MULTI-DISEASE DETECTION USING MACHINE LEARNING

Prof. Namithadevi N N[1], Chandan P K[2], Shreyas M[3], Tarun Ayappa K G[4], Adarsh K S[5]

[1] *Assistant Professor, Computer science and Engineering, Vidya Vikas Institute of Engineering & Technology, Karnataka, India*
[2] *Student, Computer science and Engineering, Vidya Vikas Institute of Engineering & Technology, Karnataka, India*
[3] *Student, Computer science and Engineering, Vidya Vikas Institute of Engineering & Technology, Karnataka, India*
[4] *Student, Computer science and Engineering, Vidya Vikas Institute of Engineering & Technology, Karnataka, India*
[5] *Student, Computer science and Engineering, Vidya Vikas Institute of Engineering & Technology, Karnataka, India*

ABSTRACT

The increasing demand for early disease diagnosis calls for innovative, accessible, and efficient healthcare solutions that can cater to a wide range of individuals, especially in areas with limited access to healthcare facilities. This paper presents a machine learning-based system designed to predict multiple diseases, including heart disease, kidney disease, thyroid disorders, Parkinson's disease, diabetes, and brain tumors. For diseases such as heart disease, kidney disease, and diabetes, Random Forest classifiers are utilized, providing accurate, real-time predictions based on relevant health data. For brain tumor detection, a Convolutional Neural Network (CNN) is employed, taking MRI images as input and delivering highly accurate diagnoses. Additionally, the system incorporates a unique symptom-based disease detection feature, where users can select at least three symptoms from a predefined list. This allows the system to predict the most probable disease based on the selected symptoms and suggest appropriate medication and treatment options.

A user-friendly, Streamlit-based web application serves as the interface, augmented with an AI-powered chatbot that provides real-time first aid guidance and personalized healthcare advice. This system improves prediction accuracy compared to existing methods, providing a transformative tool for preliminary healthcare screening and democratizing access to diagnostic technologies. By enabling early intervention, the system aims to empower users to take proactive steps in managing their health, especially in regions with limited access to healthcare services.

This solution demonstrates a significant advancement in the use of machine learning and AI for healthcare, offering a cost-effective and scalable way to support early disease detection and healthcare accessibility.

Keyword: - *Multi-Disease Detection, Machine Learning, Random Forest, CNN, Healthcare Prediction, Streamlit, AI Chatbot*

1. INTRODUCTION

Advancements in machine learning have paved the way for innovative healthcare solutions capable of early disease detection. Traditional diagnostic methods often require extensive testing, time, and expert analysis, limiting accessibility, especially in remote regions. Our project proposes a comprehensive system that predicts multiple diseases through machine learning models trained on structured patient data, symptoms, and medical images. The system features a user-friendly interface built with Streamlit and an AI chatbot for preliminary healthcare assistance.

The integration of multiple disease models into a single platform offers enhanced accessibility, quicker diagnosis, and promotes proactive health monitoring.

2. MULTI-DISEASE DETECTION SYSTEM

The proposed multi-disease detection system is a user-centric diagnostic platform powered by machine learning models to predict various diseases using two primary input methods: structured health parameters and symptom-based selections. The system is modular and includes an AI-powered chatbot for user assistance, ensuring access to guidance and suggestions throughout the interaction.

2.1 User Interaction Modes

Users can initiate diagnosis by selecting one of the following modes:

•**Value-Based Input:** Users enter medical parameters such as blood pressure, glucose levels, creatinine, TSH, etc., depending on the disease they wish to check. Each disease model (brain tumor, heart disease, kidney disease, etc.) processes these values using pre-trained machine learning models for accurate predictions.

•**Symptom-Based Input:** Users select at least three symptoms from a predefined list (e.g., chest pain, fatigue, cough, etc.) to predict likely diseases. This mode employs a multi-class classification model that processes the symptoms and outputs the most probable disease(s). Once detected, the system also suggests medication or preventive measures. For example, if a fungal infection is detected, the system may recommend antifungal medication as a precautionary treatment.

•**AI Chatbot Assistance:** Regardless of the input method, users can access the AI chatbot at any time. The chatbot provides support by answering general health-related queries, offering first aid advice, and suggesting precautionary measures based on the detected disease.

2.2 Key Features

- Flexible UI powered by Streamlit for seamless user interaction.
- Real-time predictions using pre-trained .pkl and .h5 models.
- Integrated medical guidance through AI chatbot for non-diagnostic assistance.
- Symptom-based prediction with medication suggestions to improve user awareness and self-care.

3. METHODOLOGY

The multi-disease detection system involves several stages as shown in Figure 1:

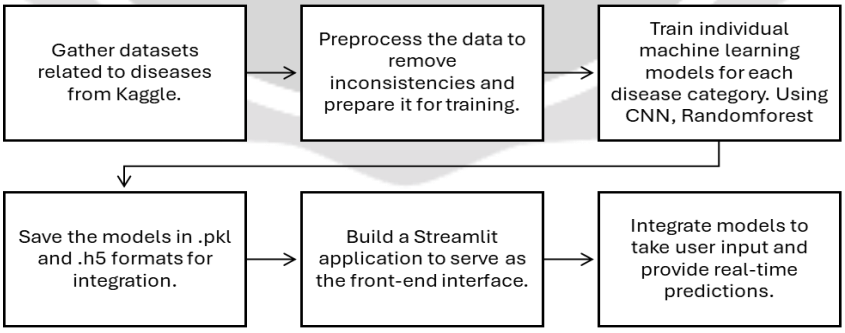

Fig -1: Methodology of Multi-Disease Detection System

i. Data Acquisition and Preprocessing

Datasets are gathered from reliable sources such as Kaggle, which include both structured patient data (CSV files with health parameters) and medical images (e.g., MRI scans for brain tumor detection). The data is preprocessed to remove inconsistencies and standardize values to ensure accurate model training.

ii.Model Training

For diseases such as heart disease, kidney disease, diabetes, and others, Random Forest classifiers are employed, using structured health parameters as input. For brain tumor detection, Convolutional Neural Networks (CNNs) are used to analyze MRI images, as these models excel in processing image data.

iii. Model Storage and Integration

Once trained, the models are saved in .pkl and .h5 formats. These models are integrated into a Streamlit application, enabling real-time predictions based on user inputs.

4. PROPOSED SYSTEM

The multi-disease detection system is designed to integrate machine learning for disease detection, medical image analysis, and symptom-based diagnosis into a unified web platform. The architecture is modular, allowing for easy expansion and updates as more diseases are added or new features are implemented.

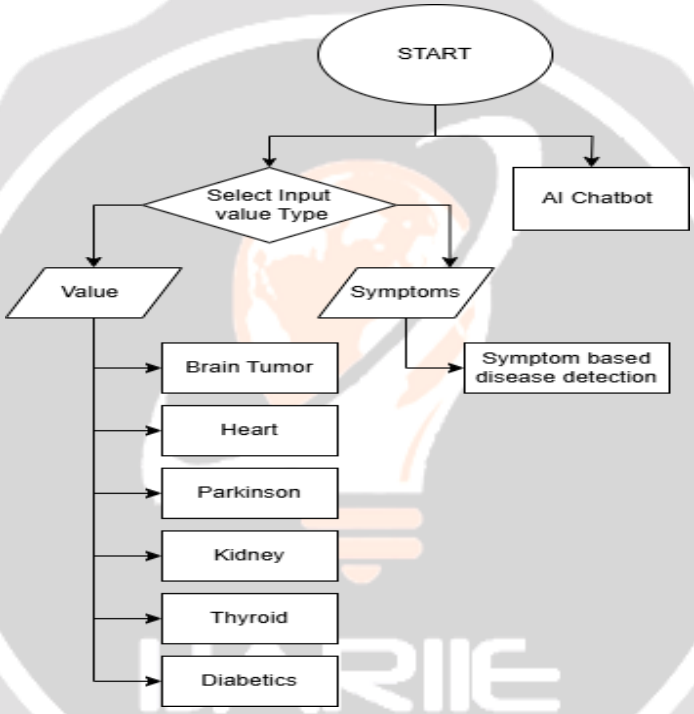

Fig -2: Workflow of Multi-Disease Detection System

4.1 System Architecture Overview

The system architecture follows a modular design, as shown in Figure 4.1. Users first select an input type—Value-Based Input (health parameters like age, blood pressure) or Symptom-Based Input (selecting symptoms).

- For **Value-Based Input**, machine learning models predict diseases such as brain tumor, heart disease, Parkinson's, kidney disease, thyroid disorders, and diabetes.
- For **Symptom-Based Input**, the system detects diseases based on user-selected symptoms.

An **AI Chatbot** provides real-time assistance throughout the process, offering medical guidance and first aid advice.

5. CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

The multi-disease detection system successfully demonstrates the potential of machine learning in democratizing healthcare diagnostics. Achieving accuracies ranging from 84% to 98.25% across various diseases, the platform significantly outperforms many traditional methods. While the current system covers six diseases, future enhancements could include expanding disease coverage, integrating real-time health data from wearables, refining

symptom-based detection using Natural Language Processing (NLP), and undergoing clinical validation for real-world deployment. Emphasis on data privacy, security, and mobile accessibility will further strengthen the impact and scalability of the platform.

6. REFERENCES

[1]. K. Arumugama, Mohd Navedb, Priyanka P. Shindec, Orlando Leiva-Chaucad, Antonio Huaman-Osorioe, Tatiana Gonzales-Yanac," Multiple disease prediction using Machine learning algorithms", https://doi.org/10.1016/j.matpr.2021.07.361

[2]. Ms. Vansika Gupta, Bhavya Tyagi, Dhruv Varshney, Samridhi Yadav, "Multiple Disease Prediction Using Machine Learning", (IJFMR) E-ISSN: 2582-2160, Volume 6, Issue 2.

[3]. Prof Vivek Pandey, Tanay Tambe, Vipin Vishwakarma, Raheem Ansari, "Multiple Diseases Prediction System using ML, Deep Learning and Big Data", ISSN 2582-7421 Vol 4, no 4, pp 2742-2745.

[4]. Kumar Bibhuti B. Singh, Ashutosh Sharma, Ashish Verma, Ranjeet Maurya, Dr. Yusuf Perwej,"Machine Learning for the Multiple Disease Prediction System", ISSN : 2456-3307 doi : https://doi.org/10.32628/CSEIT24103217

[5]. PARSHANT, DR. ANU RATHEE, "Multiple Disease Prediction Using Machine Learning", IRE Journals | Volume 6 Issue 12 | ISSN: 2456-8880.

[6]. Mastanbi Shaik, Satya Anvesh Burle, Dileep Kumar Chalumuri, Venkata Raghavendra Dalavai , Devaraj Burle, "MACHINE LEARNING BASED MULTIPLE DISEASE PREDICTION SYSTEM" IJARIIE-ISSN(O)-2395-4396, Vol-10 Issue-2 2024.

[7]. MIN CHEN, (Senior Member, IEEE), YIXUE HAO, KAI HWANG, (Life Fellow, IEEE), LU WANG, AND LIN WANG "Disease Prediction by Machine Learning Over Big Data From Healthcare Communities", IEEE, 10.1109/ACCESS.2017.2694446 .

[8]. Akkem Yaganteeswarudu, "Multi Disease Prediction Model by using Machine Learning and Flask API", IEEE, ISBN: 978-1-7281-5371-1 .

[9]. Srinidhi B, T.P.Manasa, "Multiple disease prediction using Machine Learning Algorithms", IRJET, e-ISSN: 2395-0056, Volume: 10 Issue: 02, p-ISSN: 2395-0072.

[10]. Danny A.Milner, Jr., Jochen K. Lennerz, "Technology and Future of Multi-Cancer Early Detection", Life 2024, 14, 833. https://doi.org/ 10.3390/life14070833.