International Conference on Intelligent Technologies & Science - 2021
(ICITS-2021)

A Review on Prediction of Cardiac Arrest using Machine Learning

Vicky Singh[1], Brijesh Pandey[2]
Computer Science and Engineering,
Goel Institute of Technology & Management, Lucknow, India
vicky.singh10a@gmail.com

Abstract- Cardiac Arrest which is commonly known as Heart Attack, is increasing all over the world. The modern lifestyle is acting as a catalyst and making the matter worst. The heart is the most vital organ of the body which is responsible for pumping oxygenated blood to all parts of the body. A heart attack is a condition when the heart is not able to pump an adequate amount of blood to the body. This condition arises when the heart does not get the required amount of blood supply due to the blocked arteries. In recent times huge data had been collected of patients with heart disease. The stored data can be used for predicting heart disease with significant accuracy. Few machine learning algorithms that can be used for predicting heart diseases, such as Decision Tree, k-Nearest Neighbour, Random Forest, Artificial Neural Network(ANN), and Ada Boost. This paper gives a summary of the existing works done in the field of prediction of heart disease using machine learning.

Keywords-heart attack, machine learning, k-nearest neighbour, random forest, artificial neural networks, naïve bayes, decision tree

1. Introduction

The heart is responsible for pumping blood to all parts of the body. A heart attack is a situation when the heart is not able to pump blood throughout the body. Sometimes Heart Attack is also referred to as Cardiac Arrest and Heart Stroke. A heart attack can lead to the death of the person. According to World Health Organisation (WHO), heart disease is the number cause of death in the world. Despite being a preventable disease, heart disease accounts for 31% of the total death which is approximately 17.9 million.

According to the Indian Heart Association, 25% of all heart attacks happen below 40 years of age, and 50% of heart attacks happen below the age 50 years in India. Indian rural population is comparatively less vulnerable to the Indian population[1]. The deposition of *plaque* narrows the artery. The plaque contains mainly fat and cholesterol. When plaque ruptures and forms a clot it blocks the flow of the blood and results into heart attack[2]. Fig 1.1 clearly depicts normal coronary artery and narrowed coronary artery.

The condition of heart disease does not develop all of a sudden but it is the result of having a particular lifestyle for long period. The most common habits and that make the person prone to heart disease are cholesterol-rich food like meat, and animal product, food containing unhealthy fats, high blood pressure, obesity, smoking, lack of physical activity. Modern-day culture and work are proving as an aid to heart disease. Despite being a preventable disease, it is killing more number of people in the world than any other disease.

There are two main sources of cholesterol in the body. One source of cholesterol in the body is the liver. The liver produces cholesterol to generate new cells in the body and it is also required to produce vital hormones like testosterone in the body. The other source of cholesterol in the body is animal products like meat. If the cholesterol level is high in the body then it is highly recommended to cut down on the intake of animal products.

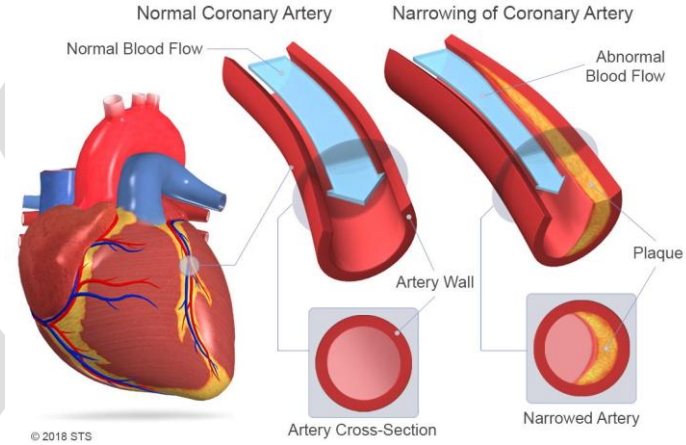

Fig 1.1: Normal Artery and Narrowed Artery [6]

High triglycerides also make the person prone to heart disease. Excess calories convert to triglycerides and are present in the blood in the form of fat cells. In order to avoid high triglycerides levels, we should avoid high calory intake at once. We should consume food that releases energy slowly so that there are never excess calories in the blood. It is highly recommended to avoid fast food that releases energy at once which can lead to an excess of energy in the blood at once.

Maintaining normal blood pressure for a heart patient is very crucial. Blood pressure puts extra pressure on the wall of the artery which could result in rupturing of the arteries that can eventually convert into heart attack.

2. Literature Review

In recent years heart disease has become a big concern as the number of cases is increasing every passing year. Early detection of heart disease can save many lives. Various research work had already been done for the prediction of heart disease using machine learning algorithms.

Hamidreza and Morteza[3] have proposed a hybrid algorithm that is based on traditional machine learning algorithms like Neural Networks and Naïve Bayes. The hybrid algorithm gives the highest accuracy of 89.1%. The

dataset used for training and testing the model is from UCI library. It contains 303 records. Although each record has 76 attributes, only 14 attributes were considered for all the experiments. The paper proposed an ensemble method that would be a combination of Neural Networks and Naïve Bayes. It has also proposed to use a weighted majority vote strategy and with the use of this strategy, the output of the ensemble classifier would be more influence by the classifier which has a higher accuracy level.

Martin Gjoreski et al[4] have proposed to use heart sound obtained using a digital stethoscope for detecting chronic heart disease using machine learning algorithms. This method achieved 96% accuracy and outperformed most of the existing classifiers. In the experiment, 152 sounds were obtained using a digital stethoscope. For obtaining these sounds 122 subjects were considered. The dataset goes through various stages like filtering, segmentation, etc.

Aditi Gavhane et al[5] have proposed to build an application that will take basic things regarding a person like age, blood pressure, sex, etc., and will predict whether the person is prone to heart disease or not. According to this research paper, Neural Network is the most accurate and reliable for building the application which predicts chronic heart disease. Cleveland dataset from the UCI library had been used for training and testing the model. Neural Networks with multi-layer perceptron had been used to train and test the dataset.

Cincy Raju et al[6] have made use of various machine learning classification algorithms to predict heart disease. Decision Tree produced an accuracy of 82.35%. SVM (Support Vector Machine) produced an accuracy of 99.3%. Neural Networks gave an accuracy of 91.1%. KNN(k-Nearest Neighbour) produced an accuracy of 87.2%. It found that SVM (Support Vector Machine) is giving the best results with 99.3% accuracy.

Sarawut Meesri et al[7] have proposed a new classification algorithm by combining the existing classification models and thus taking the advantage of each classification model. The proposed model works in two stages. In the first stage, the dataset is given to a composed technique formed by three classifier algorithms namely Naïve Bayes (NB), k-Nearest Neighbour (kNN), and Support Vector Machine(SVM). In the second stage, the dataset resulting from the first stage is fed to a multi-layer perceptron with back-propagation. If the dataset is huge and contains a lot of noise then KNN is among the most suitable machine learning algorithm.

Md. Razu Ahmed et al[8] have proposed a real-time cloud-based four-tier architecture in order to improve the prediction of heart disease significantly in comparison with the traditional methods. It made use of five prominent supervised machine learning algorithms. Tier 1 is responsible for collecting data from various sources and combining them into one. For collecting data various sensors and devices can be used. In Tier 2, real-time data are stored. In Tier 3, the classification model is trained on the real-time data. In Tier4, the output of the model is stored.

Mohan et al[9] proposed a hybrid algorithm which is a combination of many classification techniques.This algorithm produces an accuracy level of 88.7%. It makes use of Cleveland dataset. This dataset has 303 records and 14 attributes. First, the dataset is pre-processed, and then it is fed to the models. Records that had missing values were removed from the dataset. After pre-processing records were left to be 297. Out of 297, 137 records were of people with heart disease and the remaining records were of patients who do not have heart disease. Out of 14 attributes, the last attribute indicated the presence and absence of heart disease. The attribute age and sex was information about the patient and the remaining 11 attributes were important clinical records.

Repaka et al [10] has proposed a smart heart disease prediction that is built on Naïve Bayesian algorithm. The prediction is based on simple information of the patient like age, BP, blood sugar, sex, etc. The application built follows a systematic process of registering the user followed by collecting the information about the patient and then producing the result. AES is used for encrypting the data which will lead to safe data transfer. Machine learning algorithms can be used for making a meaningful decision based on the dataset collected from patients with heart disease. The main motive of developing the application is to detect heart disease at an early stage. If the heart disease is detected early then it becomes much easier to control and cure the disease. Naïve Bayesian model is trained using 80% of the dataset and the model is tested on the remaining 20% dataset.

Upretee et al [11] have proposed binary as well as a multi-class classification for detecting heart disease using heart-sound graphs of the person. Heart sound can be used to thoroughly examine heart condition. Heart sound can be used as a preliminary examination for detecting the heart disease of a person. We commonly see that a doctor uses a stethoscope to listen to heart sounds and make interpretations by listening to the sounds of the heart. Phonocardiogram is a technique of recording heart sound and plotting them to the graph. Normally, these graphs are interpreted by the doctors but the interpretation will be highly dependent upon the expertise level of the doctor. In this research paper, a machine learning algorithm is used for interpreting the graph which can give a much better result. The dataset has 1000 records, out 800 records are of normal people and the remaining 200 records show some abnormality in the heart function of the person.

Almustafa [12] has made use of various machine learning algorithms to train and test the model. It looks for the best classifier algorithm which has the highest accuracy. The dataset used contains 1025 records of the patients, although it contains 76 attributes only 14 attributes were used. It achieved 99.70% accuracy using the KNN classifier whereas 98.04% accuracy level was achieved using Decision Tree classifier and 97.26% accuracy level was achieved using JRip classifier. Heart disease depends on multiple parameters and it is a complex task to predict heart disease so a classifier algorithm should be selected carefully.

Alkhafaji et al[13] make use of various machine learning algorithms to predict heart disease. These predictions The dataset used for the prediction consists of 665 records. Out of 665 records, 300 records were of males, and the remaining 365 records were of females. Each record has 10 attributes. The target was to make the best possible prediction. First data is prepared to train the model. In data

preparation, we clean and integrate the data. It may also include data transformation and data reduction. The missing field is either replaced with the average value or that particular record is deleted. Decision Tree produced the highest accuracy of 98.85%. Naïve Bayes produced an accuracy of 98.16% and Artificial Neural Networks produced an accuracy of 91.31%.

Pranav Motarwar et al [14] has used five machine learning algorithms namely, Radom "Forest, Naïve Bayes, Support Vector Machine, Hoeffding Decision Tree, and Logistic Model Tree" [14] for predicting heart disease. First, data visualization is done, and after that feature selection is performed. After feature selection, models are trained and tested with the Cleveland dataset. The dataset has 303 records. Out of 303 records 242 records were used for training the model which approximately constitutes 80% of the dataset and the remaining 61 records were used for testing the model which approximately constitutes 20% of the dataset. Gaussian NB achieved an accuracy level of 93.44%, Support Vector Machine (SVM) achieved an accuracy level of 90.16%. Hoeffding Tree gave an accuracy of 81.24%. Logistic Model Tree gave an accuracy level of 80.69%. Random Forest gave the highest accuracy with 95.08%.

Lin et al [15] has used two machine learning algorithms, CNN (Convolutional Neural Network) and NN (Neural Network) to predict heart disease. It is found that NN has more accuracy than CNN. Regular Neural Networks has an input layer, several hidden layers, and an output layer. Convolutional Neural Networks(CNN) follows different architecture in comparison to regular Neural Networks (NNs) in several aspects like in CNN, neuron of each layer has three dimension -height, width, and depth. CNN gives a better accuracy level than regular NNs. NN produced an accuracy level of 86.40% when it is using 3 hidden layers. CNN produced an accuracy level of 78.65% when it is using 3 hidden layers.

Archana et al [16] had made use of four machine-learning algorithms to predict heart disease. SVM produced an accuracy level of 83%. Decision Tree gave an accuracy level of 79%. Linear Regression gave an accuracy level of 78%. K-Nearest Neighbour produced the highest accuracy of 87%. In the Decision Tree, we need to be cautious of the underfitting and over-fitting conditions of the model. In KNN algorithm, k nearest points are considered while making decision.Training and testing of the model are done on the dataset provided by UCI repository. For implementation purposes, Python programming language is used.

Table I: A Comparative study of various algorithms.

S.No.	Year	Author	Research Work	Algorithms	Accuracy
1	2017	Hamidreza and Morteza[3]	"Cardiovascular disease detection using a new ensemble classifier"[3]	Decision Tree	74.8%
				SVM Algorithm	75.4%
				SVM Algorithm (Poly Function)	75.4%
				Neural Network Algorithm	86.9%
				Naïve Bayes	85.4%
				Ensemble classifier	89%
2	2017	Martin Gjoreski et al[4]	"Chronic Heart Failure Detection from Heart Sounds Using a Stack of Machine-Learning Classifier" [4]	Stack of ML algorithms (J48, Naïve Bayes, kNN, SVM, Random Forest, Bagging, and Boosting)	96%
3	2018	Aditi Gavhane et al[5]	"Prediction of Heart Disease Using Machine Learning" [5]	ANN	Multi-layer perceptron algorithm
4	2018	Cincy Raju et al[6]	"A Survey on Predicting Heart Disease using Data Mining Techniques" [6]	Decision Tree	82.35%
				SVM	99.3%
				Neural Networks	91.1%
				kNN	87.2%
5	2018	Sarawut Meesri et al[7]	"Diagnosis of Heart Disease Using a Mixed Classifier" [7]	Naïve Bayes approach	84.49%
				SVM	84.82%
				KNN method	85.15%
				J48 decision tree	77.56%
				Majority Vote from Naïve Bayes, SVM,	85.48%

				and KNN	
				Mixed Classifier	86.16%
6	2018	Md. Razu Ahmed et al[8]	"A cloud based four-tier architecture for early detection of heart disease with machine learning algorithms" [8]	Naïve Bayes	82%
				Random Forest	78%
				ANN	84%
				Support Vector Machine	82%
				Decision Tree	77%
7	2018	Mohan et al[9]	"Effective heart disease prediction using hybrid machine learning techniques" [9]	Naïve Bayes	75.8%
				Generalized Linear Model	85.1%
				Logistic Regression	82.9%
				Deep Learning	87.4%
				Decision Tree	85%
				Random Forest	86.1%
				Gradient Boosted Trees	78.3%
				Support Vector Machine	86.1%
				VOTE	87.41%
				Hybrid Random Forest with Linear Model (HRFLM)	88.4%
8	2019	Repaka et al[10]	"Design and implementing heart disease prediction using naives Bayesian" [10]	Naives Bayesian	89.77%
9	2019	Upretee et at [11]	"Accurate classification of heart sounds for disease diagnosis by a single time-varying spectral feature: Preliminary results" [11]	SVM	95.07%
				kNN	96.50%
10	2020	Almustafa [12]	"Prediction of heart disease and classifiers' sensitivity analysis" [12]	kNN(k=1)	99.7.73%
				Decision Tree J48	98.0488%
				JRip classifier	97.2683%
11	2020	Alkhafaji et al[13]	"Clean medical data and predict heart disease" [13]	Decision Tree	98.85%
				Naïve Bayes	98.16%
				Artificial Neural Network	91.31%
12	2020	Pranav Motarwar et al[14]	"Cognitive Approach for Heart Disease Prediction using Machine Learning" [14]	Gaussian NB	93.44%
				Support Vector Machine	90.16%
				Random Forest	95.08%
				Hoeffding Tree	81.24%

				Logistic Model Tree	80.69%
13	2020	Lin et al[15]	"On Machine Learning Models for Heart Disease Diagnosis" [15]	Neural Networks	86.40% (3 hidden)
				Convolutional Neural Networks	78.65% (3 hidden layers)
14	2020	Archana Singh, Rakes Kumar [16]	"Heart Disease Prediction Using Machine Learning Algorithms." [16]	Support Vector machine	83%
				Decision Tree	79%
				Linear Regression	78%
				K-nearest neighbour	87%

3. Conclusion and Future Work

The availability of the records of the patients is increasing every passing year. We can make use of these data for helping the patient through the use of machine learning algorithms. Different machine learning algorithms such as k-NN, Decision Tree, etc can be used to predict the occurrence of a heart attack. By using these algorithms we can detect a patient with heart disease at an early stage. If the patient is detected at an early stage then it becomes very easy to save their life.

The number of deaths is increasing every passing year due to heart attacks. Preventive measures are more effective than curative measures for saving the life of heart patients. Heart disease is significantly dependent upon the lifestyle and food habits of the person. The prominent reason for the disease are smoking, alcohol intake, bad food habits, lack of physical activity, obesity, high blood pressure. There are certain foods that help in the reduction of the risk factor of heart disease like Walnuts, Fish such as salmon. Certain foods are very harmful to the heart for example fast foods that make use of trans fat for extra test and flavor. So, a food and lifestyle change recommendation system can be built for patients to reduce the chances of heart disease and to cure the patient of chronic heart diseases.

References

[1] Prediction of Heart Disease Using Machine Learning

[2] https://www.mayoclinic.org/diseases-conditions/heart-attack/symptoms-causes/syc-

[3] Hamidreza Ashrafi Esfahani, Morteza Ghazanfari(2017) . Cardiovascular disease detection using a new ensemble classifier. *International Conference on Knowledge-Based Engineering and Innovation (KBEI).*

[4] Gjoreski, M., Gradišek, A., Gams, M., Simjanoska, M., Peterlin, A., & Poglajen, G. (2017). Chronic heart failure detection from heart sounds using a stack of machine-learning classifiers. Proceedings - 2017 13th International Conference on Intelligent Environments, IE 2017, 2017-Janua, 14–19. https://doi.org/10.1109/IE.2017.19

[5] Aditi Gavhane, Gouthami Kokkula, Isha Pandya, Prof. Kailas Devadkar. Prediction of Heart Disease Using Machine Learning.

[6] Raju, C., Philipsy, E., Chacko, S., Suresh, L. P., & S, D. R. (2018). A Survey on Predicting Heart Disease using Data Mining Techniques. *2018 Conference on Emerging Devices and Smart Systems (ICEDSS), March*, 253–255.

[7] Meesri, S., Phimoltares, S., & Mahaweerawat, A. (2018). Diagnosis of Heart Disease Using a Mixed Classifier. *ICSEC 2017 - 21st International Computer Science and Engineering Conference 2017, Proceeding*, *6*, 118–123. https://doi.org/10.1109/ICSEC.2017.8443940

[8] Ahmed, M. R., Hasan Mahmud, S. M., Hossin, M. A., Jahan, H., & Haider Noori, S. R. (2018). A cloud based four-tier architecture for early detection of heart disease with machine learning algorithms. *2018 IEEE 4th International Conference on Computer and Communications, ICCC 2018*, 1951–1955. https://doi.org/10.1109/CompComm.2018.8781022

[9] Mohan, S., Thirumalai, C., & Srivastava, G. (2019). Effective heart disease prediction using hybrid machine learning techniques. *IEEE Access*, *7*, 81542–81554. https://doi.org/10.1109/ACCESS.2019.2923707

[10] Repaka, A. N., Ravikanti, S. D., & Franklin, R. G. (2019). Design and implementing heart disease prediction using naives Bayesian. *Proceedings of the International Conference on Trends in Electronics and Informatics, ICOEI 2019, 2019-April*(Icoei), 292–297. https://doi.org/10.1109/icoei.2019.8862604

[11] Upretee, P., & Yüksel, M. E. (2019). Accurate classification of heart sounds for disease diagnosis by a single time-varying spectral feature: Preliminary results. *2019 Scientific Meeting on Electrical-Electronics and Biomedical Engineering and Computer Science, EBBT 2019*, 1–4. https://doi.org/10.1109/EBBT.2019.8741730

[12] Almustafa, K. M. (2020). Prediction of heart disease and classifiers' sensitivity analysis. *BMC Bioinformatics*, *21*(1), 1–18. https://doi.org/10.1186/s12859-020-03626-y

[13] Alkhafaji, M. J. A., Aljuboori, A. F., & Ibrahim, A. A. (2020). Clean medical data and predict heart disease. *HORA 2020 - 2nd International Congress on Human-Computer Interaction, Optimization and Robotic*

Applications, Proceedings. https://doi.org/10.1109/HORA49412.2020.9152870

[14] Motarwar, P., Duraphe, A., Suganya, G., & Premalatha, M. (2020). Cognitive Approach for Heart Disease Prediction using Machine Learning. *International Conference on Emerging Trends in Information Technology and Engineering, Ic-ETITE 2020*. https://doi.org/10.1109/ic-ETITE47903.2020.242

[15] Lin, C. H., Yang, P. K., Lin, Y. C., & Fu, P. K. (2020). On Machine Learning Models for Heart Disease Diagnosis. *2nd IEEE Eurasia Conference on Biomedical Engineering, Healthcare and Sustainability 2020, ECBIOS 2020*, 158–161. https://doi.org/10.1109/ECBIOS50299.2020.9203614

[16] Archana Singh, Rakes Kumar. Heart Disease Prediction Using Machine Learning Algorithms. *International Conference on Electrical and Electronics(ICE3-2020)*. https://doi.org/10.1109/INCET49848.2020.9154130