

A STUDY ON CARDIOVASCULAR DISEASE PREDICTION USING DATA MINING TECHNIQUES

M.Malathi¹, Dr.T. Jayalakshmi²

Ph.D. Research Scholar¹,
PG and Research Department of Computer Science
L.R.G.Government Arts College for Women
Tirupur, Tamil Nadu, India.

Assistant Professor²,
PG and Research Department of Computer Science
Government Arts College,
Coimbatore, Tamil Nadu, India.

Abstract - In the digital era, there are abundant data created and are stored in repositories. Health sector is no exception from it. But these data are not analyzed with proper techniques in order to extract the hidden information from it. In recent days, Data mining has wide wings in its application. Medical field is one of the application areas of data mining techniques. Although there are number of diseases that affect human, not all are life threatening in nature. Heart disease being one of the serious diseases has to be predicted and treated at the earliest to avoid death rate. Every year millions of people are affected with this disease. More than one third of the world's population affected by this disease. If not treated on proper time it will lead to deaths in people under 70 years of age. Early diagnosis of cardiovascular disease are challenging task, and computer aided methods has been proposed to overcome the disease. Revealing facts about the disease supports the physician in decision making. It can be achieved by KDD (Knowledge Discovery in Databases) process using data mining techniques. The objective of this study is to carry out empirical analysis on various applications of data mining and machine learning techniques. Some of the techniques used worldwide by the researchers for the diagnosis of heart disease are support vector machine, neural network, naive Bayes, conventional neural network, k-Nearest Neighbor, Decision tree and so on.

Keywords: Heart Disease, Data Mining, KDD, Prediction.

1. INTRODUCTION

In a human system heart is the primary organ which plays a vital role. The functions of heart are pumping oxygenated blood to the other body parts, pumping hormones and other vital substances to different parts of the body, receiving deoxygenated blood and carrying metabolic waste products from the body and pumping it to the lungs for oxygenation and to maintain blood pressure [10]. The heart and its blood vessels are known as the cardiovascular system [10]. Smooth function cardiovascular system is very important for healthy life of person. In recent times it is very common to see that heart disease being the main cause for death. The death rates of recent days are alarming to find the causes for heart disease and to obtain the prevention measures at the earliest. At present, there are so many diseases like diabetes, dengue, thyroid, breast cancer, lung cancer, cardiovascular disease and so on., which affects the human worldwide. According to the World Health Organization (WHO) report more than 17.9 million

people die due to cardio vascular disease every single year around the world [11]. Especially in India, more than 3 million Indians die from cardio vascular disease every year and are expected to surpass 4 million by 2030 [11]. The medical practitioners and the researchers has found that the attributes like blood pressure, cholesterol, and pulse rate are the prime factors that influence heart disease. On the other hand, it is noticed that the elements that are irreversible like intake of nicotine in any form, excessive alcohol consumption induces heart diseases. Another factor which is found common among people is family history. In recent times medical data has gained attention of research throughout the world, which has led to number of developments in healthcare sector. The current paper, discusses the contributions of researchers worldwide in analyzing the cardio vascular disease using data mining techniques.

1.1 The Heart and its Functions

Heart is a muscular organ which operates like a pump in the human system. Heart is located slightly left side of the center chest. The organ is divided into two sides as right and left. It does the task of collecting and releasing blood to all parts of the body. The heart has four chambers: two atria and two ventricles. The right atrium receives oxygen-poor blood from the body and pumps it to the right ventricle. The right ventricle pumps the oxygen-poor blood to the lungs. The left atrium receives oxygen-rich blood from the lungs and pumps it to the left ventricle. The left ventricle is one of the four chambers of the heart. It pumps oxygen-rich blood out to the rest of the body. Blood leaves the left ventricle through the aortic valve and enters the aorta, the largest artery in the body. The following figure shows the image of heart along with the chambers and other valves in it[10].

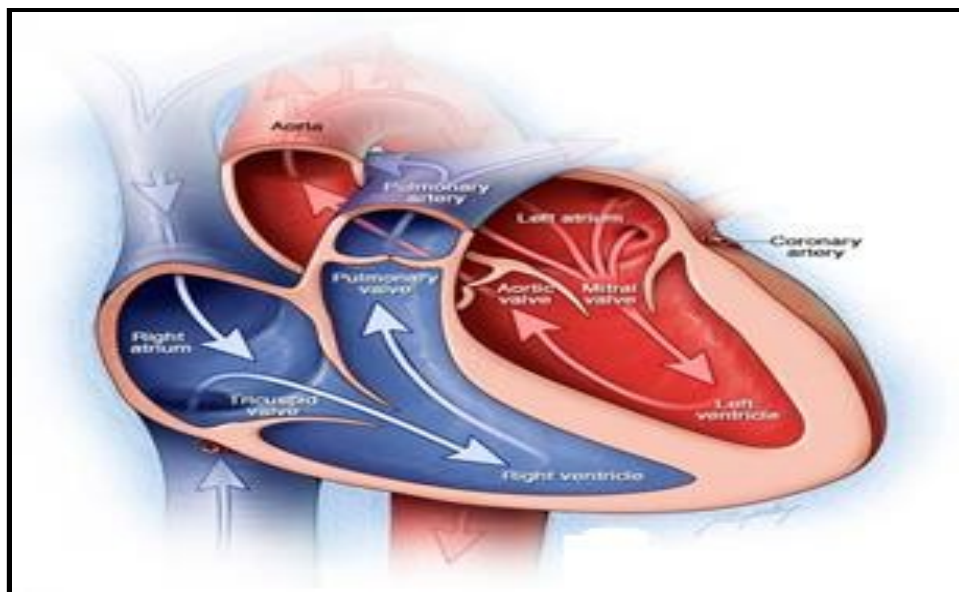


Figure1. Chambers and Valves of Heart

1.2 Causes for Heart Disease

Blood vessels are affected due the various factors like cholesterol, age, poor diet, and tobacco intake in any form, excessive alcohol consumption and family history too. Higher blood pressure results in heart diseases. Blood pressure readings are made up of two values namely Systolic and Diastolic blood pressure. Systolic blood pressure is the pressure when the heart beats that is while the heart muscle is contracting (squeezing) and pumping oxygen-rich blood into the blood vessels. Diastolic blood pressure is the pressure on the blood vessels when the heart muscle relaxes. Increase in lipids in blood also becomes cause for heart disease. Lipids are the fats like substances in blood. Two major lipids in blood are cholesterol and triglycerides. Among these two lipids, cholesterol is the primary cause for heart disease. Sediments of high cholesterol blood transported through arteries narrows the arteries which reduce the blood flow and increase the risk of heart disease. Other non-reversible factors that cause heart disease are age and family history. Among these factors smoking causes heart disease for 40% of people and sometimes leads to death [40].

1.3 Symptoms of Heart Disease

The Symptoms of coronary artery disease may vary for both the genders. For example, common symptoms for men are chest pain, whereas, women may have symptoms like chest discomfort, difficulty in breathing, nausea and extreme fatigue. The common symptoms for cardiovascular disease include, Chest tightness or pressure, Difficulty in breathing, Dizziness or fainting, Fatigue, Fluid buildup, Heart palpitations (heart pounding or racing), Pain or numbness in legs or arms and abdominal pain, nausea and vomiting [40].

1.4 Types of Heart diseases

Heart diseases are due to various reasons and it is of different types. Any discomfort which affects the functioning of heart is considered to be heart disease and also the components that affect the functioning of heart are considered to be heart disease. All the heart disease can be called in common term as cardiovascular disease. Some of the types of heart disease are Coronary heart disease, Arrhythmias, Congestive heart failure, Congenital heart disease, Cardiomyopathy, Angina pectoris, Myocarditis. Coronary heart disease (CHD) is usually caused by a build-up of fatty deposits (atheroma) on the walls of the arteries around the heart (coronary arteries). The build-up of atheroma makes the arteries narrower, restricting the flow of blood to the heart muscle. This process is called atherosclerosis. An arrhythmia is a problem with the rate or rhythm of your heartbeat. It means that your heart beats too quickly, too slowly, or with an irregular pattern. When the heart beats faster than normal, it is called tachycardia. When the heart beats too slowly, it is called bradycardia the most common causes of dilated cardiomyopathy are Heart disease caused by a narrowing or blockage in the coronary arteries. Poorly controlled high blood pressure. Angina pectoris is the medical term for chest pain or discomfort due to coronary heart disease. It occurs when the heart muscle doesn't get as much blood as it needs. This usually happens because one or more of the heart's arteries is narrowed or blocked, also called ischemia [40].

1.5 Risk factors for Heart Disease

Risk factors for developing heart disease are discussed in the following chart:

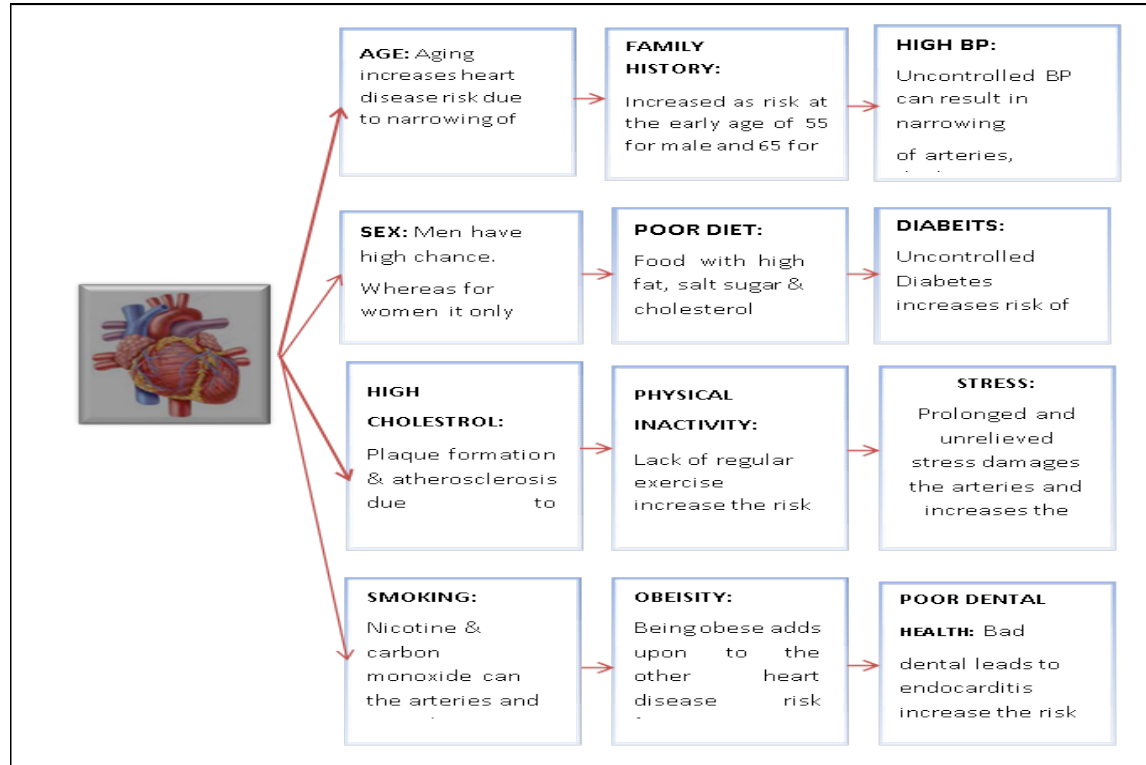


Figure 2. Risk Factors of Heart Disease

1.6 Diagnosis Methods

Though heart disease being a life-threatening disease it can be diagnosed and cured when treated in right time and with diagnosis methods. Physician examines the person with heart problems and suggests a few diagnosis options as listed below based on the health condition of the person.

- Electrocardiogram (ECG or EKG)
- Holter monitoring
- Echocardiogram
- Stress tes
- Cardiac catheterization
- Cardiac computerized tomography (CT) scan
- Cardiac magnetic resonance imaging (MRI)
- X-ray
- Blood tests

In the last few years, medical industry has been used the technological advancements to enhance the quality of health care. This paper explores various data mining and machine learning techniques to analyse and predicts the cardio vascular disease. By the use of data mining and machine learning techniques, it also gives a second opinion to doctors for faster diagnosing the disease.

2. DATA MINING AND KDD PROCESS

Data mining is the process of extracting hidden information from large dataset. The dataset can be of any form such as structured or unstructured. The extracted hidden information from the datasets can be used for analyzing various facts about the data. It also supports in decision making and forecasting the future changes. Data mining being a powerful tool has wide scope of implementation, in recent days it is commonly in fields like agriculture, market analysis, medical field, education sector and many more. The methodology used in data mining is Knowledge Discovery Database (KDD). The KDD process includes different stages like Data Cleaning, Data Integration, Data Selection, Data Transformation, Data Mining, Pattern Evaluation and Knowledge Presentation. Data mining is applied in building a model for heart disease prediction using UCI and Stalog dataset. This helps in forecasting the possibilities of getting affected by heart disease and to take precautionary measures at the earliest.

3. LITERATURE SURVEY

Heart disease is a term that includes all the abnormalities related to heart. It is seen that around 32% of deaths worldwide are due to Cardiovascular Disease (CVD) according to the reports of WHO in year 2019 [11]. Hence, heart disease being a major cause for deaths it is important to analyse the medical dataset and take necessary steps to prevent deaths. This paper focuses in making a survey on various data mining techniques that are used by different researchers worldwide for diagnosis of heart disease in people.

Amin UI Haq et al [1] the author has been used sequential backward selection (SBS) algorithm were used to select more appropriate features to increase the classification accuracy and reduced the computational time of predictive system. The proposed KNN has shown better accuracy (90%) rate. Amin UI Haq et al [2] In this paper, they have been used three feature selection algorithms Relief, MRMR and Lasso used to select the important features and compared on seven well known machine learning algorithms such as Logistic Regression, KNN, ANN, SVM, Naïve Bayes, Decision Tree and Random forest. From this result, it shows that the performance of classifiers with Relief FS algorithm is excellent as compared to mRmR and Lasso. Amin M. S et al [3] prediction model has proposed by the author using data mining algorithms like KNN, Decision Tree, Naive Bayes, Logistic Regression (LR), Support Vector Machine (SVM), Neural Network and Vote (a hybrid technique with Naïve Bayes and Logistic Regression). The paper also discloses the experimental result of various algorithms that are used in the research. The model that is proposed with the selected features has shown good performance with the accuracy rate of 87.4%.

Anitha S., & Sridevi, N. [4] has proposed a supervised machine learning algorithm for prediction of heart disease using data mining techniques. Implementation of data mining techniques in medical data analysis is very common in recent days. The author has used SVM, KNN and Naive Bayes algorithms implemented in R-programming language. Accuracy of the proposed models is commendable.

Babu. S. et al [5] has analyzed medical data using data mining techniques has the potential to reveal many hidden information. These collected data can be used to perform

clinical study. The author has used 14 attributes like age, sex, blood pressure and blood sugar etc., for the study. The attributes were given as input to the algorithms like K-means algorithms, MAFIA algorithm and Decision tree classification. The purpose of this study was to detect the disease at the earliest and provide better treatment at the affordable cost. The proposed model has produced 91% of accuracy.

Dinesh Kumar G et al [7] the author has been used five different datamining techniques such as SVM, Gradient Boosting, Random forest, Naïve Bayes classifier and logistic regression. The proposed results has better accuracy rate. Garima Choudhary et al [8] the author has been used three datamining techniques such as naïve Bayes, random tree and REP Tree algorithms. the proposed model has better accuracy rates. Gomathi.K et al [9] the author has used three different algorithms used in this paper such as Naïve Bayes, ANN and J48. The proposed Naïve Bayes model has produced 79.9%% of accuracy.

Komal Kumar.N et al [13] the author has been used five different machine learning classifiers such as Random forest, Decision tree, Logistic Regression, Support vector machine. The proposed model has highest accuracy than others. Kumar.S et al [14] the author has proposed feature selection techniques have been used with PCA, Relief-Sequential Forward Floating search (SFFS), SBFS and Genetic Algorithm for reducing the dimension. The proposed hybrid classification of Genetic Algorithm-RF model has produced 93.20% of accuracy. Maji, S., &

Arora.S[15] author has collected heart disease dataset from UCI repository and made a clear study about the various data mining techniques in heart disease perdition. The collected dataset is analyzed using WEKA, an effective data mining tool for data analysis. The dataset is validated using ten-fold test. The performance of the proposed model is analyzed using the performance analysis parameters like accuracy, sensitivity, and specificity. The proposed hybrid model gives better results when compared with other classifiers. The proposed model has produced 80.4% of accuracy.

Makumba, D. O et al [16] has proposed to support system for prediction in decision making about the presence of heart disease in a patient. The model is developed using data mining techniques like decision tree, naïve Bayes, KNN (K-Nearest Neighbors) using WEKA tool. The model takes 13 attributes like blood pressure, sex, age, cholesterol, blood sugar and so on. The dataset of 740 records were collected open access repository. Among these datasets around 296 records were used for testing and remaining are used for training. Performance of the proposed model is compared with the existing models for performance analysis purpose. The proposed method has shown better accuracy rates. Masih, N., & Ahuja, S.[17] author has taken up a complete analysis on the FHS (Framingham Heart Study) using various data mining techniques such as Decision tree, Naïve Bayes, Support vector machine (SVM) and Artificial neural network (ANN) Among these algorithms ANN has showed better results than any other classification algorithms. The results gathered from the models helps in finding the prominent feature that effects the health of a person.

Mirza, I et al [18] author has used machine learning algorithms in analyzing the medical dataset of patients which helps I finding the pattern between the data that are collected. It helps in accurate prediction of heart disease presence. The author has considered including 13 medical attributes of a person so as to identify the presence of heart disease. These attributes are given as input to RBF SVM and Linear SVM along

with KNN and Naive Bayes classifiers to classify the patient with disease and patient good health condition. MoloudAbdar et al [19] The proposed system applied preprocessing for both categorical and numerical attributes with a normalization techniques and feature selection, redundant features were eliminated using GA and PSO algorithm. The proposed model has shown better accuracy 93.08% rates. Mukesh Kumar et al [20] the author has used four different classification algorithm such as Naïve Bayes, MLP, Random forest and Decision tree. The proposed naïve Bayes has shown better accuracy (87.20%) rate. Muktevi

Srivenkatesh [21] has come up with prediction model in finding the person has cardiovascular disease or not and to provide awareness to people. The author has collected dataset from Kaggle. The collected dataset is given as input to various data mining algorithms like Support Vector Machine, Random forest, Naive Bayes classifier and logistic regression. Accuracy of these algorithms are 58.71% and 77.06% respectively. It is found that Logistic Regression has better Accuracy (77.06 %) when compared to different Machine-learning Algorithms.

Nalluri.S. et al [24] has discussed about the cardiovascular disease called chronic heart disease (CHD). The author has proposed a model using data mining techniques namely XGBoost algorithm and logistic regression algorithm to predict the CHD presence in a person. The proposed algorithm takes in various parameters blood pressure, cholesterol level, heart rate and glucose level, number of cigarettes smoked per day and BMI of a person. The proposed algorithm has shown good performance result with 87.4%. accuracy. Pavithra et al [25] has proposed a methodology using data mining techniques for predicting the heart disease at the earliest using data mining and artificial intelligence techniques. The application was developed by collecting the required data from the patients. Data includes the attributes like age, Chest Pain Type, Blood Sugar level and family history and so on. The collected data is classified to identify the disease and has been produced the accuracy of 90%

Radhimeenakshi, S. [27] the author has used Support Vector Machine (SVM) as well as Artificial Neural Network (ANN) for foreseeing the presence of Heart Disease in a person in order to safe the life of a person. The dataset for the study is collected from Cleveland Heart Database and Stalog Database taken from UCI repository. The dataset is fed into both SVM and ANN for performance analysis in which ANN shows better results over SVM. The proposed model has produced 91% of accuracy. Rahman, M et al [28] author has discussed about different heart disease like coronary aorta syndrome, arrhythmias, congenital heart defects in detail. Medical data when analyzed gives a better view in early prediction of heart disease in a patient. Thus, the author has used an effective data mining technique called decision tree for finding the presence of heart disease in a person. Decision tree provides better accuracy rates in classifying the dataset when compared with other classifiers. The proposed model has produced 80% of accuracy.

Rairikar, A et al [29] author has proposed an effective model for detecting the presence of heart disease by employing data mining techniques. The author has used genetic algorithm with the back-propagation technique. The model is trained using 13 different attributes like Gender, blood pressure, cholesterol and so on. The proposed model detects the chances of a person affected by heart disease in future. The proposed model has produced 89% of accuracy.

Repaka,A.N. et al [30] author has developed a system for prediction of heart disease by considering medical data and information. Author has proposed a model named SHDP (Smart Heart Disease Prediction) a mobile application using Navies Bayesian algorithm. And the data are transferred using AES (Advanced Encryption Standard). The results of the proposed model is discussed and compared with the other existing algorithms. The proposed model has produced 89.77% of accuracy. Ritika Chadha et al [31] the author has been used various feature selection methods like PCA, Chi-square testing, Relief and Symmetrical uncertainty and also compared a variety of classification algorithm. The proposed model has produced 85% of accuracy.

Saiful islam et al [33] the author has been used logistic Regression, decision tree, SVM and Naïve Bayes classification algorithms applied to predict heart disease. the proposed model has better accuracy rates. SenthilKumar Mohan et al [35] the author has been used novel method for applying machine learning algorithm. The proposed model shows that the Hybrid random forest with a linear model (HRLFM) achieved 88.7% accuracy.

Shaji,S.P.[36] author aims in providing, diagnostic system for different heart disease and the precautionary measures at the earliest in affordable cost. The model is developed using data mining techniques by feeding the collected attributes into SVM, Random forest, KNN, and ANN classification algorithms. The results of the proposed model gives better results when compared with other data mining algorithms. The proposed model has produced 80.4% of accuracy.

Sharmila.S et al [37] In this paper they have compared two algorithm such as Naïve Bayes and Decision tree. From this result, it shows the result is to improve the accuracy of Naïve Bayes (83.7%) Algorithm. Uma N Dulhare et al [38] the author has been used particle swarm optimization (PSO) feature selection with Naïve Bayes classifier. The proposed model has shown better accuracy level.

Vanitha.S et al [39] has proposed to two different algorithms used in this paper such as ANN and SVM. The proposed Neural Network model has produced 85.52%% of accuracy. Viren Viraj Shankar et al [40] the author has been used CNN, Naive Bayes, KNN etc. From this result, it shows that the CNN algorithm obtained the highest accuracy 80% than others.

4.METHODOLOGY

This study focuses the various phases to predict the heart disease. To build the heart disease model the machine learning process starts with Preprocessing phase, Normalization phase, Feature selection and Extraction Phase and the classification phase. The overview of the heart disease model is shown in Figure 2.



Data Collection

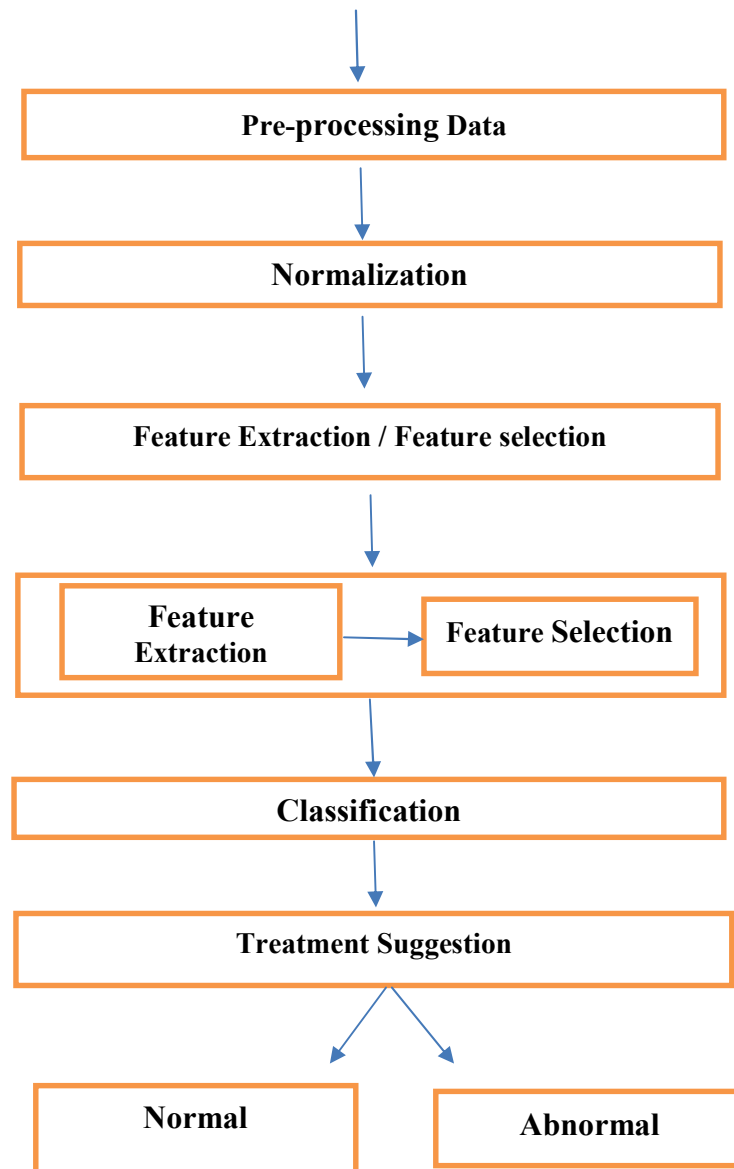


Figure 2. Framework

4.1 Dataset Description

The dataset is taken from publicly available repositories such as UCI repository and Kaggle websites. This database contains four datasets such as (Cleveland data, Hungarian data, Long-beach vs data, Switzerland data, Stalog data). Most of the researcher's commonly used in medical domain are Cleveland and Stalog datasets. Cleveland dataset has 76 attributes and 303 patient records, in which only the 14 attributes are suggested and used. The features of heart disease datasets are shown in Table 1.

Table 1. Features of heart disease datasets

Datasets	No of Instances	No of Attributes
Stalog project	270	14
Cleveland	303	14
Hungarian	294	14
V.A. Long Beach	200	14
Switzerland	123	14

4.2 Data Preprocessing

The preprocessing of data is a significant role in the data mining process. One of the challenges that face of most health care data contains eliminate unwanted noise, missing value, incomplete, inconsistent and redundant data present in the original input data and improve the quality of mining in it. The preprocessing steps includes data cleaning, data integration, data transformation, data reduction and data discretization. The problem of missing data, noisy data poses difficulty in the analysis and decision-making processes. The main objectives of preprocessing are cleaning of noise, filling up of missing values, reduce the redundancy and normalize the data. In preprocessing step, some statistical techniques are used as to find out mean, mode and median value to replace missing values [22]. To remove unwanted noise Hybrid Denoising algorithm has been used in preprocessing step to identify the breast cancer in mammogram images [6].

4.3 Normalization

After preprocessing reduce the redundancy features using normalization process to carried out for selected attributes. Data normalization can be used to scale the data in the some of values for each input features in order to minimize bias with one feature to another. There are so many existing techniques used for this method such as min-max, Z-score and decimal scaling so on [23].

4.4 Feature Extraction and Feature Selection

Feature extraction can be designed as extracting significant data from the given input. Feature selection simply tells the significance of each of the original features. There are many techniques for extracting the features such as ICA, LDA, LLE, t-SNE, Autoencoders and select the most relevant features. Some of the techniques used in this method are PCA, Chi-Square and k means clustering [12,16,32].

4.5 Classification

After extracting the appropriate features classification process is carried out. In recent days most of the existing reviews shows that classification task in datamining has plays a major role in cardiovascular disease prediction and compared with many other techniques of health care sector. This step achieves data to classify as normal, abnormal and chances for abnormality and assign it to a particular class [25]. There are so many existing techniques in data mining used for classification, some are Support Vector Machine and KNN [26,34], Decision Tree, ANN, Naïve Bayes are used.

5.CONCLUSION

Heart disease being the common problem worldwide affecting the health of the human. It is important to predict the disease at the earliest stage to avoid death rate. This paper discusses on the prediction of heart disease using data mining techniques. It helps to locate out the existing algorithms like Neural Networks, Naive Bayes, ANN, SVM, KNN and so on., This paper also discusses about the various datasets that are available in repositories such as Stalog, Cleveland and so on., Study of these algorithms and their implementation in different datasets helps in developing better model for heart disease prediction. This study paper helps to develop in future to improve the accuracy rate and also suggest the treatment opinion with the help of like datamining techniques. Hence, this study helps to know about the machine learning techniques and deep neural network techniques for prediction and classification of cardiovascular disease for further process.

References

- [1]. Amin Ul Haq, Jian Ping Li, Muhammad Hammad Memon, Shah Nazir and Ruinan Sun (2018),” A Hybrid Intelligent System Framework for the Prediction of Heart Disease Using Machine Learning Algorithms”, Hindawi Mobile Information Systems Volume, Article ID 3860146, 21 pages <https://doi.org/10.1155/2018/3860146>.
- [2]. Amin Ul Haq, Jianping Li, Muhammad Hammad Memon, (2019),” Heart Disease Prediction System Using Model of Machine Learning and Sequential Backward Selection Algorithm for Features Selection”, 5th International Conference for Convergence in Technology (I2CT) Pune, India.
- [3]. Amin, M. S., Chiam, Y. K., &Varathan, K. D. (2019). Identification of significant features and data mining techniques in predicting heart disease. *Telematics and Informatics*, 36, 82-93.
- [4] Anitha, S., & Sridevi, N. (2019). Heart disease prediction using data mining techniques. *Journal of analysis and Computation*.
- [5]. Babu, S., Vivek, E. M., Famina, K. P., Fida, K., Aswathi, P., Shanid, M., &Hena, M. (2017) Heart disease diagnosis using data mining technique, International conference of electronics, communication and aerospace technology (ICECA) vol. 1, pp. 750-753. IEEE.
- [6]. Devakumari D and Punithavathi V, (2020), “Noise Removal in Breast Cancer using Hybrid De-Noising Filter for Mammogram Images”, published by International Conference on Computational Vision and Bio-Inspired Computing (ICCVBIC 2019), pp. 109-119.
- [7]. Dinesh Kumar G, Santhosh Kumar D, Arumugaraj K, Mareeswari V, (2018) ” Prediction of Cardiovascular Disease Using Machine Learning Algorithms”, International Conference on Current Trends toward Converging Technologies, Coimbatore, India IEEE.DOI 10.1007/s40012-016-0121-0.
- [8]. Garima Choudhary, Shailendra Narayan Singh, (2019) ” Prediction of Cardiovascular Disease using Data Mining Technique” 2019 4th International Conference on Information Systems and Computer Networks (ISCON) GLA University, Mathura, UP, India. Nov 21-22,.
- [9].Gomathi.K, Dr.Shanmugapriya (IJRASET 2013), “Heart Disease Prediction Using Data Mining Classification”, International Journal for Research in Applied Science & Engineering Technology (IJRASET 2013).IEEE Xplore Part Number:CFP20K25-ART; ISBN:978-1-7281-4889-2.
- [10].https://www.medicinenet.com/what_are_the_four_main_functions_of_the_heart/article.htm
- [11].<https://www.who.int/health-topics/cardiovascular-diseases>.
- [12].Kavitha.R, E.Kannan,” An Efficient Framework for Heart Disease Classification using Feature Extraction and Feature Selection Technique in Data Mining”,
- [13].KomalKumar.N, G.Sarika Sindhu, D.KrishnaPrashanthi, A.ShaeenSulthana,” Analysis and Prediction of Cardio Vascular Disease using Machine Learning Classifiers”, International Conference on Advanced Computing & Communication Systems (ICACCS). 978-1-7281-5197-7/20/\$31.00 ©2020 IEEE.

- [14].Kumar.S, G. Sahoo (November 2017),” A Random Forest Classifier based on Genetic Algorithm for Cardiovascular Diseases Diagnosis”, IJE TRANSACTIONS B: Applications Vol. 30, No. 11, 1723-1729.
- [15].Maji, S., & Arora, S. (2019). Decision tree algorithms for prediction of heart disease. In *Information and communication technology for competitive strategies* (pp. 447-454). Springer, Singapore.
- [16].Makumba, D. O., Cheruiyot, W., &Ogada, K. (2019). A model for coronary heart disease prediction using data mining classification techniques. *Asian Journal of Research in Computer Science*, 1-19.
- [17].Masih, N., & Ahuja, S. (2018). Prediction of heart diseases using data mining techniques: application on Framingham heart study. *International Journal of Big Data and Analytics in Healthcare (IJBDAH)*, 3(2), 1-9.
- [18].Mirza, I., Mahapatra, A., Rego, D., &Mascarenhas, K. (2019, December). Human Heart Disease Prediction Using Data Mining Techniques. In *2019 International Conference on Advances in Computing, Communication and Control (ICAC3)* (pp. 1-5). IEEE.
- [19].MoloudAbdar,WojciechKsizek , U Rajendra Acharya, Ru-San Tan , Vladimir Makarenkov , PawełPławiak, ”A new machine learning technique for an accurate diagnosis of coronary artery disease”, *Computer Methods and Programs in Biomedicine* 179 (2019) 104992. <https://doi.org/10.1016/j.cmpb.2019.104992>
- [20]. Mukesh Kumar, Shankar Shambhu, Abha Sharma³ (JULY - SEPTEMBER 2018),” Classification of Heart Diseases Patients using Data Mining Techniques”, *IJRECE VOL. 6 ISSUE ISSN: 2393-9028 (PRINT) | ISSN: 2348-2281 (ONLINE)*.
- [21].MukteviSrivenkatesh,(2020).Prediction of Cardiovascular Disease using Machine Learning Algorithms. *International Journal of Engineering and Advanced Technology (IJEAT)* ISSN: 2249 – 8958, Volume-9 Issue-3.
- [22].NabaouiaLouridi, Meryem Amar, Bouabid El Ouahidi (2019),” Identification of Cardiovascular Diseases Using Machine Learning”, 78-1-7281-4420-7 / 19 / 31,00 \$ © IEEE.
- [23].Nagrajan.G,A.PMohanRaju,V.Logeshwaran,K.Nandhakumar,S.NaveenKumar,”Effective Prediction Model for Heart Disease Using Machine Learning Algorithm”*International Journal of Engineering Research & Technology(IJERT)*,RTICCT-2019 Conference Proceeding,Volume 7,Issue 01.
- [24].Nalluri, S., Saraswathi, R. V., Ramasubbareddy, S., Govinda, K., & Swetha, E. (2020). Chronic heart disease prediction using data mining techniques. In *Data engineering and communication technology* (pp. 903-912). Springer, Singapore.
- [25].Pavithra.M., Sindhana, A. M., Subajanaki, T., & Mahalakshmi, S. (2021). Effective Heart Disease Prediction Systems Using Data Mining Techniques. *Annals of the Romanian Society for Cell Biology*, 6566-6571.
- [26].Punithavathi V and Devakumari D, (2020) “ A Hybrid Algorithm with Modified SVM and KNN for Classification of Mammogram Images using Medical Image Processing with Data Mining Techniques” published by *European Journal of Molecular & Clinical Medicine*, ISSN 2515-8260, Vol.7, Issue.10.
- [27].Radhimeenakshi, S. (2016, March). Classification and prediction of heart disease risk using data mining techniques of Support Vector Machine and Artificial Neural Network. In *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)* (pp. 3107-3111). IEEE.
- [28].Rahman, M., Zahin, M. M., & Islam, L. (2019, November). Effective Prediction on Heart Disease: Anticipating Heart Disease Using Data Mining Techniques. In *2019.IEEE*.
- [29].Rairikar, A., Kulkarni, V., Sabale, V., Kale, H., &Lamgunde, A. (2017, June). Heart disease prediction using data mining techniques. In *2017 International conference on intelligent computing and control (I2C2)* (pp. 1-8). IEEE.
- [30].Repaka, A. N., Ravikanti, S. D., & Franklin, R. G. (2019, April). Design and implementing heart disease prediction using naives Bayesian. In *2019 3rd International conference on trends in electronics and informatics (ICOEI)* (pp. 292-297). IEEE.
- [31].Ritika Chadha, Shubhankar Mayank,” Prediction of heart disease using data mining techniques”, *CSIT*
- [32].Robinson Spencer, FadiThabtah , Neda Abdelhamid and Michael Thompson,” Exploring

- feature selection and classification methods for predicting heart disease”, Digital Health Volume 6: 1–10, DOI: 10.1177/2055207620914777 journals.sagepub.com/home/dhj.
- [33].Saiful Islam, Nusrat Jahan, Mst. Eshita Khatun,” Cardiovascular Disease Forecast using Machine Learning Paradigms”, Proceedings of the Fourth International Conference on Computing Methodologies and Communication (ICCMC 2020)
- [34].Satish Chandra Reddy,N, Song Shue Nee, Lim Zhi Min & Chew Xin Ying,” Classification and Feature Selection Approaches by Machine Learning Techniques: Heart Disease Prediction”, International Journal of Innovative Computing 9(1) IJIC Vol. 9:1(2019) 39-46 39-46.
- [35].Senthil Kumar Mohan, Chandrasekar Thirumalai, And Gautam Srivastava,” Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques”, Special Section on Smart Caching, Communications, Computing and Cybersecurity for Information-Centric Internet of Things, VOLUME 7, 2019, IEEE Access.
- [36].Shaji, S. P. (2019, April). Prediction and Diagnosis of Heart Disease Patients using Data Mining Technique. In 2019 international conference on communication and signal processing (ICCSP) (pp. 0848-0852). IEEE.
- [37].Sharmila.S, M.P. Indra Gandhi (Mar - Apr 2017),” Improved Heart Disease Prediction Used Data Mining Techniques”, International Journal of Information Technology (IJIT) – Volume 3 Issue 2,.
- [38].Uma N Dulhare,” Prediction system for heart disease using Naive Bayes and particle swarm optimization”, Biomedical Research 2018; 29 (12): 2646-2649, ISSN 0970-938X.
- [39].Vanitha.S, Dr.P. Balamurugan (February 2017),” Medical Data Classification Using Svm and Neural Network Classifier-A Study”, International Journal of Innovative Research in Science and Engineering Vol.No.3, Issue 02, February 2017.
- [40].VirenVirajShankar,VarunKumar,UmeshDevagade,VinayKaranth, K. Rohitaksha,” Heart Disease Prediction Using CNN Algorithm”, SN Computer Science (2020) 1:170 <https://doi.org/10.1007/s42979-020-0097-6>.
- [41] www.mayoclinic.org/diseases-conditions/heart-disease/symptoms/causes/syc-20353118