



# AN AUTOMATED DEEP LEARNING SYSTEM TO DETECT HIGH RISK CARDIOVASCULAR DISEASE USING MACHINE LEARNING

S.JENIFER BLESSY,  
PG SCHOLAR,  
DEPARTMENT OF CSE,  
VELAMMAL ENGINEERING COLLEGE.

DR.S.RAJALAKSHMI,  
ASSISTANT PROFESSOR,  
DEPARTMENT OF CSE,  
VELAMMAL ENGINEERING COLLEGE.

**ABSTRACT**-Heart attack is one of the leading causes of the death all around the world. Predicting the occurrence of disease at initial stages is a major challenge at the moment. Machine learning when applied in health care is capable of premature and accurate recognition of disease. Coronary artery calcium score (CACS) is a reliable predictor for forthcoming cardiovascular disease risk. CACS is visible in all the computed tomography (CT) scans of the chest. This information is not computed routinely as it requires time, expertise and also some specialized equipment. This paper comprises of a deep learning approach which is robust and time-efficient and it automatically computes coronary calcium on routine cardiac-gated and non-gated CT. This paper aims to develop a deep learning system which automatically identifies individuals who are at high risk for cardiovascular disease and to test the system's performance with a variety of clinical presentations and CT scanning techniques.

## I. INTRODUCTION

Cardiovascular or heart diseases are the primary reason of death and a lot of disability globally. Consequently understanding the etiology roots, distribution and trends is necessary to improve public health in all countries. Heart disease is one of the prevalent syndromes that can lead to reduction of lifespan of human beings nowadays. Each year 17.5 million people are dying because of heart diseases. Heart is one of the body's most significant organs. Basically heart is a organ that is made up of four chambers separated by valves and divided into two halves. Every half comprises one chamber

called an atrium and one called a ventricle. The oxygen produced by them provides your body with energy and is essential to keep your body healthy.

The universal term used to shield malfunctions of the heart is called cardiovascular diseases, Cardiac Disease. Developing countries, which had low rates previously are now having increased rates as economies develop, infectious diseases are dominated and life expectancy improves. Cardiovascular disease is a category of disease that affects the heart or blood vessels. The risk of these diseases may be increased by smoking, high blood pressure, high cholesterol, unhealthy diet, lack of exercise, and obesity. The utmost common cardiovascular disease is coronary artery disease (narrow or blocked coronary arteries), which can lead to chest pain, heart attacks, or stroke. Some of the other heart diseases comprises of heart rhythm problems, congenital heart disease ,congestive heart failure and endocarditis. Machine learning when applied in health care is capable of premature and accurate recognition of disease.

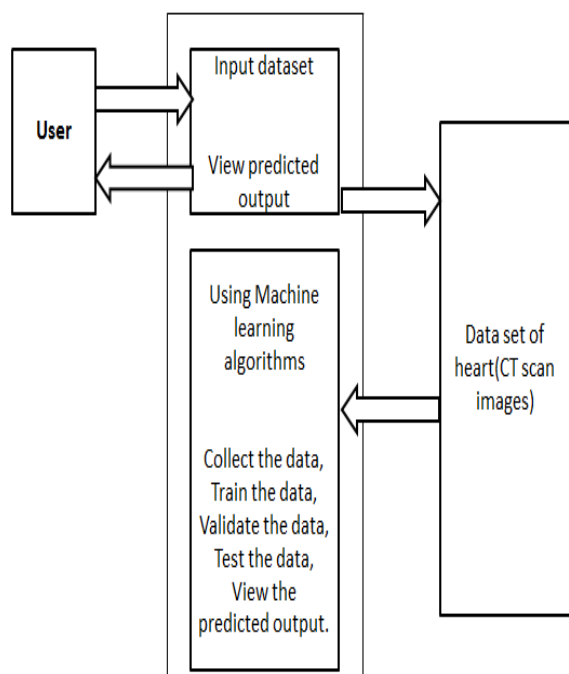
### Types of heart diseases:

- a. **Cardiac arrest** : This is a sudden loss of heart function, breathing and consciousness that occur without any warning.
- b. **Congestive heart failure** : The heart does not pump blood as good as it should, it is the condition of chronic.
- c. **Congenital heart disease** : It is one of the birth defects that affects the normal way of how the heart works.

- d. **Coronary artery disease** : It is a disease where the arteries are not able to deliver oxygen rich blood or any disease occurs in the blood vessels.
- e. **High Blood Pressure** : It is a situation in which the blood force against the walls of the artery in heart is too high.
- f. **Peripheral artery disease** :The narrowed blood vessels which decrease flow of blood in the limbs, is the circulatory condition.
- g. **Stroke** :Supply of blood is intervened occur damage to the brain.

### Machine learning:

Machine learning is a type of Artificial intelligence in which the system learns automatically from its experience and by data used .Machine learning when applied in health care is capable of premature and accurate recognition of disease. Here the user sends the input dataset to the system and the system collects the data, trains the data, validates it and tests it. Finally the predicted output of whether the person has heart disease or not is viewed to the user as shown in Fig 1,



**Fig 1: ML system to detect heart disease.**

## II. LITERATURE SURVEY

Senthilkumar Mohan et al, [1] introduced neural networks using heart rate time series and used various clinical records for prediction to find out the exact condition of the patient in relation to heart disease. Computer Aided Decision Support System and also conducted experiments used to identify the features of a machine learning algorithm with a hybrid

method. The dataset which has a radial basis function network (RBFN) is used, where 70% of the data is used for training and the remaining 30% is used for classification. The results show that the hybrid method predicts heart disease efficiently. Here the features are gathered from the UCI Dataset ,random forest algorithm and linear methods are performed for prediction.

Hybrid method predicts heart disease efficiently so it can be used for prediction of cardiovascular diseases. It does not deal with real-world datasets instead it deals with just theoretical approaches and simulations. There can be some feature selection methods that can be developed to get a broader perception of the significant features to increase the performance of heart disease prediction

K. Polaraju et al,[2] suggested the Heart Disease prediction by means of Multiple Regression Model and it shows that this model is suitable for predicting chances of heart disease. The data set is divided into two parts and then the 70% of the data is used for training and 30% is used for testing. From the results, the accuracy of Regression algorithm is better than other algorithms. Multiple Linear Regression Analysis is performed on trained data to build a model on which test data is applied. Multiple Linear Regression is more appropriate in predicting the chances of heart disease so this method can be included for predicting chances of heart diseases. This is a statistical method that is used to define data and explains the association in-between one dependent variable and two or more independent variables, but it does not achieves the accuracy of 99.2% by Decision Tree,.

Roman Zeleznik et al ,[3] suggested a deep learning system to automatically classify individuals at high risk for cardiovascular disease and tested the system's performance in four big independent held-out cohorts with a variation of medical presentations and CT scanning methods. the system automatically identifies and segments the coronary calcium and computes the CAC scores, and stratifies them into clinically relevant categories of very low (CAC = 0), low (CAC = 1–100), moderate (CAC = 101–300), and high (CAC > 300).In the deep learning approach the cohort is trained and tuned, heart localization ,heart segmentation ,calcium segmentation and coronary artery calcium score are carried out and finally the test cohorts are done.

(CACS) is a reliable predictor for that easily predicts future cardiovascular disease risk, it is visible on all computed tomography (CT) scans. So this paper provides information regarding this. Most disagreements were due to misclassification of non-coronary calcium as coronary calcium and vice versa. In a few occurrences, inaccurate heart segmentation led to coronary calcium being outside the calcium segmentation network and hence being missed.

M.A.Jabber et al,[4] claims that the Hidden Naïve Bayes (HNB) can be applied to heart disease classification (prediction), the experimental results on heart disease data set illustrate that the HNB accounts 100% in terms of accuracy and out performs naïve bayes. The input is the heart disease dataset and the output is classification based on whether the person is healthy or having some heart diseases.

Finally as a result HNB classifier is a favorable model for medical data sets like heart disease with dependent attributes for diagnosis of disease. If your test data set has a categorical variable of a category that was not present in the training data set, the Naive Bayes model will assign it probability value of zero and won't be able to make any predictions in this

Mohammed Nowshad et al ,[5] qualified the system by means of classification algorithms including Naive Bayes, Decision Tree, Support Vector Machine (SVM), Logistic Regression, K-Nearest Neighbors (KNN), etc. However the accuracy for different algorithms varies for different examples. In the dataset, SVM gave the best performance with an accuracy level of 91% for the threshold instance of the dataset. In ML techniques to collect data based on questionnaire, then reduce redundant, noisy data and create input dataset, train the dataset. Contains dividing (70 %) and test (30%) sets up and analyzes the performance of the classifier.

This paper describes our attempt to do so by indent of Real-world data collected from hospitals and healthcare industries helps in analyzing real-world data. Supervised Machine Learning's regression analysis can be used to determine vulnerabilities Instead the person's level toward heart disease in percentage There is either a heart patient or not using the classification only to classify.

M.Snehith Raja et al, [6] developed a system that requires the user to enter all the symptoms he is suffering from, on the basis of which the outcome is predicted. The training of the neural

network is done using back propagation to evaluate the prediction system. Inputs are collected, trained, validated, tested, and then visualized outputs using the Random Forest algorithm.

It is possible to obtain a high accuracy rate using the Random Forest algorithm which will be useful in predicting whether he has any heart disease. Practical use of data collected from past records is time consuming and low accuracy rate. So to overcome this Random Forest Algorithm is applied to get accurate result in less time.

Aditi Gayhane et al ,[7] suggests a neural network algorithm Multi-layer Perceptron (MLP) ,one s to train and to other to test the dataset which is a supervised neural network algorithm that has one layer for input, another for output and hidden One or more layers between these two layers. Every node of the input layer is connected to the output nodes through hidden layers. The output of the system gives a prediction result if the person has heart disease, in terms of yes or no and gives an idea about the condition of the heart causing coronary artery disease.

It gives an idea about multi-layer assumption to train and then test the dataset which gives prediction results which leads to CAD. There is a challenge in achieving greater accuracy and reliability. Big data technologies such as Hadoop can be used to store and manage large chunks of data from all users around the world or user report.

Anjan et al, [8] focuses on cost-cutting through data mining techniques and creating effective approaches to enhance DSS (Decision Support System).. It uses the Naives Bayesian - Data Mining Classification Technique to effectively enable heart disease diagnosis and thereby offer appropriate treatment. First the user registration is done through a mobile application, then by training and test datasets classification is done using Navy Bayesian and finally prediction is done and an AES to transfer the data to the database in a secure manner has been designed.

This yields a higher security performance evaluation than AES PHEA (Parallel Homomorphic Encryption Algorithm). Application developers must work closely with health care professionals and researchers to deliver disease apps that Improving health care outcomes. a holistic of the research process to Reducing delays will help ensure that application-based heart disease prevention

research is not entirely lagging behind advances in technology.

Mamta Rani et al, [9] suggested Principal Component Analysis (PCA) to reduce the features. In addition to a hybrid genetic algorithm (HGA) with foot-mean for final clustering, which may be trapped in local optima as this method is heuristic and a hybrid genetic algorithm (HGA) for clustering of data is used to avoid this problem. used which can predict early heart disease with an accuracy of 94.06%. PCA is applied to reduce the characteristics, the population is initialized and evaluated and randomly two populations are selected and crossover and mutation is done, now it definitely improves the fitness and evaluates new offspring. Final quality of the clustering is upgraded.

Accuracy is high and clustering quality is improved. Data standardization is essential before PCA which is a drawback of this paper and also leads to loss of information.

Md.Touhidul Islam et al, [10] evaluates heart disease prediction using datasets from the UCI Machine Learning Repository useful for mankind and smart healthcare system and Used both principal component analysis and CHI square feature selection techniques to reduce Number of variables to be analyzed to predict heart disease Using traditional machine learning algorithms like KNN, SVM, Naive Bayes, and Random Forest Classifier etc. This work has achieved the highest classification accuracy of 92.85% using random forest classifier and PCA..

Disease prediction classifier like this work will play a Active role in early diagnosis of heart disease which will be a great achievement in the field of medical science. Accuracy depends on the quality of the data. With big data, the forecasting phase can be slow. Sensitive to scale and irrelevant features of the data and also requires high memory.

Farzani Tasnim et al ,[11]compared the performance of a neural network, a support vector machine, a system with genetically constructed fuzzy rules, a classification and regression tree and its direct evolution that is a random forest, in analyzing our database. Both HF severity assessment and HF type prediction tasks are obtained using the Random Forest algorithm. Here a management tool lets cardiologists populate a "supervised database" suitable for machine learning during routine

outpatient consultations. In the train phase the samples and desired outputs are provided and the training and modeling is done and obtained in the use phase by the trained intelligence outputs.

Monitoring scenarios are facilitated by automatically providing readable output even by non-cardiologist physicians and nurses regarding the severity and type of HF. However, the CART achieved slightly lower performance than the RF but had the advantage of providing a sensible model. CART provides cross-validation multiclass accuracy of 81.8% in severity assessment and 87.6% in type prediction. Unfortunately these findings are difficult to generalize due to the small sample size..

G.Guidi et al ,[12] presented detailed information Discussed about existing heart disease prediction systems and Cleveland datasets, machine learning algorithms, prior heart disease prediction models, accessories, considerable metrics and research challenges, etc. Prior authors have explored the proposed disease prediction models by emphasizing their advantages and limitations. . The most reliable machine learning application development platforms such as Weka, Rapid Miner, Mahout and MATLAB are also covered briefly.

Gives information about all the existing heart disease prediction systems that needs to be known, to get through the project. While processing Heart disease data for predictors, effects on characteristics The results must be calculated individually to avoid the involvement of low impact characteristics in processing to save To increase timing and accuracy.

Dr.Lakshmi Prasad Koyi et al ,[13] proposed, a novel machine learning approach to predicting heart disease that uses the Cleveland Heart Disease dataset, and data mining techniques such as regression and classification is used. machine learning technology random Forests and Decision Trees are implemented and in the implementation, 3 Machine Learning Algorithms are used, 1. Random Forest, 2.Decision Trees and 3. Hybrid Models. The results of the experiments show that accuracy level of 88.7% is achieved through a heart disease prediction model using hybrid model. This interface is modeled to get the user's input parameters and then predict heart disease, using hybrid model Decision Trees also and Random Forests

This hybrid model Gives us better-optimized results based on both the machine learning algorithms that are considered for implementation.88% accuracy. deep learning for the heart Disease prediction gives better results. Also, it can be classified as a multi-class problem to identify the stage of the disease.

Rashmi G Saboji et al ,[14] extends the prediction of the classifier, genetic search is integrated; The genetic search resulted in 13 characteristics that contribute significantly to the prediction of heart disease. Classifier ensemble algorithms such as random forest are used to predict patients with heart disease and have demonstrated that random forest data mining techniques outperform other data mining techniques such as Naive Bayes and Decision Trees.. Random Forest performs consistently first and after reduction of characteristics with the same model build time.

It runs resourcefully on bulky data bases as well large data sets and it would handle nearly thousands of input variables without deletion of variables. The data filtering mechanism is also used. The impact of a massively supervised dataset performance and accuracy, running at high performance .The cluster is missing.

### III. CONCLUSION

In conclusion as recognized through literature surveys, there is confidence that only a modest success is achieved in building predictive models for patients with heart disease and therefore to increase the accuracy of prediction of premature onset of cardiovascular disease. Combination and more difficult models are required. In this age of machinery and population, early and accurate prediction of any disease is very important. An appropriate ML model can not only predict a disease quickly, but also predict it with acceptable level of accuracy, advance treatment, reduction of human intervention and medical laboratory tests. As the mortality rate for heart disease is increasing beyond imagination, creating proper machine learning models for heart disease prediction is one of the imperative requirements for the present time. So it is more likely that this paper will provide the result with more accuracy.

### REFERENCES

- [1] Senthilkumar Mohan; Chandrasegar Thirumalai; Gautam Srivastava,” Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques”, 19 June 2019, IEEE Access ( Volume: Page(s): 81542 – 81554.
- [2] K.Polaraju, D.Durga Prasad,” Prediction of Heart Disease using Multiple Linear Regression Model” 2017 IJEDR ,Volume 5, Issue 4 ,ISSN: 2321-9939
- [3] Roman Zeleznik , Borek Foldyna , Parastou Eslami , Jakob Weiss , Ivanov Alexander , Jana Taron , Chintan Parmar , Raza M. Alvi2 , Dahlia Banerji , Mio Uno , Yasuka Kikuchi , Julia Karady2 , Lili Zhang , Jan-Erik Scholtz , Thomas Mayrhofer , Asya Lyass , Taylor F. Mahoney , Joseph M. Massaro , Ramachandran S. Vasam , Pamela S. Douglas , Udo Hoffmann , Michael T. Lu & Hugo J. W. L. Aerts , “Deep convolutional neural networks to predict cardiovascular risk from computed tomography “ 29 January 2021 ,Article number: 715 .
- [4] M.A.Jabbar ,Shirina samreen ,” Heart disease prediction system based on hidden naïve bayes classifier “ ,02 October 2017 .
- [5] Mohammed Nowshad Ruhani Chowdhury, Ezaz Ahmed , Md. Abu Dayan Siddik , Akhlak Uz Zaman,” Heart Disease Prognosis Using Machine Learning Classification Techniques”, 2021 6th International Conference on Convergence in Technology (I2CT) Pune, India.
- [6] M.Snehith Raja, M.Anurag, Ch.Prachetan Reddy, NageswaraRao Sirisala,” machine learning based heart disease prediction system”, 2021 International Conference on Computer Communication and Informatics (ICCCI 2021), Jan 27-29,2021, Coimbatore, INDIA,
- [7] Aditi Gavhane, Isha Pandya, Gouthami Kokkula, Prof. Kailas Devadkar (PhD),” Prediction of Heart Disease Using Machine Learning” 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA),
- [8] Anjan, Nikhil Repaka, Sai Deepak, Ravikanti ,Ramy G Franklin, “Design And Implementing Heart Disease Prediction Using Naives Bayesian, 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI) .
- [9] Mamta Rani; Aditya Bakshi; Akhil Gupta ,” Prediction of Heart Disease Using Naïve bayes and Image Processing” ,2020 International

Conference on Emerging Smart Computing and Informatics (ESCI).

[10] Md. Touhidul Islam, Sanjida Reza Rafa, Md. Golam Kibria ,” Early Prediction of Heart Disease Using PCA and Hybrid Genetic Algorithm with  $\epsilon$ -Means, 2020 23rd International Conference on Computer and Information Technology (ICCIT).

[11] Farzana Tasnim, Sultana Umme Habiba,” A Comparative Study on Heart Disease Prediction Using Data Mining Techniques and Feature Selection”, 2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST).

[12] G.Guidi, M. C. Pettenati, P. Melillo, Member, IEEE, E. Iadanza, Member, IEEE,” A Machine Learning System To Improve Heart Failure Patients Assistance”, IEEE Journal of Biomedical and Health Informatics ( Volume: 18, Issue: 6, Nov. 2014), 10 July 2014.

[13] Dr. Lakshmi Prasad Koyi ,Tejaswi.Borra ,Dr. G. Lakshmi Vara Prasad,” A Research Survey on State of the art Heart Disease Prediction Systems”, 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS).

[14] Rashmi G Saboji , Prem Kumar Ramesh,” A Scalable Solution for Heart Disease Prediction using Classification Mining Technique”, 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS).

