DOI: 10.21917/ijsc.2018.0234

COMPUTATIONAL APPROACHES FOR HEART DISEASE PREDICTION - A REVIEW

R. Manavalan and S. Saranya
Department of Computer Science, Arignar Anna Government Arts College, India

Abstract

The data mining techniques can be primarily used to extract the potentially useful hidden knowledge from the large volume of health care industry databases for predicting the diseases. It is also assist to locate the relationships and patterns among the extracted data. In this paper, detailed survey of various computational techniques used in health care industry for predicting the heart diseases is presented. The issues facing by the computational models are identified while predicting the heart diseases and the same is also presented with future possible research directions.

Keywords:

Data Mining, Heart Disease, Classification, Prediction, Parameters

1. INTRODUCTION

Data mining is used for various applications such as business organizations, e-commerce, and health care industry, scientific and engineering. A literature review is a description of the study relevant to particular field or topic. It is discusses about the information in a particular subject area for the past years. The data mining is process of reading, analyzing, evaluating and summarizing materials about specific topic to accomplish particular task. The review describes summaries, evaluated and clarified various applications of heart disease prediction using computational methods. Data Mining is an essential process in medical data analysis where intelligent methods are required in order to extract valuable data patterns from huge data set. Heart Diseases remain the biggest cause of deaths for the past two decades.

Recently, computer technology and Machine Learning techniques are used to develop a system to assist doctors in decision making for heart disease prediction at the earliest stage. The clinical and pathological data could be used for the diagnosis of heart disease. Heart disease prediction system can assist medical professionals in predicting heart disease status based on the clinical data of the patients. In biomedical field data mining techniques play an essential role in the prediction of diseases. The health care industries are facing many challenges to address the issues based on patient's severity with cost effective treatment. The World Health Organization (WHO) has estimated that 12 million deaths occur worldwide every year due to the Heart diseases. In 2008, 17.3 million people died in worldwide due to Heart Disease [1] accounting for 80% of total deaths.

WHO report estimated that by 2030, almost 23.6 million people will be getting the chance to die regarding Heart disease. In India, about 25% of deaths occur only in the age group of 25 to 69 years alarm to heart disease whereas, in urban areas, 32.8% deaths occur because of heart ailments, while this percentage in rural areas is about 22.9%.

Data mining is the most efficient technique for predicting the heart diseases. The prediction of heart disease using data mining techniques is not easy task since the complexity and toughness of information is too high in medical domain data. Risk factors and symptoms for heart diseases are clearly explained in subsection 1.1 and 1.2 respectively.

1.1 RISK FACTORS FOR HEART DISEASE

Risk factors are conditions or habits that make a person more likely to have a heart disease. Further, they can also have the chances to increase the existing disease to get worse.

Several risk factors for heart disease include i) family history of heart disease, ii). smoking, iii) cholesterol, iv) high blood pressure, v) obesity, vi) diabetes, vii) stress and viii) lack of physical exercise. Heart disease is a broad term that includes all types of diseases affecting different components of the heart. Heart means 'cardio'. Therefore, all heart diseases belong to the category of cardiovascular diseases.

Some types of heart diseases are i) coronary heart disease, ii) angina pectoris, iii) congestive heart failure, iv) cardiomyopathy, v) congenital heart disease, vi) arrhythmias and vii) myocarditis.

1.2 SYMPTOMS OF A HEART ATTACK

The signs and symptoms of Coronary Heart Disease (CHD) may differ from women to men. It was identified that there was no common signs or symptoms for the silent CHD. Silent CHD is very difficult to diagnose until a woman has signs and symptoms of a heart attack, heart failure, or an arrhythmia (irregular heartbeat). Symptoms of a heart attack can include:

- Discomfort, pressure, heaviness, or pain in the chest, arm, or below the breastbone.
- Discomfort radiating to the back, jaw, throat, or arm.
- Fullness, indigestion, or choking feeling (may feel like heartburn).
- Sweating, nausea, vomiting, or dizziness.
- Extreme weakness, anxiety, or shortness of breath.
- Rapid or irregular heartbeats.

This paper is organized as follows: section 2 describes the extensive review about the computational approaches introduced for Heart Disease prediction, the challenges and issues faced by the same while prediction of the Heart Disease. Finally, the conclusion about this study is given in section 3 with further promising research directions.

2. COMPUTATIONAL MODEL BASED HEART DISEASE PREDICTION - A REVIEW

The ability to extract useful hidden knowledge in the large amount of data and to act on the same is a challenging task, since growth of the data size is enormously increasing day by day. The need to understand large, complex, enriched information in data sets has now increased in all the varied fields of technology, business, medical and science. The process of developing a novel Computer Based Information System (CBIS) for discovering potentially useful knowledge from data is necessary to increase the life time of the patient. The various data mining techniques introduced so for to heart disease predictions are reviewed extensively here under:

In 2009, Resul et al. [2] introduced Neural Network which was implemented using SAS base software. Accuracy of the algorithm is 97.4% for the Diagnosis of Valvular Heart Disease.

In 2010, Anbarasi et al. [3] proposed enhanced prediction model for Heart Disease using Feature Subset Selection using Genetic Algorithm with Decision Tree, Genetic with Naive Bayes and Genetic with classification via clustering. The enhanced prediction of heart disease is carried out by using a genetic algorithm based feature subset selection method using 10 attributes. The prediction accuracy is increased while incorporating classification via clustering with feature subset selection algorithms. Classification techniques such as Naïve Bayes, Decision Tree and Classification by clustering are used for prediction.

Kumari et al. [5] applied data mining techniques such as SVM, ANN, Decision Tree, RIPPER classifier to predict the risk of heart diseases, in 2011. The performance of these algorithms is analyzed using statistical analysis factors such as Accuracy and Error Rate. Accuracy of RIPPER, Decision Tree, ANN and SVM are 81.08%, 79.05%, 80.06% and 84.12% respectively whereas error rates of RIPPER, Decision Tree, ANN and SVM are 2.756, 0.2755, 0.2248 and 0.1588 respectively. The analysis clearly showed that out of these four classification models, SVM predicted cardiovascular disease with least error rate and highest accuracy.

Sundar, et al. [6] discussed the performance analysis of classification techniques Naive Bayes and Weighted Association Classifier (WAC) for heart disease prediction based on evaluation using classification matrix, which revels the frequency of correct and incorrect prediction. Accuracy of WAC and Naive Bayes are 84% and 78% respectively. The accuracy of WAC is 6% higher than Naive Bayes.

In 2012, Dangare et al. [1] used 13 input attributes medical terms such as sex, blood pressure, and cholesterol etc. To get more appropriate results, additionally added two more important attributes with above attributes i.e. obesity and smoking for heart disease. Multilayer Perceptron Neural Network (MLPNN), Back propagation algorithm is introduced for heart disease. The method yielded the accuracy of 99.25% for 13 attributes and nearly 100% for 15 attributes.

In 2012, Peter et al. [7] introduced the algorithm such as Naive Bayes, Multilayer, J48 and KNN and conducted experimentation on dataset of health care domain. The Correlation based Feature Selection and filter subset evaluation methods are adapted to reduce more number of irrelevant and redundant attributes thereby increases the performance of classifiers. The accuracy of these algorithms are 85.18%, 78.88%, 88.18% and 85.55% respectively and provide the evidence of significant improvement in the results.

In 2012, Shouman et al. [8] proposed the model called as K-Nearest Neighbour for the prediction of heart disease. In this model, the value of K is ranged between one and thirteen. KNN yielded the accuracy between 94% and 97.4% for different values of K. When K = 7, the method achieved the maximum accuracy and specificity are 97.4% and 99% respectively.

In 2012, Nidhi et al. [9] introduced the algorithms Naïve Bayes, Decision Tree and Neural Network and the same was analyzed. The analysis results showed that the neural network produced at maximum accuracy for 15 attributes i.e. nearly 100%. On the other hand, Decision Tree has also performed well for the same number of attributes and yielded 99.62% of accuracy. Moreover, Genetic Algorithm and Feature Subset Selection methods are used to reduce the number of attributes to six and give as input to classification models. One of the classification model Decision Tree produced almost same accuracy.

In 2012, Pethalakshmi et al. [10] introduced the algorithm such as Fuzzy Decision tree, Fuzzy Naive Bayes, Fuzzy Neural Network, Fuzzy k-means to predict the heart disease. The accuracy of these algorithms is 90.06%, 89.62%, 91.09% and 99.49% respectively. Fuzzy k-means yielded the best result which was at least 8.4% more than others.

In 2013, Abishek [11] proposed J48 and Naive Bayes to predict the Heart disease and adopted Weka tool to implement the above techniques. Accuracy of these algorithms is 95.56% and 92.42% respectively. The accuracy of j48 was 3.14% higher than Naive Bayes algorithm.

In 2013, Chitra et al. [12] introduced the algorithm such as Artificial Neural Network (ANN), K-Means Clustering and Fuzzy C Means Clustering to predict the heart disease. The accuracy of these algorithms is 85%, 88% and 92% respectively. Fuzzy C Means Clustering yielded the best result which is at least 4% more than others.

In 2013, Rashedur et al. [34] introduced an algorithm called Neural Network and fuzzy logic to predict the risk of heart diseases. Weka and Tanagara is adopted to run these techniques. Accuracy of Neural Network and Fuzzy Logic are 79.19% and 83.85% respectively.

In 2013, Patel et al. [14] introduced the algorithm such as Decision Tree, Naive Bayes and Classification to predict the heart disease. The accuracy of these algorithms are 99.2%, 96.5% and 88.3% respectively. Decision Tree yielded the best result which is at least 2.7% more than others.

In 2013, Vikas [15] applied CART Classification to predict the heart disease. Accuracy of the algorithm is 84.49% and the time consumed by the algorithm is 0.23s with an average error 0.3.

In 2014, Waghulde et al. [18] proposed the models called as Neural Network and Genetic Algorithm for the prediction of heart disease. The accuracy of the system is 98% respectively.

In 2014, Niranjana Devi et al. performed a research work, Evolutionary-Fuzzy Expert System for the Diagnosis of Coronary Artery Disease. In this, a fuzzy expert system with Genetic

Algorithm is proposed to diagnose CAD disease condition. Genetic Algorithm is used to optimize the membership function parameters. The proposed system is validated over CAD dataset and achieved an accuracy of 88.79%.

Rupali et al. [20] the Classification based algorithms Naive Bayes and Laplace smoothing are introduced to predict the heart disease in 2014. The accuracy of these algorithms is 78% and 86% respectively. Classification using Laplace Smoothing yielded the best result than others.

Venkatalakshmi et al. [21] introduced the algorithm such as Naive Bayes and Decision Tree to predict the heart disease in 2014. The accuracy of these algorithms is 85.03% and 84.01% respectively. Naive Bayes yielded the higher result which is at least 1.02% more than others.

In 2015, D'Souza [23] applied three techniques Artificial Neural Network, K Means Clustering and Apriori Algorithm to classify whether the patients have the heart disease or not and their performance are compared. The results showed that Artificial Neural Networks outperform well compare to others.

In 2015, Adbar et al. [25] proposed methods C5.0, Neural Network, SVM, and KNN to predict the risk of heart diseases. Accuracy of c5.0, Neural Network, SVM, and KNN are 93.02%, 89.4%, 86.05% and 80.23% respectively. C5.0 performed well when compared to others. Further, it was also noted that the accuracy of C5.0 is at least 3.62% higher than others.

In 2015, Kau et al. [26] proposed data mining technique using SVM Classifier with Genetic Algorithm and obtained an accuracy of 95%.

In 2015, Swati et al. [27] presented the application of Data Mining techniques in the Healthcare and Prediction of Heart attacks. Classification based data mining techniques such as Naïve Bayes and KNN approaches were analyzed over massive volume of healthcare data. Accuracy of Naive Bayes and K-NN are 84% and 76% respectively.

In 2016, Patel et al. [28] applied methods J48 based UCI and LMT algorithm to predict the heart diseases. Performances of these algorithms are analyzed in terms of accuracy and time complexity. The accuracy of J48 and LMT are 56.76% and 55.77% whereas time consumption these algorithms are 0.04s and 0.39s respectively.

In 2016, Rajalakshmi et al. [29] proposed Weighted Association Classifier (WAC) and K-Means Clustering to predict the heart disease. Accuracy of these algorithms is 93.89% and 92.84% respectively. Combination of these algorithms produced maximum accuracy of 94.54%.

In 2016, Suganya et al. [30] proposed CART classifier to predict the heart disease. Minimum distance CART classifier is to classify the data among various groups. The accuracy of the CART classifier algorithm is 83% respectively.

In 2017, Karthikeyan et al. [31] proposed Deep Belief Network (DBN) algorithm which was offered 90% accuracy in heart diseases prediction. In 2018, Wadhawan [32] applied combination of K-Means and Apriori Algorithm for heart diseases prediction. The method incurred the accuracy of 74%. The Apriori technique found strong rules for prediction. The proposed method yielded results as good as a neural network.

Summarization of the study about heart disease prediction using various computational approaches tabulated in Table.1.

Table.1. Study on Computational Model Based Heart Disease Prediction

Author	Technique Used	Dataset
Resul et al. [2]	Neural Network	UCI machine learning repository
Anbarasi et al. [3]	Genetic with Decision Tree	Cleveland Heart Disease Dataset
	Genetic with Naive Bayes	
	Genetic with Classification via Clustering	
Rajkumar et al. [4]	Naive Bayes	Statlog Heart Disease Dataset
	Decision Tree	
	KNN	
Kumari et al. [5]	Decision Tree	Cleveland Heart Disease Dataset from UCI repository
	Artificial Neural Network	
	Support Vector Machine	
	RIPPER	
Sundar, et al. [6]	WAC	Cleveland Heart Disease Database
	Naive Bayes	
Chaitrali et al. [1]	Artificial Neural Network	Cleveland Heart Disease & Statlog Heart Disease Dataset
John, et. al. [7]	Naive Bayes	Heart Disease data.arff
	Multilayer	
	J48	
	KNN	
Shouman et al. [8]	K-Nearest Neighbour	UCI machine learning repository
Nidhi et al. [9]	Naive Bayes	Cleveland Heart Disease Database, Statlog Database and American Association
	Decision Trees	
	Neural Network	
Pethalakshmi et al. [10]	Fuzzy Decision Tree	UCI Machine Learning Repository
	Fuzzy Naive Bayes	
	Fuzzy Neural Network	
	Fuzzy K-means	
Abhishek et al. [11]	J48	PGI, Chandigarh from the year 2008 to 2011
	J48	
	Naive Bayes	
Chitra et al. [12]	Artificial Neural Network	UCI Machine Learning Repository

	K-Means Clustering	
	Fuzzy C Means Clustering	
Dessai [13]	PNN	Cleveland heart disease database
	Decision Tree	
	Naive Bayes	
	BNN	
Patel et al. [14]	Decision Tree	UCI Machine Learning Repository
	Naive Bayes	
	Classification Clustering	
Vikas and Pal [15]	CART Classification	UCI repository
Methaila et al. [16]	Naive Bayes	Cleveland heart disease database
	Decision Tree	
	Classification via Clustering	
Wisaeng [17]	K-Nearest Neighbor	Medical Dataset
Waghulde et al. [17]	Neural Network & Genetic Algorithm	American Heart Association
Rupali et al. [20]	Classification using Naive Bayes	UCI Machine Learning Database
	Classification using Laplace Smoothing	
Venkatalakshmi et al. [21]	Naive Bayes	UCI Machine Learning Repository
	Decision Tree	
Jarad [22]	Naive Bayes	UCI Machine Learning Repository
	Decision Tree	
	KNN	
D'Souza [23]	ANN	Cleveland Heart Disease Database
	K-Mean Clustering	
Baiju and Janet [24]	Naive Bayesian classification Technique	Cleveland Heart Disease Database
Adbar et al. [25]	C 5.0	Statlog Dataset from UCI repository
	NN	
	SVM	
	KNN	
Kaur and Kaut [26]	SVM Classifier with Genetic Algorithm	American Heart Association
Swati et al. [27]	Naive Bayes	Hospital Dataset
	KNN	
Patel et al. [28]	J48	Cleveland Dataset from UCI repository
Rajalakshmi et al. [29]	K-Means Clustering	Cleveland Dataset from UCI repository
	WAC	
Suganya et al. [30]	CART Classifier	Statlog Heart Disease Database
Karthikeyan et al. [31]	Deep Belief Network	UCI Machine Learning Data Repository
Wadhawan [32]	K-Means Using Apriori Algorithm	Clinical

Table.2. Shows Different Data Mining Techniques used on Heart Disease Detection with Prediction Accuracy

Author	**Technique Used**	**No. of Attributes**	**Accuracy**
Resul et al. [2]	Neural Network	13	89.01%
Anbarasi et al. [3]	Genetic with Decision Tree	6	99.2%
	Genetic with Naive Bayes		96.5%
	Genetic with Classification via Clustering		88.3%
Rajkumar et al. [4]	Naive Bayes	17	52.33%
	Decision Tree		52%
	KNN		45.67%
Kumari et al. [5]	Decision Tree	14	79.05%
	Artificial Neural Network		80.06%
	Support Vector Machine		84.12%
	RIPPER		81.08%
Sundar, et al. [6]	WAC	15	84%
	Naive Bayes		78%
Chaitrali et al. [1]	Artificial Neural Network	13	99.25%
		15	99.9%
John, et. al. [7]	Naive Bayes	14	85.18%
	Multilayer		78.88%
	J48		85.18%
	KNN		85.55%
Shouman et al. [8]	K-Nearest Neighbour	13	97.4%
Nidhi et al. [9]	Naive Bayes	15	90.74%
		13	94.44%,
		6	96.5%
	Decision Trees	15	99.62%,
		13	96.66%,
		6	99.2%
	Neural Network	15	96.5%,
		13	99.2%,
		6	88.3%

Pethalakshmi et al. [10]	Fuzzy Decision Tree	13	90.06%
	Fuzzy Naive Bayes		89.62%
	Fuzzy Neural Network		91.09%
	Fuzzy K-means		99.49%
Abhishek et al. [11]	J48	15	95.56%
	J48		94.85%
	Naive Bayes		92.42%
Chitra et al. [11]	Artificial Neural Network	13	85%
	K-Means Clustering		88%
	Fuzzy C Means Clustering		92%
Dessai [13]	PNN	14	94.6%
	DT		84.2%
	NB		84%
	BNN		84.6%
Patel et al. [14]	Decision Tree	14	99.2%
	Naive Bayes		96.5%
	Classification Clustering		88.3%
Vikas and Pal [15]	CART Classification	11	84.49%
Methaila et al. [16]	Naive Bayes	15	96.53%
	Decision Tree		99.2%
	Classification via Clustering		88.3%
Wisaeng [17]	K-Nearest Neighbor	14	93%
		8	90%
Waghulde et al. [17]	Neural Network & Genetic Algorithm	13	98%
Rupali et al. [20]	Classification using Naive Bayes	14	78%
	Classification using Laplace Smoothing		86%
Venkatalakshmi et al. [21]	Naive Bayes	13	85.03%
	Decision Tree		84.01%
Jarad [22]	Naive Bayes	14	52.33%
	Decision Tree		52%
	KNN		45.67%
D'Souza [23]	ANN	14	79.38%
	K-Mean Clustering		63.299%
Baiju and Janet [24]	Naive Bayesian classification Technique	13	81%
Adbar et al. [25]	C 5.0	14	93.02%
	NN		89.4%
	SVM		86.05%
	KNN		80.23%
Kaur and Kaut [26]	SVM Classifier with Genetic Algorithm	12	95%
Swati B et al. [27]	Naive Bayes	13	84%
	KNN		76%
Patel et al. [28]	J48	13	56.76%
Rajalakshmi et al. [29]	K-Means Clustering	14	93.89%
	WAC		92.84%
Suganya et al. [30]	CART Classifier	14	83%
Karthikeyan et al. [31]	Deep Belief Network	16	90%
Wadhawan [32]	K-Means Using Apriori Algorithm	7	74%

3. DISCUSSION

In this paper, the different data mining methods introduced for heart disease prediction are extensively reviewed and discussed their computational results. Each method has its own advantages and disadvantages. The different computational model yield various accuracy, since, nature of the algorithms, number of features in the data sets, the size the training and testing data is chosen. From the review analysis, it is observed that most of the author used the dataset of Heart disease from UCI Machine Learning Repository, University of California. Hence, it might be a good choice for training. The method KNN [4] used by Rajkumar et al. for the prediction of Heart Disease using 17 attributes from Statlog Heart Disease Dataset and the same method used by Jarad et al. [22] for 14 attributes from UCI Machine Learning Repository. The method obtained minimum accuracy of 45.67% for both cases. Chaitrali et al. [1] obtained the maximum accuracy of 99.9% by ANN model to predict Heart Disease using 15 attributes from Cleveland Heart Disease & Statlog Heart Disease Dataset. For the minimum number of attribute, Decision Tree produced maximum accuracy of 99.2%. The maximum accuracy obtained using Deep Belief Network is 90% by Karthikeyan et al. [31] with maximum attributes 16 from UCI Machine Learning Data Repository. The Decision Tree algorithm produced highest accuracy of 99.2% for different number of attributes from different dataset.

3.1 ISSUES IN PREDICTION OF HEART DISEASE USING COMPUTATIONAL METHODS

From this literature survey, the following issues are identified in heart disease prediction using computational models in health care industry data.

- Prediction accuracy is low with reduced number of attributes.
- It take more time for predict.
- More False classification.
- The choice of the kernel and other parameter selections.
- The limit on ability to categorize correctly.
- Poor clinical decisions lead to mortality
- The noise and missing values make a hurdle to design classification model.

- The irrelevant or redundant attributes Removal is difficult task.
- The suitable Data transformation, discretization and reduction methods are required for the prediction.
- Accuracy and Speed of the model only for some extents.
- The time to construct the model (training time) is high
- Lack of handling noise and missing values
- Interpretability i.e. Level of understanding and insight provided by the model

4. CONCLUSION

This survey projected various computational approaches used for heart disease prediction. The areas of classification and prediction have many issues which need to be address. Certainly, the paper gives a broad overview of the existing techniques to the interested researcher and beginners in the same area of the research. In future, novel algorithms and techniques for prediction of heart disease are to be developed using neural network with Fuzzy or Rough-set or Regularize set and optimization methods to overcome the shortcomings in the existing system and to enhance result more accurately.

REFERENCES

[1] Chaitrali S. Dangare and Sulabha S. Apte, "A Data Mining Approach for Prediction of Heart Disease using Neural Networks", *International Journal of Computer Engineering and Technology*, Vol. 3, No. 3, pp. 23-29, 2012.

[2] Resul Das, Ibrahim Turkoglu and Abdulkadir Sengur, "*Diagnosis of Valvular Heart Disease through Neural Networks Ensembles*", Elsevier, 2009.

[3] M. Anbarasi, E. Anupriya and N. Iyengar, "Enhanced Prediction of Heart Disease with Feature Subset Selection using Genetic Algorithm", *International Journal of Engineering Science and Technology*, Vol. 2, No. 10, pp. 5370-5376, 2010.

[4] A. Rajkumar and G.S. Reena, "Diagnosis of Heart Disease using Data Mining Algorithm", *Global Journal of Computer Science and Technology*, Vol. 10, No. 10, pp. 788-797, 2010.

[5] Milan Kumari and S. Godara, "Comparative Study of Data Mining Classification Methods in Cardiovascular Disease Prediction", *International Journal of Computer Science and Technology*, Vol. 2, No. 2, pp. 304-308, 2011.

[6] N. Aditya Sundar, P. Pushpa Latha, M. Rama Chandra, "Performance Analysis of Classification Data Mining Techniques over Heart Disease Data Base", *International Journal of Engineering Science and Advanced Technology*, Vol. 2, No. 3, pp. 470-478, 2013.

[7] T. John Peter and K. Somasundaram, "Study and Development of Novel Feature Selection Framework for Heart Disease Prediction", *International Journal of Scientific and Research Publications*, Vol. 2, No. 10, pp. 1-7, 2012.

[8] Mai Shouman, Tim Turner and Rob Stocker "Applying K-Nearest Neighbour in Diagnosing Heart Disease Patients", *Proceedings of International Conference on Knowledge Discovery*, pp. 23-29, 2012.

[9] Nidhi Bhatla and Kiran Jyoti, "An Analysis of Heart Disease Prediction using Different Data Mining Techniques", *International Journal of Engineering Research and Technology*, Vol. 1, No. 8, pp. 1-4, 2012.

[10] A. Pethalakshmi and A. Anushya, "Effective Features Selection via Futuristic Genetic on Heart Data", *International Journal of Computational Intelligence and Informatics*, Vol. 2, No. 1, pp. 23-27, 2012.

[11] Abhishek Taneja, "*Heart Disease Prediction System using Data Mining Techniques*", Oriental Scientific Publishing, 2013.

[12] R. Chitra and V. Seenivasagam, "Heart Disease Prediction System using Supervised Learning Classifier", *Bonfring International Journal of Software Engineering and Soft Computing*, Vol. 3, No. 1, pp. 1-7, 2013.

[13] S. Indira and Fal Dessai, "Intelligent Heart Disease Prediction System using Probabilistic Neural Network", *International Journal on Advanced Computer Theory and Engineering*, Vol. 4, No. 2, pp. 33-43, 2013.

[14] Shamsher Bahadur Patel, Pramod Kumar Yadav and D.P. Shukla, "Predict the Diagnosis of Heart Disease Patients using Classification Mining Techniques", *IOSR Journal of Agriculture and Veterinary Science*, Vol. 4, No. 2, pp. 61-64, 2014.

[15] Vikas Chaurasia and Saurabh Pal, "Early Prediction of Heart Diseases using Data Mining Techniques", *Caribbean Journal of Science and Technology*, Vol. 1, pp. 208-217, 2013.

[16] Aditya Methaila, Prince Kansal, Himanshu Arya and Pankaj Kumar, "Early Heart Disease Prediction using Data Mining Techniques", *Proceedings of International Conference on Bioscience and Bioinformatics*, pp. 53-59, 2014.

[17] Kittipal Wisaeng, "Predict the Diagnosis of Heart Disease using Feature Selection and K-Nearest Neighbor Algorithm", *Applied Mathematical Sciences*, Vol. 8, No. 83, pp. 4103-4113, 2014.

[18] Nilakshi P. Waghulde and Nilima P. Patil, "Genetic Neural Approach for Heart Disease Prediction", *International Journal of Advanced Computer Research*, Vol. 4, No. 3, pp. 331-338, 2014.

[19] R. Chitra and V. Seenivasagam, "Review of Heart Disease Prediction System using Data Mining and Hybrid Intelligent Techniques", *ICTACT Journal on Soft Computing*, Vol. 3, No. 4, pp. 605-609, 2013.

[20] Rupali R. Patil, "Heart Disease Prediction System using Naive Bayes and Jelinek-Mercer Smoothing", *International Journal of Advanced Research in Computer and Communication Engineering*, Vol. 3, No. 5, pp. 515-523, 2014.

[21] B. Venkatalakshmi and M.V. Shivsankar, "Heart Disease Diagnosis using Predictive Data mining", *International Journal of Innovative Research in Science, Engineering and Technology*, Vol. 3, No. 3, pp. 223-229, 2014.

[22] A. Jarad, R. Katkar, A. Rehaman Shaikh and A. Salve, "Intelligent Heart Disease Prediction System with MONGODB", *International Journal of Emerging Trends and Technology in Computer Science*, Vol. 4, No. 1, pp. 410-417, 2015.

[23] Andrea D. Souza, "Heart Disease Prediction using Data Mining Techniques", *International Journal of Research in Engineering and Science*, Vol. 3, No. 3, pp. 74-77, 2015.

[24] B.V. Baiju and R.J. Remy Janet, "A Survey on Heart Disease Diagnosis and Prediction using Naive Bayes in Data Mining", *International Journal of Computer Engineering and Technology*, Vol. 5, No. 2, pp. 1034-1038, 2015.

[25] Moloud Adbar et al., "Comparing Performance of Data Mining algorithms in Prediction Heart Diseases", *International Journal of Electrical and Computer Engineering*, Vol. 5, No. 6, pp. 1569-1576, 2015.

[26] Rajwant Kaur and Sukhpreet Kaur, "Prediction of Heart disease Based on Risk Factors Using Genetic SVM Classifier", *International Journal of Advanced Research in Computer Science and Software Engineering*, Vol. 5, No. 12, pp. 205-208, 2015.

[27] B. Shinde Swati, "Decision Support System on Prediction of Heart Disease using Data Mining Techniques", *International Journal of Engineering Research and General Science*, Vol. 3, No. 2, pp. 1453-1461, 2015.

[28] Jaymin Patel, Teja Upadhyay and Samir Patel, "Heart Disease Prediction using Machine Learning and Data Mining Techniques", *International Journal of Computer Science and Communication*, Vol. 7, No. 1, pp. 129-137, 2016.

[29] K. Rajalakshmi and K. Nirmala, "Heart Disease Prediction with Map Reduce by using Weighted Association Classifier and K-Means", *Indian Journal of Science and Technology*, Vol. 9, No. 19, pp. 231-237, 2016.

[30] S. Suganya and P. Tamil Selvi, "A Proficient Heart Disease Prediction Method using Fuzzy-Cart Algorithm", *International Journal of Scientific Engineering and Applied Science*, Vol. 2, No. 1, pp. 1-6, 2016.

[31] T. Karthikeyan and V.A. Kanimozhi, "Deep Learning Approach for Prediction of Heart Disease using Data mining Classification Algorithm Deep Belief Network", *International Journal of Advanced Research in Science, Engineering and Technology*, Vol. 4, No. 1, pp. 3194-3201, 2017.

[32] Rishabh Wadhawan, "Prediction of Coronary Heart Disease using Apriori algorithm with Data Mining Classification", *International Journal of Research in Science and Technology*, Vol. 3, No. 1, pp. 1-15, 2018.

[33] K. Srinivas, B.K. Rani and A. Govrdhan, "Applications of Data Mining Techniques in Healthcare and Prediction of Heart Attacks", *International Journal on Computer Science and Engineering*, Vol. 2, No. 2, pp. 250-255, 2010.

[34] Rashedur M. Rahman and Farhana Afroz, "Comparison of Various Classification Techniques Using Different Data Mining Tools for Diabetes Diagnosis", *Journal of Software Engineering and Applications*, Vol. 6, No. 3, pp. 85-97, 2013.