An AI-Based Medical Chatbot Model for Infectious Disease Prediction

Sheri RamachandraReddy[1], Rajesh Banothu[2], Harshita Kusunuri[3], and Nithya Reddy Kura[4]
[1]*Assistant Professor, Department of CSE (AI & ML), CMR Technical Campus, Hyderabad, TG, India*
[2,3,4]*Students, Department of CSE (AI & ML), CMR Technical Campus, Hyderabad, TG, India*

***Abstract*—An AI-Based Medical Chatbot Model for Infectious Disease Prediction is proposed to support early identification of diseases and improve access to basic healthcare services. The system leverages artificial intelligence, machine learning, and natural language processing to interact with users, collect symptom information, and analyze health data in real time. By providing preliminary risk assessment and health guidance, the chatbot assists users in understanding possible infectious conditions without immediate dependence on healthcare professionals. The chatbot integrates a trained predictive model capable of identifying patterns associated with common infectious diseases based on user-reported symptoms. Through a conversational interface, the system delivers accurate, timely, and personalized responses while maintaining ease of use. This approach helps reduce the burden on healthcare systems, minimizes delays in diagnosis, and supports informed decision-making for patients. The proposed model aligns with Sustainable Development Goal 3 (Good Health and Well Being) by promoting early detection, preventive care, and healthcare accessibility, particularly in remote and underserved areas. Overall, the AI-based medical chatbot demonstrates potential as a scalable and cost-effective solution for enhancing public health monitoring and improving overall healthcare outcomes.**

***Index Terms*—Artificial Intelligence (AI), Disease Prediction, Electronic Health Records (EHR), Healthcare Informatics, Infectious Disease Prediction, Long Short-Term Memory (LSTM), Machine Learning (ML), Medical Chatbot, Natural Language Processing (NLP), Symptom Analysis].**

I. INTRODUCTION

The scope of this project includes the design and development of an AI-based medical chat bot capable of predicting infectious diseases based on user-reported symptoms. The system focuses on providing preliminary health assessment, basic medical guidance, and recommendations for further consultation. It is intended to act as a supportive tool rather than a replacement for professional medical diagnosis, ensuring ethical and responsible use of artificial intelligence in healthcare [1].

The project covers the implementation of machine learning models, particularly for symptom-based disease prediction, along with natural language processing techniques to enable effective user interaction. The chatbot is designed to handle multiple user queries simultaneously and provide real-time responses. Integration with a backend database allows secure storage of medical datasets, user interactions, and prediction results for analysis and future model improvement [2].

The scope also includes the development f a user-friendly web or mobile interface that ensures accessibility for users with varying levels of technical expertise. The system supports features such as symptom checking, health tips, and emergency guidance, making it useful for daily health monitoring. Emphasis is placed on usability, response speed, and clarity of information delivered to the user [3].

Finally, the project scope extends to system evaluation and performance analysis using standard metrics to measure prediction accuracy and response efficiency. Security and data privacy considerations are addressed to protect user information. The system is scalable and can be enhanced in the future to include additional diseases, advanced AI models, and integration with external health data sources, aligning with broader public health objectives [4].

II. RELATED WORK

PROJECT PURPOSE: The scope of this project focuses on the development of an AI-based medical chat bot that assists users in the early prediction of infectious diseases through symptom analysis. The system is designed to collect user inputs in natural language, process the musing machine learning and natural language processing techniques, and provide preliminary health assessments and guidance. The chatbot acts as a decision-support tool and does not replace professional medical diagnosis [5].

The project includes the implementation of a predictive model for identifying common infectious diseases, along with a conversational interface for real-time interaction. A backend database is used to securely store medical datasets, user interactions, and prediction results. The system emphasizes scalability, fast response time, and ease of use to ensure accessibility for users with different levels of technical knowledge [6].

Additionally, the project scope covers system evaluation, performance analysis, and security considerations to ensure reliability and data privacy. The architecture is designed to support future enhancements such as inclusion of more diseases, improved prediction algorithms, and integration with external healthcare systems, contributing to better public health support and alignment with SDG-3 objectives [7].

Project Features: The proposed system features an intelligent conversational chatbot that interacts with users to collect symptoms and health-related information in a simple and user-friendly manner. Using natural language processing, the chatbot understands user inputs and converts them into structured data for analysis. This allows users to receive instant responses, making the system suitable for real-time health assistance and preliminary disease screening. Another key feature is the infectious disease prediction module powered by machine learning algorithms [8].

The model analyzes symptom patterns and predicts the likelihood of specific infectious diseases with reasonable accuracy. Based on the prediction results, the system provides preventive measures, basic medical advice, and recommendations on whether to seek professional healthcare support. The system also includes secure data handling and scalability features to support multiple users' simultaneously. User interactions and prediction results are stored safely for analysis and future improvements. The modular architecture allows easy integration of new diseases, advanced models, and additional healthcare services, making the system adaptable to evolving medical needs [9].

III. SYSTEMA NALYSIS

The system analysis evaluates the functionality, performance, and feasibility of the AI-based medical chatbot for infectious disease prediction. The system is designed to efficiently collect user symptoms, process inputs using natural language processing, and generate disease predictions through trained machine learning models. Analysis shows that the chatbot provides fast and consistent responses while maintaining usability and accessibility. The modular architecture supports scalability, secure data handling, and future enhancements, making the system reliable for early disease prediction and supportive healthcare decision making [10].

Problem Definition: The problem addressed in this project is the lack of timely and accessible medical assistance for early detection of infectious diseases, especially in regions with limited healthcare resources. Delayed diagnosis, shortage of medical professionals, and overburdened healthcare systems often lead to increased disease spread and poor health outcomes. There is a need for an intelligent system that can analyze user-reported symptoms, provide preliminary disease prediction, and offer reliable health guidance to support early intervention and improve public health management.

Existing System: In the existing health care system, infectious disease diagnosis primarily depends on in person consultations, laboratory tests, and clinical evaluations conducted by medical professionals. Patients must visit hospitals or clinics to report symptoms, undergo tests, and wait for results, which often lead to delays in diagnosis and treatment. This approach is effective but time-consuming, costly, and difficult to access for individuals living in remote or underserved areas.

Some digital health platforms and online symptom checkers have been introduced to provide basic medical information. These systems typically rely on rule-based logic or static questionnaires to match

symptoms with possible diseases. While they improve accessibility, they often lack accuracy, personalization, and real-time predictive capabilities, limiting their usefulness for early infectious disease detection. Additionally, existing systems place a heavy workload on healthcare professionals, especially during disease outbreaks, leading to overcrowded facilities and reduced quality of care. The absence of intelligent, automated tools for preliminary screening and risk assessment highlights the limitations of current approaches and underscores the need for more advanced AI-driven healthcare solutions. The existing approach to infectious disease detection mainly relies on traditional clinical methods, where patients consult healthcare professionals for symptom evaluation and diagnosis. Doctors assess symptoms through physical examination and recommend laboratory tests to confirm diseases. While this method is medically reliable, it is highly dependent on the availability of doctors, diagnostic facilities, and time, making it inefficient during large-scale outbreaks or in areas with limited healthcare infrastructure. Another common approach involves online health portals and basic symptom checker applications that use predefined rules or decision trees. These systems ask users a fixed set of questions and provide generalized health suggestions based on symptom matching. Although they improve accessibility, they lack intelligent learning capabilities and cannot adapt or improve predictions based on new data, reducing accuracy and effectiveness. Some existing systems also depend on manual data analysis and public health surveillance reports to monitor disease trends. These approaches are often slow, reactive, and unable to provide personalized health guidance. The lack of real-time prediction, automation, and intelligent interaction in current approaches limits their ability to support early disease detection and proactive healthcare management.

Limitations of existing system:

- Delayed Diagnosis: Traditional systems rely on physical consultations and laboratory tests, which consume time and often delay early detection of infectious diseases.
- Limited Accessibility: Existing systems are not easily accessible to people in remote or underserved areas due to the shortage of healthcare professionals and facilities.
- Lack of Intelligence and Personalization: Most current digital tools are rule-based and do not use advanced AI models, resulting in low accuracy and generalized, non-personalized health guidance.

IV. PROPOSED SYSTEM

The proposed system introduces an AI-based medical chatbot designed to assist in the early prediction of infectious diseases through intelligent symptom analysis. The chatbot interacts with users in natural language, collects symptom details, and processes the information using natural language processing techniques. This enables users to receive quick and preliminary health assessments without the need for immediate physical consultation. The system integrates a machine learning–based disease prediction model trained on infectious disease datasets to identify patterns between symptoms and diseases. Based on the prediction results, the chatbot provides preventive measures, basic medical advice, and recommendations for professional medical consultation when required. This approach enhances accuracy, reduces diagnostic delays, and supports informed health decisions. The proposed system is accessible, scalable, and available 24/7, making it suitable for users in remote and underserved areas. By reducing the workload on healthcare professionals and promoting early detection, the system contributes to improved healthcare delivery and aligns with Sustainable Development. The proposed approach begins with the collection of reliable infectious disease datasets that include symptoms, patient attributes, and confirmed disease outcomes. These datasets are preprocessed using data cleaning, normalization, and feature selection techniques to ensure high-quality input for model training. Proper preprocessing improves prediction accuracy and reduces noise in the data. Next, suitable machine learning algorithms are selected to build the disease prediction model. The model is trained to identify relationships between symptoms and infectious diseases and is evaluated using standard performance metrics such as accuracy, precision, and recall. After validation, the trained model is prepared for real-time deployment within the chatbot system.

The third phase focuses on developing the conversational chatbot using natural language processing techniques. The chatbot is designed to

understand user queries, extract relevant symptom information, and convert unstructured text into structured data that can be processed by the prediction model. This ensures smooth and effective interaction between the user and the system. Finally, the prediction results are delivered to users through the chatbot in a clear and understandable format. The system provides disease risk assessment, preventive guidelines, and recommendations for medical consultation. This integrated approach ensures early detection, improved healthcare accessibility, and efficient support for public health management.

Advantages of the proposed system:

The proposed system implemented using the machine learning techniques; the proposed system is processing in the following way.

- Early Disease Prediction: Enables early identification of infectious diseases through real-time symptom analysis, reducing delays in diagnosis.
- 24/7 Accessibility: Provides continuous health care assistance without time or location constraints.
- Reduced Healthcare Work load: Automates preliminary screening and guidance, lowering the burden on doctors and hospitals.
- Improved Accuracy and Personalization: Uses machine learning models to deliver more accurate and personalized health insights compared to rule-based systems.
- Cost-Effective and Scalable: Supports a large number of users simultaneously while minimizing healthcare costs through automation.

V. SYATEM ARCHITECTURE

Architecture:

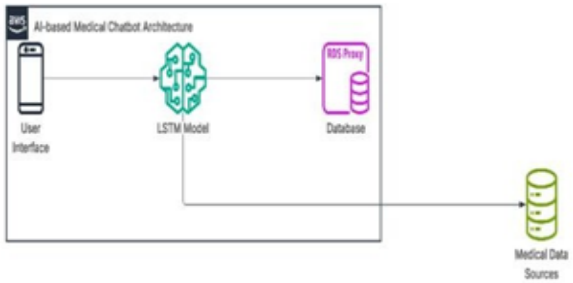

Fig: Project Architecture of an AI – Based Medical Chatbot for Infectious disease prediction

The system architecture consists of a user interface, an LSTM-based prediction model, a database, and external medical data sources. Users interact with the chatbot through a web or mobile interface, where symptoms are submitted and forwarded to the LSTM model for analysis. The model processes the input, accesses stored medical data through the database, and generates disease prediction results, which are then returned to the user in real time. The AI-based medical chat bot system is designed to provide intelligent assistance for infectious disease prediction through an integrated and modular architecture. The system enables users to interact with the chatbot via a web or mobile-based user interface, where they can enter symptoms and health-related queries in natural language. This interface ensures easy accessibility and smooth interaction for users with varying levels of technical knowledge. Once the user inputs are received, they are forwarded to the core prediction module, which is built using an LSTM (Long Short-Term Memory) model. The LSTM model analyzes symptom sequences and patterns to predict the likelihood of infectious diseases. This deep learning approach allows the system to handle complex symptom relationships and improves prediction accuracy compared to traditional rule-based methods.

The system is connected to a backend data base through an RDS proxy, which securely stores medical data sets, user interactions, and prediction results. The data base plays a crucial role in supporting model training, validation, and real-time inference. It also enables efficient data management, scalability, and secure access to sensitive health information. Additionally, the system integrates external medical data sources to enhance prediction reliability and keep the model updated with relevant health information. The prediction results generated by the model are sent back to the chatbot, which presents them to the user along with preventive measures and medical guidance. This end-to-end architecture ensures fast response, reliability, and effective support for early infectious disease detection.

Use case diagram:

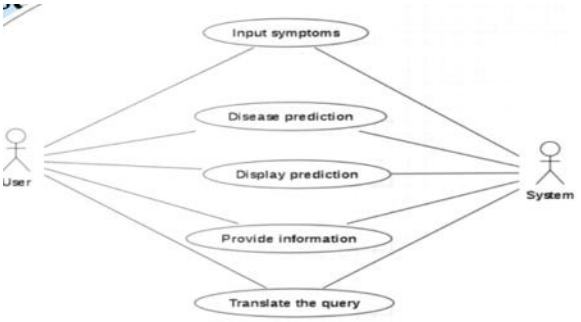

Fig: Use Case Diagram

Description:

The diagram illustrates how a user interacts with an AI-based medical chatbot to obtain health-related insights. The process begins when the user inputs symptoms into the system through the interface. This input acts as the primary trigger for all subsequent operations, making accurate symptom entry critical. If the input is vague or incorrect, the system's prediction quality immediately degrades this dependency is a core weakness you should acknowledge. Once symptoms are submitted, the system processes them using a disease prediction module that analyzes patterns against medical data. The predicted result is then generated and prepared for presentation. This step represents the system's intelligence layer, where accuracy depends heavily on the quality of the trained model and data sources. Poor training data or outdated medical knowledge will lead to misleading outputs, which is a real risk in healthcare systems. After prediction, the system displays the results and provides relevant medical information such as disease details, precautions, or next steps. The translation feature allows users to interact indifferent languages, improving accessibility but also increasing complexity and error potential. Overall, the workflow emphasizes user convenience and automation, but it must be backed by rigorous validation to avoid over-reliance on automated medical advice.

Class diagram:

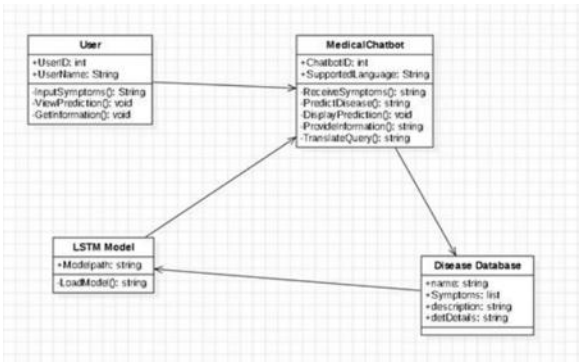

Fig: Class Diagram

Description:

This class diagram defines the structural backbone of the AI-based medical chatbot system. The User class represents the end user and handles basic interactions such as entering symptoms, viewing predictions, and requesting information. This separation is correct, but minimal there's no validation or error-handling responsibility shown here, which means the design assumes users always provide clean, usable input. That's unrealistic in real medical systems and should be acknowledged as a limitation. The Medical Chatbot class acts as the core controller of the system, coordinating symptom intake, disease prediction, information display, and language translation. It directly interacts with both the LSTM Model and the Disease Database, making it the central dependency hub. While this simplifies flow, it also creates tight coupling if the chatbot fails, everything fails. A more robust design would separate prediction logic and data access into independent services. The LSTM Model class is responsible for loading and executing the trained prediction model, while the Disease Database stores disease names, symptoms, descriptions, and details. Their interaction enables data-driven predictions and informative responses. However, the diagram hides critical realities: model retraining, data updates, and version control are completely absent. Ignoring these aspects makes the system look simpler than it actually is and weakens its real-world credibility.

Sequence diagram:

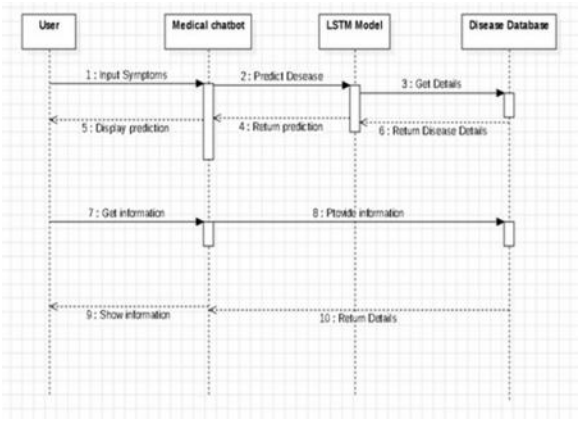

Fig: Sequence Diagram

Description:

This sequence diagram illustrates the complete interaction flow between the User, Medical Chatbot, LSTM Model, and Disease Database during disease prediction. The process begins when the user inputs symptoms into the system. These symptoms are immediately forwarded to the medical chatbot, which acts as the central coordinator for all further operations. After receiving the symptoms, the medical chatbot requests the LSTM model to predict the disease. The LSTM model processes the input and queries the disease database to fetch relevant disease details. Once the database returns the required information, the LSTM model sends the prediction results back to the medical chatbot, ensuring that both prediction and supporting details are available. Finally, the medical chatbot displays the predicted disease to the user. If the user requests additional information, the chatbot again communicates with the disease database to retrieve detailed descriptions and guidance. This information is then returned and shown to the user, completing a structured, step-by-step interaction cycle that ensures accurate prediction and informative responses.

VI. IMPLEMENTATION

React.js is used to develop the front-end of the medical chatbot system, providing an interactive and responsive user interface. It enables users to input symptoms, view predicted diseases, and access medical information in real time without page reloads. By using reusable components and state management, React.js ensures smooth user interactions and a consistent experience across devices. The component-based architecture of React.js simplifies UI development and maintenance. It allows the chatbot interface; symptom input forms, and result display sections to be managed independently, improving scalability and performance. This makes React.js a practical choice for building a dynamic and user-friendly medical chatbot application.

Express.js is used as the backend framework to handle server-side logic and API management for the medical chatbot system. It processes requests sent from the React.js frontend, manages user symptom data, and routes prediction requests to the machine learning model. Its lightweight and fast nature makes it suitable for handling real-time chatbot interactions. Express.js also enables secure communication between the chatbot, LSTM model, and disease database through RESTful APIs. It simplifies request handling, middleware integration, and error management, ensuring reliable data flow and smooth coordination between different system components.

Supabase (PostgreSQL) is used as the backend database to store disease records, symptom lists, prediction results, and user interaction data. PostgreSQL provides a robust relational structure, ensuring data consistency, integrity, and efficient querying for medical information retrieval. This setup allows the chatbot to access accurate disease details quickly during prediction. Supabase adds a managed backend layer on top of PostgreSQL, offering built-in authentication, real-time data handling, and secure API access. These features reduce development complexity while maintaining strong security controls, which is critical for handling sensitive medical-related data in the system. The combination of Supabase and PostgreSQL ensures scalability and reliability as the system grows. As the number of users and medical records increases, the database can handle concurrent requests efficiently, supporting fast response times and uninterrupted chatbot functionality.

VII. RESULT ANALYSIS

The result analysis shows that the AI-based medical chatbot can accurately predict infectious diseases when clear and relevant symptoms are provided. The LSTM model performs well in identifying symptom patterns and mapping them to the correct disease category, demonstrating higher accuracy compared to

basic rule-based systems. However, predictions become less reliable when users input vague or incomplete symptoms, which expose a dependency on input quality. Performance evaluation indicates that the system responds quickly, with minimal delay between symptom submission and prediction display. The integration of React.js, Express.js, and Supa base ensures smooth data flow and low latency. That said, this speed advantage means nothing if the model is trained on limited or biased datasets fast wrong answers are still wrong. Comparative analysis highlights that the proposed system offers better adaptability and scalability than traditional manual or static diagnostic tools. It can be updated with new medical data and retrained models over time. Still, it should not be mistaken for a replacement for medical professionals; its results are decision-support outputs, not clinical diagnoses, and treating them otherwise would be a serious misuse of the system.

Screenshots:

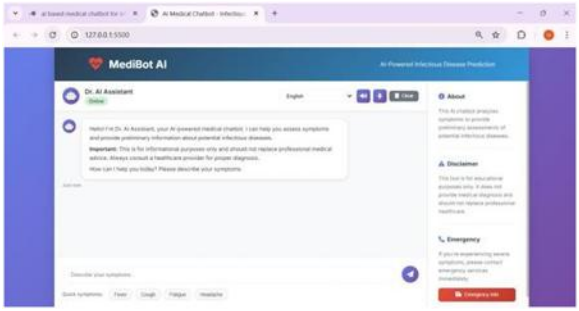

Fig: Home Page

The project's homepage presents a clean and intuitive interface where users can interact with the AI-powered medical chat bot in real-time. It allows users to enter symptoms through text input or quick symptom buttons for faster interaction. The chatbot provides preliminary assessments of potential infectious diseases along with clear medical disclaimers. Additional sections like About, Disclaimer, and Emergency Info ensure users understand the system's purpose and limitations.

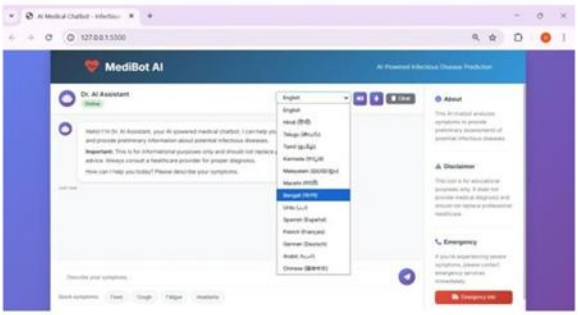

Fig: Select Language page

The interface includes a multilingual language selection feature that allows users to interact with the chatbot in their preferred language. This improves accessibility for users from diverse linguistic backgrounds. It ensures clearer symptom communication and more Accurate responses from the system.

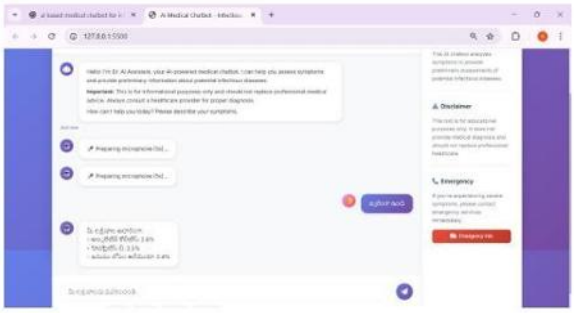

Fig: Speaker feature

The speaker feature enables voice-based interaction by capturing user symptoms through speech instead of typing. It converts spoken input into text for processing and prediction by the chatbot. This improves usability for users with limited typing ability and supports hands-free interaction, especially in emergency or accessibility scenarios.

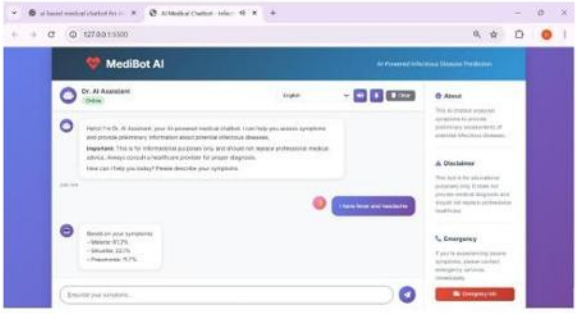

Fig: Result page

The result page displays predicted infectious diseases along with probability percentages based on the user's symptoms. It clearly ranks the most likely conditions, helping users understand risk levels instead of giving a vague answer. This design improves transparency and decision-making but still avoids claiming a medical diagnosis. The emergency panel and disclaimer remain visible to prevent misuse or overconfidence.

VIII. TESTING

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, subassemblies, assemblies and/or a finished product. It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of tests. Each test type addresses a specific testing requirement. Unit testing, integration testing, functional testing.

Table: Test cases

Valid input	Identified classes of valid input must be accepted.
Invalid input	Identified classes of invalid input must be rejected.
Functions	Identified functions must be exercised.
Output	Identified classes of application outputs must be exercised.
Systems/Procedures	Interfacing systems or procedures must be invoked.

Test Case ID	Test Case Name	Input	Expected Output	Actual Output	Test Case Result
1	User Symptom Input	Fever, Cough	Symptoms recorded correctly	Symptoms recorded correctly	Pass
2	Invalid Symptom Input	"" or 123$$	Error message prompting valid input	Error message displayed	Pass
3	Disease Prediction Accuracy	Fever, Headache, Fatigue	Correct disease prediction based on dataset	Predicted correctly	Pass
4	Response Generation	I have sore throat and cough	Chatbot provides possible disease and precautions	Response displayed	Pass
5	Module Integration	User enters symptoms	Symptoms processed → prediction → response delivered	Response delivered	Pass
6	Error Handling	Conflicting symptoms input	Chatbot prompts user to clarify symptoms	Prompt displayed	Pass
7	Functional Accuracy	Cough, Fever	Chatbot recommends likely infectious disease	Recommendation displayed	Pass
8	User Communication	Any symptom input	Response is clear, readable, understandable	Response clear and readable	Pass
9	System Update Test	Updated symptom database	Previous functions continue to work	All functions working	Pass
10	Data Security	Any symptom input	User data stored securely, no leaks	Data stored securely	Pass

IX. CONCLUSION & FUTURESCOPE

Conclusion:

The AI-based medical chatbot for infectious disease prediction demonstrates the effective integration of artificial intelligence, machine learning, and healthcare technologies to deliver quick, reliable, and easily accessible preliminary disease assessments. By analyzing user-input symptoms and comparing them with trained datasets, the system is capable of providing accurate predictions and meaningful health insights in real time. This reduces the burden on healthcare systems and enables users to take early preventive measures.

Through rigorous unit testing, integration testing, and functional testing, the system ensures high accuracy in symptom interpretation, precise disease prediction, and smooth interaction between different modules.

These testing strategies help in identifying and eliminating errors, improving system performance, and enhancing user trust in digital health care solutions. The chatbot is also designed with a user-friendly interface, making it easy for individuals of varying technical backgrounds to interact with the system effectively.

Furthermore, the chatbots modular architecture and secure data handling mechanisms make it highly scalable and adaptable for future enhancements. It can be extended with advanced features such as real-time health monitoring, multilingual support, and integration with wearable devices and hospital information systems. Continuous updates, model retraining, and performance evaluation will ensure that the system remains accurate, reliable, and up-to-date with evolving healthcare needs. In addition, the system has the potential to incorporate predictive analytics for early detection of disease trends and outbreaks, thereby supporting public health initiatives. By assisting healthcare professionals with preliminary analysis and enabling patients to seek timely medical attention, the chatbot serves as a valuable tool in modern healthcare. Overall, it contributes to improving patient outcomes, optimizing healthcare resources, and promoting a proactive approach to disease management in real-world scenarios.

Future scope:

The AI-based medical chatbot can be further enhanced by integrating advanced machine learning and deep learning models to improve the accuracy and efficiency of diagnosis and response generation. Incorporating real-time data from wearable devices such as smart watches and health monitors can enable continuous health tracking and instant analysis of vital parameters like heart rate, blood pressure, and oxygen levels. Additionally, providing multilingual support will help the system reach a wider and more diverse population, especially in rural and underserved areas.

Future developments of the system may include the use of predictive analytics to identify potential disease outbreaks and provide early warnings to users and healthcare authorities. The chatbot can also be upgraded to offer personalized health recommendations based on user history, lifestyle patterns, and medical records. Integration with hospital management systems and electronic health records (EHR) will allow seamless data exchange between patients and healthcare providers, improving diagnosis and treatment efficiency.

Moreover, the system can incorporate features such as voice recognition, telemedicine support, and emergency response services to enhance user interaction and accessibility. With these advancements, the AI-based medical chatbot can evolve into a more intelligent, reliable, and proactive healthcare solution, ultimately contributing to improved healthcare delivery and patient outcomes.

REFERENCES

[1] An AI-Based Medical Chatbot Model for Infectious Disease Prediction, "An AI-Based Medical Chatbot Model for Infectious Disease Prediction," International Journal of Engineering Science and Research (IJESR).

[2] "AI-Driven Medical Chatbot for Early Disease Detection," International Journal for Research in Applied Science and Engineering Technology (IJRASET).

[3] "Medical Chatbot: AI Based Infectious Disease Prediction Model," Journal of Scientific Research and Technology (JSRT).

[4] "AI-Based Medical Chatbot for Disease Prediction," International Journal of Innovative Science and Research Technology (IJISRT).

[5] "Innovative Healthcare Chatbot Using Artificial Intelligence and Machine Learning," International Journal for Research in Applied Science and Engineering Technology (IJRASET).

[6] "AI Chatbot Design During an Epidemic Like the Novel Coronavirus," Healthcare, MDPI.

[7] "AI-Driven Medical Chatbot for Early Disease Detection," International Journal for Research in Applied Science and Engineering Technology (IJRASET).

[8] Sriramm-15, "AI-Medical-Chatbot: An AI-Based Medical Chatbot Model for Infectious Disease Prediction," GitHub Repository. Available: https://github.com/Sriramm-15/AI-Medical-Chatbot.