A Sequence-Aware Medical Chatbot Framework for Intelligent Healthcare Interaction

Y.Akhila[1], M.Anusha[2]
[1]MCA Student , Dept of MCA , Ellenki College of Engineering and Technology , Hyderabad
[2]Assistant Professor , Dept of MCA , Ellenki College of Engineering and Technology , Hyderabad

Abstract— The increasing demand for timely and reliable healthcare information has encouraged the development of intelligent conversational systems capable of assisting users in obtaining medical guidance. This paper presents an AI-driven medical chatbot designed to answer queries related to infectious diseases through both text and voice-based interaction. The proposed system employs a Long Short-Term Memory (LSTM) neural network trained on a structured medical question–answer dataset to recognize user queries and retrieve the most appropriate responses. The application incorporates user authentication, multilingual communication support, speech recognition, voice response generation, and conversation history management to improve accessibility and user experience. A sequence of preprocessing and training operations is performed to prepare the dataset and optimize the learning model before deployment. The trained chatbot is capable of responding in both English and Telugu, enabling wider usability among diverse users. Experimental evaluation indicates that the LSTM model achieves approximately 99% training accuracy, demonstrating its effectiveness in identifying relevant responses for disease-related questions while maintaining low prediction error. The integrated framework offers a practical solution for delivering instant medical information, reducing response time, and providing accessible healthcare assistance through an interactive web-based environment. The proposed system can serve as a supportive tool for preliminary health guidance, public awareness, and digital healthcare services while improving communication between users and intelligent medical assistance platforms.

Keywords— Medical Chatbot, Infectious Disease Prediction, Long Short-Term Memory (LSTM), Natural Language Processing (NLP), Voice-Based Interaction, Text-Based Chatbot, Machine Learning, Healthcare Assistance, Multilingual Chatbot, Conversational Artificial Intelligence.

I.INTRODUCTION

The growing demand for digital healthcare services has accelerated the adoption of intelligent systems capable of providing immediate medical assistance to users. Infectious diseases continue to pose significant public health challenges worldwide, making timely access to reliable health information increasingly important [26]. Many individuals seek preliminary medical guidance through online platforms before consulting healthcare professionals. However, conventional information retrieval methods often require extensive manual searching and may provide inconsistent or unreliable responses. These challenges have encouraged the development of intelligent conversational systems that can deliver accurate and contextually relevant medical information through natural language interaction [17].

Recent advances in artificial intelligence and natural language processing have enabled chatbots to become effective virtual assistants across healthcare applications [5]. By learning from historical medical question–answer datasets, machine learning models can recognize user intentions and generate suitable responses with minimal human intervention. Among various deep learning techniques, Long Short-Term Memory (LSTM) networks have demonstrated strong capabilities in understanding sequential text data, preserving contextual information, and improving response prediction accuracy [25]. These characteristics make LSTM-based architectures well suited for developing conversational systems that support healthcare information services and patient assistance [18].

Although several healthcare chatbot solutions have been proposed, many existing systems primarily focus on text-based communication or require complex computational frameworks and large training datasets [9]. Some applications also provide responses in only a single language, limiting accessibility for users from diverse linguistic backgrounds. Integrating voice-based interaction, multilingual communication, and efficient response generation within a unified platform remains an important research objective [22]. A practical healthcare chatbot should offer accurate information retrieval while maintaining ease of use, quick response time, and an intuitive user experience [13].

The proposed system introduces an intelligent medical chatbot for infectious disease prediction and healthcare assistance using a Long Short-Term Memory (LSTM) model. The framework is trained using a structured medical question–answer dataset and incorporates user authentication, dataset training, text-based conversation, voice-enabled interaction, multilingual response generation, and conversation history management within a web-

based environment [6]. The trained model delivers relevant responses in both English and Telugu while supporting speech recognition and voice synthesis for improved accessibility. Experimental evaluation demonstrates that the developed chatbot achieves high prediction accuracy and provides an effective platform for delivering preliminary healthcare information, enhancing user engagement, and supporting digital healthcare services [11].

II.RELATED WORK

R. Dharwadkar and N. A. Deshpande (2018) introduced a medical chatbot designed to provide users with instant responses to common healthcare-related questions through conversational interaction [9]. The proposed system demonstrated how natural language processing and chatbot technologies can reduce the burden on healthcare professionals by offering preliminary medical guidance. The study highlighted that conversational interfaces improve accessibility to health information and enable users to obtain basic medical advice without requiring continuous human supervision. The authors concluded that intelligent chatbot systems have significant potential in supporting digital healthcare services and patient engagement.

Madhu et al. (2017) presented a chatbot-based medical assistance system that utilizes a trained conversational model to respond to healthcare queries efficiently [10]. Their work focused on improving the interaction between patients and automated healthcare applications by providing accurate responses based on previously learned medical knowledge. The researchers emphasized that integrating machine learning techniques into chatbot systems enhances response quality while minimizing manual intervention. Experimental observations indicated that such intelligent assistants can serve as reliable support tools for preliminary healthcare consultation.

M. Almalki and F. Azeez (2020) conducted a comprehensive review of health-oriented chatbot applications developed during the COVID-19 pandemic to support public healthcare services [15]. The study discussed how conversational agents assist users by providing health education, symptom-related guidance, and awareness regarding disease prevention. The authors observed that AI-powered chatbots improve healthcare accessibility by delivering immediate responses while reducing unnecessary hospital visits. They concluded that healthcare chatbots can become valuable components of future digital health ecosystems when combined with reliable medical knowledge and intelligent learning techniques.

U. Bharti et al. (2020) developed Medbot, an artificial intelligence-powered conversational system intended to facilitate telehealth services during the COVID-19 outbreak [17]. The proposed framework enabled users to communicate with the chatbot using natural language to obtain information related to symptoms, disease awareness, and healthcare recommendations. The authors demonstrated that conversational AI can effectively support remote healthcare delivery by offering continuous assistance without requiring direct physical interaction. Their findings suggested that intelligent medical chatbots can improve healthcare accessibility during public health emergencies.

N. Battineni, N. Chintalapudi, and F. Amenta (2020) investigated the role of artificial intelligence chatbots in managing healthcare communication during epidemic situations such as the COVID-19 pandemic [18]. The study emphasized that AI-driven conversational systems can provide accurate health information, reduce misinformation, and assist users in obtaining timely medical guidance. The authors discussed the importance of integrating machine learning and natural language processing techniques to improve response quality and user interaction. Their research concluded that healthcare chatbots represent an effective digital solution for supporting public health communication and emergency response services.

III.DATASET DETAILS

The proposed intelligent medical chatbot is developed using a publicly available COVID-19 medical question–answer dataset available in JSON format. The dataset contains a collection of healthcare-related questions and their corresponding responses covering topics such as COVID-19 symptoms, preventive measures, medical assistance, oxygen support, emergency helpline information, and other frequently asked healthcare queries. Each question–answer pair represents a conversational example that enables the chatbot to understand user intentions and retrieve the most appropriate response during interaction. The structured nature of the dataset makes it suitable for training sequence-based deep learning models capable of learning conversational patterns.

Before training the prediction model, the dataset undergoes several preprocessing operations to improve its quality and compatibility with the learning algorithm. The collected question and answer pairs are organized into structured text sequences, where unnecessary symbols and irrelevant characters are removed to obtain meaningful input. The textual information is then tokenized and converted into numerical

representations that can be processed by the Long Short-Term Memory (LSTM) network. During preprocessing, the vocabulary is prepared, sequence lengths are standardized through padding, and the processed data is divided into training and validation samples to ensure reliable model learning and unbiased performance evaluation. These preprocessing steps help the chatbot understand contextual relationships between user questions and corresponding medical responses.

The processed dataset is subsequently used to train the LSTM-based conversational model for automated medical response generation. During the testing phase, the chatbot accepts either text or voice input from users, identifies the most relevant question pattern learned during training, and generates the corresponding medical response. The application further translates the generated reply into both English and Telugu, improving accessibility for multilingual users. Experimental evaluation reported in the developed system demonstrates that the trained LSTM model achieves approximately 99% training accuracy, indicating its capability to learn conversational patterns effectively and provide accurate responses for infectious disease-related queries. The availability of diverse healthcare questions allows the proposed framework to deliver fast, consistent, and user-friendly medical assistance through an interactive web-based platform.

IV.PROPOSED METHODOLOGY

The proposed framework presents an intelligent medical chatbot that employs a Long Short-Term Memory (LSTM) neural network to provide automated responses for infectious disease-related queries. The primary objective of the system is to deliver accurate and timely healthcare information through an interactive conversational platform, thereby reducing the dependency on manual information retrieval. The chatbot is trained using a publicly available medical question–answer dataset containing information related to COVID-19 symptoms, preventive measures, medical assistance, emergency services, and general healthcare guidance. The developed application supports both text-based and voice-based communication, allowing users to interact with the chatbot in a natural and convenient manner.

The methodology begins with dataset collection, followed by preprocessing to prepare the textual data for model training. During this stage, irrelevant characters and incomplete records are removed, and the question–answer pairs are organized into structured sequences. The text is tokenized, converted into numerical representations, and padded to obtain a uniform sequence length suitable for the LSTM network. A vocabulary dictionary is created to encode each word into a unique numerical index, enabling efficient sequence learning. The processed dataset is then divided into training and validation sets for reliable model development.

The prepared data is supplied to the LSTM model, which learns contextual relationships between user questions and corresponding responses through its memory cell architecture. After training, the chatbot predicts appropriate replies for unseen queries received through text or speech. Voice inputs are converted into text using speech recognition, while the generated responses are provided in both English and Telugu using translation and text-to-speech modules. The trained model achieves approximately 99% training accuracy, demonstrating effective conversational learning and reliable healthcare assistance through a multilingual web-based platform.

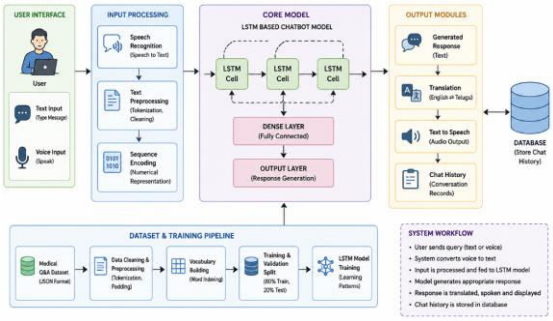

Figure 1: System Architecture of the Proposed LSTM-Based Medical Chatbot for Infectious Disease Prediction

Figure 1 presents the architecture of the proposed LSTM-based medical chatbot designed to provide intelligent responses to infectious disease-related queries. The system begins with a user interface that accepts both text and voice inputs. Voice inputs are converted into text through a speech recognition module, while textual queries are directly processed. The input undergoes preprocessing, including text cleaning, tokenization, and sequence encoding, before being supplied to the Long Short-Term Memory (LSTM) model. The LSTM network analyzes the processed sequence, captures contextual information, and generates the most appropriate response. The generated reply is displayed as text, translated into English and Telugu, and converted into speech using a text-to-speech module. User conversations are stored in a database for future reference. The lower section of the architecture represents the dataset preparation and model training process, where medical question–answer data is preprocessed and used to train the LSTM model. This integrated framework provides an efficient, accurate, and user-friendly healthcare assistance system.

V.RESULT AND DISCUSSION

The proposed LSTM-based medical chatbot was evaluated using the developed web-based

application under various healthcare query scenarios. The application successfully performed user registration, login, model training, text-based conversation, voice-based interaction, multilingual response generation, and conversation history management within a single platform. During execution, the medical question–answer dataset was imported and preprocessed by organizing conversational pairs, cleaning textual content, tokenizing words, converting text into numerical sequences, and padding the sequences before model training. The processed dataset was then utilized to train and validate the LSTM network for conversational response prediction.

The Long Short-Term Memory (LSTM) model was trained using the prepared dataset to learn contextual relationships between healthcare questions and their corresponding responses. Experimental observations showed that the trained model achieved approximately 99% training accuracy, demonstrating its ability to recognize conversational patterns effectively. The developed application successfully accepted both text and voice inputs, generated appropriate responses for previously unseen user queries, translated the responses into English and Telugu, and produced speech output through the integrated text-to-speech module. In addition, all user interactions were stored in the database, allowing previous conversations to be retrieved through the chat history module.

The experimental results indicate that the proposed framework provides accurate, responsive, and user-friendly healthcare assistance for infectious disease-related queries. By integrating dataset preprocessing, LSTM-based sequence learning, multilingual communication, speech recognition, voice synthesis, and conversation management into a unified application, the developed system offers an efficient and scalable solution for intelligent medical assistance, enabling users to obtain reliable healthcare information through both text and voice interactions.

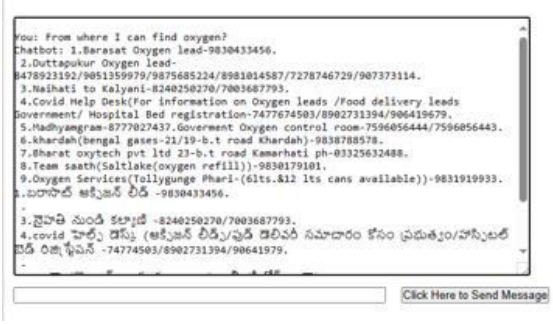

Figure 2: Text-Based Chatbot Response for Oxygen Availability Query

Figure 2 illustrates the output of the proposed LSTM-based medical chatbot when a user submits the query, *"From where can I find oxygen?"* through the text-based conversation interface. After receiving the input, the chatbot analyzes the query using the trained LSTM model and retrieves the most relevant response from the learned medical knowledge base. The system provides a list of oxygen suppliers and emergency contact details in English, followed by the corresponding information translated into Telugu. This demonstrates the chatbot's capability to support multilingual communication and improve accessibility for users with different language preferences. The generated response is displayed instantly within the chat window, enabling users to obtain healthcare-related information without manual searching. The result highlights the effectiveness of the proposed framework in delivering accurate, context-aware, and user-friendly assistance for infectious disease emergencies through an intelligent conversational interface.

DISCUSSION

The proposed LSTM-based medical chatbot integrates deep learning with natural language processing to deliver automated healthcare assistance for infectious disease-related queries. The framework is trained using a publicly available medical question–answer dataset that contains information on symptoms, preventive measures, treatment guidance, emergency support, and other healthcare topics. Before model training, the dataset undergoes several preprocessing operations, including text cleaning, tokenization, vocabulary generation, sequence encoding, and padding to transform conversational data into a format suitable for the LSTM network. The processed data is then divided into training and validation sets, enabling the model to learn meaningful relationships between user questions and corresponding responses. The developed web application combines user authentication, model training, text and voice interaction, multilingual response generation, and conversation history management within a single user-friendly platform, making healthcare information easily accessible to users.

Experimental evaluation demonstrates that the LSTM model effectively learns conversational patterns and generates contextually appropriate responses for a wide range of healthcare-related questions. The trained model achieves approximately 99% training accuracy, indicating its capability to provide reliable responses while maintaining conversational consistency. The integrated speech recognition and text-to-speech modules further enhance user interaction by supporting both voice and text communication, while multilingual response generation in English and Telugu improves accessibility for diverse users. By combining deep learning, speech processing, language translation, and intelligent response generation within a unified framework, the proposed system offers an efficient, scalable, and practical

solution for digital healthcare assistance. The developed chatbot can support users by providing timely preliminary medical information, improving healthcare accessibility, and promoting the adoption of intelligent conversational technologies in modern healthcare environments.

VI.CONCLUSION

This work presents an intelligent LSTM-based medical chatbot developed to provide automated healthcare assistance for infectious disease-related queries through an interactive conversational platform. The proposed system utilizes a publicly available medical question–answer dataset to train the Long Short-Term Memory (LSTM) model, enabling it to understand user queries and generate contextually appropriate responses. The developed application integrates user registration, authentication, dataset preprocessing, model training, text-based and voice-based communication, multilingual response generation, and conversation history management within a single web-based interface. During preprocessing, the conversational dataset is cleaned, tokenized, transformed into numerical sequences, and padded before being supplied to the LSTM network for effective sequence learning. The trained model demonstrates approximately 99% training accuracy, indicating its capability to learn meaningful relationships between healthcare questions and corresponding responses. The chatbot successfully processes both text and voice inputs, generates responses in English and Telugu, and delivers speech output through text-to-speech technology, thereby improving accessibility for a wider range of users. By integrating deep learning, natural language processing, speech recognition, language translation, and intelligent response generation, the proposed framework provides an accurate, scalable, and user-friendly solution for digital healthcare assistance. The developed system can support users by offering timely preliminary medical guidance while reducing the effort required to obtain reliable health information. Future work may focus on expanding the medical knowledge base, incorporating advanced transformer-based language models, supporting multiple regional and international languages, integrating real-time hospital and healthcare services, enabling electronic health record connectivity, and deploying the chatbot on cloud and mobile platforms to further enhance prediction accuracy, personalization, and accessibility in modern healthcare environments.

REFERENCES

[1] J. Cahn, Chatbot: Architecture, Design, Development. 2017.

[2] E. Tebenkov and I. Prokhorov, "Machine learning algorithms for teaching AI chatbots," Procedia Computer Science, vol. 190, pp. 735–744, 2021

[3] A. S. Lokman and M. A. Ameedeen, "Modern chatbot systems: A technical review," in Proceedings of the Future Technologies Conference, pp. 1012–1023, Nov. 2018.

[4] A. Kumar, P. K. Meena, D. Panda, and M. Sangeetha, "Chatbot in Python," International Research Journal of Engineering and Technology (IRJET), vol. 6, no. 11, 2019.

[5] S. Raj and K. Raj, Building Chatbots with Python. New York, NY, USA: Apress, 2019.

[6] H. Koundinya, A. K. Palakurthi, V. Putnala, and K. A. Kumar, "Smart college chatbot using machine learning and Python," in Proceedings of the International Conference on System, Computation, Automation and Networking (ICSCAN), pp. 1–5, Jul. 2020.

[7] S. A. Sheikh, "Artificial Intelligence-Based Chatbot for Human Resource Using Deep Learning," 2019.

[8] M. Hossain, S. Krishna Pillai, S. E. Dansy, and A. A. Bilong, "Mr. Dr. Health: Assistant chatbot," International Journal of Artificial Intelligence, vol. 8, no. 2, pp. 58–73, Dec. 2021.

[9] R. Dharwadkar and N. A. Deshpande, "A medical chatbot," International Journal of Computer Trends and Technology, vol. 60, no. 1, pp. 41–45, 2018.

[10] A. Madhu, C. J. N. Jain, E. Sebastain, S. Shaji, and A. Ajayakumar, "A novel approach for medical assistance using trained chatbot," in Proceedings of the International Conference on Inventive Communication and Computational Technologies (ICICCT), pp. 243–246, Mar. 2017.

[11] Mehfooz, S. Jha, S. Singh, S. Saini, and N. Sharma, "Medical chatbot for novel COVID-19," in ICT Analysis and Applications. Singapore: Springer, pp. 423–430, 2021.

[12] J. Herriman, E. Meer, R. Rosin, V. Lee, V. Washington, and K. G. Volpp, "Asked and answered: Building a chatbot to address COVID-19-related concerns," NEJM Catalyst Innovations in Care Delivery, vol. 1, pp. 1–13, Jun. 2020.

[13] S. Altay, A. S. Hacquin, C. Chevallier, and H. Mercier, "Information delivered by a chatbot has a positive impact on COVID-19 vaccine attitudes and intentions," Journal of Experimental Psychology: Applied, vol. 27, pp. 1–11, Oct. 2021.

[14] S. Amiri and E. Karahanna, "Chatbot use cases in the COVID-19 public health response," Journal of the American Medical Informatics Association, vol. 29, no. 5, pp. 1000–1010, Apr. 2022.

[15] M. Almalki and F. Azeez, "Health chatbots for fighting COVID-19: A scoping review," Acta Informatica Medica, vol. 28, no. 4, p. 241, 2020.

[16] M. Weber and T. Ludwig, "(Non-)interacting with conversational agents: Perceptions and motivations of using chatbots and voice assistants," in Proceedings of the Conference Mensch und Computer, pp. 321–331, Sep. 2020.

[17] U. Bharti, D. Bajaj, H. Batra, S. Lalit, S. Lalit, and A. Gangwani, "Medbot: Conversational artificial intelligence powered chatbot for delivering tele-health after COVID-19," in Proceedings of the 5th International Conference on Communication and Electronics Systems (ICCES), pp. 870–875, Jun. 2020.

[18] N. Battineni, N. Chintalapudi, and F. Amenta, "AI chatbot design during an epidemic like the novel coronavirus," Healthcare, vol. 8, no. 2, pp. 1–8, Jun. 2020.

[19] A. Shawar and E. Atwell, "Using dialogue corpora to train a chatbot," in Proceedings of the Corpus Linguistics Conference, pp. 681–690, Mar. 2003.

[20] M. Ciotti, M. Ciccozzi, A. Terrinoni, W. C. Jiang, C. B. Wang, and S. Bernardini, "The COVID-19 pandemic," Critical Reviews in Clinical Laboratory Sciences, vol. 57, no. 6, pp. 365–388, 2020.

[21] S. J. Daniel, "Education and the COVID-19 pandemic," Prospects, vol. 49, no. 1, pp. 91–96, 2020.

[22] N. Rosruen and T. Samanchuen, "Chatbot utilization for medical consultant system," in Proceedings of the 3rd Technology Innovation Management and Engineering Science International Coference (TIMES-iCON), pp. 1–5, Dec. 2018.

[23] K. Rarhi, A. Bhattacharya, A. Mishra, and K. Mandal, Automated Medical Chatbot. 2017.

[24] S. Majumder and A. Mondal, "Are chatbots really useful for human resource management?" International Journal of Speech Technology, vol. 24, no. 4, pp. 969–977, Dec. 2021.

[25] S. Ashour, A. El-Attar, N. Dey, H. A. El-Kader, and M. M. A. El-Naby, "Long short-term memory based patient-dependent model for FOG detection in Parkinson's disease," Pattern Recognition Letters, vol. 131, pp. 23–29, Mar. 2020.

[26] S. J. Fong, N. Dey, and J. Chaki, "AI-enabled technologies that fight the coronavirus outbreak," in Artificial Intelligence for Coronavirus Outbreak. Singapore: Springer, pp. 23–45, 2021.

[27] S. Chakraborty and L. Dey, "The implementation of AI and AI-empowered imaging systems to fight against COVID-19: A review," in Smart Healthcare System Design: Security and Privacy Aspects, pp. 301–311, 2022.

[28] L. Dey, S. Chakraborty, and A. Mukhopadhyay, "Machine learning techniques for sequence-based prediction of viral-host interactions between SARS-CoV-2 and human proteins," Biomedical Journal, vol. 43, no. 5, pp. 438–450, 2020.

[29] K. Tripathi and A. S. Jalal, "A robust approach based on local feature extraction for age-invariant face recognition," Multimedia Tools and Applications, vol. 81, pp. 21223–21240, Mar. 2022.

[30] A. Sharma and A. S. Jalal, "A framework for visual question answering with the integration of scene-text using PHOCs and Fisher vectors," Expert Systems with Applications, vol. 190, Mar. 2022.