INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

e-ISSN : 2583-1062

Impact Factor : 5.725

www.ijprems.com
editor@ijprems.com

Vol. 03, Issue 10, October 2023, pp : 357-360

HEART DISEASE PREDICTION WITH MACHINE LEARNING HEART DISEASE PREDICTION WITH MACHINE LEARNING USING PYTHON

Mr. Aniket Navale[1], Mr. Siddharth Sapkal[2], Mrs. Swapna Patil[3]

[1,2,3]Student, Department of Electronics and Telecommunication Engineering, Shree L. R. Tiwari College of Engineering, India.

ABSTRACT

Heart disease is a leading cause of mortality worldwide, emphasizing the need for accurate and timely diagnosis. Machine learning techniques have shown promise in predicting heart disease risk, and this study focuses on the application of the Machine Learning algorithms for heart disease prediction.

We explore a wide range of machine learning techniques, including decision trees, support vector machines, random forests, and K-Nearest Neighbour to assess their performance in heart disease prediction. Our study involves data preprocessing, feature selection, and model evaluation to ensure the robustness and accuracy of the predictions. The results reveal promising outcomes in terms of predictive accuracy, specificity. We compare the performance of these algorithms and discuss their strengths and limitations in the context of heart disease prediction. In this system, to train and test the Machine Learning algorithms, we use a dataset containing clinical and demographic data. Using these algorithms, we divide patients into two categories: those who are at risk of heart disease and those who are not.

1. INTRODUCTION

One of the leading causes of morbidity and mortality among the global population is heart disease. One of the most crucial topics in the data analysis area is predicted cardiovascular disease. Since a few years ago, the prevalence of cardiovascular disease has been rising quickly throughout the world. Many studies have been carried out in an effort to identify the most important risk factors for heart disease and to precisely estimate the overall risk. Heart disease is also referred to as a silent killer because it causes a person to pass away without any evident signs. Cardiovascular disease must be detected early. aiding high-risk patients in making decisions regarding lifestyle changes will help to reduce the difficulties. Making choices and predictions from the vast amounts of data generated by the healthcare sector is made easier with the help of machine learning. By evaluating patient data that uses a machine-learning algorithm to categorise whether a patient has heart disease or not, this study hopes to predict future cases of heart disease. Machine learning methods can be extremely helpful in this situation. There is a common set of basic risk factors that determine whether or not someone will ultimately be at risk for heart disease, despite the fact that heart disease can manifest itself in various ways. By gathering information from numerous sources, organising it into categories that make sense, and then performing analysis to get out the desired information based on statistics, we may conclude that this technique is quite adaptable. The main difficulty with heart disease is detecting it. There are tools that can forecast heart disease, but they are either expensive or ineffective at calculating the likelihood of heart disease in a human. The mortality rate and total consequences can be reduced by early identification of heart disorders. Since it takes more intelligence, time, and knowledge, it is not always possible to accurately monitor patients every day, and a doctor cannot consult with a patient for a whole 24 hours. As can use a variety of machine learning methods to search for hidden patterns. The underlying patterns may utilised in medical data for health diagnosis.

Problem Definition- The major challenge in heart disease is its detection. There are instruments available which can predict heart disease but either they are expensive or are not efficient to calculate chance of heart disease in human. Early detection of cardiac diseases can decrease the mortality rate and overall complications. However, it is not possible to monitor patients every day in all cases accurately and consultation of a patient for 24 hours by a doctor is not available since it requires more sapience, time and expertise. Since we have a good amount of data in today's world, we can use Machine Learning Algorithms to analyze and train the data.

2. LITERATURE SURVEY

These literature surveys explore various applications of machine learning for heart disease prediction. The studies focus on the significance of leveraging machine learning techniques to address the critical issue of heart disease, a major cause of mortality globally.

- Heart Disease Prediction Using Machine Learning- The multi-layer perceptron neural network approach is used for dataset training and testing in the study. "Prediction of Heart Disease using Machine Learning algorithm" that Aditi Gavhane proposed. There will be one input layer , one output layer and perhaps more hidden layers in this algorithm between the two input and output layers.

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

e-ISSN : 2583-1062

Impact Factor : 5.725

www.ijprems.com
editor@ijprems.com

Vol. 03, Issue 10, October 2023, pp : 357-360

- Heart Disease Prediction Using Machine Learning- The proposed methodology is also critical in a healthcare organization with experts that have no more knowledge and skill. It uses different medical attributes such as blood sugar and heart rate, age, sex are some of the attributes are included to identify if the person has heart disease or not. Analyses of the dataset are computed using WEKA software.
- Heart Disease Prediction Using Machine Learning- Random Forest algorithm is an efficient algorithm which is an ensemble learning method for regression and classification techniques. The algorithm constructs N of Decision trees and outputs the class that is the average of all decision trees output. So accuracy of prediction at early stages is achieved effectively.
- Heart Disease Prediction Using Machine Learning- To achieve better accuracy and to make ththe more efficient so that it can predict the chances of heart attack.

SOFTWARE REQUIREMENT

- **Google colab :** Google Colaboratory, or Colab, is an as-a- service version of Jupyter Notebook that enables you to write and execute Python code through your browser.
- **Python Libraries** :

1. Numpy **:** A library for the Python programming language called NumPy adds support for big, multi-dimensional arrays and matrices as well as a tonne of high-level mathematical operations that can be performed on these arrays.
2. Matplotlib **:** For the Python programming language and its NumPy numerical mathematics extension, Matplotlib is a graphing library like Tkinter, wxPython, Qt, or GTK, it offers an object-oriented API reference to both "Panel Data", and "Python Data Analysis" and was created by Wes McKinney

FLOWCHART

- Seaborn :- A matplotlib-based Python data visualisation library is called Seaborn. It offers a sophisticated user interface for creating visually appealing and educational statistical visuals. Python's Seaborn package is mostly used to create statistical visuals.
- SciPy:- SciPy includes modules for common tasks in science and engineering like optimisation, linear algebra, integration, interpolation, special functions, FFT, signal and image processing, and ODE solvers.
- Pandas : Pandas is a Python library used for working with data sets. It has functions for analyzing, cleaning, exploring, and manipulating data. The name "Pandas" has

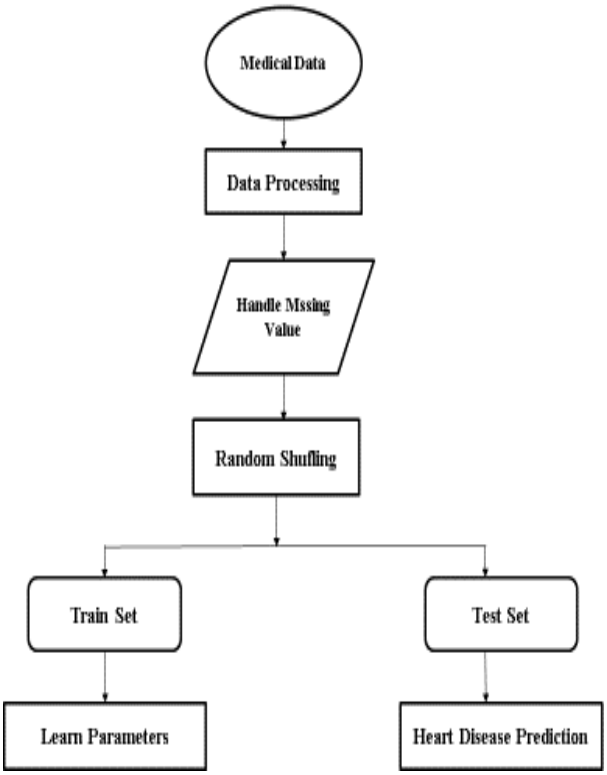


Figure 1: Flowchart of Proposed System

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

e-ISSN : 2583-1062

Impact Factor : 5.725

www.ijprems.com
editor@ijprems.com

Vol. 03, Issue 10, October 2023, pp : 357-360

DESCRIPTION

Start **:** This is the beginning of the process, indicating that we are about to embark on the journey of creating a heart disease prediction model.

Medical Data **:** Medical data is collected, which can include a wide range of information about patients. This data may consist of attributes such as age, gender, cholesterol levels, blood pressure , family history, lifestyle factors, and more. Data Preprocessing**:** Data preprocessing involves several tasks to prepare the collected medical data for machine learning. This can include cleaning the data to remove errors and feature extraction, where relevant information is derived from the raw data inconsistencies, normalizing features to bring them to a common scale

Handle Missing Values**:** In real-world datasets, there are often missing values. Handling missing values may involve techniques like imputation, where missing values are filled in based on other data, or in some cases, rows with missing values may be removed.

Random Shuffling**:** To avoid any potential biases in the dataset, the data is randomly shuffled. This step helps ensure that the order of data points doesn't impact the performance of the machine learning model.

Splitting Data into Train and Test Sets**:** The dataset is divided into two subsets: the training set and the test set. The training set is used to train the machine learning model, while the test set is used to evaluate its performance. For Train Set, It Learns Parameters**:** Using the training set, the machine learning algorithm is applied to learn the parameters that define the model. This process involves adjusting the model's internal settings to make accurate predictions based on the retraining data Test Set, Heart Disease Prediction: Once the model has been trained, it is used to make predictions on the test set. In the context of heart disease prediction, the model uses the information in the test set to predict whether or not a patient is at risk of heart disease.

3. RESULTS

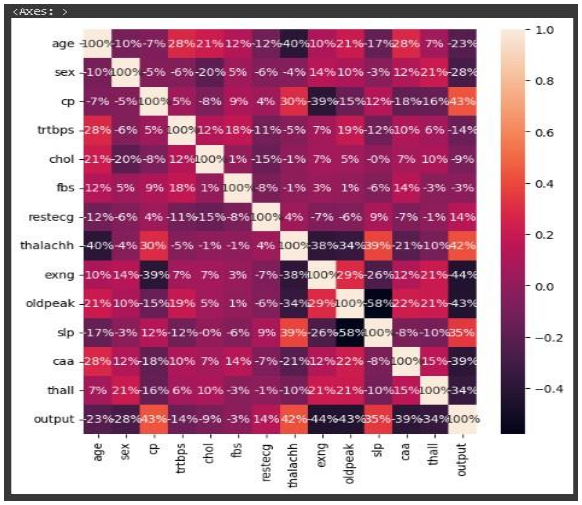
This is the parameters which are use to predict the heart disease.

	age	sex	cp	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak	slp	caa	thall	output
0	63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
1	37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
2	41	0	1	130	204	0	0	172	0	1.4	2	0	2	1
3	56	1	1	120	236	0	1	178	0	0.8	2	0	2	1
4	57	0	0	120	354	0	1	163	1	0.6	2	0	2	1

This is the parameters description table. In this, we find the mean, standard deviation , min, max

	age	sex	cp	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak	slp	caa	thall	output
count	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000
mean	54.366337	0.683168	0.966997	131.623762	246.264026	0.148515	0.528053	149.646865	0.326733	1.039604	1.399340	0.729373	2.313531	0.544554
std	9.082101	0.466011	1.032052	17.538143	51.830751	0.356198	0.525860	22.905161	0.469794	1.161075	0.616226	1.022606	0.612277	0.498835
min	29.000000	0.000000	0.000000	94.000000	126.000000	0.000000	0.000000	71.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	47.500000	0.000000	0.000000	120.000000	211.000000	0.000000	0.000000	133.500000	0.000000	0.000000	1.000000	0.000000	2.000000	0.000000
50%	55.000000	1.000000	1.000000	130.000000	240.000000	0.000000	1.000000	153.000000	0.000000	0.800000	1.000000	0.000000	2.000000	1.000000
75%	61.000000	1.000000	2.000000	140.000000	274.500000	0.000000	1.000000	166.000000	1.000000	1.600000	2.000000	1.000000	3.000000	1.000000
max	77.000000	1.000000	3.000000	200.000000	564.000000	1.000000	2.000000	202.000000	1.000000	6.200000	2.000000	4.000000	3.000000	1.000000

This is a confusion matrix.



INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

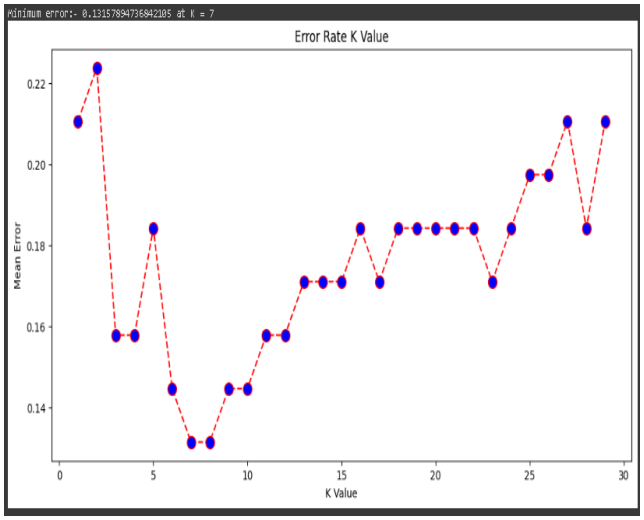
e-ISSN : 2583-1062

Impact Factor : 5.725

www.ijprems.com
editor@ijprems.com

Vol. 03, Issue 10, October 2023, pp : 357-360

This graph shows the optimal "k" value with the minimum error rate. The minimum error rate with k = 7 is 0.1315.



4. CONCLUSION

The project on heart disease prediction using the various Machine Learning algorithms has provided valuable insights into the potential of machine learning in healthcare.

Through the analysis of a relevant dataset and the implementation of the machine learning algorithms, we have demonstrated that it can be an effective tool for predicting heart diseases with reasonable accuracy.

5. REFERENCES

[1] Mr. Valle Harsha Vardhan, Mr. Uppala (2023) Rajesh Kumar Heart Disease predictio System using Machine Learning algorithms.

[2] Mr. Santhana Krishnan J , Dr.Geeta S Prediction of Heart Disease using KNN algorithm – International Journal of Recent Technology and Engineering (IJRTE)

[3] M. Marimuthu Gayatri. R S. Deivarani “Analysis of heart disease prediction using KNN algorithms – Journal of Engineering sciences.

[4] Prof. Dr. Sulochana Sonkamble Snehal Dange “Hear Disease Prediction system using SVM algorithm.”- International Journal of Recent Technology and Engineering (IJRTE)

[5] M. Senith Raja M. Anurag (2021) Machine learning based Heart Disease prediction system

[6] Baban. U. Rindhe, Nikita Ahire, Rupali Patil (2021) Heart Disease Prediction Using Machine Learning.- “International Journal of Recent Technology and Engineering (IJRTE).