



ISSN: 0976-3031

Available Online at <http://www.recentscientific.com>

CODEN: IJRSFP (USA)

*International Journal of Recent Scientific Research*  
Vol. 9, Issue, 2(H), pp. 24350-24354, February, 2018

**International Journal of  
Recent Scientific  
Research**

DOI: 10.24327/IJRSR

## Research Article

### A STUDY ON VARIOUS DATA MINING TECHNIQUES USED FOR HEART DISEASES

**KolluruVenkata Nagendra and Ussenaiah M**

Department of CS, VS University, Nellore

DOI: <http://dx.doi.org/10.24327/ijrsr.2018.0902.1646>

#### ARTICLE INFO

##### Article History:

Received 16<sup>th</sup> November, 2017  
Received in revised form 25<sup>th</sup>  
December, 2017  
Accepted 23<sup>rd</sup> January, 2018  
Published online 28<sup>th</sup> February, 2018

#### ABSTRACT

The rapid growth of the Data Mining Techniques in the Health care is used to take appropriate decisions in diagnosis. Due to the experience and knowledge of the Doctors clinical diagnosis was done. But from the last decade decision support system with computers plays a vital role in health care industry. In this paper the commonly used Data Mining Technique for the Heart disease prediction are summarized. This survey helps to understand the Data Mining Techniques, which are used to predict the heart disease. At last, the survey summarizes the heart disease factors, symptoms and performance study of Data Mining.

##### Key Words:

Data mining, disease, SVM, Neural Networks, Classification, Medical Data Mining.

**Copyright © KolluruVenkata Nagendra and Ussenaiah M, 2018**, this is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

#### INTRODUCTION

Data mining is the process of extracting hidden knowledge from large volume of data. This knowledge new, not obvious and one must be able to use it. Data Mining usage witnessed unprecedented growth in the last few years. Data Mining Techniques which are applied to medical data include association rule, classification, clustering etc.

The diseases prediction plays an important role in Data Mining. There are different types diseases predicted in Data mining namely Hepatitis, Lung Cancer, Liver Disorder, Breast Cancer, Thyroid diseases and Heart diseases. This paper analysis the Heart Diseases.

Our next section presents causes and Impact of Heart diseases. Section III describes symptoms of Heart attacks. Types of Diseases discussed in IV. Section V summarizes the Literature Survey and in section VI Issues and Challenges, health care applications in section VII and in the final section conclusion followed by references..

##### Causes and Impact of Heart Diseases

According to WHO report Global atlas on cardiovascular disease prevention and control states that cardiovascular disease (CVDs) are the leading causes of death and disability in the world. Although a large proportion of CVDs is preventable,

they continue to rise mainly because preventive measures are inadequate.[1][2].

- Protect heart
- Cardiovascular Diseases (Cvds) Key Facts
- Cardiovascular Diseases

The term Heart disease refers to disease of heart & blood vessel system within it. There are number of factors which increase the risk of Heart disease [3]:

- Family history of heart disease
- Smoking
- Cholesterol
- Poor diet
- High blood pressure
- High blood cholesterol
- Obesity
- Physical inactivity
- Hyper tension

##### Symptoms of a Heart Attack

Symptoms of a heart attack can include: [4]

- Discomfort, pressure, heaviness, or pain in the chest, arm, or below the breastbone.
- Discomfort radiating to the back, jaw, throat, or arm.

\*Corresponding author: **KolluruVenkata Nagendra**  
Department of CS, VS University, Nellore

- 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 Types

### Types of Heart Diseases

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. [5].Some types of Heart diseases are

S.No	Heart Disease	Description
1	Coronary heart disease	It also known as coronary artery disease (CAD), it is the most common type of heart disease across the world. It is a condition in which plaque deposits block the coronary blood vessels leading to a reduced supply of blood and oxygen to the heart.
2	Angina pectoris	It is a medical term for chest pain that occurs due to insufficient supply of blood to the heart. Also known as angina, it is a warning signal for heart attack. The chest pain is at intervals ranging for few seconds or minutes.
3	Congestive heart failure	It is a condition where the heart cannot pump enough blood to the rest of the body. It is commonly known as heart failure.
4	Cardiomyopathy	It is the weakening of the heart muscle or a change in the structure of the muscle due to inadequate heart pumping
5	Congenital heart disease	It also known as congenital heart defect, it refers to the formation of an abnormal heart due to a defect in the structure of the heart or its functioning.
6	Arrhythmias	It is associated with a disorder in the rhythmic movement of the heartbeat. The heartbeat can be slow, fast, or irregular.
7	Myocarditis	It is an inflammation of the heart muscle usually caused by viral, fungal, and bacterial infections affecting the heart.

### LITERATURE SURVEY

K. Srinivas *et al.*[6] proposed Application of Data Mining Technique in Healthcare and Prediction of Heart Attacks. The prospective use of classification based data mining techniques such as Rule based, Decision tree, Naive Bayes and Artificial Neural Network to the massive Volume of healthcare data.

In Latha Parthiban *et al.*[7] presented a model on basis of Coactive Neuro-Fuzzy Inference System (CANFIS) for prediction of heart disease. The CANFIS model diagnosed the presence of disease by merging different techniques that includes the neural network adaptive capabilities, the fuzzy logic qualitative approach and further integrating with genetic algorithm.

Mrs.G.Subbalakshmi (M.Tech), *et al.* [8], developed a Decision Support in Heart Disease Prediction System (DSHDPS) using data mining modelling technique, namely, Naïve Bayes. Using medical profiles such as age, sex, blood pressure and blood sugar it can predict the likelihood of patients getting a heart disease.

N.Deepika, *et al.* [9], presents about the various effective heart attack prediction system using PCAR: an Efficient Approach for mining Association Rules. A proficient methodology for the generation of association rules from the heart disease warehouses for heart attack prediction has been presented. Mai Shouman, *et al.* [10] proposed k-means clustering with the decision tree method to predict the heart disease. In their work

they suggested several centroid selection methods for kmeans clustering to increase efficiency.

The below Table illustrates different data mining techniques applied to Heart Disease data. The different Data Mining Techniques were applied very efficiently to all the Medical Data Mining.

### Issues and Challenge

Applying data mining in the medical field is a very challenging task in medical profession.[11] In medical research the data mining begins with a hypothesis and results are adjusted to fit the hypothesis. This differs from standard data mining practice, which simply starts with datasets without an apparent hypothesis. [12]Patterns and trends in dataset are mainly concerned with traditional data mining, but in medical data mining they are not conformed. According to the doctor intuition the clinical decision are often made. The quality of service provided to patients is affected due to unwanted bias, errors and excessive medical cost. Data mining have the capacity to generate a knowledge-rich environment. It can help to improve the significant quality of clinical decision. [13] In the survey of [14] the three supervised machine learning algorithms are used. These algorithms have been used for analyzing the heart disease dataset. The Classification Accuracy should be compared for this algorithm. This work should be extended to predict the heart disease with reduced number of attributes. In the survey of the heart disease is predicted by using association rule data mining technique. The author introduced an algorithm that uses search constraint to decrease the number of rules. In future this work should be extended by using fuzzy learning models to find the accuracy of time to decrease the number of rules. In the survey of the author proposed a new concept that uses weighted association rule for classification. In future this work can be extended by using association rule hiding technique in data mining. In the survey of [13] the author proposed the minimal subset of attributes for predicting heart disease. In future this work can be expanded and enhanced for the automation of heart disease prediction. Real data should be collected from health care organizations and agencies are taken to compare the optimum accuracy with all data mining technique. In the survey of [15] the author predicts attributes of a diabetic patient getting a heart disease. Weka tool is performed as a result bayes model was able to classify 74% of the input instances correctly. In future this work is extended by using other data mining techniques.

### Data Mining Applications in Health Care

To take the accurate decisions in health care the Data Mining is very needful. The following are the different applications [16].

- Data mining provides support for constructing a model for managing the hospital resources which is an important task in healthcare. Using data mining, it is possible to detect the chronic disease and based on the complication of the patient disease prioritize the patients so that they will get effective treatment in timely and accurate manner.
- Different data mining approaches are used to analyze the various hospital details in order to determine their ranks [18]. Ranking of the hospitals are done on the basis of their capability to handle the high risk patients.

- Data Mining helps the healthcare institute to understand the needs, preferences, behavior, patterns and quality of their customer in order to make better relation with them.
- A system for inspection is constructed using data mining techniques to discover unknown or irregular patterns in the infection control data. Association rules are used to produce unexpected and interesting information from the public surveillance and hospital control data.
- Using Data Mining, physicians and patients can easily compare among different treatments technique. They can analyze the effectiveness of available treatments and find out which technique is better and cost effective. Data Mining also helps them to identify the side effects of particular treatment, to make appropriate decision to reduce the hazard and to develop smart methodologies for treatment.
- Data mining helps the healthcare providers to identify the present and future requirements of patients and their preferences to enhance their satisfaction levels. Milley has also recommended that data mining are useful to determine the requirement of particular patients for enhancing the services provided by healthcare organization [19].
- Healthcare insurer develops a model to detect the fraud and abuse in the medical claims using data mining techniques. This model is helpful for identifying the improper prescriptions, irregular or fake patterns in medical claims made by physicians, patients, hospitals etc.
- American Health ways system construct a predictive model using data mining to recognize the patients having high risk. The main concern of this system is to handle the diabetic patients, improve their health quality and also offers cost savings services to the patient. Using Predictive model, healthcare provider recognize the patient which require more concern as compare to other patients [20].
- Data mining play an important role for making effective policy of healthcare in order to improve the health quality as well as reducing the cost for health services.

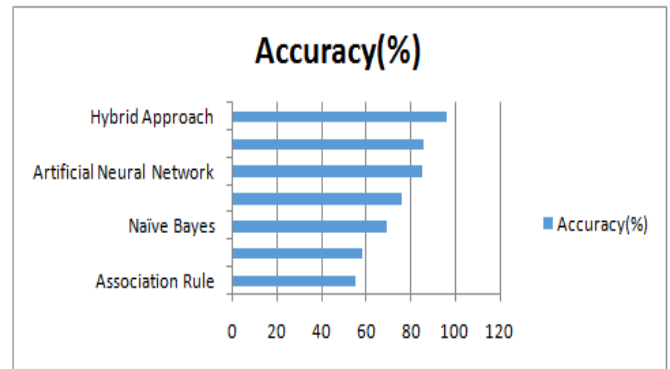
**Performance Study of Data Mining Algorithms**

According to G.Purushothaman and P Krishna Kumari in his survey of data mining techniques on risk prediction for Heart Diseases explained the performance study as shown in below figure.[17]. The below Table 2 shows that the accuracy of algorithms based performance on Heart disease data.

**Table 2** Performance Study of Data Mining Algorithms

S.No	Algorithm Used	Accuracy (%)
1	Association Rule	55
2	K-NN	58
3	Naïve Bayes	69
4	Decision Tree	76
5	Artificial Neural Network	85
6	SVM	86
7	Hybrid Approach	96

The Performance of different algorithms in Data Mining as shown in below chart.



**CONCLUSION**

The prediction of Heart Dieses in earlier is very important aspect in Health care sector. Especially in the heart disease time is precious. In this paper, various Data Mining Techniques are used to predict Heart Diseases. It include that there is no single DMT which gives consistent results for all type of Heart Diseases data. The performance of the DMT depends on the type of data set is used in medical diagnosis. The main idea of this survey focuses on using different DMT on Heart Diseases data.

**References**

1. Aqueel Ahmed and Shaik Abdul "Data mining techniques to find out Heart Diseases: An Overview" *IJITEE*, Vol-1(4) Sep2012.
2. Srinivas, K., Analysis of coronary heart disease and prediction of heart attack in coal mining regions using data mining techniques, *IEEE Transaction on Computer Science and Education (ICCSE)*, p(1344 - 1349), 2010.
3. Yanwei Xing-Combination Data Mining Methods with New Medical Data to Predicting Outcome of Coronary Heart Disease, *IEEE Transactions on Convergence Information Technology*, pp(868 - 872), 21-23 Nov. 2007.
4. K.Sudhakar and Dr.M.Manimekhala "Study of the heart Diseases prediction using data mining"-*IJARCSSE*, Vol-4(1), Jan-2014.
5. Ordonez C. Improving heart deseasesprediction using constraint association rules. Seminar presentation at university Tokyo.2004.
6. Srinivas K, Rani KB, Govrdhan A. Application of data mining techniques in healthcare and prediction of heart attacks. *International Journal on Computer Science and Engineering*. 2011; 2(2):250-5.
7. Parthiban L, Subramanian R. Intelligent heart disease prediction system using CANFIS and genetic algorithm. *International Journal of Biological, Biomedical and Medical Sciences*. 2008; 3(3).
8. G.Subbalakshmi et al., "Decision Support in Heart Disease Prediction System using Naive Bayes", *Indian Journal of Computer Science and Engineering (IJCSE)*.
9. N.DEEPIKA et al., "Association rule for classification of Heart-attack patients", *International Journal of Advanced Engineering Sciences and Technologies*, Vol No. 11, Issue No. 2, 253 - 257.
10. Mai Shouman, Tim Turner and Rob Stocker, "Integrating Decision Tree and K-Means Clustering with Different Initial Centroid Selection Methods in the

- Diagnosis of Heart Disease Patients”, Proceedings of the International Conference on Data Mining, 2012.
11. S.vijaiyarani and S sudha” Diseases Prediction in Data mining Techniques-A Survey”,-IJCAIT-Jan-2013.
  12. Ruben D. Canlas Jr.”Data Mining In Healthcare: Current Applications And Issues”, August 2009.
  13. M. anbarasi, E. Anupriya, N.Ch.S.N.iyengar, “Enhanced Prediction of Heart Disease with Feature Subset Selection using Genetic Algorithm”, *International Journal of Engineering Science and Technology* Vol. 2(10), 2010, 5370- 5376
  14. Jyoti Soni, Ujma Ansari, Dipesh Sharma, Sunita Soni “Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction” *IJCSE* Vol. 3 No. 6 June 2011.
  15. G. Parthiban, A. Rajesh, S.K.Srivatsa “Diagnosis of Heart Disease for Diabetic Patients using Naive Bayes Method”.
  16. Divya Tomar and Sonali Agarwal “ A survey on Data Mining approaches for Healthcare” *IJBSBT-Vol5(5)*.2013.
  17. G.Purushothaman and P Krishna Kumari “survey of data mining techniques on risk prediction for Heart Diseases” *IJST-Vol8(12)* June2015.
  18. K. Srinivas, G. Raghavendra Rao and A. Govardhan, “Survey on prediction of heart morbidity using data mining techniques”, *International Journal of Data Mining & Knowledge Management Process*, Vol. 1, No. 3, pp. 14 - 34, 2011.
  19. K. Usha Rani, “Analysis of heart diseases dataset using neural network approach”, *International Journal of Data Mining and Knowledge Management Process* ,Vol. 1, No. 5, pp. 1 - 8, 2011.
  20. E. P. Ephzibah and V. Sundarapandian, “A neuro fuzzy expert system for heart disease diagnosis”, *An International Journal Computer Science & Engineering*, Vol. 2, No. 1, pp. 17 - 23, 2012.
  21. S. Floyd, “Data Mining Techniques for Prognosis in Pancreatic Cancer”, (2007).
  22. M.-J. Huang, M.-Y. Chen and S.-C. Lee, “Integrating data mining with case-based reasoning for chronic diseases prognosis and diagnosis”, *Expert Systems with Applications*, vol. 32, (2007), pp. 856-867.
  23. S. Gupta, D. Kumar and A. Sharma, “Data Mining Classification Techniques Applied For Breast Cancer Diagnosis And Prognosis”, (2011). [24] K. S. Kavitha, K. V. Ramakrishnan and M. K. Singh, “Modeling and design of evolutionary neural network for heart disease detection”, *IJCSI International Journal of Computer Science Issues*, ISSN (Online): 16940814, vol. 7, no. 5, (2010) September, pp. 272-283.
  24. S. H. Ha and S. H. Joo, “A Hybrid Data Mining Method for the Medical Classification of Chest Pain”, *International Journal of Computer and Information Engineering*, vol. 4, no. 1, (2010), pp. 33-38.
  25. R. Parvathi and S. Palaniammali, “An Improved Medical Diagnosing Technique Using Spatial Association Rules”, *European Journal of Scientific Research* ISSN 1450-216X, vol. 61, no. 1, (2011), pp. 49-59.
  26. S. Chao and F. Wong, “An Incremental Decision Tree Learning Methodology Regarding Attributes in Medical Data Mining”, (2009).
  27. A. Habrard, M. Bernard and F. Jacquenet, “Multi-Relational Data Mining in Medical Databases”, SpringerVerlag, (2003).
  28. S. B. Patil and Y. S. Kumaraswamy, “Intelligent and Effective Heart Attack Prediction System Using Data Mining and Artificial Neural Network”, *European Journal of Scientific Research* ISSN 1450-216X, © EuroJournals Publishing, Inc., vol. 31, no. 4, (2009), pp. 642-656.
  29. Sanjeev Rao and Priyanka Gupta, “Implementing Improved Algorithm Over APRIORI Data Mining Association Rule Algorithm”, *International Journal of Computer Science and Technology (IJCST)*, Volume 3, Issue1, Jan. - March 2012, ISSN : 0976-8491 (Online) | ISSN : 2229-4333 (Print).
  30. S Stilou, P D Bamidis, N Maglaveras, C Pappas, “Mining association rules from clinical databases: an intelligent diagnostic process in healthcare”, *Stud Health Technol Inform* 84: Pt 2. 1399-1403, 2001.
  31. Kaur, H., Wasan, S. K.: “Empirical Study on Applications of Data Mining Techniques in Healthcare”, *Journal of Computer Science* 2(2), 194-200, 2006.
  32. Raut R, Dudul SV. Maximum heart rate resting blood pressure scatter plot for the prominent features abnormal normal design and performance analysis of MLP NN based binary classifier for heart diseases. *Indian Journal of Science and Technology*. 2009 Aug; 2(8):43-8.
  33. Venkatesan P, Yamuna NR. Treatment response classification in randomized clinical trials: a decision tree approach. *Indian Journal of Science and Technology*. 2013 Jan; 6(1):3912-17.
  34. Kumari M, Godara S. Comparative study of data mining classification methods in cardiovascular disease prediction. *International Journal of Computer Science and Technology*. 2011 Jun; 2(2):304-8.
  35. Shen, Z., Clarke, M. Jones, R. Alberti, T. -A neural network approach to the detection of coronary artery disease, *IEEE Transaction on Computers in Cardiology*, 5-8 Sep 1993.
  36. Karaolis, M.A. Moutiris, J.A. Hadjipanayi, D. Pattichis, C.S., -Assessment of the Risk Factors of Coronary Heart Events Based on Data Mining With Decision Trees| *IEEE Transactions on Information Technology in Biomedicine*, pp 559 -566, May 2010.
  37. Frawley and Piatetsky-Shapiro, 1996. Knowledge Discovery in Databases:An Overview. The AAAI/MIT Press, Menlo Park, C.A.
  38. Miller, A., B. Blott and T. Hames, 1992. Review of neural network applications in medical imaging and signal processing. *Med. Biol. Engg. Comp.*, 30: 449-464.
  39. Chen, J., Greiner, R.: Comparing Bayesian Network Classifiers. In Proc. of UAI-99, pp.101-108,1999.
  40. Shaikh Abdul Hannan, V. D. Bhagile R. R. Manza, R. J. Ramteke, -Diagnosis and Medical Prescription of Heart Disease Using Support Vector Machine and Feed forward Back propagation technique, *International Journal on Computer Science and Engineering*, pp 2150-2159, 2010.

41. Huy Nguyen Anh Pham and Evangelos Triantaphyllou "Prediction of Diabetes by Employing a New Data Mining Approach Which Balances Fitting and Generalization" Department of Computer Science, 298 Coates Hall, Louisiana State University, Baton Rouge, LA 70803
42. Ms.S.Sapna, Dr.A.Tamilarasi "Data mining - Fuzzy Neural Genetic Algorithm in predicting diabetes" Department Of Computer Applications (MCA), K.S.R College of Engineering "BOOM 2K8" *Research Journal on Computer Engineering*, March2008.
43. Delen Dursun, Walker Glenn and Kadam Amit, "Predicting breast cancer survivability: a comparison of three data mining methods," *Artificial Intelligence in Medicine*, vol. 34, pp. 113-127, June 2005.
44. Ruben D. Canlas Jr., "Data Mining In Healthcare: Current Applications And Issues", August 2009

**How to cite this article:**

KolluruVenkata Nagendra and Ussenaiah M.2018, A Study on Various Data Mining Techniques Used For Heart Diseases. *Int J Recent Sci Res.* 9(2), pp. 24350-24354. DOI: <http://dx.doi.org/10.24327/ijrsr.2018.0902.1646>

\*\*\*\*\*