6CS014 - Complex System

Heart Disease Prediction using Neural Network

Student ID: 2042471

Student Name: Sujan Pyakurel

University Email: S.Pyakurel@wlv.ac.uk

Award/Course: BSc Hons Computer Science

Cohort/Batch: 4

Submission Date: November 28, 2020

Word Count:2,710

Acknowledgement

Abstract

In this research paper, I have studied research papers, journals related to heart disease prediction system using Neural Network. This report explains what heart disease is and different ways to diagnose heart disease using Neural Networks. The main goal of the report is to know about heart disease and ways to detect it.

Contents

Table of Figures

1. Introduction

1.1 Heart Disease

Heart is one of the most significant part of our body. Heart disease might be any conditions that affect functionality of human's heart. The term "heart disease" can be interchanged with the term "cardiovascular disease." Heart disease is a condition of heart in which blood vessels gets constricted or blocked that can cause heart attack, chest pain or stroke. (Mayo Foundation for Medical Education and Research (MFMER), 2020)

Today, the leading cause of deaths in Nepal is Heart disease. The most common type of Cardiovascular disease that causes 34,167 or 20.53% of Nepal's total deaths is Coronary Heart Disease (CHD). The age adjusted death rate is 184.5 per 100,000 of population in the world that ranks Nepal #38 in the world. (World Health Rankings, 2018)

To reduce the risk of heart disease, prediction must be done. Diagnosis is based on signs, symptoms and patient's physical examination. Prediction can only be made by doctors by learning and experience. Significant quantities of complex data about patients, disease diagnosis, electronic records of patients are generated in health care industry. This large volume of data is a precious asset that can be collected and analyzed for developing the system that can make precise prediction of heart disease.

1.2 Artificial Neural Network

One of the major aspects of machine learning is Artificial Neural Network (ANN). "Neural" in NN suggests that they are brain-inspired mechanisms which are inspired to mimic the way that we humans acquire knowledge. Neural Network consists of many hidden layers that has units that can transform inputs into significant thing that superficial layer can use. Patterns that are really too much diverse or extensive for human beings to retrieve and comprehend or guide the computer to identify through coding can be derived through Neural Networks. (Dormehl, 2019)
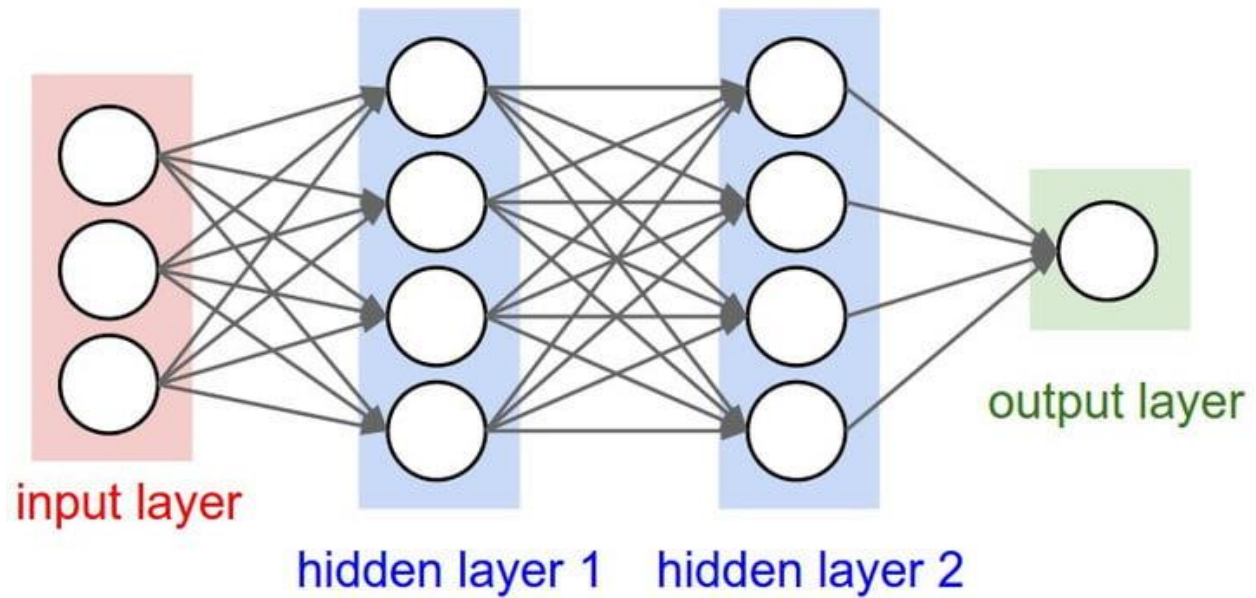
(Dormehl, 2019)

Figure 1. Artificial Neural Network

1.3 Aims

- To help in the precise detection of heart diseases.

1.4 Objectives

- To save time and effort in detection of Cardiovascular Diseases and generate more accurate results in less time
- To avoid false presumptions and unpredictable output.
- To make early prediction of heart diseases so that it can be cured in time before condition of patient gets serious.

2. Literature Review

2.1 Neural Network Based Heart Disease Prediction System

In this analysis article, prediction analytics application is integrated for the detection of patient's heart disease. The diagnosis is premised on a set of training data of heart disease. As 13 parameters traits are used the system utilizes medical terminologies such as sex, blood pressure, and cholesterol. Two more conditions i.e. obesity and smoking are used to achieve more acceptable outcomes, since these attributes are regarded as essential characteristics for heart disease. Multilayer Perception Neural Network (MLPNN) with Back Propagation algorithm (BP) is the method used for system design. (Singh P, 2018)
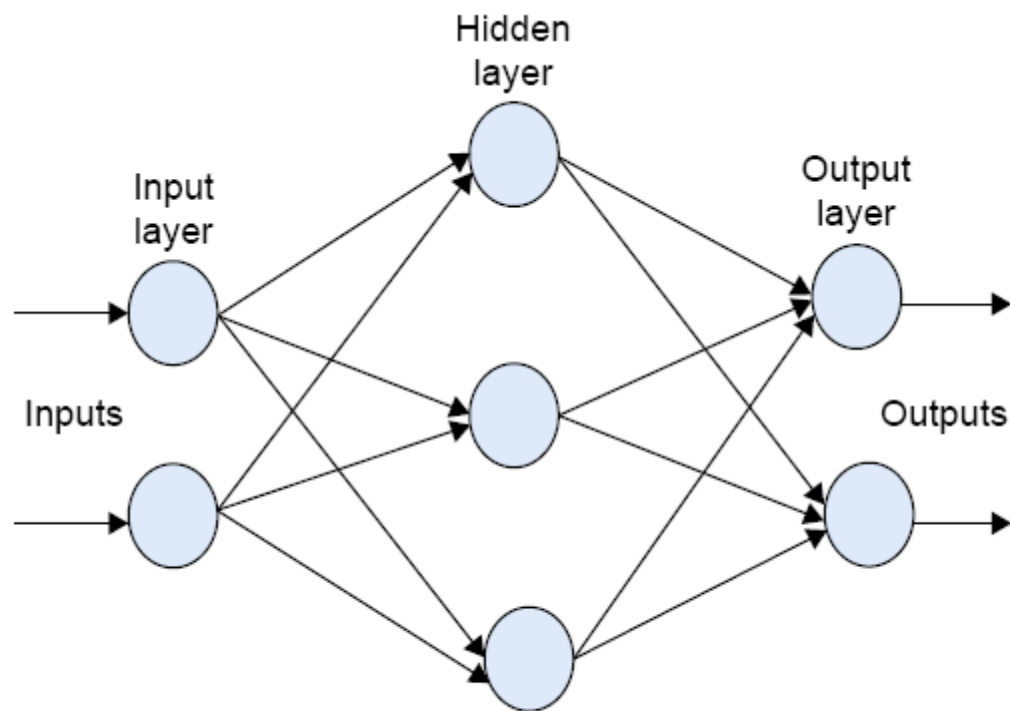
Figure I Multilayer perceptron neural network.

(Singh P, 2018)

Figure 2 Multilayer perception neural network 1Multilayer perception neural network

2.2 Genetic Neural Approach for Heart Disease Prediction

For the detection, analization and extraction of data via complex algorithms, data mining methods are implemented for the discovery of hidden trends in the process of knowledge exploration. Cardiovascular disease is a major risk to human health that might result in death and serious long-term disability. The seriousness of case affects the time required for patients to recover. Diagnosis of Heart disease is convoluted task that needs much experience and knowledge. Advanced data mining approaches along with computer generated data are used to extract information from huge health care data from medical field. The Neural Networks and Genetic Algorithms are used to predict heart disease. Thus, developed system computes the number of concealed nodes for NN that instruct the framework to pick the NN frame properly and uses the worldwide enhancement of genetic algorithm for commencement of neural network. For identification, 12 different characteristics such as age, sex, blood cholesterol etc. are used. This system is believed to be near 98% effective in extent. (Waghulde & Patil, 2014)
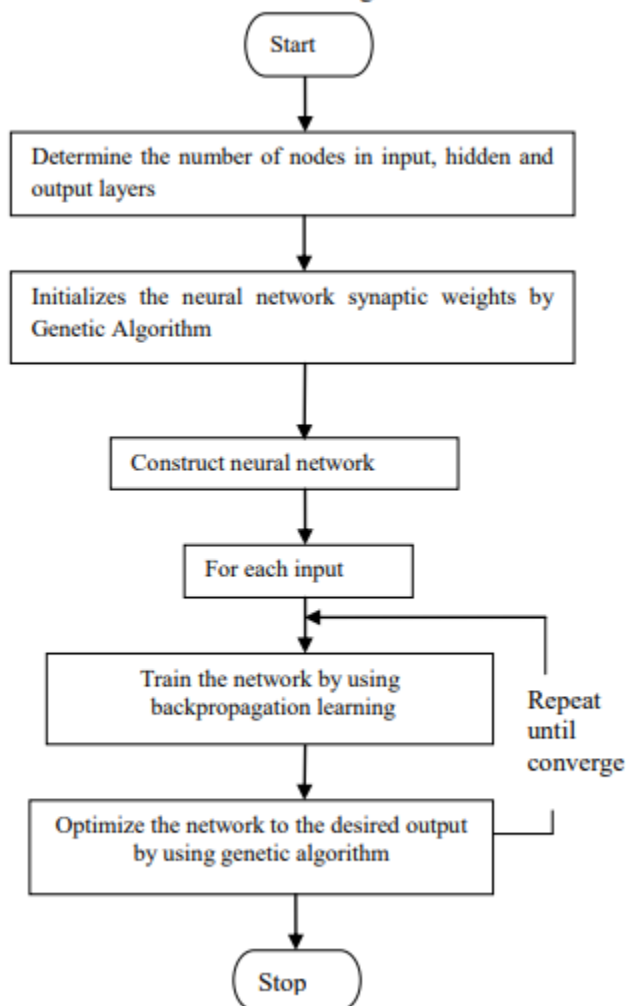
Figure 3 Flowchart of Genetic-Neural Approach for Heart Disease Prediction
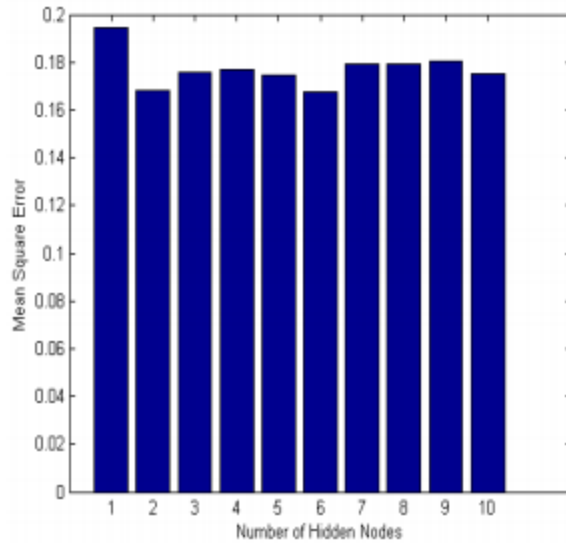
Figure 4Resulted hidden number of nodes for NN

2.3 An Efficient Convolutional Neural Network for Coronary Heart Disease Prediction

This is the study that aims an productive NN with convolutional layers to categorize substantially class-imbalanced observations from NHANES with the only aim to predict Coronary Heart Disease. In class-based success, the used double layered CNN shows responsiveness to the inconsistency with reliable harmony. To develop and generate much efficient system, neural network is trained with high data and then tested on a comparatively smaller amount of data. In this study, two-step approach is adopted to generate more efficient prediction: least absolute shrinkage and selection operator (LASSO) accomplished by large proportion identification of significant attributes. Then, the significant features are fused with the help of fully connected layer that generates the layer to successive convolution stages. The efficiency of the output is 77% to bring the nature of CHD into sets and 81.8% the omission of CHD cases on data checking, that is 85.70% of the datasets taken. (Aniruddha Dutta1, 2020)

Variable Name	Description	Code	Meaning
Gender	Gender of the participant	1	Male
		2	Female
Vigorous Activity 12 years and above	Vigorous activity in last one week or 30 days	1	Yes
		2	No
		3	Unable to do activity
Moderate Activity 12 years and above	Moderate activity in last one week or 30 days	1	Yes
		2	No
		3	Unable to do activity
Diabetes 1 yr and above	Doctor told that the participant has diabetes	1	Yes
		2	No
		3	Borderline
Blood Relative Diabetes 20 yrs and above	Biological blood relatives ever told that they have diabetes	1	Yes
		2	No
Blood Relative Stroke 20 yrs and above	Biological blood relatives ever told that they have hypertension or stroke before the age of 50	1	Yes
		2	No
Coronary Heart Disease 20 yrs and above	Ever told that the participant had coronary heart disease	1	Yes
		2	No

Figure 5 Description of the risk factor independent variables and the dependent variable.
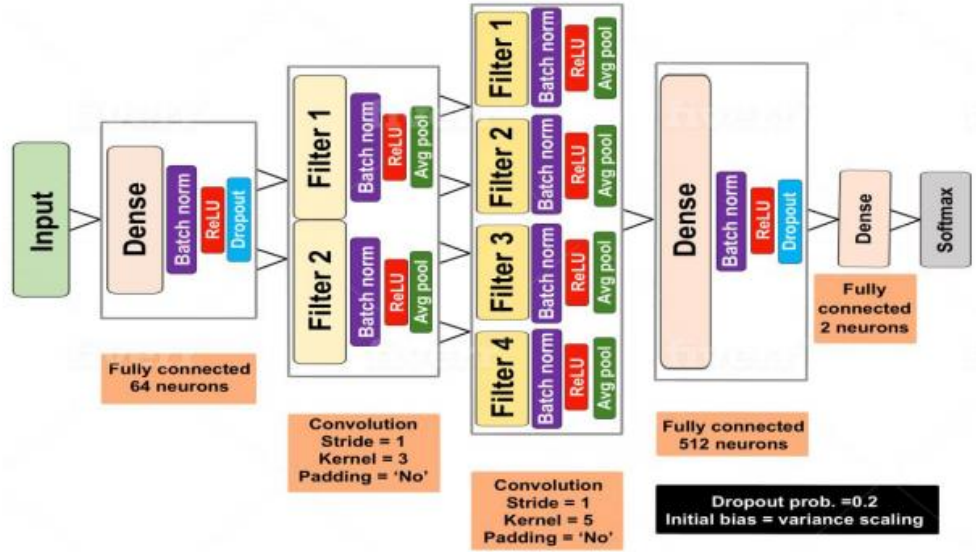
Figure 6 Proposed convolutional neural network architecture.

2.4 An Automated Diagnostic System for Heart Disease Prediction Based on χ2 Statistical Model and Optimally Configured Deep Neural Network

The studies in heart disease prediction using automated artificial NN has been widely proposed. This paper claims that the focus has been made on optimizing functionalities and mitigating of the under fitting and over fitting problems rather than focusing only on preprocessing of features which helps in displaying good results in across both training and testing datasets documentation. This paper gas proposed using χ^2 mathematical formula, thus using exhaustive search method the optimally designed deep NN is scanned. The efficiency of this conceptual hybrid approach which is called χ^2-DNN is calculated by analyzing its achievement with traditional artificial NN and deep NN. The prediction accuracy of the proposed method is more than 93% which is more promising to many of other methods used. (Ali, et al., 2019)
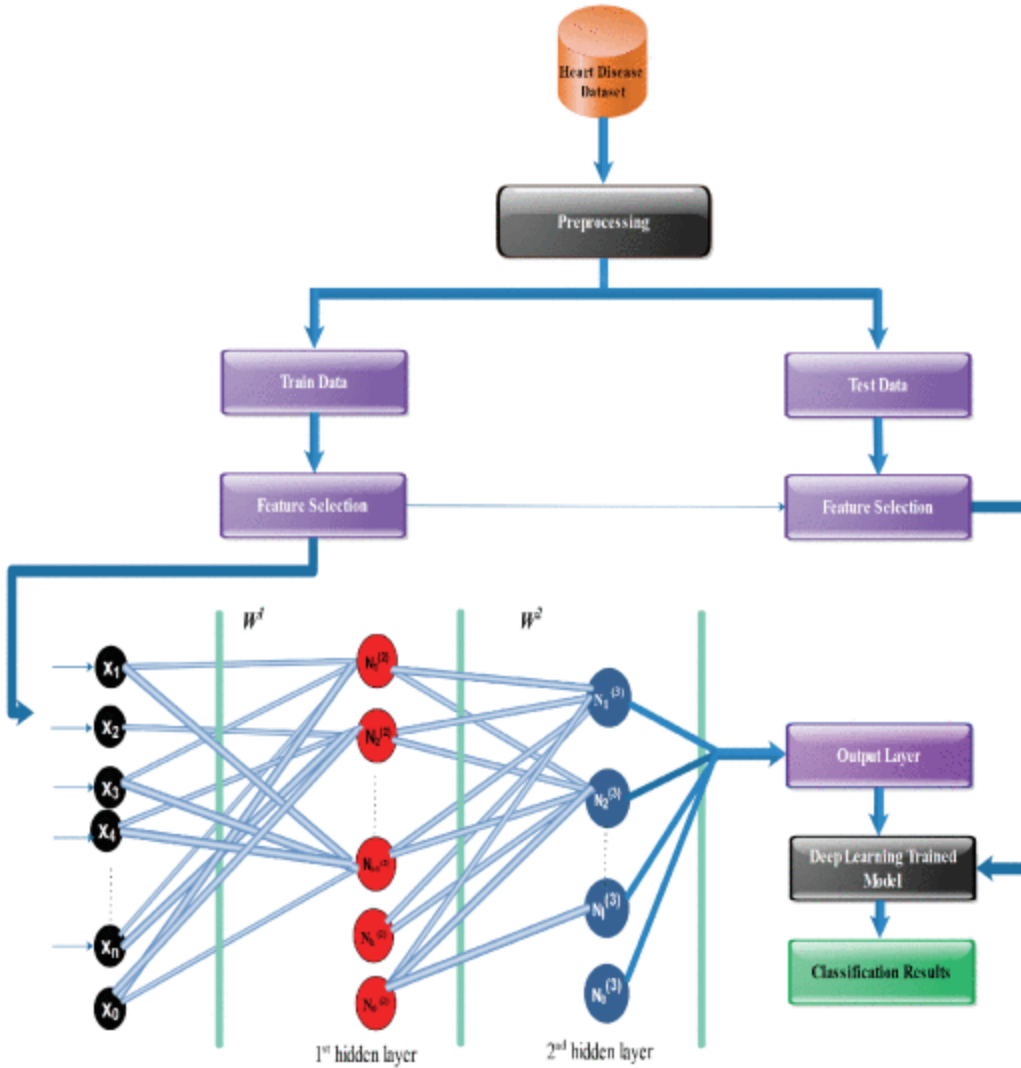
Figure 7 Block diagram of the proposed diagnostic system.

2.5 A Data Mining Approach for Prediction of Heart Disease Using Neural Networks

Experience and knowledge are must to diagnose heart disease. Primitive way of diagnosing heart diseases were doctor's checkup or several examinations such as ECG, Heart MRI and Stress Test etc. which cannot me much effective. To generate effective predictions, the raw data from health care are useful. Information and data stored and managed through computer and advanced data mining techniques are helpful in generating efficient outcomes. NN is commonly used method for the forecasting of heart disease. This research study uses 13 parameters like sex, blood, pressure, cholesterol for the detection of cardiovascular disease

using NN. Extending the parameters to 15 adding obesity and smoking the accuracy increase to reach perfection that is 100%. (Miss. Chaitrali S. Dangare1, 2012)

Sr. no	Attribute	Description	Values
1	age	Age in years	Continuous
2	sex	Male or female	1 = male 0 = female
3	cp	Chest pain type	1 = typical type 1 2 = typical type angina 3 = non-angina pain 4 = asymptomatic
4	thestbps	Resting blood pressure	Continuous value in mm hg
5	chol	Serum cholesterol	Continuous value in mm/dl
6	Restecg	Resting electrographic results	0 = normal 1 = having_ST_T wave abnormal 2 = left ventricular hypertrophy
7	fbs	Fasting blood sugar	$1 \geq 120$ mg/dl $0 \leq 120$ mg/dl
8	thalach	Maximum heart rate achieved	Continuous value
9	exang	Exercise induced angina	0= no 1 = yes
10	oldpeak	ST depression induced by exercise relative to rest	Continuous value
11	slope	Slope of the peak exercise ST segment	1 = unsloping 2 = flat 3 = downsloping
12	ca	Number of major vessels colored by floursopy	0-3 value
13	thal	Defect type	3 = normal 6 = fixed 7 = reversible defect

Figure 8 Description of 13 parameters used.

Sr. no	Attribute	Description	Values
14	obes	obesity	1 = yes 0 = no
15	smoke	smoking	1= past 2 = current 3 = never

Figure 9 Description of newly added parameters

2.6 Heart Disease Prediction using Deep Neural Networks

One of the most severe disease in today's world is heart disease. This paper uses the Neural Network to build a self-operating diagnostic framework for the detection of cardiac dysfunction diseases. Large amount of data from healthcare industry are formulated so that machine learning models can provide effective conclusion in prediction of heart related diseases. A framework that can understand the concepts of evaluating the risk of patients with the statistics obtained from healthcare industry is designed by using the Convolutional Neural Network and X^2 – statistical framework. This model enhances and improves the effectiveness of the data during its training and testing. Thus, under-fitting and over-fitting issues are fixed. This approach has improved performance than Artificial Neural Network as the Deep Neural Network with multiple hidden layers is used.
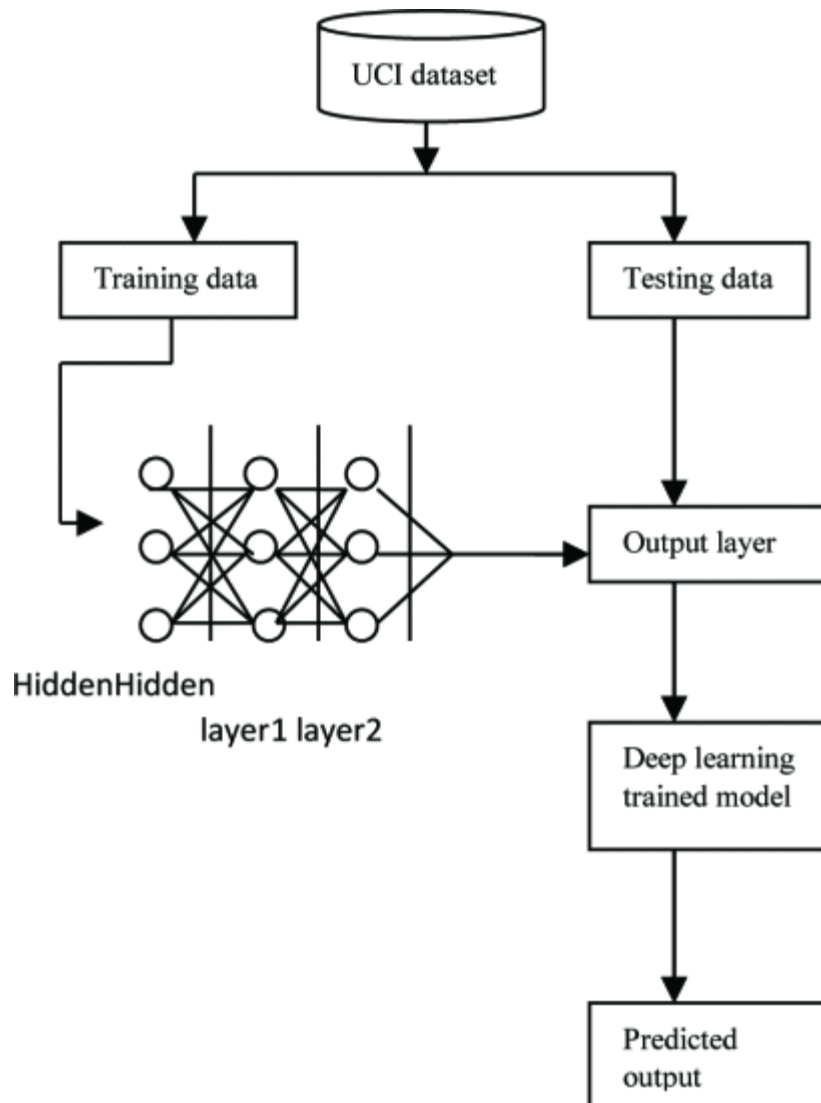
Figure 10 Block diagram of proposed model

2.7 Neural Network-Based Coronary Heart Disease Risk Prediction Using Feature Correlation Analysis

To boost the accuracy of performance in the diagnosis of coronary heart diseases, NN is used. Neural Network is set in a "black-box" technique due to which health researchers are not very pleased with the detection results. Here, in this research paper, the prediction is attempted implying feature correlation analysis (FCA) through two phases. One being the phase of feature selection, in which attributes plays great role in predicting Coronary Heart disease risk and other is the phase of feature correlation analysis. At this point, the existence of associations between

the relationships of properties and the data/information of every other output of the Neural Network generator is calculated. (Jae Kwon Kim, 2017)

Figure 11 Study design using Feature Correlation Analysis(FCA)

Figure 12 A schematic diagram of the sensitivity analysis of the application using NN.

Figure 13 An representation of the NN generator that uses the study of feature correlation.

.

Figure 14 Method of estimation of feature sensitivity.

Figure 15

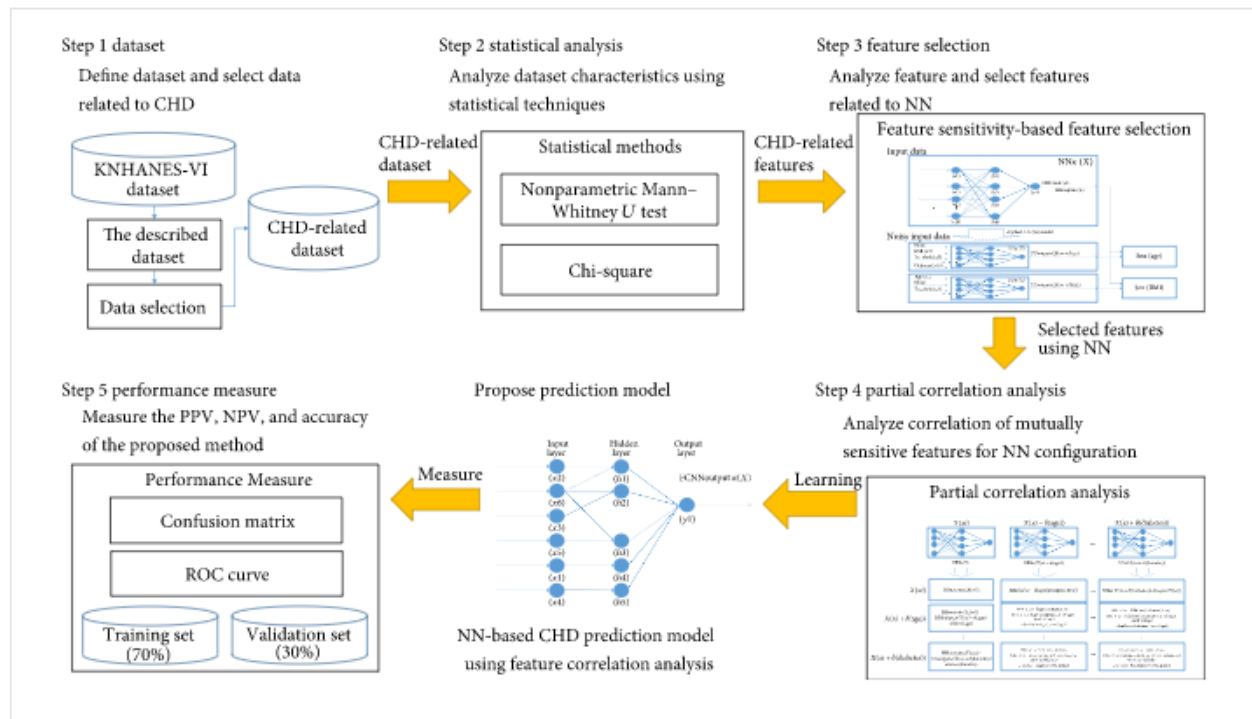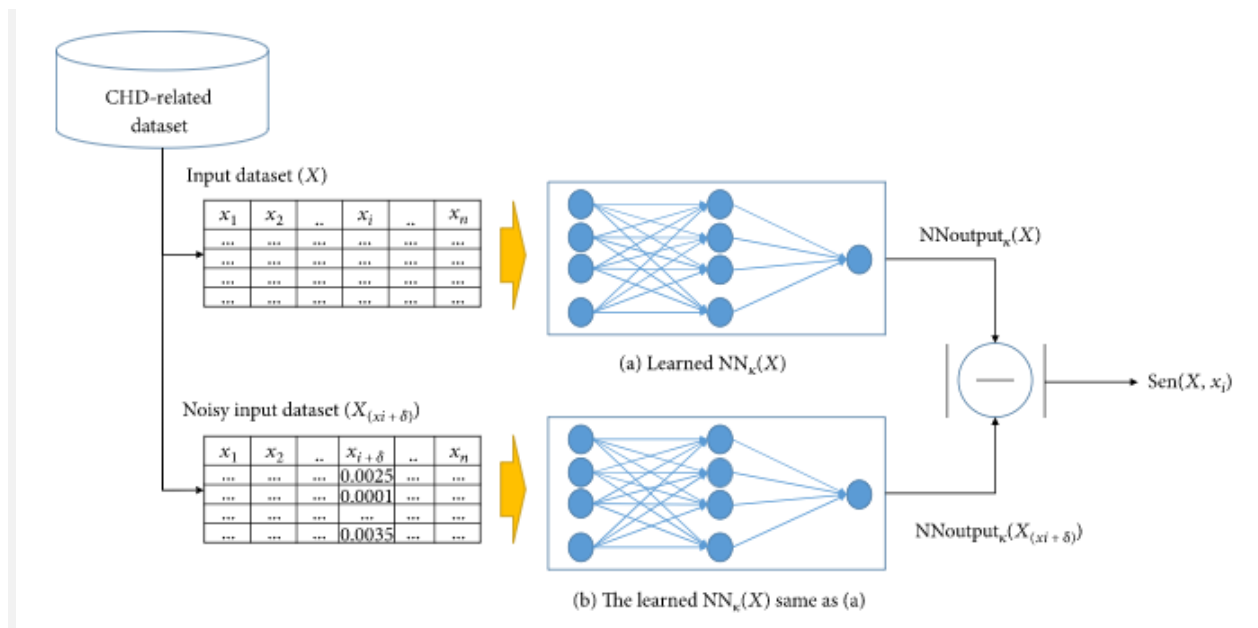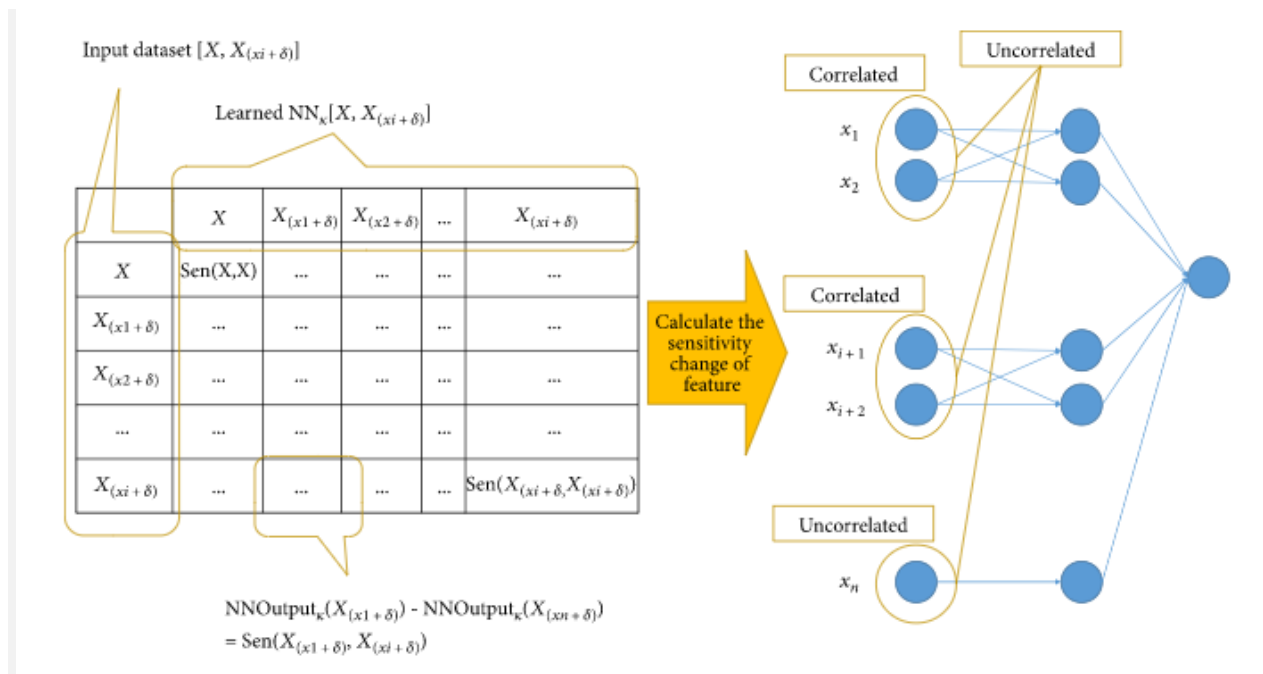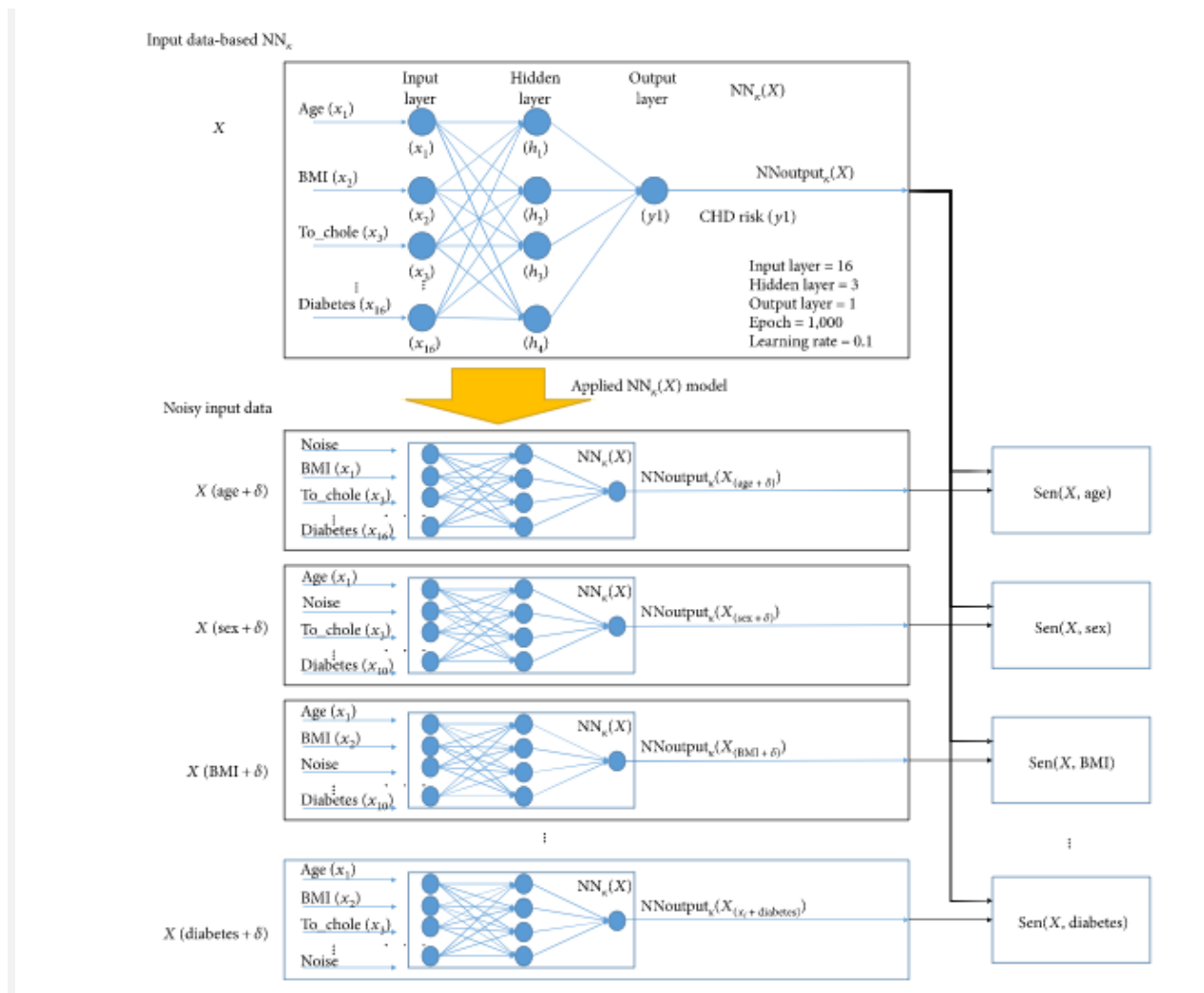| X	NNoutputκ (X)	Sen $(X \,(age + \delta, X)$	Sen $(X \,(diabetes + \delta, X)$	$\cdots$	Sen $(X \,(diabetes + \delta, X)$
$X \,(age + \delta)$	$\|$NNoutputκ $(X)-$ NNoutputκ $(X_{(age+\delta)})\|$ $=$Sen $(X, X_{(age+\delta)})$	$\|$Sen $(X_{(age+\delta, X)}$ $-$ Sen $(X, X_{(age+\delta)})\|$ $=$Sen $(X_{(age+\delta)}, X_{(age+\delta)})$	$\|$Sen $(X \,(diabetes + \delta, X)$ $-$ Sen $(X, X_{(age+\delta)})\|$ $=$Sen $(X_{(diabetes+\delta)}, X_{(age+\delta)})$	$\cdots$	$\|$Sen $(X \,(diabetes + \delta, X)$ $-$ Sen $(X, X_{(age+\delta)})\|$ $=$Sen $(X_{(diabetes+\delta)}, X_{(age+\delta)})$
$X \,(BMI + \delta)$	$\|$NNoutputκ $(X)-$ NNoutputκ $(X_{(BMI+\delta)})\|$ $=$Sen $(X, X_{(BMI+\delta)})$	$\|$Sen $(X_{(age+\delta, X)}$ $-$ Sen $(X, X_{(BMI+\delta)})\|$ $=$Sen $(X_{(age+\delta)}, X_{(BMI+\delta)})$	$\|$Sen $(X_{(diabetes+\delta)}, X)$ $-$ Sen $(X, X_{(BMI+\delta)})\|$ $=$Sen $(X_{(diabetes+\delta)}, X_{(BMI+\delta)})$	$\cdots$	$\|$Sen $(X_{(diabetes+\delta)}, X)$ $-$ Sen $(X, X_{(BMI+\delta)})\|$ $=$Sen $(X_{(diabetes+\delta)}, X_{(BMI+\delta)})$
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$
$X \,(diabetes + \delta)$	$\|$NNoutputκ $(X)-$ NNoutputκ $(X_{(diabetes+\delta)})\|$ $=$Sen $(X, X_{(diabetes+\delta)})$	$\|$Sen $(X_{(age+\delta)}, X))$ $-$ Sen $(X, X_{(diabetes+\delta)})\|$ $=$Sen $(X_{(age+\delta)}, X_{(diabetes+\delta)})$	$\|$Sen $(X_{(diabetes+\delta)}, X)$ $-$ Sen $(X, X_{(diabetes+\delta)})\|$ $=$Sen $(X_{(diabetes+\delta)}, X_{(diabetes+\delta)})$	$\cdots$	$\|$Sen $(X_{(diabetes+\delta)}, X)$ $-$ Sen $(X, X_{(diabetes+\delta)})\|$ $=$Sen $(X_{(diabetes+\delta)}, X_{(diabetes+\delta)})$

Top row labels: X → NNκ (X); $X \,(age + \delta)$ → NNκ $(X \,(age + \delta))$; $X \,(BMI + \delta)$ → NNκ $(X \,(BMI + \delta))$; $X \,(diabetes + \delta)$ → NNκ $(X \,(diabetes + \delta))$

Figure 15 The feature correlation analysis method.

Figure 16

Correlation feature list:
(1) BMI, DBP
(2) DBP, To_chole
(3) DBP, SBP

Coupled

Uncorrelation feature list:
(1) Age
(2) HDL
(3) Triglyceride
(4) Smoking
(5) Diabetes

Uncoupled

Input layer: BMI $(x2)$ (x_2), DBP $(x6)$ (x_6), To_chole $(x3)$ (x_3), SBP $(x5)$ (x_5), Age $(x1)$ (x_1), HDL $(x4)$ (x_4), Triglyceride $(x7)$ (x_7), Smoking $(x8)$ (x_8), Diabetes $(x9)$ (x_9)

Hidden layer: (h_1), (h_2), (h_3), (h_4), (h_5), (h_6), (h_7), (h_8)

Output layer: NN_FCAoutputκ (X) $(y1)$ CHD risk $(y1)$
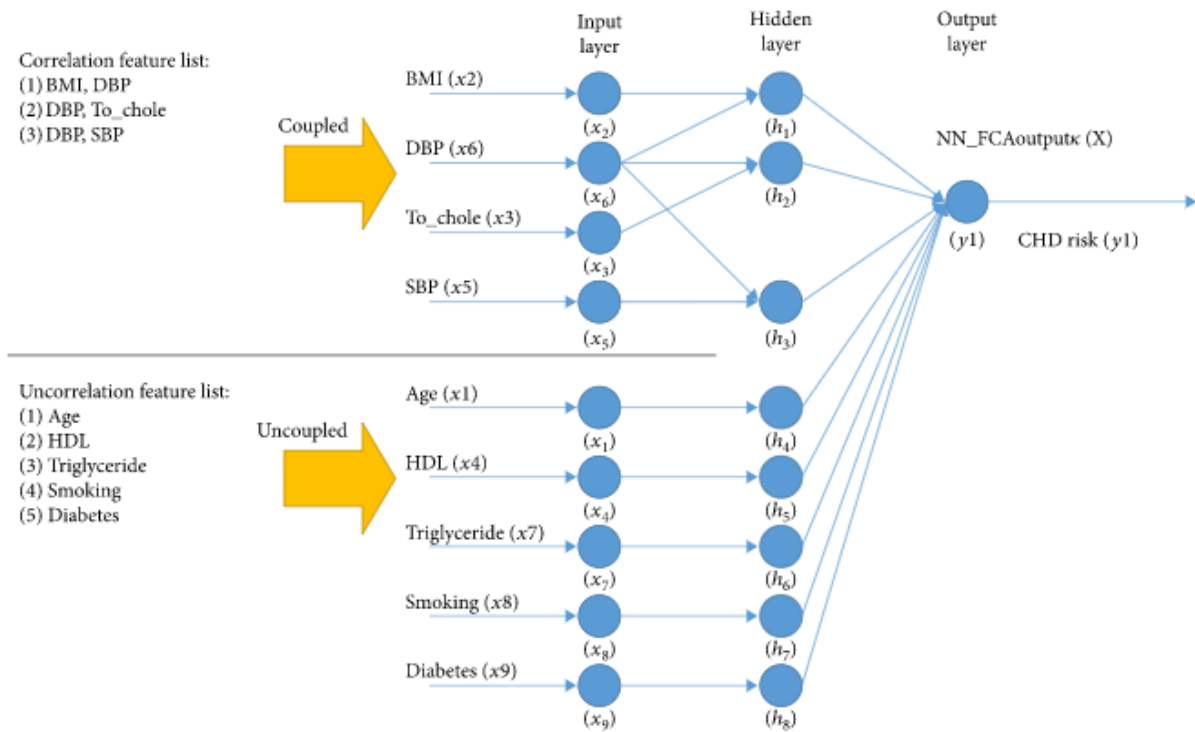
Figure 16 NN-based prediction of CHD using correlation analysis of features.

2.8 PREDICTION OF HEART DISEASE USING ARTIFICIAL NEURAL NETWORK

In this research report, machine learning and pattern matching techniques are used to diagnose heart disease in early stage so that cardiac attacks and other health hazards that might arouse from health disease can be prevented. A single neural network is known as perceptron. Multi-level Perceptron, Feed Forward Neural Network (FFNN), Recurrent Neural Network (RNN), etc are the types of Neural Network. Knowledge Discovery from Data (KDD) which is the well-known model of bioinformatics is the method that has been applied in this research paper. KDD involves the collection of datasets, its preprocessing and developing patterning matching classifier model with Artificial Neural Network.
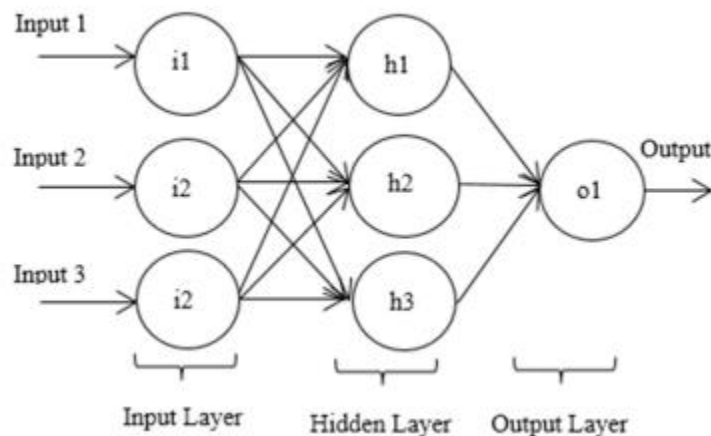

Figure 17 ANN with Input layer where the attribute is given

3. Analysis of findings

The heart is the vital organ that control blood flow across the body through blood vessels sufficient oxygen and nutrients that are necessary for the proper functioning of body parts. Proper functioning of heart is compulsory for the life of any organisms. If heart gets affected or couldn't pump blood, all other vital organs can be adversely affected and may lead to death.

Heart disease refers to the abnormal conditions of heart that disrupt the overall cardiovascular system, blood and blood vessels of the circulatory system. Cardiovascular disease is the condition of heart in which heart may not be able to pump and circulate blood throughout different body parts. Coronary heart disease is another kind of heart disease in which the heart doesn't get sufficient blood it requires due to the storage of cholesterol and fats within the artery wall that supplies deoxygenated blood back to the heart from different parts of body. Myocardial infarctions also known as heart attack is also caused due to the blockage of route in coronary artery due to blood clotting. The other heart diseases are high blood pressure, stroke etc. (Reena, 2010)

Disease diagnosis is amongst the most essential aspect in healthcare profession. If a doctor or medical experts cannot diagnose disease a person cannot be cured. Diagnosis of unknown and new cases and predicting conditions that may occur in future is more complicated even for the medical experts who have been in the field for the many years with practice and experience. Thus, to help them in diagnosis of disease in more precise way in short period of time, less costly and easy, the process can be automated for more efficiency and effectiveness.

In modern environment, huge communities are associated with numerous kinds of heart diseases. The number is increasing every year by great extent. The suffering and dying patients can be of any age group. The risk is increasing in youths and children as well. Thus, early and precise detection of heart diseases is significant to cure patients and people in risk before the situation goes critical.

From the literature review of several research papers and journals I came to conclusion that various techniques, methods and algorithms are proposed and implemented for the prediction of different diseases which are also similar to prediction of heart diseases. Even after being specific in heart disease prediction system using Neural Network various methods and techniques are proposed for its detection. Singh P has used Multilayer Perception Neural Network (MLPNN) with Back Propagation algorithm (BP) for developing heart disease diagnosing system. (Singh P, 2018)

For the detection, analization and extraction of data via complex algorithms, data mining techniques are used for the discovery of unknown patterns in the process of knowledge exploration. Advanced data mining approaches along with computer generated data are used to extract information from huge health care data from medical field. Waghulde & Patil has used Neural Networks and Genetic Algorithms to predict heart disease which is up to 98% effective. (Waghulde & Patil, 2014)

"An Efficient Convolutional Neural Network for Coronary Heart Disease Prediction" is the study that aims an productive NN with convolutional layers to categorize significantly class-

imbalanced data from NHANES with the only aim to predict Coronary Heart Disease. The used double layered CNN shows resilience to the imbalance with trustworthy harmony in class-based performance. To develop and generate much efficient system, neural network is trained with high data and then tested on a comparatively smaller amount of data by Aniruddha Dutta and co. The efficiency of the output is 77% to bring the nature of CHD into sets and 81.8% the omission of CHD cases on data checking, that is 85.70% of the datasets taken. (Aniruddha Dutta1, 2020)

Another method such as "Automated Diagnostic System for Heart Disease Prediction Based on χ2 Statistical Model and Optimally Configured Deep Neural Network" which is about 93% effective and Neural Network related coronary heart disease threat projections using Feature Correlation Analysis are also proposed and implemented.

A Data Mining Approach for Prediction of Heart Disease Using Neural Networks is implemented and proposed by (Miss. Chaitrali S. Dangare1, 2012).To generate effective predictions, the raw data from health care are useful. Information and data stored and managed through computer and advanced data mining techniques are helpful in generating efficient outcomes. For the anticipation of cardiovascular disease, NN is commonly used mechanism. For the prediction of heart disease using NN this research study adopted 13 parameters like sex, blood, pressure, cholesterol. This method can be up to 100% effective.

Various types of datasets are taken that can be formulated for the diagnosis of heart diseases more effectively with great results. The most common sets of data collected for the purpose are age, sex, chest pain type, fasting blood sugar, serum cholesterol, resting blood pressure, electrographic results, maximum heart rate achieved and others. For the more accurate results smoking and obesity are the additional parameters based on which heart diseases can be predicted. Following different measures for the diagnosis of cardiovascular diseases, the accuracy of the prediction can reach from 80% up to perfect detection (i.e. 100%).

4. Conclusion

The heart is the body's vital organ. Proper functioning of heart is compulsory for the life of any organisms. Heart disease refers to the abnormal conditions of heart that disrupt the overall cardiovascular system, blood and blood vessels of the circulatory system. Cardiovascular diseases, coronary heart disease, strokes, high blood pressure etc. are the types of heart diseases. Primitive way of diagnosing heart diseases are doctor's checkup or several examinations such as ECG, Heart MRI and Stress Test etc. which cannot me much effective. various techniques, methods and algorithms are proposed and implemented for the prediction of different diseases which are also similar to the prediction of heart diseases. Neural Network method for the detection of heart diseases is also effective which can be carried out through different techniques.

After researching different journals, research papers and articles, I got to know a lot about heart disease and several ways to detect it using NN. Also, I also got chance to know well about deep neural network, data mining and other machine learning techniques. Theses ways will help to detect and predict the occurrence of heart diseases and help to cure them in time so that future risks and critical conditions can be minimized and save life of millions of people worldwide which is a great achievement for Science itself.

Bibliography

Ali, L., Rahman, A., Khan, A., Zhou, M., Javeed, A., & Khan, J. A. (2019). *An Automated Diagnostic System for Heart Disease Prediction Based on χ2 Statistical Model and Optimally Configured Deep Neural Network.* IEEE.

Aniruddha Dutta1, 2. T. (2020). *An Efficient Convolutional Neural Network for Coronary Heart Disease Prediction.* 1Department of Pathology & Molecular Medicine, Queen's University, Kingston, ON K7L 3N6, Canada. SSRN.

Centers for Diseaes Control and Prevention. (2020). *Heart disease.* Retrieved October 04, 2020, from https://www.cdc.gov/heartdisease/index.htm

Dormehl, L. (2019). *What is an artificial neural network? Here's everything you need to know*. Retrieved 11 2020, from https://www.digitaltrends.com/cool-tech/what-is-an-artificial-neural-network/

Jae Kwon Kim, S. K. (2017). *Neural Network-Based Coronary Heart Disease Risk Prediction Using Feature Correlation Analysis.* Department of Computer Engineering, Inha University, Incheon, Republic of Korea. Incheon, Republic of Korea: Journal of Healthcare Engineering.

Mayo Foundation for Medical Education and Research (MFMER). (2020). *Heart disease.* Retrieved October 03, 2020, from https://www.mayoclinic.org/diseases-conditions/heart-disease/symptoms-causes/syc-20353118

Miss. Chaitrali S. Dangare1, D. M. (2012). *A DATA MINING APPROACH FOR PREDICTION OF HEART DISEASE USING NEURAL NETWORKS .* Walchand Institute of Technology, Computer Science and Engineering Department. Solapur, Maharashtra: ResearchGate.

Reena, A. R. (2010). *Diagnosis of heart disease using datamining algorithm.* Global Journal of Computer Science and Technology, vol. 10, pp. 38-43.

Singh P, S. S.-J. (2018, March 15). *Effective heart disease prediction system using data mining techniques.* J. Institute of Engineering and Technology, Gujarat Technological University, 2Institute of Life Sciences, School of Science and Technology, Ahmedabad University, Ahmedabad, Gujarat, India. Ahmedabad: Dovepress ltd.

Waghulde, N. P., & Patil, N. P. (2014). *Genetic Neural Approach for Heart Disease Prediction.* International Journal of Advanced Computer Research. Bhopal: Proquest.

World Health Rankings. (2018). *NEPAL: CORONARY HEART DISEASE*. Retrieved October 04, 2020, from https://www.worldlifeexpectancy.com/nepal-coronary-heart-disease

World Health Rankings. (2018). *NEPAL: RHEUMATIC HEART DISEASE*. Retrieved October 04, 2020, from https://www.worldlifeexpectancy.com/nepal-rheumatic-heart-disease