iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VI **Month of publication:** June 2021

DOI: https://doi.org/10.22214/ijraset.2021.35498

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 VI Jun 2021- Available at www.ijraset.com

Prediction of Cardiovascular Disease using Artificial Neural Network

G Naga Sujini[1], Vedavathi K[2]
Asst .Professor, Department of CSE, MGIT

Abstract: ***The determination of coronary illness much of the time relies upon an unpredictable blend of clinical and neurotic information. Inlight of this intricacy, there exists a lot of revenue among clinical experts and analysts in regards to the productive and exact expectation of coronary illness. In this paper, we foster a coronary illness foresee framework that can help clinical experts in anticipating coronary illness status dependent on the clinical information of patients. AI grouping methods are incredibly helpful in the clinical field by giving exact outcomes and fast conclusion of illnesses. Consequently, these procedures save part of time for the two specialists and patients. The neural organizations can be utilized as classifiers anticipate the determination of Cardiovascular Heart sickness.***
Keywords: Classification, Cardio vascular disease, Artificial neural network, Categorical model and Binary model.

I. INTRODUCTION

The medical services ventures gather enormous measures of information that contain some secret data, which is valuable for settling on compelling choices.

For giving proper outcomes and settling on viable choices on information. In this investigation, a Heart Disease Prediction System (HDPS) is created utilizing Deep Neural Network for forese00ing the presence of coronary illness. Neural organizations have enormous number of attributes and these properties impacts the presentation of the model. Thus, it turns out to be vital to consider their belongings whether it is positive and negative. The cleveland coronary illness dataset is utilized to foresee the finding in this task. This zeroed in on neural organizations and their conduct to acquire most ideal outcomes.

In our methodology we utilize two models to be specific clear-cut model and twofold model. In all out model we group the coronary illness into five classes and in parallel model we order if there is coronary illness.

There are a few models accessible to foresee the presence of coronary illness utilizing AI strategies like Random Forest, SVM, and so forth A large portion of the past models can foresee the presence of coronary illness. But, their exactness is low. Some of the time, their outcomes are erroneous because of the imbalanced information.

A. Limitations

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

In our proposed model we utilize the cleveland dataset which comprise of 14 clinical qualities which are prepared and tried to a fake neural organization calculation. Neural organizations have enormous number of attributes, and these properties impacts the presentation of the model.

The Cleveland coronary illness dataset is utilized to anticipate the finding in this task. This zeroed in on neural organizations and their conduct to acquire most ideal outcomes. In our methodology we utilize two models in particular straight out model and twofold model. In straight out model, we order the coronary illness into five classes dependent on the seriousness utilizing softmax and in twofold model we characterize if there is coronary illness.

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 VI Jun 2021- Available at www.ijraset.com

II. PROPOSED WORK

The beneath figure portrays the means in building a model for coronary illness forecast. As demonstrated in above figure, first information is recovered from datasets. The information needs to arrange for preparing. Thus, the information is gone through into preprocessing step. Presently, the model is built with profound learning algorithms (i.e. Ann.)[12]. presently, the model is prepared with dataset. At long last, the model is tried with test dataset to know whether the model is grouping heart sicknesses or anticipating the presence of coronary illness.

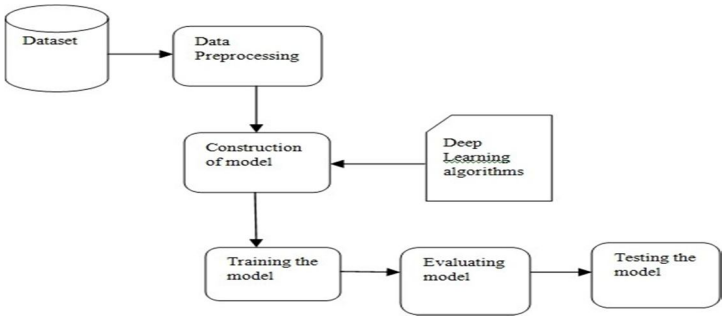


Fig 2.1BasicArchitecture

A. Dataset

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

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 quality and 6 numeric qualities. The above Table contains the total data about the dataset. The age property in the dataset contrasts from 29 to 79. Periodical investigations have appeared that individuals who are more prominent than 65 are not really experiencing heart problem. The male patient has sexual orientation worth of one and female patient has sexual orientation worth of zero. Male patient are at the high danger of coronary illness than that of female patients. It is tracked down that the female patients with diabetes are bound to experience the ill effects of coronary illness than that of male patient with diabetics.

Normal angina, abnormal angina, non-angina torment and asymptotic are the various sorts of chest torment. Angina is the chest torment caused because of the shortfall of blood that is rich in oxygen which is provided the heart. Normal angina is caused because of the decreased blood stream to the heart muscle. Abnormal angina[5] is brought about by enthusiastic or mental pressure. Asymptotic is a not side effect of coronary illness.

TRESTBPS shows the resting circulatory strain worth of an individual the unit of TRESTBPS is mmHg. Serum Cholesterol is the absolute degree of cholesterol accumulated. (LDL)Low-thickness lipoprotein which is named as dreadful cholesterol. Presence of undeniable level of LDL limits the corridors. (HDL) addresses for High thickness lipoprotein which is named as worthy cholesterol. The presence of undeniable degree of HDL diminishes the hazard of cardiovascular failure.
FBS demonstrates the fasting glucose worth of a person. On the off chance that the FBS is under 120mg/dl the worth is 1. On the off chance that the FBS is more than 120mg/dl then the worth allocated to the trait is 0. Not reacting as expected to the insulin emitted prompts the increment in the glucose level which increments the danger of coronary illness. RESTECG shows the resting electrocardiographic outcome the worth is allocated to 0 if the RESTECG is typical, the worth is appointed to 1 if the RESTECG have ST-T wave irregularity. The worth is appointed to 2 if RESTECG have left ventricular hypertrophy.
Most extreme Heart Rate Achieved shows the greatest heartbeat rate accomplished by a person. Expansion in heartbeat rate by 10% expansion the heart passing by in any event 20%. EIA is recorded as 0 if there is no agony and recorded as 1 if there is torment. Angina is typically felt in the focal point of the chest it might even spread to both of the shoulders. The significant thought is the span of ST-section sorrow as the recuperation after the pinnacle pressure results in positive ECG stress test. Slant addresses the incline of the ST section. THAL addresses span in practice test in minutes which shows the thalassemia. ECG stress test is considered as unusual when there is down inclining ST section depression> 1mm at 60-80ms. Target is the class trait it is recorded as zero for non-infected individual and one for the individual enduring with coronary illness.

B. Algorithm

In the coronary illness forecast framework, there are input factors, which are illness hazard factors which are gotten from dataset, and yield factors, which are a classification, for example, "illness nonappearance" and "infection presence". Forecast of heart illness is called regulated learning issue. In view of having yield factors are in classification type, the expectation heart illness is "order sort of directed learning".
Back propagation Algorithm [3] which is usually utilized Fake Neural Network learning method, was utilized for creating coronary illness expectation framework. Since Back propagation Algorithm is the solitary strategy which is utilized for nonlinear connections which implies it is the awesome order calculation for coronary illness expectation.

C. Artificial Neural Network:

Artificial neural networks are one of the principal devices utilized in AI. As the "neural" a piece of their name proposes, they are cerebrum motivated frameworks that are expected to imitate the way that we people learn. Neural networks comprise of information and yield layers, just as (by and large) a secret layer comprising of units that change the contribution to something that the yield layer can utilize. They are brilliant instruments for seeing examples that are far as excessively unpredictable or various for a human software engineer to concentrate and show the machine to perceive.
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 "back propagation[1]," 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 proposed coronary illness forecast framework was planned as a multi-facet perceptron neural network. The planned ANN has three layers: to be specific an info layer, a secret layer and a yield layer.

1) *Input Layer* was intended to contain 13 neurons. Number of neurons was chosen to be equivalent to the number of qualities in the informational collection.
2) *Hidden Layer* was intended to contain 3 neurons. This number was chosen as a startup point. The number was changed expanding individually until it came to the quantity of neurons of the info layer by contrasting execution of them and afterward choosing the best one. This methodology depends on one of machine learning best practices that the quantity of neurons of secret layer ought to be the mean of the quantity of the neurons of information and yield layers.
3) *Output Layer* was intended to contain 2 neurons. The planned NN is a classifier going running in Machine Mode which means returning a class name (e.g., "Infection Presence"/"Illness Absence"). Choosing 2 neurons depends on thought that the yield layer has one hub per class name in model.

Back propagation Algorithm was utilized for the proposed framework as learning calculation. 13 of the characteristics of Cleveland dataset was utilized as info information for the planned neural organization.

The dataset was parted into two sections: preparing and testing. At that point preparing was finished with Back propagation Algorithm. In the wake of preparing measure, the exhibition of the proposed framework was registered by testing the neural network with test information by various measurements including precision, accuracy and recall.

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 data appropriately, we can splitit into training and testing data sets. We will use Sklearn's train_test_split() function [3]to generate a training dataset (80 percent of the total data)and testing data set(20percentofthe total data).Furthermore, the class values in this dataset contain multiple types of heart disease with values ranging from 0 (healthy) to 4 (severe heart disease). Consequently, we will need to convert our class data to categorical labels. For example, the label2 will become [0,0, 1,0,0].

b) *Building and Training the neural network:* Now that we have our data fully processed and split into training and testing datasets, we can begin building a neural network to solve this classification problem. Using keras, we will define a simple neural network with one hidden layer. Since this is a categorical classification problem, we will use a soft max activation function in the final layer of our network and a categorical_cross entropy loss [2] during our training phase. Although we achieved promising results, we still have a fairly large error. This could be because it is very difficult to distinguish between the different severity levels of heart disease (classes 1-4).In this step the simplification of the problem done by converting the data to a binary classification problem-heart disease .

III. RESULTS

In the proposed heart disease prediction system model the performance evaluation is done by using measurements like precision, recall, accuracy ,f1-score and support.

A. Categorical Model

In the categorical model the class values in this dataset contain multiple types of heart disease with values ranging from 0 (healthy) to 4 (severe heart disease). Consequently, we will need to convert our class data to categorical labels. For example, the label2 will become [0,0, 1,0,0].

In categorical model when we evaluate the 20% of the dataset which is used to test the model we see that the accuracy 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 3.1Results for categorical model

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

B. Binary model

Although we achieved promising results, we still have a fairly large error. This could be because it is very difficult to distinguish between the different severity levels of heart disease (classes 1-4).In this step the simplification of the problem done by converting the data to a binary classification problem-heart disease .

In binary model when we evaluate the 20% of the dataset which is used to test the model, we see that the accuracy 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: 3.2 Result for binary model

C. Results with other Deep Learning Algorithms

While the test data is subjected for evaluation using other deep learning algorithms we got the following results:

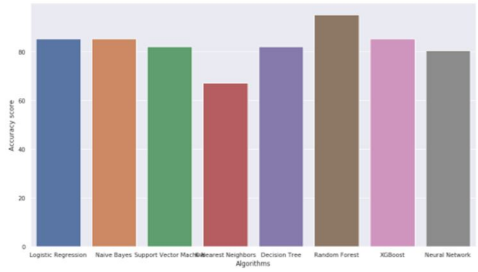


Fig 3.3: Comparison of accuracies

When SVM classifier is used for evaluating test data the accuracy given is 80.32%, Naive Bayes classifier gives the accuracy of 78.6%, Logistic Regression gives the accuracy of 80.32%, Decision tree gives the accuracy of 77%, Random forest gives the accuracy of 95%, Light GBM gives the accuracy of 77%, XGBoost gives accuracy of 85% and Neural network with 1 hidden layer gives 80.3%.Although Random Forest classifier gives maximum accuracy we chose ANN algorithm because it gives good accuracy of 85% and it is efficient and even works with incomplete data.

IV. CONCLUSION

From the proposed framework it is presumed that artificial neural network algorithm is best for characterization of information from huge measure of clinical information. Great execution with expansion in productivity is acquired from neural network when furnished with standardized information. The information is standardized utilizing a classifier. It is strong framework to the specialist's choice. The Artificial neural organization is truly outstanding for coronary illness forecast.

V. FUTURE WORK

The proposed framework is easy to use, versatile, dependable and an expandable framework. The proposed working model can likewise help in decreasing treatment costs by giving Initial diagnostics on schedule. The model can likewise effectively train apparatus for clinical understudies and will be a delicate demonstrative instrument accessible for doctor and cardiologist. General doctors can use this apparatus for starting finding of cardio-patients. There are numerous potential enhancements that could be investigated to improve the versatility and exactness of this forecast framework. As we, have fostered a summed up framework, in future we can utilize this framework for the examination of various informational indexes. The presentation of the wellbeing's conclusion can be improved altogether by taking care of various class names in the expectation interaction, and it tends to be another positive heading of examination.

REFERENCES

[1] A.N.Repaka,S.D.RavikantiandR.G.Franklin," Design And Implementing Heart Disease Prediction Using Naives Bayesian," 2019 3rd International Conference on Trends in Electronics and Informatics(ICOEI),Tirunelveli,India,2019,pp.292-297,doi:10.1109/ICOEI.2019.8862604.

[2] Purushottam,K. Saxenaand R.Sharma," Efficient heart disease prediction system using decision tree," International Conference on Computing, Communication & Automation, Greater Noida, India, 2015,pp.72-77, doi:10.1109/CCAA.2015.7148346.

[3] Javeed, S. Zhou, L. Yongjian, I. Qasim, A. Noor and R. Nour, "An Intelligent Learning System Based on Random Search Algorithm and Optimized Random Forest Model for Improved Heart DiseaseDetection,"inIEEEAccess,vol.7,pp.180235-180243,2019,doi:10.1109/ACCESS.2019.2952107.

[4] H. Chen, S. Y. Huang, P. S. Hong, C. H. Cheng and E. J. Lin, "HDPS: Heart disease prediction system,"2011ComputinginCardiology, Hangzhou,China,2011,pp. 557-560.

[5] Z. Wang, J. Chen, Z. Zheng and J. Wang, "Sample expansion-oriented LDBN heart disease riskforecast model," 2018 14th International Wireless Communications & Mobile Computing Conference(IWCMC),Limassol,2018,pp.13271332,doi:10.1109/IWCMC.2018.8450376.

[6] Tarleand S. Jena, "An Artificial Neural Network Based Pattern Classification Algorithm forDiagnosis of Heart Disease," 2017 International Conference on Computing, Communication, Control andAutomation(ICCUBEA),Pune,India,2017,pp.1-4,doi:10.1109/ICCUBEA.2017.8463729.

[7] C.Peng,C. -W.HuangandY. -C.Lai, "Heart Disease Prediction Using ArtificialNeuralNetworks:ASurvey,"2020IEEE2ndEurasiaConferenceonBiomedicalEngineering,HealthcareandSustainability(ECBIOS),Tainan,Taiwan,2020,pp .147-150,doi:10.1109/ECBIOS50299.2020.9203604.

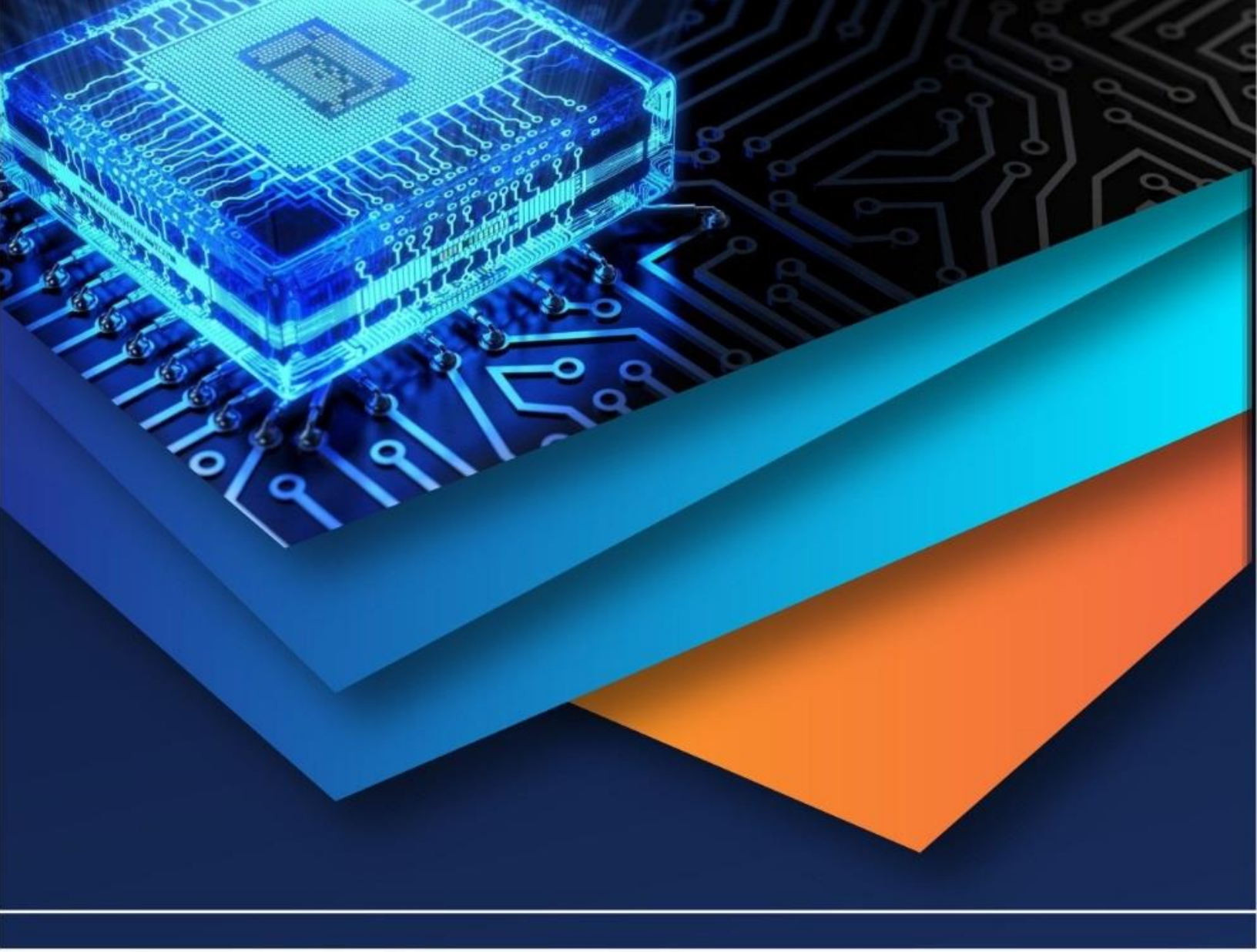
[8] R. R. Ema and P. C. Shill, "Integration of Fuzzy C-Means and Artificial Neural Network withPrinciple Component Analysis for Heart Disease Prediction," 2020 11th InternationalConference onComputing, Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2020, pp. 1-6,doi:10.1109/ICCCNT49239.2020.9225366.

[9] M. Rani, A. Bakshi, K. Munjal and A. S. Tomar, "Heart Disease Prediction Using Artificial NeuralNetworkand Image Processing,"2020 International Conference on Emerging Smart Computing andInformatics(ESCI),Pune,India,2020,pp.203-208,doi:10.1109/ESCI48226.2020.9167669.

[10] P.Ramprakash, R.Sarumathi, R.Mowriya and S. Nithyavishnupriya, "Heart Disease Prediction Using Deep Neural Network," 2020 International Conference on Inventive Computation Technologies (ICICT), 2020, pp. 666-670, doi: 10.1109/ICICT48043.2020.9112443.

[11] T. Karayılan and Ö. Kılıç, "Prediction of heart disease using neural network," 2017 International Conference on Computer Science and Engineering (UBMK), 2017, pp. 719-723, doi: 10.1109/UBMK.2017.8093512.

[12] Vedavathi K and A Swapna" An Approach to Identify Human based upon Teeth Recognition using ANN" Turkish Journal of Computer and Mathematics Education Vol.12 No.9 (2021).



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY