# A STUDY OF THE LITERATURE ON CARDIOVASCULAR DISEASE PREDICTION METHODS

**Dr.M.Deepa[1], E.Tamizhan[2], M.P. Venkat Vijay[3], S. Sri Ranjani[4], V. Sowmiya[5]**

[1]Associate Professor,   Dept. of Computer Science and Engineering, Paavai College of Engineering, Pachal, Namakkal-TamilNadu, India.

[2, 3, 4, 5] Student, Dept. of Computer Science and Engineering, Paavai College of Engineering, Pachal, Namakkal-TamilNadu, India.

-----------------------------------------------------------------***---------------------------------------------------------------------

**Abstract -** Among the main causes of death in the modern world is cardiovascular disease. *An important clinical problem is the ability to forecast cardiac disease. In order to forecast cardiac disease, this study discusses several data mining, big data, and machine learning techniques. Designing a crucial model for a healthcare profession to forecast heart illness or cardiovascular disease uses data mining and machine learning. This paper includes an overview of the prior work and offers insight into the current algorithm.*

*Key Words***:**   Data mining, prediction, cardiovascular disease, heart disease, machine learning

## 1. INTRODUCTION

 One of the common diseases that might shorten a person's lifetime nowadays is heart disease. Heart disease claims the lives of 17.5 million individuals worldwide every year [1]. Heart disease symptoms are rising quickly every day, thus it's crucial and worrisome to predict any potential illnesses in advance. This diagnosis is a challenging process that requires accuracy and efficiency. Heart disease comes in many different forms. Heart Failure (HF) and Coronary Artery Disease are the two most comparable forms (CAD). Heart Failure (HF) is mostly brought on by a blockage or narrowing of the coronary arteries.

The signs of heart illness, such as high blood pressure, chest discomfort, hypertension, cardiac arrest, etc., can be used to diagnose the condition. Birth defects, high blood pressure, diabetes, smoking, narcotics, and alcohol are all causes of cardiovascular diseases. Oftentimes, infections that damage the inner mitochondrial of the heart can also cause symptoms including fever, tiredness, a dry cough, and skin rashes. There are now too many automated tools, such as data mining, machine learning, deep learning, etc., to identify cardiac disease. Therefore, we shall give a basic overview of machine learning approaches in this work. Using machine learning resources, we train the datasets in this. There are certain risk variables that are used to make predictions about heart disease. Age, sex, blood pressure, cholesterol level, diabetes, family medical history of coronary disease, smoking, alcohol use, being overweight, heart rate, and chest pain are risk factors.

The structure of this article is as follows. Section 2 provides a comprehensive evaluation of the body of previous research. Conclusion and future work are presented in Section 3.

## 2. LITERATURE REVIEW

In healthcare institutions, a lot of progress has been made on illness prediction systems utilizing various data mining, machine learning, and algorithmic approaches.

Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques is a concept put out by Senthil Kumar Mohan et al. (2019) with the aim of enhancing the accuracy of cardiovascular disease prediction by identifying essential components by utilizing machine learning. With many combinations of highlights and a few well-known arranging techniques, the expectation model is produced. We develop a prediction model for heart disease with a precision level of 88.7% using a hybrid random forest with a linear model (HRFLM). They also received training on a variety of data mining techniques and expectation methods, including KNN, LR, SVM, NN, and Vote, which have recently gained considerable notoriety for their ability to distinguish and predict heart disease. [2].

Using data mining techniques, Mamatha Alex P and Shaicy P Shaji (2019) created "Prediction and Diagnosis of Heart Disease Patients." KNN, Random Forest, Support Vector Machine, and Artificial Neural Network methods are used in this article. Artificial Neural Networks have a greater accuracy rating for identifying heart disease in data mining when compared to the previously described categorization techniques [3].

Bo Jin, Chao Che, and colleagues (2018) suggested a neural network-based model for "Predicting the Risk of Heart Failure With EHR Sequential Data Modeling." This study conducted an attempt to foretell congestive heart disease using electronic health record (EHR) data from real-world datasets connected to the condition. To represent the diagnostic events and predicted coronary failure events using the fundamental tenets of an extended memory network model, we typically utilize one-hot cryptography

and word vectors. The significance of maintaining the sequential character of clinical data is generally made clear by our analysis of the results [4].

Prediction and Analysis the Occurrence of Heart Disease Using Data Mining Techniques was advised by Chala Beyene et al. (2018). The major goal is to foresee the development of cardiac disease in order to quickly and automatically diagnose the condition. The suggested technique is crucial in healthcare organizations with professionals that lack current knowledge and expertise. To determine if a person has heart disease or not, many medical characteristics are used, such as blood sugar and heart rate, age, and sex. WEKA software is used to compute dataset analyses. [5]

For the purpose of predicting cardiac illness, R. Sharmila et al. (2018) suggested using a non-linear classification system. For the prediction of heart disease using an optimal attribute set, bigdata techniques like Hadoop Distributed File System (HDFS), MapReduce, and SVM are suggested. This study investigated the application of several data mining approaches for the prognosis of cardiac disorders. It advises utilising HDFS to store vast amounts of data across several nodes and running the SVM-based prediction algorithm across multiple nodes at once. When SVM is utilised in parallel, calculation times are faster than when SVM is used sequentially. [6]

Heart disease prediction using data mining and machine learning algorithms was proposed by Jayami Patel et al. in 2017. The purpose of this project is to use data mining methodologies to reveal the underlying tendencies. In comparison to LMT, the optimal technique, J48, has the highest accuracy rate. [7]

Heart disease prediction using an ANN technique in data gathering was proposed by P. Sai Chandrasekhar Reddy et al. (2017). There was a need to develop a new approach that can anticipate heart illness due to the rising costs of diagnosing heart disease. After evaluating the patient based on numerous factors like metabolism, blood volume, cholesterol, etc., a forecast model is utilized to forecast the patient's symptoms. Java is used to demonstrate the system's correctness. [8]

To analyses and forecast various cardiac diseases, S. Prabhavathi et al. (2016) introduced the Decision tree based Neural Fuzzy System (DNFS) approach. The studies on diagnosing heart disease is reviewed in this essay. Decision tree-based Neuron Fuzzy System is known as DNFS. The goal of this research is to develop a system that is both clever and economical while also enhancing the functionality of the current system. Data mining techniques are specifically applied in this work to improve heart disease prediction. This study's findings demonstrate that SVM and neural networks do remarkably well in predicting coronary heart disease. These data mining methods are still not promising for predicting heart disease. [9]

K-means and naive bayes were employed by Sairabi H. Mujawar et al. (2016) to predict cardiac disease. With the use of a historical cardiac database that provides diagnoses, this study will create a system. 13 characteristics were taken into account when developing the system. Data mining techniques like clustering and classification algorithms may be used to extract knowledge from databases. 300 entries total with 13 characteristics were taken from the Cleveland Heart Database. Based on the values of 13 characteristics, this model is intended to determine whether or not the patient has heart disease. [10]

An examination of metabolic syndrome was suggested by Sharan Monica.L et al. (2015). In order to anticipate the sickness, this research suggested data mining approaches. It is intended to offer an overview of the most recent methods for extracting data from datasets, which will be helpful to healthcare professionals. The decision tree for the system's performance can be built in a certain amount of time. The main goal is to forecast the illness with the fewest possible attributes. [11]

The effectiveness of the diagnosis of cardiac disease using various approaches is described in Table 1 hereunder.

**Table -1:** A comparison of several approaches for reviewed literature

| YEAR | AUTHOR | PAPER | METHODS | RESULT |
|---|---|---|---|---|
| 2019 | Senthil Kumar et al, [2] | Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques | HRFLM | 88.7% |
| 2019 | Mamatha et al, [3] | Prediction and Diagnosis of Heart Disease Patients | KNN, SVM, ANN | 79.4% |
| 2018 | Bo Jin et al, [4] | Predicting the Risk of Heart Failure With EHR Sequential Data Modeling | EHR | 82.6% |
| 2018 | Chala Beyene et al [5] | Prediction and Analysis the Occurrence of Heart Disease Using Data Mining Techniques | J48, Naïve Bayes, Support Vector Machine | WEKA Tool |
| 2018 | R. Sharmila et al [6] | Non-linear Classification System | HDFS, Map Reduce, SVM | 84.669% |
| 2017 | Jayami Patel et al [7] | Heart disease prediction using data mining and machine learning algorithms | J48, LMT | 91.38% |
| 2017 | Sai Chandrasekhar Reddy et al [8] | Heart disease prediction using an ANN technique | ANN | 78% |

## 3. CONCLUSIONS

To forecast the occurrence of heart disease, many data mining and machine learning algorithms have been compiled. Find out how well each algorithm performs predictions, then implement the suggested system where it is needed. Use more precise attribute selection techniques to increase the precision for programs. In the event that a patient is diagnosed with a certain type of heart disease, there are several treatment options available. Such an appropriate dataset can yield considerable information via data mining.

The notion of many strategies that have been researched for detecting cardiovascular illness is presented in this paper through a systematic review of the literature. Big data, machine learning, and data mining may be used to great success to analyses the prediction model with the highest level of accuracy. The primary goal of this study is to diagnose cardiovascular disease or heart disease utilizing a variety of techniques and procedures to obtain a prognosis.

Even after further analysis, only sporadic accuracy is attained for more complicated forecasting model for heart disease Models are required to improve the precision of forecasting the early-stage heart disease In the future, we'll suggest a technique. For the highly accurate early diagnosis of heart disease and lowest complexity and expense.

## REFERENCES

[1]  M. Preethi, Dr. J. Selvakumar ,"A Literature Survey of Predicting Heart Disease ", International Research Journal of Engineering and Technology, Volume: 07 ,Issue: 05 , May 2020.

[2]  Senthilkumar Mohan, Chandrasegar Thirumalai, Gautam Srivastava,Effective Heart Disease Prediction Using Hybrid Machine Learning,Techniques||, Digital Object Identifier,10.1109/ACCESS.2019.2923707, IEEE Access, VOLUME 7, 2019.

[3]  Mamatha Alex P and Shaicy P Shaji,"Prediction and Diagnosis of Heart Disease Patients using Data Mining Technique", International Conference on Communication and Signal Processing 2019.

[4]  Bo Jin ,Chao Che, Zhen Liu, Shulong Zhang, Xiaomeng Yin, And Xiaopeng Wei, "Predicting the Risk of Heart Failure With EHR Sequential Data Modeling" ,IEEE Access 2018.

[5]  Mr. Chala Beyene, Prof. Pooja Kamat, "Survey on Prediction and Analysis the Occurrence of Heart Disease Using Data Mining Techniques", International Journal of Pure and Applied Mathematics, 2018.

[6]  R. Sharmila, S. Chellammal, "A conceptual method to enhance the prediction of heart diseases using the data techniques", International Journal of Computer Science and Engineering, May 2018.

[7]  Jayami Patel, Prof. Tejal Upadhay, Dr. Samir Patel, "Heart disease Prediction using Machine Learning and Data mining Technique", March 2017.

[8]  Mr.P.Sai Chandrasekhar Reddy, Mr.Puneet Palagi, S.Jaya, "Heart Disease Prediction using ANN Algorithm in Data Mining", International Journal of Computer Science and Mobile Computing, April 2017, pp.168-172.

[9]  S.Prabhavathi, D.M.Chitra, "Analysis and Prediction of Various Heart Diseases using DNFS Techniques", International Journal of Innovations in Scientific and Engineering Research, vol.2, 1, January 2016, pp.1-7.

[10] Sharan Monica.L, Sathees Kumar.B, "Analysis of CardioVasular Disease Prediction using Data Mining Techniques", International Journal of Modern Computer Science, vol.4, 1 February 2016, pp.55-58.

[11] Sairabi H.Mujawar, P.R.Devale, "Prediction of Heart Disease using Modified K-means and by using Naïve Bayes", International Journal of Innovative research in Computer and Communication Engineering, vol.3, October 2015, pp.10265-10273.

[12] Animesh Hazra, Arkomita Mukherjee, Amit Gupta, Asmita Mukherjee, "Heart Disease Diagnosis and Prediction Using Machine Learning and Data Mining Techniques: A Review", Research Gate Publications, July 2017, pp.2137-2159.

[13] V. Krishnaiah, G. Narsimha, N. Subhash Chandra, "Heart Disease Prediction System using Data Mining Techniques and Intelligent Fuzzy Approach: A Review", International Journal of Computer Applications, February 2016.

## BIOGRAPHIES

Dr.M.Deepa received her Ph.D Degree in Computer Science from Periyar University, Salem, in the year 2021. She is presently working as an Associate Professor in Department of Computer Science and Engineering, Paavai College of Engineering, Pachal, Namakkal(DT),TamilNadu, India.

Tamizhan.E. Currently pursuing 2nd year CSE in Paavai college of Engineering, Namakkal.

Venkat Vijay. MP. Currently pursuing 2nd year CSE in Paavai college of Engineering, Namakkal

Sri Ranjani.S. Currently pursuing my 2nd year CSE in Paavai college of Engineering, Namakkal

Sowmiya.V. Currently pursuing my 2nd year CSE in Paavai college of Engineering, Namakkal.