

Medical Chatbot: AI Based Infectious Disease Prediction Model

Srujana Biradar¹, Dr. Swaroopa Shastri²

¹Student, Department of Computer Science & Engineering (MCA), Visvesvaraya Technological University Kalaburagi, Karnataka, India. srujanabiradar25@gmail.com

²Assistant Professor, Department of Computer Science & Engineering (MCA), Visvesvaraya Technological University Kalaburagi, Karnataka, India. swaroopas04@gmail.com

ABSTRACT

A medical chatbot is an artificial intelligence (AI) application specifically designed to interact with users in a conversational manner, providing medical information, assistance, and guidance. These chatbots are built using natural language processing (NLP) and machine learning techniques to understand user queries and respond with appropriate and relevant answers related to healthcare and medical topics. These chatbots are constructed using a combination of NLP and ML techniques to understand user queries and respond with appropriate and relevant answers related to healthcare and medical topics. The intent is to enhance accessibility to medical knowledge through this innovative solution. Computer applications called chatbots which use natural language to communicate with users. To identify keywords in sentences, make query decisions, and provide answers, these chatbots store information in a database. To obtain accurate prediction and effectiveness of the model, To achieve the conversational style, the technology makes advantage of NLP and using SVM as the primary supervised learning algorithm, it is trained on a large dataset containing labeled symptom descriptions and corresponding infectious disease diagnoses, and SVM classifier predicts the most probable infectious disease related to the recognized symptoms and historical data. The model achieved an impressive accuracy of 97.4%. It provides or helps users with immediate access to reliable and relevant medical information.

Keywords: AI, Diseases, Chatbot, Healthcare.

1. INTRODUCTION

An infectious disease is a condition caused by the invasion and growth of harmful microorganisms, such as bacteria, viruses, fungi, or parasites, in the body. These pathogens can be transmitted from one individual to another through various means, including direct contact, airborne particles, contaminated food or water, insect bites, or sexual contact. Disease from mild and self-limiting illnesses, such as the common cold or flu, to severe and life-threatening conditions, such as HIV/AIDS, tuberculosis, or Ebola. Some infectious diseases can be acute, with a rapid onset and short duration, while others may be chronic, persisting for an extended period of time or recurring intermittently. Maintaining public health relies heavily on the vital task of preventing the spread of infectious diseases. Measures such as good hygiene practices (e.g., handwashing), vaccination programs, safe food handling, vector control, and proper sanitation can help reduce the transmission of pathogens. Additionally, timely diagnosis, appropriate treatment, the emphasis on infectious diseases has significantly grown over the last 200 years, extremely due to global outbreaks like COVID-19, underscoring the consequence of developing effective antimicrobial drugs to manage these diseases and prevent their complications, which emerged in late 2019 and continues to impact populations worldwide. These events highlight the importance of surveillance, preparedness, and international cooperation in addressing infectious

diseases and minimizing their impact on individuals and communities. To efficiently manage and prevent infectious diseases, which constitute a serious threat to global health, early detection and rapid action are essential. Medical chatbots that use AI to diagnose and predict infectious diseases represent an attractive new tool for both patients and healthcare professionals. These chatbots can be added to already-existing healthcare platforms or made available as stand-alone programmes that can be accessed by an assortment of gadgets, including smartphones, PCs, and smart speakers. The chatbot serves as a teaching tool by giving current and accurate data about several infectious diseases, their means of transmission, accessible therapies, and prevention strategies. It can support healthy behaviours and assist people in making educated decisions regarding their health. Here we used datasets which is downloaded from kaggle repository. The datasets are symptom-Description which describes the disease, symptom- precaution it predicts the precaution of disease, symptom-severity symptoms with its effect on body per two day. Each dataset with its record symptom- Description with 41 rows 2columns records, symptom-precaution with 41rows 5columns records, symptom-severity with 132 rows 2 columns and overall dataset with 4921 rows 7 columns records.

1.1.1 Methodology

NLP is a field of AI that helps in designing a program to progress and analyse natural language data. It permits to set up communications among PCs and people in a characteristic language. The proposed framework is a talk interface that depends on the Retrieval based model of NLP where the bot is prepared inquiries with a set. Such a wise chatbot can manage the patients by comprehension and surveying their negative impacts that are features of the Proposed System. The working of the system is as follows:

The chatbot will take symptoms as input from the user and provides you with the related searches from the dataset to know the extent of the symptom experienced. Then it will provide users with possible symptoms to answer the question in a yes or no so that it can analyze all the symptoms the user is experiencing and on that basis will perform calculations for predicting the disease that user could most probably have. To be able to distinguish the user's intention or purpose behind their query, helping the chatbot determine the appropriate response and route the user to the relevant module.

SVM Algorithm

A supervised learning system is a SVM and is employed to deal with classification and regression issues. And is used to classify user queries into different intent categories, such as "Get Symptoms," "Diagnose Disease," "Treatment Options," etc. The chatbot can better identify the health issues the user is experiencing and provide relevant information or preliminary advice.

CNN Algorithm

A Convolutional Neural Network can be used for image analysis and diagnosis in a medical chatbot. Due to its efficiency in gathering features from images, CNNs are frequently utilised in tasks associated with computer vision, include medical image analysis. A collection of medical visuals relevant to the chatbot's determination, such as X-rays, MRI scans, CT scans, or pathological images, should be compiled to facilitate divergence during training, reduce and normalise the pixel values of the medical images as a component in the preprocessing phase. The system can benefit from increased diagnostic features for medical image processing, giving users relevant information and advice about their health issues.

2. LITERATURE SURVEY

- [1] F. Mehfooz, S. Jha, S. Singh, S. Saini, and N. Sharma, "Medical chatbot for novel COVID- 19," in *ICT Analysis and Applications*. Singapore: Springer, 2021, pp. 423–430. The prime focus of a paper is to show implementation of a retrieval-based chatbot with voice support, and we will investigate other standing chatbot and in what way it is useful in helping the patients fetching all the necessary details about COVID-19. The primary goal of this work is to determine the creation of a retrieval-based chatbot with voice support. We will explore different existing chatbots and their potential to aid patients in gathering comprehensive information about COVID-19.
- [2] M. 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 Innov. Care Del.*, vol. 1, pp. 1–13, Jun. 2020. Penn Medicine's administrators saw the need for institution-specific replies to respond the unique information patients sought for when they launched a Covid-19 chatbot to aid patients. It is crucial to modify the answers so that they are in line with the regional capabilities and paths of the health system upon which the initiative is based, even when utilizing standard responses for symptom checker questions. Leaders worked with skilled ML and NLP companies to build a specialized chatbot, which reached to people in few days after the initial planning meeting. After comprehending the need for distinct response mapping, delicate contextualization, and dynamic, human-guided content validation.
- [3] S. Altay, A. S. Hacquin, C. Chevallier, and H. Mercier, "Information delivered by a chatbot has a Positive Impact on COVID-19 vaccines attitudes and intentions," *J. Exp. Psychol., Appl.*, vol. 27, pp. 1–11, Oct. 2021. The Coronavirus disease; COVID-19 vaccines will not end the pandemic if they stay in freezers. In many countries, such as France, COVID-19 vaccines hesitancy is high. It is crucial that governments approach is to facilitate easy access to vaccinations for those willing to receive them while simultaneously creating communication strategies that effectively address the concerns for individuals who are reluctant to get vaccinated.
- [4] P. Amiri and E. Karahanna, "Chatbot use cases in the Covid-19 public health response," *J. Amer. Med. Inform. Assoc.*, vol. 29, no. 5, pp. 1000–1010, Apr. 2022. A total of 3334 items had been found after our search. After the search the certainty that 61 of them indicated the acceptance of 61 deployed chatbots in 30 various nations meant that they all corresponded our inclusion criterion. Six forms of public health response use cases were identified in our study, which organized these chatbots relative to their design and public health response use case(s). These categories included 15 distinct use cases.
- [5] M. Almalki and F. Azeez, "Health chatbots for Fighting COVID-19: A scoping review," *Acta Inf. Medica*, vol. 28, no. 4, p. 241, 2020. This study aims to review the current literature on COVID-19 related chatbots in healthcare, identify and characterize these emerging technologies and their applications for combating COVID-19, and describe related challenges.
- [6] P. Weber and T. Ludwig, "(Non-)Interacting with Conversational Agents: Perceptions and Motivations of Using Chatbots and Voice Assistants," in *Proc. Conf. Mensch Comput.*, vol. 1, Sep. 2020, pp. 321–331. Because voice-controlled assistants like Siri, Alexa, and Google Assistant are forming more prevalent in daily life, it is problematic to design them from the standpoint of Human-Computer Interaction (HCI). Although there are fundamental design principles for human-AI interaction, little is known about how users see conversational agents (CAs) in their day-to-day lives or what inspires them to utilize such tools.

[7] U.Bharti, D.Bajaj, Hunar Batra, S.Lalit, S.Lalit, A.Gangwani, “Medbot: Conversational artificial intelligence powered Chatbot for delivering telehealth after COVID-19,” in Proc. 5th Int. Conf. Commun. Electron. Syst. (ICCES), Jun. 2020, pp. 870–875. During the recent Coronavirus outbreak, medical professionals may interact with patients by telemedicine, while attempting to reduce COVID-19 transmission among patients and clinicians. Amidst the pandemic, Telemedicine has the potential to help by permitting patients to receive supportive care without having to physically visit a hospital by using a conversational artificial intelligence-based application for their treatment.

[8] G. 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. It exhibits the creation of a sophisticated chatbot with artificial intelligence that can assess patients exposed to nCOV- 19 and offer assistance right away. Additionally, virtual assistants can determine the infection severity and connect patients with registered doctors when symptoms are acute.

[9] B. A. Shawar and E. Atwell, “Using dialogue corpora to train a chatbot,” in Proc. Corpus Linguistics Conf., Mar. 2003, pp. 681–690. The paper presents two chatbot systems, ALICE and Elizabeth, illustrating the dialogue knowledge representation and pattern matching techniques of each.

[10] M. Ciotti, M. Ciccozzi, A. Terrinoni, W. C. Jiang, C. B. Wang, and S. Bernardini, “The COVID-19 pandemic,” *Crit. Rev. Clin. Lab. Sci.*, vol. 57, no. 6, pp. 365–388, 2020. We focus on the available therapies to fight COVID-19, the increased development of vaccines, the role of artificial intelligence in the management of the pandemic and limiting the spread of the virus, the impact of the COVID-19 epidemic on our lifestyle, and preparation for a possible next second wave.

2.1 EXISTING AND PROPOSED SYSTEM

2.1.1 Existing system

In this existing system, implemented an AI Chat bot interaction and prediction model using a deep feed forward multi-layer perceptron.

Disadvantages

Feed forward systems can't be accurate without an approximate process model. They can't process or consider unmeasured variables. These systems might struggle to keep up with too many individual variables present. One still has to use feedback controls to provide a proper backup.

2.1.2 Proposed system

The feature learning process in CNNs is mainly automated from the input images, which alleviates the effort of extracting multiple features for traditional sign recognition classifiers. This was the impetus for the creation of the deep CNN model for recognising Indian traditional dance. The learned characteristics are viewed in a higher level abstract representation of low level sign pictures due to the deep architecture. As a result, we develop a deep CNN model for dance identification in this study. The variety of poses available is to be increased. The findings are contrasted to the results of other conventional state-of-the-art techniques like as SVM to emphasise CNN's recognition abilities.

Advantages

- Instant Answers
- Programmability
- Reduce care costs
- Reducing waiting time
- Improve patient’s satisfaction

3. SYSTEM DESIGN

3.1 SYSTEM PRESPECTIVE

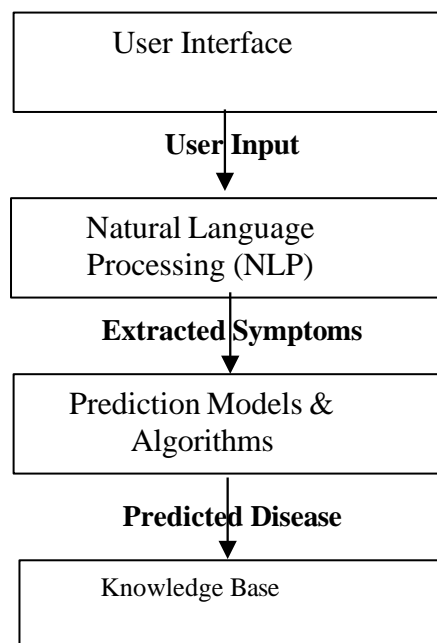


Figure 1: Methodology

User Interface: The user interface is the front-end of the medical chatbot where users interact with the system. Which could be a web-based interface, a mobile app, or embedded into communication platforms such as Fb Messenger or WhatsApp. The UI allows users to input their queries, ask questions, and receive responses from the chatbot.

Natural Language Processing(NLP): NLP is a vital aspect of the medical chatbot system. It enables the chatbot to be understand and process human language, making it capable of interpreting user inputs and generating appropriate responses. NLP approach, like intent recognition, named entity recognition, and sentiment analysis are strategies used for figuring out relevant information from user searches.

Prediction Models and Algorithms: The medical chatbot may incorporate various prediction models and algorithms to enhance its functionalities. These models can include disease prediction models, treatment recommendation algorithms, drug interaction models, and more. Prediction models help the chatbot offer

personalized healthcare insights and responses based on user-specific information and medical history

Knowledge Base: The knowledge base is a repository of medical information and data. It contains a deep range of medical concepts, including symptoms, diseases, treatments, medications, and preventive measures. The knowledge base is curated from reputable medical sources, clinical guidelines, research papers, and expert databases, ensuring the accuracy and reliability information by the chatbot.

4. IMPLEMENTATION

A number of steps must be taken in completing the creation of an AI-based medical chatbot model for infectious disease prediction. Here is a summary of the method:

Data collection: Compile precise and relevant data on infectious diseases, include symptoms, risk factors, as well as past patterns of illness. Access to medical documents, research studies, public health databases, and online medical records can be essential.

Extraction of Features: Identify the most essential characteristics from the preprocessed data. In concentrate on the elements that are most beneficial to detecting infectious diseases, this process assists in decreasing the complexity of the data.

Model development: Apply deep learning or machine learning to develop an AI model. You can choose from an assortment of techniques, such as decision trees, random forests, support vector machines (SVM), or neural networks, depending on the variety of data available and the complexity of the assignment.

Evaluation and Validation: To evaluate your model's performance, separate the data set into testing and training sets. To assess the model's estimation ability, employ suitable evaluation metrics like accuracy, precision, recall, or F1-score. The model's parameters have to be adjusted and modified as appropriate.

Data Preprocessing: Preprocessing the collected data comprises organising and preparing it. In this step, the data are cleaned up, missing values are addressed, data formats are organised, and the data is transformed into a machine-learning-ready format.

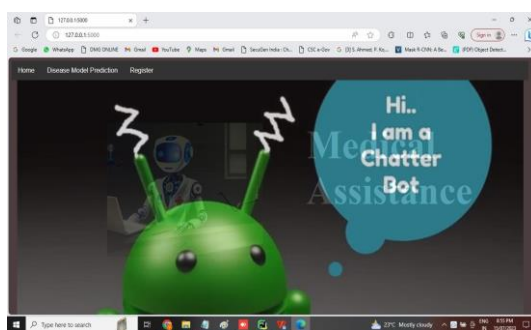


Figure 2: Main page

The navigation menu enables users to access various parts of the website links to the "Home" page, the "Disease Model Prediction" page, and the "Register" page are present in this instance. The header section contains the logo and name of the medical chatbot, providing branding and easy identification.



Figure 3: User Register

The title of the web page, indicating that this is the registration page for new users. Name, A text input field where the user can enter their full name. Email ID, A text input field where the user can enter their email address. Contact, A text input field where the user can enter their contact number or phone number. Username, A text input field where the user can choose a unique username for their chatbot account. Password, A password input field where the user can enter a secure password for their account. The characters will be masked for security.

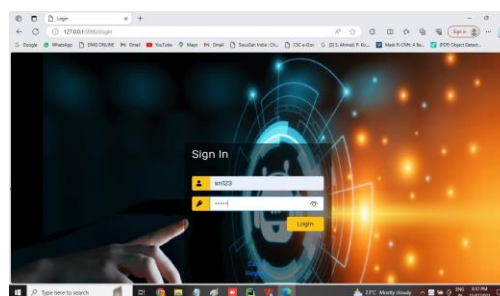


Figure 4: User Login

The title of the web page, indicating that this is the login page for existing users to access their accounts. Username: A text input field where the user can enter their registered username or email ID. Password: A password input field where the user can enter their account password. The characters will be masked for security.

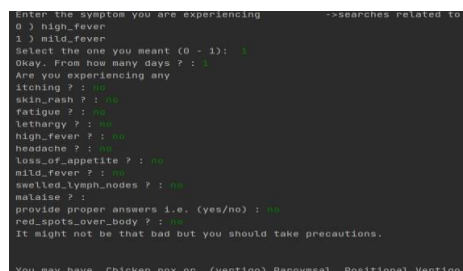


Figure 5: Symptoms

The title of the web page, indicating that this page is for users to input their symptoms. Text input fields where the user can enter their symptoms. There are many different fields provided for users to input up to different symptoms they are experiencing and the user should answer in yes or no format for the symptoms asked. For ex, experiencing fever the system ask for the symptom you are experiencing and you provided fever the very next it ask for high fever or mild fever and the datasets are distributed in training and testing where 0 and 1 are used for choosing options after selecting the type of fever it takes for different symptoms related to that.

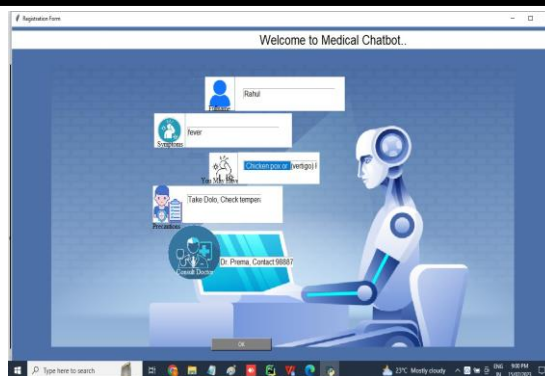


Figure 6: Chatbot GUI

The GUI starts with the chatbot greeting the user and asking for their full name. The user can enter their full name in the provided text field. The chatbot proceeds to ask the user to input their symptoms. The user can type their symptoms in the text field provided. The chatbot responds by suggesting possible symptoms based on the user's input after verifying the categories of that particular symptom. The chatbot then provides precautions based on the selected symptoms to guide the user on what to do next and also suggests to consult doctor with their and contact number.

Test Number	Form	Description	Expected Result	Actual Result
1	Unit Test	Verify NLP processing of user query	Relevant information is extracted	Relevant information is extracted
2	Unit Test	Test prediction model accuracy	Accurate disease prediction is made	Accurate disease prediction is made
3	Unit Test	Validate preprocessing techniques	Input data is cleaned and transformed	Input data is cleaned and transformed
4	Integration Test	Verify seamless integration of NLP, prediction models, and data sources	User queries are correctly routed and processed	User queries are correctly routed and processed
5	Validation Test	Evaluate chatbot's performance with diverse user scenarios	Accurate disease predictions and appropriate recommendations are provided	Accurate disease predictions and appropriate recommendations are provided

Fig 7: Table of test Cases

RESULTS

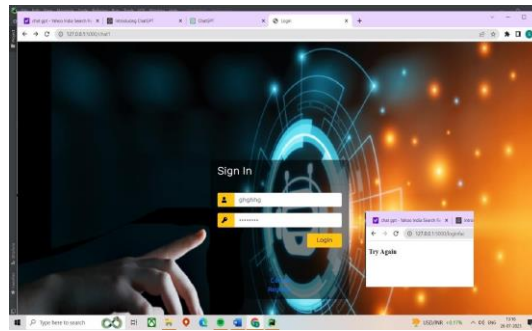


Fig 8: Incorrect password or Username

The title of the web page, indicating that the login credentials provided by the user were incorrect. The message requesting the user to try again after notifying them that their username or password had been entered incorrectly. It displays try again with correct username and password to the user at new page after entering incorrect password.

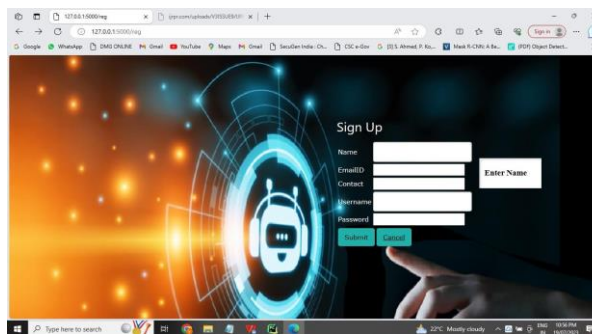


Figure 9: Blank field validation for name

The title of the web page, indicating that this page validates the name field for blank entries. It is kind of error, the user receives a message informing them that they must provide their full name for them to proceed and a text input field is displayed in the text box beside the name which indicates Enter Name.

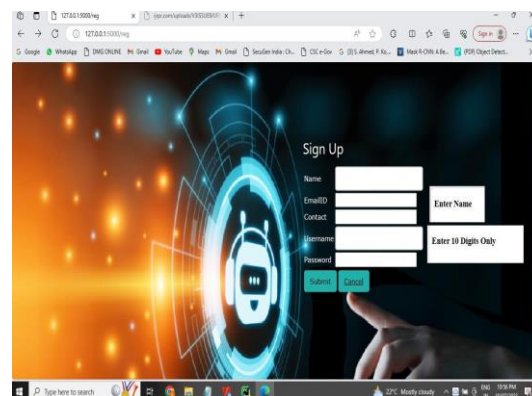


Figure 10: Contact digit validation

The title of the web page, indicating that this page validates the contact number for exactly 10 digits it displays error if the contact number contains more than 10 digits in the text input. It shows to Enter 10 Digits Only beside the contact text field.

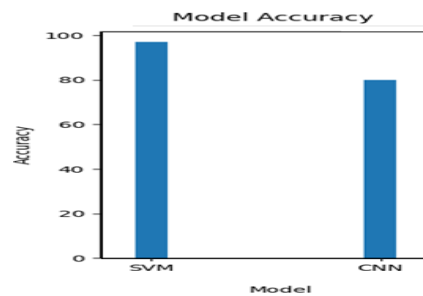


Figure 11: Model accuracy graph

SVM is a supervised machine learning algorithm which has a 97% accuracy percentage, which implies that 97% of the test data were accurately predicted by the SVM model. This shows that SVM was successful in classifying the data into the appropriate classes and is able to predict results with accuracy. CNN is a deep learning algorithm has a 81% accuracy percentage. In terms of total classification accuracy, the SVM performed better than the CNN. SVM and CNN's results may vary in various cases or when used with different datasets.

CONCLUSION

In conclusion, an AI-based medical chatbot model for infectious disease prediction can be a valuable tool in healthcare. Such a chatbot can assist in early detection and prediction of infectious diseases, enabling proactive measures to be taken to prevent their spread and mitigate their impact on public health.

By leveraging machine learning algorithms and natural language processing capabilities, the chatbot can analyze user input, symptoms, and relevant data to provide accurate predictions and recommendations. It can utilize historical data, epidemiological trends, and real-time information to assess probability of a user being affected by a particular infectious disease.

The benefits of an AI-based medical chatbot for infectious disease prediction include:

Early detection: The chatbot can identify potential cases of infectious diseases at an early stage by analyzing user-reported symptoms and risk factors. This allows for prompt medical intervention and preventive measures to limit the spread of the disease.

FUTURE ENHANCEMENT

Rapid response: The chatbot can provide real-time information and guidance to users, enabling them to take appropriate actions quickly. It can offer advice on self-isolation, seeking medical attention, or following specific preventive measures depended for the predicted risk level.

Accessibility: A chatbot can be easily accessible to a more in number of users, presenting them with precise data and personalized recommendations. It can be available 24/7, eliminating requirement for individuals to wait for a healthcare professional's availability or visit a medical facility unnecessarily.

Public health monitoring: Aggregated and anonymized data collected by the chatbot can contribute to monitoring infectious disease patterns on a broader scale. Health representatives can use this data to identify

hotspots, track outbreaks, and allocate resources more effectively. However, it's important to note that an AI-based medical chatbot model for infectious disease prediction should not be considered a substitute for certified health advice or a diagnosis. It should be designed to work in collaboration with healthcare providers and serve as a complementary tool in the healthcare ecosystem.

Furthermore, the accuracy and reliability of the chatbot's predictions depend on the quality of the data it is trained on. Regular updates and continuous training of the AI model with up-to-date medical information and research findings are crucial to maintaining its effectiveness.

In summary, an AI-based medical chatbot model for infectious disease prediction has the potential to enhance public health surveillance, provide timely guidance to individuals, and contribute to proactive measures in controlling the spread of infectious diseases. It can serve as a valuable resource in the healthcare industry, supporting healthcare professionals and empowering individuals empowering them to make decisions on their own preferences.

BIBLIOGRAPHY

- 1) F. Mehfooz, S. Jha, S. Singh, S. Saini, and N. Sharma, "Medical chatbot for novel COVID- 19," in *ICT Analysis and Applications*. Singapore: Springer, 2021, pp. 423–430.
- 2) M. 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 Innov. Care Del.*, vol. 1, pp. 1–13, Jun. 2020.
- 3) S. Altay, A. S. Hacquin, C. Chevallier, and H. Mercier, "Information delivered by a chatbot has a positive impact on COVID-19 vaccines attitudes and intentions," *J. Exp. Psychol., Appl.*, vol. 27, pp. 1–11, Oct. 2021
- 4) P. Amiri and E. Karahanna, "Chatbot use cases in the COVID-19 public health response," *J. Amer. Med. Inform. Assoc.*, vol. 29, no. 5, pp. 1000–1010, Apr. 2022.
- 5) M. Almalki and F. Azeez, "Health chatbots for fighting COVID-19: A scoping review," *Acta Inf. Medica*, vol. 28, no. 4, p. 241, 2020.
- 6) P. Weber and T. Ludwig, "(Non-) interacting with conversational agents: Perceptions and motivations of using chatbots and voice assistants," in *Proc. Conf. Mensch Comput.*, vol. 1, Sep. 2020, pp. 321–331.
- 7) U. Bharti, D. Bajaj, H. Batra, S. Lalit, S. Lalit, and A. Gangwani, "Medbot: Conversational artificial intelligence powered chatbot for delivering telehealth after COVID-19," in *Proc. 5th Int. Conf. Commun. Electron. Syst. (ICCES)*, Jun. 2020, pp. 870–875.
- 8) M. Ciotti, M. Ciccozzi, A. Terrinoni, W. C. Jiang, C. B. Wang, and S. Bernardini, "The COVID-19 pandemic," *Crit. Rev. Clin. Lab. Sci.*, vol. 57, no. 6, pp. 365–388, 2020.
- 9) Priyanka Kulkarni, & Dr. Swaroopa Shastri. (2024). Rice Leaf Diseases Detection Using Machine Learning. *Journal of Scientific Research and Technology*, 2(1), 17–22. <https://doi.org/10.61808/jsrt81>
- 10) Shilpa Patil. (2023). Security for Electronic Health Record Based on Attribute using Block-Chain Technology. *Journal of Scientific Research and Technology*, 1(6), 145–155. <https://doi.org/10.5281/zenodo.8330325>
- 11) Mohammed Maaz, Md Akif Ahmed, Md Maqsood, & Dr Shridevi Soma. (2023). Development Of Service Deployment Models In Private Cloud. *Journal of Scientific Research and Technology*, 1(9), 1–12. <https://doi.org/10.61808/jsrt74>
- 12) Antariksh Sharma, Prof. Vibhakar Mansotra, & Kuljeet Singh. (2023). Detection of Mirai Botnet Attacks on IoT devices Using Deep Learning. *Journal of Scientific Research and Technology*, 1(6), 174–187.
- 13) Dr. Megha Rani Raigonda, & Shweta. (2024). Signature Verification System Using SSIM In Image Processing.

Journal of Scientific Research and Technology, 2(1), 5–11. <https://doi.org/10.61808/jsrt79>

- 14) Shri Udayshankar B, Veeraj R Singh, Sampras P, & Aryan Dhage. (2023). Fake Job Post Prediction Using Data Mining. Journal of Scientific Research and Technology, 1(2), 39–47.
- 15) Gaurav Prajapati, Avinash, Lav Kumar, & Smt. Rekha S Patil. (2023). Road Accident Prediction Using Machine Learning. Journal of Scientific Research and Technology, 1(2), 48–59.
- 16) Dr. Rekha Patil, Vidya Kumar Katrabad, Mahantappa, & Sunil Kumar. (2023). Image Classification Using CNN Model Based on Deep Learning. Journal of Scientific Research and Technology, 1(2), 60–71.
- 17) Ambresh Bhadrashetty, & Surekha Patil. (2024). Movie Success and Rating Prediction Using Data Mining. Journal of Scientific Research and Technology, 2(1), 1–4. <https://doi.org/10.61808/jsrt78>
- 18) Dr. Megha Rani Raigonda, & Shweta. (2024). Signature Verification System Using SSIM In Image Processing. Journal of Scientific Research and Technology, 2(1), 5–11. <https://doi.org/10.61808/jsrt79>