iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VI **Month of publication:** June 2021

DOI: https://doi.org/10.22214/ijraset.2021.35935

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI June 2021- Available at www.ijraset.com

Heart Attack Prediction Using Neural Networks

Chandu Nereeksha[1], Amitha S[2], Poornima V[3], Manasa G Hiremath[4], Dr. A B Rajendra[5]
[1, 2, 3, 4.]*B.E Student,* [5]*HOD, Dept. Of Information Science and Engineering, VVCE, Mysuru*

Abstract: ***Today, Heart disease seems to be a great cause for the increasing rate of immortality especially taken the current health situation under consideration. Improving the health conditions using the latest technology makes an enormous amount of contribution to the healthcare industry. One such, improvement is the use of Machine Learning in determining the heart diseases. Machine learning has a wide range of advancement in Neural Networks (NN). Artificial Neural Networks are basically inspired by the working of neural network inside a human brain. Our study aims to use the different algorithms and technologies to predict heart diseases at an early stage. Different data mining algorithms namely, Decision Tree, K-means clustering, Back-propagation and Random Forest are being used. The system classifies data into different stages such as normal and mild or extreme.***

Keywords: Heart disease, k-means clustering, neural networks, back propagation, random forest

I. INTRODUCTION

Our heart is one of the most important organs in our body. Basically, our heart is a muscle which comprises of chambers which are separated by valves and divided into two halves. There are two chambers in each half that is the atrium and the ventricle. The atria are responsible for collecting the blood and the ventricle pushes the blood out of the heart. The right part of the heart pushes the oxygen-poor blood to the lungs so that the cells can acquire more amount of oxygen. This blood then proceeds from the lungs to the left atrium and left ventricle. The left ventricle is then responsible for pushing this oxygen rich blood to the body organs. This oxygen acts as a fuel for creating energy and keeping us healthy.

The common term which is used to describe heart impairment is heart disease. In our county more than 17L people die every year due to cardiovascular disease and it is estimated that by 2030 more than 2 crore people will die from heart disease. Heart disease can take place in various forms but there are repeated set of key factors that whether or not someone is at the chance for heart disease. These common factors include gender, age, obesity, cholesterol, smoking, high blood pressure, diabetes, alcohol consumption.

Machine learning algorithms are efficacious for forecasting heart diseases at higher accuracy. These algorithms include Decision Tree, Naive Bayes, KNN, Artificial Neural Network where each one of them has its own speciality in predicting whether the person is at the risk of CVD. This prediction helps the doctor to speculate CVD at early stages and suggest necessary precautions

Prediction must be done during the initial levels to reduce the severe risks involved with it. Machine Learning plays a vital role in detecting the hidden layer patterns through back propagation algorithm. Further analysis of the data is carried out by decision tree algorithm to obtain the classified data.

This survey paper gives an overall review about various machine learning methods used in detection and prevention of CVD and their parallel comparison on several specification.

II. LITERATURE SURVEY

Tulay Karayilan and Ozkan Kilic [1] proposed a method for heart disease forecast. This method is build from an robotized medical diagnosis which is purely based on machine learning techniques and algorithms. Proposed system planned as a multilayer perceptron neural network and most popular Clevaland dataset is used. The neural network in the system had an input of 13 clinical data. It was trained with Back Propogation algorithm, which is the known as the best Artificial Neural Network learning algorithm to predict whether the patient is having any kind of heart disease present or not.

Apurb Rajdhan and teammates [2] proposed a model to predict heart attack at an early stage. It does so by analysing four codification algorithms namely, decision tree, Random Forest, Logistic Regression and Naive Bayes techniques based on f-measure score, accuracy, precision recall. Random Forest can be applied to large datasets and provides exact output even if the data has gone missing. Decision tree is applied as they are reliable, quick, and very little data preparation is needed..Logistic Regression is used for binary classification problems. Naive Bayes proves to be the best when the supposition of independence put away. Dataset is a combination of 4 different database which consists of 14 attributes. The experimental result of this research designate that the Random Forest algorithm has an accuracy score of 90.16% and hence proves to be the best and most streamlined algorithm.

A.Sankari Karthiga and teammates [3] proposed a novel procedure to prognosticate heart disease using Decision Tree and Naive Bayes Algorithm. Decision tree technique incorporates constructing a tree and operating the tree onto the dataset. Then,Entropy and Information Gain is calculated .On the other hand, Naive Bayesian model is undemanding to build, with uncomplicated iterative parameter evaluation. Prior probability and likelihood is calculated Performance Analysis is made between Naive Bayes and Decision tree algorithm. From the outcome acquired in this study, it is clear that Decision tree proves to be more error-freein producing the results compared to Naive Bayes.

Ketuk Agung Enriko and 2 more researchers [4] in their research used K-nearest Neighboring Algorithm with instant parameters. KNN is one of the most effective data mining techniques which often used in heart attack prognostication. The parameters are body indispensable signs that can be computed immediately. The advantage of immediate parameters obtained is that they can can be capturedover sensors for long-sufferers treated at home. In this research, dataset is obtained from UCI Database known as Hungarian dataset which has the highest records and 8 parameters are considered. Data mining analysis is done using KNN and results demonstrate that having 8 attributes is more sufficient to prognosticate heart attack.

Mrs .K.Uma Maheswari and Ms.J.Jasmine [5] proposed a method to predict heart attack. In this, they integrated Artificial neural network based technique and Logistic Regression model to foresee heart disease. Clevaland Dataset containing 297 observations taken into account.This system consists of 2 parts. The first part is about finding very much necessary risk factors to foresee heart disease. And another part is to separate the dataset into two sets namely, training and testing dataset. The neural network is developed for the training dataset.The trained neural network is now be able to foresee the testing dataset. Thus, the reconciliation of logistic regression and neural network gives the novel methodology in anticipating the heart disease of a person.

M.Chitrali S Dangare and DR. Sulabha S Apte [6] carried out research to predict heart attack in advance. This research involves developing decision support system for anticipating heart attack at an early stage. This prediction build upon historical heart disease database. Two additional attributes, obesity and smoking is added as an important prediction tool for heart attack forecasting. Taking ANN into consideration, back propagation algorithm and multi-layer perceptron neural network is used to build the model.Thus the experimental result proves that the model foresee heart attack at a very early stage with 100% accuracy because of usage of neural networks.

III. METHODOLOGY

In this project, real time dataset is collected from the hospitals and included 14 attributes namely, Smoking, Alcohol Consumption, tobacco chewing and some more. And in this section steps and methods involved in the processing of data is briefly explained.

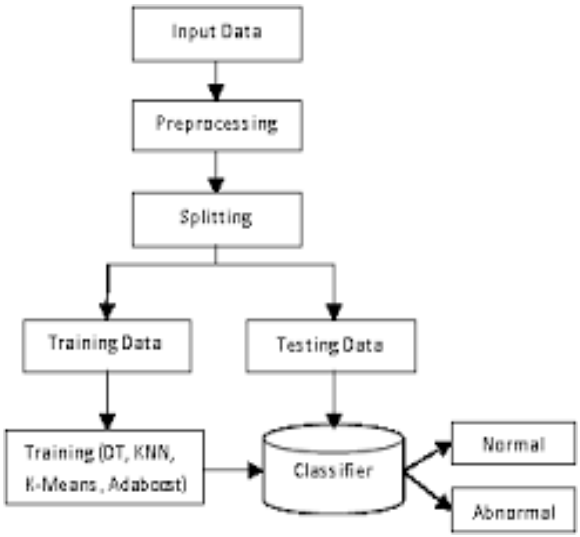


Figure 1: System Architecture

Figure 1 represents the system architecture. First step is the collection of dataset. The collected dataset is then preprocessed. Training and testing samples are obtained followed by the feature selection. Training data is modified through various algorithms and the required output is obtained. These steps are further explained in detail below:

A. Pre Processing

This is the process of converting the given dataset into clean dataset. This uses feature to normalize data which has been collected from user. This dataset contains NaN (Not a number) values which are converted into numerical values. This conversion is critical because the data collected from various sources are intermediate and by doing this, we can analyse the data more cautiously.

B. Splitting

The entire dataset is broken down into two sets i.e, testing and training sets. The training dataset includes a group of examples which is used for learning and to assign weights. On the other hand, the testing dataset is a group of samples used to examine the execution of fully enumerated classifier.

C. Classification

There are various machine learning algorithms that are applied to train the training data. Most of the papers that we have studied mainly uses Decision Tree, Neural Network, KNN, Random forest and K-mean clustering.

IV. ALGORITHMS

A. Decision Tree

This approach mainly composes of two steps i.e., building a tree and mapping this built tree to the dataset. The models in which a target variable can take on a possible set of values are called classification trees. In these models the nodes are the class labels and the branches constitutes the set of attributes that leads to that label. The most common DT is the gain ratio-decision tree. In this type of system, the root node is designated by the attribute that reduces entropy and utilizes the information gain. This is believed to give maximum accuracy and flexibility.

B. K-nearest Neighbour

KNN belongs to the type of supervised learning which preserves each and every cases and sorts new measure build from similarity measure. It basically comprises of two steps: to find k number of instances that are nearest to undetermined instance and choose the most recurring classification for these instances(k). The precise location of KNN is determined with the help of the training database. The distance between the target and each of instances in the training dataset is calculated using Euclidean distance. After locating the nearest neighbours, this will be assigned to the row that was examined in the dataset. This process is repeated for all the rows in the target set.

C. K-means clustering

This is an unsupervised learning algorithm which is applied when we have unknown class label. The main goal is to determine positions inside the cluster that reduces the sum of square distances from the center of the clusters.

Steps involved in k-means are:

1) Set the center for k cluster from n points that will be separated in k clusters
2) Assign the nearest cluster to each point using Euclidean distance
3) By using the mean of all points residing in that cluster initialize the position of each cluster.
4) Repeat the steps until confluence.

This algorithm is robust and easily understandable and provides best result when the examples are disconnected from one another.

D. Back Propagation

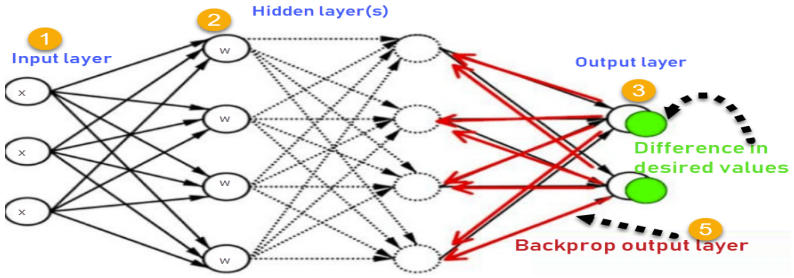


Figure 2: Back Propagation Algorithm

Figure 2 represents the working of back propagation algorithm. This is a machine learning algorithm used for training feedforward neural networks. It is mainly used to compute the gradient of loss function with reference to each weight by the chain rule. Each layer is effectively computed at a time. Steps carried out are:

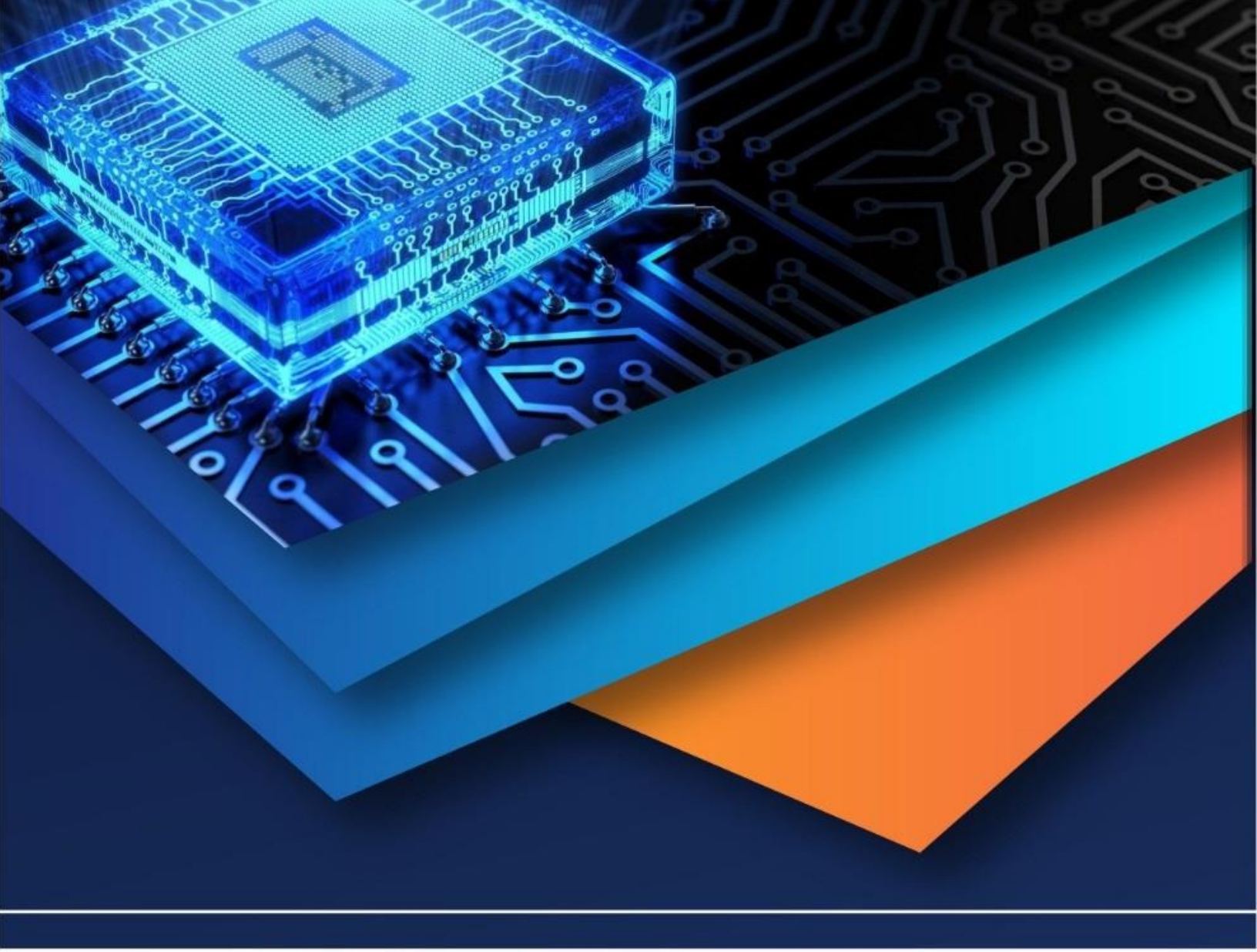
1) Arrival of inputs(x) by the previously connected path.
2) Input is patterned using the real weights W. Here weights are usually selected in a random manner.
3) Through the input layer, the hidden layer and the output layer output for every neuron is calculated.
4) This steps involves the calculation of errors in the output.
5) Finally, from output layer the output is sent back to the hidden layer to modify the weights accordingly so that the errors are decreased.

V. CONCLUSION

This research paper presented heart attack prediction system at an early stage which is build upon machine learning algorithms and neural network techniques. The main advantage of neural network is that it needs few formal statistical training to advance and can implicitly diagnose complex non-linear relationship that exists between dependent and independent variables. The experimental results show that neural network does better forecast heart attack comparison with others.

REFERENCES

[1] Tulay Karayilan , Ozkan Kilic, "Prediction Of Heart Disease Using Neural Network".
[2] Apurb Rajdhan , Milan Sai, Dundigalla Ravi , Avi Agarwal "Heart Disease Prediction Using Machine Learning", Internatinal Journal Of Engineering Research and Technology(IJERT) ISSN: 2278-0181 Vol. 9 Issue 04, April-2020.
[3] A. Sankari Karthiga, M. Safish Mary, M. Yogasini "Early Prediction Of Heart Disease Using Decision tree Algorithm" , (IJARBEST) ISSN: 2395-695X
[4] Ketut Agung Enriko, Muhammad Suryanegara, Dadang Gunawan, "Heart Disease Prediction System using k-Nearest Neighbor Algorithm with Simplified Patients Health Parameters" ISSN: 2180-1843.
[5] Mrs. K Uma Maheshwari, Ms. J Jasmine "Neural Network based Heart Disease Prediction (IJERT) ISSN: 2278-0181.
[6] Miss Chitrali S. Dangare, DR. Sulabha S Apte "A Data Mining Approach For Prediction Of Heart Disaese Using Neural Networks" (IJCET) ISSN:0976-6367.
[7] S. Kavitha, K. R. Baskaran, S. Sathyavathi "Heart Disease with Risk Prediction Using Machine Learning Algorithms" (IJRTE) ISSN:2227-3878.
[8] Shahid Mehmood Awan, Muhammad Usama Riaz, Abdul Ghaffar Khan, "Prediction Of Heart Disease Using Artificial Neural Network", VFAST Transactions on Software Engineering, ISSN(e): 2309-6519, ISSN(p): 2411-6327, Volume 13, September-December, 2018.
[9] Animesh Hazra, Subrata Kumar Mandal, Amit Gupta, Arkomita Mukherjee, Asmita Mukherjee, "Heart Disease Diagnosis and Prediction Using Machine Learning and Data Mining Techniques: A Review", Advances in computational Science and Technology, ISSN 0973-6107, Volume 10, Number 7 (2017) pp. 2137-2159.
[10] Shrinivas D Desai, Indira Fal Dessai, Linganagouda Kulkarni, "Intelligent Heart Disease Prediction System Using Probabilistic Neural Network", IJATCSE, Vol.2, September 20, 2013, Bangalore, India.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY