Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)

Volume 5- Issue 2, Paper 8 August 2022

Early Prediction of Cardiac Disease using Expert Systems

AJAAY KRISHNA P , ARAVIND C J ,
AKHILESH R, DR RAMA ABIRAMI K
Computer Science and Engineering
Sri Krishna College of Engineering and Technology
Coimbatore, India

***Abstract*-- Machine learning is effective in helping and making selections from the high volumes of data created by the healthcare business. In this work, completely different classification algorithms are applied with their own advantage on separate databases of malady accessible for disease prediction. The results of the study strengthen by using Artificial Intelligence in the early detection of diseases and this will increase the survival rate of patients considerably. The motivation of this paper is to develop efficacious treatment of data processing techniques that will facilitate remedial things. Data processing classification algorithms are used to diagnose heart diseases.**

I. INTRODUCTION

Cardiac disease describes a range of heart-related conditions. Heart conditions embrace a variety of diseases, like arteria malady, rhythm diseases (arrhythmias), and heart defects (congenital heart defects), among others. The term vessel malady is usually used interchangeably. The process of Data Mining is the discovery of the unknown hidden patterns (knowledge) of enormous pre-existing knowledge sets involving data mining, machine teaching, statistics, and info systems. The information discovered might be used to develop intelligent prophetical deciding systems for correct designation in completely different fields such as health care to provide reasonable service and save precious lives. The ability to learn from planned data and improve performance while not being an interference with human beings and then using what they have learned to build a wise call is provided by artificial intelligence. It is, therefore, vital to document the maximum vital signs, symptoms and fitness behavior that make a contribution to CVD. Various tests are executed previous to analysis of CVD, which

includes auscultation, ECG, blood strain, ldl cholesterol and blood sugar. These exams are frequently lengthy and lengthy whilst an affected person's situation can be important and they need to begin taking medicine right away, so

it will become vital to prioritize the tests. Several fitness behaviors contribute to CVD. Therefore, it's also vital to understandwhichfitnessbehaviormakes a contribution to CVD. Machine gaining knowledge is now a rising subject because of the growing quantity of information. Machines gaining knowledge makes it viable to collect the required amount of data from a large quantity of information, which may be very heavy for a person and occasionally impossible. The goal of this paper is to prioritize the diagnostic check and to look at a number of the fitness behaviors that make a contribution to CVD. Moreover, and above all, the unique devices gaining knowledge of algorithms are in comparison the usage of wise optimization algorithms. In this article, manually categorized information is used. Intelligent optimization algorithms are evolved through simulating or revealing phenomena and are broadly used in lots of research fields due to their versatility. Classification is one of the information mining approaches to categorize the affected person elegance as regular or coronary heart ailment however type use all attributes both applicable or inappropriate functions which can also additionally lessen the type performance. Feature subset choice is one of the dimensionality discount strategies used to enhance the accuracy. Our proposed version identifies the applicable functions and eliminates the inappropriate functions, to be expecting the coronary heartailmenteffectivelyusingawebapp.

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)

Volume 5- Issue 2, Paper 8 August 2022

II. LITERATURE SURVEY

Year	Author	Algorithm	Inference
2015	L. Sarangi, M. N. Mohanty, S. Pattnaik[4]	1.Radial Basis 2.Function Neural Networks 3.Gaussian Kernel function	RBFN with Gaussian Kernel function detects the problem with high accuracy
2018	N.Singh and S.Jindal[2]	1.Genetic Algorithm 2.Naive Thomas Bayes Hybrid Technique of Data Mining Algorithms	Obtained accuracy of 97.14%
2018	N.Shirwalkar and T.Tak [3]	1.Naive Bayes 2.Improved K-means algorithm	K-means: Simple to build, useful if information is unorganized, produces tighter clusters when variables are large.
2018	V.V. Ramalingam[5]	1.Ensemble Model 2.K-Nearest Neighbor 3.Naive Bayes 4.Random Forest 5.Support Vector Machines (SVM) 6.Decision trees	SVM is more accurate than the other techniques with 98.9%
2018	Abhay Kishore [8]	1.Convolutional Neural Network(CNN)& Decision Tree 2.Support Vector Machine (SVM) 3.Naive Bayes 4.Kohonen Self-Organizing Map (KSOM) 5.C4.5 MAFIA K-means Cluster 6.DMN 7.Recurrent Neural Network(RNN)	RNN has increased the heart attack accuracy of 92% and has proved to be an excellent source in predicting heart attacks.

2019	Mr. S. J. Krishnan and Dr. S. Geetha[6]	1.Naive Bayes 2.Decision Trees 3.ID3 Algorithms	Decision tree - 91% accuracy
2019	Avinash Golande [1]	1.K-Mean Clustering 2.K- Nearest Neighbor 3.Decision Tree	Decision Tree is better than the other algorithms with 86.6% accuracy
2019	M.Mari muthu, S.Deivarani, R. Gayathri [7]	1. Support vector machine (SVM) 2. Decision Tree Naïve Bayes approach 3.K- Nearest neighbor	K-nearest neighbor scored better with high accuracy of 83.6%
2019	M P Alex and S P Shaji [9]	1.Artificial Neural Network 2.KNN 3.Random Forest 4.SVM	Artificial neural networks provide the highest accuracy among all techniques

III. EXISTING SYSTEM

As consistent with the latest examination through WHO, coronary heart associated illnesses are growing. 17.9 million humans die every-year because of this. With a developing population, it is in addition hard to diagnose and begin remedy at an early level. But because of the latest development in technology, Machine Learning strategies have multiplied the fitness quarter through more than one research. Thus, the goal of this paper is to construct a ML version for coronary heart ailment prediction primarily based totally at the associated parameters. We have used a benchmark dataset of UCI Heart ailment prediction for this study's paintings, which include 14 unique parameters associated with Heart Disease. Machine Learning algorithms along with Random Forest, Support Vector Machine (SVM), Naive Bayes and Decision tree were used for the improvement of version. In the studies additionally attempted to locate the correlations among the unique attributes to be had within the dataset with the assistance of general Machine Learning strategies after which the usage of them successfully within the prediction of possibilities of Heart ailment. Result suggests that in comparison to different ML strategies, Random Forest offers greater accuracy in much less time for the prediction. This version may be useful to the scientific practitioners at their clinics as a selectional guide system.

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)

Volume 5- Issue 2, Paper 8 August 2022

IV. PROPOSED SYSTEM

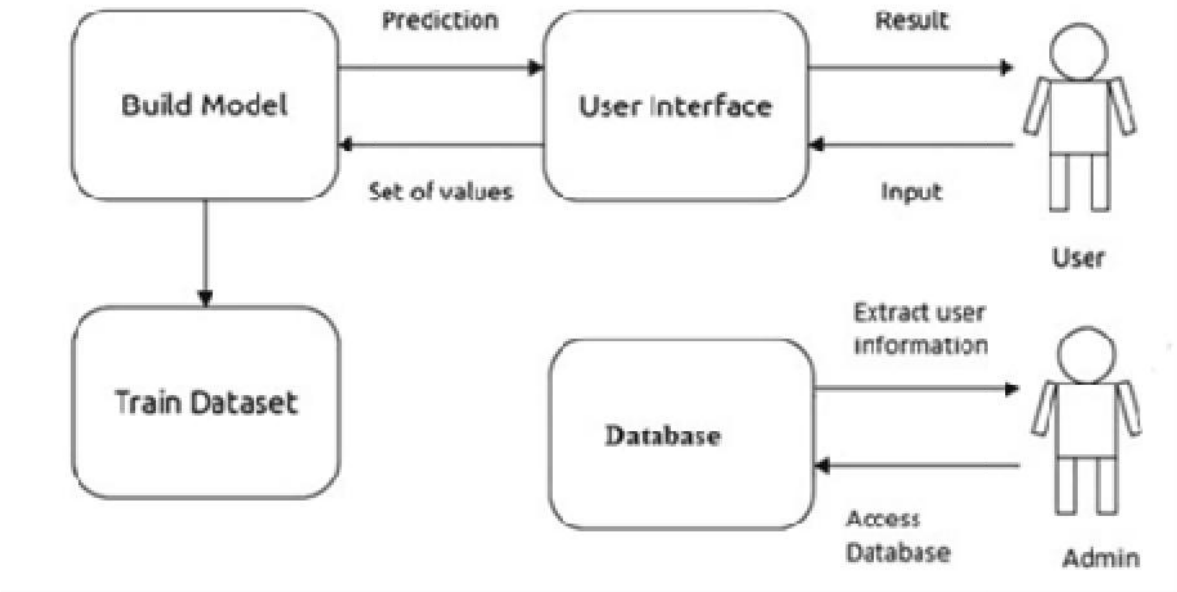

Figure 1. Block Diagram

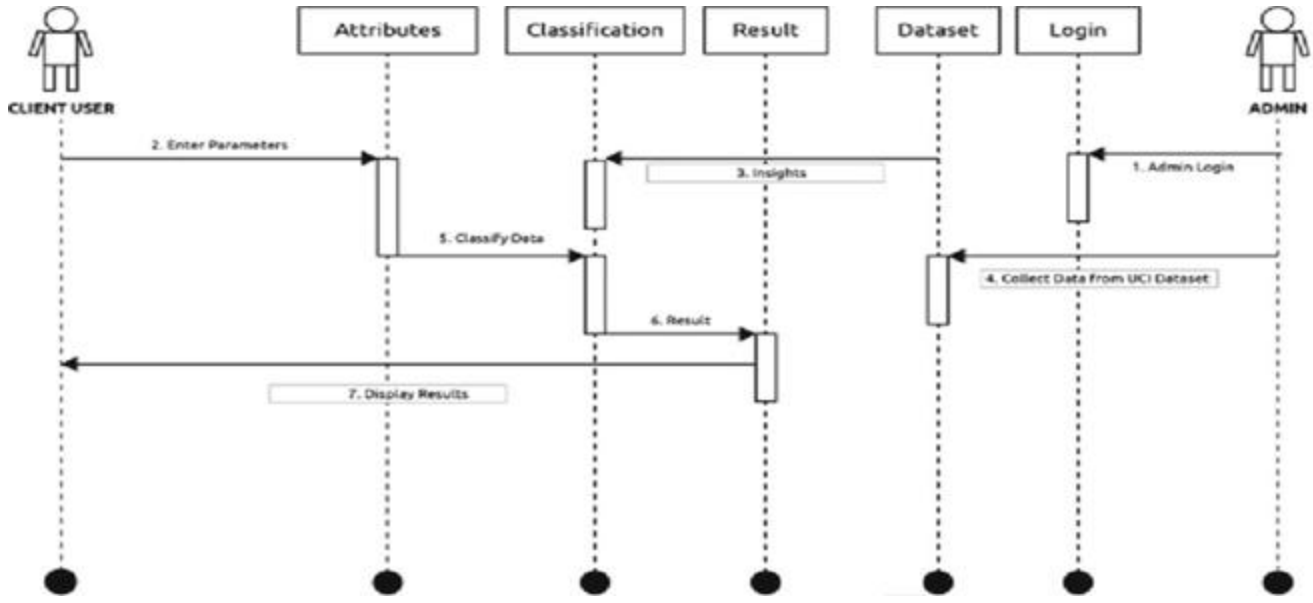

Figure 2. Architecture diagram

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)

Volume 5- Issue 2, Paper 8 August 2022

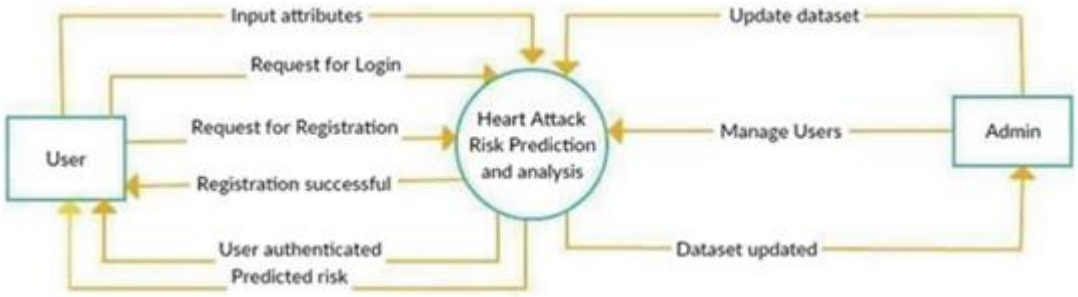

Figure 3. workflow of the proposed

Cardiovascular ailment is a chief purpose of loss of life worldwide. Data mining is an important part of the scientific quarter in permitting functional frameworks to utilize statistics and studies to pick out the vulnerabilities which enhance remedy at decreasing costs. One data mining approach as type is a supervised gaining knowledge used to as it should be expecting the goal elegance for every case with inside the information. Heart disorder type entails figuring out wholesome and unwell individuals. The user can login into the web app and make the registration and predict the risk with their readings by comparing with the provided dataset.

V. MODULES OF THE PROPOSED SYSTEM

The webapp consists of a login page, signup page, prediction page, details entering page, log out.

Clinical choices are frequently performed primarily based on surgeons' trip about the knowledge unseen between the data. This leads to errors that affect the characteristics regarding medical services. Using analytic equipment or statistics modeling perform assistance among improving the clinical decisions. Thus, the aim right here is to construct a web application to help medical practitioners diagnose and solve the diseases.

1. Data Analysis: This method entails facts cleaning, records statistics, getting insights from the dataset.

2. Algorithm Implementation: This entails four machine discipline algorithms which will end result within overall performance metrics on the model.

3. Model Implementation: The well-doing algorithm is applied within the model yet checking results together with the real-time data.

A. Data Preprocessing

The dataset chronic of its project incorporates 14 variables. The unbiased moving so wishes to stand predicted, 'diagnosis', determines whether or not a man or woman is healthy then go through out of bravery disease. Experiments along the Cleveland database bear gray on endeavors to individualize disease arrival beyond non-existence (value 0). There are various missing characteristic values, distinguished along the image. The header range is missing into that dataset, consequently the stagnation names bear in conformity with remain inserted manually. The missing values are numerical entries then we bear crammed the values by means of taking Mode

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)

Volume 5- Issue 2, Paper 8 August 2022

regarding the columns. We take a look at additionally to that amount there is an inverse proportion of mettle ailment and maximum mettle degree (thalch).

B. Train and split

We cut up the whole dataset between the train set and test set as it carries 75% teach and 25% test.We execute this train put in between classifiers after educating our model and taking a look at embark is beneficial for predicting the overall performance concerning the model by exclusive classifiers.

VI. EXPERIMENTAL RESULTS

age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
52	1	0	125	212	0	1	168	0	1	2	2	3	0
53	1	0	140	203	1	0	155	1	3.1	0	0	3	0
70	1	0	145	174	0	1	125	1	2.6	0	0	3	0
61	1	0	148	203	0	1	161	0	0	2	1	3	0
62	0	0	138	294	1	1	106	0	1.9	1	3	2	0
58	0	0	100	248	0	0	122	0	1	1	0	2	1
58	1	0	114	318	0	2	140	0	4.4	0	3	1	0
55	1	0	160	289	0	0	145	1	0.8	1	1	3	0
46	1	0	120	249	0	0	144	0	0.8	2	0	3	0
54	1	0	122	286	0	0	116	1	3.2	1	2	2	0
71	0	0	112	149	0	1	125	0	1.6	1	0	2	1
43	0	0	132	341	1	0	136	1	3	1	0	3	0
34	0	1	118	210	0	1	192	0	0.7	2	0	2	1
51	1	0	140	298	0	1	122	1	4.2	1	3	3	0
52	1	0	128	204	1	1	156	1	1	1	0	0	0
34	0	1	118	210	0	1	192	0	0.7	2	0	2	1
51	0	2	140	308	0	0	142	0	1.5	2	1	2	1
54	1	0	124	266	0	0	109	1	2.2	1	1	3	0
50	0	1	120	244	0	1	162	0	1.1	2	0	2	1
58	1	2	140	211	1	0	165	0	0	2	0	2	1
60	1	2	140	185	0	0	155	0	3	1	0	2	0

Figure 4. Dataset

Figure 5. Count of each target class

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)

Volume 5- Issue 2, Paper 8 August 2022

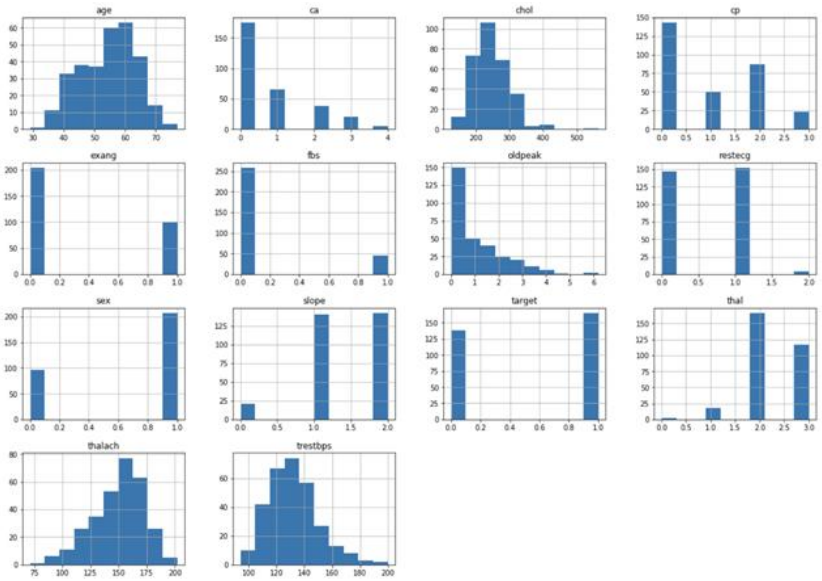

Figure 6. histogram equalization

```
[52] from sklearn.ensemble import RandomForestClassifier
     model = RandomForestClassifier(n_estimators=20)
     model.fit(X_train, y_train)

     RandomForestClassifier(bootstrap=True, ccp_alpha=0.0, class_weight=None,
                            criterion='gini', max_depth=None, max_features='auto',
                            max_leaf_nodes=None, max_samples=None,
                            min_impurity_decrease=0.0, min_impurity_split=None,
                            min_samples_leaf=1, min_samples_split=2,
                            min_weight_fraction_leaf=0.0, n_estimators=20,
                            n_jobs=None, oob_score=False, random_state=None,
                            verbose=0, warm_start=False)

[53] from sklearn.metrics import accuracy_score

     print(accuracy_score(y_test, model.predict(X_test))*100)

     93.75
```

FIGURE 7. ACCURACY PREDICTION

CONCLUSION

In this paper, datasets are used to get correct results with high accuracy regarding cardiovascular disease. Extraction of categorized accuracy beneficial for coronary heart ailment prediction. Random forest is distinctly solid with recognition of small versions or adjustments within the function normalization. The collected dataset is trained and data preprocessing and modeling is carried out before training and the model is predicted using random forest. Then

we deploy it as a web application. Identification of the first-rate performance-primarily based totally set of rules for disease prediction.

REFERENCES

[1] Avinash Golande, Pavan Kumar T Heart Disease Prediction Using Effective Machine Learning Techniques, International Journal of Recent Technology and Engineering (IJRTE) ISN: 2277-3878, Volume-8, Issue-1S4, June 2019.

[2]Navdeep Singh and Sonika Jindal "Heart Disease Prediction System using Hybrid Technique of Data Mining Algorithms", International Journal of Advance Research, Ideas and Innovations in Technology, Vol.4, Issue 2, 2018.

[3] Nikita shirwalkar, Tushar Tak "Human heart disease prediction system using data mining techniques", (IJIACS), vol.7 Mar 2018.

[4] Sarangi, Lokanath et al. "An intelligent decision support system for cardiac disease detection." *International Journal of Control Theory and Applications* 8.5 (2015): 2137-2143.

[5] Ramalingam, V. V. et al. "Heart disease prediction using machine learning techniques: a survey." *International Journal of Engineering & Technology* 7.2.8 (2018): 684-687.

[6] Krishnan, Santhana, and S. Geetha. "Prediction of Heart Disease Using Machine Learning Algorithms." 2019 1st international conference on innovations in information and communication technology *(ICIICT)*. IEEE, 2019.

[7] Marimuthu, M. et al. "Analysis of heart disease prediction using various machine learning techniques." *Advances in Computerized Analysis in Clinical and Medical Imaging*. Chapman and Hall/CRC, 2019. 157-168.

[8] Kishore, Abhay, et al. "Heart attack prediction using deep learning." *International Research Journal of Engineering and Technology (IRJET)* 5.04 (2018): 2395-0072.

[9] Shaji, Shaicy P. "Prediction and Diagnosis of Heart Disease Patients using Data Mining Technique." *2019 international conference on communication and signal processing (ICCSP)*. IEEE, 2019.

[10] Ghosh et al. "Efficient Prediction of Cardiovascular Disease Using Machine Learning Algorithms With Relief and LASSO Feature Selection Techniques." *IEEE Access* 9 (2021): 19304-19326.

[11] Limbitote et al. "A Survey on Prediction Techniques of Heart Disease using Machine Learning." *International Journal of Engineering Research & Technology(IJERT)* (2020): 2278-0181.