IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

# Enhanced Disease Prediction in Cardiovascular Diseases using Machine Learning Techniques

Sharwari Patil[1], Tejaswini Tekale[2], Shreya Sinha[3], Usha Prasad[4], Mrs. Smita Khot[5]

[1]Student, [2]Student, [3]Student,[4]Student,[5]Professor

[1]Computer Engineering,

Dr. D. Y. Patil Institute of Technology, Savitribai Phule Pune University, Pimpri, Pune, India

*Abstract:* Myocardial infarction forecast is treated as most dumbfounded task in the field of medical sciences. There emanates a need to build up a network for identifying heart problems of a patient. This system suggest effective calculation in conjunction with machine learning approach for heart disease analysis. Today medical field have made substantial progress to heal patients with different sort of medical disorders. There are various medical clinics using clinical data to deal with medical services. This data produces colossal measures of information. Appallingly, this information is whimsically used to help the medical progression. There is a greater part of concealed data in this information that isn't yet scrutinized that would lead to significant probe of how to make valuable data out of the information. Henceforward it's required to make an incredible undertaking which will benefit experts with anticipating the heart issues before it happens. The fundamental objective of this paper is to build up a model which is capable of establishing and extracting enigmatical information related with heart problems from a past heart information base record.

***Keywords*** **--** Machine learning, heart disease prediction, prediction Model, classification algorithms, cardiovascular disease.

## I. INTRODUCTION

Acute myocardial infarction is the most widely recognized reason for abrupt passing in urban and country regions. Ideal recognition of a coronary failure is vital on the grounds that deferred expectations can make serious harm the heart muscle, called cardiomyopathy, which can prompt bleakness and passing. At the point when cell phones become one of the most broadly utilized advances today, building up an application for anticipating a coronary episode will deliver productive outcomes in diagnosing somebody who has endured a chest pain. This will lead to an early prediction of a heart attack, which will lead to doctors' early diagnosis and treatment as well as early treatment. Chest pain is the most common and obvious symptom of a heart attack, although some other characteristics can easily cause a heart attack. In this era, modern medicine has enriched many modern technologies and biological equipment, greatly reducing the overall mortality rate. But Cardiovascular disease (CVD) or Ischemic heart disease, Stroke, Chronic obstructive pulmonary disease, Lower respiratory infections, Diabetes mellitus are precariously fatal. Predicting a heart attack in a timely manner is critical, because delayed detection can cause severe damage to the heart muscle, called cardiomyopathy, which can lead to morbidity and death. Acute myocardial infarction occurs when the coronary artery is suddenly and completely blocked, supplying blood to the heart area (also called "heart attack"). It may be due to the accumulation of plaque, which is mainly made of fat, cholesterol and cellular waste. As the blood supply starts to drop, after a time it drops far below normal level and sometimes even stops, and this leads to weakening of heart muscle and eventually they began to die. If not treated early, this damage may be irreparable. The medical department has a wealth of information, but the main problem of medical data mining is their quantity and complexity, poor mathematical classification and standardized form. The proposed system discovered insights from medical data sets using modern data mining techniques. A big issue is reducing the time between having a heart attack and getting care. Individuals who are busy with daily work at home or in the office and rural people who are ignorant of heart attack symptoms may ignore chest discomfort. They may not have a clear intention to ignore it, but they may waste time and decide to see a doctor or be hospitalized after a period of time. However, in the case of a heart attack, the most crucial factor is time. Consumers can use many mobile health tools to prevent CVD, such as self-monitoring mobile applications. The use of a large number of mobile devices, such as mobile phones for communication and feedback, as well as smartphone applications, is supported by current science. Because the medical diagnosis of heart disease is very important, but the task is complex and costly, the system proposes a medical diagnosis system to improve the quality of medical treatment and reduce costs. The goal of the system is to provide a ubiquitous service that is both feasible and sustainable, and enable people to assess their risk of heart attack at that point in time or later. Acute myocardial infarction, often introduced as Heart Attack is the most cause for sudden deaths in city and village areas. It is one the most dangerous disease among men and women and early identification and treatment is the best available option for the people. Identifying the cardiovascular diseases in antecedent stages. To predict the onset of high-risk cardiovascular diseases for patients with heart attack, hypertension or hyperlipidemia with better accuracy.

## II. RELATED WORK

M. A. Jabbar et.al, proposed another method of applying affiliation management procedures in the medical field to find heart disease predictions. Medical data is collected in vast quantities by the human services business. Unfortunately, these services have not been tapped to find successful enveloped data. Choose hidden examples and find that hidden examples and relationship softening have not been fully utilized. Information mining strategies can help solve this problem. Data mining has discovered various applications in business and science. Affiliation rules, arrangements, and aggregation are important areas of passion for information mining [1]. Ms.M.C.S.Geetha et.al, analyzed the commonly used classification algorithms in the medical data collection that aids in the prediction of cardiac illnesses, the leading cause of mortality worldwide. Professionals are required by doctors to forecast heart attacks based on experience and expertise, which is a difficult task. Today's healthcare field contains secret but meaningful information to make decisions. The experiments carried out reveal this algorithm. J48, SIMPLECART, and REPTREE have higher prediction precision than other algorithms, as predicted [2]. M. Akhil Jabbar et.al, pointed out that the nearest neighbor (KNN) is a basic, well-known, proficient and powerful design confirmation method. KNN is a direct classifier, where the arrangement of parts depends on the category of its nearest neighbors. The clinical information base is substantial in nature. If the informational set contains too high and irrelevant attributes, grouping may produce less accurate results. Coronary disease is a key contributor to India's rising death rate. Coronary heart disease is the leading cause of mortality in Andhra Pradesh, accounting for 32% of all fatalities, comparable to Canada (35%), and the United States (35%). Following that, a selection of emotional support network must be described. This choice will lead the clinician to make great strides. In this work, another strategy is proposed that combines KNN with genetic programs to achieve a strong order. Genetic strategies perform global queries in complex huge and multi-mode scenarios and provide ideal arrangements [3].

Chaitrali S Dangare used increasing information quality to investigate the expected framework of cardiac disease. This work uses clinical terms such as gender, circulatory system strain, cholesterol and 13 credits to predict the likelihood of patients suffering from heart disease. Until recently, 13 traits have been used for expectations. The inspection work also includes two additional functions, such as robustness and smoking The estimation of the data mining arrangement was checked in the coronary illness database, including specific decision trees, naive Bayes and neural networks [4]. A high-precision hybrid approach for identifying coronary artery disease was suggested by Zeinab Arabasadi et al. In fact, the suggested strategy may improve the performance of a neural network by almost tenfold by increasing its initial weights with a genetic algorithm [5] . Sahar H. El-Khafifand Mohamed A. El-Brawany introduced that the ECG signal is famous for its non-linear changing behavior and is the key trademark used in this inspection. Under normal and abnormal conditions, the non-linear part of its elements changes more naturally than straight conditions. Due to the high measurement (HOS) required to maintain the phase data, this work utilizes the one-dimensional shear from the terrible areas that are more demanding for typical and ischemic subjects. The feed forward multi-layer neural system (NN) has a Negligent Backlash (BP) learning method and is used as a computerized ECG classifier to discover the chance of discovering ischemic coronary heart disease from common ECG signals [6].

Senthil Kumar Mohan et.al, proposed a novel method which focuses on discovering important features by applying ML techniques to improve the precision of cardiovascular disease prediction model. This prediction model presents different combinations of features and various popular classification techniques [7].I. S. Siva Rao, T. Srinivasa Rao predicted that heart disease is the most widespread driving force for humans to kick buckets. Constantly, 7.4 million people are associated to heart disease of which 52% die from stroke and 47% die from coronary heart disease. Subsequent determination of evidence of various heart diseases at the basic stage is of great significance for the safety of cardiovascular-related diseases. Current conventional ECG examination strategies (such as RR span, wavelet mutation, and group calculation) (for example, Support Vector Machine(SVM), K-nearest neighbors(KNN) algorithm, and Levenberg-Marquardt neural networks) are used to find cardiovascular arrhythmias. Even after extracting a large number of highlights, the problem identification will not be accurate using these programs [8].Amma, N.G.B proposed that medical diagnostic systems undertake important work in clinical practice and are used by clinical professionals for analysis and treatment. In this work, the clinical estimation system is portrayed by the normal danger of cardiovascular infection. This structure works by consolidating genetic programs with the general preferences of the sensory system. The multi-layer feed forward neural system has been specially designed for complex layout problems. The load of the nervous system can be solved by genetic methods because it can find an acceptable load arrangement in fewer cycles [9]. Saba Bashir et.al, describes the use of data science in prediction of the heart disease in the medical field. Since a large number of studies have conducted research on this problem, the accuracy of prediction still needs to be improved. Therefore, this study focuses on feature selection techniques and algorithms, in which multiple data sets on heart disease are used for experimental analysis and show higher accuracy in Cardiovascular disease prediction [10].

## III. PROPOSED SYSTEM

This system proposes a novel Heart disease prediction mechanism as well as Hypertension and Hyperlipidemia prediction model that first learns deep features and then trains these learned features. This proposed system uses dataset available on kaggle. After verifying the dataset, the next step will be data preprocessing. The results show that when using all attributes and the same training sample for training, the classifier is superior to all other classifiers. Since the samples are not enough to learn the precise mapping between features and category labels, using low population, high-dimensional data sets to predict heart attacks is challenging. Current literature usually accomplishes this task by manually creating and selecting features. Compared with other technologies, it is found that random forest can identify the basic structure of the data. Random Forest is used for both classification and Regression analysis. It consists of many decision tree. Multiple decision trees are constructed and the output will be the mean prediction of multiple decision trees.

## IV. METHODOLOGY

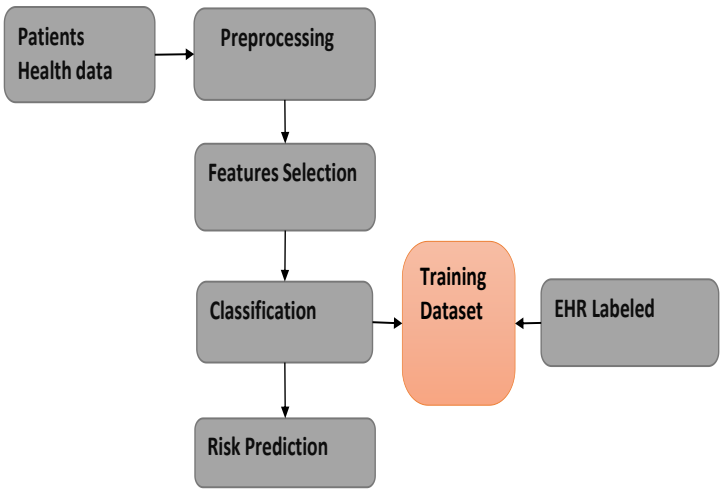


Figure 1. Proposed System Architecture

Python is a computer programming language often used to build websites and software, and conduct data analysis. The project is on python as it is easy to learn and use and it has an extensive library. In this paper, the dataset available on Kaggle have been used. The dataset covers various aspects like age, cholesterol, blood pressure, etc. which will help in accurate prediction of cardiovascular diseases. The data is collected after which preprocessing is done i.e taking cleaned and organized raw data for training models, then the relevant features were selected for feature selection step and then classification model is applied on it. After training the dataset with help of EHR labelled data the result is generated and the risk is predicted.

## V. RESULTS AND DISCUSSION

Cardiovascular Disease dataset is taken and analysed to predict the asperity of the disease. Random Forest approach is used to predict the disease it also has better accuracy than other machine learning techniques. The data is pre-processed to make it acceptable for classification. This system will help the doctors for early prediction of cardiovascular diseases and risk factors related to heart disease such as hypertension and hyperlipidemia. Early detection of high risk factors the person can be referred to the cardiologists for further testing.

.

## VI. CONCLUSION AND FUTURE SCOPE

In this work a unique methodology has been introduced for grouping coronary illness. As an approach to approve the proposed strategy, it includes the patient heart testing result subtleties to foresee the sort of coronary illness utilizing Machine Learning. Train informational collections taken from UCI Repository. This methodology utilizes different Machine learning techniques like Naive Bayes algorithm and Random Forest which are considered as aggressive techniques for classification. This model can help medical specialist perform an effective heart disease diagnosis process with fewer attributes. Cardiovascular diseases are the most widely recognized as a contributor of mortality rate in major parts of India. Distinguishing proof of significant hazard factors and creating choice emotionally supportive network, and successful control measures and wellbeing instruction projects will decrease in the coronary illness mortality. The working model can also help in decreasing the treatment costs by catering Initial diagnostics in time. General physicians can take advantages of this technology for initial diagnosis of patients.

## VII. REFERENCES

[1] M. A. Jabbar, B.L Deekshatulu, Priti Chandra, "An evolutionary algorithm for heart disease prediction"CCIS,PP 378-389 , Springer(2012).

[2] Ms.M.C.S.Geetha, Dr.I.Elizabeth Shanthi, Ms.N. Sanfia Sehnaz," Ana-lyzing the Suitability Of Relevant Classification Techniques On Medical Data Set For Better Prediction" International conference on I-SMAC 2017.

[3] M.Akhil jabbarB.L Deekshatulua Priti Chandra International"Classification of Heart Disease Using K- Nearest Neighbor and Genetic Algorithm"Conference on Computational Intelligence: Modeling Techniques and Applications (CIMTA) 2013.

[4] Chaitrali S Dangare "Improved Study Of Heart Disease Prediction System Using Data Mining Classification Techniques", International Journal Of Computer Applications, Vol.47, No.10 (June 2012).

[5] Zeinab Arabasadi , Roohallah Alizadehsani , Mohamad Roshanzamir , Hossein Moosaei , Ali Asghar Yarifard " Computer aided decision making for heart disease detection using hybrid neural network-Genetic algorithm" Science Direct 2017.

[6] Sahar H. El-Khafifand Mohamed A. El-Brawany, "Artificial Neural Network-Based Automated ECG Signal Classifier", 29 May 2013.

[7] Senthilkumar Mohan, Chandrasegar Thirumalai, And Gautam Srivastava" Effective Heart Disease Prediction Us-ing Hybrid Machine Learning Techniques"IEEE ACCESS 2019

[8] I. S. Siva Rao, T. Srinivasa Rao, "Performance Identification of Different Heart Diseases Based On Neural Network Classification".ISSN 0973-4562 Volume 11, Number 6 (2016) pp 3859-3864.

[9] Amma, N.G.B "Cardio Vascular Disease Prediction System using Genetic Algorithm", IEEE International Conference on Computing, Communication and Applications, 2012.

[10] Saba Bashir, Zain Sikander Khan, Farhan Hassan Khan, Aitzaz Anjum, Khurram Bashir," Improving Heart Disease Prediction Using Feature Selection Approaches", International Bhurban Conference on Applied Sciences Technology 2019.