ADVANCED AND SMART HEART DISEASE PREDICTION USING HYBRID MACHINE LEARNING TECHNIQUES

Miss. Srushti N. Raut
Department of Computer Engineering
Zeal College of Engineering and Research
Researchsrushti.raut25@gmail.com

Prof. Vikas M. Patil
Department of Computer Engineering Zeal
Zeal College of Engineering and
vikas.patil@zealeducation.com

ABSTRACT

Heart disease is one of the leading cause of death globally because of change in lifestyle of human being.90% of cardiovascular disease can be prevented. Expectation of cardiovascular infection is a basic test in the region of clinical information examination. Health care fields have a vast amount of data, for processing those data certain technologies are used. Machine learning strategies are one of the effective for prediction. Different investigations give just a look into anticipating heart ailment with ML systems. predict the cause and disease is one of the major challenge now a days. In this paper, we propose a novel strategy that targets finding huge includes by applying AI systems bringing about improving the exactness in the forecast of cardiovascular disease. ANN is one of the ML procedures which can be utilized to accomplish productive conclusion results and also we are using hybrid machine learning techniques for prediction.

Keywords: Artificial Neural Network, SVM, Heart Disease, Machine Learning.

I. INTRODUCTION

Heart disease is the leading Causes of death in the world over the past ten years. It is must that diagnosing system must generate accurate results that proper treatment can be available for the patient. In earlier systems various techniques was used for predicting the reason behind the disease which lacks in accuracy of the output. Various heart diseases like heart attacks, chest pain etc. This thesis is based on the heart disease diagnosis of patients. Heart disease is a prevailing disease nowadays. Now due to increasing expenses of heart disease, there was a need to develop a new system which can predict heart diseases in an easy and cheaper way. The proposed work divides proposed system in two parts such as performance model and prediction model. Performance model is designed to evaluate the overall performance of the application. Prediction Model is used to predict the condition of the patient after evaluation on the basis of various parameters like heart beat rate, blood pressure, cholesterol etc. The accuracy of the system is proved in python. A significant test confronting social insurance associations and businesses (clinics, therapeutic testing focus) is the arrangement of value administrations at reasonable expenses and Quality assistance infers diagnosing patients accurately also, directing

medicines that are successful. In the Poor clinical choices can prompt lamentable outcomes which are accordingly unsuitable. Emergency clinics should likewise limit the expense of clinical tests and reports. They can accomplish these outcomes by utilizing proper PC based data or potentially choice emotionally supportive networks. Diagnosing is a procedure which is finished via robotized frameworks or machines in human framework to discover the idea of the infection by observing the different side effects of the ailment. Diagnosing is a most confused assignment to perform. Diagnosing machines or frameworks hush up accommodating in this procedure in light of the fact that not each specialist must have the

information on every single sort of issue of ailment. In this diagnosing physically by specialists can prompt off base outcomes a few times. Accordingly a mechanized diagnosing machine is utilized by them to analyze the issue precisely. The WHO consortium has shared this data that ten a large number of passing's happen in this world is a direct result of heart disease. so it was a very risks issue in world. These frameworks regularly produce colossal measures of information which appear as numbers, diagrams and pictures. Tragically, these information are seldom used to help clinical basic leadership.

II. LITERATURESURVEY

Senthilkumar Mohan ,Chandrasegar Thirumala and Gautam Srivastava[1] In this paper, they proposed a novel method that aims at finding significant features by applying machine learning techniques resulting in improving the accuracy in the prediction of cardiovascular disease. The prediction model is introduced with different combinations of features and several known classification techniques. They produced an enhanced performance level with an accuracy level of 88.7% through the prediction model for heart disease with the hybrid random forest with a linear model (HRFLM). In HRFLM, they used a computational approach with the three association rules of mining namely, apriori, predictive and Tertius to find the factors of heart disease on the UCI Cleveland dataset. HRFLM makes use of ANN with back propagation along with 13 clinical features as the input. The obtained results are comparatively analyzed against traditional methods**.**

Mamatha Alex P and Shaicy P Shaji [2] state that we are living in a post present day time and there are colossal changes happening to our day by day schedules which make an effect on our wellbeing emphatically and adversely. Because of these progressions different sort of sicknesses are gigantically expanded. Particularly, coronary illness has become increasingly normal these days. The life of individuals is at a hazard. Variety in Blood pressure, sugar, and beat rate and so on can prompt cardiovascular illnesses that incorporate limited or blocked veins. It might cause Heart disappointment, Aneurysm, Peripheral vein infection, Heart assault, Stroke and even abrupt heart failure. Numerous types of heart illness can be distinguished or determined to have distinctive therapeutic tests by thinking about family therapeutic history and different variables. Be that as it may, the forecast of heart maladies without doing any therapeutic tests is very troublesome.

Jyotismita Talukdar, Bhupesh Kumar Dewangan [3] an investigation has been made on the probability and precision of early forecast of a few Coronary illness utilizing Artificial Neural Network. (ANN).The study has been made in both clamor free and uproarious condition. The information gathered for this examination is from five Hospitals. Around 1500 heart patient's information has been gathered and examined. The information is broke down and the results have been contrasted and the Doctor's determination. It is discovered that, in clamor free condition, the precision changes from 74%
to 92%.and in loud condition (2dB),the consequences of precision fluctuates from 62% to 82%. In the present examination, four essential traits considered are Blood Pressure (BP), Fasting Blood Sugar (FBS), Thallic (THAL) and Cholesterol (CHOL).

Balasaheb Tarle and Sudarson Jena [4] proposed that definitive point of the proposed strategy is to build up a model for order of medicinal information. In this system we present ANN for Diagnosis of Heart infection. ANN works like the neural course of action of mind. The mind examines from quite a while ago encounters. It can tackle the issues that are not process able or resolvable by current processing frameworks. Highlight determination is most significant procedure for choosing a pertinent property among the colossal informational collection. This decreased arrangement of ascribes is then passed to fake neural system. ANN utilizes back engendering calculation for better expectation of result, likewise the ANN depends on the information this preferred position can be taken for increasingly precise outcome.

Dr.C.Sowmiya and Dr.P.Sumitra [5] introducing Heart disease is the main issue for world. Heart diseases more than individuals passing's happen during the principal coronary failure .But not just for cardiovascular failure have a few issues assaulted for bosom malignant growth, lung disease, ventricle. Valve, and so forth… It is basic to have a casing work that can usefully perceive the commonness of coronary illness in thousands of tests in timeously. In this paper the capability of nine (9) characterization strategies was assessed of forecast of coronary illness.

Tahira Mahboob et.al [6] proposed Heart maladies may maybe outcome in debility, serious issue, and small nature of life expectancy. Moreover, it could likewise be deadly. Henceforth gathering coronary illness has transform into first misery as of now. This paper focuses on different machine learning rehearses which help determining and seeing countless heart maladies. Diverse AI approaches chatted here are Concealed Markov Models, Support Vector Machine, Highlight Selection, Computational smart classifier, expectation framework, information mining procedures what's more, hereditary calculation. Investigating each approach altogether enabled us to choose most apt one.

III. PROBLEM STATEMENT AND GOALS

Heart disease can be managed effectively with a combination of lifestyle changes, medicine and in some cases, surgery. With a right treatment, the symptoms of heart disease can be reduced and the functioning of the heart improved. The early and accurate prediction result of heart disease is used to prevent it and thus reduce cost for surgical treatment. So we are giving more accurate solution by using SVM classifier and artificial neural network (ANN) To detect heart diseases and make predictions using hybrid approach.

1. Main objective behind to develop a system helps the doctors to cross verify their diagnosed results which gives promising solution over existing death rates.

2. by using our proposed work try to invent unique platform and most promising solution for early diagnosis of heart diseases.

3. Existing work analysis accuracy is reduced when the quality of medical data is incomplete. Moreover, different regions exhibit unique characteristics of certain regional diseases, which may weaken the prediction of heart disease wrong.

4. So we are giving more accurate solution by using SVM classifier and artificial neural network (ANN) to detect heart diseases and make predictions using hybrid approach.

IV. METHODOLOGY USED IN PROPOSED SYSTEM

This proposed system includes a novel heart detection and Stage prediction mechanism hybrid machine learning Approaches. Proposed which initially adapts profound highlights and afterward prepares a fake neural system with these scholarly highlights. Exploratory outcomes show the profound educated classifier beats every single other classifier when prepared with all qualities and same preparing tests. It is additionally shown that the presentation improvement is factually huge. Characterization of heart disease utilizing a low populace, high dimensional informational index is trying because of lacking examples to become familiar with an exact mapping among highlights and class names. In Proposed system consist of hybrid learning approaches. Current writing for the most part handles this undertaking through high quality element creation and choice. Deep learning is found to be able to identify the underlying structure of data through the use of ANN and other techniques. This shows that application of machine learning has the potential to significantly detect and classify with almost high accuracy for the low population in India .High dimensional heart disease image data-set without requiring any hand-crafted, case specific features.
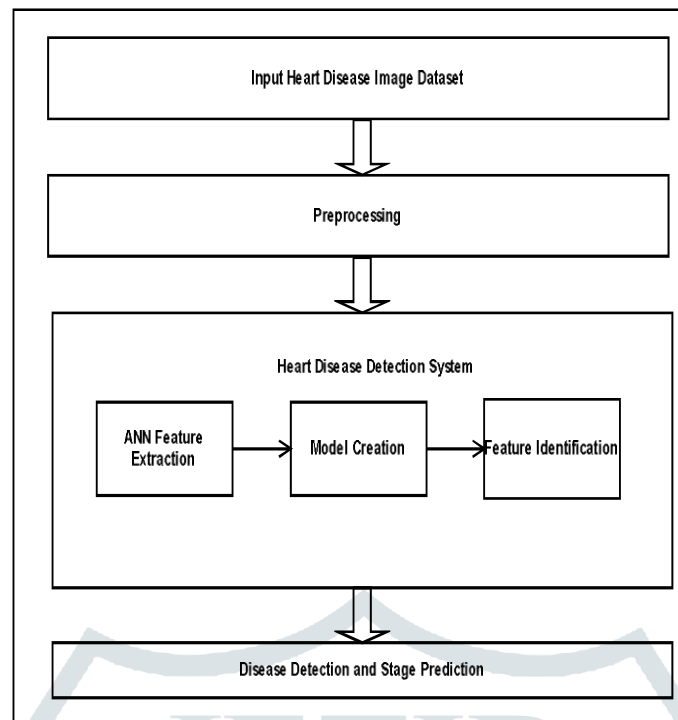
Fig. System Architecture

In ANN model consist of following stages

•   **Image Processing:-**

An image is made up of RGB colors. Pre-processing unit consists of noise removal, gray scale conversion, and binary conversion of images followed by feature extraction. In future extraction five steps followed in which fingertips searches by eccentricity. Next Elongations of images are measured by considering pixel segmentation as well as rotation of images.

•   **Image Filtering:-**

Filtering is a technique to modify or enhance the image, i.e. to highlight certain features or remove other features. It includes smoothing, sharpening and edge enhancement. Image filtering algorithms generate an output pixel by observing the neighborhood of the input pixel in an image. Image filtering algorithms are used to remove different types of noise from the image.

•   **Feature Extraction:-**

In feature extraction, algorithmic study used to find the feature vectors of systematic results combines K curvature and convex hull algorithms. In present work "K convex hull" algorithm which is used to detect fingertip with greater accuracy. In our system, ANN is used for future recognition in which we having the input unit of training data set of images.

•   **Segmentation:-**

Image segmentation is the way toward apportioning an advancedpicture into various portions (sets of pixels). All pixels in an area share a typical property. Least complex property that pixel can share power. The objective is to disentangle and change the portrayal of the picture into something that is increasingly important and less demanding to break down.

•   **Edge Detection:-**

Edge defines the boundaries between regions in an image which helps in object detection. There are many edge detection operators and algorithms available. Edge Detection Operators and Algorithms used in our research like convex hull method.

•

The Support Vector Machine with RBF Kernel:

SVM can be utilized as a relapse show, keeping up all the primary highlights that contribute to maximal edge. SVM utilizes indistinguishable standards from the traditional SVM for grouping, with just a couple of minor changes. As a matter of first importance, since yield is a genuine number it turns out to be exceptionally hard to anticipate the current data, which has limitless conceivable outcomes. On account of relapse, an edge of resistance (epsilon) is set in estimation to the SVM which would have effectively asked for from the issue. Yet, other than this reality, there is likewise a more confused reason, the calculation is more convoluted in this manner to be taken in thought. Notwithstanding, the principle thought is dependably the same: to limit mistake, individualizing the hyperplane which augments the edge, remembering that piece of the blunder is endured.

Kernel functions:

$$k\left(xi,xj\right)=\left(xi,xj\right)^{d} \qquad \text{For Polynomial}$$

$$k\left(xi,xj\right)=\exp\left(-\frac{|xi-xj|^2}{2\sigma^2}\right) \text{ For Gaussian RBF}$$

SVM with ANN Kernel

If our features contain nonlinear i.e. non-discriminative data, then classification problem become more complex. To tackle this problem our kernel functions transform the data into a higher dimensional feature space to make it possible to perform the linear separation.

$$y=\sum_{i=1}^{N}(\alpha_i-\alpha_i^*)g[f(x_i),f(x)]+b$$

(6)

$$y=\sum_{i=1}^{N}(\alpha_i-\alpha_i^*).K(x_i,x)+b$$
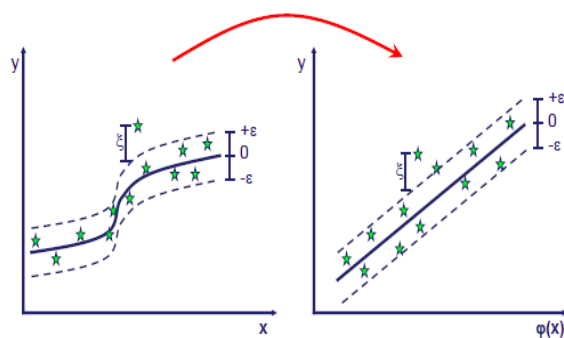
(7)

Figure : *In case of non-discriminative feature.*

.

MATHEMATICAL MODEL

- System Description: S= I, F, O S be the heart disease detection and stage prediction system.

INPUT:

- F=F1,F2,F3...FN Functions to be execute modules
- I=C1,C2,C3... input of systems heart images
- O=R1,R2Rn prediction results in the form of stages
- I=result access by User

F:

- F1=Image processing applied on heart images
- F2=feature extraction from image
- F3=Stage detection results

O:

- R1= model creation from training.
- R2= model based image testing

Success:
1. High accuracy achieved by using all type of CT image dataset.
2. User gets result very fast according to their needs.

Failures:
1. Huge database can lead to more time consumption to get the information.
2. Hardware failure.
3. Software failure.

PSEUDO CODE:

```
Data    : Dataset with p* variables and binary outcome.
Output: Ranked list of variables according to their relevance.

Find the optimal values for the tuning parameters of the SVM model;
Train the SVM model;
p ← p*;
while p ≥ 2 do
    SVMp ← SVM with the optimized tuning parameters for the p variables and
    observations in Data;
    wp ← calculate weight vector of the SVMp (wp1,...,wpp);
    rank.criteria ← (w²p1,...,w²pp);
    min.rank.criteria ← variable with lowest value in rank.criteria vector;
    Remove min.rank.criteria from Data;
    Rankp ← min.rank.criteria;
    p ← p - 1 ;
end
Rank1 ← variable in Data ∉ (Rank2,...,Rankp*);
return (Rank1,...,Rankp*)
```

V. RESULTS AND DISCUSSION

We get prediction result in the form of prediction score matched with every prediction. We got final result basis of higher prediction score of test images with testing dataset to training.

From the below table, it is shown that the SVM-ANN Hybrid classifier gives the Highest Accuracy, Specificity and Sensitivity.

Models	Accuracy	Classification Error	Precision	F-measure	Sensitivity	Specificity
Naïve Bayes	75.8	24.2	90.5	84.5	79.8	60.0
Generalized Linear Model	85.1	14.9	88.8	91.6	94.9	20.0
Logistic Regression	82.9	17.1	89.6	90.2	91.1	25.0
Deep Learning	87.4	12.6	90.7	92.6	95	33.3
Decision Tree	85	15.0	86.0	91.8	98.8	0.0
Random Forest	86.1	13.9	87.1	92.4	98.8	10.0
Gradient Boosted Tress	78.3	21.7	94.1	86.8	80.7	60.0

Support Vector Machine	86.1	13.9	86.1	92.5	100	0.0
VOTE	87.41	12.59	90.2	84.4	-	-
HRFLM	88.4	11.6	90.1	90	92.8	82.6
ANN + SVM (Proposed)	91. 3	26.72	94.7	93.1	90.02	87.3

Result Screenshots:
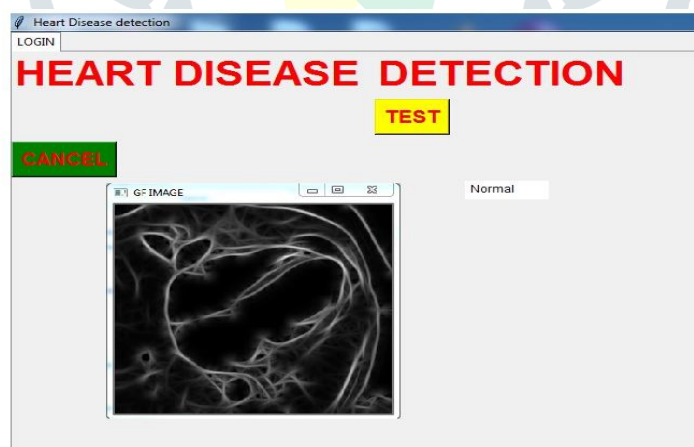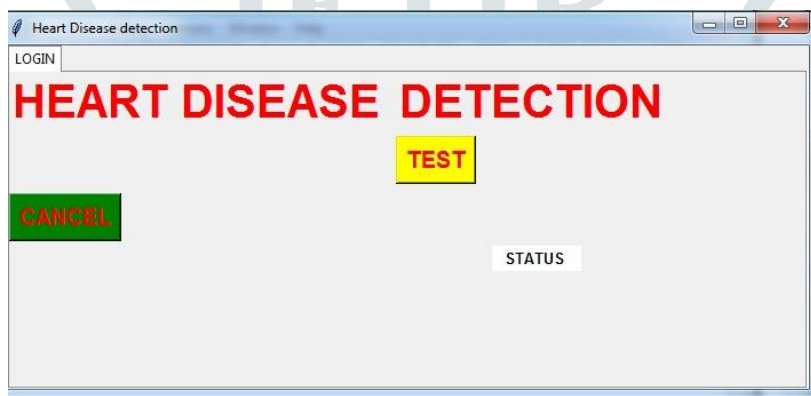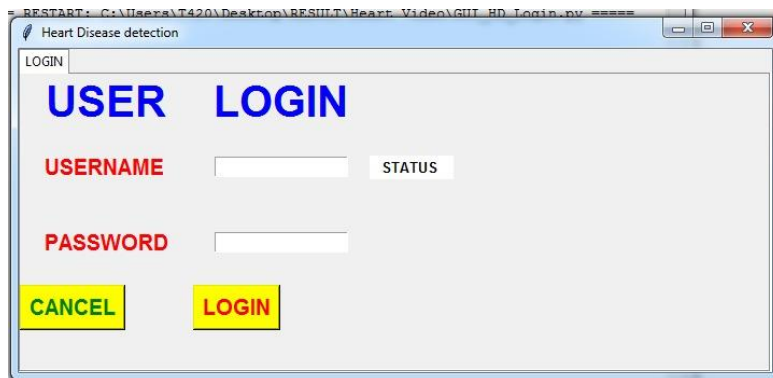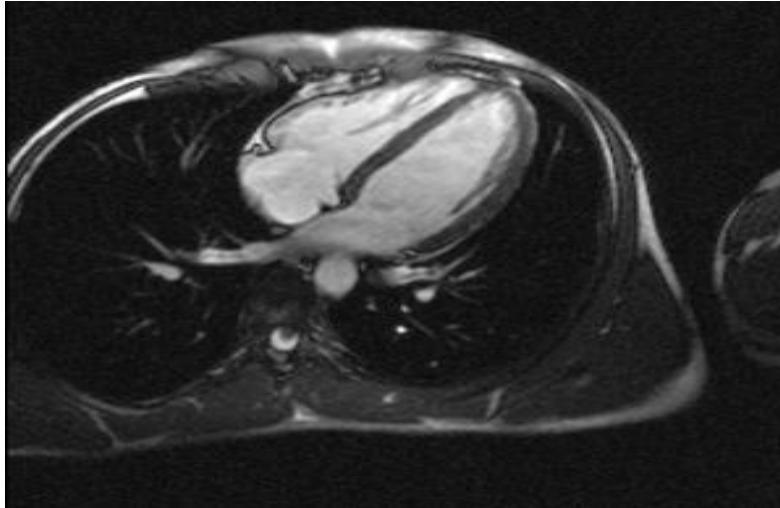
VI. CONCLUSION

In this technique, we have anticipated coronary illness analysis in view of ANN; standardization is done on informational collection. This dataset is passed to BPNN for preparing, testing and grouping .The proposed framework with ANN and five- overlap cross approval gives 83% grouping exactness. We trust this might be contribute towards reliable and snappier programmed mind bug supposition framework, distinctive bug might be dealt with by this framework adequately if extra examination will empowered in future.

Future work will be based on real time heart disease image dataset with considering high accuracy over proposed work.

VII.   REFERENCES

[1] Senthilkumar Mohan ,Chandrasegar Thirumala and Gautam Srivastava,(Member, IEEE) "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques"June 19,2019,IEEE

[2] Mamatha Alex P and Shaicy P Shaji "Prediction and Diagnosis of Heart Disease Patients using Data Mining Technique" International Conference on Communication and Signal Processing, April 4-6, 2019, India

[3] Jyotismita Talukdar, Bhupesh Kumar Dewangan, "Analysis of cardiovascular diseases using artificial neural network" 5th IEEE International Conference on Parallel, Distributed and Grid Computing(PDGC-2018), 20-22 Dec, 2018, Solan, India

[4] Balasaheb Tarle and Sudarson Jena "An Artificial Neural Network Based Pattern Classification Algorithm for Diagnosis of Heart Disease" 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA) 2017

[5] C.Sowmiya and Dr.P.Sumitra, "Analytical Study of Heart Disease Diagnosis Using Classification Techniques." 2017 IEEE INTERNATIONAL CONFERENCE ON INTELLIGENT TECHNIQUES IN CONTROL, OPTIMIZATION AND SIGNAL PROCESSING

[6] Abbas Khosravi, Amin Khatami, "Lung cancer classification using deep learned features on low population dataset." Canadian Conference on Electrical and Computer Engineering (CCECE) 978-1-5090-5538-8/17 2017 IEEE.