Review on Hearth diseases prediction on Neural Network

Vijay yadav, Ashish Tiwari

P.G Scholar, Department of Computer Science and Engineering,vindhya institute of technology Indore(mp), India

Professor, Department of Computer Science and Engineering, vindhya institute of technology Indore(mp), India

Vijayyadav0105@gmail.com ; ashishtiwari205@gmail.com

ABSTRACT: *N*eural **network (ANN) tool has been used for solving many decision modeling problems. A multilayer perception is a feed forward ANN model that is used extensively for the solution of heart diseases problems. An ANN is the same as human brain. This is a non-linear classification monitoring learning technique used for coronary heart disease is the biggest epidemic the world is at risk of coronary heart disease. Most doctor's diagnosis is made by clinical expertise and patients were asked to take someone. Diagnostic testing. But all the tests will not contribute to diagnosis of the disease effectively. Feature subset selection, dimensionality used to remove irrelevant data to reduce a preprocessing step. In this paper we have a classification approach to the classification of heart disease ANN uses features from subset selection. Properties which indirectly reduces. Diagnostic tests which need to be taken by the patient. We applied our approach to heart disease data base. Our experimental results show that the accuracy improves with conventional classification techniques. This system is possible and for faster and more accurate diagnosis of heart disease.**

K*EYWORDS:* **Classification , artificial neural networks, symptoms, Ann tool**

(I). INTRODUCTION

A medical condition is a broad term that includes all diseases and disorders. A disease is an abnormal condition affecting the body of an organism. A disorder is a functional abnormality or disturbance. The human body is remarkably well designed. Most of its organs have a great deal of extra capacity or reserve: They can still function adequately even when damaged. For example, more than two thirds of the liver must be destroyed before serious consequences occur, and a person can usually live with only one lung or one kidney. Other organs can tolerate little damage before they malfunction and symptoms occur. For example, if an artery in the brain becomes blocked or ruptures (stroke) and even a small amount of tissue in a vital part of the brain is destroyed, a person may be unable to speak, move a limb, or maintain balance. If a heart attack destroys a small amount of tissue in the part of the heart that creates or carries the signals to beat, the heart rate may become dangerously slow and the person may even die. Disease often affects anatomy, and changes in anatomy can cause disease. If the blood supply to a tissue is blocked or cut off, the tissue dies (called infarction), as in a heart attack (myocardial infarction) or stroke (cerebral infarction). An abnormal heart valve can cause heart malfunction. Trauma to the skin may damage its ability to act as a barrier, which may lead to infection. Abnormal growths, such as cancer, can directly destroy normal tissue or produce pressure that ultimately destroys it.

(II) DISEASE

Cancer :Cancer is a group of diseases that involve abnormal increases in the number of cells, with the potential to invade or spread to other parts of the body. Not all tumors or lumps are cancerous; benign tumors are not classified as being cancer because they do not spread to other parts of the body. There are over 100 different known cancers that affect humans. Cancers are often described by the body part that they originated in. However, some body parts contain multiple types of tissue, so for greater precision, cancers are additionally classified by the type of cell that the tumor cells originated from.

Coetaneous conditions: Many conditions affect the human integumentary system—the organ system covering the entire surface of the body and composed of skin, hair, nails, and related muscle and glands. The major function of this system is as a barrier against the external environment. The skin weighs an average of four kilograms, covers an area of two square meters, and is made of three distinct layers: the epidermis, dermis, and subcutaneous tissue. The two main types of human skin are: glabrous skin, the hairless skin on the palms and soles (also referred to as the "palm plantar" surfaces), and hair-bearing skin. Within the latter type, the hairs occur in structures called pilosebaceous units, each with hair follicle, sebaceous gland, and associated arrestor pile muscle. In the embryo, the epidermis, hair, and glands form from the ectoderm, this is chemically influenced by the underlying mesoderm that forms the dermis and subcutaneous tissues.

Endocrine disease: Endocrine diseases are disorders of the endocrine system. The branch of medicine associated with endocrine disorders is known as endocrinology. Broadly speaking, endocrine disorders may be subdivided into three group

1. Endocrine gland hypo secretion (leading to hormone deficiency)
2. Endocrine gland hyper secretion (leading to hormone excess)
3. Tumors (benign or malignant) of endocrine glands

Endocrine disorders are often quite complex, involving a mixed picture of hypo secretion and hyper secretion because of the feedback mechanisms involved in the endocrine system. For example, most forms of hyper thyroidismare associated with an excess of thyroid hormone and a low level of thyroid stimulating hormone.

Eye disease :This is a partial list of human eye diseases and disorders. The World Health Organization publishes a classification of known diseases and injuries, the International Statistical Classification of Diseases and Related Health Problems, or ICD-10.

Human gastrointestinal tract: Thehuman gastrointestinal tract (GI tract or GIT) is an organ system responsible for transporting and digesting foodstuffs, absorbing nutrients, and expelling waste**.** The tract consists of the stomach and intestines, and is divided into the upper and lower gastrointestinal tracts. The GI tract includes all structures between the mouth and the anus, forming a continuous passageway that includes the main organs of digestion, namely, the stomach, small intestine, and large intestine. In contrast, the human digestive system comprises the gastrointestinal tract plus the accessory organs of digestion (the tongue, salivary glands, pancreas, liver, and gallbladder).The tract may also be divided into foregut, midget, and hindgut, reflecting the embryological origin of each segment.

Brain Tumors: Some times, tumors form in the brain and can be very dangerous. These are called primary brain tumors. In other cases, cancer somewhere else in your body spreads to your brain. These are called secondary or metastatic brain tumors. Symptoms of brain tumors depend on the size and location of the tumor. The most common symptoms of brain tumors are:

- headaches
- seizures
- numbness or tingling in your arms or legs
- nausea
- vomiting
- changes in personality
- difficulty with movement or balance
- changes in your hearing, speech, or vision

The type of treatment you'll receive depends on many different factors, such as the size of the tumor and your age and overall health. he main types of treatment for brain tumors are surgery, chemotherapy (medication), and radiation therapy.

Cardiovascular disease: Cardiovascular disease (CVD) is a class of diseases that involve the heart or blood vessels. Cardiovascular disease includes coronary artery diseases (CAD) such as angina and myocardial infarction (commonly known as a heart attack). Other CVDs are stroke, hypertensive heart disease, rheumatic heart disease, cardiomyopathy, heart arrhythmia, congenital heart disease, valvular heart disease, cordites, aortic aneurysms, peripheral artery disease, and venous thrombosis. The underlying mechanisms vary depending on the disease in question. Coronary artery disease, stroke, and peripheral artery disease involve atherosclerosis. This may be caused by high blood pressure, smoking, diabetes, lack of exercise, obesity, high blood cholesterol, poor diet, and excessive alcohol consumption, among others. High blood pressure results in 13% of CVD deaths, while tobacco results in 9%, diabetes 6%, lack of exercise 6% and obesity 5%. Rheumatic heart disease may follow untreated strepthroa.

The heart is important organ or part of our body. Life is itself dependent on efficient working of heart. If operation of heart is not proper, it will affect the other body parts of human such as brain, kidney etc. It is nothing more than a pump, which pumps blood through the body. If circulation of blood in body is inefficient the organs like brain suffer and if heart stops working altogether, death occurs within minutes. Life is completely dependent on efficient working of the heart. The term Heart disease refers to disease of heart & blood vessel system within it.

There are number of factors which increase the risk of Heart disease [2]:

- Family history of heart disease
- Smoking
- Cholesterol
- Poor diet
- High blood pressure
- High blood cholesterol
- Obesity
- Physical inactivity
- Socioeconomic disadvantage
- Air pollution

Symptoms of a Heart Attack

Symptoms of a heart attack can include:

- Discomfort, pressure, heaviness, or pain in the chest, arm, or below the breastbone.

- Discomfort radiating to the back, jaw, throat, or arm.

- Fullness, indigestion, or choking feeling (may feel like heartburn).

- Sweating, nausea, vomiting, or dizziness.

- Extreme weakness, anxiety, or shortness of breath.

- Rapid or irregular heartbeats

Types of Heart diseases :

There are many cardiovascular diseases involving the blood vessels. They are known as vascular diseases:

- **Coronary artery disease**- (also known as coronary heart disease and ischemic heart disease)
- **Peripheral arterial disease** – disease of blood vessels that supply blood to the arms and legs
- **Cerebrovascular disease** – disease of blood vessels that supply blood to the brain (includes stroke)
- **Renal artery steno sis**
- **Aortic aneurysm**

There are also many cardiovascular diseases that involve the heart.

- **Cardiomyopathy** – diseases of cardiac muscle
- **Hypertensive heart disease** – diseases of the heart secondary to high blood pressure or hypertension
- Heart failure
- **Pulmonary heart disease** – a failure at the right side of the heart with respiratory system involvement
- **Cardiac dysrhythmias** – abnormalities of heart rhythm
- **Inflammatory heart disease**
    - o **Endocarditic** – inflammation of the inner layer of the heart, the endocardium. The structures most commonly involved are the heart valves.
    - o **Inflammatory cardiomegaly**
    - o **Myocarditis** – inflammation of the myocardium, the muscular part of the heart.
- **Valvular heart disease**
- **Congenital heart disease** – heart structure malformations existing at birth
- **Rheumatic heart disease**– heart muscles and valves damage due to rheumatic fever caused by *Streptococcus pyogenes* a group A streptococcal infection.

(III) LITERATURE SURVEY

Numerous studies have been done that have focus on diagnosis of heart disease. They have applied different techniques on artificial neural network data mining for diagnosis & achieved different probabilities for different methods

- Bahadur Patel et al. [3] designed a system to identify the chances of a coronary heart disease. They have divided all parametersin two levels according to criticality and each level is assigned separate weight age. Finally both levels are considered to derive afinal decision. They have implemented neuro-fuzzy integrated approach at two levels. So, error rate is very low and work efficiencyis high. In this paper, they have performed the analysis for coronary heart disease

- The prediction of Heart disease, Blood Pressure and Sugar with the aid of neural networks was proposed by Niti Guru et al. [4]. The dataset contains records with 13 attributes in each record. The supervised networks i.e. Neural Network with back propagation algorithm is used for training and testing of data.

- To develop the multi-parametric feature with linear and nonlinear characteristics of HRV (Heart Rate Variability) a novel technique was proposed by HeonGyu Lee et al. [5]. To achieve this, they have used several classifiers e.g. Bayesian Classifiers, CMAR (Classification based on Multiple Association Rules), C4.5 (Decision Tree) and SVM (Support Vector Machine).

- The healthcare industry collects huge amounts of healthcare data which, unfortunately, are not "mined" to discover hidden information, to take decisions effectively, to discover the relations that connect parameters in a database is the subject of data mining. As large amount of data is generated in medical organizations (hospitals, medical centers) but as this data is not properly used. There is a wealth of hidden information present in the datasets.[6]

- Latha Parthiban et al. [7] projected an approach on basis of coactive neuro-fuzzy inference system (CANFIS) for prediction of heart disease. The CANFIS model uses neural network capabilities with the fuzzy logic and genetic algorithm.

- Naïve Bayes Rule is the basis for many machine learning and data mining methods. The rule (algorithm) is used to create models with predictive capabilities.
  It provides new ways of exploring and understanding data. A naive Bayes classifier is a term dealing with a simple probabilistic classification based on applying Bayes theorem[8]

- To have focus on diagnosis of heart disease different studies have been done. Different data mining techniques has been used by them for diagnosis & achieved different probabilities for different methods. Using data mining techniques an Intelligent Heart Disease Prediction System (IHDPS) is developed Sellappan Palaniappan et al[9] proposed Naive Bayes, Neural Network, and Decision Trees For appropriate results each method has its own strength Hidden patterns and relationship between them is used to build this system It is user friendly, expandable & web based.

(VI). DATA MINING TECHNIQUES USED FOR PREDICTIONS

The three different data mining classification techniques, i.e. Neural Networks, artificial neural network , Decision Trees, and Naive Bays are used to analyze the dataset.

4.1. Neural Networks An artificial neural network (ANN), often just called a "neural network" (NN), is a mathematical model or computational model based on biological neural network. In other words, it is an emulation of biological neural system [10]. A Multi-layer Perception Neural Networks (MLPNN) is used. It maps a set of input data onto a set of appropriate output data. It consists of 3 layers input layer, hidden layer & output layer. There is connection between each layer & weights are assigned to each connection. The primary function of neurons of input layer is to divide input x_i into neurons in hidden layer. Neuron of hidden layer adds input signal x_i with weights w_{ji} of respective connections from input layer. The output Y_j is function of $Y_j = f(\Sigma w_{ji} x_i)$ Where f is a simple threshold function such as sigmoid or hyperbolic tangent function

4.2. Decision Trees The decision tree approach is more powerful for classification problems. There are two steps in this techniques building a tree & applying the tree to the dataset. There are many popular decision tree algorithms CART, ID3, C4.5, CHAID, and J48. From these J48 algorithm is used for this system. J48 algorithm uses pruning method to build a tree. Pruning is a technique that reduces size of tree by removing over fitting data, which leads to poor accuracy in predications. The J48 algorithm recursively classifies data until it has been categorized as perfectly as possible. This technique gives maximum accuracy on training data. The overall concept is to build a tree that provides balance of flexibility & accuracy.

4.3. Naive Bayes Naive Bayes classifier is based on Bayes theorem. This classifier algorithm uses conditional independence, means it assumes that an attribute value on a given class is independent of the values of other attributes. The Bayes theorem is as follows: Let $X=\{x1, x2, ...,xn\}$ be a set of n attributes. In Bayesian, X is considered as evidence and H be some hypothesis means, the data of X belongs to specific class C. We have to determine P (H|X), the probability that the hypothesis H holds given evidence i.e. data sample X. According to Bayes theorem the P (H|X) is expressed as P (H|X) = P (X| H) P (H) / P (X) ..
:

   4.4   CLASSIFICATION TECHNIQUE he proposed system is to build the Decision Support System for precluding Heart Disease using classification in data mining techniques to identify the level of risk in coronary heart diseases. Coronary heart disease can be caused due to risk factors like high blood pressure, high blood cholesterol, tobacco use, obesity, unhealthy diet, physical inactivity, diabetes, advancing age, inheriteddisposition. This system helps the patients in take precautionary actions to stretch their life span and to assist medical practitioners to diagnose and predict the probable complications well in advance. Identifying themajor risk factors of CHD categorizing the risk factors in an order which causes high damages such as high blood cholesterol, diabetes, smoking, poor diet, obesity, hyper tension, stress, etc. Primary prevention is recommended as promoting healthy lifestyle and habits through increased awareness and consciousness, to prevent development ofany risk factors. This system which predicts the possibility of heart disease risk of patient for the next ten years for prevention using clustering algorithm.
To assist medical practitioners to diagnose and predict the probable complications and predictaccuracy.
1) Generate class based association rules from a training data set
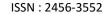 2) Classify the test data set into predefined class labels

(V). ANALYSIS AND RESUILT

5.1 Comparison on the basis of accuracy

For comparing the performance of the proposed algorithms we implement the single linkage and complete linkage method. Our first comparison is based on execution time and number of objects.

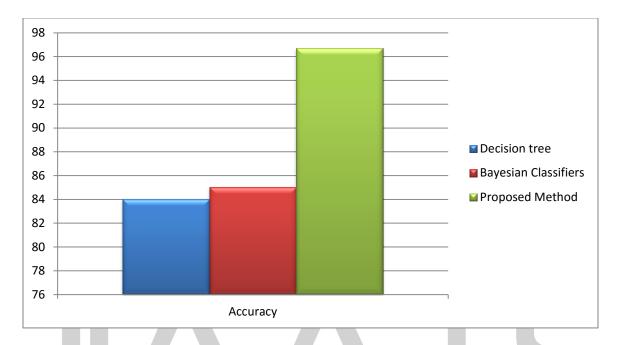Table 5.1 Accuracy in percentages

Decision Tree (weka)	Bayesian Classifiers(weka)	Proposed Method

84.01	85.03	98.02

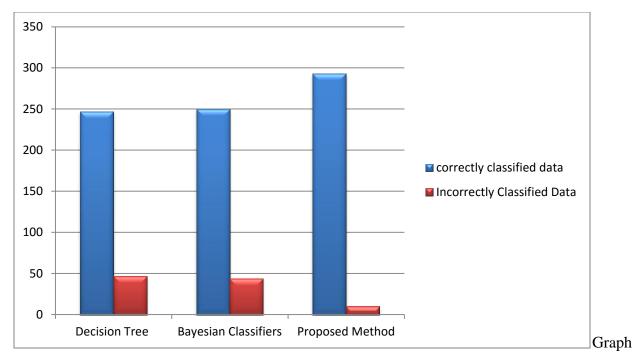Table 5.1 Number of Record and accuracy in percentages

Graph 5.1 Accuracy Comparison between various method

Table 6.2 Tested classified Data

	Correctly Classified	Incorrectly Classified	Total Test Data
Decision Tree	247	47	294
Bayesian Classifiers	250	44	294
Proposed Method	293	10	303

Graph
6.2 Comparison of various methods

VI. CONCLUSION

There are several algorithms and methods have been developed for classify heart attack problem accurately. But problem are always arises for finding a new algorithm and process for extracting knowledge for improving accuracy and efficiency. The most popular classification methods are Decision Tree and Naïve Bays Classifier. From the experiment it clear that proposed method is more accurately classify the recodes as compared to previous method. Proposed method considers all attribute given to heart attack condition. Proposed method is also simple to under stands and calculation is also easy. Heart disease is a fatal disease by its nature and misdiagnosis of this disease can cause serious, even life threatening complications such as cardiac arrest and death. The best model selected for predicting heart disease could not exceed a classification accuracy of 96.69% and still much remains to fill the gap of 3.31% misclassified cases.

REFERENCES:

[1]https://en.wikipedia.org/wiki/Lists_of_**diseases**

[2] https://en.wikipedia.org/wiki/Cardiovascular_disease#Types

[3] Bahadur Patel, Ashish Kumar Sen D P, ShamsherShukla, "A Data Mining Technique for Prediction of Coronary HeartDisease Using Neuro-FuzzyIntegrated Approach Two Level", International Journal Of Engineering And Computer ScienceISSN:2319-7242, Page No. 2663-2671, Volume 2 Issue 9 Sept., 2013

[4] Niti Guru, Anil Dahiya, NavinRajpal, *"Decision Support System for Heart Disease Diagnosis Using Neural Network"*, Delhi Business Review, Vol. 8, No. 1 (January - June 2007).

[5] HeonGyu Lee, Ki Yong Noh, KeunHoRyu, "Mining Biosignal Data: Coronary Artery Disease Diagnosis using Linear and Nonlinear Features of HRV," LNAI 4819: Emerging Technologies in Knowledge Discovery and Data Mining, pp. 56-66, May 2007.

[6] K. Cinetha, Dr. P. Uma Maheswarion Heart Disease Diagnosis and Prediction using Naive Bayes inData MiningK. Cinetha*et al*, International Journal of Computer Science and Mobile Computing, Vol.3 Issue.2, February- 2014,

[7}LathaParthiban and R.Subramanian, *"Intelligent Heart Disease Prediction System using CANFIS and Genetic Algorithm"*, International Journal of Biological, Biomedical and Medical Sciences, Vol. 3, No. 3, 2008.

[8] B.V. Baiju and R.J. Remy Janet Heart Disease Diagnosis and Prediction using Naive Bayes in Data MiningInternational Journal of Current Engineering and Technology10 April2015, Vol.5, No.2 (April2015)

[9] SellappanPalaniappan, RafiahAwang, "Intelligent Heart Disease Prediction System Using Data Mining Techniques" IJCSNS International Journal of Computer Science and Network Security, Vol.8 No.8, August 2008 .

[10]*https://en.wikipedia.org/wiki/Artificial_**neural_network***