



Machine Learning Based Disease Prediction from User Symptoms

Ms. P. Yaswanthi (M.C.A), Rajeev Gandhi Memorial college Of Engineering and Technology, Nandyal

* Mr. V. V. Nagendra Kumar, M. Tech, (Ph.D.), Rajeev Gandhi Memorial College of Engineering and Technology, Nandyal

Abstract

The requirement for health information is altering knowledge-seeking behavior, which should be noted globally. Many of us struggle with finding health information online about illnesses, diagnoses, and various treatments. It will save a lot of time if a suggestion system is frequently used for physicians and medications. Because the users of a system like this are laypeople, the user has several difficulties understanding the basic medical terminology. The customer is perplexed since there is an excessive amount of medical information available in various formats. A method called Disease Prediction using Machine Learning forecasts the disease based on the data the user provides. Additionally, it accurately forecasts the user's or the patient's disease based on the data or symptoms entered into the system and returns findings accordingly. If the customer simply wants to know the type of ailment the patient has experienced and the condition is not particularly significant.

Keywords: Random Forest Algorithm, Naive Bayes, Support Vector Machine, Logistic regression etc.

1. INTRODUCTION

1.1 Introduction

A method called Disease Prediction using Machine Learning forecasts the disease based on the data the user provides. Additionally, it accurately forecasts the user's or the patient's disease based on the data or symptoms

entered into the system and returns findings accordingly. If the customer simply wants to know the type of ailment the patient has experienced and the condition is not particularly significant. It is a system that gives users advice on how to keep their health systems in good shape and offers a technique to identify diseases using this prediction. The health sector now plays a significant role in treating patients' illnesses, hence this is The user can understand the disease they are affected by by simply entering the symptoms and any other useful information, and the health industry can also benefit from this technique by simply asking the user for their symptoms and entering them into the system and in just a few short minutes. This is useful for the user in case they don't want to travel to the hospital or other clinics. It is a system that gives users advice on how to keep their health systems in good shape and offers a technique to identify diseases using this prediction. The health sector today is very important in treating people' illnesses, so Simply by entering the symptoms and any other helpful information, the user can understand the disease he or she is affected by, and the health industry can also benefit from this technique by just asking the user to provide the symptoms and entering them in the system and in just the same way that it would if they were going to a hospital or other clinic. Programming computers to perform better using example data or historical data is known as machine learning. The study of computer systems that learn from data and experience is known as machine learning. The training and testing tracks of a machine learning algorithm. Illness diagnosis based on

the symptoms and medical history of the patient. Machine learning technology has improved throughout the years. The medical profession now has an incomparable platform thanks to machine learning technology, making it possible to handle healthcare challenges quickly. Machine learning is being used to keep entire hospital data. With the aid of machine learning technology, which enables constructing models to analyze data quickly and give answers more quickly, clinicians may more accurately diagnose and treat patients, which improves patient outcomes.

2. LITERATURE SURVEY

• [1] **Colorful machine learning algorithms were streamlined for the effective prophecy of a habitual complaint outbreak by Chen et al:** created a statistical model to predict whether an illness was caused by the flu or not. They included 3744 adult and adolescent cases of influenza who were unvaccinated and had fever and at least two additional influenza symptoms. Out of 3744 people, 2470 had influenza that was confirmed by a lab. Based on these facts, their model predicted a delicacy of 79. The information gathered for the purpose of training was insufficient. A model with idle factors was applied to get around this. A novel multimodal complaint trouble vaticinator (CNN-MDRP) with a convolutional neural network foundation was created. The algorithm arrived at a delicateness of roughly 94.8%.

[2] **The DNN model performed more in terms of average performance and the LSTM model gave close prognostications when circumstances were large. Haq et al:** utilized a database that had data on patients with any heart-related complaints. They used the least absolute loss and selection motorist, relief, minimum redundancy and maximum connection (mRMR), and these three selection algorithms to extract features, which were then cross-verified by the K-fold system. The extracted features were fed into six different machine learning algorithms, and they were also categorized according to whether or not there was a cardiac complaint. Employing ML techniques, classified the diabetic complaint. Using logistic regression (LR), the problem variables for diabetes complaints were found. The ML-based system's overall delicacy was 90.62%. The system evaluates the user's symptoms as input and outputs the likelihood that the disease will occur. Utilizing a decision tree classifier, disease prediction is accomplished. The probability of the sickness is calculated by a decision tree classifier. Accurate medical data analysis aids in early disease identification and patient care as big data usage increases in the biomedical and healthcare sectors.

3. OVERVIEW OF THE SYSTEM

3.1 Existing System

Since the advent of powerful computing, doctors have continued to use technology to perform numerous procedures, such as x-ray photography and surgical representation. However, technological advancements have perceptually lagged behind. Due to other aspects including weather, atmosphere, blood pressure, and several other parameters, the approach still requires the doctor's knowledge and expertise. Although there are a great deal of factors that are acknowledged as being necessary to comprehend the total functioning process, no model has ever been able to analyze them adequately. Medical decision support systems must be employed to address this problem. The doctors can use this technique to help them choose wisely.

3.1.1 Disadvantages of Existing System

Low Accuracy
High complexity.
Highly inefficient.
Requires skilled persons

3.2 Proposed System

Using symptoms, this approach is used to forecast disease. This system evaluates the model using a decision tree classifier. End users make use of this system. Based on symptoms, the system will be able to forecast disease. The technology used by this system is machine learning. The decision tree classifier method is used to forecast diseases. This system is known as "AI Therapist" by us. This system is designed for those people who are constantly worried about their health, so we have included certain features that recognize them and improve their mood as well. As a result, the function "Disease Predictor" for health awareness can identify diseases based on their symptoms.

3.3 Methodology

In this project work has two modules

Admin:

1. System:

Pre-processing:

In this step data cleaning and data filling is done.

Training:

Use the pre-processed training dataset is used to train our machine learning algorithms.

Generate accuracy

System generates accuracy for our model and dataset. This tells us how much efficiently model is working.

Generates results:

The results will be displayed are which type job class.

2. Patient:**Data Collection**

The user has to upload an image which needs to be classified.

Model Building

User builds the models to fit our data for prediction of job class.

View Accuracy

Users view the generated accuracy from the system.

View Results

Users can view the generated classification from the user.

4 Architecture

Dataset is available in the system and when the user uploads the data and, in this project, we are using machine learning and the algorithm is decision tree and the patients are coming to check their disease so for the sake we are using training and dataset provide we should entire the data.

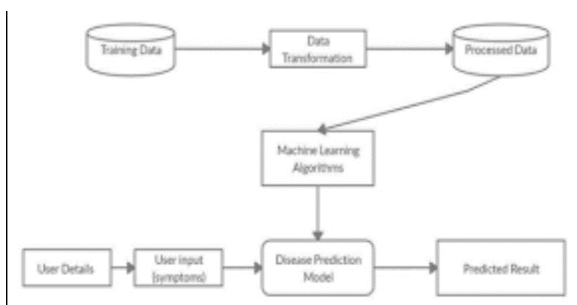


Fig 1: Frame work of proposed method

By using these data, we should now which disease we are facing and the patients have the training periods and the testing for patients we now the condition of

patients. The testing will give the results of predicting personality. System module is used to pre-process the data .it is used to clean the data and the data filling can be done. System generates accuracy for our model and dataset. This tells us how much efficiently model is working.

To build a machine learning model for symptom-based disease prediction, you can follow the steps outlined below:

1. **Data Collection:** Gather a large and diverse dataset that contains information about symptoms and corresponding disease diagnoses. This data can be obtained from healthcare databases, medical records, or even through surveys and online platforms.
2. **Data Preprocessing:** Clean and preprocess the collected data to ensure it is in a suitable format for machine learning algorithms. This step may include handling missing values, removing outliers, normalizing numerical features, and encoding categorical variables.
3. **Feature Selection/Extraction:** Identify relevant features from the dataset that will help in predicting diseases based on symptoms. This can be done using techniques such as correlation analysis, feature importance, or domain knowledge.
4. **Model Selection:** Choose an appropriate machine learning algorithm for disease prediction. Depending on the nature of the problem, you can consider using algorithms like logistic regression, support vector machines (SVM), random forests, or gradient boosting.

The Information about in Datasets:

In my project. the dataset we have the attributes 193 and we have the implemented the diseases smoking, vomiting, fatigue, weight loss, muscle pain, stomach pain, skin rash, itching, anxiety, Shivering. chills, smoothing, lethargy etc..... we have 193 attributes and the cells we have 4920.

Mainly we are splitting the percentage into 20 percentage then we have the 23 attributes and we can collect the data 132456 and the training algorithm are decision tree, random forest, lvd these are the algorithm with accuracy we have the decision tree is 65.99. random forecast is accuracy is having 100 and the lvd algorithm accuracy is 99.98. we have the count 64553 and having the attributes are 193. After training the prediction should be have the data should be present in yes and no type and then submit the file in data. then you can get the result.

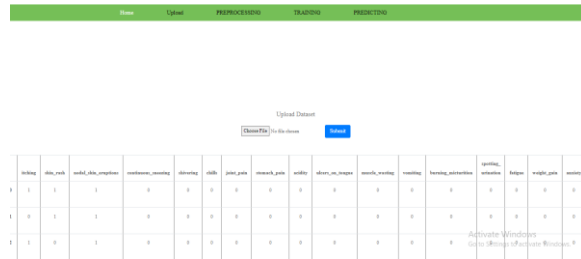
Sl. No	Splitting	Decision Tree	Random Forest	Support Vacca Method
1	0-10	65.99	59.98	67.99
2	10-20	78.95	86.56	76.56
3	20-30	98.56	99.56	67.98
4	30-40	67.98	89.67	56.89
5	40-50	56.98	67.98	56.87
6	50-60	68.96	98.67	67.95
7	60-70	56.98	76.97	89.67
8	70-80	99.65	67.98	78.96
9	80-90	67.89	87.65	78.90
10	90-100	78.98	78.98	89.65

5. RESULTS SCREEN SHOTS

Home Page:



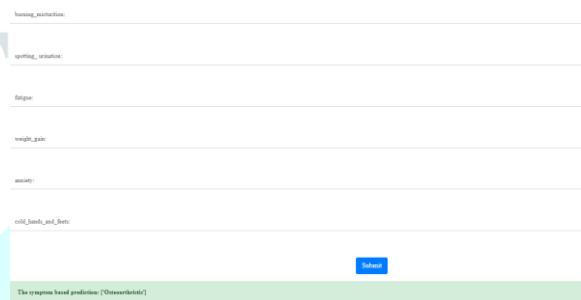
Upload image:



Choose options:



Predict Result:



7. CONCLUSION

I have therefore concluded that machine learning can be utilized to track our health in an efficient manner. We can periodically check our health for free and stay healthy. After constructing the machine learning model, I deployed it using Flask (a Python web framework), and by converting that domain into a website in the future, it will be freely accessible to everyone. In order for our model to forecast the optimal outcome, the user only needs to visit the relevant page and choose 5 to 8 diseases. After receiving the prediction, the user will gain insight into their health and, if necessary, contact the appropriate doctor. Any person in our planet can become healthy person.

Future Enhancement

Everyone wants to have a good healthcare system, and doctors are expected to be knowledgeable and considerate in all situations. However, it is highly improbable to review all the information, medical history, and documents required for each circumstance. Even if they have a

vast amount of data and information, comparing and analyzing all of the illnesses' symptoms in order to predict their outcomes is delicate. Therefore, incorporating facts into a case's unique profile and doing in-depth research are outside of the purview of a croaker. As a result, the outcome is always described as a supported healthcare plan that was simply created for an actual person. By evaluating actual data, prophetic analytics is the process of making predictions about the future. For medical care, making the fashionable opinions in the case of every existent would be possible. Using artificial intelligence, prophetic modeling creates a vatic nation from previously recorded data, trends, individuals, and conditions. The model is set up such that a new existence can obtain a vatic nation without difficulty. These predictive algorithms can be used by Medicare and Health departments to determine when a patient can be released safely.

8. References

[1] Pingale, Kedar, et al. "Disease Prediction using Machine Learning." (2019). Mr. Chala Bayne, Prof. Pooja Kamat, "Survey on Prediction and Analysis the Occurrence of Heart Disease Using Data Mining Techniques", International Journal of Pure and Applied Mathematics, 2018.

[2] Pingale, K., Surwase, S., Kulkarni, V., Sarage, S., & Karve, A. (2019). Disease Prediction using Machine Learning.

[3] Aiyasha Sadiya, Differential Diagnosis of Tuberculosis and Pneumonia using Machine Learning (2019)

[4] S. Patel and H. Patel, "Survey of data mining techniques used in healthcare domain," Int. J. of Inform. Sci. and Tech., Vol. 6, pp. 53-60, March, 2016.

[5] Rajeswari D and Kumar, "Analyzing Prognosis Methods using Machine Learning Algorithms for Detecting COPD", International Conference on Applied Artificial Intelligence and Computing (ICAAIC 2022), ISBN: 978-1-6654-9709-1, Pages 595-602

[6] Balasubramanian, Satyabhama, and Balaji Subramanian. "Symptom based disease prediction in medical system by using Kmeans algorithm." International Journal of Advances in Computer Science and Technology 3.

[7] Dhenakaran, K. Rajalakshmi Dr SS. "Analysis of Data mining Prediction Techniques in Healthcare Management System." International Journal of Advanced Research in Computer Science and Software Engineering 5.4 (2015).

[8] Maniruzzaman, M., Rahman, M., Ahammed, B. and Abedin, M., 2020. Classification and prediction of diabetes disease using machine learning paradigm. Health information science and systems, 8(1), pp.1-14.

[9] Chen, M., Hao, Y., Hwang, K., Wang, L. and Wang, L., 2017. Disease prediction by machine learning over big data from healthcare communities. Ieee Access, 5, pp.8869-8879.

[10] Haq, A.U., Li, J.P., Memon, M.H., Nazir, S. and Sun, R., 2018. A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms. Mobile Information Systems, 2018.