iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VIII **Month of publication:** August 2021

DOI: https://doi.org/10.22214/ijraset.2021.37727

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VIII Aug 2021- Available at www.ijraset.com

Artificial Neural Network for Cardiovascular Disease Prediction

Abhay Patil[1], Pallavi Thorat[2], Shreyash Agrawal[3]
[1]*Student, Department of Computer Engg., Zeal College of Engineering and Research, Pune, India*
[2]*Student, Department of Computer Engg., NBN Sinhgad College of Engineering, Pune, India*
[3]*Student, Department of Electronics and Telecommunication Engg, Shrimati Kashibai Navale College of Engineering, Pune, India*

Abstract: ***The assurance of coronary ailment a large part of the time depends upon an eccentric mix of clinical and masochist data. Considering this multifaceted nature, there exists a ton of income among clinical specialists and experts with respect to the useful and careful assumption for coronary sickness. In this paper, we cultivate a coronary disease prediction system that can help clinical specialists in expecting coronary ailment status reliant upon the clinical data of patients. Man-made intelligence-gathering strategies are amazingly useful in the clinical field by giving accurate results and quick finishes of ailments. Thusly, these techniques save part of the ideal opportunity for the two trained professionals and patients. The neural associations can be used as classifiers to expect the assurance of Cardiovascular Heart disorder.***
Keywords: Cardio Vascular disease, Classification, Artificial neural network, Categorical model and Binary model

I. INTRODUCTION

The clinical benefits adventures assemble tremendous proportions of data that contain some restricted information, which is significant for making convincing decisions. Pardoning legitimate results and making practical decisions on the data. In this examination, a Heart Disease Prediction System (HDPS) is made using a Deep Neural Network for foreseeing the presence of coronary sickness. Neural associations have a huge number of traits and these properties sway the introduction of the model. Consequently, it ends up being fundamental to consider their effects whether it is positive and negative. The Cleveland coronary sickness dataset is used to anticipate the finding in this undertaking. This focused on neural associations to gain the best results. In our technique, we use two models be a particularly obvious model and a twofold model. In the full-scale model, we bunch the coronary sickness into five classes and in the equal model, we request in case there is a coronary ailment. There are a couple of models open to predict the presence of coronary ailment using AI systems like Random Forest, SVM, etc. A huge part of the past models can anticipate the presence of coronary ailment. Be that as it may, their precision is low. A portion of the time, their results are incorrect in light of the imbalanced data.

A. Limitations

1) Misleading if there should arise an occurrence of imbalanced information
2) Low exactness

In our proposed model we use the Cleveland dataset which includes 14 clinical characteristics which are ready and attempted to a phony neural association estimation. Neural associations have a colossal number of characteristics, and these properties sway the introduction of the model.

The Cleveland coronary ailment dataset is used to expect the finding in this undertaking. This focused on neural associations to gain the best results. In our strategy, we use two models specifically a straight-out model and a twofold model. In the straight out model, we request the coronary sickness into five classes subject to the earnestness using softmax, and in a twofold model, we describe in case there is a coronary ailment.

B. Benefits

1) Able to determine the seriousness of coronary illness utilizing absolute model.
2) Accuracy is high in the double model.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VIII Aug 2021- Available at www.ijraset.com

II. PROPOSED WORK

The underneath figure beneath depicts the means in building a model for coronary ailment gauge. As exhibited in the above figure, the primary data is recuperated from datasets. The data needs to organize planning. In this way, the data is gone through into preprocessing step. As of now, the model is worked with significant learning calculations. As of now, the model is ready with a dataset. Finally, the model is attempted with a test dataset to know whether the model is gathering heart afflictions or expecting the presence of coronary disease.

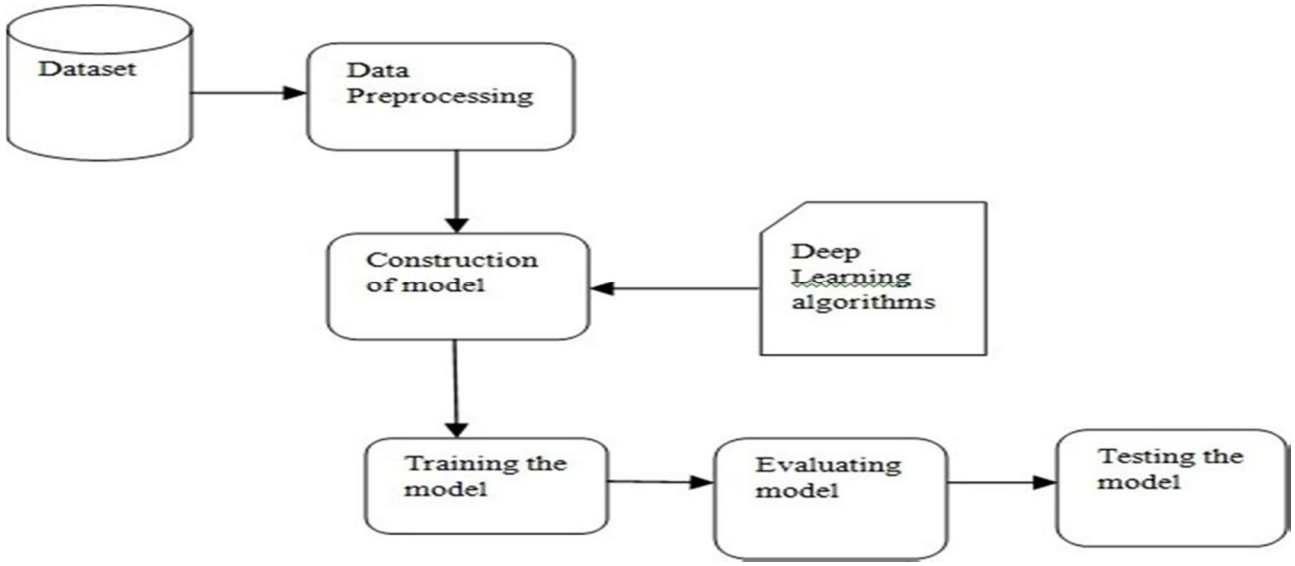


Fig 1: Basic Architecture

A. Dataset

The Cleveland coronary problem dataset is used which was taken from an online UCI AI document. This dataset is utilized for a research study. The dataset has 303 cases and 14 traits.

Table I: Analysis of dataset

Clinical Features	Description
Age	Age
Ca	Number of major vessels (0-3) colored by flourosopy
Chol(mg/dl)	Serum cholesterol
Cp	Chest pain type
Exang	Exercise induced angina
Fbs	Fasting blood sugar
Num	Diagnosis of heart disease
Oldpeak	ST depression induced by exercise relative to rest
Restecg	Resting electrocardiographic results
Sex	Gender
Slope	The slope of the peak exercise ST segment
Thal	3=normal ; 6 = fixed defect; 7= reversible defect
Thalach	Maximum heart rate achieved
Trestbps(mmHg)	Resting Blood Pressure

The dataset contains 8 unmitigated characteristics and 6 numeric characteristics. The above table contains the complete information about the dataset. The age property in the dataset contrasts from 29 to 79. Periodical examinations have created the impression that people who are more unmistakable than 65 are not actually encountering heart issues. The male patient has a sexual direction worth one and the female patient has a sexual direction worth nothing. Male patients are at a high risk of coronary sickness than female patients. It is found that female patients with diabetes will undoubtedly encounter the evil impacts of coronary sickness than that of male patients with diabetics. Ordinary angina, unusual angina, non-angina torture, and asymptotic are the different kinds of chest torture. Angina is the chest torture caused due to the setback of blood that is wealthy in oxygen that is given to the heart. Ordinary angina is caused on account of the diminished circulatory system to the heart muscle. Strange angina is achieved by excited or mental pressing factors. Asymptotic is a not symptom of coronary disease.

B. Algorithm

In the coronary ailment estimate structure, there are input factors, which are ailment risk factors that are gotten from the dataset, and yield factors, which are a grouping, for instance, "ailment nonappearance" and "contamination presence". A figure of heart disease is known as a controlled learning issue. Considering having yield factors are in grouping type, the assumption heart disease is "request kind of coordinated learning". Backpropagation Algorithm which is normally used Fake Neural Network learning technique was used for making coronary ailment assumption structure. Since Backpropagation Algorithm is the single procedure that is used for nonlinear associations which suggests it is the magnificent request computation for coronary sickness assumption.

C. ANN

Artificial neural networks are one of the key gadgets used in AI. As the "neural" a piece of their name proposes, they are frontal cortex propelled structures that are relied upon to copy the way that we individuals learn. Neural networks include data and yield layers, similarly as (all around) a mysterious layer containing units that change the commitment to something that the yield layer can use. They are splendid instruments for seeing models that are far as unnecessarily erratic or different for a human programmer to think and show the machine to see.

While neural networks (likewise called "perceptrons") have been around since the 1940s, it is just over the most recent quite a few years where they have become a significant piece of computerized reasoning. This is because of the appearance of a procedure called "backpropagation," which permits networks to change their secret layers of neurons in circumstances where the result doesn't coordinate with what the maker is expecting — like a network intended to perceive canines, which misidentifies a feline, for instance. Another significant development has been the appearance of profound learning neural networks, in which various layers of a multi-facet network remove various highlights until it can perceive what it is searching for.

-The planned ANN has three layers:

1) Input Layer was expected to contain 13 neurons. Various neurons were picked to be identical to the number of characteristics in the educational assortment.

2) Hidden Layer was expected to contain 3 neurons. This number was picked as a startup point. The number was changed growing independently until it went to the number of neurons of the information layer by differentiating execution of them and a short time later picking the best one. This system relies upon one of AI best practices that the number of neurons of the mysterious layer should be the mean of the number of neurons of data and yield layers.

3) Output Layer was expected to contain 2 neurons. The arranged NN is a classifier going running in Machine Mode which means returning a class name (e.g., "Disease Presence"/"Ailment Absence"). Picking 2 neurons rely upon the prospect that the yield layer has one center for every class name in the model.

A backpropagation Algorithm was used for the proposed structure as a learning computation. 13 of the qualities of the Cleveland dataset were used as information data for the arranged neural association. The dataset was separated into two segments: getting ready and testing. By then, planning was done with the Backpropagation Algorithm. In the wake of setting up the action, the presentation of the proposed structure was enlisted by testing the neural organization with test data by different estimations including exactness, precision, and review.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VIII Aug 2021- Available at www.ijraset.com

D. Data Preprocessing

In Data preprocessing the steps are:

1) Finding All the null values

2) Empty rows are eliminated

a) *Creating Training and Testing Datasets:* Now that we have preprocessed the information suitably, we can split it into training and testing informational collections. We will use Sklearn's train_test_split() function to produce a training dataset (80% of the absolute data)and testing information set(20percentofthe all-out information). Besides, the class values in this dataset contain numerous types of coronary illness with values going from 0 (beneficial) to 4 (severe coronary illness). Consequently, we should change our class information over to unmitigated labels. For instance, the label2 will become [0,0, 1,0,0].

b) *Building and Training the Neural Organization:* Since we have our information completely prepared and split into preparing and testing datasets, we can start fabricating a neural organization to tackle this characterization issue. Utilizing Keras, we will characterize a straightforward neural organization with one secret layer. Since this is a categorical arrangement issue, we will utilize a softmax actuation work in the last layer of our organization and a categorical_cross entropy misfortune during our preparation stage. In spite of the fact that we accomplished promising outcomes, we actually have a genuinely enormous mistake. This could be on the grounds that it is truly challenging to recognize the diverse seriousness levels of coronary illness (classes 1-4).In this progression, the improvement of the issue is finished by changing the information over to a parallel arrangement issue coronary illness.

III. RESULT

A. Categorical Model

In the categorical model, the class esteems in this dataset contain different kinds of coronary illness with values going from 0 (beneficial) to 4 (extreme coronary illness). Thusly, we should change our class information over to categorical marks. For instance, the label2 will become [0,0, 1,0,0]. In the categorical model when we assess the 20% of the dataset which is utilized to test the model, we see that the exactness is 66.67%.

```
Results for Categorical Model
0.6666666666666666
              precision    recall  f1-score   support

           0       0.82      0.92      0.87        36
           1       0.33      0.09      0.14        11
           2       0.00      0.00      0.00         6
           3       0.35      1.00      0.52         6
           4       0.00      0.00      0.00         1

    accuracy                           0.67        60
   macro avg       0.30      0.40      0.31        60
weighted avg       0.59      0.67      0.60        60

E:\anaconda\lib\site-packages\sklearn\metrics\classification.py:1437:
  'precision', 'predicted', average, warn_for)
```

Fig 2: Categorical Model Result

B. Binary Model

Despite the fact that we accomplished promising outcomes, we actually have a genuinely huge blunder. This could be on the grounds that it is undeniably challenging to recognize the diverse seriousness levels of coronary illness (classes 1-4).In this progression, the rearrangements of the issue are finished by changing the information over to a double arrangement issue coronary illness. In the twofold model when we assess the 20% of the dataset which is utilized to test the model, we see that the exactness is 85%.

```
Results for Binary Model
0.85
              precision    recall  f1-score   support

           0       0.90      0.82      0.86        34
           1       0.79      0.88      0.84        26

    accuracy                           0.85        60
   macro avg       0.85      0.85      0.85        60
weighted avg       0.86      0.85      0.85        60
```

Fig 3: Binary Model Result

C. Other Deep Learning Algorithms Results

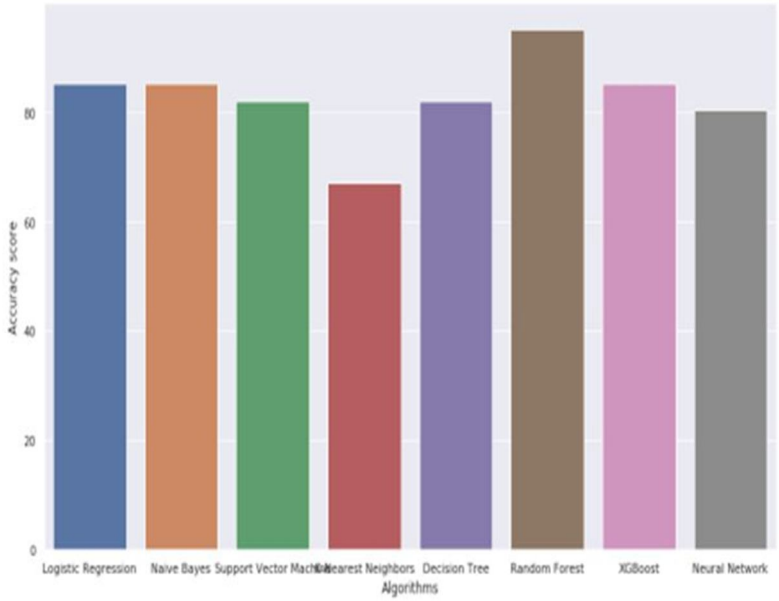


Fig 4: Comparision of different algorithms

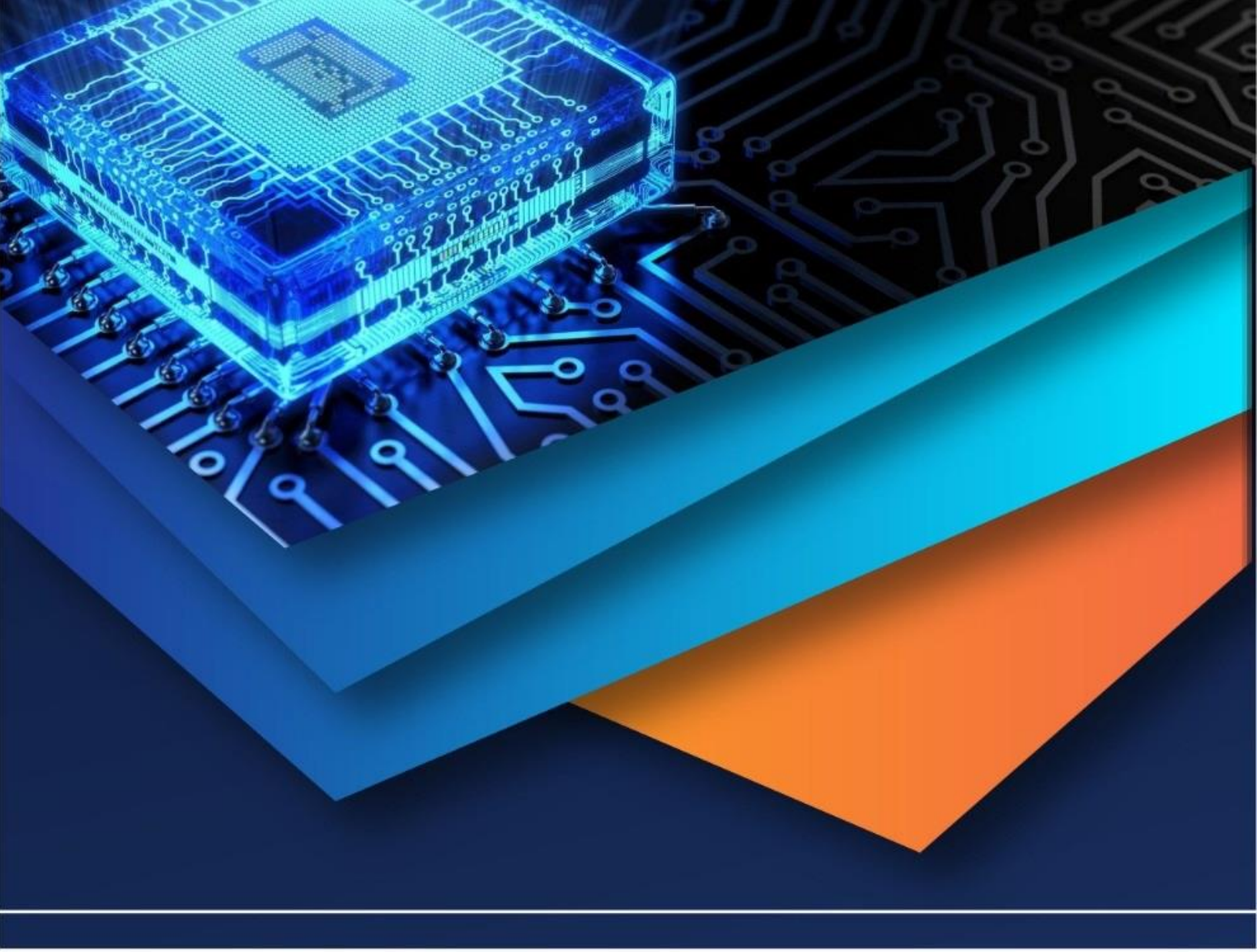
At the point when SVM classifier is utilized for assessing test information, the exactness given is 80.32%, Naive Bayes classifier gives the precision of 78.6%, Logistic Regression gives the exactness of 80.32%, Decision tree gives the precision of 77%, Random forest gives the precision of 95%, Light GBM gives the exactness of 77%, XGBoost gives the precision of 85% and Neural organization with 1 secret layer gives 80.3%. Although Random Forest classifier gives the greatest exactness we picked ANN calculation since it gives great exactness of 85% and it is productive and even works with deficient information.

IV. CONCLUSION

From the proposed system it is assumed that a fake neural organization calculation is best for the portrayal of data from the enormous proportions of clinical data. Extraordinary execution with development in efficiency is obtained from neural organizations when outfitted with normalized data. The data is normalized using a classifier. It is a solid system for the expert's decision. The Artificial neural association is genuinely exceptional for coronary sickness gauges.

REFERENCES

[1] A Swapna, Vedavathi K, ” An Approach to the Identify Human based upon Teeth Recognition using ANN” , 2021
[2] R.Sarumathi, P.Ramprakash,,S.Nithyavishnupriya, “Heart disease prediction using the deep-learning method”, 2020
[3] Tarleand Jena, "ANN Based Pattern Classification Algorithm forDiagnosis of Heart Disease," India, 2017
[4] K. Saxenaand, R.Sharma," Efficient heart disease prediction system using decision tree", India, 2015
[5] Franklin, A.N.Repaka," Design And Implementing Heart Disease Prediction Using Naives Bayesian", 2019



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY