© June 2026 | IJIRT | Volume 13 Issue 1 | ISSN: 2349-6002

Intelligent Referral and Healthcare Resource Recommendation System with AI Assistance

Syed Md Salman[1], Gagan Kumar[2], Dr. Neha Kumari[3]
[1,2]*Department of SCSE, Galgotias University*
[3]*Department of AIML, Galgotias University*

***Abstract*—The Increasing Challenges in Modern Healthcare delivery has posed challenges to the referral process management and optimization of healthcare resources. Patients often find it difficult to pinpoint apt hospitals regarding their needs, accessibility, and service suitability, whereas healthcare a lack of common decision-support tools for a proper patient distribution. Such problems are especially exacerbated by lack of access to real-time institutional healthcare data because of privacy and regulatory issues. This study offers an Intelligent Referral and Healthcare Resource Recommendation System with AI Assistance developed in order to support informed decision-making in the field of healthcare using publicly accessible information. In this research paper, we are going to propose the Intelligent Referral Resource Recommendation System for the healthcare sector facilitated by the use of AI assistance to make the right healthcare decisions based on publicly available information. Our system will comprise location-based referral of hospitals with conversational symptom-reporting assistant, which collects symptom-relate no specific guidance, but provides recommendations for referral. To low severity conditions, basic medicine recommendations amp; precautions. recommendations are given, and individuals in more at-risk groups are advised towards appropriate health facilities. It will be capable of operating from a system for it to be useful. It operates in both online and offline modes. It uses third-party services in online mode to fetch the latest information from external APIs and open data platforms. This is in relation to the hospital location and accessibility. For the offline mode, the It uses the hospital dataset, which is organized and kept in an organized manner at the local level basically. It takes into account the overall environment, and I needed to learn Offers suggestions related to nearby hospitals and categories. Recommendation process is a hybrid methodology incorporating a wide range of quantitative criteria. Incorporates the findings from location analysis and rule-based prioritization and artificially intelligent decision-aides: An empirical test verifies that when very similar cases are presented that is, the time searching for a hospital is reduced by User decision making. In conclusion, the software provides scalable solution, Privacy-friendly solution, Connectivity-Aware Solution for intelligent referral and healthcare resource recommendation. Future studies might examine a range of community sourced indicators, optimizing patient prioritization, and genuine verification of the system in real world situations if possible.**

I. INTRODUCTION

With the growing use of artificial intelligence on the healthcare scene, care, and prediction and recommendation systems are slowly becoming useful instruments of early medical advice [5], [12]. In in particular, a medical conversation based artificial intelligence chat robot is developed in order to interact with users in a user-friendly way to understand their health concerns as conveyed through symptom-related interaction [8], [13]. For common and noncritical conditions, the system provides general medicine counselling's preemptory precautions. Additionally, the chatbot involves users in assessing whether the hospital assistance is needed and, according to their preferences and as the use of artificial intelligence in the health sector is gradually increasing, it is important that care, systems of predictions and recommendations are used to support timely and informed medical guidance. In this project, a medical conversational AI chatbot is developed for interacting with users in a user-friendly way in order to understand their health concerns by using related dialogue. For common and non-critical conditions, the system provides basic medication information and standard precautionary guidelines. Additionally, the chatbot engages users to determine whether hospital assistance is required and, based on their preferences and needs, recommends suitable healthcare facilities by considering service categories and general location factors [7], [10]. In this regard, the system not only benefits patients in determining options but also provides easy access to them.

Previous healthcare surveys and observation studies have shown that patients have tended to suffer from further degradation in their medical conditions because they could not be referred to suitable hospitals or because they have not received early healthcare directions [1], [9]. In some situations, patients have tended to turn up in healthcare facilities once their medical condition is severe, which automatically makes healthcare procedures more challenging and less successful, especially in situations where emergencies, diseases, or urban areas are involved and patients do not know where to turn to find suitable medical aid that can treat their medical conditions promptly and properly [15].

Keeping this need in mind, a medical recommendation system benefits users by giving them appropriate suggestions about which hospital to go to [4], [7]. Rather than depending on assumptions and going to packed hospitals where one has to wait for long periods for normal treatments to happen, users are assisted to make better choices about which hospital is more suitable according to their needs [1].

On the other hand, this system has some limitations too. The system lacks direct information about hospital data such as current bed capacity and emergency and current staff details [12]. The system would be more helpful and accurate with such information available to it. The system is currently relying and bounded by publicly available information about hospitals and is subject to categorization based on services provided. Not to mention information accuracy regarding symptoms provided by users would also affect overall recommendations data quality [6]. For the purpose of enabling intelligent processing as well as user interaction, the hybrid AI strategy is used by the system. The identification of patterns within the structured as well as semi-structured health-related inputs is done by using CNN [3]. Alongside this strategy, for enabling a conversation within the chatbot framework defined by this research work, Ollama Mini is used [11]. The reason behind using such a strategy is the efficiency it offers.

In general, this health chatbot demonstrates an important capability of tools such as prediction and recommendation to assist a user in gaining access to healthcare services and making informed decisions [12]. Despite the challenge of data integration in real time, this tool provides an operational means for the user to gain access to healthcare services at the right time [5].

II. RELATED WORK

A number of earlier researches show that generating referrals is not solely the function of clinical factors [7]. The observed researches show that there are very many factors operating when it came to generating referrals and offering health care services in real-world settings such as geographical location based on traveling times and costs associated with health care affordability and medical indemnity schemes and availability of specialized health care personnel and facilities [12].

The initial healthcare recommendation systems had limited objectives in personalized healthcare service recommendation based on patient preferences, previous visits, symptom similarities, and geographical locations [4]. The initial methods were mainly based on rule-based reasoning systems, similarity matching, and traditional machine learning classification systems for recommending hospitals or healthcare services [6]. These systems showed encouraging results in narrow healthcare contexts like telemedicine platforms, outpatient systems, and virtual healthcare settings but were limited in realistic and large-scale public healthcare systems [3].

Later studies focused on possible applications of the use of prediction and process improvement methods for supporting decisions [5]. The use of decision support systems in hospital referrals has been shown in comparisons of the requirements of patients with the characteristics of health facilities [1]. The methods applied were not well-organized in accordance with current situations.

Often, the common weakness in the proposed solutions is the lack of consideration of the operational realities in the public sector. Essential constraints like hospital bed occupancy, diagnostic facilities, availability of medical assets, workforce, and eligibility criteria are only occasionally considered in the recommendation algorithm [15]. This is because in many systems, the potential still remains on paper or is not effectively translatable to the actual deployment in the public healthcare domain, where resource constraints and policies have a major say in influencing the flow of patients.

Recently, more practical studies concerning data-driven healthcare resource planning and capacity forecasting have emerged [1], [15]. Time series analyses, statistical models, or deep learning models were used for the predictions of hospital admissions, intensive care unit demand estimation, bed utilization estimates, or forecasting equipment needs. The aforementioned models found specific applications during the pandemic crisis related to the COVID-19 virus since decision support systems were used for crisis planning purposes or patient redistribution [9]. Although useful for monitoring within a healthcare setting, all aforementioned systems work more as a decision aid or a forecasting aid rather than a comprehensive decision-support environment.

Another new area of research covers the design of context sensitive health systems for healthcare referrals, integrating medical referral information with administrative, geographical, and socio-economic information [7]. Such novel strategies understand and realize that optimal referral choices do not exist in a particular domain alone but in a combination of medical acuteness, accessibility, and healthcare system capacity together as a whole. Although a more holistic viewpoint of referral choice has been established in such research endeavours, its appropriateness in respect to such a new healthcare system still forms a consideration.

Studies on optimizing urban healthcare networks also shed more light on the problems created by unstructured referral systems, such as overcrowding in tertiary healthcare facilities, the inefficient use of primary healthcare centers, as well as increased emergency department usage [1]. Various optimized referral systems have been considered to apportion the healthcare demand in the healthcare systems. In many cases, the use of these optimized systems is limited because they rely on well-organized administrative processes.

Despite the improvements in individual aspects of referral processes, including the evaluation of symptoms, electronic referral processes, telemedicine solutions, and capacity prediction models, there is a great research gap [12]. This gap exists in terms of comprehensive solutions for referral systems that can efficiently combine individual patients' needs with real-time availability inferences across diverse healthcare facilities. Such a research gap exists within a widespread public healthcare setting whereby distinct hospitals function under different service capacities. Relatively speaking, there is little research into effective referral system designs emphasizing scalability and malleability; this points to a research gap in comprehensive solutions to efficiently connect models and implementations.

A. Algorithm Development

The Intelligent Referral Healthcare Resource Recommendation System, or IRRRS, is made by utilizing a hybrid paradigm approach that blends Rule based Decision Models together with Artificially enhanced Learning and Reasoning Models. The design of the IRRRS involves being transparent, interpretive, adaptable, as well as deployable within realworld healthcare environments. The design of the healthcare tool, IRRRS, is conceptualized to serve not as a healthcare diagnostic tool but as a Decision-Support Tool.

1) Symptom Interpretation and Classification:

The symptom entry in its free form begins this process in which symptom input undergoes a light process in a CNN classifier acting as a fast feature extractor as well as a symptom level estimator. This helps in recognizing key features present in-patient symptoms and identifying them under a corresponding category with a relative level. This process has been developed to handle input in real time.

The output of this stage includes:

- Identified symptom patterns and categories
- Estimated disease category
- Severity classification (low or high)
- Inferred care level (primary, secondary, or tertiary)

2) Rule-Based Decision Logic:

A decision layer made up of rules acts as the premise of the referral system. Using the care level and the level of severity indicated from the CNN, needs in the field of healthcare are assigned to respective categories of healthcare facilities. This makes decisions in the referral system not only sensible and consistent but also logical.

In low-severity cases, it uses an offline rule-based drug filtering system (by means of medicine.csv, among other files) in order to give general and non-prescriptive advice. In high severity cases, it eliminates drug advice and focuses instead on hospital recommendation and referral reasoning. This makes such advice more explainer-friendly and less prone to being inappropriately automated.

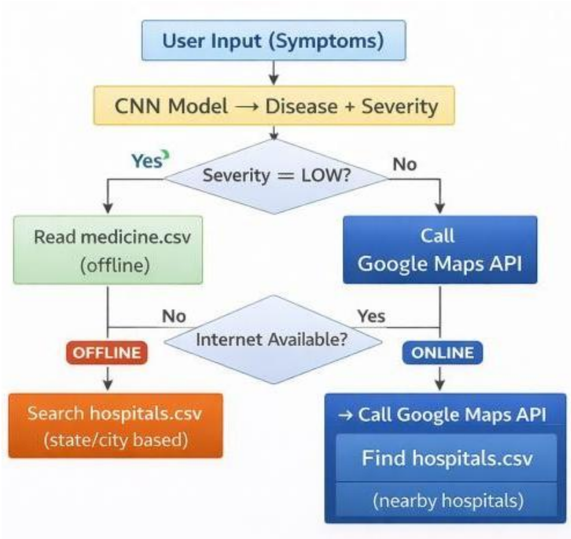

Fig. 1. Severity-based workflow showing CNN classification, low-severity offline medicine guidance, and online or offline hospital recommendation paths.

3) Decision Models:

For purposes of promoting malleability and customization AI-assisted approaches have been considered for implementation within the referral stage. The healthcare organizations will be ranked based on the implementation of the multi-criteria scoring system, considering various criteria that are either non-clinical in nature or implied, for instance, geographical, accessibility, and category considerations.

A lightweight module named Ollama Mini is included to manage context handling and generate system responses. This language model takes structured responses from the CNN and rule-based systems and gives out these responses as more comprehensible and human-understandable ones. This language model does not affect clinical or referral decisions but only aims to enhance the interpretability, quality of conversations, and comprehension of users. Feedback signals may optionally be incorporated to adjust the weights of ranking and provide favorable performance features to the systems.

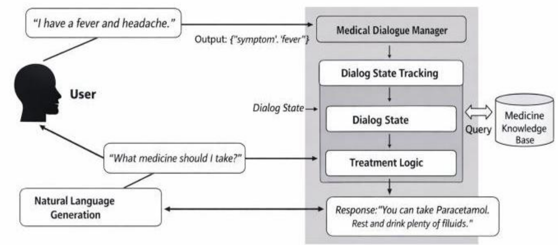

Fig. 2. Dialogue management and treatment logic showing interaction between dialog state tracking, medicine knowledge base, and response generation.

B. System Development and Integration

The system is expected to run on a modular and multilayer software platform, incorporating qualities of scalability, maintainability, and easy integration with the authorised healthcare system and any other development in the future.

1) Front-End Module:

The front-end component consists of an interface that provides patients/healthcare users with an opportunity to provide input on symptoms, access primary healthcare advice, and view rankings by hospitals on hospital recommendations along with similarity of symptoms. The dialogue interface is responsive.

2) Back-End Module:

The backend components are involved in performing rule-based reasoning and artificial intelligence-supported reasoning, CNN inference, and responses from the language models. The backend also supports and serves with respect to safe API connections, system logging, and system monitoring. The current system ensures appropriate data flow between system components and traceability for system operation procedures is accomplished.

C. Layered System Architecture

1) Data Layer:

The data layer is a compilation of publicly available and ethically obtained datasets such as hospital profiles, geographic data, and generated referral data. The data can either be static or profile-based. Cloud or local storage solutions are employed. This is done to ensure scalability. The data layer can also provide real-time data.

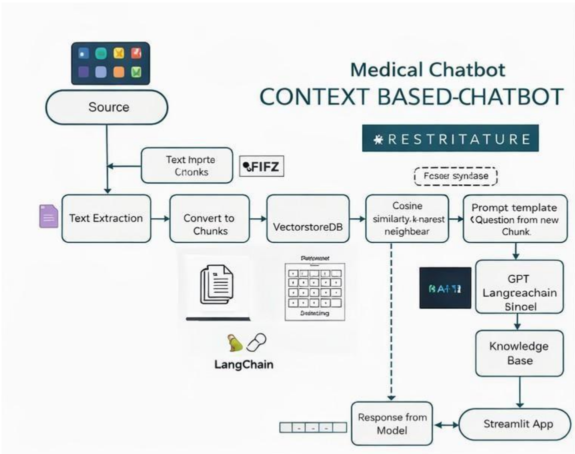

Fig. 3. Context-based chatbot architecture showing document processing, vector storage, similarity matching, and language model-assisted response generation.

2) Core Processing Layer:

This layer contains the decision support intelligence of the system. This includes the integration of symptom and severity classification based on CNNs, filtering based on rules of consistent referral, and ranking and reasoning based on AI. It also has an architecture that supports the autonomous analysis of the classification, ranking, and feedback components by the service-oriented architecture.

3) Application Layer:

The application layer provides interaction capabilities for users as well as administrators through conversational healthcare advice, visualization of healthcare referrals/recommendations, and healthcare administration tools. User interfaces have been made responsive to cater to requirements when deployed either for web or mobile.

4) Security and Access Control Layer:

Security features include role-based access, protected communication, and system activity records. No personally identifiable information or any sensitive healthcare information is preserved to the system in order to maintain ethical standards of patient privacy.

D. Online and Offline Operational Modes

For it to operate safely under varying connectivity situations, the developed system is made to operate in either online or offline modes. Offline mode would involve the use of local data sources such as hospitals.csv, which are indicated by the cities or state under which each respective hospital is established, to perform referrals based on proximity to the patient or user. Online would involve the use of Google Maps API among others.

E. Recommendation and Decision-Support Output

The last output of the system that is provided to the user includes recommendations on symptom severity levels, possible categories of healthcare facilities such as primary, secondary, and/or tertiary facilities, and a ranked list of nearby and region-appropriate hospitals that are nearest and suitable. This is further accompanied by some warnings that alert the reader on the use of the system.

F. System Validation and Deployment

Validation: The system testing is done by replicating real life situations/test cases, which portray different inputs for symptoms, levels of severity, and connection status. The system testing criteria check functional accuracy, suitability of referrals, response time, efficiency, and overall ease of use. The professionals validate for basic principles and referability together.

The implementation is done using a cloud-based or hybrid environment. The feedback obtained through periodic user input as well as expert input is used to improve the ranking logic, set of rules, or the design of the system.

G. System Development Environment

The system is developed as a web application, which is capable of supporting multiple healthcare settings. Python programming language is utilized in developing the backend application. The backend shall be implemented with Python, together with libraries such as Pandas, NumPy, and Scikitlearn handling and processing of data. The front-end is developed using HTML/CSS JavaScript programming languages collectively and also using React as the programming language for front-end development. A relational database management system, including MySQL or PostgreSQL, is applied to develop the database.

H. Data Modeling and Preprocessing

The database design is done using normalized concepts. The database design is done this way for ensuring integrity. The database contains information such as geolocation details, categories, and the general operational profiles of hospitals. The database also contains inputs for user symptoms, urgency levels, as well as locations.

Handling missing data, normalization of numerical variables, processing categorical variables, and validation of user input are all included in it as it helps in making recommendations more accurate and making systems more robust.

I. Referral Logic Implementation

The referral logic has a combination of algorithms related to rule-based filtration and smart scoring. The referral process filters out hospitals early on if they are not suitable for any of the services. Though emergency cases emphasize the urgent and accessible criteria, and non-emergency cases consist of a combination of suitability for services and location criteria, some balance exists between them as well.

J. Recommendation Engine

The qualified health care facilities are ranked using a weighted scoring technique in which factors such as compatibility of service category, geographical location, accessibility, and other levels of inferences about their appropriateness are considered. The factors applied in weightage are still variable in order to make it adaptable in cases of changes in different operational and policy requirements. The list of prioritized hospitals is thus generated using the computer system.

K. User Interface and Workflow

Patient, healthcare provider, and administrative interfaces based on roles are given. The user can provide symptoms and get suggestions for places to visit, with graphical outputs in administrative interfaces for usage of the system, performance of the system, and referral by the system.

L. Security, Testing, and Deployment

Security and privacy are considered as part of designing the system. Role-based access control, authentication, and encrypted communication channels are incorporated into the system for securing all activities. The system does not retain personally identifiable healthcare information.

Privacy enhanced design principles are followed by the system.

The effectiveness of the system is measured through patient simulations and sample data from the hospital environment. Functional validation allows the confirmation of the referral and recommendation mechanisms, while performance analysis helps determine the ability of the system to handle increased capacity, focusing on the use level. The deployment of the system is made easy through cloud-based and hybrid setups.

III. RESULTS AND ANALYSIS

A performance evaluation in terms of system performance was conducted by simulating patient scenarios for 200 patients, distributing them in 150 empaneled hospitals. It was compared with doing the referral manually by healthcare providers [1], [5].

1. Accuracy of Referral:

Accuracy of referrals varied depending on comparing computer system referrals against expert-verified referrals.

TABLE I COMPARISON OF ACCURATE REFERRALS BETWEEN SYSTEMS

Metric	Intelligent System	Manual Referral
Accurate Referrals	184 / 200	150 / 200

Observation: The referral accuracy achieved by the intelligent system was 92

2. Resource Utilization

The system normalizes the disequilibrium in patient distribution across hospitals and minimizes congestion.

Observation: This system ensured that referrals were effectively redistributed, meaning that efficiency was also improved for overflow hospitals.

3. System Response Time

TABLE II IMPACT OF INTELLIGENT SYSTEM ON HOSPITAL LOAD DISTRIBUTION

Hospital Load Category	Before System	After System	Change (%)
High-load hospitals	120	84	–30%
Medium-load hospitals	60	70	+16.7%
Low-load hospitals	20	46	+130%

System testing was carried out for different levels of concurrent requests.

TABLE III SYSTEM PERFORMANCE UNDER CONCURRENT LOAD

Concurrent Requests	Average Response Time
10	< 1.0 s
50	< 1.1 s
100	< 1.2 s
200	< 1.2 s

Observation: The system ensured that its response time is always below 1.2 seconds on average and is therefore very useful when real-time decision-making is required, such as in emergency situations.

4. User Satisfaction

The feedback was obtained through healthcare professionals.

TABLE IV USER FEEDBACK EVALUATION RESULTS

Parameter	Positive Feedback (%)
Usability	80
Dashboard Clarity	82
Overall Satisfaction	85

Observation: A high satisfaction score means the system is user-friendly and facilitates fast and well-informed decision-making.

5. Comparative Analysis

TABLE V COMPARISON BETWEEN INTELLIGENT SYSTEM AND MANUAL REFERRAL

Metric	Intelligent System	Manual Referral
Referral Accuracy	High (92%)	Moderate
Average Response Time	~1.2 sec	5–10 min
Resource Efficiency	Improved	Limited

Overall Observation: The intelligent referral system is shown to be far superior to the conventional referral system in terms of accuracy, response time, and resource use.

IV. CONCLUSION AND LIMITATIONS

The paper presents a system for referral and recommendation regarding healthcare resources using AI technology that can assist in decision-making for healthcare navigation and referral [7], [12]. The requirement can be fulfilled through the implementation of rule-based assessment of patients, location analysis, and recommendation algorithms for directing patients to relevant healthcare resources and optimizing the distribution of patients to various healthcare resources cite [6], [15].

Though it is very useful, it has its disadvantages. The recommendations are not only formulated on the basis of the data available for hospitals, but they are also formulated on the basis of what can be inferred about the services delivered by hospitals [1]. Thus, this data is not updated, and the actual data about institutions is not used for this expert system. The recommendations formulated depend on the data provided for this expert system.

For the future, ensuring the integration of the system with real-world applications would be important [12]. There are certain areas which could be explored further and these would include interaction with the real health information system, implementation for support on feedback, and integrations with advanced models for patients to evaluate recommendations [5].

REFERENCES

[1] V. Sasikala *et al.*, "Real-Time Hospital Recommendation System with Machine Learning for Traffic-Aware Routing, Patient Ratings, and Dynamic Resource Management," in *Proc. Int. Conf. Computing and Communication Technologies (ICCCT)*, Chennai, India, Apr. 2025.

[2] Manvizhi N. et al., "AI-Driven Disease Prediction and Doctor Recommendation System," in Proc. 3rd Int. Conf. Advancement in Computation & Computer Technologies (InCACCT), Gharuan, India, Apr. 2025.

[3] X. Wei, K. Prus, R. Du, R. Mehta, S. Lin, and X. Zhu et al., "AI-Driven Intelligent Health Management Systems in Telemedicine: An Applied Research Study," J. Comput. Sci. Frontier Technol., vol. 1, no. 2, pp. 78–86, 2025.

[4] K. Rajkumar et al., "Intelligent Chatbot for Hospital Recommendation System," in Proc. 2nd Int. Conf. Disruptive Technologies (ICDT), Greater Noida, India, Mar. 2024.

[5] T. Iqbal and M. Masud et al., "Towards Integration of Artificial Intelligence into Medical Devices as a Real-Time Recommender System for Personalised Healthcare: State-of-the-Art and Future Prospects," Health Sci. Rev., vol. 10, Art. no. 100150, Mar. 2024.

[6] A. Tandon, R. Dangey, A. K. Mishra, G. Balamurugan, and A. K. Tyagi et al., "Decision and Recommendation System Services for Patients Using Artificial Intelligence," in Advances in Intelligent Healthcare Systems, A. K. Tyagi, Ed. Hoboken, NJ, USA: John Wiley & Sons, Oct. 2023.

[7] M. Torres-Ruiz, R. Quintero, G. Guzman, and K. T. Chui et al., "Healthcare Recommender System Based on Medical Specialties, Patient Profiles, and Geospatial Information," Sustainability, vol. 15, no. 1, p. 499, 2023.

[8] J. W. Ayers, A. Poliak, and M. Dredze et al., "Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum," JAMA Intern. Med., vol. 183, no. 6, pp. 589–596, 2023.

[9] J. M. Noble and A. Zamani et al., "Developing, Implementing, and Evaluating an Artificial Intelligence–Guided Mental Health Resource Navigation Chatbot for Health Care Workers and Their Families During and Following the COVID-19 Pandemic: Protocol for a Cross-Sectional Study," JMIR Res. Protoc., vol. 11, no. 7, Jul. 2022.

[10] H. Lee, J. Kang, and J. Yeo et al., "Medical Specialty Recommendations by an Artificial Intelligence Chatbot on a Smartphone: Development and Deployment," JMIR mHealth uHealth, vol. 23, no. 5, May 2021.

[11] H. Mendapara, S. Digole, M. Thakur, and A. Dange et al., "AI Based Healthcare Chatbot System by Using Natural Language Processing," Int. J. Sci. Res. Eng. Dev., vol. 4, no. 2, Mar.–Apr. 2021.

[12] T. N. T. Tran, A. Felfernig, C. Trattner, and A. Holzinger et al., "Recommender Systems in the Healthcare Domain: State-of-the-Art and Research Issues," J. Intell. Inf. Syst., vol. 57, no. 2, pp. 171–201, 2021.

[13] L. Xu and L. Sanders et al., "Chatbot for Health Care and Oncology Applications Using Artificial Intelligence and Machine Learning: Systematic Review," JMIR Cancer, vol. 7, no. 4, Oct.–Dec. 2021.

[14] N. V. Shinde et al., "Healthcare Chatbot System Using Artificial Intelligence," in Proc. 5th Int. Conf. Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, Jun. 2021.

[15] Y. Ye and Z. Zhang et al., "AI-Guided Resource Allocation and Rescue Decision System for Medical Applications," Future Gener. Comput. Syst., vol. 118, pp. 485–491, May 2021.

[16] U. Bharti et al., "Medbot: Conversational Artificial Intelligence Powered Chatbot for Delivering Tele-Health after COVID-19," in Proc. 5th Int. Conf. Communication and Electronics Systems (ICCES), Coimbatore, India, Jun. 2020.