International Journal On Engineering Technology and Sciences – IJETS™
ISSN(P): 2349-3968, ISSN (O): 2349-3976
Volume X, Issue III, March - 2023

DOCASSIST: PREDICTIVE ANALYTICS AND CHATBOT FOR MEDICAL DIAGNOSIS

Ashitha Angel D
Computer Science and Engineering
Francis Xavier EngineeringCollege, Tirunelveli – Tamil Nadu - India
ashithaangeld.ug.21.cs@francisxavier.ac.in

Raga Sudha M
Computer Science and Business System
Francis Xavier EngineeringCollege, Tirunelveli – Tamil Nadu – India
ragasudham.ug.21.cb@francisxavier.ac.in

Dheiva Sri M
Computer Science and Business System
FrancisXavier Engineering College Tirunelveli –Tamil Nadu – India
dheivasrim.ug.21.cb@francisxavier.ac.in

Dr. R. Ravi
Professor / Dept. of Computer Science and Engineering,
Francis Xavier Engineering College
Tirunelveli – Tamil Nadu – India
fxhodcse@gmail.com

Abstract:
The aim of this project is to develop a machine learning-based system that can predict and diagnose various medical conditions related to the kidney, eye, heart, lungs, breast, liver, brain tumor, and other diseases. The system will utilize a chatbot interface that will allow patients to interact with the system in a natural and intuitive way. The system will be designed using various machine learning algorithms and techniques to predict the likelihood of various medical conditions based on a set of symptoms, medical history, and other relevant factors. The system will also provide patients with personalized recommendations based on their specific conditions, such as dietary advice, lifestyle changes, and treatment options. The chatbot interface will be developed using natural language processing (NLP) techniques to enable patients to communicate with the system in a conversational manner. The chatbot will be trained using a large dataset of medical conversations and will be able to provide personalized recommendations and advice to patients.The system will be designed to be scalable and modular, allowing for easy integration with other healthcare systems and technologies. The system will also be designed with security and privacy in mind, ensuring that patient data is protected at all times.Overall, this project aims to develop a powerful and effective tool for predicting and diagnosing various medical conditions using machine learning and NLP techniques. The chatbot interface will make it easy and convenient for patients to interact with the system, providing personalized recommendations and advice for better health outcomes.

Keywords: Artificial Intelligence, Data Mining, Deep Learning, Machine Learning, Data Processing, Disease Prediction.

Introduction:

The purpose of this project is to develop a machine learning-based system for disease prediction and a doctor's chatbot, with a focus on several vital organs and conditions such as kidney, eye, heart, lungs, breast, liver, and brain tumors. A. Lavanya Mathiyalagi, R. Mallika@pandeeswari, S. Srihari Seenivasan and Dr. R. Ravi (2021) stated that the advantages of cloud computing in healthcare are scalability of the required service and the provision to upscale or downsize the data storage collaborating with Artificial Intelligence [1].The goal of the project is to improve the accuracy of disease prediction and provide patients with quick and efficient access to medical advice and guidance, potentially improving their overall health outcomes.

The project will use a range of machine learning algorithms, including deep learning, to analyze medical

data and predict the likelihood of a particular disease or condition. G. Prince Devaraj, J. Zahariya Gabriel, R. Kabilan, J. Monica Esther, U. Muthuraman, and R. Ravi (2022) suggested a display design for accessible home control, emphasising on the use of home area networks to foster the independence of disabled individuals at home [2]. The system will be trained on large datasets of medical records and will learn to identify patterns and links between different diseases and their symptoms. The algorithms used will be optimized to provide the most accurate and reliable predictions possible.

The system will be designed to be user-friendly and accessible to patients of all ages and backgrounds. The doctor's chatbot component will serve as an interface between the patient and the system, allowing patients to input their symptoms and receive personalized advice and recommendations based on the analysis of their medical data.. R. Kabilan, R. Ravi, J. Monica Esther, U. Muthuraman, J. Zahariya Gabriel, and G. Prince Devaraj (2022) claimed that a reusable and resilient verification environment was necessary because it teaches people how to validate intellectual property and create an effective verification environment. Traditional verification and UVM-based verification were compatible in a SoC case study [3].

Edwin Raja S and Ravi R (2020) proposed to use the DMLCA approach to increase the detection accuracy utilizing a variety of factors, including detection accuracy based on true positive ratio, precision, and recall [4]. The system will be capable of predicting a range of diseases and conditions, including kidney disease, eye disease, heart disease, lung disease, breast cancer, liver disease, and brain tumors. The system will also be capable of providing recommendations for treatment and follow-up care based on the patient's medical history and symptoms. The likelihood of people developing many disease can be predicted using medical profiles including age, sex, blood pressure, and blood sugar.

The project will require a large amount of medical data to be collected and analyzed, which may require collaboration with healthcare providers and hospitals. The system will also need to comply with all relevant privacy regulations, ensuring that patient data is protected at all times. D. Priyadharshini and R. Ravi (2020) noted that there has been a late development in natural language processing. The deep learning research is still being conducted [5].

In summary, this project aims to develop a machine learning-based system for disease prediction and doctor's chatbot, focusing on several critical organs and conditions such as kidney, eye, heart, lungs, breast, liver, and brain tumor. The system will use various machine learning algorithms, including deep learning, to analyze a patient's medical history and symptoms, allowing it to make accurate predictions about the likelihood of a specific disease or condition. The doctor's chatbot component will serve as an interface between the patient and the system, allowing them to input their symptoms and receive personalized advice and recommendations based on the analysis of their medical data.

Algorithms:

In this project, a number of machine learning algorithms can be utilized to forecast ailments and create a chatbot for doctors. Some of the algorithms that can be used in this project are:

Logistic Regression: Logistic regression is a statistical method used for binary classification problems. It can be used to determine whether or not a patient has a particular disease.

Decision Tree: Decision trees are used to create models that predict the probability of a particular outcome. Decision trees can be used in this project to predict which disease a patient is likely to develop based on their medical history and symptoms.

Random Forest: Random forest is an ensemble method that uses multiple decision trees to improve the accuracy of predictions. It can be used in this project to predict diseases based on various patient parameters.

Support Vector Machine (SVM) is a technique for supervised learning that can be applied to classification or regression issues. SVM can be used in this project to classify patients based on their medical history and symptoms.

Artificial Neural Networks (ANN): ANN is a type of deep learning algorithm that mimics the structure and function of the human brain. It can be used to predict diseases and develop a doctor's chatbot based on a patient's medical history and symptoms.

Convolutional Neural Networks: CNN is a deep learning method for categorizing images. It can be used to analyze medical images such as X-rays, MRI scans, and CT scans,

and predict diseases based on these images. Using structured and unstructured hospital data, we suggest a new multimodal disease risk prediction algorithm based on convolutional neural networks (CNNs).

Recurrent Neural Networks (RNN): RNN is a deep learning algorithm used for sequence modeling. It can be used in this project to analyze a patient's medical history and symptoms and predict the likelihood of developing a specific disease.

These are some of the algorithms that could be applied to this project's disease prediction and chatbot doctor development. The choice of algorithm(s) will depend on the specific problem being addressed and the data available for training and testing the model.

Proposed System:

Development of Disease Prediction Model:-

The development of a disease prediction model involves several steps, including data collection, preprocessing, feature selection, model selection, and evaluation. Here are the general steps involved in developing a disease prediction model:

Data Collection: The first step in developing a disease prediction model is to collect relevant medical data. This data can be obtained from electronic health records, medical databases, or clinical studies. The data collected should include relevant medical history, demographic information, laboratory test results, imaging studies, and other clinical parameters.

Data Preprocessing: Once the data is collected, it needs to be preprocessed to remove any missing or irrelevant data. This involves data cleaning, normalization, and transformation. The engineering study of extracting previously unidentified patterns from a chosen set of data is known as data mining. In this study, Naive Bayes and J48 algorithms are evaluated in order to evaluate their performance and accuracy on training medical datasets.

Feature Selection: The next step is to select relevant features from the preprocessed data. This involves identifying the most important features that are highly correlated with the outcome variable (i.e., disease). This step helps reduce the number of features and improves the performance of the model.

Model Selection: The next step is to select the appropriate machine learning algorithm to build the disease prediction model. Several algorithms, as discussed earlier, can be used to predict diseases. The choice of algorithm(s) depends on the type of data and the specific problem being addressed.

Model Training: The selected model is then trained on the preprocessed data, using a supervised learning approach. The model is trained to learn the patterns and correlations between the input features and the outcome variable (i.e., disease).

Model Evaluation: Once the model is trained, it needs to be evaluated to determine its performance. This involves using various evaluation metrics, such as accuracy, precision, recall, and F1-score, to assess the model's performance.

Model Optimization: If the model's performance is not satisfactory, it may need to be optimized by tuning the model hyperparameters or changing the feature selection approach.

Deployment: Once the model is optimized, it can be deployed in the production environment to predict diseases based on new patient data.

In summary, the development of a disease prediction model involves several steps, including data collection, preprocessing, feature selection, model selection, training, evaluation, optimization, and deployment. The success of the model depends on the quality of the data collected, the choice of algorithm(s), and the performance evaluation metrics used.

Development of Doctor Chat Bot Model:-

Developing a doctor's chatbot using machine learning involves the following steps:

Data Collection: The first step in developing a doctor's chatbot is to collect data related to medical conditions, treatments, and diagnosis. This can be done by scraping medical websites, blogs, and forums. The data collected should include questions and answers related to medical conditions.

Data Preprocessing: Once the data is collected, it needs to be preprocessed to remove any irrelevant data, duplicate entries, and spelling errors. There are 2 types of data mining 1. Descriptive data mining; 2. Predictive data mining. The general characteristics of the data in the database are summarised or generalized via descriptive data mining. This step involves cleaning, normalization,

and transformation of the data.

Intent Classification: The next step is to classify the user's intent by analyzing their text input.

Response Generation: Once the user's intent is classified, the chatbot needs to generate an appropriate response. This involves using machine learning algorithms such as rule-based systems, decision trees, or neural networks to generate a response.

Dialogue Management: The next step is to manage the dialogue between the user and the chatbot. This involves keeping track of the conversation context, maintaining the chat history, and handling user inputs and queries.

Model Training: The selected model is then trained on the preprocessed data, using a supervised or unsupervised learning approach. The model is trained to learn the patterns and correlations between the input features and the output response.

Model Evaluation: Once the model is trained, it needs to be evaluated to determine its performance. This involves using various evaluation metrics such as accuracy, precision, recall, and F1-score, to assess the model's performance.

Model Optimization: If the model's performance is not satisfactory, it may need to be optimized by tuning the model hyperparameters, adding new features, or using a different algorithm.

Deployment: Once the model is optimized, it can be deployed in the production environment to interact with users and provide medical advice and support.

In summary, developing a doctor's chatbot using machine learning involves several steps, including data collection, preprocessing, intent classification, response generation, dialogue management, model training, evaluation, optimization, and deployment. This study examines the patient risk prediction issue within the framework of active learning, with comparatively similar results. Active learning has been thoroughly investigated and successfully used to address practical issues. Active learning techniques are typically used to explore absolute questions. In a medical application, the use of electronic health records (EHR) is used to forecast patients' propensity for a particular disease. The success of the chatbot depends on the quality of the data collected, the choice of algorithm(s), and the performance evaluation metrics used.

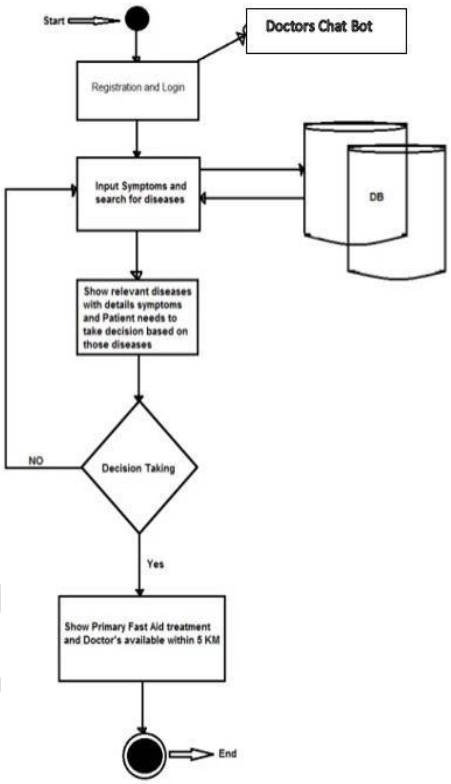

Figure:1 Flow diagram of the disease predictor and doctors chatbot.

Figure 1 shows the flow of the program. It starts with registering and login in with the database and then inputting the symptoms or the scan reports. Then it searches and compares with the data set used and predicts the disease**.**

Result and Discussion:

The result and discussion for a disease predictor and doctor's chatbot using machine learning project will depend on the specific implementation and performance of the system. Here are some general points that may be discussed:

Disease Predictor:

Accuracy: The accuracy of the disease predictor model on the test data can be evaluated to determine how well the model performs.

Feature importance: The most important features that contribute to the prediction can be identified, which can provide insight into the factors that are most influential in the disease diagnosis.

False positives/negatives: The number of false positives and false negatives can be discussed to understand how well the model performs in correctly diagnosing patients and avoiding misdiagnoses.

Doctor's Chatbot:

Intent classification accuracy: The accuracy of the chatbot in correctly classifying user intent can be evaluated to determine how well the NLP techniques perform.
Response generation: The appropriateness and relevance of the responses generated by the chatbot can be discussed, and any errors or areas for improvement can be identified.
Dialogue management: The flow and management of the conversation between the user and chatbot can be evaluated to determine how well the chatbot maintains context and keeps the conversation on track.
User feedback: Any feedback received from users on the chatbot's performance and suggestions for improvement can be discussed and incorporated into future versions of the chatbot.
Overall, the result and discussion for a disease predictor and doctor's chatbot project will aim to evaluate the performance of the system and identify areas for improvement. By analyzing the accuracy, feature importance, false positives/negatives, intent classification accuracy, response generation, dialogue management, and user feedback, the project team can identify areas for improvement and make updates to the system to enhance its performance.

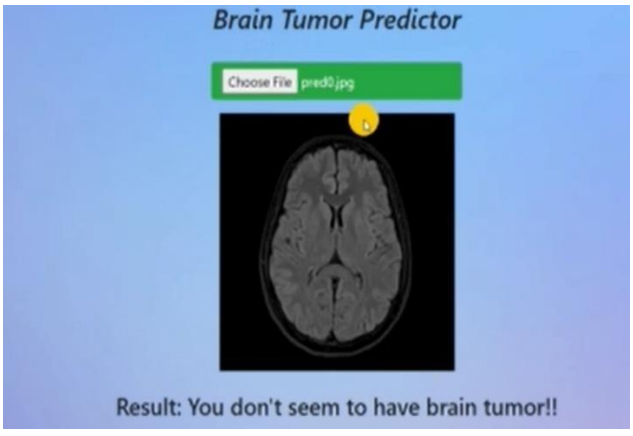

Figure:2 Output of the brain tumor predictor

Figure 2 shows the interface of the brain tumor predictor. This figure shows that this particular scan has no brain tumor.

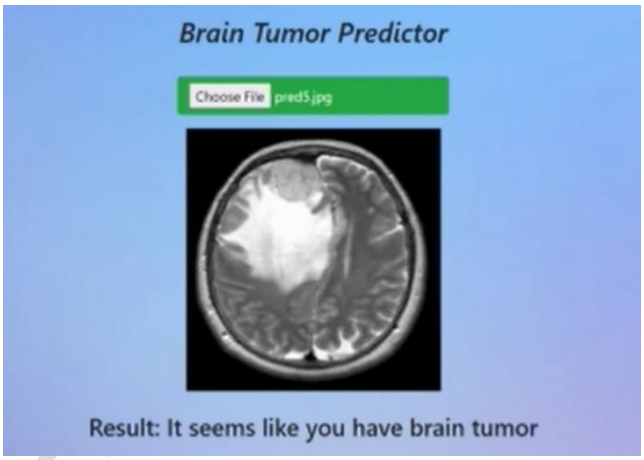

Figure:3 Output of the brain tumor predictor predicting the tumor

Figure 3 shows the interface of the brain tumor predictor. This figure shows that this particular scan has brain tumor.

Figure:4 Output of Breast Cancer Predictor

Figure 4 shows the interface of the breast cancer predictor. When you input the necessary details it compares with the data set and produces the result.

Conclusion:

In conclusion, a disease predictor and doctor's chatbot using machine learning project can have significant potential in improving healthcare services. By leveraging the power of machine learning algorithms and NLP techniques, these systems can aid in disease diagnosis and provide a convenient and accessible way for patients to receive medical advice.

The disease predictor can provide accurate and reliable predictions of various diseases, helping doctors and healthcare professionals to make informed decisions and improving patient outcomes. The doctor's chatbot can provide quick and convenient medical advice, reducing the burden on healthcare systems and providing patients with instant access to information and advice.

However, it is important to note that these systems are not a replacement for medical professionals. They are intended to aid in medical decision-making and provide support to patients, but should not be relied upon as a substitute for professional medical advice.

Overall, the development of disease predictors and doctor's chatbots using machine learning can contribute to the improvement of healthcare services and provide patients with more accessible and convenient access to medical advice. With ongoing refinement and development, these systems have the potential to become valuable tools in the healthcare industry.

Reference:

1. A. Lavanya Mathiyalagi, R. Mallika@pandeeswari, S. srihari Seenivasan and Dr. R. Ravi, “Securing Data using Deduplication in E-Healthcare System”, International Journal on Engineering Technology and Sciences, vol. 8, no. 9, pp. 30-34, 2021.

2. G. Prince Devaraj, J. Zahariya Gabriel, R. Kabilan, J. Monica Esther, U. Muthuraman, and R. Ravi, “ Multipurpose Intellectual Home Area Network Using Smart Phone”, IEEE Proceedings of the Second International Conference on Artificial Intelligence and Smart Energy, pp.1464-1469, 2022.

3. R. Kabilan, R. Ravi, J. Monica Esther, U. Muthuraman, J. Zahariya Gabriel, and G. Prince Devaraj, “Constructing Effective UVM Testbench By Using DRAM Memory Controllers”, IEEE Proceedings of the Second International Conference on Artificial Intelligence and Smart Energy, pp.1034-1038, 2022.

4. Edwin Raja S and Ravi R, “A performance analysis of Software Defined Network based prevention on phishing attack in cyberspace using a deep machine learning ith CANTINA approach(DMLCA)”, Computer Communications, vol. 152, pp.0-6, 2020.

5. D. Priyadharshini, and R. Ravi, “Deep learning: a survey and techniques for language processing, image, speech and text”, Francis Xavier Journal of Science Engineering and Management, vol. 1, no. 1 , pp.11-14, 20