International Journal of Engineering Inventions
e-ISSN: 2278-7461, p-ISSN: 2319-6491
Volume 12, Issue 6 [June. 2023] PP: 216-218

Heart Disease Prediction Using Machine Learning

[1]Ms.Kavitha.G, [2]Dr C.Kumuthini

[1]II Msc Computer Science, Dr.N.G.P Arts and Science College,Coimbatore-48,Tamil Nadu, India.
[2]Professor, Department of Computer Applications,Dr.N.G.P Arts and Science College, Coimbatore-48,Tamil Nadu, India.

ABSTRACT:
Heart disease is a major health concern globally. Early diagnosis and preventative measures can help reduce mortality and improve overall health outcomes. Machine learning provides a powerful tool that can be used to identify patterns and help predict the risk of heart disease. This abstract describes a machine learning approach to predict the presence or absence of heartdisease in a dataset of patient records. The dataset contains patient demographics, medical history, lifestyle, and laboratory test results. The machine learning algorithms employed include logistic regression, support vector machines, decision trees, and neural networks. The results of the machine learning predictions were evaluated using a variety of performance metrics. The results indicated that the machine learning approach was successful in predicting the presence or absence of heart disease in the patient dataset. The results suggest that machinelearning can be a useful tool for clinicians in the early detection and prevention of heart disease.

KEYWORDS: *Machine Learning, Supervised learning, Data Analysis*

Date of Submission: 16-06-2023 Date of acceptance: 02-07-2023

I. INTRODUCTION:

The heart is the human body's most vital organ. The heart's job is to circulate blood throughout the body via pumping blood. The biggest problem facing the medicalsector today is to provide higher level facilities to health infrastructure so that diseases can be detected early on and treated promptly to improve the quality of life through quality of service[1].It is difficult for doctors in poor and underdeveloped nations to diagnose diseases at an early stage and prevent complications, which increases mortality.

Patients of all income levels have benefited from the development of information and communications technology since it allowspatients to receive real-time information at lower diagnostic and monitoring costs the patients' health.

The research has access to a massive collection of medical records. The use of the massive medical data presents the medical industry with enormous challenges. Machines quickly transform the enormous amount of data to produceaccurate and useful information. Thus, machine learning is the important area.

Machine learning can forecast heart disease more accurately, which will aid healthcare professionals in patient diagnosis and treatment in early stages help many individuals quickly diagnose their illnesses.

II. LITERATURE REVIEW:

One of the most popular databases in this field of study is the UCI Heart DiseaseDataset from the UCI Machine Learning Repository [2].

Another commonly used dataset is Statlog [3]. Such ML models seek to increase precision and lower overall computationcosts in the clinical detection of diseases.

Verma et al. (2016), for instance, suggested a hybrid model for the prediction of heart disease using particle swarm optimisation (PSO) and two machine learning classifiers, namely K-nearest neighbour and multi- layer perceptron (MLP), which obtained a 90.28% accuracy [4].

For the classification of heart disease, FirdaAnindita Latifah et al. suggested a comparative study of two machine learning models, logistical regression and random forest. The study used the 3656-record Framingham dataset and a 70:30 training toassessment ratio. The model was able to attain an accuracy of 85.04% [5].

In 2015 a research is done by Nguyen Cong Long et al. on disease prediction using firefly algorithm. The classifier is trained by using rough set theory. The results are compared with other classification techniques such as Naïve Bayes and SVM.Proposed work overcomes convergence speed, processing time and increaseaccuracy to 87.2%. Limitation of this study is that rough set attribute is unmanageable when there is large number of

attributes [6].

EXISTING SYSTEM:

The existing system modules generates comprehensive report by implementing the strong prediction algorithm[9].The main aims of the existing system to compare and check the before patient whose having disease outputs and new patient illnesses are determined, as are the likelihood that a specific patient may develop heart disease in the future.

Implementing the aforementioned model will help us achieve our aim of creating a system that is more accurate at predicting the likelihood that a new patient will experience a heart attack. The model which is proposed for Heart Attack Prediction System is invented for using Deep learning algorithms and approach[7]. But the accuracy is relatively low whenemploying all the current systems.

PROPOSED SYSTEM:

This proposed system have a data which classified if patients have heart disease or not according to features in it. This proposed system can try to use this data to create a model which tries predict if a patient has this disease or not. Use the logistic regression (classification) technique in the suggested system. By usingsklearn library to calculate score.Implements Naive Bayes algorithm togetting accuracy result.

From the data we are having, it should be classified into different structured data based on the features of the patient heart. We must build a model that, using the logistic regression technique, predicts the patient's disease based on the data[8].First, we have to import the datasets. Read the datasets, the data should contain different variables like age, gender, sex, cp(chestpain),slope, target. Create a temporary variable and also build a model for logisticregression. By using logistic regression, naive bayes the accuracy rate increases.

MODULE DESCRIPTION:

The model generation subsequently occurs, followed by analysis of the results. Each step of the model is discussed in detail in subsequent sections.

Data acquisition:

The cardiac disease dataset obtained fromthe Kaggle. It contains 14 features and 1026 records.

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slop	ca
0	63	1	1	145	233	1	2	150	0	2.3	3	0
1	67	1	4	160	286	0	2	108	1	1.5	2	3
2	67	1	4	120	229	0	2	129	1	2.6	2	2
3	37	1	3	130	250	0	0	187	0	3.5	3	0
4	41	0	2	130	204	0	2	172	0	1.4	1	0
...	...	...	...	...	...	...	...	...	...	...	...	...
292	57	0	4	140	241	0	0	123	1	0.2	2	0
293	45	1	1	110	264	0	0	132	0	1.2	2	0
294	68	1	4	144	193	1	0	141	0	3.4	2	2
295	57	1	4	130	131	0	0	115	1	1.2	2	1
296	57	0	2	130	236	0	2	174	0	0.0	2	1

297 rows × 14 columns

Data pre-processing:

Cardiovascular disease UCI dataset is first loaded and then data cleaning and finding missing values was performed on allrecords. The dataset contains complete information.

Splitting dataset:

The splitting of the dataset in the ratios oftraining and testing set in percentile.

Classification Algorithms:

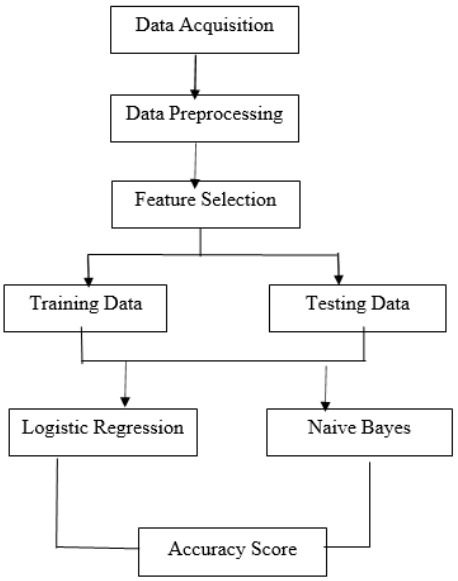
Following classification algorithms arethen applied on the pre-processed dataset.

Logistic Regression:

Logistic regression is assessing the parameters of logistic model in regression analysis.

Naive Bayes:
Naïve Bayes classifiers perform classification by calculating the probability of given dataset. Each attribute in given data is considered as independent of other. High probability class is the output class ofgiven instance.

ARCHITECTURAL DESIGN:



III. CONCLUSION:

Heart disease prevalence can rise above the control line and reach its highest point. Heart disease is a difficult condition that claims the lives of many people every year. One of the main weaknesses of these works, which use all of these systems, is that they primarily concentrate on the application of classify techniques and algorithms for heart disease prediction, while all of these systems study different data cleaning and mining techniques that prepare and build a dataset suitable for data mining. So that I can use machine learning to predict if a patient has heart disease or not using logistic regression techniques. Thissoftware can be used by any non-medical worker to forecast cardiac problems, whichwill save doctors' time.

FUTURE WORK:

In today's society, the majority of data is computerized, dispersed, and not adequately utilised. We can also use the data analysis to look for undiscovered patterns. This study's main goal is to predictheart illnesses with a high degree ofaccuracy. We may utilise the logisticregression algorithm, navie bayes, and sklearn in machine learning to predict cardiac disease. The paper's future focus will be on predicting cardiac illnesses with cutting-edge methods and algorithms in a minimal amount of time.

REFERENCES:

[1]. Rifki Wijaya,2013.Preliminary design of estimation heart disease by using machine learning ANN within on year; communication technology and electricvehicle technology, bandung-bali,Indonesia.
[2]. UCI Heart Disease Data set. [Online] Available:https://archive.ics.uci.edu/ml/datasets/heart+disease.
[3]. Verma, L., Srivastava, S. and Negi, P.C.(2016) "A Hybrid Data Mining Model to Predict Coronary Artery Disease Cases Using Non-Invasive Clinical Data." J Med Syst 40: 178.
[4]. A. Chauhan, A. Jain, P. Sharma and V. Deep. (2018) "Heart Disease Prediction using Evolutionary Rule Learning" 4th International Conference on ComputationalIntelligence & Communication Technology(CICT): 1-4.
[5]. F.A. Latifah, I. Slamet, Comparison of heart disease classification with logistic regression algorithm and random forest algorithm, in: Proceedings of the AIPConference,2020,p.2296,doi:10.1063/5.00 30579.
[6]. Nahar, Jesmin, Tasadduq Imam, Kevin
[7]. S. Tickle, and Yi-Ping Phoebe Chen. "Association rule mining to detect factors which contribute to heart disease in males and females." Expert Systems with Applications 40, no. 4 (2013): 1086-1093.
[8]. Sarath Babu,2017.Heart disease diagnosis using data mining technique, international conference on electronics,communication and aerospace technology ICECA2017.
[9]. Monika Gandhi,2015. Prediction in heart disease using techniques of data mining, International conference on futuristic trend in computational analysis and knowledge management(ABLAZE- 2015).
[10]. Sana Bharti,2015 .Analytical study of heart disease prediction comparing with different algorithms; International conference on computing, communication and automation(ICCA2015).