DOI: 10.56452/2021SP-8-009

ISSN: 0974-5823 Vol. 6 (Special Issue 3, November 2021)

International Journal of Mechanical Engineering

Cardiovascular Diseases Detection using Machine Learning Algorithms

Rohini Mahajan[1], Parminder Pal Kaur[1], Rini Saxena[1] and Manjeet Kaur[2]

[1]Department of Computer Science & Engineering, Chandigarh Engineering College, Jhanjeri, Mohali-140307, Punjab, India

[2]Department of Agriculture, Chandigarh School of Business, Jhanjeri, Mohali-140307, Punjab, India

Email id: er.rohinimahajan @gmail.com; sukh.pammi.pk@gmail.com; rini.cgctc@gmail.com

Abstract: Cardiovascular diseases are one of the most well-known things causing deaths worldwide over the span of the recent couple of decades in majority of the nation's worldwide, irrespective of a nation being developed, developing or an underdeveloped one. Early detection of heart issues or any other related problem and the constant support and advice by the doctors can diminish the huge death rate that has been constantly increasing for the past few decades now. Right now, a speculative plan of a heart health prediction system framework has been proposed to recognize approaching heart illness utilizing Machine learning algorithms. For the precise recognition of the heart diseases and health, an effective machine learning system ought to be utilized. Some of the information extraction and system learning strategies that is used to predict the heart health inclusive of "Artificial Neural Network" (ANN), "Decision tree", "K-Nearest Neighbour" (KNN), "Naïve Bayes" and "Support Vector Machine" (SVM).

Keywords: Heart, Machine Learning, Human, cardiovascular

Introduction

The Heart is the following primary essential organ comparing to the brain/neural system which has greater priority in the Human frame [1]. The heart circulates the blood to all organs located across the human body. The heart is a crucial organ that has the major task of circulating the flow of blood all over the human body and is the focal point of the human body's cardiovascular system which is also comprised of the contains lungs. The cardiovascular system hence contains a system of veins, for instance, veins, arteries, and vessels. The veins transport blood everywhere throughout the body.

In our daily life, anxiety and pressure tags along with people while they are simply undergoing a habitual life which is inclusive of a busy and stressful agenda. In addition to this, the graph of the share of those who are dependent and addicted to cigarettes goes up increasing drastically as we breathe. These finally become the causes and results in illnesses like coronary heart disorder, most cancers, and so forth. The undertaking behind those illnesses is their assumption. Every individual has extraordinary values of pulse rate / heart beat rate and blood pressure. But medically demonstrated, the heartbeat rate ought to be in the range of 60 - 100 beats in a minute and the blood pressure ought to be within the upper limit of 120/80 in and around. Heart disorders are the cause of the majority of deaths caused all over the globe. The number of people laid low with heart sickness keeps increasing regardless of age in each women and men. But different elements other than gender like diabetes, BMI also make a considerably significant amount of contributions to this disease. In this paper, we have attempted prediction and evaluation of heart ailment through considering the specifications like age, sex, blood pressure, coronary heart price, diabetes and so forth. Since severe elements are involved in heart ailment, the conclusion of this disorder is difficult.

Few of the common signatures of coronary heart attack are [2][4]:

1. Tightened torso.
2. Shortened breath.
3. Vomiting
4. Dyspepsia
5. Heartburn
6. In some cases, abnormal aches.
7. Pain that spreads to the arm.

The below mentioned are the kind of heart sickness: Heart manner “cardio”. Hence all coronary heart sicknesses falls under the category of cardiovascular sicknesses. The extraordinary kinds of heart disease are :

1. Coronary heart Diseases (CHD): If a person is affected by this disease then it implies that his coronary arteries tend to become narrower, reducing the free flow of blood
2. Angina pectoris: This medical keyword is used to refer the chest pain and tidiness because of CHD.
3. Congestive heart failure: A critical failure that causes the reduction of pumping power of heart muscles because of fluids accumulation.
4. Cardiomyopathy: In the case of cardiomyopathy, the heart fails to pump blood to various other parts.
5. Congenital heart disease: Since birth, the patients tend to have one or more forms of abnormalities. The most common defect being the alteration of the path in which blood flows.

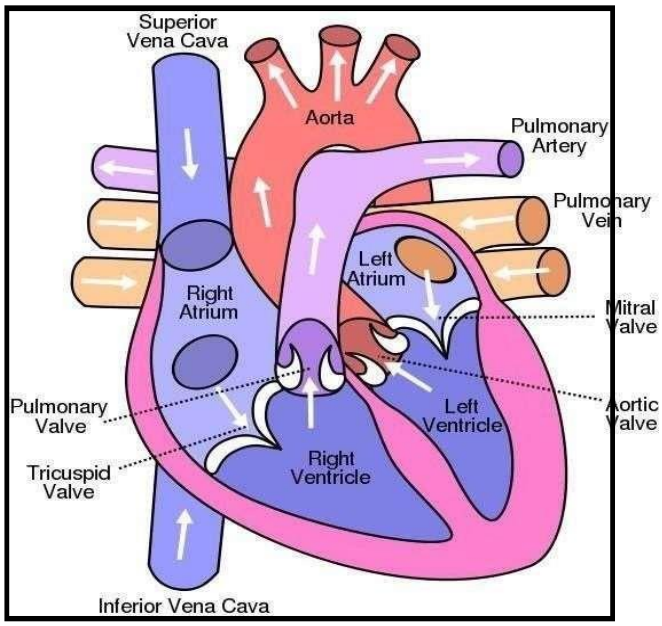

Fig. 1 Heart, the pumping machine [1]

Coronary heart disease (CHD) is the shrinking of the coronary arteries which restricts the easy flow of blood [4]. The purpose of the coronary artery is to carry oxygen and blood into the coronary heart. It is the cause of a significantly huge quantity of people to turn out to be ill or to witness another loss of life. It is as life-threatening as any other form of heart disorders. High blood glucose levels as a result of diabetes can harm blood vessels and the nerves that manage the heart and the blood vessels. Having contracted with diabetes for a considerably long period increases the probability of contracted coronary heart disease regardless of the gender the person involved. With diabetes, there are different reasoning for contracting coronary heart disorder which varies from person to person. And smoking also raises the chance of employing a heart disease, along with it, Increase in blood pressure level makes the heart's functioning poorer to pump blood and it can strain the coronary heart and harm blood vessels. Abnormal i.e mostlty increase cholesterol levels and obesity also make contributions to coronary heart sickness. Also, one's family history of contracting with heart disease can be the purpose of having any heart ailment. In this paper for the prediction of heart disorder,

family history isn't available for all patients. The different hazard factors consist of age, sex, strain, and an unhealthy diet. The chance of having a heart disorder increases as the person grows older. Leading a harassed life can also harm the arteries and increase the chance of having a coronary heart ailment. So, based on several previously published by various researchers we try to are expecting the danger of coronary heart disease. Already there is a massive amount of work that has been done related to coronary heart prediction by using the diverse strategies and algorithms. These predicting ways are based on deep-learning, machine-learning, Information mining and so on. The major objective of considering all the above-mentioned works is to obtain better results with utmost accuracy and to build the system extra useful and efficient so that it can predict the probabilities of a heart attack. Few of the data mining and machine learning manoeuvrings are used to await the coronary heart sickness, inclusive of "Artificial Neural Network (ANN), Decision tree, Fuzzy Logic, K-Nearest Neighbour (KNN), Naïve Bayes and Support Vector Machine (SVM)." The significance and focal point of the use of Machine learning-based coronary illness location and expectation framework were examined in a few research discoveries. The use of man-made consciousness in sickness location framework particularly the cardiovascular infection framework recognition improves the presentation of other existing The work points in building up a Decision Support System in coronary illness identification that utilizes the information mining method having the best exactness and execution among Naïve Bayes, Support Vector Machine, Simple Logistic Regression, Random Forest and Artificial Neural Network (ANN) and so on. By utilizing a few cardiovascular framework parameters, for example, age, circulatory strain, ECG results, sex, and glucose, it is convincing to calculate the likelihood of getting affected by coronary illness. For understanding the calculation with the best accuracy in the discovery and calculation of coronary illness, a near examination of picked AI computations has emerged. This calculation takes the clinical parameters, for example, age, circulatory strain, heartbeat, sex, ECG results, glucose and so on as information and shows the likelihood of getting influenced by coronary illness as yield. This proposed framework involves the strategy and formation of machine learning algorithms to recognize coronary illness. It can fill in as an exceptionally helpful instrument for specialists, patients, and clinical understudies to analyze coronary illness.

Proposed work

In the world of data, whenever we hear the term "Big Data" the first thing that comes to our mind above all thoughts is "Machine Learning". From identifying trends to handling multi- dimensional data, ML can perform a lot of task. Another important thing is its continuous improvement and its automation i.e. no intervention of humans prevents the chances of errors which is indeed an advantage. So, whatever maybe the problem or how big the problem maybe, Machine Learning can play a great role in solving it. Hence, we started our research by implementing various ML algorithms to our dataset of patients. The ML algorithms that we implemented in the process are Logistic Regression, K-nearest neighbour (KNN), Decision tree, K-Fold method, SVM, Naïve Bayes and Voting classifier. The complete algorithms have been discussed in the below sections.

To start up the task, a productive ML technique was selected from some the set of available learning methods and techniques to cast the chances of suffering from coronary disorder or illness from a large dataset of about 300 patients. Then a continuous heart health detecting system design has been planned by making use of "Arduino based microcontroller" framework. The bit by bit configuration strategies of the framework and the work process of the framework is referenced underneath:

1. First step dealt with finding and collecting different heart health-based data, to which we can apply the ML algorithms to train our framework.
2. Comparing the accuracy and performance of various algorithms on the dataset to choose the best one amongst them was the task.
3. Picking and sorting the best technique based on the performance of the ML models build in the initial phases of the project.

DOI: 10.56452/2021SP-8-009

Results

The dataset that we used for this proposed idea is Heart Disease UCI dataset which had 76 attributes, but we used a small subset of them. We used 14 of them to do the task allotted in this project.

The dataset is as shown below-

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
0	63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
1	37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
2	41	0	1	130	204	0	0	172	0	1.4	2	0	2	1

Fig. 2 Data set

The dataset shown above in fig. 2 is just a part of it. It specifically had 14 rows and 303 entries i.e the columns. The various attributes were age, blood sugar level, chest pain type, serum cholesterol, blood sugar level etc.

The results of the 5 methods of machine learning implemented are as follows-

Logistic Regression model

```
Result of Logisitic Regression Model
              precision    recall  f1-score   support

           0       0.80      0.78      0.79        41
           1       0.82      0.84      0.83        50

    accuracy                           0.81        91
   macro avg       0.81      0.81      0.81        91
weighted avg       0.81      0.81      0.81        91

Classification Table
            Predicted (1)  Predicted (0) Total
Expected (1)     42              8           50
Expected (0)     9              32           41
Total            51              40           91
Sensitivity= 0.84
Specificity= 0.7804878048780488
Precision= 0.8235294117647058
Accuracy= 0.8131868131868132
```

Fig. 3 Logistic Regression Results

DOI: 10.56452/2021SP-8-009

Decision Tree model

```
Result of Decision Tree Model
              precision    recall  f1-score   support

           0       0.67      0.76      0.71        41
           1       0.78      0.70      0.74        50

    accuracy                           0.73        91
   macro avg       0.73      0.73      0.72        91
weighted avg       0.73      0.73      0.73        91

Classification Table
           Predicted (1)  Predicted (0) Total
Expected (1)     35             15            50
Expected (0)     10             31            41
Total            45             46            91
Sensitivity= 0.7
Specificity= 0.7560975609756098
Precision= 0.7777777777777778
Accuracy= 0.7252747252747254
```

Fig. 4 Decision tree results

K Nearest Neighbour model

```
Result of K Nearest Negihbors Model
              precision    recall  f1-score   support

           0       0.65      0.68      0.67        41
           1       0.73      0.70      0.71        50

    accuracy                           0.69        91
   macro avg       0.69      0.69      0.69        91
weighted avg       0.69      0.69      0.69        91

Classification Table
           Predicted (1)  Predicted (0) Total
Expected (1)     35             15            50
Expected (0)     13             28            41
Total            48             43            91
Sensitivity= 0.7
Specificity= 0.6829268292682927
Precision= 0.7291666666666666
Accuracy= 0.6923076923076923
```

Fig.5 K-Nearest Neighbour model results

Voting Classifier model

```
Result of VotingClassifier
              precision    recall  f1-score   support

           0       0.79      0.76      0.77        41
           1       0.81      0.84      0.82        50

    accuracy                           0.80        91
   macro avg       0.80      0.80      0.80        91
weighted avg       0.80      0.80      0.80        91

Classification Table
           Predicted (1)  Predicted (0) Total
Expected (1)     42             8             50
Expected (0)     10             31            41
Total            52             39            91
Sensitivity= 0.84
Specificity= 0.7560975609756098
Precision= 0.8076923076923077
Accuracy= 0.8021978021978022
```

Fig. 6 voting classifier results

DOI: 10.56452/2021SP-8-009

Now applying K-Fold Cross Validation model on the results of Voting Classifier

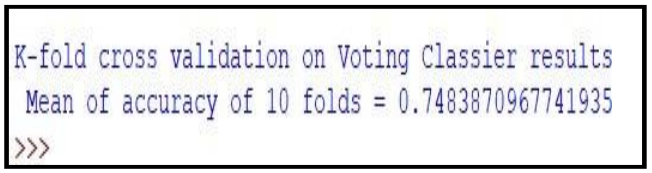

Fig. 7 K-Fold cross validation results The above results can be summarized as shown-

Table 1: Comparison analysis of all models in terms of accuracy and performance

Parameters	Accuracy	Precision	Sensitivity	Specificity
Logistic Regression	0.8131	0.8235	0.84	0.7804
Decision Tree	0.7252	0.7777	0.7	0.7560
K-nearest Neighbors	0.6923	0.7291	0.7	0.6829
Voting Classifier	0.8021	0.8076	0.84	0.756
K-fold Cross Validation (mean accuracy)	0.7483	-	-	-

From the above results we can conclude that we got highest accuracy and precision in prediction when using Logistic Regression model as compared to other models when performed on the same dataset.

Conclusion

Heart ailments are of the most injurious ailment for human health. Heart being the most important organ needs to be taken care the most. Heart health as we have seen is dependent on a lot of factors and activities of one does. In this paper we have implemented machine learning algorithms using logistic Regression model, Decision tree, K-nearest neighbors model, voting classifier and K-fold cross validation model on different dataset for identifying heart disease. After implementing these machine learning algorithms we provided a comparison of all ML algorithms. In this method Logistic Regression model shows better in accuracy and others parameters on same dataset. Overall, the system aims in contributing to the wide heart health monitoring system. The system also records and saves data and can gives the compared results to track what's the progress or degradation in the health.

Future scope

In future work we will work on IOT, in which we will done things in real-time and predicting the outcomes. We will use various signs like ECG, heart beat etc. and also various environmental factors like temperature and humidity which also play a significant role in the overall health of an individual. We assume that this implementation will definitely lead to increased accuracy and precision as compared to ML based algorithms which does not take into account the environmental conditions as well as real-time data for the predictions.

Future work that can be involved with the system is increasing the number of sensors and considering more factors to give results with more accuracy

References

[1] M Marimuthu, M Abinaya, K S Hariesh, K Madhankumar and V Pavithra. A Review on Heart Disease Prediction using Machine Learning and Data Analytics Approach. International Journal of Computer Applications 181(18):20-25, September 2018.

[2] https://my.clevelandclinic.org/health/s ymptoms/16821-coronary-artery- disease- symptoms https://www.medicalnewstoday.com/article s/237191 Everything you need to know about heart disease.

[3] https://www.medicalnewstoday.com/articl es/184130 What to know about coronary heart disease?

[4] Madani, Arnaout, Mofrad, et al. Fast and accurate view classification of echocardiograms using deep learning. NPJ Digit Med. doi: 10.1038/s41746-017-0013- 1.: 21 March 2018.

[5] Playford, Bordin, Talbot, et al. Analysis of aortic stenosis using artificial intelligence. Heart Lung Circ. 2018;27:S216–S216.

[6] Seah, Tang, Kitchen, et al. Chest radiographs in congestive heart failure: visualizing neural network learning. Radiology. 2019;290:514–522.

[7] Tao Q, Yan W, Wang Y, et al. Deep learning– based method for fully automatic quantification of left ventricle function from cine MR Images: a multivendor, multicenter study. Radiology. 2019;290:81–88.

[8] Ngo TA, Lu Z, Carneiro G. Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance. Med Image Anal. 2017;35:159–171.

[9] Isin A, Ozdalili S. Cardiac arrhythmia detection using deep learning. Procedia Comput Sci. 2017;120:268–275.

[10] Przewlocka-Kosmala, et al. Contribution of Cardiovascular Reserve to Prognostic Categories of Heart Failure with Preserved Ejection Fraction: A Classification Based on Machine Learning.

[11] Kwon J, Lee Y, Lee Y, et al. An algorithm based on deep learning for predicting in- hospital cardiac arrest. J Am Heart Assoc. doi: 10.1161/JAHA.118.008678. Published Online First: 26 June 2018.

[12] Mortazavi BJ, Downing NS, Bucholz EM, et al. Analysis of machine learning techniques for heart failure readmissions. Circ Cardiovasc Qual Outcomes. 2016;9:629–640.

[13] Alaa A., Bolton T, Di Angelantonio E, et al. Cardiovascular disease risk prediction using automated machine learning: a prospective study of 423,604 UK Biobank participants. PLoS One. 2019;14:e0213653–e0213653.

[14] http://archive.ics.uci.edu/ml/datasets/Heart+Disease

[15] Poorejbari, Vahdad, Mansoor. Diabetes patients monitoring by cloud computing Cloud Computing Systems and Applications in Healthcare 2016: 99 DOI: 10.4018/978-1- 5225- 1002-4.ch006. Pdf available Jan 2017

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

[17] T. Mythili, D. Mukherji, N. Padaila and A. Naidu, "A Heart Disease Prediction Model using SVM Decision Trees- Logistic Regression (SDL)", International Journal of Computer Applications, vol. 68, 16 April 2013. International Journal of Computer Applications (0975 – 8887) Volume 181 – No. 18, September 2018 25

[18] A. Dwivedi, "Evaluate the performance of different machine learning techniques for prediction of heart disease using ten-fold cross- validation", Springer, 17 September 2016.

[19] A. Dwivedi, "Evaluate the performance of different machine learning techniques for prediction of heart disease using ten-fold cross- validation", Springer, 17 September 2016.

[20] J. Patel, Prof. T. Upadhay, Dr. S. Patel, "Heart disease Prediction using Machine Learning and Data mining Technique", March 2017.

[21] Purush ottam, Prof. K. Saxena, R. Sharma, "Efficient Heart Disease Prediction System", 2016, pp.962-969.

[22] N. Ajam, "Heart Disease Diagnoses using Artificial Neural Network", The International Insitute of Science, Technology and Education, vol.5, No.4, 2015, pp.7- 11.

[23] Pavithra B.., Rajalakshmi V.,(2020) "Heart Disease Detection Using Machine Learning Algorithms" , Springer Nature DOI:doi.org/10.1007/978-3-030- 32150-5_115 , pp 1131-1137