabilities.

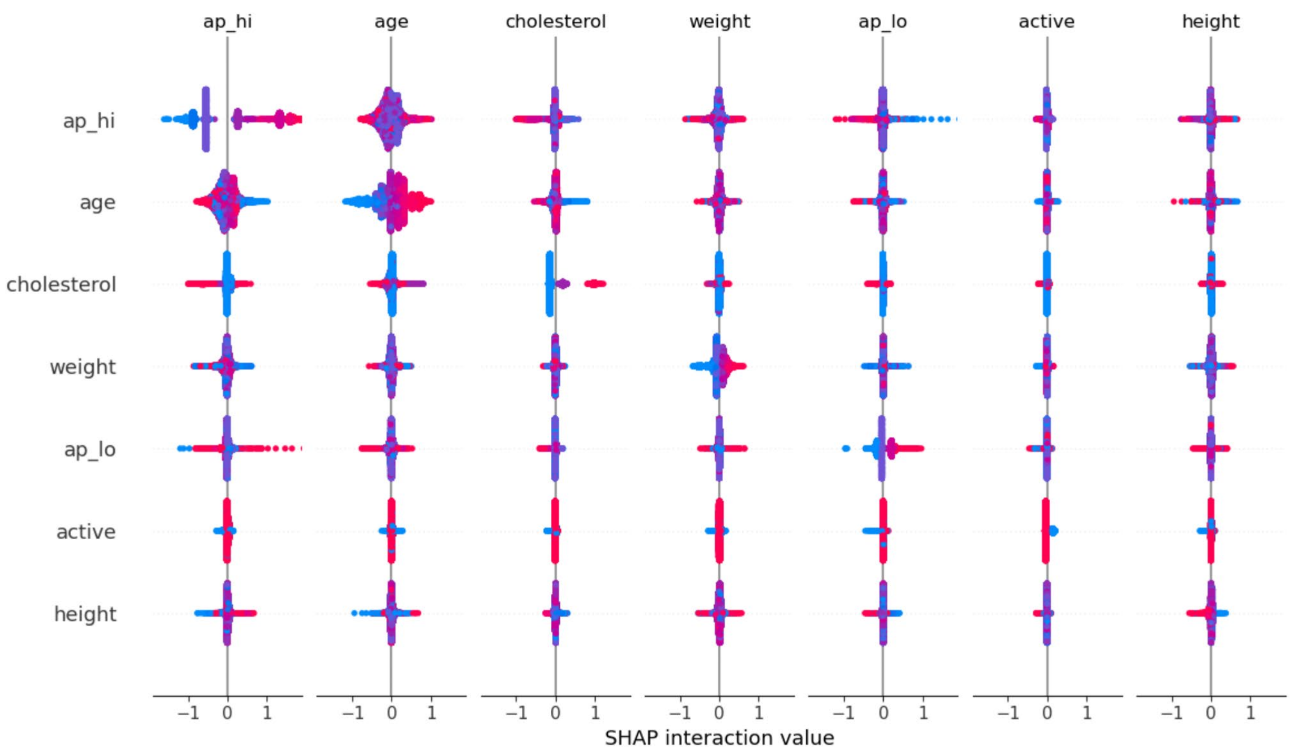

Fig. 9. SHAP interaction value.

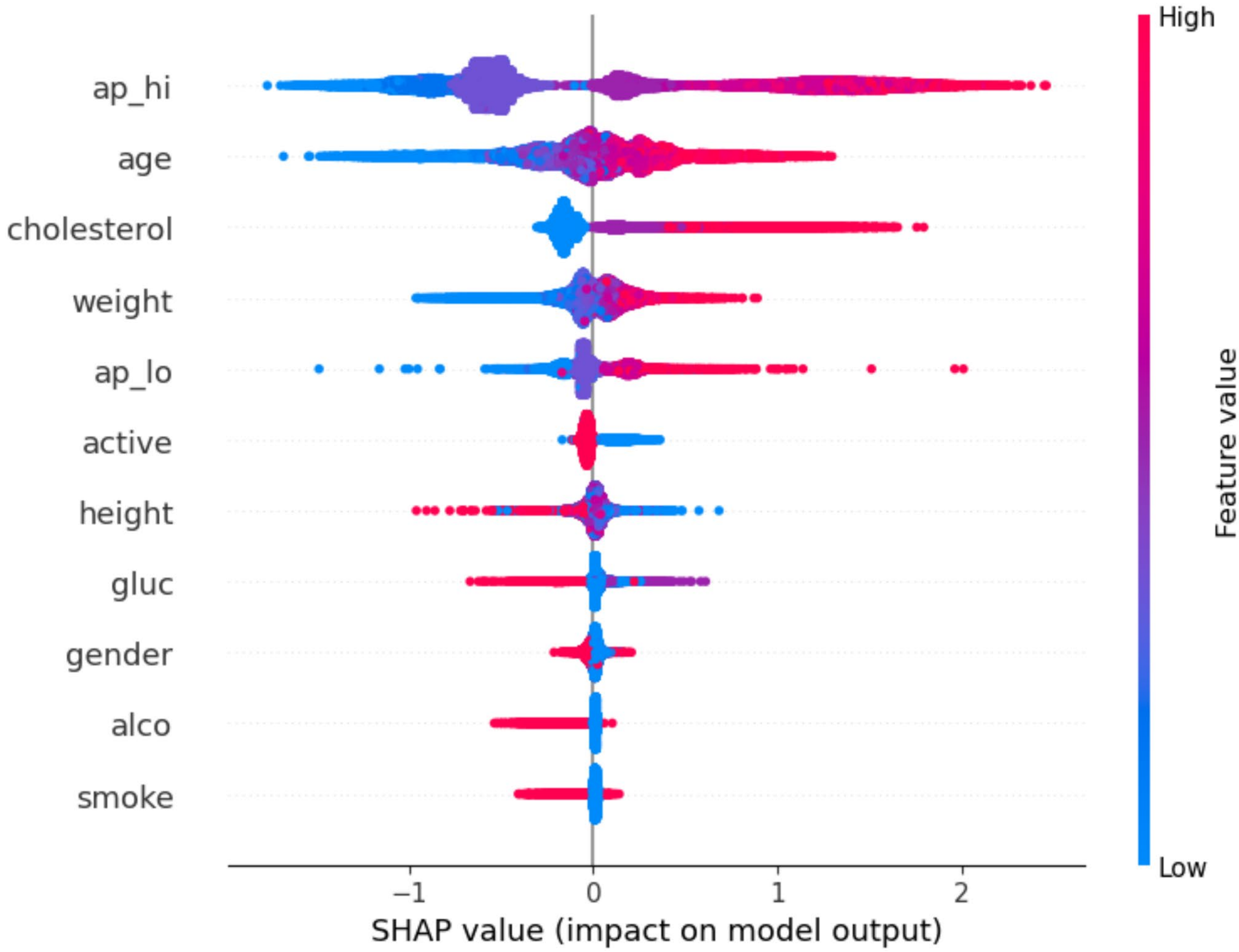

Fig. 10. SHAP value (Impact on model output).

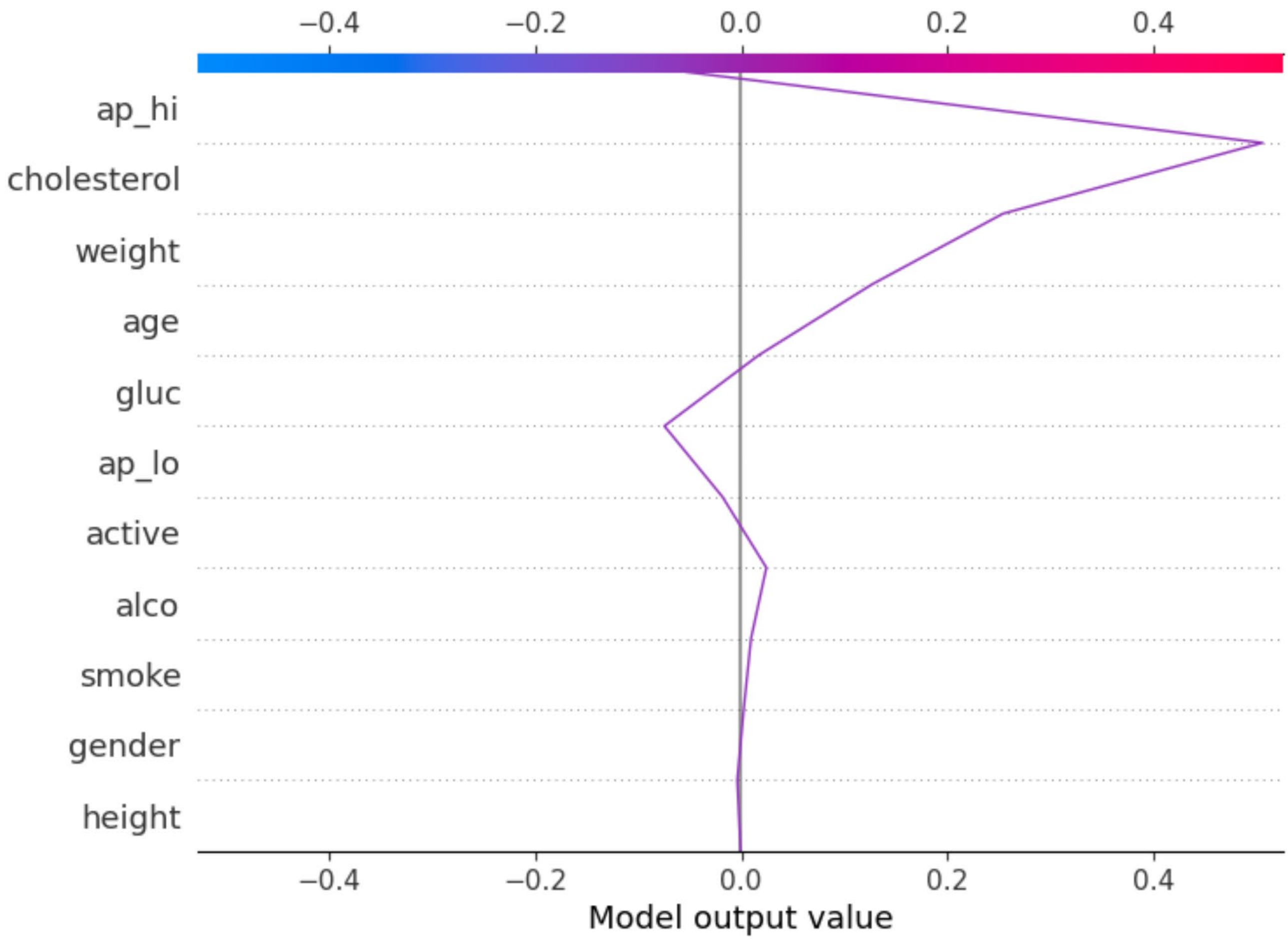

Fig. 11. Decision model.

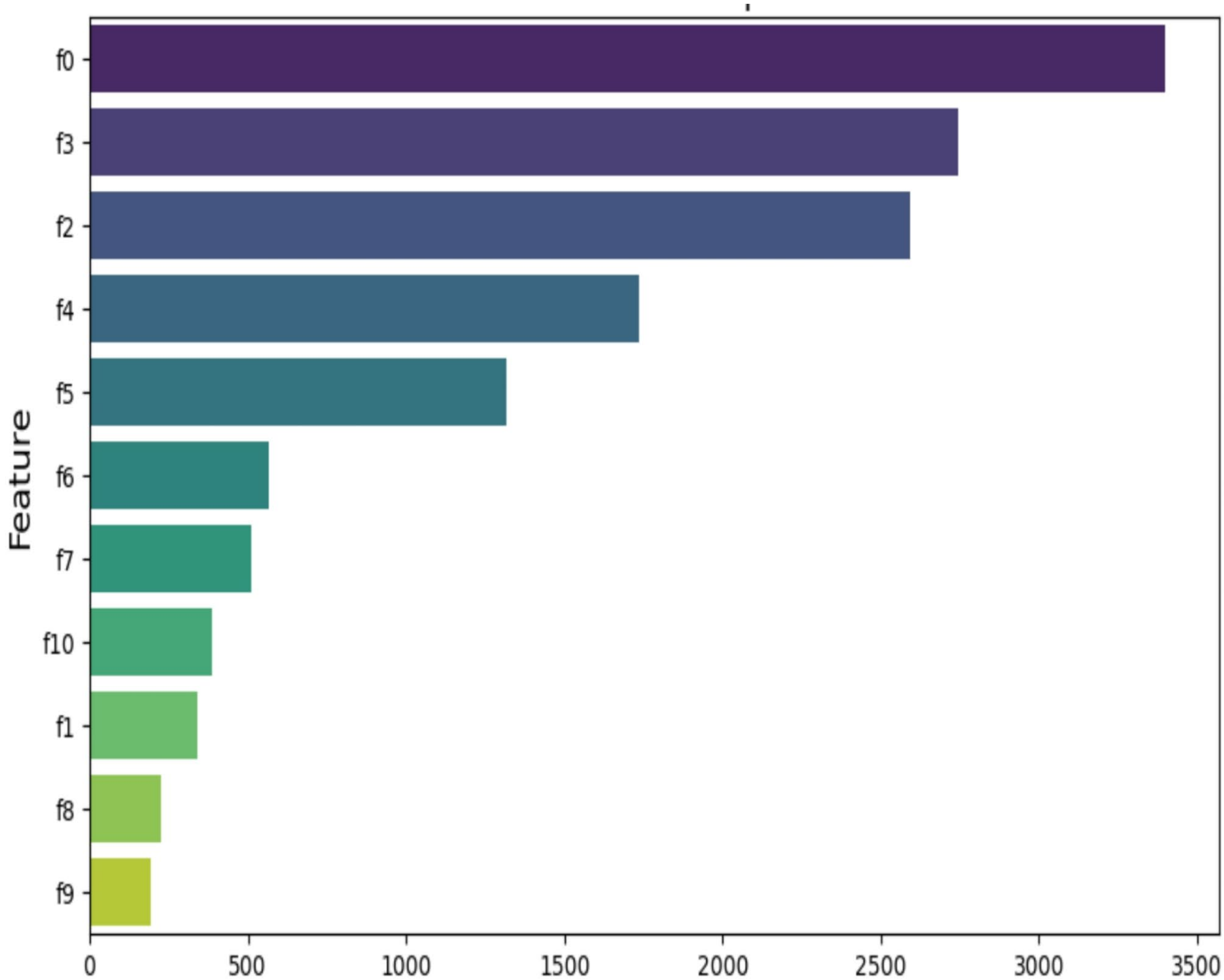

Fig. 12. Feature importance.

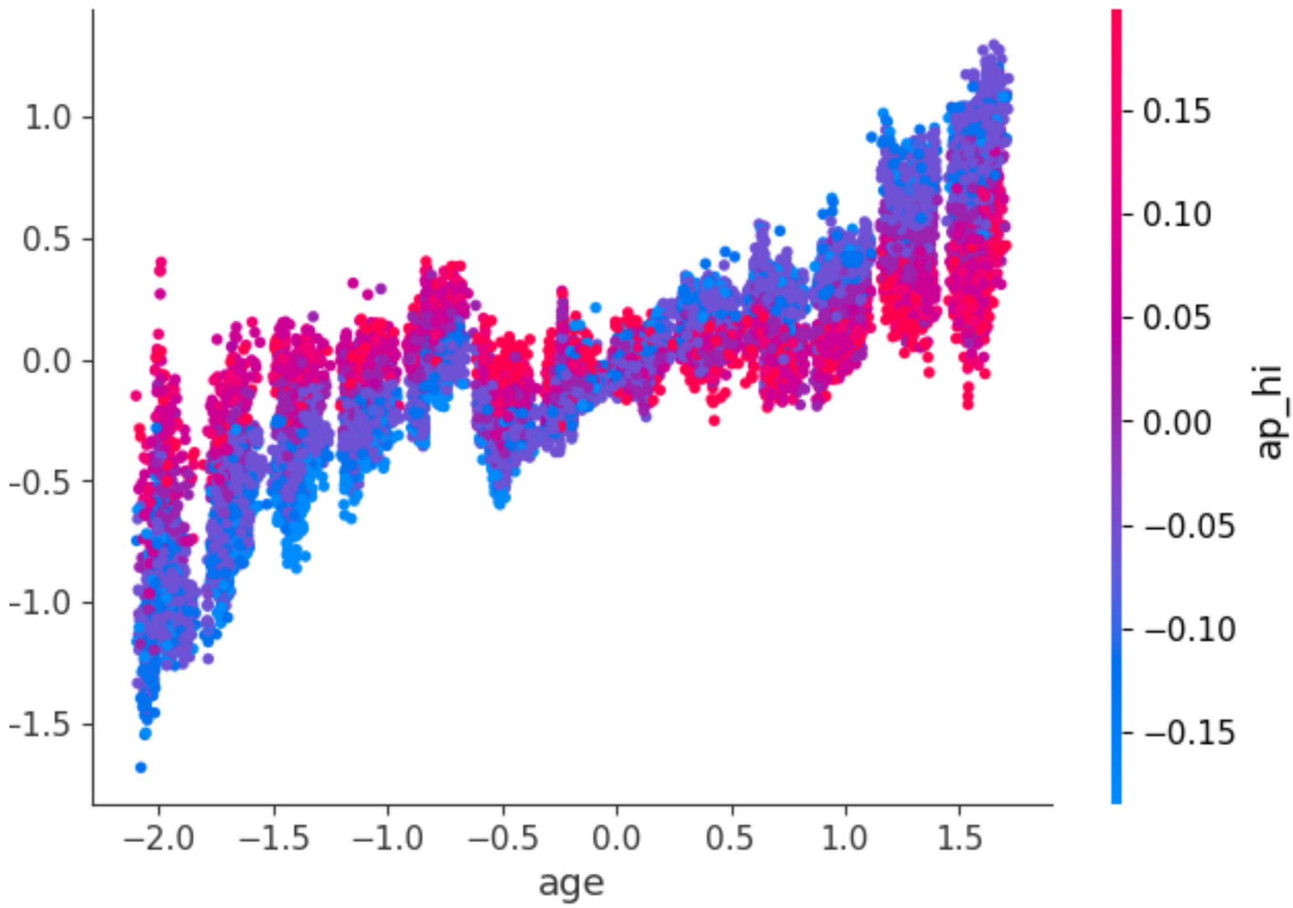

Fig. 13. SHAP value for age.

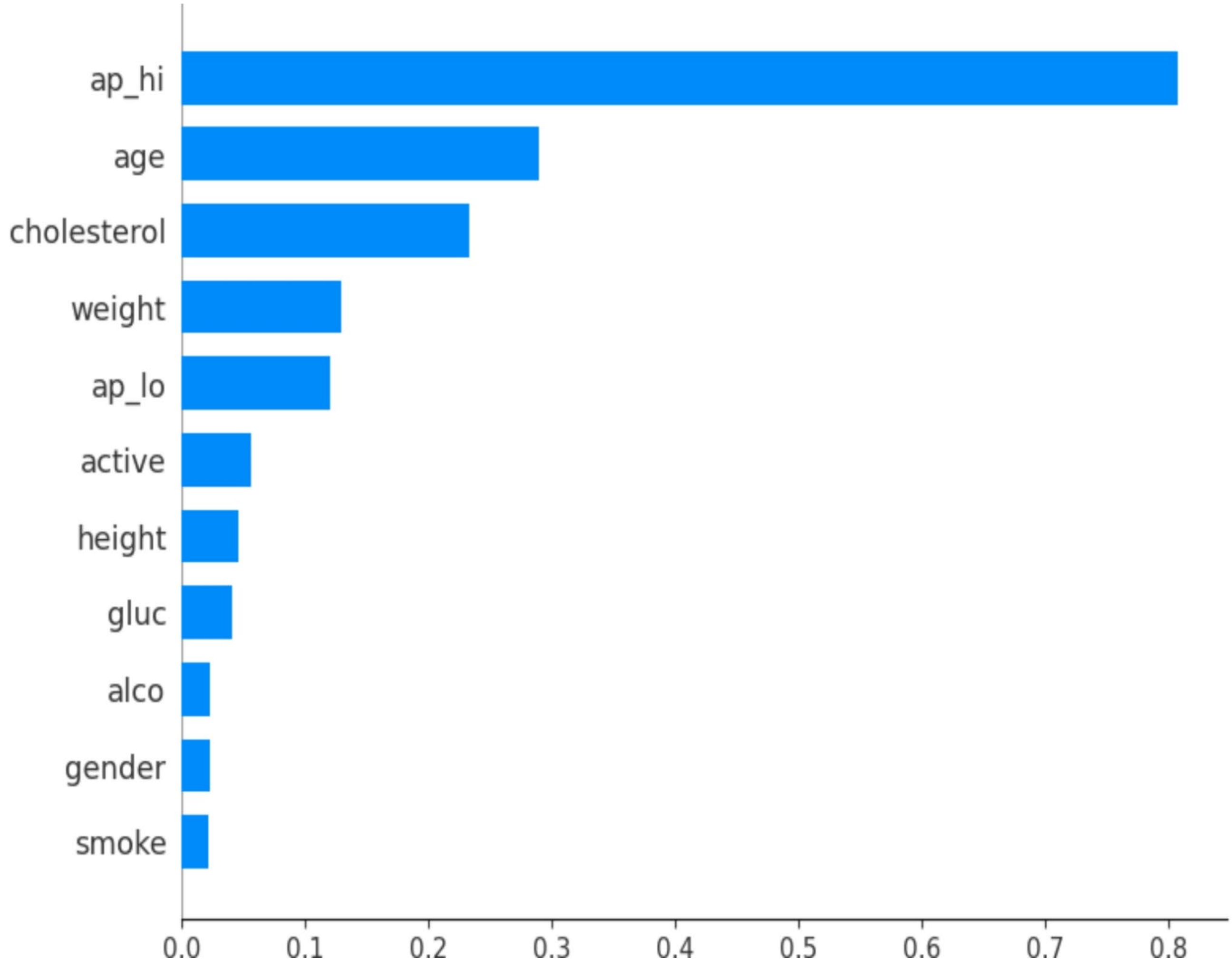

Fig. 14. SHAP summary plot for the training phase.

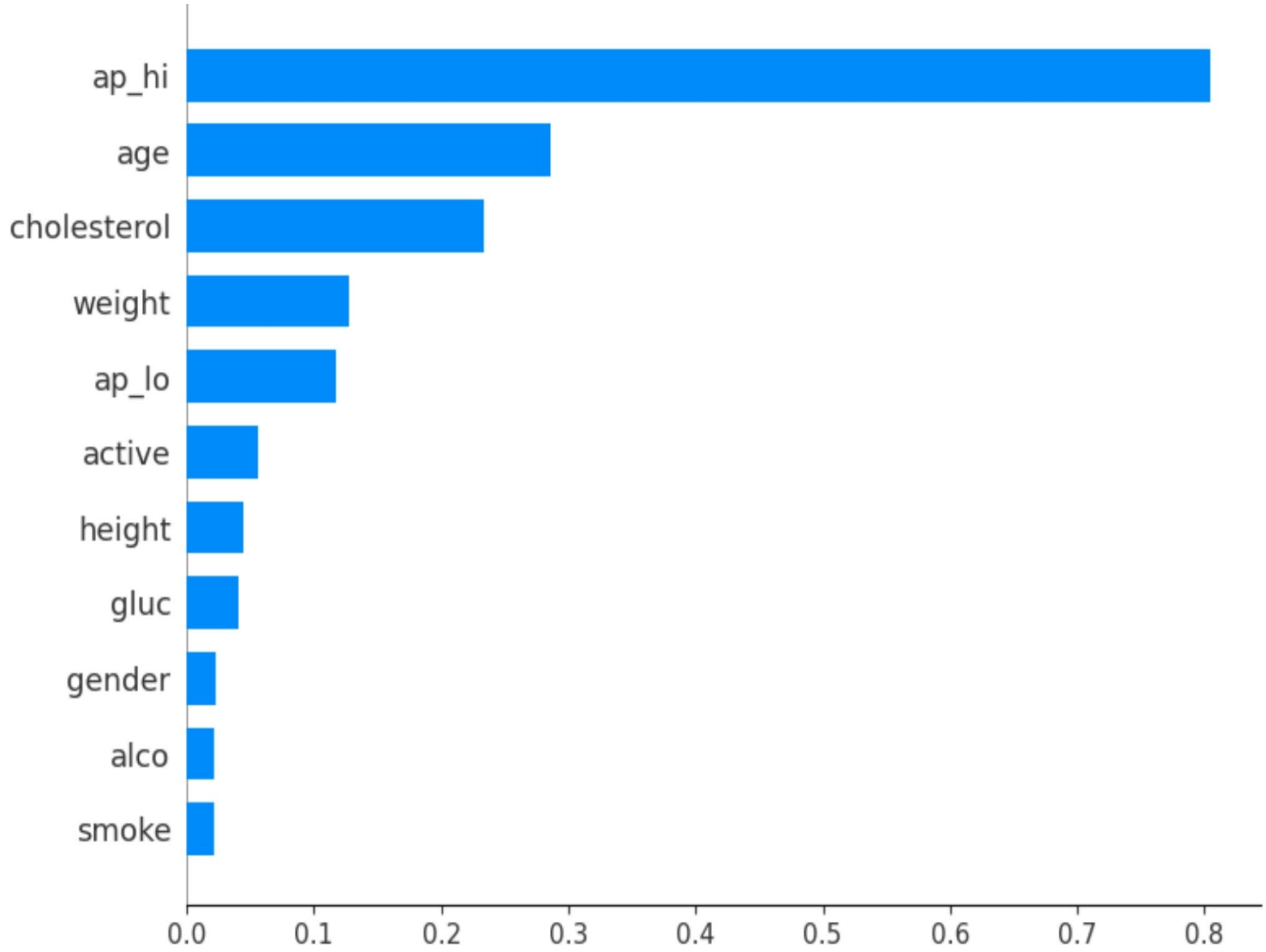

Fig. 15. SHAP summary plot for the validation phase.

XGBoost	Acc (%)	TPR (%)	TNR (%)	FNR (%)	FPR (%)	LR+	LR−	PPV (%)	NPV (%)
Training	98.17	97.47	98.65	1.83	1.35	72.2	0.019	98.05	98.26
Validation	97.40	97.83	96.84	2.60	3.16	30.96	0.027	97.61	97.12

Table 6. Proposed responsible healthcare chatbot approach performance in training and validation.

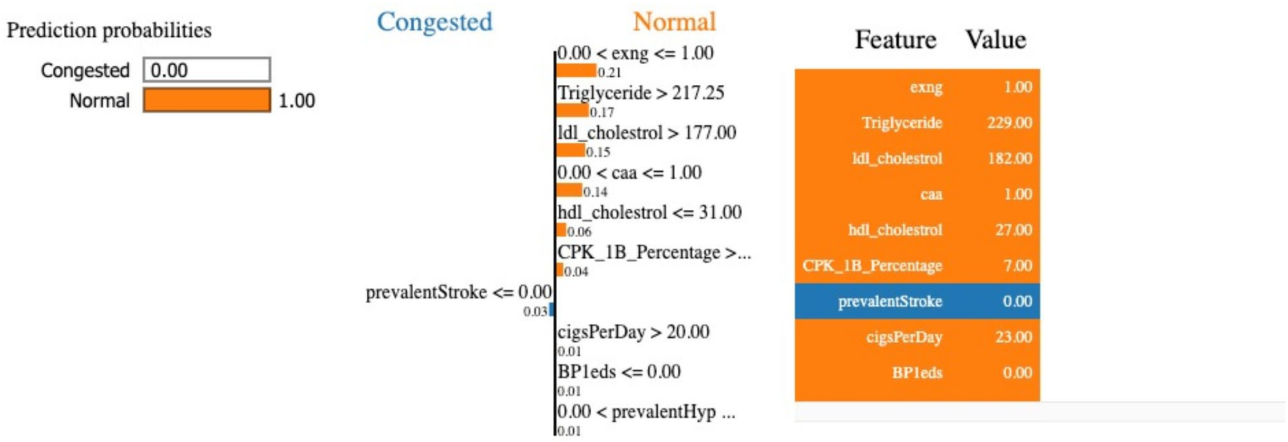

Fig. 16. Proposed responsible healthcare chatbot approach explanation with CVD prediction (No).

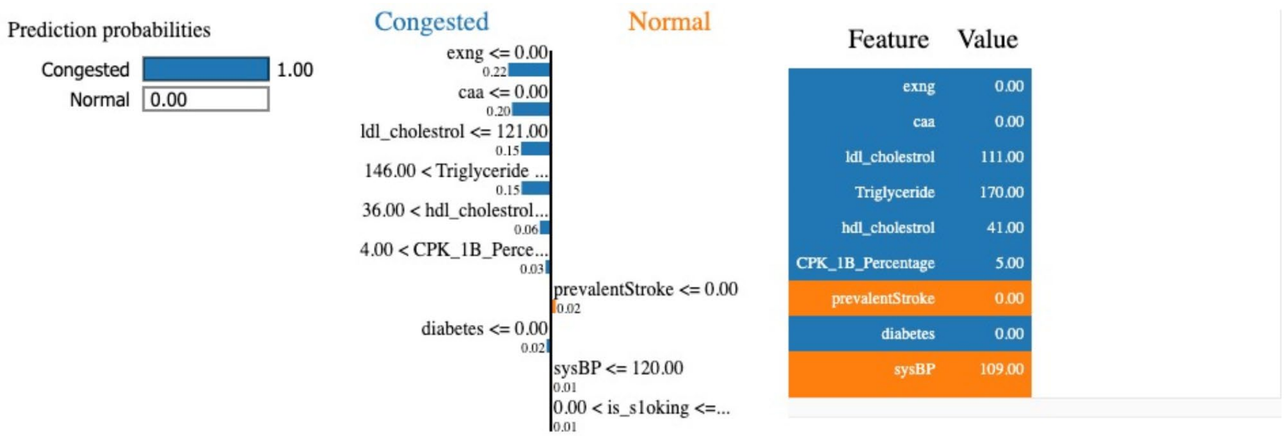

Fig. 17. Proposed responsible healthcare chatbot approach explanation with CVD prediction (Yes).

Approaches	Accuracy (%)	Miss-rate
Hybrid random forest with a linear approach (HRFLM)[52]	88.70	11.30
SVM[53]	85.40	14.60
KNN[54]	90.789	9.211
Logistic regression and KNN[55]	87.5	12.5
Multilayer perceptron with cross-validation[56]	87.28	12.72
Proposed approach	97.12	2.88

Table 7. Comparing the proposed responsible healthcare chatbot approach with previous approaches.

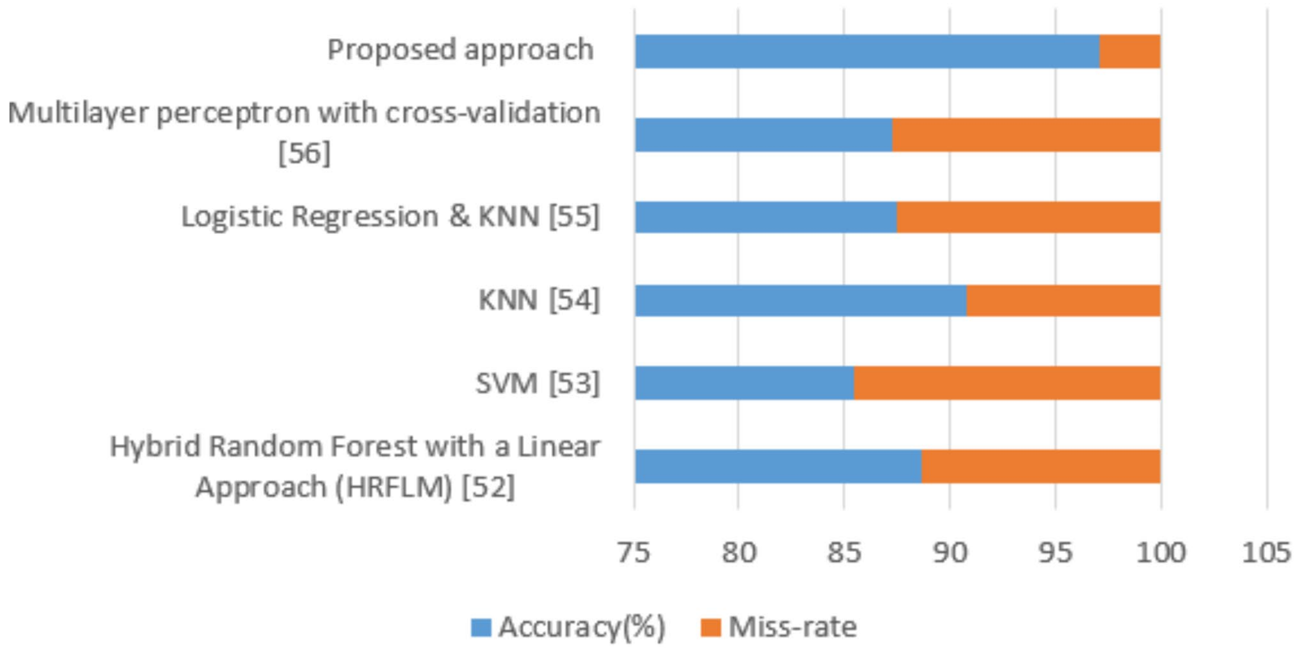

Fig. 18. Graphical representation of the previous approaches with the proposed approach.

Data availability

The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding authors.

Received: 31 January 2025; Accepted: 31 March 2025
Published online: 04 April 2025

References

1. Aslam, F. The impact of artificial intelligence on chatbot technology: A study on the current advancements and leading innovations. *Eur. J. Technol. Aug.* **15** (3), 62–72 (2023).
2. .Wang, J. et al. Machine learning assisted CALPHAD framework for thermodynamic analysis of CVD SiOxNy thin films. *Calphad* **88**, 102806 (2025).
3. Suta, P. et al. An overview of machine learning in chatbots. *Int. J. Mech. Eng. Rob. Res. Apr*. **9** (4), 502–510 (2020).
4. Nath, M. P. & Sagnika, S. Capabilities of chatbots and its performance enhancements in machine learning. In *Machine Learning and Information Processing: Proceedings of ICMLIP* 183–192 (Springer Singapore, 2020).
5. Ayanouz, S., Abdelhakim, B. A. & Benhmed, M. A smart chatbot architecture based NLP and machine learning for health care assistance. In *Proceedings of the 3rd International Conference on Networking, Information Systems & Security* 1–6 (2020).
6. Burr, C. & Floridi., L. The ethics of digital well-being: A multidisciplinary perspective. Ethics of digital well-being: A multidisciplinary approach. 1–29 (2020).
7. Ahmad, M. et al. Stylized heterogeneous dynamic links among healthcare expenditures, land urbanization, and CO_2 emissions across economic development levels. *Sci. Total Environ. Jan*. **20**, 753:142228 (2021).
8. Patrício, L. et al. Leveraging service design for healthcare transformation: Toward people-centered, integrated, and technology-enabled healthcare systems. *J. Serv. Manag.* **31**(5), 889–909 (2020).
9. Liu, T. et al. Cluster infections play important roles in the rapid evolution of COVID-19 transmission: a systematic review. *Int. J. Infect. Dis*. **99**, 374–380 (2020).
10. Bates, D. W. et al. The safety of inpatient health care. *N. Engl. J. Med*. **388**(2), 142–153 (2023).
11. Bachina, L. & Kanagala, A. Health revolution: AI-powered patient engagement.
12. Wang, X. et al. Revealing the complexity of users' intention to adopt healthcare chatbots: A mixed-method analysis of antecedent condition configurations. *Inf. Process. Manag*. **60**(5), 103444 (2023).
13. Shahsavar, Y. & Choudhury, A. *The Role of AI Chatbots in Healthcare: A Study on User Intentions To Utilize ChatGPT for Self-Diagnosis* (JMIR Preprints, 2023).
14. Solanki, R. K. Building a conversational chatbot using machine learning: towards a more intelligent healthcare application. In *Handbook of Research on Instructional Technologies in Health Education and Allied Disciplines*, 285–309 (IGI Global, 2023).
15. Jacquier-Bret, J. & Gorce, P. Prevalence of body area work-related musculoskeletal disorders among healthcare professionals: A systematic review. *Int. J. Environ. Res. Public Health.* **20**(1), 841 (2023).
16. Wenhua, Z. et al. Blockchain technology: security issues, healthcare applications, challenges and future trends. *Electron*. **20** (3), 546 (2023).
17. Vishwakarma, A. et al. Adoption of blockchain technology enabled healthcare sustainable supply chain to improve healthcare supply chain performance. *Manag. Environ. Qual. Int. J.*. **17** (4), 1111–1128 (2023).
18. Misra, D. K., Kumar, P. & Khari, M. Achieving the security in cyber world through blockchain: A taxonomy. In *2023 10th International Conference on Computing for Sustainable Global Development (INDIACom)*, 1552–1557 (IEEE, 2023).
19. Chaddad, A. et al. Survey of explainable AI techniques in healthcare. *Sens.*. **5** (2), 634 (2023).
20. Bharati, S., Mondal, M. R. & Podder, P. A review on explainable artificial intelligence for healthcare: Why, how, and when?? *IEEE Trans. Artif. Intell.* **13**. (2023).
21. Bharati, S. et al. Explainable artificial intelligence (XAI) with IoHT for smart healthcare: A review. *Interpret. Cogn. Internet Things Healthc.* **27**, 1–24 (2023).
22. Swanepoel, D. W., Manchaiah, V. & Wasmann, J. W. The rise of AI Chatbots in hearing health care. *Hear. J.* **76**(04), 26–30 (2023).
23. Hueso, M. et al. Artificial intelligence for the artificial kidney: pointers to the future of a personalized Hemodialysis therapy. *Kidney Dis*. **4**, 1–9 (2018).
24. Schario, M. E. et al. Chatbot-assisted care management. *Prof. Case Manag*. **27**, 19–25 (2022).
25. Chow et al. Design of an educational chatbot using artificial intelligence in radiotherapy. *AI* **4**, 319–332. https://doi.org/10.3390/ai4010015 (2023).
26. Kovacek, D. & Chow, J. C. An AI-assisted chatbot for radiation safety education in radiotherapy. *IOP Sci. Notes*. **2**, 034002. https://doi.org/10.1088/2633-1357/ac1f88 (2021).
27. Rebelo, N. et al. Learning the treatment process in radiotherapy using an artificial intelligence–assisted chatbot: development study. *JMIR Form. Res.* **6**, e39443. https://doi.org/10.2196/39443 (2022).
28. El Hajal, G., Daou, R. A. Z. & Ducq, Y. Human firewall: Cyber awareness using WhatApp AI chatbot. In *Proceedings of the IEEE 3rd International Multidisciplinary Conference on Engineering Technology, Beirut, Lebanon, 8–10 December*. (2021).
29. Purohit, S. et al. HonestChain: consortium blockchain for protected data sharing in health information systems. *Peer-to-peer Netw. Appl.* **14** (5), 3012–3028 (2021).
30. Kurdi, H. & Alsalamah, S. HealthyBroker: a trustworthy blockchain-based multi-cloud broker for patient-centered eHealth services. *Electronics* **8** (6), 602 (2019).
31. Agbo, C. C. & Mahmoud, Q. H. Blockchain in health-care. *Int. J. Healthc. Inf. Syst. Inf.* **15** (3), 82–97 (2020).
32. Brownlee, J. What is machine learning: A tour of authoritative definitions and a handy One-Liner you can use. https://www.machinelearningmastery.com (accessed 25 Nov 2023).
33. Knevel, R. & Liao, K. P. From real-world electronic health record data to real-world results using artificial intelligence. *Ann. Rheum. Dis.* **82**, 306–311 (2023).
34. Tandon, A. et al. Blockchain in healthcare: A systematic literature review, synthesizing framework and future research agenda. *Comput. Ind.* **122**, 103290 (2020).
35. Sadiku, M. N. O., Eze, K. G. & Musa, S. M. Block chain technology in healthcare. *Int. J. Adv. Sci. Res. Eng.* **4**, 154–159 (2018).
36. Deloitte *Breaking Blockchain Open-Deloitte's 2018 Global Blockchain Survey*, 48 (Deloitte Insights, 2018).
37. Shubbar, S. Ultrasound medical imaging systems using telemedicine and blockchain for remote monitoring of responses to neoadjuvant chemotherapy in women's breast cancer: Concept and implementation. (Master's Thesis), Kent State University, Kent, OH, United States (2017).
38. Ianculescu, M. et al. Innovative, adapted online services that can support the active, healthy and independent living of ageing people. A case study. *Int. J. Econ. Manag. Syst.* **2**, 321–329 (2017).
39. Laranjo, L. et al. Conversational agents in healthcare: a systematic review. *J. Am. Med. Inform. Assoc.* **25**(9), 1248–1258 (2018).
40. Adam, M., Wessel, M. & Benlian, A. AI-based chatbots in customer service and their effects on user compliance. *Electron. Markets.* **17**, 1–19. https://doi.org/10.1007/s12525-020-00414-7] (2020).

41. Androutsopoulou, A. et al. Transforming the communication between citizens and government through AI-guided chatbots. *Gov. Inform. Q.* **36** (2), 358–367. https://doi.org/10.1016/j.giq.2018.10.001] (2019).
42. Georgescu, A. Chatbots for education - trends, benefits and challenges. In *2018 Presented at: eLearning & Software for Education*; April 20, Bucharest, 195–200. https://doi.org/10.12753/2066-026x-18-097] (2018).
43. Nadarzynski, T. et al. Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: A mixed-methods study. *Digit. Health*. **5**, 2055207619871808 (2019).
44. Miner, A. S., Laranjo, L. & Kocaballi, A. B. Chatbots in the fight against the COVID-19 pandemic. *NPJ Digit. Med.* **3**(1), 65. https://doi.org/10.1038/s41746-020-0280-0] (2020).
45. Ahmad, R. W. et al. The role of blockchain technology in telehealth and telemedicine. *Int. J. Med. Inform.* **148**, 104399 (2021).
46. Fan, K. et al. Medblock: efficient and secure medical data sharing via blockchain. *J. Med. Syst.*. **42**, 1–1 (2018).
47. Quasim, M. T. et al., A blockchain framework for secure electronic health records in healthcare industry. In *International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE)*, 605–609. (IEEE, 2020).
48. Zarour, M. et al. Evaluating the impact of blockchain approachs for secure and trustworthy electronic healthcare records. *IEEE Access*. **27**, 157959–157973 (2020).
49. Görtz, M. et al. An artificial intelligence-based chatbot for prostate cancer education: design and patient evaluation study. *Digit. Health*. **9**, 20552076231173304 (2023).
50. Budler, L. C., Gosak, L. & Stiglic, G. Review of artificial intelligence-based question-answering systems in healthcare. *Wiley interdiscip. Rev. Data Min. Know. Discov.* **13**(2), e1487 (2023).
51. https://ieee-dataport.org/documents/cardiovascular-disease-data#files
52. Mohan, S., Thirumalai, C. & Srivastava, G. Effective heart disease prediction using hybrid machine learning techniques. *IEEE Access*. **19**, 7:81542–81554 (2019).
53. Ahamed, J. et al. CDPS-IoT: cardiovascular disease prediction system based on IoT using machine learning. (2022).
54. Shah, D., Patel, S. & Bharti, S. K. Heart disease prediction using machine learning techniques. *SN Comput. Sci.* **1**, 1–6 (2020).
55. Jindal, H. et al. Heart disease prediction using machine learning algorithms. In *IOP Conference Series: Materials Science and Engineering*, vol. 1022, No. 1, 012072. (IOP Publishing, 2021).
56. Bhatt, C. M. et al. Effective heart disease prediction using machine learning techniques. *Algorithms*. **6** (2), 88 (2023).
57. Muneer, S. et al. Explainable AI-Driven chatbot system for heart disease prediction using machine learning. *Int. J. Adv. Comput. Sci. Appl.* **15**(12). (2024).
58. Verma, G. A secure blockchain-based health prediction framework using Mantaray optimization. *Secur. Priv.* **8** (2), e70010 (2025).

Author contributions

Salman Muneer, Asghar Ali Shah, Sagheer Abbas, Meshal Alharbi and Haya Aldossary, have collected data from different resources and contributed to writing—original draft preparation. Khan Muhammad Adnan., Areej Fatima and Sagheer Abbas performed formal analysis and Simulation, Taher M. Ghazal, Asghar Ali Shah, Meshal Alharbi and Haya Aldossary; writing—review and editing, Asghar Ali Shah, and Khan Muhammad Adnan; performed supervision, Salman Muneer, Sagheer Abbas, Asghar Ali Shah, Ahmad Alshamayleh and Meshal Alharbi.; drafted pictures and tables, Khan Muhammad Adnan, Haya Aldossary and Areej Fatima; performed revisions and improve the quality of the draft. All authors have read and agreed to the published version of the manuscript.

Funding

No funding involved in this study.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to A.A.S. or K.M.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.