International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 06, 2020
ISSN: 1475-7192

HEART DISEASE PREDICTION USING MACHINE LEARNING

[1]A.K.S.C. Bhaskar, [2]Ms. J. Briskilal

ABSTRACT— *There are many deadly diseases present in the huge population of the people in the world, Heart diseases are some of them. When death rates are considered it is known that many people are suffering from hear diseases since these diseases are very dangerous early diagnosis is very important. The major cause of death is Heart diseases according to WHO (World Health Organization). Using a standard method for diagnosis is not good enough fo these diseases. We need to develop a good medical diagnosis model which uses machine learning algorithms and techniques for prediction of diseases an his gives an accurate diagnosis and results than the standard method. By predicting the disease in an earlier stage it reduces the cost of treatment and it also plays an essential role in the treatment. This prediction system's application is to take the input of the patient's data and heartbeat sound recordings and predict the diseases.*

Keywords*—Machine learning, Heart disease, Heart Sounds, machine learning algorithms, Heart disease prediction.*

I. INTRODUCTION

Machine learning is very useful in the health care industry, it analyses large datasets and discovers patterns and use those for predicting diseases. This is one of the better ways to check the heart diseases whenever you feel panicked by some pain or gastric pains in the chest to get clarified that it is not an abnormality in the heart. Our goal is to detect heart disease based on given heartbeat sound and patients history. First, we enter the details of the patient to the predicting system. Then we'll give the heartbeat sound recordings file recorded with an electronic stethoscope or other devices. Now we apply different machine learning algorithms and check which one is getting more accuracy and we use that algorithm for predicting system. Client/patient receives the output which allows him to see whether he/she has a disease or not and if yes what kind of disease it is. This model makes things easier and simpler to find some complicated predictions about heart diseases.

II. LITERATURE SURVEY

We found a huge number of research papers which provided information about methods which are helpful to predict diseases of the heart using patient's details and heartbeat sounds. The summaries of them have been presented in this paper.

[1]*Department of computer science and engineering SRM institute of science and technology, Kattankulathur,kumarbhaskar2009@gmail.com*

[2]*Assistant Professor, Department of computer science and engineering, SRM institute of science and technology, Kattankulathur. briskilal.j@ktr.srmuniv.ac.in*

Received: 22 Feb 2020 | Revised: 13 Mar 2020 | Accepted: 05 Apr 2020

A Tülay Karaylan [1] This paper states about predicting heart disease using neural network this system takes input as disease risk factors and output variable as disease presence or absence. This technique is a classification technique and it is of supervised learning type. An ANN also known as artificial neural network technique is used for prediction with backpropagation algorithm. This is solved using multilayer perceptron neural network with input, output, and hidden layer. Input layer has 13 neurons and hidden layer has 3 neurons (The number will be changed comparing performance), output layer contains 2 neurons since it has 2 classes. The database used is Cleveland database with 303 instances and 14 attributes.

The presence of disease tells on a scale of (0 to 4). Missing values in the dataset are filled with interpolated values and split into training (70%), testing (15%), validation (15%). The results obtained in this paper are good, it is about 95%. This paper studies only ANN for prediction of diseases of the heart but a lot more study can be done using other machine learning algorithms.

A V.V Ramalingam [2] In this paper the author used different algorithms such as SVM (support vector machine), Naïve Bayes, Decision tree and random forest. Random forest gives an out fitted results. Decision tree got an accuracy of

78.4 % which is the least. Out of applied algorithms SVM performed well with 92.1 % accuracy. This paper does not mention any information about what dataset they used. The results obtained using SVM are good. As the results obtained after using random forest are over fitted so we can say this is the drawback.

A Sanjay Kumar Sen [3] In this paper he mentioned few diseases related to heart. he used Weka tool which is open source for all the analytics. They took a dataset from UCI, which is a Machine Learning Repository containing 303 samples and 14 input features and one output feature. They applied algorithms such as Naive Bayes, k – nearest, SVM and decision tree. Out of all these Naïve Bayes performed well with 83.49 % of accuracy. He used various machine learning algorithms. The results obtained after the work are low when compared to the papers.

A research citation by Animesh Hazra, Subrata Kumar Mandal [4] This paper discussed about the symptoms while there are heart diseases such as having a pain in the chest, sweating etc. They different algorithms like k-NN, Decision list, and Naive Bayes algorithms. Out of applied algorithms they got a better accuracy for Naive Bayes and that is 52.3%, all other algorithms did not perform well.

Table 1: Table with the comparison of Accuracy and Time taken for the applied algorithms [4].

Algorithm	Accuracy	Time taken(ms)
Naïve Bayes	52.33%	609
Decision list	52%	719
k-NN	45.67%	1000

A Mohan Mishra, Anushikha Singh, Malay Kishore Dutta [5] This is supervised classification of audio signals. In this samples of audio files were taken and processed in terms of frames. MFCC which is also known

as Mel frequency cepstral coefficients were obtained for each frame of audio sample. They are classified into normal and abnormal cases with feature analysis and supervised classification. Svm classification is done to find the abnormal heart beats. A Wilcoxon hypothesis test is used to calculate the difference between two populations. This is proposed method achieved

97.5 % of accuracy, It obtained good results with 97.5% of accuracy. This paper classifies only between normal and abnormal, it doesn't say what kind of abnormality.

A Peter Bentley, Glenn Nordehn, Miguel Coimbra, Shie Mannor, Rita Getz [6] According to this there are two datasets dataset A and dataset B they are collected from iStethoscope Pro iPhone app and digital stethoscope DigiScope used in hospitals. There are two sounds s1(lub) and s2 (dub). This is to find beat classification into different categories:

1. Normal
2. Murmur
3. Extra Heart Sound
4. Artifact

For e.g.: …lub……. dub……. Lub which is normal

...dub……. lub. **..dub comes under murmur category.

This talks about different kinds of abnormality in the heart beat. This doesn't state information about the results or accuracy what they obtained.

An Anatoly Yakovlev [7] In this paper logistic regression, K means, Gaussian Kernel-Based SVM are used. Logistic regression HSMM (hidden semi-Markov model) to segment time series heart sound into the four heart cycle states. The Dataset consists of 3,541 heart sound recordings in .wav format, each lasting from 5 to over 120 seconds. The data is split into 80% training, 20% testing data. SVM achieved highest performance with an accuracy of 88%. This paper used a good dataset with 3541 heart sound recordings. The amount of accuracy they obtained is less compared to other papers.

A Purushottama, Prof. (Dr.) Kanak Saxenab, Richa Sharma [8] the author of the paper used Data mining techniques for finding the diseases of the heart using thisPrediction System. Knowledge Extraction Evolutionary Learning which is also known as KEEL technique is used, it is a java programming for data mining issues. The authors used some of the algorithms like SVM, C4.5, 1-NN, PART, MLP, RBF and TSEAFS for prediction and acquired an accuracy of 78.53% as highest for RBF and 86.7% after performing 10-fold method.

A Tanvi Sharma, Sahil Verma, Kavita [9] In this paper they used both supervised and unsupervised learning techniques. In supervised they used Support Vector Machine, Naive Bayes, Decision tree and Ann. They used unsupervised learning to find unknown trends since there are no labels. They are further classified into Cluster Analysis and Dimensionality Reduction. They obtained a good accuracy with 99.2% for Naïve Bayes algorithm.

An Amin Ul Haq, Jian Ping Li, Muhammad Hammad Memon [10] In this paper they used Logistic regression, SVM, naïve bayes, Artificial neural network, Decision tree classifier. They used K-fold classification with k = 10. They checked performance of the classifiers using confusion matrix. SVM with

Radial based function kernel performed well with 86% accuracy. ANN and Decision tree resulted with 74% accuracy.

A Jayshril S. Sonawanel, D. R. PatiI [11] Learning vector quantization, it is also known as LVQ algorithm, it is used in this paper and its purpose is it will make prototypes. The respective domain experts can understand this easily. It is created using 2 layers. One is Linear layer and the other is Competitive layer. The prediction system is implemented in MATLAB with random sets of training and testing data. This system got an accuracy of 85.5% for 200 epochs. This paper used a model which is very rarely used and performed well.

A Jayshril S. Sonawane, D. R. Patil [12] In this paper they used multilayer perceptron neural network with input, hidden and output layers. They also used Back- Propagation Network to train multilayer perceptron neural network. They got an accuracy of 98.58% and this is the highest, here hidden layer is made with 20 neurons. They also plotted a graph of accuracy verses number of neurons. They produce a SOM (self-organizing map) map for feature extraction and the training is done with more number of iterations which improved the performance.

A Mrs. S. Radhimeenakshi [13] In this paper they used both Cleveland and Statlog database. When ANN is applied for statlog data they got an accuracy of 89.3% and the performance of SVM in terms of accuracy is 85.06% by their research they SVM performance can be still improved.

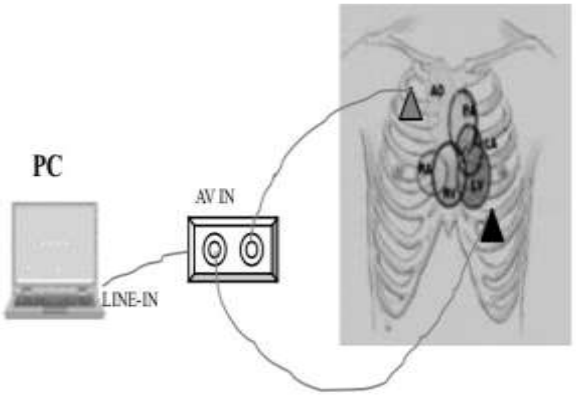

Figure 3: Performance metrics of Cleveland dataset [13]

A Jagdeep Singh, Amit Kamra, Harbhag Singh [14] the objective of this paper is to generate classification association rules (CAR's) for predicting more accurately. The process contains different stages they are Selection, Preprocessing and Transformation, Selection of Associative rules, Performance evaluation and predict disease. Their proposed method is IBk with Aprior Algorithm and for this they acquired an accuracy of 99%. The results for this were calculated using data mining tool, WEKA.

A Kaan Uyara, Ahmet Ilhan [15] In this paper they used a trained RFNN to study the input data, this is assessed by root means square error which is also known as RMSE, probability of the misclassification error, specificity, sensitivity, F-score, precision. They used a powerful machine to perform this model. They got an accuracy of 97.78% for testing set and with ANN-Fuzzy they got an accuracy of 91%. Their suggested future

work is use different data sets for useful prediction and make good data set jointly handled with medical experts.

A S.M. DEBBAL, F. BEREKSI-REGUIG [16] This paper is for prediction of abnormality in the heartbeat sounds. They discussed about the Fourier transform method (FR), it is used to analyze different PCG signals. 8000 samples/sec sample rate used to obtain better reconstitution. Fourier transform is applied to the heart sounds (S1, S2, S3, S4) to analyze the frequency. They did not discuss anything about the results obtained.

A Koredianto Usman, Muhammad Athar Sadiq, Hiroshi Juzoji, and Isao Nakajima [17] a paper on heartbeat sound classification. In this they are not using any sound files to detect the heartbeat abnormality. Independent Component Analysis is explored here for separation of sound signals of heartbeat and a FastICA technique used in this simulation. However, this is helpful only to investigate the heartbeat of the small or newly born babies.

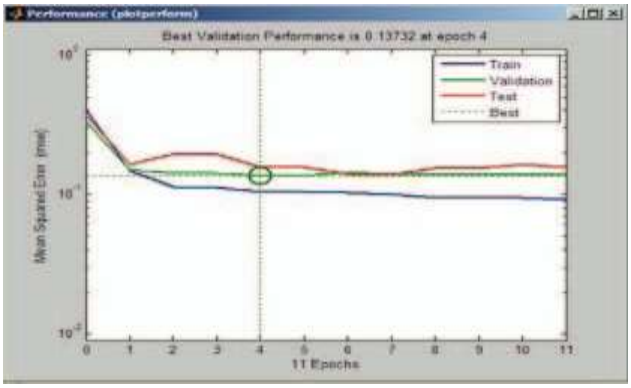

Figure 4: The heartbeat sound recording diagram.

A E. Safar Khorasani, S. Doraisamy, A. Azman, M. Azmi Murad [18] a paper on heartbeat classification. In this paper they used HMMfor segmentation of heartbeat sound. In this they used recorded sound files for predicting the abnormality. The heartbeat sound files are transformed into MIDI format. The Longest common subsequence (LCS) used for identifying matching in the database. Different kinds of heartbeat sounds recognized they are Split Second Sound, Normal Sound, Clicks, Ejection Sound, Gallop Rhythm, Systolic Murmurs, Continues Murmurs, Diastolic Murmurs.

A Amir Mohammad Amiri, Giuliano Armano [19] a paper on heartbeat sound analysis in newly born babies. They used nonparametric spectral estimation which is smoothed through cepstrum thresholding, also bispectrum and Wigner bispectrum were used to analyze the heart murmur. The approach is completely different from other papers and this didn't discuss about the results obtained.

A Simarjot kaur Rhandhawa, and Mandeep Singh [20] A paper on classification of sounds. This paper they used PCG signal to classify sounds, the process contains steps like signal acquisition, feature extraction, feature reduction and then classification. They acquired a highest accuracy of 99.6% with K-NN and 98.8% with Fuzzy K-NN classifier.

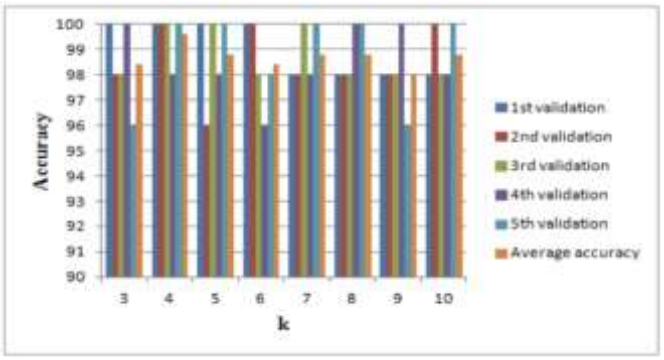

Figure.5 :5-fold validation and average accuracy using K-NN as classifier [20]

III. TABLE

Survey Table

Author	Techniques	Publication	Highest accuracy
Uday Kamylan [1]	Multilayer perceptron neural network	IEEE. (2017)	98%
V V Ramalingam [2]	Support vector machine, Naive Bayes, Decision tree and random forest	Science pub co.(2018)	92.1%
Sanjay Kumar Sen [3]	Naive Bayes, k – nearest, support vector machine, decision tree	International Journal of Engineering and Computer Science (2017)	83.48%
Aniruddh Hazra, Subrata Kumar Mandal, [4]	Naive Bayes, Decision list, k-NN	Research publication India. (2017)	92.33%
Mohan Mishra, Arushiksha Singh, Malay Kishore Dutta. [5]	Svm for Heartbeat sound classification	IEEE. (2017)	97.8%
A Peter Bentley, Glenn Nordehn, Miguel Coimbra [6]	-	-	-
Anatoly Yakovlev [7]	logistic regression, K means, Gaussian Kernel Based SVM	Stanford. (2017)	88%
Purushottama, Prof. (Dr.) Kanak Saxena, Richa Sharma [8]	SVM, C4.5, 1-NN, PART, MLP, RBF and TSEAFS	science direct. (2016)	86.7%
Tanvi Sharma, Sahil Verma, Kavita [9]	Support Vector Machine, Decision tree, Naive Bayes, Artificial neural network, Cluster Analysis and Dimensionality Reduction	IJRA. (2017)	99.2%
Amin Ul Haq, Jian Ping Li, Muhammad Hammad Memon [10]	SVM with Radial based function kernel, ANN and Decision tree	Hindawi. (2018)	86%
Jayshril S. Sonawane, D. R. Patil [11]	Learning vector quantization (LVQ)	IEEE. (2014)	85.5%
Jayshril S. Sonawane, D. R. Patil [12]	multilayer perceptron neural network and produced Self –organizing map	IEEE. (2014)	98.88%
Mrs. S. Radhimeenakshi [13]	ANN, SVM	IEEE. (2016)	89.3%
Jagdeep Singh, Amit Kamra, Harbhag Singh [14]	IBk with Apriori Algorithm	IEEE. (2016)	98%
Kaan Uyanı, Ahmet Ilhan [15]	ANN-Fuzzy approach, GA based trained RFNN	Science Direct (2017)	97.78%
S.M. DEIBIAL, F. BEREKSI-REGUIG [16]	Fourier transform method (FR)	IJIMCMS. (2016)	-
Kowshalaiy Usman, Muhammad Athar Sadiq, Hasnain Jawaj, Isac Nakajima. [17]	Independent Component Analysis	IEEE. (2004)	-
E. Safar Khorasani, S. Doraisamy, A. Azman, M. Azmi Murad [18]	Longest common subsequence (LCS), HMM	ISSN. (2017)	-
Amir Mohammad Amiri, Godiana Armono [19]	cepstrum thresholding, Inspection and Wigner Hipercrum	Science Direct. (2015)	-
Simrandeep Kaur Bhandhawa, and Mandeep Singh [20]	signal acquisition, feature extraction, feature reduction, K-NN, Fuzzy K-NN	Science Direct. (2015)	99.6%

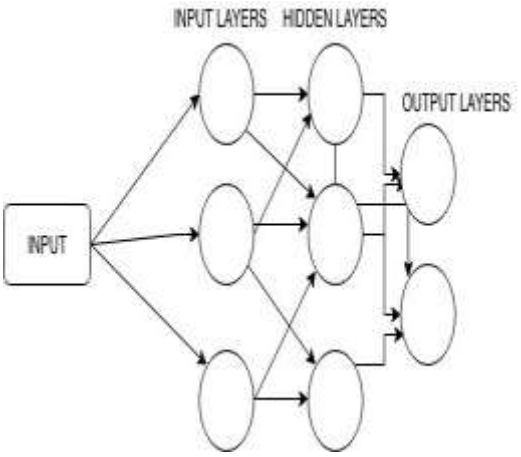

Figure 1: Artificial Neural Network

B. Support Vector Machines (SVM)

SVM is the most useful technique for supervised machine learning, it is also used as both predictor and classifier. It differentates classes with the help of hyperplane. Feature space represents data points which are trained which are seperated by margin . In the same way the test data points will also be mapped in space and then they are classified.

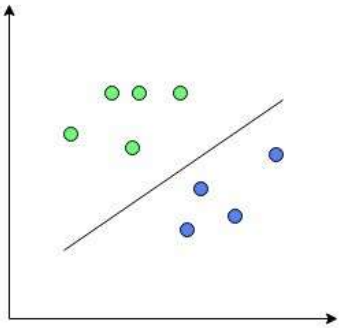

Figure 2: Support Vector Machine

III. ALGORITHMS

Some of the algorithms used for classifying the patient's data and heartbeat sounds to find the disease and to find the abnormality in the heartbeat sound are:

A. Artificial Neural Networks

Artificial neural network is a duplication of human brain and has good processing with interconnected neurons. They are designed by perceptron which solve problems. (MLP) it is the multiple layer perceptron neural network which has input, hidden and output layer.

Input layer: It will be designed with 'n' number of neurons
which are equal to the no. of attributes in the dataset.

Received: 22 Feb 2020 | Revised: 13 Mar 2020 | Accepted: 05 Apr 2020

Hidden layer: It is designed with 'k' number of neurons which is not fixed. It can be increased one by one until we get the best performance.

Output layer: It is designed with m neurons which returns the class label that is the presence or absence of disease.

C. Naive Bayes Algorithm

Naïve Bayes algorithm is a classification technique which is based on the Bayes Theorem and it is very simple. This algorithm uses predictors which are independent, those predictors can be attributes or any other features which should not be correlated to another in anyway. If there are any dependents, then independent features donate to probability that's the reason this algorithm is called naive.

D. K-Means Clustering Algorithm

K-means clustering algorithm is an unsupervised machine learning technique which aims to group similar objects into several cohesive clusters. After the algorithm brings together, new examples can be classified as normal or abnormal by determining which cluster they belong to. The k-means algorithm consists of iterations with more than two steps. [6] First, every example is "colored" assigned to a cluster and it is based on the minimum distance to all cluster centroids. Second, each cluster centroid is updated to the mean value of all the examples in the cluster.

E. Decision Tree Algorithm

A Decision tree [21] is a machine learning algorithm and it uses a tree. It comes under supervised learning, it will create models and will use them to predict. The decision of the obtained from previous data by using some decision rules. The process is first it will take best value and assign it to the root of a tree then it splits them into many sets which are known as subsets. This process will be repeated until they reach the leaf node. Mostly attributes are selected by two methods they are Information gain and Gini index.

F. Random Forest

Random forest is a machine learning algorithm and is supervised. The name itself says that it is made up of a forest. It creates many decision trees and combines them to predict accurately. The best thing about this algorithm is it is useful for both classification as well as regression types. It has both advantages and disadvantages. Its major problem is overfitting and advantage is easy to implement and obtain good results. This algorithm can handle many types of features like numerical etc. This can be implemented by both R and python programming languages by using packages.

IV. TOOLS USED

Some of the tools used for implementing this prediction system are as follows:

A. WEKA

Weka [23] is a tool used for visualization and analyzing data. This tool can be used for free of cost. Since this tool is made using java, it is platform independent so, this tool can be used on mac, windows, Linux etc. We will have an option to install packages which are used for classification, regression and many more. Weka's GUI is very user friendly and easy to use. It bolsters distinctive standard data mining assignments. In our case this tool is very helpful to visualize dataset of heart disease and to find the correlation between the attributes used in the dataset. The input file is of .arff extension.

B. MATLAB

Mathworks developed the Matlab [24]. The programming language used is matrix programming. This is very useful for computation and also visualization. There are many built in commands for performing mathematical calculations. It has different panels such as current folder, workspace, command window, command history. Matlab is an inexpensive software which provides accurate solution to the problem.

C. R-Programming

R is a programming language and it very well may be utilized in R-studio for free without any cost. It is a decent domain for factual processing and illustrations. It is used for data analysis and machine learning algorithms. There is something called packages in R where there are numerous packages available which can be installed in R and can be used for machine learning. CRAN is the name of library where all the packages are available, we install any package for free without any cost. R handles the data in an effective way and allow access to tools for analysis data.

D. Python

Python is a programming language and it very well may be utilized in IDLE or Anaconda Jupyter notebook. It has numerous libraries like scipy, numpy, matplotlib, pandas, sklearn and significantly more valuable for AI execution. In contrast to different dialects, Python is intended to be an effectively intelligible language. Its designing is outwardly uncluttered, and it regularly utilizes English catchphrases. Indeed, even data visualization is significantly more simple with the Python programming language [26].

V. DATASETS

Many datasets are available for the heart disease prediction and for finding heart beat abnormality. Out of all it is found that these are good datasets available for future works and current works.

A dataset called Cleveland dataset is used for the prediction of heart diseases. There are around 76 characteristics in the dataset but only 14 of them are used here. Many machine learning scientists use this dataset even now. The presence of heart disease will be known on a scale of 0 to 4. This is very much useful for recognizing heart infections.

A heart sound recordings dataset is utilized from Physio Net [27] which have five datasets named from A up to E, with an aggregate of 3,541 recorded files of heartbeat in .wav format and they last between 5 secs to 120 secs.

Another dataset [6] observed to be useful for heartbeat variation from the norm and its information has been accumulated from two sources: one from the overall population by means of the I Stethoscope Pro iPhone application, gave in Dataset "A" and the other from a facility preliminary in emergency clinics utilizing the computerized stethoscope called as Digi Scope, gave in Dataset "B".

Dataset A has about 176 .wav records and Dataset B has almost 656 .wav documents.

VI. CONCLUSION

As we all know heart diseases are very dangerous and are complicated, they are leading to many deaths every year. They should be treated in the early stages to reduce the risk of life. Now a day's people are addicted to smoking and consuming fat foods which are very dangerous for the heart. For this purpose, we are proposing a prediction system which will predict the heart disease by taking the patient's data and heartbeat sound recorded file as input. This paper shows all the relevant works done for implementing a prediction system. We can say the main drawback of these works is most of them are focused on implementing classification rather than to focus on cleaning and preprocessing of data. If the data is clean and has no discrepancies, then the algorithms work perfectly and can acquire good accuracies. A good system will be developed in the future that can lead to proper implementation. A lot of work has been done already in predicting patient's data and finding an abnormality of heartbeat separately. I suggest if they both were combined it will be very useful and can produce more accurate results.

REFERENCES

1. A Tülay Karaylan, "Prediction of Heart Disease Using Neural Network," 2nd International Conference on Computer Science and Engineering. (2017)
2. A V.V Ramalingam, "Heart disease prediction using machine learning," from Science pubco. (2018)
3. A Sanjay Kumar Sen, "Predicting and Diagnosing of Heart Disease Using Machine Learning," International Journal of Engineering and Computer Science. (2017)
4. A research citation by Animesh Hazra, Subrata Kumar Mandal from research publications India. (2017)
5. A Mohan Mishra, Anushikha Singh, Malay Kishore Dutta, "Classification of Normal and Abnormal Heat Sounds for Automatic Diagnosis" IEEE. (2017)
6. A Peter Bentley, Glenn Nordehn, Miguel Coimbra, Shie Mannor, RitaGetz. (2017)
7. A Anatoly Yakovlev, "Application of Machine Learning Techniques for Heart Sound Recording Classification", 2017 Stanford, Vincent Lee. (2017)
8. A Purushottama, Prof. (Dr.) Kanak Saxenab, Richa Sharma, "Efficient
9. Heart Disease Prediction System", science direct. (2016)
10. A Tanvi Sharma, Sahil Verma, Kavita, "Intelligent Heart DiseasePrediction System using Machine Learning", IJRA. (2017)
11. A Amin Ul Haq, Jian Ping Li, Muhammad Hammad Memon, "A Hybrid Intelligent System Framework for the Prediction of Heart Disease Using Machine Learning Algorithms", Hindawi. (2018)
12. A Jayshril S. Sonawanel, D. R. PatiI, "Prediction of Heart DiseaseUsing Learning Vector Quantization

Algorithm", IEEE. (2014)

13. A Jayshril S. Sonawane, D. R. Patil, "Prediction of Heart Disease Using
14. Multilayer Perceptron Neural Network", IEEE. (2014)
15. A Mrs. S. Radhimeenakshi, "Classification and Prediction of Heart Disease Risks using Data Mining Techniques of Support Vector Machine and Artifical neural network", IEEE. (2016)
16. A Jagdeep Singh, Amit Kamra, Harbhag Singh, "Prediction of Heart
17. Diseases Using Associative Classification", IEEE. (2016)
18. A Kaan Uyara, Ahmet Ilhan, "Diagnosis of heart disease using genetic algorithm based trained recurrent fuzzy neural networks", Science Direct. (2017) A S.M. DEBBAL, F. BEREKSI-REGUIG, "Frequency analysis of the heartbeat sounds", IJBSCHS. (2014)
19. A Koredianto Usman, Muhammad Athar Sadiq, Hiroshi Juzoji, and Isao Nakajima, "A Study of Heartbeat Sound Separation Using Independent Component Analysis Technique", IEEE. (2004)
20. A E Safar Khorasani, S Doraisamy, A Azman, M Azmi Murad, "An Approach towards a Heartbeat Sound Information Retrieval System", ISSN 0276−6574. (2010)
21. A Amir Mohammad Amiri, Giuliano Armano, "Heart Sound Analysis for Diagnosis of Heart Diseases in Newborns", Science Direct. (2013)
22. A Simarjot kaur Rhandhawa, and Mandeep Singh, "Classification of heart sound signals using Multi – modal features", Science Direct. (2015)
23. A Animesh Hazra, Subrata Kumar Mandal, Amit Gupta, Arkomita Mukherjee and Asmita Mukherjee, "Heart Disease Diagnosis and Prediction Using Machine Learning and Data Mining Techniques: A Review", ISSN 0973-6107 Volume 10, Number 7. (2017
24. https://en.wikipedia.org/wiki/Random_forest.
25. https://en.wikipedia.org/wiki/Weka_(machine_learning)
    https://en.wikipedia.org/wiki/MATLAB.
26. https://en.wikipedia.org/wiki/R_(programming_language)
27. https://machinelearningmastery.com/machine-learning-in-python-step-by-step/"Index of /ml/machine-learning-databases," Index of/ml/machine- learning-databases. [Online]. Available: http://archive.ics.uci.edu/ml/machine- learning-databases/ heart- disease/heart-disease.names. [Accessed: 03-Jul- 2017].
28. G. D. Clifford et al., "Classification of normal/abnormal heart sound recordings: The PhysioNet/Computing in Cardiology Challenge 2016," 2016 Computing in Cardiology Conference (CinC), Vancouver, BC,2016, pp. 609-612. http://ieeexplore.ieee.org/document/7868816/Dataset URL: https://physionet.org/challenge/2016/