

Heart Disease Analysis Using Crossover with Multi-Class Support Vector Machine Method

Nishtha Pahwa¹, Ashima Kalra²

¹M.Tech (Scholar), ²Associate Professor

Department of ECE, Chandigarh Engineering College, Landran, Mohali (Punjab)

Abstract:- In the current era, the human beings are spending their lots of time while sitting and followed the job cycle of minimum 8 hours or more. Therefore, the activeness of humans declined gradually. Science is emerging in every field and gained an extreme attention of researcher and the doctors. The technology changed the way of lifestyle and utilized in the medical field, computer applications, business criteria and so much other. It is assisted for the prediction of various diseases. The existing work is performed by data mining techniques as Naïve Bayes and the Decision Tree (J48). The shortcomings of previous work are expelled in the current research. The focus was at the heart disease, its signs and symptoms. The prediction of heart disease was essential because of the majority of people is surviving through this disease and the early detection is being essential. The proposed techniques are k mean clustering, feature extraction (TCA as textual component analysis) with the help of the multi class SVM classifier. The performance is enhanced in the terms of performance parameters as FAR (False Acceptance Rate), FRR (False Rejection Rate), Accuracy and Specificity and Sensitivity rate.

Keywords: HD(Heart Disease Detection), MC-SVM (Multi-class Support Vector Machine), FRR(False Rejection Rate), FAR(False Acceptance Rate) and CGA (Crossover Genetic Algorithm) and TCA(Texture Component Analysis).

I. INTRODUCTION

Due to the unhealthy diet and work out for a long time caused certain diseases. People do not have much time for the physical activities. The diseases are raised in the entire world. There were various diseases related to heart, liver, kidney, brain and skin diseases. Usually, a disease is an abnormal condition in the body which related to the weaknesses and lack of crucial minerals in the body [1]. Heart disease is one of them that required to be cured in the early stages of the diseases because the number of deaths is climbed up sharply. A heart is a basic and essential component of the human body and its main work was to pump. Its weight is approximately between 250 grams to 350 grams. The size of the heart has been just like a fist. The heart is beating nearly 2.5 billion times in the overall lifespan of 65 to 70 years. The position of the heart is between the lungs and on the left side of the chest.

The heart is made up of special units as known as sino atrial node [2].

The life is fully dependent upon the work of the heart. The blood is circulated in the entire body that pumped by heart organ. The structure of the heart is composed of four chambers, two atria along with two ventricles. The four chambers of the heart are managed in the double walled sac generally known as pericardium. In pericardium, there are three layers as namely outermost, middle myocardium and the inner endocardium. Valves are carried out for the partition of chambers. The other major component of heart is the blood vessels. Generally, blood vessels are also made up of three layers as the endothelium. Smooth lining for the blood flow and the middle layer that is created from soft muscle tissues [3].

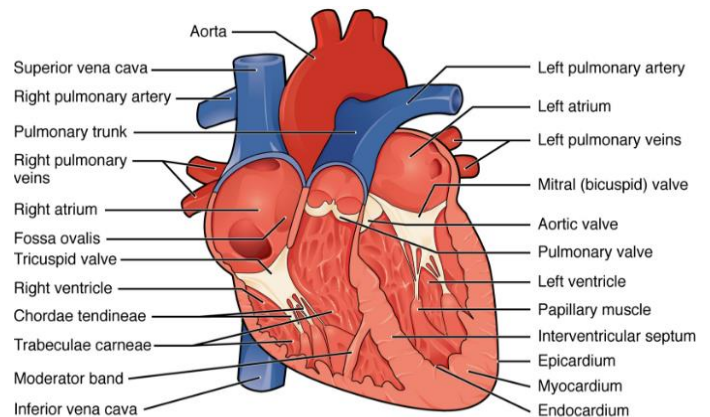


Fig 1. Internal Structure of Heart [3]

In figure 1 the internal parts and structure of the heart is shown that described the each part of the heart with its name.

A. Blood Circulation

In heart the blood is pumped and sent to the other body parts through the two pathways as below-

- Pulmonary Circuit: the deoxygenated blood from the left ventricle through the pulmonary artery and moved to the lungs.
- Systemic Circuit: The filtered blood leaves the heart organ by the aorta from the left side ventricle. The blood is required for heart to fulfill the demands in the body.

The heart contains an electrical conduction system (ECS) for the beating process and moves in a regular rhythm. There are generally two types of the sounds as lube and dub. The sounds are created when the contraction and the relaxation of the chambers are performed in the cardiac cycle of flowing blood.

B. Heart Disease

Heart diseases are the deadly and harmful form of diseases and it is occurring while the processing of blood flow is not normal and there are some abnormal processes under the working criteria of heart. In the whole world, people are surviving from the heart diseases as heart attack. The classification of heart disease is described in the following section.

Classification of Heart Disease

- The angina, heart disease is related to the unmanaged area where the lack of oxygen and not reached to the muscles.
- Arrhythmia occurred when the heart beat is not proper and in irregular processes. The heart beats are varying from slow to fast and not in the moderate form.

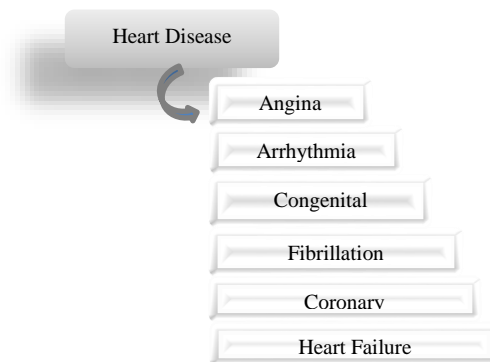


Figure 2. Classification of Heart Diseases [5]

- Fibrillation occurs in the same situation as arrhythmia.
- The congenital heart disease is initialized from the birth time.
- Coronary disease damaged because of the cholesterol and the heart got the shortage of oxygen.

- Heart failure is the major category of heart diseases in which the pumping of blood not processed well and affects the full body [4] [5].

The common causes are described as below [6]. -

Table 1. Causes of Heart Disease

Causes	Description of Causes
Obesity	It is excessive fat in the body and nearly 80 cm for females and 90% for the males and the body mass index become higher than 25.
Blood Pressure	The patients have the range of blood pressure between the 140/70mm.
Swelling	The legs and foot are swelled.
Fatty foods	The patients are moving towards the fatty food which later caused the heart disease.
Fatigue	The feelings are related to stiffness in both head and neck.
Shortness of breath	The shortage of breathing occurred when people walk or run.
Weight	The weight of the body suddenly increased.
Diabetes	The diabetes is the another symptom of patients.
History	Heart disease is based on the family history, it comes in the genes.
Age	The age also varied for the heart diseases [7].

C. Diagnosis and preventions of Heart Disease

The diagnosis of heart disease is a necessity. From a few years ago, the rate of deaths of heart patients inclined drastically and the diagnosis of heart patients is required to be done in the early stages of heart disease. There are some preventions are also given in the paper that helped to less survive and better as compared to the cure. The diagnosis of heart disease is completed by medication, surgeries and other blood tests by electronic machines. The preventions are essential to reduce the survival of heart disease. It is worked by taking a balanced diet, which included low fat and higher fiber food. Exercises are also prevention method do the work out of all body parts, a healthy body weight that is not higher or lower, restrict the habit of smoking and usage of alcoholic drinks [6].

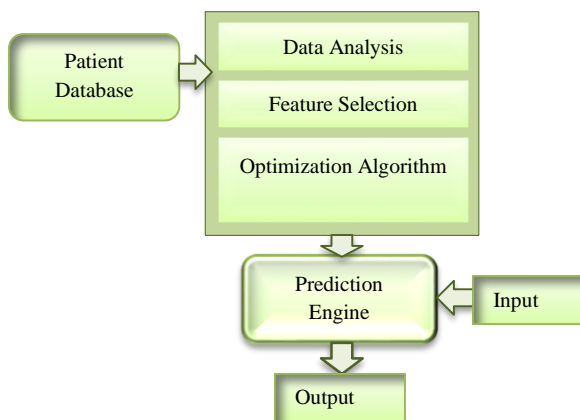
In the previous work, the proposed techniques were based on the intelligent systems and preferred for the detection of diseases by computer vision applications. The techniques were

relying on the data mining methods which composed of other techniques as Naïve Bayes and Decision Tree (J48). The purpose is to create a more precise and efficient classifier for the medical field.

The current research work was partitioned into several sections. These sections are containment of the relevant description of the section headings. Section I included the overview of heart organ and the heart diseases. Section II is considered to describe the previous techniques that utilized for the prediction of diseases. Section III is all about the literature reviews. Section IV and more explained the current research work with results. The last one section considered for the overall summary of research.

II. BACKGROUND

In this section, the detailed description of previous techniques is discussed about the detection of heart disease. The intelligent systems are carried out for the principal work to diagnose the heart related diseases. The structure of heart disease prediction is shown in the following figure 2.



The process of Heart Disease prediction on the gathered database of patients. It included the personal as well as medical information about the patients such as age, blood sugar, gender, cholesterol and heart beat rate. Secondly, the data are analyzed, and the un-relevant data is deducted and performed data integration with the relevant data. After this, the features are selected and prefer a better algorithm for enhancing the performance of the technique. Further, the prediction engine generates the output.

A. Explanation of Previous used Techniques

In the medical field, the technology enhanced the process of detection of diseases. The major techniques that utilized and performed well for the detection of heart related diseases are described as follows-

- a) **GA (Genetic Algorithm):** The concept of genetic algorithm was found by Johan Holland in the 1970s. It is a meta heuristic approach and relied on the natural selection process. The hard problems such as traveler salesman problems are sorted by this algorithm. It has an important place in the medical science to select and search the best solution from the given population of strings (Chromosomes). The process of GA is entirely dependent on its operators as selection, crossover and the mutation. The fitness value is calculated to determine the best value for a particular problem. The solutions are described in the binary form as 0 and 1. The sum of probabilities of each string in a mating pool must be 1 and the probability of selection of string is shown in the following equation.

$$E = \sum_{i=1}^N e_i / N \quad \dots (i)$$

Where the probability e_i is proportional to the fitness value of the string and N is the size population [9].

- b) **Neural Networks:** The neural network (NN) is a kind of artificial neural network that's based on the structure of the layers.

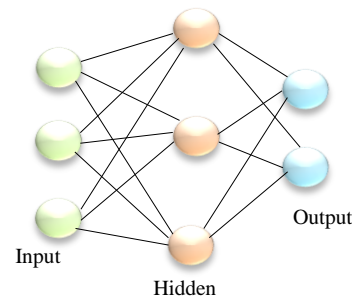


Figure 4. Architecture of Neural Networks [8]

The nodes that gathered the information is neurons and linked with each other through the assignment of weights. It's composed of three layers as input, hidden and the output layer.

In figure 4 three layers are shown. The input layer is associated with the external sources, and the data are proceeding in the hidden layer. The output layer is based on the outcomes of a hidden layer and generates the results. It is supportive of supervised learning, unsupervised learning, that is, the tasks of machine learning.

- c) **Naïve Bayes:** Naive bayes are a classification technique which utilized in statistics and based on the predictions. It basically came from the family of linear classifiers and obtained better results. The determination of class probability for a particular data set is as below-

$$R(G_K/Y) = R(G_K) Y(Y/G_K) / R(Y) \dots (ii)$$

In equation (ii), Y is the instance that required to be classified and G_k belongs to the class and $R(G_k/Y)$ related to the probability of the vector Y to class G_k [10].

- d) **DWT (Discrete Wavelet Transform):** DWT is a transformation approach which has the ability to alter the images from the spatial domain into the frequency domain. It simply partitions the images into different groups with their relative coefficients. The groups are later called as high pass and low pass of DWT [11],
- e) **PSO (Particle Swarm Optimization):** The optimization approach is founded via Kennedy and Eberhart that influenced by the behavior of the bird flocking. The process is initiated by the collection of random particles P in a dimensional space. The particles have their own velocity and the location. The location is determined by below written equations (iii) and (iv).

$$S^j_0 = S_{\text{minimum}} + \mu (S_{\text{minimum}} - S_{\text{maximum}}) \dots (iii)$$

$$W^j_0 = S_{\text{minimum}} + \mu (S_{\text{minimum}} - S_{\text{maximum}}) \dots (iv)$$

In each process the iterative particle is modified with different values and the required to be recorded. The first one of them always best solution and the value of it stored in the memory. W stands for the velocity whereas S considered as the position. For all the results the best one called as global value and denoted by GV_{best} [8].

III. LITERATURE REVIEW

The literature survey is a process to know about the existing work on the topic of research and it gives a detailed description of existing works with their used techniques.

Gomathi, K., et al., (2016) [12] purposed a prediction method for multiple diseases through the data mining approaches. The major work was performed on the heart disease. Generally, the heart diseases caused the deaths. The disease was raised in any age during the lifetime. The factors that caused the heart disease were diabetics, smoking, high cholesterol and high blood pressure. In the current study, the key motive was to predict the heart diseases. The planned approach was data mining. Data mining was a procedure to select, discover and model the extreme amount of data. The demand of data mining was flourished due to the applications in computer vision systems and the medical field also. The purpose in medical areas was to detect and prevent the diseases. It was completed by attaining some medical tests from the patient. Usually, the data mining was categorized into two forms as Naïve Bayes and J48 (Decision Tree). The data mining in naïve bayes were dependent upon the bayes theory. The

implementation was easy and came under the more sophisticated classification approach. On the other side, the decision trees were common methods in data mining. It is simply discovering the relevant data and converted it into a tree format. The nodes of the tree represent the target attributes for the prediction. The results demonstrated that the accuracy, precision and recall inclined continuously. The obtained accuracy by naïve bayes was 79% and by J48 it was recorded at 77.033%. It overcomes the challenge of precise and efficient classifiers. **Rairikar, A., et al., (2017) [13]** described the process of prediction of heart related diseases by using data mining. The medical field acquired the huge amount of data of patients. But it was harder to manage the entire data in a compressed form. It was necessary for the searching criteria of hidden patterns and the decision making purposes. Therefore, data mining was considered and worked well in the medical field. The chief work of data mining was to accomplish the entire data in compressed form by applying several classification techniques. It gives access to various methods as KNN k-nearest neighbors, GA (Genetic Algorithm) and BPNN (Back Propagation Neural Network). To obtain the best results from data mining, the planned technique was utilized with the combination of GA and BPNN specifically to predict heart disease. The motive behind the use of KNN was for the regression and classification due to the characteristics of non-parametric. The research gap was presented due to the shortcomings of majority voting classification. GA and BPNN worked and enhance the performance of prediction process. The rate of parameters such as accuracy and time complexity was better in this work. The accuracy was improved whereas the time complexity was fallen down. Further the description of other methods of data mining was also discussed as decision tree, naïve bayes, knowledge based description. The searching time was expelled more by KNN. On contrast, it was higher in the naïve bayes and decision tree. **Gandhi, M., et al., (2015) [14]** recommended the various methods that were best for the prediction of heart disease. Heart was a crucial part in the human body. Generally, the disease was raised due to several reasons and bad health. In the current era, enormous diseases were obtained such as cancer, brain tumor, liver disease, kidney diseases and so on. Heart disease was one of them which being a cause of more deaths. The work of heart was to pump and without pumping the body could not survive. In the research, several methods were approached for the prediction of heart diseases [17-18]. Data mining was an enhanced system of information science which gained attention in almost each area even private, medical or industrial. The reason of using data mining was to capture the information and managed in a compressed form. The prediction of diseases was mainly based on the methods of data mining. To acquire the better results for prediction, neural networks and decision trees was considered. The neural network was an

artificial network and processed the information in the form of layers. The architecture of NN (Neural Network) was organized in three fundamental layers such as input layer for gaining information, hidden layers for processing and the output layer worked to generate the output based on the processing in hidden layers. But the description of other methods was also depicted in this research. **Thomas, J., et al., (2016) [15]** proposed research for the prediction of heart related diseases. From several years ago, the heart related problem was not even seen in patients. But the chances of heart related diseases were raised rapidly and the diagnosis of this disease was crucial and it was harder. The life of heart patients was at high risk. Each person had different blood pressures and cholesterol levels. But as per the medical proven that the normal range of blood pressure was 120/90 and the pulse was at 72. The current work was a brief survey of different classification techniques for the prediction of heart diseases. The data was acquired on the basis of the patient's age, gender, blood pressure, pulse and cholesterol levels. The risk levels of patients were categorized through data mining as naïve bayes, KNN, neural networks, ANN (Artificial neural networks) and so on. The accuracy was raised of risk levels as it increased. The risk levels were determined through various approaches to predict the level of the heart disease. The process of determining the risk level was initialized by reading the input attributes and then the details of patients were extracted as their age, gender. After acquiring details, the tests were conducted as blood pressure checking, cholesterol and reading of the pulse rate. Further, the patient history was compared with the new reports and the detailed of results were sent for the prediction of risk levels. **Singh, J., et al., (2016) [16]** predicts the heart disease predictions by associative classification. In the entire world, the chances of heart attacks were raised. The prediction of heart diseases was being crucial. The planned approach was associative classification that was better in performance than other methods of classification. Association was defined as the link of various data items in the similar transaction and later, it utilized to search out the hidden patterns. On the other hand, classification was assisted to segment the attributes and managed them in the data which relied on the training set and to determine the measures for different values. The overall rules of classification were based on the prediction rules that means to describe the future situation whereas association was just described the current situation. The planned process worked by describing attributes in terms of gender, chest pain, blood sugar, blood pressure and age range. The experiment demonstrated the accuracy 99.19% which was higher than other methods.

Table 2: Comparison of Literature Survey

Author's name and year	Title of paper	Technique used	Parameters
Gomathi, K., et al., (2016)	Multi disease prediction using data mining techniques	Naïve Bayes Decision tree (J78)	Accuracy Precision Recall.
Rairikar, A., et al., (2017) [2]	Heart disease prediction using data mining techniques	Genetic algorithm KNN BPNN	Accuracy Time complexity
Gandhi, M., et al., (2015) [3]	Predictions in heart disease using techniques of data mining.	Neural networks Naïve Baye's	-
Thomas, J., et al., (2016) [4]	Human heart disease prediction system	Naïve Bayes Decision tree ANN	-
Singh, J., et al., (2016) [5]	Prediction of heart disease using associative classification	Association classification	Accuracy

IV. PROPOSED WORK

In this section, we implement a novel approach to analyze the medical system. We study the various issues in the Heart Disease Prediction System. Heart disease is another leading killer disease in the world. It is also a major cause of disability. Generally, doctors and health professionals use their knowledge to make a judgment for the diagnosis of heart disease in patients. Usually, most of the medical data collected from patients are just saved in files or kept in folders. Generally, those huge amounts of messy medical registers have no meaning for operators. The technique which solves a new problem by remembering a previous case and by reusing information and knowledge of that case, go for those data into useful info that can help to make decision support system for the diagnosis of heart disease. This system can be used to assist doctors and support education for the scholar and postgraduate fresh physicians as a tool to improve the quality of care for the patients. This system can be used as a reference for those students and new doctors. Currently, doctors have problems in determining heart disease in a new patient who does not have existing medical record. Therefore, these data can be used to diagnose heart sickness for new patients who do not have existing medical proceedings. This system is designed to assist doctors and health professionals in determining the diagnosis of patient data. Therefore, this system could assist doctors and health experts to determine the diagnosis and examination of the patient health status.

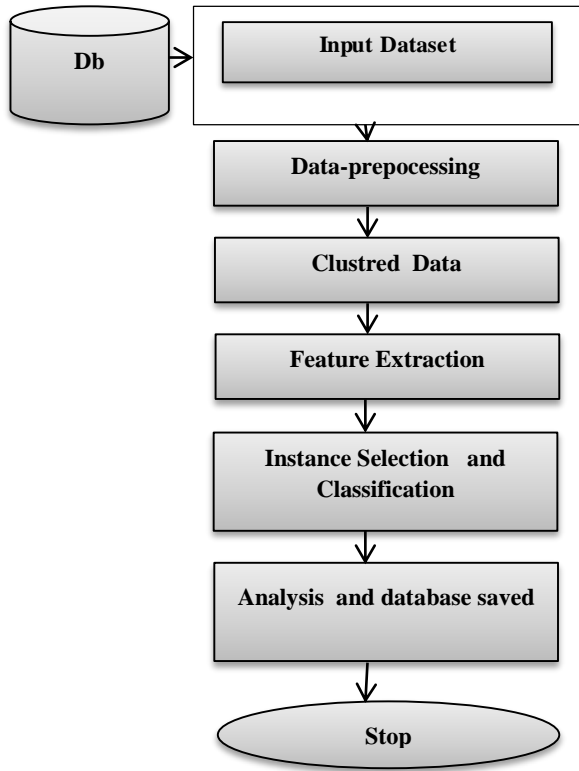


Figure 5. Proposed Flow Chart

The main objectives are described in research work: -

- The study and implement the existing technique of diagnosis of heart diseases using classification methods.
- To propose and implement the K-mean clustering algorithm, feature extraction (Texture Component Analysis), classify the heart disease attributes using a Crossover genetic algorithm and Multi-class SVM algorithm.
- To compare existing and proposed technique using the performance parameters as, FAR, FRR and Accuracy rate.

V. RESULT DISCUSSIONS

In this section, the accuracy rate of five DM (Data Minig) approaches is compared. The main goal is to have a high accuracy rate, False Rejection rate, False rejection Rate, Sensitivity and Specificity metrics.

Although, some performance metrics are utilized more often in the area of IR (Image Retrieval), here we have measured them as they are reviewed to the some other previous metrics like as accuracy. Metris could be derived from the Matrix

Value and can be normally converted to FPR and TPR parameters.

False Acceptance and Rejection Rate

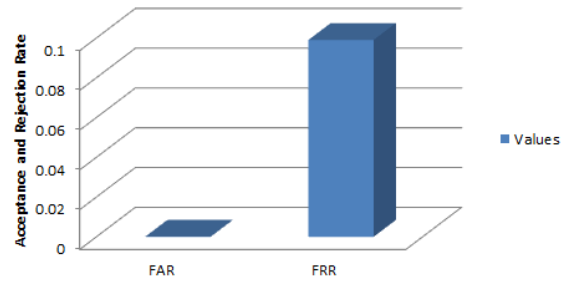


Fig 6. False Acceptance and Rejection Rate (Proposed Work)

The above figure define that the false acceptance rate and false rejected data is reduces in heart disease prediction. It means false data is less acceptable and mitigate the error rates in proposed work.

Table 2. Proposed Parameters

Metrics	Values
FAR	0.0001
FRR	0.099
ACC	90.09
Specificity	0.99
Sensitivity	0.90

Performance

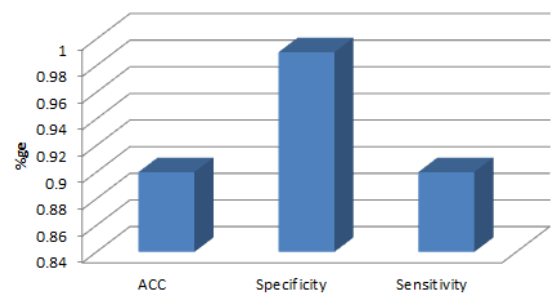


Figure 7. Performance

The above figure show that the performance metrics like as an accuracy, specificity and sensitivity means True Positive rate and True Negative Rate. It enhance the performance metrics using M-SVM and Crossover Optimization method.

The proposed work in MC-SVM algorithm performance metrics FAR (False Acceptance Rate) value is 0.0001, FRR (False Rejection Rate) Value is 0.099, Accuracy value is 90.09, Specificity value is 0.99 and Sensitivity value is 0.90.

Table 3. Comparison

Techniques	Accuracy rate
Naïve Bayes	79 %
J48	77.0330
Multi-class SVM	90.09

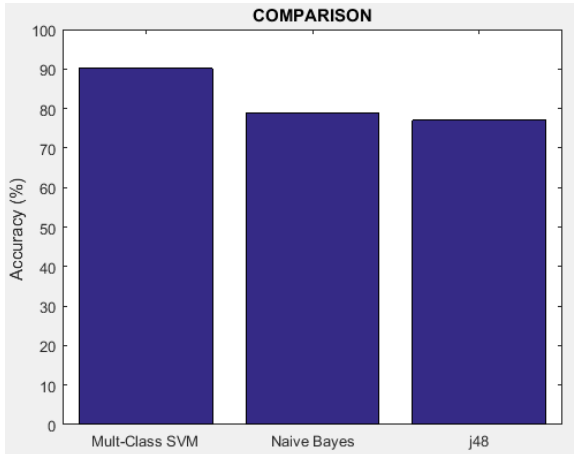


Figure 8. Comparison – Accuracy (%)

Improve the accuracy rate in the proposed work with various techniques like as a Naïve Bayes, j48 and Multi-class SVM.

VI. CONCLUSION AND FUTURE SCOPE

In this study MC-SVM, NB and j48 data mining, classification methods were utilized for the analysis and prediction of several diseases and their performance metrics was compared in order to compute the better classifier. The significant main challenge in DM and ML fields is to construct precise data and evaluation effective classification methods for medical applications. To summarize, the entire work is based on the current emerging disease known as heart disease. The majority of people are surviving from heart attacks. The previous used techniques for the detection and prediction of heart diseases are briefly defined in research work. In the previous work, Naïve Bayes and decision tree (J48) are assisted they performed well, although the requirement of more effectiveness in the process is essential. Hence, the planned approaches are k mean clustering, textual component analysis and multi class SVM that is predicted to be better for the enhancement of prediction of heart disease.

In future, the benefits of new techniques easily enhanced the diagnosis and prediction of diseases at the early stages of the disease. Subsequently, the survival of patients is gradually declining.

VII. REFERENCES

- [1]. Suvarna, C., Sali, A., and Salmani, S. (2017, July). Efficient heart disease prediction system using optimization technique. In *Computing Methodologies and Communication (ICCMC)*, 2017 International Conference on (pp. 374-379). IEEE.
- [2]. Ravish, D. K., Shanthi, K. J., Shenoy, N. R., and Nisargh, S. (2014, November). Heart function monitoring, prediction and prevention of heart attacks: Using artificial neural networks. In *Contemporary Computing and Informatics (IC3I)*, 2014 International Conference on (pp. 1-6). IEEE.
- [3]. Human Heart - Structure, Layers, facts, Functions and Basic Anatomy. (2018). Retrieved from <http://www.sciencestudent.net/human-heart-anatomy-functions/>.
- [4]. Karayilan, T., and Kılıç, Ö. (2017, October). Prediction of heart disease using neural network. In *Computer Science and Engineering (UBMK)*, 2017 International Conference on (pp. 719-723). IEEE.
- [5]. Sowmiya, C., of heart disease diagnosis using classification techniques. In *Intelligent and Sumitra, P.* (2017, March). *Analytical study Techniques in Control, Optimization and Signal Processing (INCOS)*, 2017 IEEE International Conference (pp. 1-5). IEEE.
- [6]. Human Heart - Structure, Layers, facts, Functions and Basic Anatomy. (2018). Retrieved from <http://www.sciencestudent.net/human-heart-anatomy-functions/>.
- [7]. Zennifa, F., Kamil, H., and Iramina, K. (2014, August). Prototype early warning system for heart disease detection using Android application. In *Engineering in Medicine and Biology Society (EMBC)*, 2014 36th Annual International Conference of the IEEE (pp. 3468-3471). IEEE.
- [8]. Bharti, S., and Singh, S. N. (2015, May). Analytical study of heart disease prediction comparing with different algorithms. In *Computing, Communication & Automation (ICCCA)*, 2015 International Conference on (pp. 78-82). IEEE.
- [9]. Malik, A. (2014). Different Genetic Operator Based Analysis and Exploration of TSP. *International Journal of Computer Science and Mobile Computing*, 3(4), 677-687.
- [10]. Jabbar, M. A., Deekshatulu, B. L., and Chandra, P. (2015, March). Computational intelligence technique for early diagnosis of heart disease. In *Engineering and Technology (ICETECH)*, 2015 IEEE International Conference on (pp. 1-6). IEEE.
- [11]. Pugazhenthii, D. (2016, September). Detection of Ischemic Heart Diseases from Medical Images. In *Micro-Electronics and Telecommunication Engineering (ICMETE)*, 2016 International Conference on (pp. 355-360). IEEE.
- [12]. Gomathi, K., and Priyaa, D. S. (2016) Multi Disease Prediction using Data Mining Techniques.
- [13]. Rairikar, A., Kulkarni, V., Sabale, V., Kale, H., and Lamgunde, A. (2017, June). Heart disease prediction using data mining techniques. In *Intelligent Computing and Control (I2C2)*, 2017 International Conference on (pp. 1-8). IEEE.
- [14]. Gandhi, M., and Singh, S. N. (2015, February). Predictions in heart disease using techniques of data mining. In *Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE)*, 2015 International Conference on (pp. 520-525). IEEE.
- [15]. Thomas, J., and Princy, R. T. (2016, March). Human heart disease prediction system using data mining techniques.

In Circuit, Power and Computing Technologies (ICCPCT), 2016 International Conference on (pp. 1-5). IEEE.

- [16]. Singh, J., Kamra, A., and Singh, H. (2016, October). Prediction of heart diseases using associative classification. In *Wireless Networks and Embedded Systems (WECON), 2016 5th International Conference on* (pp. 1-7). IEEE.
- [17]. Kalra, A., Sharma, S., Sharma, S., 2017, Supervised learning technique for prediction of diseases, Springer- ICICCD-2017, Advances in Intelligent Systems and Computing book series (AISC0 on "Intelligent Communication, Control and Devices", vol 624, springer sinagpore., pp 357-369.
- [18]. Kalra, A., Tomar, R., 2017, heart risk prediction system based on supervised ANN, Springer -ICRAC 2017, LNNS springer book series- Lecture Notes in Networks and System, volume 34, chapter 42, pp 405-413 , ,Springer.