iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** III **Month of publication:** March 2025

DOI: https://doi.org/10.22214/ijraset.2025.67504

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue III Mar 2025- Available at www.ijraset.com

Early Intervention through AI: A CNN based Approach to Medical Virtual Assistants

Jinka Neha[1], Sri Vaishnavi Koduri[2], Sai Kiran Reddy[3], L. Chandra Sekhar Reddy[4], M. Parameswar[5]
[1, 2, 3]*Artificial Intelligence and Data Science, CMR College of Engineering and Technology, Hyderabad*
[4, 5]*Computer Science and Engineering (Data Science), CMR College of Engineering and Technology, Hyderabad*

Abstract: ***Accessibility in Healthcare continues to be a major challenge owing to time availability, cost and limited availability of medical professionals. Most of the people, especially the ones with a fast-paced life do not take routine check-ups which results in late diagnosis and serious health issues. In order to solve this problem, we present a chatbot-based disease diagnosis system that uses artificial intelligence for preliminary medical assessments. This chatbot allows users to insert symptoms using a conversational interface, and then the chatbot predicts possible diseases and gives proper dietary recommendations. This chatbot is designed to analyze symptoms using a CNN model and deliver results efficiently and accurately. This system provides a quick, economical, and easily accessible alternative for basic health knowledge compared to conventional hospital visits. Chatbot technology improves the efficiency of the health care system by removing the need to burden the hospitals with minor health issues, while encouraging preemptive measures in monitoring health. It is a 24/7 service, offering healthcare advice around the clock, anytime and anywhere. furthermore, its simplistic interface invites people to talk about their health freely.***
Keywords: Chatbot Diagnosis, Convolutional Neural Network

I. INTRODUCTION

A healthy population is a key indicator of the wealth of a country. After all, people are most valuable in any society, so their well-being will always be the priority. In the bustling modern society, however, people often miss out on taking care of their well-being due to lifestyle constraints. As a result, people tend to avoid regular visits to the doctor, only seeking help when their situation becomes dire. Such too-late approaches can have drastic consequences, but they can also be remedied with utmost care and attention along with timely diagnosis. Apart from the need and the consequences of such neglect, the advancement of services like AI and machine learning has made strides in healthcare too. Medical Virtual Assistants, commonly known as medical chatbots, are an example of this progress. The proposed chatbot system is capable of engaging with the public, gathering symptom data from them, and even predicting diseases. Using machine learning speaks to the competency of the bot. The diet suggestion system, in addition to the symptom and disease prediction system, suggests relevant doctors and hospitals and operates using the CNN algorithm. The system that serves patients with a pre-trained dataset is innovative and modern. The patients speak, the chatbot extracts symptoms, processes them and matches them against credible pre - set symptoms. The easy to use interface ensures that people of all ages find it simple and affordable so as to promote the prevention of dire health conditions through early intervention.

II. RELATED WORKS

Many in-depth studies have investigated the use of AI-driven chatbots in healthcare settings. This section reviews several research papers that highlight the improvements and challenges in disease diagnosis using chatbot technology.

[1] Chatbot for Disease Prediction and Treatment Recommendation using Machine Learning - Published in 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI) : An AI chatbot was developed which allows the user to give input symptom, the given input symptoms are processed by NLP techniques and ML algorithms are used to predict the disease. All treatments are recommended according to each predicted disease, this chatbot helps to reduce every single hospital visit for each minor ailment.

[2] Med-bot: Conversational AI-Powered Chatbot for Telehealth after COVID-19 - Published In: 2020 5th International Conference on Communication and Electronics Systems (ICCES) : The paper particularly focuses on Telehealth, Building Upon Kentucky's COVID-19 Experience employs a multilingual NLP chatbot to provide patients with remote help. It is a serverless application based on GCP provides preventative direction, diagnostics based on symptoms, and healthcare tips.

[3] A Medical Chatbot using Natural Language Processing - Published In: 2018 International Journal of Computer Trends and Technology: A system which is used to develop a chatbot conversation for patient assistance. Assists users in diagnosing symptoms and giving general medical advice. Decreases unnecessary visits to the hospital. Trained on data processed until October 2023.
[4] Disease Prediction Using Machine Learning Algorithms - Published In: 2020 International Journal of Scientific & Technology Research : Naïve Bayes, Decision Tree, and Support Vector Machines (SVM) are some of the machine learning algorithms studied that applied on predicting the disease. The paper uses dataset which consists of patient symptom data to diagnose possible diseases and measurers model performances with their accuracy and efficiency. This study presents comparative analysis of various algorithms in disease diagnosis which will help identify the best approach for accurate diagnosis.
[5] AI-Powered Chatbots in Healthcare: Systematic Review - Published In: 2021 Journal of Medical Internet Research (JMIR): This paper focus on different chat bots implementations using NLP to engage better with the patient, helping patients get automated diagnosis, symptoms analysis, and health recommendations. The paper emphasizes how AI chatbots can enhance access to medical advice and alleviate the workload of healthcare professionals. It also tackles important issues like data privacy, security risks, and the ethical dilemmas that come with AI in medicine, stressing the importance of responsible AI use in the healthcare sector.

III. IMPORTANCE OF DISEASE DIAGNOSIS

An accurate diagnosis of a disease at its early stage is the most important factor in preventing serious clinical consequences and improving patient outcomes. Late Diagnosis Timely diagnosis of conditions like these can help in identifying these diseases in their infancy stages, preventing an illness from worsening and making treatment more effective while lowering healthcare costs. The World Health Organization (WHO) also reported that delayed detection of chronic illnesses including diabetes, cancer, and cardiovascular conditions were responsible for a considerable number of avoidable deaths globally [11]. Many people resort to seeking medical attention once they can no longer bear their symptoms, which makes the recovery more challenging and costly.
Early detection of diseases is a challenge for better global health. Early detection allows clinicians to implement appropriate management techniques, which leads to a decrease in mortality and improved quality of life. Diseases when detected early have a much higher treatment success rate, for example breast cancer and hypertension [6]. On top of that, preventative care, including check-ups and lifestyle changes, is far easier to initiate if conditions are diagnosed early on.
But the traditional medical check-up comes with a few challenges that makes early diagnosis difficult. Long wait times, expensive consultation fees, and hospital visits deter people from getting routine screenings. Along with this, the availability of healthcare professionals is extremely poor in remote or underprivileged areas. Indeed, the aversion to seeking care even for mild symptoms frequently leads to the exacerbation of conditions that are easily treatable if diagnosed timely as sought.
An AI-powered chatbot would be able to let users instantly check their symptoms, provide quick preliminary diagnoses, etc. Such systems help in bridging the gap between the people and the doctors using technology, which leads to early diagnosis and helps in the better management of health.

IV. PROPOSED SYSTEM

The proposed system is An AI-powered medical chatbot that helps users find a possible disease based on their symptoms. Due to its interactive and user-friendly interface, anyone with a health anxiety can use the chatbot to describe their symptoms and obtain a preliminary diagnosis instantly without going through the long queues of hospitals. Using Convolutional Neural Networks (CNNs) which provides with a mapping of symptoms to diseases, as well as NLP the chatbot tracks symptoms (Disease mapper) and helps in predicting the diseases and offer personalized suggestions.
The chatbot application allows users to register and log in to the app and enter their symptoms in natural language. These symptoms are processed by the system using NLP techniques and finally, through a CNN-based machine learning model, it predicts the disease with the highest probability [7]. The chatbot also offers dietary suggestions and shows nearby doctors and hospitals, making it a full-fledged healthcare helper.Furthermore, since the chatbot is trained over the structured medical dataset, it is more accurate in mapping symptoms and diseases.
Advantages of the proposed system:
1. Provides accurate disease prediction as per the symptoms entered by the user.
2.Also provides personalized dietary recommendations based upon the predicted disease.
3.Provides information about relevant healthcare facilities based on user location.
4. It can be accessed anytime, avoiding the burden of hospital visits for basic health problems.

V. METHODOLOGY

This proposed system leverages a chatbot integrated with CNNs and NLP to analyze symptoms input by users and predict potential diseases. By utilizing machine learning models trained on structured medical datasets, the system enhances diagnostic accuracy and provides personalized health recommendations. The following section provides an in-depth understanding of the data collection and preprocessing, algorithm selection, system architecture, and chatbot workflow.

A. Data Collection and Preprocessing

The chatbot is trained using a well-structured medical dataset consisting of two key columns: Disease Title, representing the name of a disease, and Symptoms, listing the associated symptoms for each disease. Each row corresponds to a specific disease along with the commonly observed symptoms. To enhance the model's accuracy, data preprocessing is crucial as it ensures that symptom descriptions are cleaned, standardized, and transformed into a format suitable for machine learning. The preprocessing process includes several essential steps.

Cleaning and normalization is the first step, addressing issues such as misspellings, redundant entries, special characters, and inconsistent formats [8]. This involves stopword removal, which eliminates non-essential words like "a," "the," and "is" that do not contribute to disease identification. Additionally, all symptom descriptions are converted to lowercase for uniformity. Techniques such as lemmatization and stemming further refine the data by reducing words to their root form, for example, changing "running" to "run" for better generalization. Missing values are also handled either through imputation techniques to fill in incomplete records or by removing highly inconsistent data.

To transform textual symptom descriptions into numerical representations suitable for machine learning, tokenization and feature extraction are applied. Tokenization breaks down input text into individual words or phrases, enabling the system to extract meaningful symptom keywords. Furthermore, the Term and Document Frequency (TF-IDF) is used to assign appropriate importance to symptoms, ensuring that rare but significant symptoms like "jaundice" are given more weight than frequently occurring symptoms such as "fever."The significance of each symptom in the dataset is calculated using Term Frequency (TF) and Inverse Document Frequency (IDF), which help determine the weight of each term. These measures ensure that rare yet crucial symptoms receive higher importance compared to commonly occurring ones. The formulas for TF and IDF are given below:

$$\mathrm{tf}(t,d) = \frac{f_{t,d}}{N} \tag{1}$$

$$\mathrm{idf}(t, D) = \log \frac{}{|\{d : d \in D \text{ and } t \in d\}|} \tag{2}$$

Once the data is preprocessed by them, it is divided up into training and testing sets, with 80% given for training and 20% for testing. The model is able to understand connections involving indications along with conditions through the training data, as the testing data guarantees the model performs well regarding new circumstances, which improves its prediction capabilities in addition to its reliability.

B. Algorithm Selected

The chatbot employs a CNN for disease prediction, leveraging its powerful ability to recognize patterns and perform classification tasks with high accuracy. CNNs are particularly effective in identifying complex relationships between multiple symptoms and diseases, making them a superior choice over traditional machine learning models such as KNN or Decision Trees. Unlike KNN, which relies on manual feature selection, CNNs are good at automated feature extraction, detecting important symptom patterns without human intervention. Additionally, CNNs demonstrate high accuracy and scalability, improving their performance as more data becomes available while efficiently adapting to new medical conditions. To prevent overfitting, the model incorporates pooling layers and dropout techniques, ensuring it does not simply memorize training data but generalizes well to unseen cases [9].

The CNN model architecture follows a structured multi-layer design, with each layer playing a crucial role in disease prediction. The input layer receives the processed symptom data, transforming it into a structured input vector for the network. This transformation can be mathematically represented as follows:

$$X' = X * W + b \tag{3}$$

where X is the input symptom vector, W is the learnable weight filters, b is bias term and * is convolution operation

Next, the convolutional layers apply kernels to detect patterns between symptoms and diseases, extracting meaningful relationships within the data. The Rectified Linear Unit or the ReLU activation function is used to introduce non-linearity, allowing the model to capture complex symptom interactions effectively.The activation function is defined as follows:

$$f(x) = \max(0, x) \tag{4}$$

To enhance efficiency, pooling layers, particularly max pooling, are implemented to reduce the dimensionality of feature maps while retaining essential information. The fully connected layer then processes the extracted features, converting them into a probability distribution of potential diseases. Finally, the softmax output layer generates the final classification, assigning probability scores to different diseases, ultimately predicting the most likely diagnosis based on the given symptoms.

$$P(y = i) = \frac{e^{z_i}}{\sum e^{z_j}} \tag{5}$$

Where P(y=i) is the probability of disease i, and z_i represents the CNN output before activation.

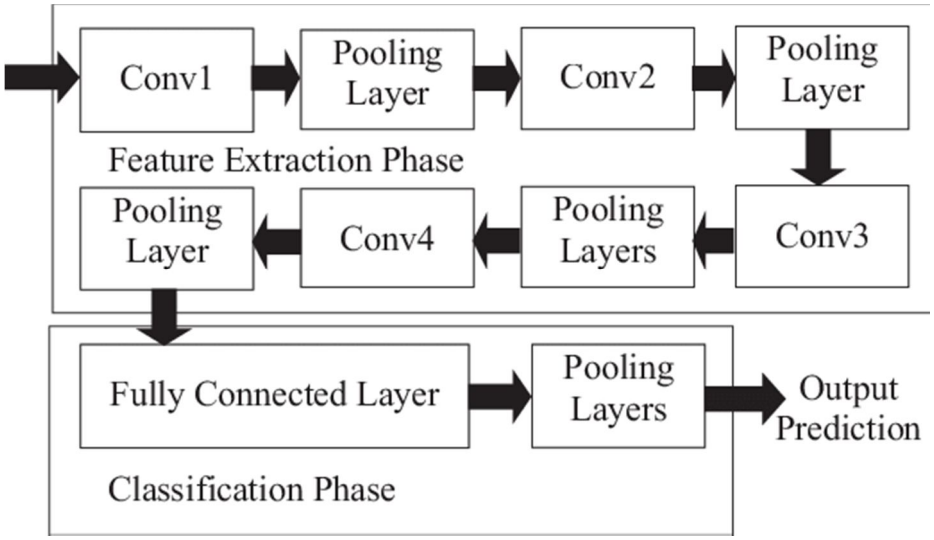


Fig. 1 Disease Classification Using CNN

C. System Architecture

The chatbot's system architecture is designed to efficiently process user inputs, extract meaningful symptoms, predict diseases using machine learning, and provide relevant medical recommendations. The system begins with a user-friendly chatbot interface, allowing individuals to enter their symptoms in natural language. These inputs are then processed by a NLP module, which utilizes techniques such as Named Entity Recognition (NER) and Dependency Parsing to accurately extract symptom-related keywords from user text. Once the symptoms are identified, they are passed to the CNN-based disease prediction model, which analyzes the data and determines the most probable disease based on trained medical datasets[10]. The recommendation system then generates personalized health suggestions, including diet plans, treatment options, and doctor specializations, ensuring users receive comprehensive medical guidance. To support these functionalities, the system integrates a medical database that securely stores disease-related information, user interaction history, and symptom-disease mappings. Additionally, a response generator is responsible for delivering chatbot responses in real time, offering users an interactive and informative experience. To enhance usability, the chatbot also sends an email report summarizing the predicted disease and recommended actions, ensuring users have access to their health insights beyond the chat interface. This structured and intelligent system architecture ensures efficiency, accuracy, and accessibility in preliminary disease diagnosis and medical assistance.

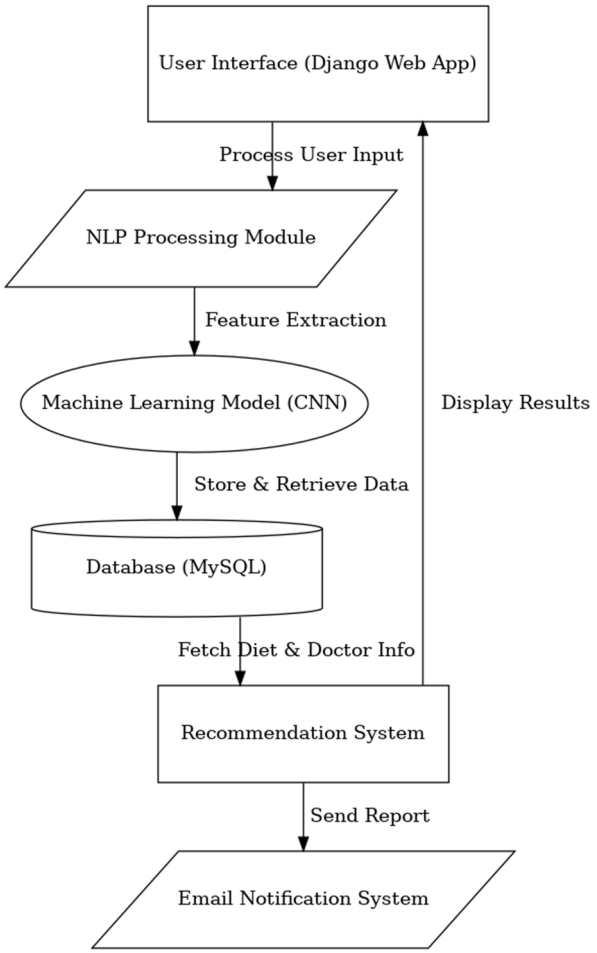


Fig. 2 Disease Diagnosis Chatbot Architecture

D. Model Evaluation

For assessing the effectiveness of the CNN-based chatbot at predicting many diseases, multiple accuracy metrics were used upon it, including Precision, Recall, and F1-score. These metrics can be helpful to judge the model's success in finding true diseases and guaranteeing few false alarms. The trained model underwent testing with a validation dataset during the assessment. This dataset differed from the original dataset, providing a fair way to evaluate the model's predictive capabilities. Precision calculates the percentage of accurately forecasted diseases out of all the diseases found by the model. A greater precision score shows that the model is creating fewer incorrect positive predictions, indicating it is differentiating well between multiple illnesses using patient symptoms. Recall, on the other hand, evaluates how well the model identifies actual diseases among all the cases that should have been detected. A high recall score signifies that the chatbot is successfully recognizing diseases when relevant symptoms are provided [12]. Providing an equilibrium between exactitude and recollection, the F1-score represents a special measurement that thoroughly considers both false positives and false negatives. Considering that a dependable disease diagnosis system has to give correct forecasts and also confirm that no serious ailments are overlooked, a suitable F1-score displays the chatbot's dependability. The CNN model's ability for recognition of subtle patterns along with relationships between symptoms notably contributes to its high accuracy as well as efficiency within disease prediction.

VI. RESULT

The CNN-based chatbot that was suggested was then carefully assessed using common precision measurements, displaying how that kind of model is able to correctly forecast diseases. The presented outcomes point to the improved working of the chatbot helped by deep learning methods over regular setups that adhered to specific rules when talking. The model precisely distinguishes between different symptom types, as with its strong accuracy in properly forecasting illnesses. The rare amount of false positives and false negatives betters the dependability throughout its diagnosis.

The minimal number of false positives and false negatives makes it a preferred choice over typical systems. Also, the software retains a clear ability when it comes to differentiating correct entries from mistakes. The architecture in the model lets it understand several relationships that are hard to see between symptoms and in diseases, which is why it can make better predictions for many diseases, as shown in the evaluation results. Also, its demonstrated capability of notably decreasing false positives and false negatives greatly betters its general dependability within the diagnosis, making it a helpful addition to the usual systems, with an important guide for correct diagnosis. The results from that detailed analysis stress that the careful structure of the framework lets it record quite detailed connections linking signs as well as illnesses, thus making far better forecasts across quite a few health aims.

Performance Metrics: The trained model passed through a validation dataset separated from the initial dataset to test its performance for the chatbot. Here are the main evaluation metrics:

TABLE I

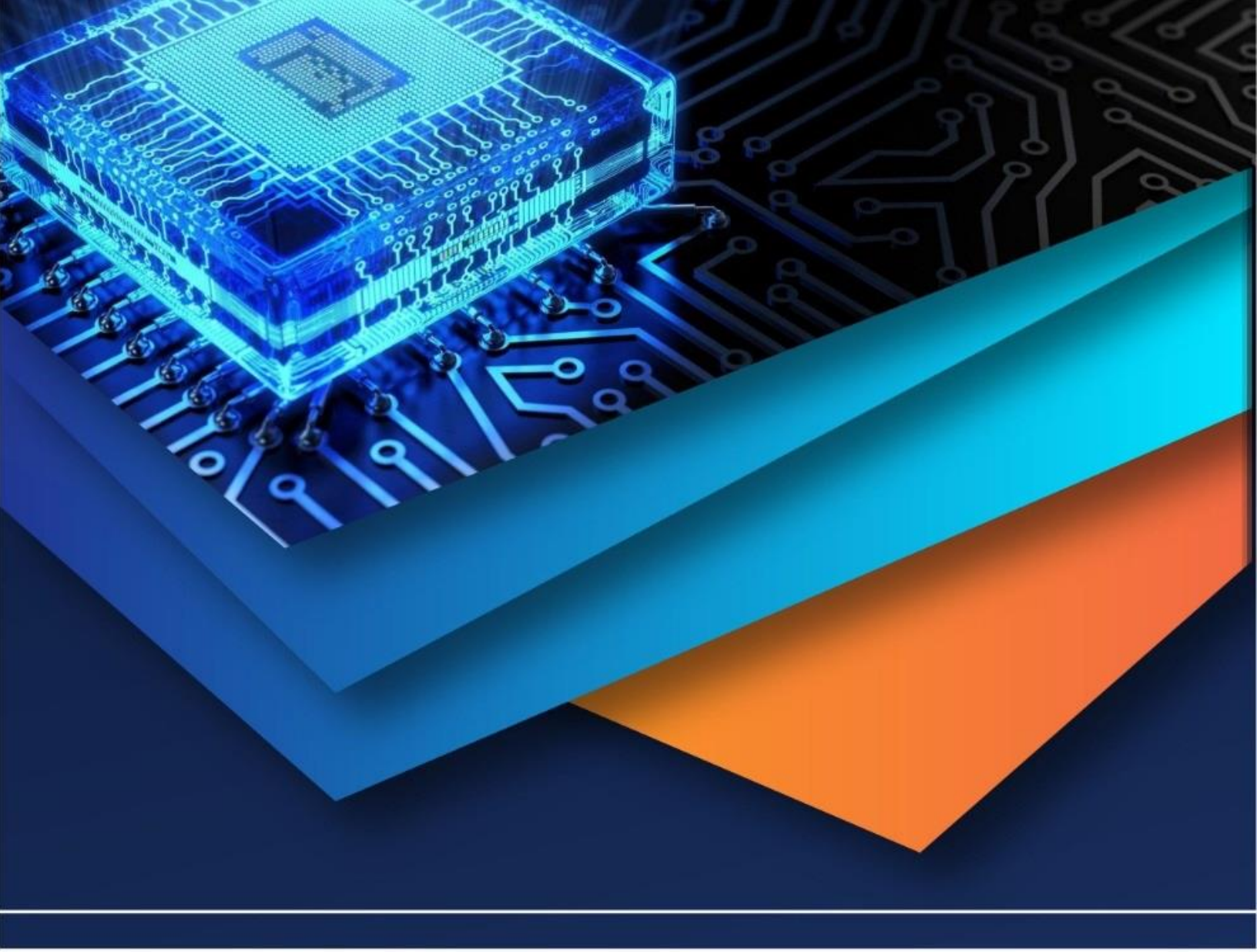
METRIC	VALUE(%)
Precision	91.2%
Recall	89.7%
F1-Score	90.4%
Accuracy	92.5%

VII. CONCLUSION

The proposed medical chatbot is a way to automate the process of diagnosing illnesses and making recommendations for treatments. Using Convolutional Neural Networks (CNN), a type of machine learning, the chatbot effectively interprets user-provided symptoms and makes precise predictions about potential illnesses. By providing real-time diagnostic support and tailored health recommendations, it lessens the need for frequent hospital stays. Additionally, by offering a private and intuitive platform where users may freely discuss their health concerns, the chatbot fosters greater user connection.This system's cost-effectiveness and accessibility are two of its main benefits, as they enable a wider population to have access to high-quality healthcare information. In contrast to traditional approaches that necessitate medical assistance, this chatbot autonomously offers reliable disease predictions, food recommendations, and doctor details, guaranteeing consumers get prompt direction. Consequently, it serves as a preliminary diagnostic tool that promotes proactive healthcare management in addition to raising health awareness. This chatbot acts as an effective and useful healthcare assistant, bridging the gap between people and timely medical advice in a fast-paced environment where routine medical consultations can be difficult.

REFERENCES

[1] R. B. Mathew, S. Varghese, S. E. Joy, and S. S. Alex, "Chatbot for Disease Prediction and Treatment Recommendation using Machine Learning," Proc. 3rd Int. Conf. Trends Electron. Inform. (ICOEI), Tirunelveli, India, 2019, pp. 851-856.
[2] M. B. Devi, G. G. Kumar, and P. Rajalakshmi, "Med-bot: Conversational AI-powered chatbot for telehealth after COVID-19," Proc. 5th Int. Conf. Commun. Electron. Syst. (ICCES), 2020, pp. 870-87
[3] R. Dharwadkar and N. A. Deshpande, "A Medical ChatBot," Int. J. Comput. Trends Technol. (IJCTT), vol. 60, no. 1, pp. 41-45, June 2018. ISSN:2231-2803. Available: www.ijcttjournal.org.
[4] I. Jovovic, D. Babic, T. Popovic, S. Cakic, and I. Katnic, "Disease Prediction Using Machine Learning Algorithms," Proc. Int. Conf. Inf. Technol. (IT), 2023,
[5] N. Ruggiano, E. L. Brown, L. Roberts, C. V. F. Suarez, Y. Luo, Z. Hao, and V. Hristidis, "Chatbots to Support People With Dementia and Their Caregivers: Systematic Review of Functions and Quality," J. Med. Internet Res., vol. 23, no. 6, pp. e25006, 2021
[6] Z. Yu, K. Wang, Z. Wan, S. Xie, and Z. Lv, "Popular deep learning algorithms for disease prediction: a review," Cluster Comput., vol. 26, no. 2, pp. 1231-1251, 2023,
[7] S. Grampurohit and C. Sagarnal, "Disease Prediction using Machine Learning Algorithms," Proc. Int. Conf. Emerg. Technol. (INCET), Belgaum, India, 2020, pp. 1-7
[8] Divya Madhu,Neeraj Jain C. J, Elmy Sebastain, Shinoy Shaji, Anandhu Ajayakumar. A Novel Approach for Medical Assistance Using Trained Chatbot, International Conference on Inventive Communication and Computational Technologies (ICICCT 2017).
[9] Mrs. Rashmi Dharwadkar, Dr.Mrs. Neeta A. Deshpande "A Medical ChatBot"" in International Journal of Computer Trends and Technology (IJCTT) V60(1):41-45 , June 2018.
[10] S. Divya, V. Indumathi, S. Ishwarya, M. Priyasankari, S. Kalpana Devi, "A Self-Diagnosis Medical Chatbot Using Artificial Intelligence", J. Web Dev. Web Des., vol. 3, no. 1, pp. 1-7, 2018.
[11] Amiya Kumar Tripathy , Rebeck Carvalho , Keshav Pawaskar , Suraj Yadav, "Mobile based healthcare management using artificial intelligence" in International Conference on Technologies for Sustainable Development (ICTSD), 2015, 4-6 Feb, 2015.
[12] S. du Preez, M. Lall, S. Sinha, "An intelligent web-based voice chat bot", EUROCON 2009 EUROCON'09. IEEE, pp. 386-391, 2009.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY