

# Survey Paper on Effective Heart Disease Prediction Using Hybrid Machine learning

PRIYANKA KESHAV GAIKWAD.

PG Student, Department of Information Technology, Bharati Vidyapeeth (Deemed to be university) College of Engineering Pune.

PROF. YOGINI C. KULKARNI, PROF. PRAMOD A. JADHAV

Department of Information Technology, Bharati Vidyapeeth (Deemed to be university) College of Engineering Pune.

**Abstract**—Big data analytics has started to play an important role in the evolution of healthcare practices and research. It has provided tools to collect, manage, analyze, and assimilate large volumes of disparate, structured, and unstructured data produced by current healthcare systems. Big data analytics has been recently applied towards aiding the process of care delivery and disease exploration. Medical data mining has great potential to explore the hidden models in data sets of the medical domain. These models can be used for make a clinical diagnosis these data should be collected in a standardized form. Of the medical profiles six attributes are extracted, such as age, sex, blood pressure and blood sugar etc. can predict the likelihood of a patient contracting heart disease. These attributes are introduced in the machine learning algorithms, classification of decision tree in heart disease prediction, applying the technique of data mining to cardiopathies treatment; it can provide a reliable performance like that achieved in the diagnosis of heart disease. For these medical industries it could offer a better diagnosis and treatment of the patient to be achieved good quality of services the main advantages of this document are: timely detection of heart disease and its diagnosis in time and provide treatment at an affordable cost.

**Keywords**-Decision Tree, Machine Learning, QA System, heart disease prediction.

## I. INTRODUCTION

The main reason for death worldwide, including South Africa is heart attack diseases and possible detection at an earlier stage will prevent these attacks. Medical practitioners generate data with a wealth of concealed information present, and it's not used effectively for predictions. For this reason, the research converts the unused data into a dataset for shaping using different data mining techniques. People die having encountered symptoms that were not taken into considerations. There is a requirement for medical practitioners to defined heart disease before they occur in their patients.

The features that increase the chances of heart attacks are smoking, lack of physical exercises, high blood pressure, high cholesterol, unhealthy diet, detrimental use of alcohol, and high sugar levels . Cardio Vascular Disease (CVD) constitutes coronary heart, cerebro-vascular or Stroke, hypertensive heart disease, congenital heart, peripheral artery, rheumatic heart disease, and inflammatory heart disease.

Data mining is a knowledge discovery technique to examine data and encapsulate it into useful information. The

current research intends to forecast the probability of getting heart disease given patient data set. Prophecies' and descriptions are principal goals of data mining; in practice Prediction in data mining involves attributes or variables in the data set to locate unknown or future state values of other attributes. Description emphasize on discovering patterns that describes the data to be interpreted by humans.

### A. Motivation:

The huge data growth in biomedical and healthcare businesses, accurate analysis of medical data benefits from early detection, patient care, and community services. The analysis accuracy is reduced if the quality of the medical data is incomplete.

## II. RELATED WORK

Literature survey is the most important step in any kind of research. Before start developing we need to study the previous papers of our domain which we are working and on the basis of study we can predict or generate the drawback and start working with the reference of previous papers.

In this section, we briefly review the related work on Heart disease prediction and their different techniques.

Nearest neighbor (KNN) is very simple, most popular, highly efficient and effective technique for pattern recognition. KNN is a straight forward classifier, where parts are classified based on the class of their nearest neighbor. Medical data bases are big volume in nature. If the data set contains excessive and irrelevant attributes, classification may create less accurate result. Heart disease is the best cause of death in INDIA. In Andhra Pradesh heart disease was the best cause of mortality accounting for 32% of all deaths, a rate as high as Canada (35%) and USA. Hence there is a need to define a decision support system that helps clinicians to take precautionary steps. In this work proposed a new technique which combines KNN with genetic technique for effective classification. Genetic technique perform global search in complex large and multimodal landscapes and provide optimal solution [1].

Image de-noising includes the manipulation of the image data to produce a visually high quality image. The Non Local means filter is originally designed for Gaussian noise removal and the filter is changed to adapt for speckle noise reduction. Speckle noise is a firstly source of medical ultrasound imaging noise and it should be filtered out. This

work reviews the existing Non-Local Means based filters for image de-noising [2].

This work has analyzed prediction systems for Heart disease using more number of input attributes. The work uses medical terms such as sex, blood pressure, cholesterol like 13 attributes to predict the likelihood of patient getting a Heart disease. Until now, 13 attributes are used for prediction. This research work added two more attributes i.e. obesity and smoking. The data mining classification algorithms, namely Decision Trees, Naive Bayes, and Neural Networks are analyzed on Heart disease database [3].

Medical Diagnosis Systems play important role in medical practice and are used by medical practitioners for diagnosis and treatment. In this work, a medical diagnosis system is defined for predicting the risk of cardiovascular disease. This system is built by combining the relative advantages of genetic technique and neural network. Multilayered feed forward neural networks are particularly adapted to complex classification problems. The weights of the neural network are determined using genetic technique because it finds acceptably good set of weights in less number of iterations [4].

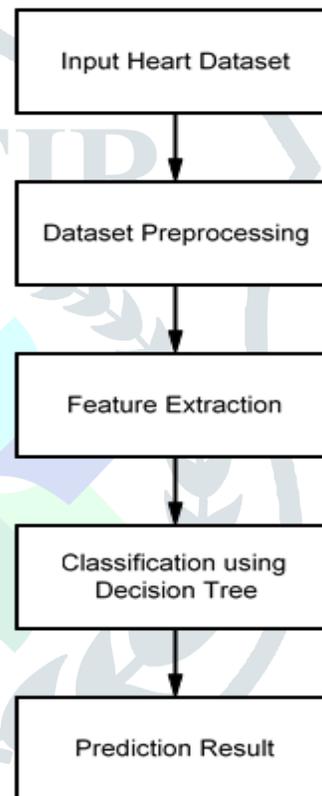
A wide range of heart condition is defined by thorough examination of the features of the ECG report. Automatic extraction of time plane features is valuable for identification of vital cardiac diseases. This work presents a multi-resolution wavelet transform based system for detection 'P', 'Q', 'R', 'S', 'T' peaks complex from original ECG signal. 'R-R' time lapse is an important minutia of the ECG signal that corresponds to the heartbeat of the related person. Abrupt increase in height of the 'R' wave or changes in the measurement of the 'R-R' denote various anomalies of human heart. Similarly 'P-P', 'Q-Q', 'S-S', 'T-T' also corresponds to various anomalies of heart and their peak amplitude also envisages other cardiac diseases. In this proposed method the 'PQRST' peaks are marked and stored over the entire signal and the time interval between two consecutive 'R' peaks and other peaks interval are measured to find anomalies in behavior of heart, if any [5].

algorithms provides better performance. After classification, performance criteria including accuracy, precision, F-measure is to be calculated. The comparison measure reveals that Decision tree is the best classifier for the diagnosis of heart disease on the existing data.

Advantages:

- Predict heart disease for Structured Data using machine learning algorithms i.e., Decision Tree (DT).
- Find reliable answer using this system
- To achieve better accuracy.

System Architecture:



**Table1: Accuracy Graph for Heart Dataset**

	Existing System	Proposed System
<b>Precision</b>	<b>0.825</b>	<b>0.9</b>
<b>Recall</b>	<b>0.825</b>	<b>0.9</b>
<b>F-Measure</b>	<b>0.825</b>	<b>0.9</b>

**Fig. System Architecture**

Conclusion

The experiment is organized with the dataset of Heart Disease by machine learning algorithms. Heart Disease dataset is taken and analysed to predict the asperity of the disease. A Decision tree approach is used to predict the disease. The data in the dataset is pre-processed to make it suitable for classification. The Decision tree approach to generate efficient classification rules is proposed. To perform classification task of medical data, the network is trained using Convolutions technique. Machine learning technique is a multilayer perceptron that is the special design for identification of two-dimensional image information. Always have more layers: input layer, convolution layer, sample layer and output layer.

III. PROPOSED SYSTEM:-

This work is used for finding heart diseases. Based on risk factor the heart diseases can be predicted very easily. The main aim of this paper is to predict the heart diagnosis. First, the heart numeric dataset is extracted and preprocess them. After that using extract the features that is condition to be find to be classified by Decision Tree (DT).Compared to existing;

## REFERENCES

- [1] Algorithm M.Akhiljabbar B.L DeekshatuluaPriti Chandra International “Classification of Heart Disease Using K-Nearest Neighbor and Genetic Algorithm” Conference on Computational Intelligence: Modeling Techniques and Applications (CIMTA) 2013.
- [2] Beshiba Wilson, Dr.JuliaPunitha Malar Dhas “A Survey of Non-Local Means based Filters for Image Denoising” International Journal of Engineering Research & Technology, Vol.2 - Issue 10 (October – 2013).
- [3] Chaitrali S Dangare “Improved Study Of Heart Disease Prediction System Using Data Mining Classification Techniques”, International Journal Of Computer Applications, Vol.47, No.10 (June 2012).
- [4] Amma, N.G.B “Cardio Vascular Disease Prediction System using Genetic Algorithm”, IEEE International Conference on Computing, Communication and Applications, 2012.
- [5] Sayantan Mukhopadhyay<sup>1</sup> , Shouvik Biswas<sup>2</sup> , AnamitraBardhan Roy<sup>3</sup> , Nilanjan Dey<sup>4</sup> ‘ Wavelet Based QRS Complex Detection of ECG Signal’ International Journal of Engineering Research and Applications (IJERA) Vol. 2, Issue 3, May-Jun 2012, pp.2361-2365

