

An Analysis of Data Mining Techniques in Aggregation with Real Time Dataset for the Prediction of Heart Disease

Priyanka P. Shinde* and Kavita S. Oza**, Rajanish K. Kamat**

ABSTRACT

Data is a collection of facts and figures available in huge volume which needs to be organized into meaningful way to get useful information out of it. Data mining helps to get useful patterns out of big data. Healthcare field has wide scope for research since vast amount of unattended data is available which needs to be investigated for the benefit of human health. A big data related with heart disease can be investigated by applying different data mining techniques and come up with predictive model that will help patients to diagnose the disease at early stage. Most of the studies have stated that Decision Tree is an efficient technique to identify useful statistics out of available data. Accuracy of result set can be enhanced further using Genetic Algorithm in conjunction with data mining techniques. Real time dataset can help to understand the significant trends in heart disease which could play instrumental role in prediction of heart disease. The aim of study is to revise the existing findings, analyze and co-relate heart disease data, find better solution and minimize unwanted medical costs for heart disease diagnosis and save many lives in right time.

Keywords: Data Mining techniques, Heart disease, Decision Tree, Prediction, Medical dataset.

I. INTRODUCTION

In today's age of internet, information has been scattered in the form of Statistics, Reports, and Forms etc. Using this huge available data it is benefited to deliver insights that can enable decision making in real time situations.

Though plenty of research has already been performed in different areas, health care has wide scope to use already available data and derive results which will be novel to the world.

1.1. Heart Disease

Efficient working of heart is very important in human life. If function of heart is not good then it influences other parts of body. Functioning of heart and brain are interdependent, when heart and brain stops working death occurs within minutes. If blood circulation is not proper then heart and brain suffers. When blood gets blocked in brain, it is called as brain stroke and when blood gets blocked in heart, it is called as heart attack. So the heart plays an important role in human body. Chronic heart disease has new opportunities to investigate a numerous clinical and administrative questions.

The World Health Organization (WHO) survey has recorded 12 million deaths due to Heart disease every year and also observed that Heart disease kills one person every 34 seconds [1]. So there is a need to analyze available Data related with chronic heart disease and organize in such a manner that some significant facts and patterns will be derived out of it for the benefit of human being.

Many factors cause risk of Heart disease like [2]:

* Government College of Engineering, Karad, 415124, Maharashtra, India, Email: ppshinde.gcek@gmail.com

** Shivaji University, Kolhapur, 416004, Maharashtra, India, Emails: skavita.oza@gmail.com, rkk_eln@unishivaji.ac.in

Table 1
Risk Factors of heart disease

<i>Family History</i>	<i>Obesity</i>	<i>Unhealthy diet</i>	<i>Cholesterol</i>	<i>High sugar level</i>	<i>High B. P.</i>
Smoking	Hyper tension	Harmful use of alcohol	Physical inactivity	Heredity	etc.

There are many types of heart disease exist like:

1. Coronary heart disease (due to plaque supply of blood is reduced)
2. Angina pectoris (chest pain and it's a warning of attack)
3. Congestive heart failure (pumping of heart is not well)
4. Cardiomyopathy (weakening of heart muscle due to high BP, use of alcohol etc.)
5. Congenital heart disease (formation of abnormal heart)
6. Arrhythmias (abnormal heart beats like slow, fast, irregular)
7. Myocarditis (viral, fungal infections affecting the heart) [2] etc.

2. DATA MINING

Data mining provide techniques and tools which are used to analyze the data and find meaningful patterns and provide the knowledge based decisions. In healthcare, diagnosis is very complicated task so it is needed to use data mining technique and discover the useful information to prevent the risk of heart disease, find the better solutions and minimize the unwanted medical cost for heart disease diagnosis and save many lives in proper time.

Data Mining Techniques are:

Classification

Decision Tree

Neural Network

Clustering

K Nearest Neighbor

Genetic Algorithm

2.1. Classification

This technique involves learning and classification – In learning; training data are analyzed by classification algorithm. Depends on test data, classification rule will be estimated, and if enough accuracy is their then it will be used to apply on new data. Classification technique could be helpful to find out more accurate results for chronic heart disease since it helps to split data into useful groups.

2.2. Decision Tree

Decision tree is used as decision support tool e.g. in operation research it is used for the purpose of decision analysis. As in tree, we have Node, branch and leaf node – “test” is the node, outcome of the test is branch and decision taken after analyzing all attributes becomes leaf node. The path taken from root node to leaf node can be considered as classification rules.

2.3. Neural Network

Artificial Neural Network (ANN) OR just Neural Network is inspiration of biological Neural Network and basically a mathematical model. A Neural Network is group of artificial neurons which are interconnected

to each other. Neural Network has ability of pattern recognition, data classification from complicated data. Neural Network could be one of the most efficient techniques of data mining.

2.4. Clustering

In technique of clustering data is assembled into a meaningful sub modules called as clusters. So in one module or cluster the data will be bundled which has natural similarities. Like classification technique, clustering technique could play vital role for the prediction of heart disease well in advance.

2.5. K Nearest Neighbor (KNN)

This is non-parametric classification method which classifies each case based on similarity measure. Two types are used, in first type – structure less nearest neighbor, data is classified into training and test sample data, distance evaluated from training data to sample data and data with lowest distance called as NN. In second type- structure based nearest neighbor is based on structure of data like axis tree, k-d tree etc.

2.6. Genetic Algorithm

Genetic Algorithm is the most popular technique in evolutionary computing used for problem optimization. In genetic algorithm the term genes (individual element) represent the problem, chromosome represents the solution which contains number of genes. There are basically three operations that are performed on data in GA – Selection, Crossover and Mutation.

3. OBJECTIVE OF THE STUDY

The objective of this study is to examine different data mining techniques that could play crucial role for the prediction of heart disease, discover the useful information from available data to prevent the risk of heart disease, find the better solutions and minimize unwanted medical costs for heart disease diagnosis to save many lives in appropriate time.

4. LITERATURE REVIEW

A paper on “Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques” [3] by Chaitrali Dangare and Sulabha Apte analyzed prediction system by using more number of attributes and added obesity and smoking attributes and demonstrated that neural network is more accurate than other data mining techniques like Decision Tree and Naïve Bayes. A big amount of data has been analyzed using three different techniques and concluded that Neural Network has highest accuracy.

A paper presented on “An Efficient Method for Predicting Heart Disease Problem using Fitness Value” [4] by Prachi Paliwal and Mahesh Malviya introduced a new method based on fitness value by of the attributes to find out more accurate results in order to predict the heart disease problem. 10 attributes were considered which are responsible for heart attack problem and given dataset converted into binary formats. Further dividing the datasets into subsets, common attributes satisfying the conditions were detected. It has been claimed that proposed method is simple and more accurate.

A study on the subject “Predictive Data mining for Medical Diagnosis: An overview of Heart Disease Prediction” [1] conducted by Jyoti Soni, Ujma Ansari, Dipesh Sharma and Sunita Soni have reported that Decision Tree is much accurate and for a time Bayesian classification is having similar correctness but practices like KNN, Neural Networks, Classification based on clustering are not achieving the accuracy up to the mark. Another conclusion was that by applying genetic algorithm the precision of the Decision Tree and Bayesian Classification improves.

A research performed on “Classification of Heart Disease Using K-Nearest Neighbour and Genetic Algorithm” [5] by M. Akhil Jabbar, B. L Deekshatulu and Priti Chandra have stated that there exists lot of other Data Mining techniques in which “Classification” is one of them. By using less number of attributes efficient prediction of heart disease diagnosis is possible with the help of Classification method. A new algorithm has been proposed which combines KNN (K-Nearest Neighbour) and Genetic Algorithm for effective classification. Experiments conducted by using this algorithm showed that accuracy is enhanced in diagnosis of heart disease.

Another research performed by M. Akhil Jabbar, B. L Deekshatulu and Priti Chandra with the title “Heart Disease Prediction System using Associative Classification and Genetic Algorithm” [6] have proposed efficient associative classification algorithm practicing genetic approach for heart disease prediction. Gini Index used to produce a compact rule set and filter rules by applying z-statistics and Genetic Algorithm. Informative attributes used for rule generation and hypothesis testing z-statistics for heart disease prediction. In future work it is proposed that number of attributes will be reduced and only those attributes will be considered which contributes towards the diagnosis of disease using genetic algorithm.

A research performed on “Decision Support System for Heart Disease Prediction using Data Mining Classification Techniques” [7] by Ankur Makwana and Jaymin Patel have stated that to get better classification precision of Heart disease dataset, combination of Naive Bayes and Genetic Algorithm is used . Proposed algorithm has two parts, first managing attributes using genetic search and second is building classifier and measuring accuracy.

A paper presented on “Modelling and design of evolutionary neural network for heart disease detection” [8] by K. S. Kavitha, K. V. Ramakrishnan and Manoj Kumar Singh introduced a new method based on feed forward neural network architecture and genetic algorithm for detection of heart disease. Using genetic algorithm hybridization has applied on neural network and proved that results are more stable compare to back propagation.

A study on “Comparison of ANNs, Fuzzy Logic and NeuroFuzzy Integrated Approach for Diagnosis of Coronary Heart Disease: A Survey” [9] by Nitin Kumari, Sunita, Smita have compared data mining techniques for healthcare data are ANNs, Fuzzy logic and NeuroFuzzy integrated approach. Result of comparison tells that NeuroFuzzy integrated approach is best for diagnosis of heart disease. ANN and fuzzy logic gives lot of advantages but have some disadvantages too. The technique NeuroFuzzy combines features of artificial neural network and fuzzy logic and give the best result.

In “Review of Heart Disease Prediction System using Data Mining and Hybrid Intelligent Techniques” [10] R. Chitra, Dr. V. Sreenivasagam have stated that accuracy of heart disease prediction has improved using Hybrid Intelligent Algorithm. Prediction System runs in three types, in first analysis of data mining technique – Classification technique, in second type intelligent technique such as Neural Network are analyzed and in third type discussion on role of feature subset in heart disease prediction. Some data mining techniques are analyzed in this review paper and concluded that Neural Network on offline training is an effective method for prediction of heart disease at an early stage.

A study on “Predict the Diagnosis of Heart Disease Patients Using Classification Mining Techniques” [11] by Shamsher Bahadur Patel, Pramod Kumar Yadav, Dr. D. P. Shukla have used initially 14 attributes for the prediction of heart disease which was reduced to 6 attributes that are mainly responsible for the heart disease. Attributes were reduced from 14 to 6 by using Genetic Algorithm. Genetic Algorithm was useful to identify attributes that contribute more towards the diagnosis of heart disease and hence reducing extra attributes that are not crucial for the prediction. Patient need to take less number of tests by using this approach. Three data mining techniques used namely Naïve Bayes, Decision Tree and Classification by Clustering. All classifiers were feed with the reduced 6 attributes and found that Decision Tree outperforms than other two techniques with 99.2% accuracy while Naïve Bayes performs consistently before and after

Table 2
Comparative studies from year 2010 to 2015

<i>Author</i>	<i>Topic</i>	<i>Year</i>	<i>Techniques used</i>	<i>Conclusion</i>
Ankur Makwana and Jaymin Patel, 2015 [7]	Decision Support System for Heart Disease Prediction using Data Mining Classification Techniques	2015	Naïve Bayes and Genetic Algorithm combined. Thesis divided in two parts—managing attributes using Genetic search and building classifier.	Combination of Naïve Bayes and Genetic Algorithm will improve the classification accuracy of Heart disease dataset.
Prachi Paliwal and Mahesh Malviya, 2015 [4]	An Efficient Method for Predicting Heart Disease Problem using Fitness Value	2015	Introduced a new method based on fitness value. 10 attributes were used which are responsible for heart disease.	A new method introduced to find predict heart disease has been claimed as simple and more accurate.
S. U. Amin, K. Agarwal and R. Beg, 2015 [14]	Genetic neural network based data mining in prediction of heart disease using risk factors	2015	Hybrid of NN and Genetic Algorithm	Proposed system gives 89% correctness to predict the risk of heart disease.
Jimson Mathew, Ashutosh Kumar Singh, Ilayaraja M., Meyyappan T. 2015 [15]	Efficient Data Mining Method to Predict the Risk of Heart Diseases Through Frequent Itemsets	2015	Association rule frequent itemset	Proposed method gives risk level efficiently from frequent itemsets
Nitin Kumari, Sunita, Smita, 2013 [9]	Comparison of ANNs, Fuzzy Logic and NeuroFuzzy Integrated Approach for Diagnosis of Coronary Heart Disease: A Survey	2013	Compared data mining techniques for healthcare data are ANNs, Fuzzy logic and NeuroFuzzy integrated approach.	Comparison of different data mining technique tells that NeuroFuzzy gives best results since it combines features of artificial neural network and fuzzy logic.
R. Chitra and V. Seenivasagam, 2013 [10]	Review of Heart Disease Prediction System Using Data Mining And Hybrid Intelligent Techniques	2013	“Classification” technique used as data mining technique, “Neural Network” used as intelligent technique.	Proposed hybrid intelligent algorithm enhance accuracy of heart disease prediction.
Shamsher Bahadur Patel, Pramod Kumar Yadav, Dr. D. P. Shukla, 2013 [11]	Predict the Diagnosis of Heart Disease Patients Using Classification Mining Techniques	2013	Naïve Bayes, Decision Tree and Classification by Clustering	14 attributes reduced to 6 and found that decision Tree gives more accurate results, where Naïve Bayes does not have any difference and Classification by Clustering gives poor result than other two techniques.
M.A.Nishara Banu, B Gomathy, 2013 [12]	Disease Predicting System Using Data Mining Technique's	2013	Association rule mining, Classification, Clustering	MAFIA algorithm used with Classification technique has given more accurate results.
Chitra, Dr. V. Sreenivasagam, 2013 [13]	Heart disease Prediction System Using Supervised Learning Classifier	2013	CNN (Cascaded Neural Network) and SVM (Support Vector Machine) classifiers were analyzed.	13 attributes of 270 patients were analyzed and found that CNN performs with better accuracy than SVM.
M. Akhil Jabbar, B. L Deekshatulu and Priti Chandra, 2013 [5]	Classification of Heart Disease Using K-Nearest Neighbour and Genetic Algorithm	2013	“Classification” prediction model helps efficient heart disease analysis with minimum attributes. Proposed new algorithm which combines KNN and genetic Algorithm for effective classification	A new algorithm which combines KNN and GA gives greater accuracy.
M. Akhil	Heart Disease Predi-	2012	Gini Index used and further applied Z-	Number of attributes will be reduced to

(contd...)

(Table 1 contd...)

Author	Topic	Year	Techniques used	Conclusion
Jabbar, B. L Deekshatulu and Priti Chandra, 2012 [6]	ction System using Associative Classification and Genetic Algorithm		statistics and Genetic Algorithm. Only informative attributes used for rule generation.	only those that are responsible for heart disease and Genetic Algorithm will be applied to get more accurate results.
Jyoti Soni, Ujma Ansari, Dipesh Sharma and Sunita Soni, 2011 [1]	Predictive Data mining for Medical Diagnosis: An overview of Heart Disease Prediction	2011	Practices like Decision Tree, Bayesian Classification, KNN, Neural Networks, Classification based on clustering are evaluated.	1. Decision Tree is more accurate than other techniques and sometimes Bayesian Classification has similar accuracy. 2. In additionally Genetic Algorithm applied, it improves precision of prediction by lessening actual data size.
K.S.Kavitha, K.V. Ramakrishnan and Manoj Kumar Singh, 2010 [8]	Modelling and design of evolutionary neural network for heart disease detection	2010	New method introduced based on feed forward neural network architecture and genetic Algorithm for detection of heart disease.	Applying genetic algorithm hybridization on neural network has proved that results are more constant compare to back propagation.

reducing the number of attributes where Classification by Clustering is poor compared to other two techniques.

A paper presented on “Disease Predicting System Using Data Mining Technique’s” [12] by M. A. Nishara Banu, B. Gomathy have stated that medical field is information rich but feeble in knowledge. It means lack of tools that are used for powerful analysis and hence has wide scope for research. Different data mining techniques like association rule mining, classification, clustering is implemented and tried to find out different problems related with heart. A MAFIA algorithm used for data classification which has given more accurate results. K-means clustering algorithm has been used for the clustering of heart disease database. 16 attributes were considered for the prediction of heart disease well in advance. In this paper, an efficient approach has been presented which helps to find and extract considerable forms from the heart disease database for the prediction of heart disease well in advance.

A study on “Heart disease Prediction System Using Supervised Learning Classifier” [13] by R. Chitra, Dr. V. Sreenivasagam have presented for the reason of testing, records of 270 patients are collected and this information classified using CNN (Cascaded Neural Network) classifier. Total 13 attributes are used as input to CNN classifier to know the risk of heart disease. Supervised Learning algorithm has been adopted for heart disease prediction at the earliest stage. CNN and SVM (Support Vector Machine) are analyzed by applying data obtained from 270 patients and it has shown that CNN has better accuracy than SVM, because CNN does not uses predefined set of hidden units, instead it will minimize an error by adding hidden units one by one.

A paper presented on “Genetic neural network based data mining in prediction of heart disease using risk factors” [14] by S. U. Amin, K. Agarwal and R. Beg have stated that there is a large amount of data available and data mining techniques can be used for the diagnosis of many diseases with more correctness. There are many common visible risk factors which can cause the heart disease, so author has presented the work of prediction of heart disease focused on these risk factors. Hybrid of NN and Genetic Algorithm techniques were used to combine advantages of both. MATLAB was used to implement proposed system which gives 89% correctness to predict the risk of heart disease.

A study on “Efficient Data Mining Method to Predict the Risk of Heart Diseases through Frequent Itemsets” [15] by Jimson Mathew, Ashutosh Kumar Singh, Ilayaraja M., Meyyappan T. have generated

frequent itemsets using mining association rule based on symptoms and least support value. Based on chosen factors, projected method can be used to forecast risk of heart disease at an early stage. Experimentally stated that proposed method efficiently predicts risk level from frequent itemsets.

5. HEART DISEASE DATASET

There are many different attributes exist but following are recorded to find useful information in heart disease dataset [16]:

Table 3
Dataset Description

Attribute	Explanation and values
Age	Recorded in years
Sex	1=male0=female
Cp (chest pain type)	1=typical angina2=atypical angina3=non-angina pain4=asymptomatic
Trestbps(resting blood pressure)	In mm Hg on admission to the hospital
Chol	Serum cholesterol in mg/dl
Fbs(fasting blood sugar>120 mg/dl)	1=true0=false
Restecg(resting electrocardiographic results)	0=normal1=having ST-T wave abnormality2=showing probable left ventricular hypertrophy
Thalach	Maximum heart rate achieved
Exang(exercise induced angina)	1=yes0=no
Oldpeak	ST depression included by exercise related to rest
Slope(slope of the peak exercise ST segment)	1=unsloping2=flat3=down sloping
Ca	Number of major vessels (0-3)
Thal	3=normal6=fixed defect7=reversible defect

6. DATASET ANALYSIS

A dataset of 303 records related with heart disease has been analyzed by considering attributes in above table.

Table 4
Analysis of dataset

Attribute	Explanation and values	No. of Patients	Percent (%)
Age (Recorded in years)	0-30	1	0.3
	31-40	18	6
	41-50	76	25
	51-60	129	43
	61-70	73	24
	71-Above	6	2
Sex	1=male	207	68
	0=female	96	32
Cp (chest pain type)	1=typical angina	23	8
	2=atypical angina	50	17
	3=non-angina pain	87	29
	4=asymptomatic	143	47

(contd...)

(Table 2 contd...)

Attribute	Explanation and values	No. of Patients	Percent (%)
Trestbps(resting blood pressure)	In mm Hg on admission to the hospital	101 - 200	
Chol	Serum cholesterol in mg/dl	126-417	
Fbs(fasting blood sugar>120 mg/dl)	1=true 0=false	45 258	15 85
Restecg(resting electrocardiographic results)	0=normal 1=having ST-T wave abnormality 2=showing probable left ventricular hypertrophy	152 4 147	50 1 49
Thalach	Maximum heart rate achieved	71-202	
Exang(exercise included angina)	1=yes 0=no	99 204	33 67
Oldpeak	ST depression included by exercise related to rest	0-6	
Slope(slope of the peak exercise ST segment)	1=unslowing 2=flat 3=down sloping	142 140 21	47 46 7
Ca	Number of major vessels (0-3)	0	0
Thal	3=normal 6=fixed defect 7=reversible defect	166 18 117	55 6 39

Some observations found on the dataset are:

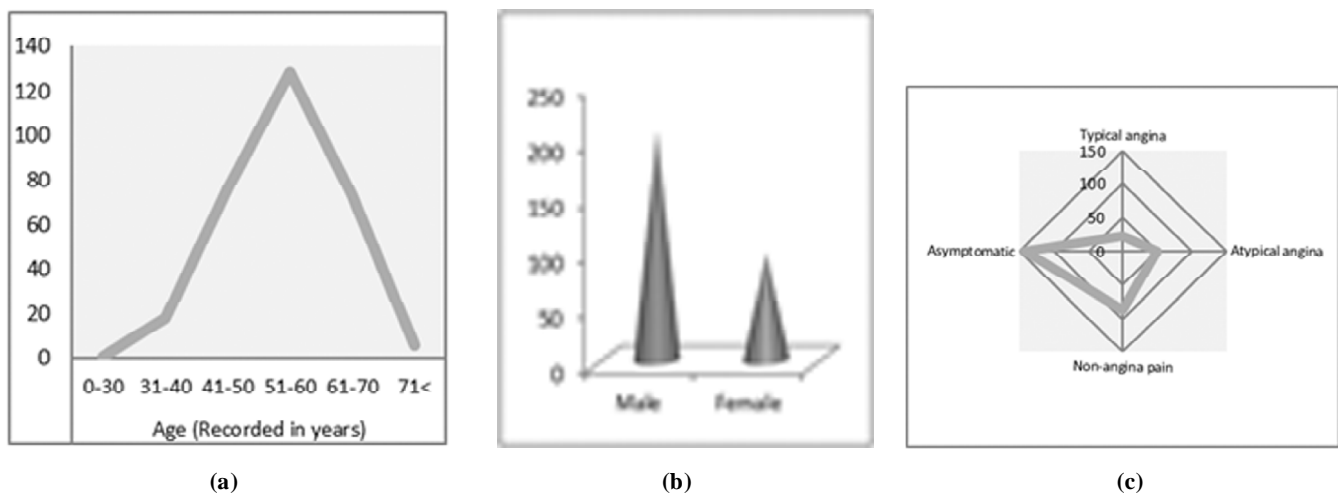


Figure 1: (a) Risk analysis for different age group (b) Risk analysis for gender (c) Chest Pain Risk Analysis

In age group of 51 to 60, 43 % people suffer with heart disease whereas Age group of 41-50 and 61-70, around 25% people are casualty to heart disease. It shows that, people above 40 years age should be more alert and learn symptoms of heart disease to identify it at early stage and get immediate help to save life, avoid serious disability.

As compared to female near about double percent of males are suffered by heart disease. Also it can be concluded that heart disease symptoms in women can be different than men.

As per records of chest pain due to non-angina, 29% people suffer by heart disease and 47% people suffer due to asymptomatic. Available data also shows that 85 % patients have normal blood sugar and only 15% patients have high level of fasting blood sugar (greater than 120 mg/dl).

7. CONCLUSION

This survey presents a study on approaches of data mining techniques towards heart disease prediction, reuse of available data for risk assessment and diagnosis of heart disease. Numerous studies have stated that Decision Tree can be instrumental data mining technique to explore the research around heart disease prediction. In several studies, Genetic Algorithm has been applied in combination with other data mining techniques to find out more accurate results which concludes that Genetic Algorithm plays vital role while working on prediction of heart disease and one should necessarily consider Genetic Algorithm to get more viable results. An analysis of dataset related with heart disease gives indication that people in age group of 40-60 years should be aware and alert of heart disease symptoms since this age group carries more risk of heart disease. Males have higher risk of heart disease than female gender. According to study of heart disease dataset it can be concluded that diverse attributes of human body should be considered while depicting heart disease predictive model. Though moderate number of studies has been performed around heart disease prediction area, still there is a huge scope of research to predict heart disease at early stage with the help of different data mining techniques.

REFERENCES

- [1] J. Soni, U. Ansari, D. Sharma and S. Soni, "Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction", *International Journal of Computer Applications*, vol. 17, no. 8, pp. 43-48, 2011.
- [2] K.Sudhakar, Dr. M. Manimekalai , "Study of Heart Disease Prediction using Data Mining", *International Journal of Advanced Research in Computer Science and Software Engineering*, Volume 4, Issue 1, January - 2014, pp. 1157-1160 ISSN: 2277 128X
- [3] C. S.Dangare and S. S. Apte, "Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques", *International Journal of Computer Applications*, vol. 47, no. 10, pp. 44-48, 2012.
- [4] Prachi Paliwal, Mahesh Malviya, "An Efficient Method for Predicting Heart Disease Problem Using Fitness Value", (*IJCSIT*) *International Journal of Computer Science and Information Technologies*, Vol. 6 (2), 2015, 1290-1293
- [5] M. Akhil Jabbar, B.L. Deekshatulub, Priti Chandra, "Classification of Heart Disease Using K- Nearest Neighbour and Genetic Algorithm", *First International Conference on Computational Intelligence: Modelling Techniques and Applications (CIMTA) 2013* Volume 10, 2013, Pages 85-94
- [6] M. Akhil Jabbar, B. L Deekshatulu and Priti Chandra, "Heart Disease Prediction System using Associative Classification and Genetic Algorithm", *International Conference on Emerging Trends in Electrical, Electronics and Communication Technologies-ICECIT, 2012*
- [7] A. Makwana and J. Patel, "Decision Support System for Heart Disease Prediction using Data Mining Techniques", *International Journal of Computer Applications*, vol. 117, no. 22, pp. 1-5, 2015.
- [8] K. Kavitha, K. Ramakrishnan and M. Singh, "Modelling and design of evolutionary neural network for heart disease detection", *IJCSI International Journal of Computer Science Issues*, Vol. 7, Issue 5, September 2010, vol. 7, no. 5, p. 272, 2010.
- [9] Nitin Kumari, Sunita, Smita, "Comparison of ANNs, Fuzzy Logic and NeuroFuzzy Integrated Approach for Diagnosis of Coronary Heart Disease: A Survey", *International Journal of Computer Science and Mobile Computing, IJCSMC*, Vol. 2, Issue. 6, June 2013, pg. 216 – 224
- [10] R. Chitra and V. Seenivasagam, "Review of heart disease prediction system using data mining and hybrid intelligent techniques", *ICTACT journal on soft computing*, July 2013, VOLUME: 03, ISSUE: 04, ISSN: 2229-6956(ONLINE)
- [11] Shamsheer Bahadur Patel, Pramod Kumar Yadav, Dr. D. P.Shukla, "Predict the Diagnosis of Heart Disease Patients Using Classification Mining Techniques", *IOSR Journal of Agriculture and Veterinary Science (IOSR-JAVS)* e-ISSN: 2319-2380, p-ISSN: 2319-2372. Volume 4, Issue 2 (Jul. - Aug. 2013), PP 61-64 www.iosrjournals.org
- [12] M.A.Nishara Banu, B Gomathy, "Disease Predicting System Using Data Mining Technique's" *International Journal of Technical Research and Applications* e-ISSN: 2320-8163, www.ijtra.com Volume 1, Issue 5 (Nov-Dec 2013), PP. 41-45]

- [13] R. Chitra, Dr. V. Sreenivasagam, "Heart disease Prediction System Using Supervised Learning Classifier" *Bonfring International Journal of Software Engineering and Soft Computing*, Vol. 3, No. 1, March 2013
- [14] S. U. Amin, K. Agarwal and R. Beg, "Genetic neural network based data mining in prediction of heart disease using risk factors," *Information & Communication Technologies (ICT), 2013 IEEE Conference on*, JeJu Island, 2013, pp. 1227-1231. doi: 10.1109/CICT.2013.6558288
- [15] Jimson Mathew, Ashutosh Kumar Singh, Ilayaraja M., Meyyappan T., "Efficient Data Mining Method to Predict the Risk of Heart Diseases Through Frequent Itemsets, Procedia Computer Science", *Proceedings of the 4th International Conference on Eco-friendly Computing and Communication Systems* Volume 70, 2015, Pages 586-592, ISSN 1877-0509, <http://dx.doi.org/10.1016/j.procs.2015.10.040>.
- [16] Lichman, M. (2013). UCI Machine Learning Repository [<http://archive.ics.uci.edu/ml>]. Irvine, CA: University of California, School of Information and Computer Science.