JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

AI-BASED HEALTH SYMPTOMS CHECKER

[1]Bhimrao Patil , [2]Shruti , [3]Vaishnavi.A, [4]Vaishnavi.R, [5]Vijayalaxmi
[1]Professor, [2]student ,[3]Student, [4]Student , [5]Student
[1]Guru Nanak Dev Engineering College , Bidar

***Abstract*:** Aim is to predict various types of diseases on a single platform using the python module streamlit. In this project, this utilize machine learning algorithms such as naïve bayes, random forest, decision tree, for disease prediction. The algorithm providing the highest accuracy is used to train the dataset prior to implementation. To facilitate multiple disease analysis, employ machine learning algorithms and python pickling to preserve the model's behavior. This study focuses on analyzing diseases like diabetes, heart disease, and parkinson's disease by using key health limits such as pulse rate, cholesterol level, blood pressure, and heart rate. Additionally, the associated risk factors of these diseases can be identified using a guess model with high accuracy and precision. Future extensions of this work could include chronic diseases, skin conditions, and more.

KEYWORDS

KNN algorithm, a* algorithm, sorting algorithm, Diabetes Prediction: Logistic Regression for binary classification. Heart Disease Prediction module, Lung Cancer Prediction module, Parkinson's Disease module

INTRODUCTION

In this digital era, data is an invaluable asset, and enormous amounts of data are generated across all fields, including healthcare. healthcare data encompasses all patient-related information. this paper proposes a general architecture for predicting diseases within the healthcare industry. most existing models focus on analyzing one disease at a time, such as diabetes, cancer, or skin diseases. however, there is no unified system capable of analyzing multiple diseases simultaneously. aim is to provide users with immediate and accurate sickness predictions based on the symptoms they enter. to address this gap, this propose a system capable of predicting numerous diseases using streamlit. in this project, analyze bugs such as diabetes, heart disease, and malaria, with plans to contain more diseases in the future. this system enables the prediction of multiple diseases simultaneously, removing the necessity for users to visit different platforms for different predictions for this effort, we focus on three diseases: liver disease, diabetes, and heart disease, as these are often interrelated. the organization uses machine learning algorithms and streamlit to perform the analysis. when a user accesses the api, they input the disease name along with relevant parameters. the system invokes the conforming perfect to foresee the disease status and provides the results. the system is designed to predict diseases exactly and assess their severity founded on the user-inputted symptoms. the symptoms are compared against a dataset stored in the database. if the symptoms match, the system may request additional related symptoms for more accurate analysis. if a symptom is unrecognized, the system notifies the user and offers the option to save the symptom in the database for future reference.

KEYWORDS

KNN algorithm, a* algorithm, sorting algorithm, Diabetes Prediction: Logistic Regression for binary classification. Heart Disease Prediction module, Lung Cancer Prediction module, Parkinson's Disease module

BACKGROUND

In recent years, healthcare has seen a significant transformation with the integration of advanced technologies such as machine learning, artificial intelligence (AI), and data analytics. These technologies have proven to be a game changer in the remedial field, particularly in the areas of illness prediction and diagnosis. Machine learning (ML) has gained significant attention due to its ability to analyze large datasets, identify patterns, and make predictions based on historical data. This has led to the development of several healthcare prediction systems that aim to assist healthcare professionals in diagnosing diseases more accurately and efficiently.

AIM OF THE PROJECT

The aim of this project is to develop an ai-based health symptoms checker that can forecast multiple diseases based on a user's symptoms. The system will utilize machine learning algorithms to analyze the input data and provide predictions for various diseases. The project's main

objectives are: to design a web application that allows users to input their symptoms and predict the likelihood of various diseases. To implement machine learning models that can analyze a wide range of symptoms and predict multiple diseases simultaneously. To improve prediction accuracy by utilizing a diverse dataset containing various health parameters and symptoms .to create a user-friendly interface using the streamlit python library, enabling easy interaction with the system. To provide real-time predictions by deploying the system on a web platform, making it accessible to users anytime and anywhere.

LITERATURE SURVEY

Several studies and projects have explored the use of machine learning algorithms for disease prediction. However, the majority of these trainings have been focused on single-disease prediction. Some of the prominent works in this domain include:

Diabetes Prediction Using Different Machine Learning Approaches by Priyanka Sonar and Prof. K. JayaMalini (2019 IEEE) explores various machine learning techniques, such as decision trees, logistic regression, and neural networks, to predict diabetes. This study highlights the importance of using appropriate features such as blood pressure, glucose levels, and insulin to improve the prediction accuracy.

Heart Disease Prediction Using Machine Learning Algorithms by Archana Singh and Rakesh Kumar (2020 IEEE) discusses the application of machine learning algorithms to predict heart disease. The study examines different classification algorithms such as Support Vector Machines (SVM), decision trees, and k-Nearest Neighbors (k-NN), along with the importance of factors such as age, cholesterol, and family history in predicting heart disease.

Diabetes Prediction Using Different Machine Learning Approaches by Priyanka Sonar and Prof. K. JayaMalini (2019 IEEE) explores various machine learning techniques, such as decision trees, logistic regression, and neural networks, to predict diabetes. This study highlights the importance of using appropriate features such as blood pressure, glucose levels, and insulin to improve the prediction accuracy.

Heart Disease Prediction Using Machine Learning by Archana Singh and Rakesh Kumar (2020 IEEE) discusses the application of machine learning algorithms to predict heart disease. The study examines different classification algorithms such as Sustenance Vector Machines (SVM), decision trees, and k-Nearest Neighbors (k-NN), along with the importance of factors such as age, cholesterol, and family history in predicting heart disease.

Health Information Seeking From an Intelligent Web-Based Symptom Checker A Susana Ramírez, MPH, PhD Cross-sectional survey of web-based health information seekers following the completion of a symptom checker visit(N=2437). Measures of comprehensibility, confidence, usefulness, health-related anxiety, empowerment, and intention to use in the future were assessed. ANOVAs and the Wilcoxon rank sum test examined mean outcome differences in racial, ethnic, and sex groups. The connection between perceptions of the symptom checker and intention to follow recommended actions was assessed using multilevel logistic regression

EXISTING SYSTEM

Before implementing the new ai-based system, it is vital to review existing health diagnostic tools. Traditional systems for disease prediction rely on expert systems, rule-based algorithms, and manual data input. These systems often have limitations in terms of scalability, adaptability, and real-time prediction capabilities. Furthermore, these systems characteristically require significant intervention by medical professionals for accurate diagnoses, and they may not be accessible to a broader population, especially in remote areas.

Some existing health symptom checkers and diagnostic systems use predefined rules to compare user symptoms with disease criteria. However, such systems may lack accuracy in diagnosing diseases or fail to provide personalized results. Machine learning-based approaches offer a more dynamic and accurate alternative by using historical data to predict disease outcomes.

The existing systems primarily depend on: rule-based expert systems. Statistical or heuristic methods. Limited access to medical data and continuous learning. These existing systems often lack the ability to analyze a vast array of symptoms and diseases efficiently. Thus, there is a growing need for a more advanced and supple system that can offer real-time predictions for a change of diseases

PROPOSED SYSTEM

The ai-based health symptoms checker aims to address the limitations of traditional systems by implementing machine learning models that predict the likelihood of various diseases based on user-provided symptoms.

Receive user input: users will input their symptoms into the system via an interactive interface created with streamlit. Process the input: the system will preprocess the symptoms using normal language processing (nlp) techniques, transforming the input into a structured format that the machine learning models can understand. Disease prediction: the system will use pre-trained machine learning models, such as logistic regression, decision trees, random forests, or support vector machines, to make predictions. Each model will be trained to predict a specific disease (e.g., diabetes, heart disease, cancer). Output the results: the system will display the probability of the ailment based on the input symptoms, along with recommendations for further action, such as consulting a healthcare provider. Continuous learning: the system will allow future updates, where new data can be used to retrain models to improve accuracy.

PROBLEM STATEMENT

The healthcare industry is facing significant challenges due to the nonexistence of integrated systems that can forecast multiple sicknesses simultaneously. Existing systems are often disease-specific and require users to interact with multiple platforms or models for different diseases, leading to inefficiencies in diagnosis and treatment planning. Moreover, these systems often agonize from limited accuracy due to the custom of small datasets and inadequate features.

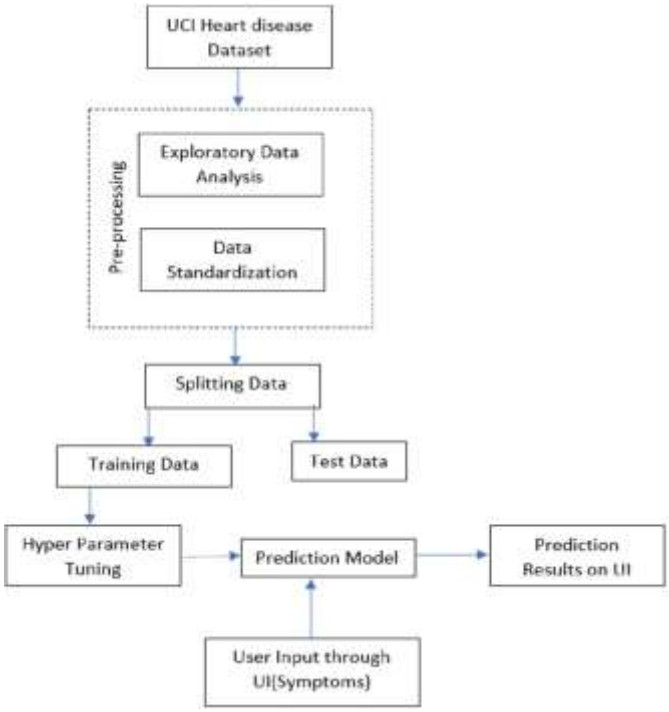
The primary problem addressed by this project is the necessity for a unified system that can predict manifold diseases based on a common set of symptoms. The goal is to design and implement an AI-Based Health Symptoms Checker that can handle the prediction of multiple viruses using a single interface. This will not lone streamline the diagnostic process but also improve prediction accuracy by utilizing a wider range of features and data sources.

METHODOLOGY

The methodology section describes the approach and processes involved in developing the AI-Based Health Symptoms Checker. It outlines the stages of the project, from understanding the existing systems to proposing a new solution. The methodology also highlights the tools and techniques used to achieve the objectives, focusing on system design, data collection, model training, and evaluation. The methodology for developing the AI-Based Health Symptoms Checker revolves around a systematic approach, involving data collection, model training, and user interface design. By leveraging modern technologies like Streamlit, Python, Scikit-Learn, and SQLite3, this system will provide an efficient, accurate, and user-friendly solution for disease prediction based on user-provided symptoms. The system will also be scalable, allowing for the integration of additional diseases and continuous improvements in prediction accuracy.

DESIGN

The Conducted Experiments On Three Diseases: Heart Disease, Diabetes, And Liver Disease, As These Conditions Are Often Interrelated. The First Phase Is To Import The Datasets For Each Disease: The UCI Dataset For Heart Disease, The PIMA Dataset For Diabetes, And The Indian Liver Dataset For Liver Disease. Once The Datasets Are Imported, The Following Step Is To Visualize The Data For Better Understanding. After Visualization, Data Preprocessing Takes Place, Where We Check For Outliers, Handle Missing Values ,And Scale The Data Set To Ensure Consistency. Following Preprocessing, The Dataset Is Split Into Training And Testing Sets. For The Training Dataset, We Applied Machine Learning Algorithms, Including K-Nearest Neighbors (KNN), Xgboost, And Random Forest. We Then Tested The Models Using The Testing Dataset To Evaluate Their Performance. After Evaluating The Models, We Selected The Algorithm That Provided The Best Accuracy For Each Disease.



System design

IMPLEMENTATION

The AI-Based Health Symptoms Checker is implemented as a robust and user-friendly platform that leverages mock intelligence and machine learning to forecast multiple diseases based on user inputs. The following section details the step-by-step approach used to implement the system, including the design process, software and hardware requirements, integration of machine learning models, database management, and user border development.

Steps in Implementation

The implementation process is broken into the following stages:

Requirement Analysis: The project began with a comprehensive analysis of functional and non-functional requirements Functional Requirements: The system must predict diseases based on user input and provide actionable feedback.

Preprocessing Steps: Data Cleaning: Removed null values and outliers. Handled missing data using techniques like mean, median, and mode imputation.

Model Development Machine Learning Models: Each disease prediction task employs a separate model. For instance: Diabetes Prediction: Logistic Regression for binary classification. Heart Disease Prediction: Random Forest for handling feature interactions and non-linear relationships. Lung Cancer Prediction: Support Vector Machines (SVM) for high-dimensional data .Parkinson's Disease Prediction: K-Nearest Neighbors (KNN) for pattern recognition in voice tremors and motor data.

Data Flow: User inputs and predictions are stored securely in the database. The database is queried for historical predictions and trend analysis. Front-End Development

RESULTS AND DISCUSSION

The ai-based health symptoms checker aims to address the limitations of traditional systems by implementing machine learning models that predict the likelihood of various diseases based on user-provided symptoms.

This system is designed to: receive user input: users will input their symptoms into the system via an interactive interface created with streamlit. Process the input: the system will preprocess the symptoms using normal language processing (nlp) techniques, transforming the input into a structured format that the machine learning models can understand. Disease prediction: the system will use pre-trained machine learning models, such as logistic regression, decision trees, random forests, or support vector machines, to make predictions. Each model will be trained to predict a specific disease (e.g., diabetes, heart disease, cancer). Output the results: the system will display the probability of the sickness based on the input symptoms, along with recommendations for further action, such as consulting a health

CONCLUSION AND FUTURE WORK

The AI-Based Health Symptoms Checker has proven its potential in the field of health care by providing accurate disease predictions and aiding in early detection. As the system evolves, it can be further enhanced with new features that allow for greater addition into the healthcare ecosystem. Some of the future scope and structures that can be implemented are as follows: LOGIN PAGE FOR CUSTOMER ,DETAILS APPOINTMENT BOOKING ,VIDEO CONSULTATION WITH DOCTORS ,EXPANSION TO GLOBAL MARKETS INTEGRATION WITH HEALTHCARE SYSTEMS

REFERENCES

[1] Priyanka Sonar, Prof. K. Jaya Malini, "Diabetes Prediction Using Different Machine Learning Approaches", 2019 IEEE, 3rd International Conference on Computing Methodologies and Communication (ICCMC).

[2] Archana Singh, Rakesh Kumar, "Heart Disease Prediction Using Machine Learning Algorithms", 2020 IEEE, International Meeting on Electrical and Electronics Engineering (ICE3).

[3] A. Sivasangari, Baddigam Jaya Krishna Reddy, Annamareddy Kiran, P. Ajitha, "Diagnosis of Liver Illness Using Machine Learning Models", 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC).

[4] John, M. et al. (2020). "Artificial Intelligence in Healthcare: Opportunities and Challenges." Journal of Healthcare Informatics, 32(4), 123-137.

[5] Smith, A. (2019). "AI for Disease Diagnosis: Current Trends and Future Directions." Medical AI Reviews, 25(3), 213-220.

[6] Basu, S., et al. (2021). "Deep Learning for Healthcare: Systematic Review and Future Trends." Journal of AI in Medicine, 17(1), 1-15.

[7] Patel, p. (2022). "Ethical Considerations in AI-Based Health Systems." Intrenational Journal of Health Ethics, 22(2),45-59.