IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 9, September 2021

DOI: 10.17148/IJARCCE.2021.10918

A Case Study on Expert System for Diagnosis of Heart Disease

Ali Mir Arif Mir Asif[1]

[1]Assistant Professor, I.M.S.I.T., Aurangabad, India

Abstract: The use of neural network on different diseases has been used on large scale since last two decades. This expert system offers a helping hand for the accurate decision over a certain diagnosis. A medical training may not have enough experience to deal and tackle with some high risk diseases like heart, kidney and brain. This case study includes details about patient's data, coding, normalization and tabulation. It describes various heart disease diagnostic techniques such as Feed Forward Back- propagation (FFBP), Support Vector Machine (SVM), Generalized Regression Neural Network (GRNN) and Radial Basis Function (RBF) has been applied over the data for the experiment. Additionally it represents different tables of symptoms used for heart disease diagnosis. In this case study, expert system for diagnosis of heart disease useful for the new researcher to understand how to collect the data and perform experimental analysis using different neural network techniques.

Keywords: Data Mining, Expert system, Heart disease, SVM, RBF, GRNN, FFBP

I. INTRODUCTION

Nowadays we know that a heart disease becomes a common one. If we have to live luxurious life we need to work like machines in order to earn lot of money hence we have to forget to take care of our health. So due to this, there is change in the food which we have consume, our lifestyle changes and finally, leads to diabetes, blood pressure and various other diseases at young age. All these reasons lead to negligence of our health which increases the changes of heart diseases. We know that heart is the most essential organ of the human body and if it gets affected then it also affects the other major organs of the body that's why we need a doctor's advice for good health and avoid major heart diseases. As human beings are mostly inactive in physical work, it becomes main cause of suffering from prediabetes. Physically inactive healthy person usually are seen having excess fat. This is another cause of being suffered from prediabetes and if the excess fat is in the abdominal fat then it becomes the main cause of being suffered from prediabetes [1-3].

An expert system is defined as software that attempts to reproduce the performance of one or more human experts, most commonly in a specific problem domain. It basically uses an inference engine connected to the knowledge base. A wide variety of methods can be used to simulate the performance of the expert however common to most or all are: • the creation of a so-called "knowledgebase" which uses some knowledge representation formalism to capture the Subject Matter Expert's (SME) knowledge;

• a process of gathering that knowledge from the SME and codifying it according to the formalism, which is called knowledge engineering [4]. Expert systems may or may not have learning components but a third common element is that once the system is developed it is proven by being placed in the same real world problem solving situation as the human SME, typically as an aid to human workers or a supplement to some information system.

The examples of Expert Systems are explained as following:

• MYCIN: To identify bacteria causing severe infections, such as bacteremia and meningitis, and to recommend antibiotics, with the dosage adjusted for patient's body weight.

• DENDRAL: Its primary aim was to help organic chemists in identifying structure of unknown organic molecules, by analyzing their mass spectra and using knowledge of chemistry.

• SAINT: To solve problems based on Integral Calculus, and eventually generate simpler expressions for complex functions.

According to WHO report Global atlas on cardiovascular disease prevention and control states that cardiovascular disease (CVDs) are the leading causes of death and disability in the world. Although a large proportion of CVDs is preventable, they continue to rise mainly because preventive measures are inadequate. Over 17.3 million An estimated 17.3 million people died from CVDs in 2008, Over 80% of CVD deaths take place in low- and middle-income countries, 23.6 million By 2030, almost 23.6 million people will die from CVDs. A rule-based system is a system to store and manipulate knowledge to interpret information in a useful way. They are often used in artificial intelligence applications and research. Rule-based programming attempts to derive execution instructions from a starting set of data

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 9, September 2021

DOI: 10.17148/IJARCCE.2021.10918

and rules, which is a more indirect method than using a imperative programming language which lists execution steps straightforwardly [5-8].

A formal definition of Knowledge discovery in databases is given as follows: "Data mining is the non trivial extraction of implicit previously unknown and potentially useful information about data" [52]. Data mining technology provides a user-oriented approach to novel and hidden patterns in the data. The discovered knowledge can be used by the healthcare administrators to improve the quality of service. The discovered knowledge can also be used by the medical practitioners to reduce the number of adverse drug effect, to suggest less expensive therapeutically equivalent alternatives. Anticipating patient's future behaviour on the given history is one of the important applications of data mining techniques that can be used in health care management[11][58].

Data is very important factor in every field of today life. There is no hard and fast rule about exactly what size a database needs to be in order for the data inside of it to be considered "big." Instead, what typically defines big data is the need for new techniques and tools in order to be able to process it. In order to use big data, you need programs which span multiple physical and/or virtual machines working together in concert in order to process all of the data in a reasonable span of time. The big data techniques can also be used for heart disease diagnosis as this data can also be in the various forms.[54-56].

Robots have entered into many aspects of human life, including medical science. The impact of robotics on medicine is undeniable. Robots can be defined as "automatically controlled multitask manipulators, which are freely programmable in three or more axes." The success of robots is based on their precision, lack of fatigue, and speed of action [57].

A major challenge facing healthcare organizations (hospitals, medical centers) is the provision of quality services at affordable costs. Quality service implies diagnosing patients correctly and administering treatments that are effective. Poor clinical decisions can lead to disastrous consequences which are therefore unacceptable. Hospitals must also minimize the cost of clinical tests. They can achieve these results by employing appropriate computer-based information and/or decision support systems. Health care data is massive. It includes patient centric data, resource management data and transformed data. Health care organizations must have ability to analyze data. Treatment records of millions of patients can be stored and computerized and data mining techniques may help in answering several important and critical questions related to health care.[9]

Neural network is a set of connected input/output units and each connection has a weight present with it. During the learning phase, network learns by adjusting weights so as to be able to predict the correct class labels of the input tuples[10]. Neural networks have the remarkable ability to derive meaning from complicated or imprecise data and can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. These are well suited for continuous valued inputs and outputs. For example handwritten character reorganization, for training a computer to pronounce English text and many real world business problems and have already been successfully applied in many industries [12-14][59].

II. RELATED WORK

Up to now, various authors who diagnosed the heart diseases and achieved dissimilar probabilities for different methods of data mining techniques. The current novel variety of work is connected and initiated to diagnose heart diseases by data mining techniques. Heon Gyu Lee et al. [15] proposed the latest utilization techniques of Statistical and classification for the extension of the multi-parametric feature with linear and nonlinear characteristics of Heart Rate Variability (HRV), which have also been assessed for the three recumbent positions, to be accurate the supine, the left and right side locations of the HRV. There are several classifiers such as Support Vector Machine (SVM) [16], Classification based on Multiple Association Rules (CMAR) [17], Bayesian classifiers [18], Decision Tree (C4.5) and which has been surmounted the other classifiers too, have been experimented for the assessment of the linear and nonlinear characteristics of the HRV indices[19].

There are various classification algorithms have been employed on Turkoglu's valvular heart disease data set and high classification accuracies have been reported in the last decade [20-25]. Turkoglu's valvular heart disease data set was obtained from Firat Medical Center. A detailed description for the data set will be given in the next section. The valvular heart disease data set was firstly utilized in [20] where Turkoglu et al. fulfilled an expert diagnosis system which uses backpropagation artificial neural networks (BPANN) classifier. The performance evaluation of the realized system was evaluated by classification accuracy and the correct classification rate was about 94% for normal subjects and 95.9% for abnormal subjects. Later, Turkoglu et al. suggested an intelligent system for detection of heart valve disease based on wavelet packet neural networks (WPNN) [21]. The reported correct classification rate was about 94% for abnormal and normal subjects. Recently, Comak et al. investigated the use of least-square support vector machines (LS-SVM) classifier for improving the performance of the Turkoglu's proposal [22]. They intended to realize a comparative study. Classification rates of the examined classifiers were evaluated by ROC curves based on the terms of sensitivity and specificity. The application results showed that according to the ROC curves, the LS-SVM classifier

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 9, September 2021

DOI: 10.17148/IJARCCE.2021.10918

performance was comparable with ANN, but the training time of LS-SVM is shorter than that of the ANN and it can always converge the same solution while ANN cannot. According to these results, LS-SVM's training time is about 13 times shorter than ANN's training time. This is an important difference. Because, LS-SVMs are trained only depending on support vectors, not by whole training data set. In addition, LS-SVM can overcome the overfitting much successfully than ANN.

More recently, Uguz et al. performed a biomedical system based on Hidden Markov Model for clinical diagnosis and recognition of heart valve disorders [23]. The fulfilled methodology was also used the database of Turkoglu et al. In the presented study, continuous HMM (CHMM) classifier system was used. Single Gaussian model was preferred to determine emission probability. The methodology was composed of two stages. At the first stage, the initial values of average and standard deviation were calculated by separating observation symbols into equal segments according to the state number and using observation symbols appropriate to each segment. At the second stage, the initial values of average and standard deviation were calculated by separating observation symbols into the clusters (FCM or K-means algorithms) that have equal number of states and using observation symbols appropriate to the separated clusters. The implementations of the experimental studies were carried out on three different classification systems such as CHMM, FCM–K-means/CHMM and ANN. These experimental results were obtained for specificity and sensitivity rates 92% and 94% for CHMM, 92% and 97.26% for FCM–Kmeans/CHMM), respectively. Finally, Sengur et al. investigated the use of principal component analysis (PCA), artificial immune system (AIS) and fuzzy k-NN to determine the normal and abnormal heart valves from the Doppler heart sounds [24]. For reducing the complexity, PCA was used. In the classification stage, AIS and fuzzy k-NN were used. To evaluate the performance of the methodology, a comparative study was realized by using a data set containing 215 samples. The validation of the method was measured by using the sensitivity and specificity parameters; 95.9% sensitivity and 96% specificity rate was obtained. Sengur et al. also investigated the use of Linear Discriminant Analysis (LDA) and Adaptive neuro-fuzzy inference system (ANFIS) for clinical diagnosis and recognition of heart valve disorders [25]. The validation of the method is measured by using the sensitivity and specificity parameters. 95.9% sensitivity and 94% specificity rate was obtained.

Carlos Ordonez [26] describes the inhibited problem to identify and to predict the rules of association for the heart disease. It includes medical records of the heart disease patients with attributes for the risk factors, measurements of heart perfusion and narrowed artery. The three constraints were introduced to reduce the number of patterns, they are as follows:

a) The attributes have to appear on one side of the rule only.

b) The rule separates the attributes into the uninteresting groups.

c) The number of attributes from the rule is controlled by the medical records of the people of heart disease finally.

Further falling the running time as per the experiments illustrated the constraints of showing rules have been extremely decreased the number. The author, Carlos Ordonez, anticipated the presence or the absence of heart disease in four specific heart arteries into the two groups of rules. According to the work of Franck Le Duff et al. [27], it is executed for each medical procedure or medical problem. Boleslaw Szymanski et al. [28] projected a novel heuristic for the ability of computation of sparse kernel in SUPANOVA. It was applied to a standard Boston housing market dataset and to the discovery of the heart diseases in the population generally major topic of striking with the aid of a novel, non-invasive measurement of the heart activities on the basis of attractive field generated by the human heart. 83.7% predictions on the results were correct, in this manner outperforming the results which were obtained through Support Vector Machine and equivalent kernels. The spline kernel yielded good results equally on the standard Boston housing market dataset.

Kiyong Noh et al. [29] has been placed forth a classification method for the extraction of multi parametric features by assessing HRV from ECG, the data pre-processing and the heart disease pattern. The proficient FP-growth method was the foundation of this method which is an associative. They accessible a rule cohesion measure that allows a strong press on pruning patterns in the pattern of generating method as the volume of patterns created could probably be huge. Kasbe [30] has also given most of the information about dengue fever, one of the most dangerous diseases of the world. They have also mentioned first vaccine of dengue.

III. EXPERT SYSTEMS IN THE FIELD OF MEDICINES

To diagnose the heart failure cause different popular methods used are MRI, Doppler and Expert System. MRI can provide clear three dimensional images of the heart. Doppler technique has gained much more interest since Satomura first demonstrated the application of the Doppler effect to the measurement of blood velocity in 1959. However the factor such as calcified disease or obesity often results in a diagnostically unsatisfactory. Doppler techniques assessment and therefore, it is sometimes necessary to assess the spectrogram of the Doppler Shift signals to elucidate m the degree of the disease. Many studies have been implemented the classify Doppler signals in the pattern recognition field [31-33].

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 9, September 2021

DOI: 10.17148/IJARCCE.2021.10918

An expert system is a computer program that incorporates knowledge to solve complex problems and can either replace or assist a human expert. Many expert systems have been developed to diagnose diabetes and heart diseases, where diagnosis is complex and involves up-todate parameters. Conigliaro et al. in proposed an expert system for venous insufficiency in the human body using statistical evaluation to detect a set of particular symptoms among others [34].
Expert system is a intelligent program which holds the accumulated knowledge of one or more domain experts. There are many types of expert system currently exist. MYCIN expert system is used in medical field for diagnosis of blood disorders. DESIGN ADVISOR is another expert system used in processor chip design to give advice to designer about component placement, minimizing chip size etc. PUFF expert system is also used in medical system for diagnosis of respiratory condition of patient. PROSPECTOR expert system is used by geologists to identify sites for drilling or mining. DENDRAL expert system is used to identify the structure of chemical compounds. LITHIAN expert system gives advice to archaeologists to examiner stone tools. Expert system having three main components knowledge base, inference engine and user interact. Knowledge base is the collection of facts and rules which describe all the knowledge about problem domain. The inference engine is used to choose the appropriate facts and rules to apply during user query. Where as user interface takes the user query in a readable form and passes it to the inference engine. It then displays the result to the user. Based on much more useful but it has some limitation like limited domain, no current updation, no system self learning, no common sense, expert needed to setup and maintain. But even though it is used world wide because they are not always available, can be used anytime anywhere, human experts are not 100% reliable or consistent. Expert may not good for explanation of decisions and cost effective. While using the expert system some legal and ethical issues we need to follow to set the responsibility [35].

i) Support Vector Machine (SVM):

Support Vector Machines (SVMs) are state of the art pattern recognition techniques whose foundation stem from statistical learning theory. However, the scope of SVMs goes beyond pattern recognition because they can also handle two more learning problems i.e. regression estimation and density estimation. An SVM is a general algorithm based on guaranteed risk bounds of statistical learning theory i.e. the so called structural risk minimization principle. It is a learning machine capable of implementing s set of functions that approximate best the supervisor's response with an expected risk bounded by the sum of the empirical risk and Vapnik – Chevonenkis (VC) confidence. Recent advances in statistics, generalization theory, computational learning theory, machine learning and complexity have provided new guidelines and deep insights into the general characteristics and nature of the model building/learning/fitting process. Some researchers have pointed out that statistical and machine learning models are not all that different conceptually. Many of the new computational and machine learning methods generalize the idea of parameter estimation in statistics. Among these new methods, Support Vector Machines have attracted most interest in the last few years.
Support vector machine (SVM) is a novel learning machine introduced first by Vapnik. It is based on the Structural Risk Minimization principle from computational learning theory. Hearst et al. positioned the SVM algorithm at the intersection of learning theory and practice: ''it contains a large class of neural nets, radial basis function (RBF) nets, and polynomial classifiers as special cases. Yet it is simple enough to be analyzed mathematically, because it can be shown to correspond to a linear method in a high dimensional feature space nonlinearly related to input space.'' In this sense, support vector machines can be a good candidate for combining the strengths of more theory-driven and easy to be analyzed conventional statistical methods and more datadriven, distribution free and robust machine learning methods[36].
In the last few years, there have been substantial developments in different aspects of support vector machine. These aspects include theoretical understanding, algorithmic strategies for implementation and real life applications. SVM has yielded excellent generalization performance on a wide range of problems including bioinformatics, text categorization, image detection, etc. These application domains typically have involved high-dimensional input space, and the good performance is also related to the fact that SVM's learning ability can be independent of the dimensionality of the feature space.
The SVM approach has been applied in several financial applications recently, mainly in the area of time series prediction and classification. A recent study closely related to our work investigated the use of the SVM approach to select bankruptcy predictors. They reported that SVM was competitive and outperformed other classifiers (including neural networks and linear discriminant classifier) in terms of generalization performance [37]. In this study, we are interested in evaluating the performance of the SVM approach in the domain of heart disease in comparison with that of Radial Basis Function in neural networks.[43-45]

ii) Feedforward Backpropagation (FFBP) :

The Feedforward Backpropagation is one of the most studied neural network by the scientific community and the most common used in many medical applications. The feedforward, back-propagation architecture was developed in the early 1970's by several independent sources (Werbor; Parker; Rumelhart, Hinton and Williams) [38]. This independent codevelopment was the result of a proliferation of articles and talks at various conferences which stimulated the entire

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 9, September 2021

DOI: 10.17148/IJARCCE.2021.10918

industry. Currently, this synergistically developed back-propagation architecture is the most popular, effective, and easy to earn model for complex, multi-layered networks [39-41]

This network is used more than all other combined. It is used in many different types of applications. This architecture has spawned a large class of network types with many different topologies and training methods. Its greatest strength is in non-linear solutions to ill-defined problems. The typical back-propagation network has an input layer, an output layer, and at least one hidden layer. There is no theoretical limit on the number of hidden layers but typically there is just one or two [42].

iii) Radial Basis Function (RBF)

Radial basis function (RBF) networks wereintroduced into the neural network literature by Broomhead and Lowe Broomhead, Lowe, 1988). TheRBF network is similar to a general feed- forward neural network trained using the back-propagation scheme . It has three layers of neurons, namely input, hidden and

output. However it uses only one hidden layer, each neuron in which operates as the Gaussian transfer function, as against the sigmoid function.[51]

The major task of RBF network design is to determine centers. The easiest way to do so is to choose the centers randomly from the training set. Another approach is to use the k-means technique consisting of clustering the input training set into groups and choose the mean of each group as the center. Also, the centers can be treated as a network parameter along with wi and adjusted through error-correction training. After the center is determined, the connection weights wi between the hidden layer and output layer can be obtained through ordinary back-propagation-based training [46-47].

iv) Generalized Regression Neural Network (GRNN)

A GRNN is a variation of the radial basis neural networks, which is based on kernel regression networks. A GRNN does not require an iterative training procedure as back propagation networks. It approximates any arbitrary function between input and output vectors, drawing the function estimate directly from the training data. In addition, it is consistent that as the training set size becomes large, the estimation error approaches zero, with only mild restrictions on the function [52].

A GRNN consists of four layers: input layer, pattern layer, summation layer and output layer. The number of input units in input layer depends on the total number of the observation parameters. The first layer is connected to the pattern layer and in this layer each neuron presents a training pattern and its output. The pattern layer is connected to the summation layer. The summation layer has two different types of summation, which are a single division unit and summation units. The summation and output layer together perform a normalization of output set. In training of network, radial basis and linear activation functions are used in hidden and output layers. Each pattern layer unit is connected to the two neurons in the summation layer, S and D summation neurons. S summation neuron computes the sum of weighted responses of the pattern layer.[53]

IV. DATABASE USED TO IMPLEMENT NEURAL NETWORK TECHNIQUES

The data is collected from daily OPD session while doctor examining the patients. The symptoms and information about patients details like Previous History(P1), Present History(P2), Personnel History(P3), Physical Examination(P4), Cardio Vascular System(CVS), Respiratory Rate(RS), Per Abdomen(PA), Central Nervous system(CVS), ECG and Blood Investigation(BI). The main point is ECG from which the patient can easily diagnose whether the patient is having heart problem or not.

All 300 patients data collected regarding heart disease and the data are prepared in different Excel Sheets which contains codes of each individual disease, history and symptoms. In one excel file 13 sub-sheets are taken for each field of information such as for Previous History (P1), for Present History the second sub-sheet and the name is given (P2), for Personnel History (P3) the third subsheet is taken, like this the data collection has 13 different subsheets for different fields. All the fields are taken under the supervision of the Cardiologist The code is given to each symptoms, physical examination parameter or diseases in each sub-sheet for experimental work. On this data some pre-processing i.e. normalization, coding and decoding methods are applied for the expected output [48-49].

In table 1, the Previous History (P1) has represented with 1 to 18 different diseases of total 300 heart patients and it is defined by the codes respectively from 1 to 18. The code 1 which represents Hypertension, Code 2 represents Diabetes Mallitus like this it contains 18 different diseases. Some of them are as shown in table 1.

Table 1. Previous History of Patients

Code	Name of Disease
1	Hypertension
2	Diabetes Mallitus

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 9, September 2021

DOI: 10.17148/IJARCCE.2021.10918

3	TB
4	Bronchial Asthama
5	Hyperthyroidism

In table 2, Present History (P2) and the symptoms present in P2 are represented by Codes. The Code 1 which represents Chest Pain/Discomfort, Code 2 represents Retrosternal Pain like this it contains 29 different symptoms. Some of the symptoms are shown in table 2.

Table 2. Present History of patients

Code	Symptoms
1	Chest Pain/Discomfort
2	Retrosternal Pain
3	Palpitation
4	Breathlessness
5	Sweating

In table 3, Personnel History (P3) and the information present in P3 are represented by codes for different bad habits. The Code 1 which represents Smoking, Code 2 represents Tobacco like this 4 different bad habits are taken and specified by 1 to 4 codes. Some of the personnel history parameters are given below.

Table 3. Personnel History

Code	Symptoms
1	Smoking
2	Tobacco
3	Alcohol
4	Nil

In table 4, Physical Examination (P4) and the information present in P4 are represented by codes for different physical parameters. The Code 1 which represents Consciousness, Code 2 represents Orientation like these 25 different physical parameters and specified by 1 to 25 codes for each parameter. Some are as shown below in table 4.

Table 4. Physical Examination

Code	Physical Examination
1	Consciousness
2	Orientation
3	Dyspnoea
4	Fever
5	Low Pulse Rate

In table 5, Cardio Vascular System (CVS) and the information present in CVS are represented by codes for different symptoms. The Code 1 which represents Heart Sound, Code 2 represents Normal Heart Rate like this 8 different symptoms and specified by 1 to 8 codes for each symptom. Some are as shown below in table 5.

Table 5. Cardio Vascular System

Code	Symptoms
1	Heart Sounds
2	Normal Heart Rate
3	Tachycardia

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 9, September 2021

DOI: 10.17148/IJARCCE.2021.10918

4	Bradycardia
5	Regular Heart Rhythm

In table 6, Respiratory System (RS) and the information present in RS are represented by codes for different symptoms. The Code 1 which represents Breath Sound preserved, Code 2 represents Breath Sound Reduced like this 5 different symptoms are found and specified as shown in table 6.

Table 6. Respiratory System

Code	Symptoms
1	Breath Sounds Preserved
2	Breath Sound Reduced
3	Basal Crepts
4	No Abnormality Detected (NAD)
5	Ranchi

In table 7, Per-Abdomen (PA) and the information present in PA are represented by codes for different symptoms. The Code 1 which represents Liver (Hepatomegaly), Code 2 represents Spleen (Splenomegaly) like these 6 different symptoms have found and specified by 1 to 6 codes for each symptom. Some are as shown below in table 7.

Table 7. Per Abdomen

Code	Symptoms
1	Liver(Hepatomegaly)
2	Spleen (Splenomegaly)
3	Free Fluid Present
4	Abdominal Distension
5	No Abnormality Detected (NAD)

In table 8, Central Nervous System (CNS) and the information present in CNS are represented by codes for different symptoms. The Code 1 which represents Consciousness, Code 2 represents Orientation like this 5 different symptoms are found and specified by 1 to 5 codes for each symptom. Some are as shown below in table 8.

Table 8. Central Nervous System

Code	Symptoms
1	Consciousness
2	Orientation
3	Focal Deficit
4	No Abnormality Detected (NAD)
5	Restlessness

In table 9, Electro Cardio Gram (ECG) and the information present in ECG are represented through codes for different finding which points to different problems of heart. The Code 1 which represents ST Elevation, Code 2 represents Anterior Wall like this 21 different heart findings are found and specified by 1 to 21 codes for each finding. Some are as shown below in table 9.

Table 9. Electro Cardio Gram (ECG)

Code	ECG Point
1	ST Elevation
2	Anterior Wall
3	Antero Septal

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 9, September 2021

DOI: 10.17148/IJARCCE.2021.10918

4	Inferior
5	Infero Posterior

In table 10, Blood Investigation (BI) and the information present in BI are represented through codes for blood investigation. The Code 1 which represents Cardiac Enzymes (High), Code 2 represents Blood Sugar Test like this 24 different investigations has found and specified by 1 to 24 codes for each investigation in all patient. Some are as shown below in table 10.

Table 10. Blood Investigation

Code	Symptoms
1	Lipid Profile normal
2	Lipid Profile Abnormal
3	Complete Blood Count Normal
4	Leucocytosis
5	Anaemia

In table 11, all the medicines names along with their codes i.e. MID which are prescribed by the doctor to the patients. The medicine sheet contains 52 different medicines which are prescribed by the doctor in different 300 stages. Some are as shown below in table 11.

Table 11. Medicine Names

Code	Medicine Name
1	Alprazolam
2	Amlodepine
3	Aspirin
4	Atenolol
5	Atorvastatin

In table 12, all Patients information such as previous history(P1), P2(Present History), P3(personnel History), P4(Physical Examination), CVS(Cardio Vascular System), RS(Respiratory System), PA(Per Abdomen), CNS(Central Nervous System), ECG(Electrocardiography) and BI(Blood Investigation) which contains all the represented codes that are present in the individual patients.

Table 12. Collection of different details of the individual Heart Patients

Sr. No.	Patient Name	Symptoms and Findings										
		Age	P1	P2	P3	P4	CVS	RS	PA	CNS	ECG	BT
1	A	55M	2	1,2,5,13	4	7,10	8	4	5	4	1,3	14
2	B	58M	2	1,2,8	2	7,8,13,14	8	4	5	4	2	7
3	C	60M	8	5,7,13	4	1,6,12	8	4	5	4	9	14
4	D	60M	1,2	4,5	4	1,2,7,13,14	3,5	3	5	4	12	4
5	E	56F	1	15,16	4	6,9,12	8	4	5	4	10	2

In table 13, different 52 medicines were used by the doctor on total 300 patients. All the medicines are prescribed by the doctor. In this table the medicines codes are used as the description given in the table 11.

Table 13. All the Medicine codes provided by the doctor to the individual patients.

Sr. No.	Patient Name	MID 1	MID 2	MID 3	MID 4	MID 5	MID 6	MID 7	MID 8	MID 9	MID 10	MID 11	MID 12	MID 13
1	A	2	3	5	6	14	17	19	21	23	25	26	27,29	36
2	B	2	3	5	6	14	16	17	21	23	25	26	27	28

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 9, September 2021

DOI: 10.17148/IJARCCE.2021.10918

3	C	1	5	6	14	25								
4	D	3	5	7	10	11	13	14	17	19	30			
5	E	5	14	15	19									

V. EXPERIMENTAL ANALYSIS

Table 14. The individual data of the patient 1 is defined in binary form

Sr. No.	00000001
Age	0110111
P1	010000000000000000
P2	11001000000010000000000000000
P3	0001
P4	000000100100000000000000
CVS	00000001
RS	00010
PA	000010
CNS	00010
ECG	101000000000000000000
BT	000000000000010000000000

For further training of neural network process the proposed information is coded in binary form (0 or 1). If the symptom is present in the patients at particular position at that point it is defined by one (1) and if the symptoms or disease is not present at that position it is placed by Zero (0). Suppose for example in the field P2 (present history) there are total 29 symptoms present and the patient no 1 is having the symptom 1,2,5 and 13 so these locations are defined by 1 (one) and all other symptoms are 0 (zero). In such a way all the fields are defined. All the parameter that we consider in medical prescription like Sr. No., ,age , P1, P2,P3,P4,CVS, RS, PA, CNS, ECG and BT are converted in binary number where this is used in neural network for train the neurons for achieving better result.
Using this sequence of binary format we were not getting appropriate result. Therefore we have change the order of fields as per suggestion of the doctor because the doctors are prescribing the medicines on the basis of the ECG and blood investigation. So the order of ECG is changed from field no. 9 to field no. 1 and after ECG we have taken Blood Investigation and rest of the fields are same and at last age is placed. Due of reshuffling of the fields we got satisfactory result upto 97% by using Radial Basis Function. For this expert system total 52 different medicine are prescribed by the Doctor and if the medicine is present at that position it is defined by one (1) and if it is absent at that position it is defined by Zero(0). Similarly for patient one the prescribed medicine are defined as:[50]

Original Medicines given by doctor:
A) 1,3,5,6,14.17,19,21,23,25,26,27,29,36
B) 2,3,5,6,14,16,17,21,23,25,26,27,28
C) 1,5,6,14,25
D) 3,5,7,10,11,13,14,17,19,30
E) 3,14,15,19

Medicines given by GRNN
A) 1,3,5,6,14, 16,17,21,23,25,26,27
B) 1,3,5,6,11,14,16,17,21,22,23,24,25,26,27,36
C) 1,3,5,6,11,17,21,25
D) 1,3,5,6,11,13,14,17,21,25
E) 3,5,6,14

Medicines given by RBF :
A) 3,5,6,14,16,17,21,23,25,26,27,28,29
B) 1,3,5,6,11,14,16,17,21,22,23,24,25,26,27,28,36
C) 1,3,5,6,14,21,25
D) 1,3,5,6,11,13,14,17,21,25,30
E) 3,5,6,14,15

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 10, Issue 9, September 2021
DOI: 10.17148/IJARCCE.2021.10918

Medicines given by the Expert system using SVM

A) 1,3,5,6, 16,1718,21,23,25,26,27,28,29
B) 1,3,5,6,16,17,18,21,23,25,26,27,28,29
C) 1,3,5,6,11,14,21,22,23,24,25,26,27
D) 1,3,5,6,13,14,17,2122,23,25,26,27,28
E) 1,2,3,5,14.

Medicines given by the Expert system using FFBP

A) 1,3,5,6, 14, 16, 17,21,23,25,26,27,28
B) 1,3,5,6,14,25
C) 1,3,5,6,14,25
D) 1,3,5,6,14,25
E) 1,3,5,6,14,25

VI. CONCLUSION

This case study has presented all the overview and basic introductory discussion for diagnosis of heart disease using different neural network techniques such as Feedforward Backpropagation, Support Vector Machine, Generalized Regression Neural Network, Radial Basis function which are more important for the researcher to understand as beginners. The diagnosis performance of the reviewed papers shows the advantage of the system that it is rapid, easy to operate, non invasive and not expensive. The working prototype model in the field of heart diagnosis can use the system. This expert system also helps in rural area for training beginner's doctors and medical students who work in the field of heart diagnosis.

It is found that the output of the training and testing data by using Feedforward Backpropagation, Support Vector Machine, Generalized Regression Neural Network, is not satisfactory as per the result verified by the doctor. So in this case study, found that only diagnosis of heart disease using Radial Basis Function technique is satisfactory as per the result verified by the doctor. In future the researcher can extend this work using Recurrent Neural Network Techniques as this research work is not done by this technique.

REFERENCES

[1] Frase, H.S.F.et.al., "Comaring complex diagnoses a formative evaution of the heart disease.
[2] Anupriya Kamble, Shaikh Abdul Hannan, Ramesh Manza and Mohammad Eid Alzahrani, "Prediction of Prediabetes, No Diabetes and Diabetes Mellitus -2 using Simple Decision Tree Classification" Springer FICR International Conference on Rising Threats in Expert Applications and Solutions. 2020 at IIS University, 17-19 Jan, 2020 Jaipur.
[3] Eliane Rich and KevinKnight – Artificial Intelligence – Second Edition Mcgraw Hill , 1983
[4] Anupriya Kamble, Shaikh Abdul Hannan, Yogesh Rajput and Ramesh Manza, "Prediction of Prediabetes, No Diabetes and Diabetes Mellitus-2 using Pattern Recognition", Springer FICR International Conference on Rising Threats in Expert Applications and Solutions. 2020 at IIS University, 17-19 Jan, 2020 Jaipur.
[5] Long, W., et.al., "Developing a program for Tracking Heart Failure", MIT Lab for Computer Science, Cambridge, MA(2001).
[6] I. Turkoglu, A. Arslan, E. Ilkay, "An expert system for diagnosis of the heart valve diseases", Expert Systems with Applications vol.23, pp. 229–236, 2002.
[7] Anupriya Kamble, Abdul Hannan, Yogesh, Dnyaneshwari, "Association Detection of Regular Insulin and NPH Insulin Using Statistical Features", Second International Conference on Cognitive Knowledge Engineering, 21-23 December 2016 (ICKE-2016) , Aurangabad, Maharashtra, India ISBN 978-93-80876-89-4.
[8] Shaikh Abdul Hannan, V. D. Bhagile, R. R. Manza, R. J. Ramteke, "Diagnosis and Medical Prescription of Heart Disease Using Support Vector Machine and Feed forward Back propagation technique", International Journal on computer science and engineering, IJCSE – August 2010, Vol. 2, Issue 6, ISSN: 0975–3397.
[9] Aqueel Ahmed, Shaikh Abdul Hannan, "Data Mining Techniques to Find Out Heart Diseases: An Overview", International Journal of Innovative Technology and Exploring Engineering (IJITEE), An ISO 9001:2008 Certified International Journal, Volume-1, Issue-4, September 2012, ISSN: 2278-3075, New Delhi, India.
[10] Eliane Rich and Kevin Knight – Artificial Intelligence – Second Edition Mcgraw Hill, 1983.
[11] Long, W., et. al., "Developing a program for Tracking Heart Failure", MIT Lab for Computer Science, Cambridge, MA (2001).
[12] Frase, H.S.F., et. al., "Comparing complex diagnoses a formative evaluation of the heart disease program", MIT Lab for Computer Science, Cambridge, MA (2001).
[13] Salem, A.M. Raushdy, M and HodHod, R.A., " A rule based expert system for diagnosis of Heart disease", 8th International conference on soft Computing MENDEL, Brno, Czech Republic, June 5-7, PP 258- 263(2002).
[14] Turkoglu, A. Arslan, E. Ilkay, "An expert system for diagnosis of the heart valve diseases", Expert Systems with Applications vol.23, pp. 229–236, 2002.
[15] Y. M. Rajput, Shaikh Abdul Hannan, Mohammed Eid Alzahrani, R. R. Manza, Dnyaneshwari D. Patil, "EEG-Based Emotion Recognition Using Different Neural Network and Pattern Recognition Techniques – A Review" Vol 7, Issue 1, Jan 2019, EISSN: 2347-2693, India.
[16] Yogesh Rajput, Shaikh Abdul Hannan, Mohammed Eid Alzaharani, D. Patil, Ramesh Manza, "Design and Development of New Algorithm for person identification Based on Iris statistical features and Retinal blood Vessels Bifurcation points", International Conference on Recent Trends in Image Processing & Pattern Recognition (RTIP2R), December 21-22, 2018, India.

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 9, September 2021

DOI: 10.17148/IJARCCE.2021.10918

[17] Santosh K. Maher, Sumegh Tharewal, Abdul Hannan, K. V. Kale "Review on HRV based Prediction and Detection of Heart Disease", International Journal of Computer Application (IJCA), Vol. 179, Number 46, June 2018, ISSN 0975-8887, USA.
[18] Yogesh, Shaikh Abdul Hannan, Rahul Sagar, Kishor Jave, "Identification and Counting Trees from Oil Palm Plantations Using Digital Image Processing Techniques", International Journal of Engineering Research & Technology (IJERT), Vol. 6 Issue 05, May – 2017, ISSN: 2278-0181.
[19] Heon Gyu Lee, Ki Yong Noh and Keun Ho Ryu, "Mining Bio-signal 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.
[20] Cristianini, N. and Shawe-Taylor, J. "An introduction to Support Vector Machines", Cambridge University Press, Cambridge, 2000.
[21] Li, W., Han, J. and Pei, J., "CMAR: Accurate and Efficient Classification Based on Multiple Association Rules", In: Proc. of 2001 International Conference on Data Mining. 2001.
[22] Chen, J., Greiner, R., "Comparing Bayesian Network Classifiers", In Proc. of UAI-99, pp. 101– 108, 1999.
[23] Quinlan, J., "C4.5: Programs for Machine Learning", Morgan Kaufmann, San Mateo, 1993.
[24] I. Turkoglu, A. Arslan, E. Ilkay, "An expert system for diagnosis of the heart valve diseases", Expert Systems with Applications vol.23, pp. 229–236, 2002.
[25] I. Turkoglu, A. Arslan, E. Ilkay, "An intelligent system for diagnosis of heart valve diseases with wavelet packet neural networks", Computer in Biology and Medicine vol.33 pp.319–331, 2003.
[26] E. Comak, A. Arslan, I. Turkoglu, "A decision support system based on support vector machines for diagnosis of the heart valve diseases", Computers in Biology and Medicine vol.37, pp.21–27, 2007.
[27] H. Uguz, A. Arslan, I. Turkoglu, "A biomedical system based on hidden Markov model for diagnosis of the heart valve diseases", Pattern Recognition Letters vol.28 pp.395–404, 2007.
[28] A. Sengur, "An expert system based on principal component analysis, artificial immune system and fuzzy k-NN for diagnosis of valvular heart diseases", Computers in Biology and Medicine vol.38 pp.329– 338, .2008.
[29] A. Sengur, "An expert system based on linear discriminant analysis and adaptive neuro-fuzzy inference system to diagnosis heart valve diseases", Expert Systems with Applications vol.35 pp.214–222, 2008.
[30] Carlos Ordonez, "Improving Heart Disease Prediction Using Constrained Association Rules," Seminar Presentation at University of Tokyo, 2004.
[31] Franck Le Duff, Cristian Munteanb, Marc Cuggiaa, Philippe Mabob, "Predicting Survival Causes After Out of Hospital Cardiac Arrest using Data Mining Method", Studies in health technology and informatics, Vol. 107, No. 2, pp. 1256-9, 2004.
[32] Boleslaw Szymanski, Long Han, Mark Embrechts, Alexander Ross, Karsten Sternickel, Lijuan Zhu, "Using Efficient Supanova Kernel For Heart Disease Diagnosis", proc. ANNIE 06, intelligent engineering systems through artificial neural networks, vol. 16, pp:305-310, 2006.
[33] Kiyong Noh, Heon Gyu Lee, Ho-Sun Shon, Bum Ju Lee, and Keun Ho Ryu, "Associative Classification Approach for Diagnosing Cardiovascular Disease", Springer,Vol:345, pp. 721- 727, 2006.
[34] Tanmay Kasbe, Ravi Singh Pippal, "Dengue Fever: State-of-the-Art Symptoms and Diagnosis," International Journal of Computer Sciences and Engineering, vol. 04, no. 06, 2016, pp. 26-30.
[35] Shaikh Abdul Hannan, R.R. Manza and R.J. Ramteke, "Heart Disease relationship between Disease, Symptoms, Medicine and its side effects", Journal of Advance Research In Computer Engineering: An International Journal ", July to December 2009, Serials Publication, New Delhi, India, ISSN 0973-6794.
[36] J. Galindo, P. Tamayo, Credit risk assessment using statistical and machine learning: basic methodology and risk modeling applications, Computational Economics 15 (1 – 2) (2000) 107–143.
[37] Shaikh Abdul Hannan, Pravin Yannawar, R.R. Manza and R.J. Ramteke, "Expert System Data Collection Technique for Heart Disease", in International Journal of Innovative Research in Science and Techniques (IJIRST), Vol 1, No.1 , Jan – June 2010, PP 31-35, ISSN:2229-3116, India.
[38] D. Michie, D.J. Spiegelhalter, C.C. Taylor, Machine Learning, Neural and Statistical Classification, Elis Horwood, London, 1994.
[39] Shaikh Abdul Hannan, Ramesh Manza, R. J. Ramteke, "Relationship between Heart Disease and Symptoms", International Journal of Computational Intelligent, Vol. 3, No.2, JulyDecember 2009, pp. 289-292, ISSN 0974-5807.
[40] Suresh Nagar, Arvind Kumar Jain, "Expert System for Diagnosis of Heart Disease : A Review", International Journal of Computer Sciences and Engineering, Vol. 4, Issue 8, August 2016, pp. 45-47, E-ISSN 2347-2693.
[41] http://elearning.najah.edu/OldData/pdfs/Intro%20Expert% 20 Systems%20test-me.co.uk. ppt#262,5, Components of an Expert System.
[42] Shaikh Abdul Hannan, V.D. Bhagile, R. R. Manza and R.J. Ramteke, "Expert System for Diagnosis and Appropriate Medical Prescription of Heart Disease Using Radial Basis Function", CiiT International Journal of Artificial Intelligent Systems and Machine Learning, August 2010, ISSN 0974–9667 & Online: ISSN 0974–9543.
[43] N. Cristianini, J. Shawe-Taylor, An Introduction to Support Vector Machines, Cambridge Univ. Press, Cambridge, New York, 2000.
[44] Veerendra Singh and S. Mohan Rao, "Application of image processing and radial basis neural network techniques for ore sorting and ore classification". Mineral. Eng., 18: 1412-1420, 2005.
[45] http://cat.inist.fr/?aModele=afficheN&cpsidt=172892 10.
[46] Shaikh Abdul Hannan, V. D. Bhagile, R. R. Manza, R. J. Ramteke, "Development of an Expert System for Diagnosis and appropriate Medical Prescription of Heart Disease Using Support Vector Machine and Radial Basis Function", International Journal of Computer Science and Information Security, (IJCSIS) August issue (Vol. 8 No. 5), 2010, Pages/record No.: 245-254. ISSN: 19475500.
[47] Haralambos Sarimveis, Philip Doganis, Alex Alexandridis, "A classification technique based on radial basis function neural networks", Advances in Engineering Software vol.37, pp.218–221, 2006.
[48] T. Joachims, Text categorization with support vector machines, Proceedings of the European Conference on Machine Learning (ECML), 1998.
[49] E. Osuna, R. Freund, F. Girosi, Training support vector machines: an application to face detection, Proceedings
[50] Shaikh Abdul Hannan, R. R. Manza, R. J. Ramteke, "Generalized Regression Neural Network and Radial Basis Function for Heart Disease Diagnosis", International Journal of Computer Applications (IJCA) Vol. 7, No. 13, October 2010 Edition. New York, USA. ISSN: 09758887.
[51] Mohammad Eid Alzaharani,Shaikh Abdul Hannan, " Diagnosis and Medical Prescription of Heart disease using FFBP, SVM and RBF, KKU Journal of Basic and Applied Sciences, Vol 5(1), 2019.
[52] Frawley and Piatetsky-Shapiro, 1996. Knowledge Discovery in Databases:An Overview. The AAAI/MIT Press, Menlo Park, C.A.
[53] Kim B, Lee DW, Parka KY, Choi SR, Choi S. Prediction of plasma etching using a randomized generalized regression neural network. Vacuum 2004;76:37–43
[54] https://opensource.com/resources/big-data

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 9, September 2021

DOI: 10.17148/IJARCCE.2021.10918

[55] Shaikh Abdul Hannan, "An Overview of Big Data and Hadoop", International Journal of Computer Application", Volume 154, Number 10, ISSN – 0975-887, November 2016, New York, USA.

[56] Jang JSR, Sun CT, Mizutani E. Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence, Prentice Hall, Upper Saddle River, New Jersey, USA; 1997 [Chapter 9].

[57] Swati Saxena, Shaikh Abdul Hannan, "A Quaitative Review on Intervention of Robotics in Medical Science", International Journal of Computer Science and Information Security, Vol 19, No. 7, July 2021.

[58] Long, W. J., Naimi, S., & Criscitello, M. G. (1992). Development of a knowledge base for diagnostic reasoning in cardiology. Computers in Biomedical Research, 25, 292–311.

[59] Azuaje, F., Dubitzky, W., Lopes, P., Black, N., & Adamsom, K. (1999). Predicting coronary disease risk based on short- term RR interval measurements: A neural network approach. Artificial Intelligence in Medicine, 15, 275–297.

