A Hybrid Pre-Processing Technique for Stacking Ensemble with Random Forest as a Meta Classifier for Heart Disease Classification

Shakowat Zaman Sarker[a], Sheikh Farhana Binte Ahmed[a*], Ali Ahmed Ave[a], Tahsin Abrar Nabil[a]

[a]*Uttara University, Uttara, Dhaka 1230, Bangladesh*

Abstract

Heart disease is one of the widespread causes of human mortality, and it is important to regularly screen the condition of this organ. Electrocardiography is a popular tool to visualize heart activity, and this data can be used to differentiate a healthy versus a detrimental heart using machine learning techniques. This paper uses the Statlog heart dataset from the UCI repository to test the machine learning classification algorithms. The stacking ensemble classifier outperforms by giving 94.118% accuracy as the data is pre-processed using MinMax technique; while the algorithms KNN, GNB, LR, RF, SVM, DT, Bagging, AdaBoost, Gradient Boosting, and XGBoost are used as base learners and Random Forest is used as the meta classifier. It was cross-validated having almost similar accuracy. Also, different evaluation metrics as sensitivity, specificity, precision, logloss, coefficient of determination, root mean square error and computation time.

Keywords: heart; ECG; stacking ensemble; machine learning; minmax; PCA

1. Introduction

According to World Health Organization (WHO), one person dies from cardiovascular disease every 36 seconds [1]. Heart disease is easier to treat when detected early. Healthcare professionals' quality service is not affordable to many people. Also, getting their appointment is time-consuming. In this condition, automated classification using machine learning can become life-saving.

In order for the heart to do its job of pumping blood to the lungs and to the body, nutrients and oxygen must be supplied to the cells of the heart. The heart also needs to coordinate its contractions so that all parts are working together to pump blood effectively. To understand how all of this works together to give the heart its ability to pump blood, we will examine three interdependent aspects of heart function.

The heart pumps blood to two distinct but linked circulatory systems: the pulmonary and systemic circuits. The pulmonary circuit transports blood to and from the lungs, picking up oxygen and removing carbon dioxide. The systemic circuit transports freshly oxygenated blood to virtually all of the body's tissues and returns relatively deoxygenated blood and carbon dioxide to the heart to be sent back to pulmonary circulation. Figure 1 shows the different parts of a heart. The top panel shows the human heart with the arteries and veins labeled (from top, clockwise): aorta, left pulmonary arteries, pulmonary trunk, left atrium, left pulmonary veins, aortic semilunar valve, mitral valve, left ventricle, inferior vena cava, right ventricle, tricuspid valve, right atrium, pulmonary semilunar valve, right pulmonary veins, right pulmonary arteries, superior vena cava. The bottom panel shows a rough map of the the human circulatory system. Labels read (from top, clockwise): systemic capillaries of upper body, systemic arteries to upper body, pulmonary trunk, left atrium, left ventricle, systemic arteries to lower body, systemic capillaries of lower body, systemic veins from lower body, right ventricle, right atrium, pulmonary capillaries in lungs, systemic veins from upper body.

* Corresponding author

Email address: farhana@uttarauniversity.edu.bd

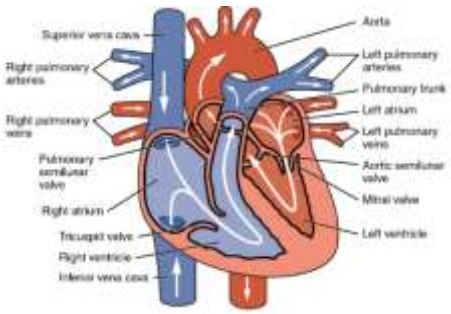


Fig. 1. The physiology of a heart [2]

The process of pumping and circulating blood is active, coordinated and rhythmic. Each heart beat represents one cycle of the heart receiving blood and ejecting blood. Diastole is the portion of the cycle in which the heart is relaxed and the atria and ventricles are filling with blood. The AV valves are open, so that blood can move from the atria to the ventricles. Systole is the portion of the cycle in which the heart contracts, AV valves slam shut, and the ventricles eject blood to the lungs and to the body through the open semilunar valves. Once this phase ends, the semilunar valves close, in preparation for another filling phase. An electrocardiogram records the electrical signals in the heart. It's a common and painless test used to quickly detect heart problems and monitor the heart's health. Figure 2 shows a typical ECG curve.

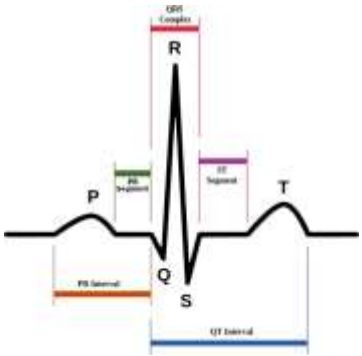


Fig. 2. The ECG curve

A lot of work [3] has been done by machine learning algorithms on detecting and classifying heart diseases. Machine learning is a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy. The three machine learning types are shown in Figure 3.

Ensemble learning is a general meta approach to machine learning that seeks better predictive performance by combining the predictions from multiple models.The three ensemble learning types are shown in Figure 4. Bagging involves fitting many decision trees on different samples of the same dataset and averaging the predictions. Stacking involves fitting many different models types on the same data and using another model to learn how to best combine the predictions. Boosting involves adding ensemble members sequentially that correct the predictions made by prior models and outputs a weighted average of the predictions.

This paper aims to explore the stacking ensemble method with different pre-processing techniques and compare the accuracy, precision, recall, sensitivity, coefficient of determination and others. In section 2, the literature review is mentioned, in section 3, the materials and methods is described. In section 4, the experimental results are mentioned whereas section 5 draws the conclusion.

2. Literature Review

Numerous studies have demonstrated the effective application of ML models in detecting heart diseases. The UCI Heart Disease Dataset from UCI Machine Learning Repository is open to the public and is one of the most used datasets in this research area [4]. The Statlog dataset is also widely used [5]. In the clinical detection of diseases, such ML models aim to improve accuracy and reduce the total cost of the computation.

In [6], Particle swarm optimization based stacked sparse autoencoder was done on Framingham and Cleveland heart disease dataset. Many works are being done on feature learning. In [7], proposes an ensemble framework based on stacking model fusion, from Support Vector. In [8], SVM, K-Nearest Neighbor (KNN), Logistic Regression (LR), Random Forest (RF), Extra Tree (ET), Gradient Boosting Decision Tree (GBDT), XGBoost, LightGBM, CatBoost, and Multilayer Perceptron (MLP) (10 classifiers to select the optimal base learners). In order to avoid the overfitting phenomenon generated by the base learners, Logistic Regression (LR) was used as the meta learner. In [6], In this study, a two level stacking based model is designed in which level 1 is base-level and level 2 is metalevel. The predictions of base-level classifiers are selected as the input of meta-level. The Pearson correlation coefficient and maximum information coefficient are first calculated to find the classifier with the lowest correlation. Then an enumeration algorithm is used to find the best combining classifiers which acquire the best result in the end. The limitation is that, the model parameters were not optimal. In [9], a cross-comparative study was done using K fold validation. In [10], two support vector machine (SVM) models for the effective prediction of HF.

3. Materials and Methods

3.1. Dataset Collection

In this research work, the dataset is taken from the UCI Machine Learning Repository, named The Statlog heart disease database which is available in kaggle. The dataset consisting of 1090 records and 11 attributes and 1 target attribute were taken into consideration. Table 1 describes the attributes of the UCI dataset, specifically 8 categorical and 6 numeric attributes. The dataset combination of different clinical test result data, such as serum cholesterol, fasting blood sugar, vessel count, and thalassemia detected from blood work. ST depression and slope of ST-segment were obtained from the electrocardiogram.

Table 1. Dataset Attributes.

Sl. No.	Attribute	Description
1	Age(age)	Age of the patient (in years)
2	Sex(sex)	Gender (0= Female and 1=Male)
3	Chest Pain(cp)	1=Typical angina,2=Atypical angina,3:Non-anginal pain,4:Asymptomatic pain
4	Resting Blood Pressure(trestbps)	Resting Blood Pressure(in mmHg)
5	Serum Cholesterol(Chol)	Serum Cholesterol level (in mg/dl)
6	Fasting Blood Sugar(fbs)	Fasting Blood Sugar(>120mg/dl 0=False , 1-True)
7	Rest Electrocardiograph(restecg)	Resting ECG (0=Normal,1=ST-T wave abnormality, 2=LV Hypertrophy)
8	Maximum Heart Rate(thalach)	Maximum heart rate achieved
9	Exercise-Induced angina(exang)	Exercise-Induced angina (0=No, 1=Yes)
10	Slope of ST segment(slope)	Slope of peak exercise ST segment(1=up sloping, 2= flat, 3=down sloping)
11	Heart Disease(target)	0= negative of disease,1=positive for heart disease

3.2. Data Pre-processing

Feature extraction is a process that extracts a subset of new features from the original set by means of some functional mapping. In this paper, three data pre-processing techniques were employed: min-max, PCA and z-score.

3.2.1 Min Max

This technique transforms each feature (x) by adapting it on a given range (by default [0, 1]). The aim of min-max normalization is to linearly transform the original data

$$X_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

3.2.2 PCA

The PCA is a method for feature extraction which generates new linear combinatorial features of the initial features. It maps each example of a given dataset present in a d dimensional space to a k dimensional subspace such that k < d and the new set of generated k dimensions referred to as the Principal Components (PC). Each PC is directed towards a maximum variance with the exception of the variance which has already been accounted for in all its preceding components. Subsequently, the first component covers the maximum variance while each subsequent component covers lesser value of variance. The PC can be represented thus:

$$PC_i = a_1X_1 + a_2X_2 + \ldots. + a_dX_d \quad (2)$$

where PC_i is Principal Component 'i'; X_j is the original feature 'j'; aj is the numerical coefficient for X_j.

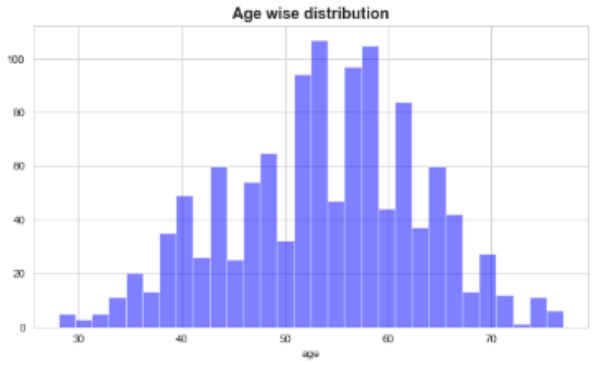
3.2.3 Z score

For Z-score normalization, the normalization of the values for an attribute is based on the mean and standard deviation of the attribute. The z-score is computed as follows where σ is the standard deviation and μ is mean.

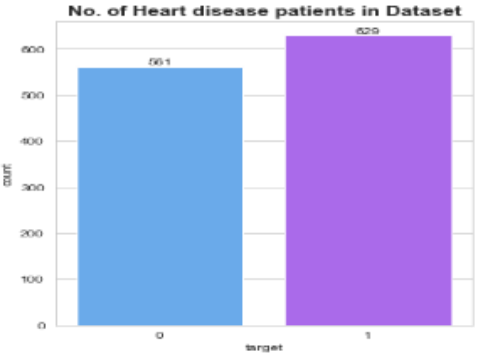
$$Z = \frac{x - \mu}{\sigma} \quad (3)$$

3.2.4 Correlation of the Dataset

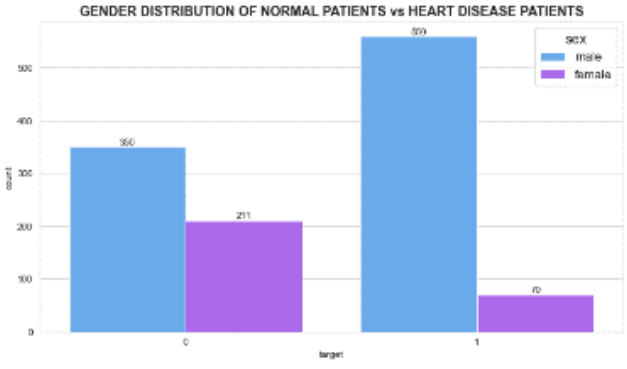
The dataset is visualized with their correlation values and hence the attributes are selected for the process.

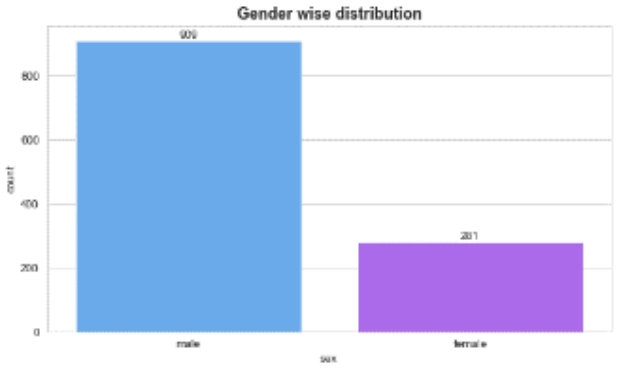


(a)

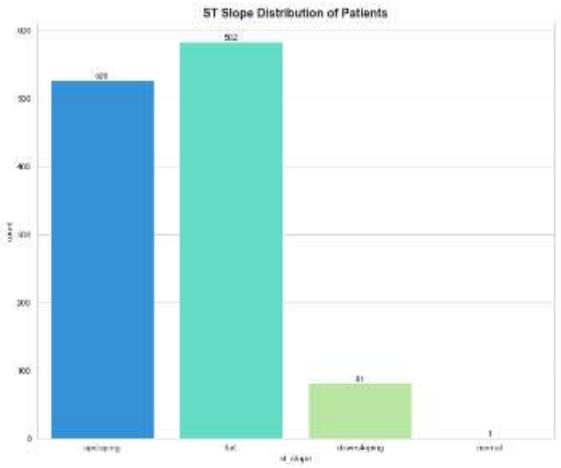


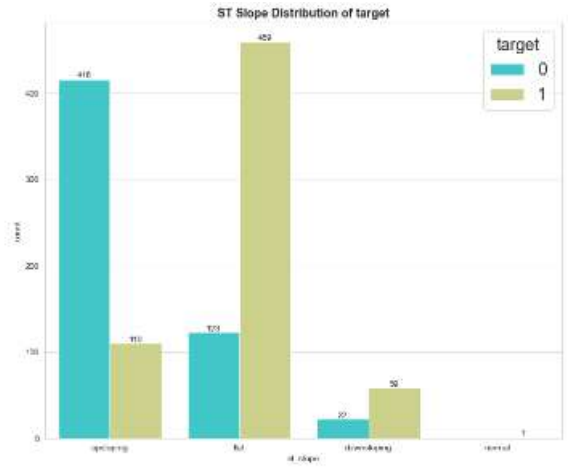
(b)


GENDER DISTRIBUTION OF NORMAL PATIENTS vs HEART DISEASE PATIENTS
sex
male
female
350
211
500
70
count
target
0
1

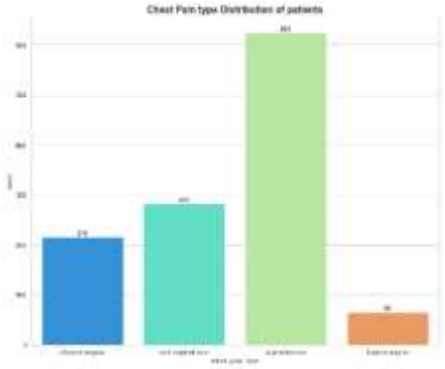

Gender wise distribution
900
281
count
male
female
sex

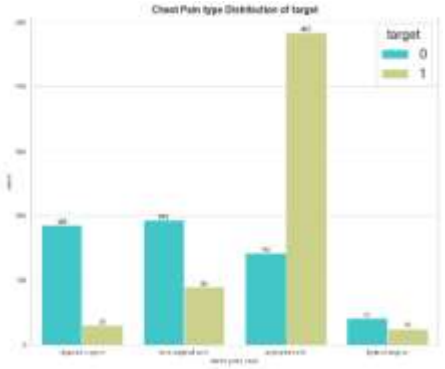
(c) (d)


ST Slope Distribution of Patients
502
81
1
count

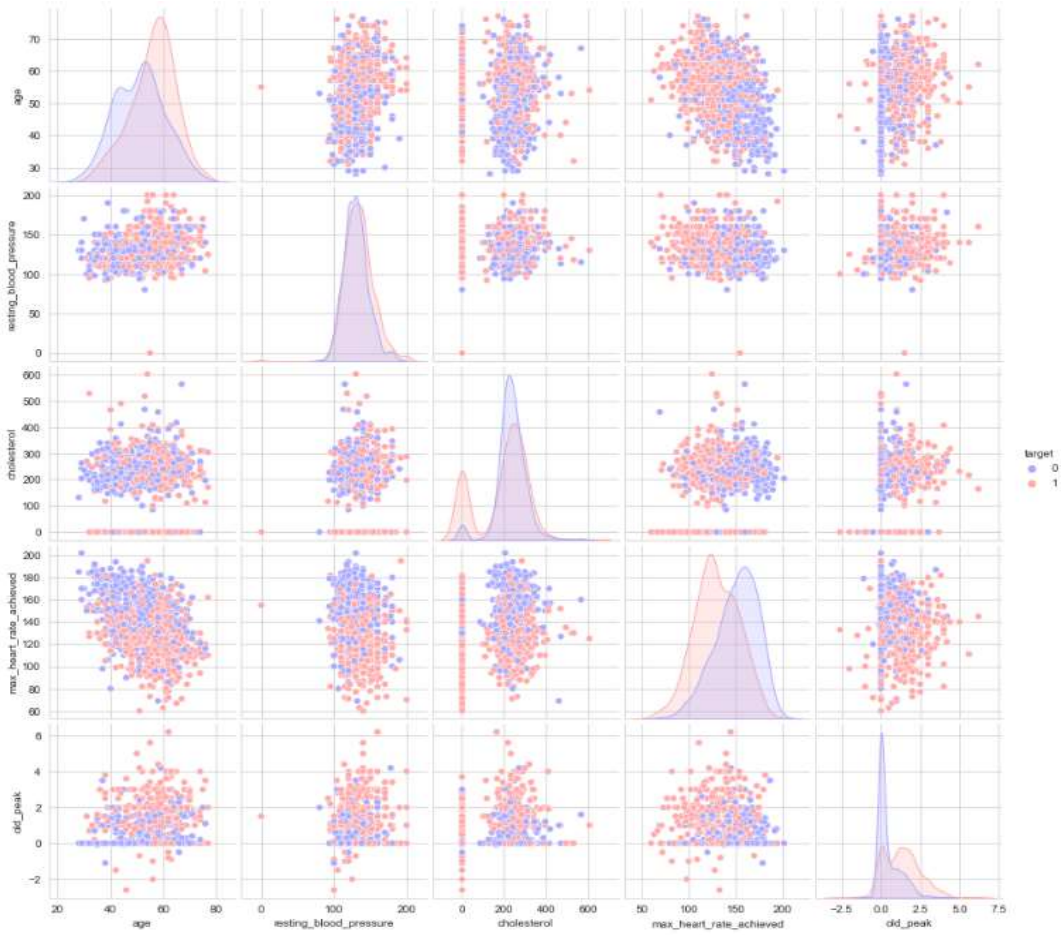

ST Slope Distribution of target
target
0
1
459
count

(e)

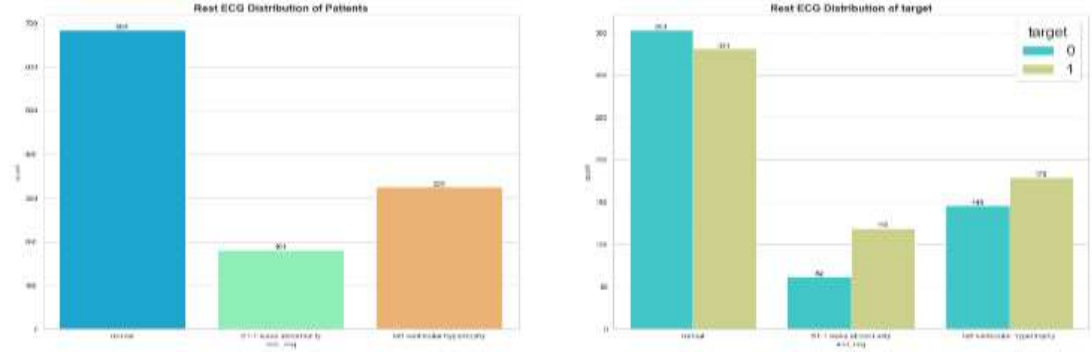

Chest Pain type Distribution of patients


Chest Pain type Distribution of target
target
0
1

(f)



(g)



(h)

Fig.3. Correlation between different features of the data

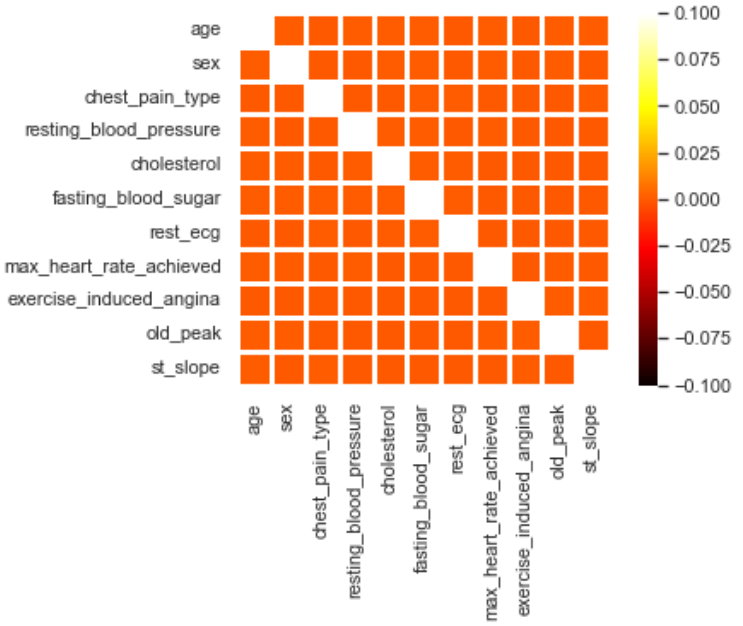


Fig. 4. Heatmap of the dataset

3.3. Model Building

3.3.1 K-Nearest Neighbor

K-NN is ML algorithm. K-NN algorithm predicts the class label of a new input; K-NN utilizes the similarity of new input to its input's samples in the training set. If the new input is same the samples in the training set. The K-NN show high performance or the K-NN classification performance is not good.

3.3.2 Gaussian Naïve Bayes

GNB is based on the concept of Bayes Theorem.

$$P(xi|y)=1/(\surd (2\ \pi\ \sigma\ \hat{}2\ y)\ )\ e\hat{}(-(x\_i-\mu y)\hat{}2/(2\ \sigma\ \hat{}2\ y)) \quad (4)$$

3.3.3 Logistic Regression (LR)

Logistic Regression is a classification algorithm for the probability of occurrence of an event, whether that event will occur or not. It is used to portray a binary or a categorical outcome with only 2 classes. If p is the probability, then, the logit function for p is defined as:

$$Logit(p) = \ln(p/(1-p)) \quad (5)$$

3.3.4 Random Forest (RF)

Random forest algorithm firstly collects random samples from the dataset and then it creates decision trees for each sample. Then from those available trees it selects the tree which produces the best prediction results.

3.3.5 Support Vector Machine (SVM)

A classifier for maximizing the margin using hyperplane is known as support vector machine. It has its roots in the field of machine learning which fits both classification and regression. The training set taking binary classification into consideration is classified as:

$$T=(xi, yi), i=1, \cdots, N, xi \in \{1, -1\} \quad (6)$$

where x_i=M dimension feature of ith case, y_i=class identifier

3.3.6 Decision Tree (DT)

Decision trees are tree like structure that are used to manage large datasets. Entropy changes when training examples are divided into smaller groups using a decision tree node. The measurement of entropy is as follows:

$$\text{Entropy}(S)=\Sigma\_(i=1)\text{^c}\ 〖-〖(P〗\_i\ \log\_2 〖P\_i)〗\ 〗 \quad (7)$$

3.3.7 Bagging

Bagging randomly selects some patterns from the training set with replacement. The newly created training set will have the same number of patterns as the original training set with a few omissions and repetitions. The new training set is known as Bootstrap replicate. In bagging, bootstrap samples are fetched from the data and the classifier is trained with each sample.

3.3.8 Adaptive Boosting (AdaBoost)

Adaptive Boosting involves the conversion of a weak classifier into a strong one using the ensemble technique. For this purpose, the prediction of each weak classifier is merged using weighted average or by taking into account their prediction accuracy as a metrics. Initially, all the attributes are given equal weights, then the algorithm assigns a higher weightage to the inaccurate observation. The error is then propagated with every prediction and multiple iterations are done to reduce it until the prediction become accurate

3.3.9 Gradient Boosting (GB)

This algorithm builds an ensemble of trees in a serial approach, where a weak model, e.g., a tree with only a few splits, is trained first and consecutively improves its performance by maintaining to generate new trees. Each new tree in the sequence is responsible for repairing the previous prediction error.

3.3.10 Extreme Gradient Boosting (X-GBoost)

The rationale of the algorithm is to seek the fine-tuned learning parameters iteratively in order to reduce a cost function. The objective function is

$$O(t)=\Sigma\_(i=0)\text{^n}\ 〖Q(y\_i,\ \ y\text{^}('t-1)\ )+f\_t\ (x\_i))+K\ 〗 \quad (8)$$

3.3.11 Ensemble Technique of Stacking

Stacking is an ensemble technique in which multiple classification models are combined via a meta classifier. Multiple layers are placed one after the other, where each of the models pass their predictions to the model in the layer above, and the model in the topmost layer makes decisions based on the models below. The bottom layer models receive input features from the original dataset. The top layer model takes the output from the bottom layer and makes the prediction.

3.3.12 Proposed Methodology

In this paper, the dataset acquired from the public repository [4] is passed through pre-processing stages. Here, four different pre-processing techniques has been used, viz z-score, min-max, hybrid of min-max and PCA and hybrid of z-score and PCA. ten machine learning models are used as base classifiers, namely K-Nearest Neighbor, Gaussian Naive Bayes, Logistic Regression, Random Forest, Support Vector Machine, Decision Tree, Bagging, Adaptive Boosting, Gradient Boosting and Extreme Gradient Boosting. Then the results from these methods are fed into a stacking method where meta learners were used namely Random Forest, Logistic Regression, K-Nearest Neighbor and Support Vector Machine.

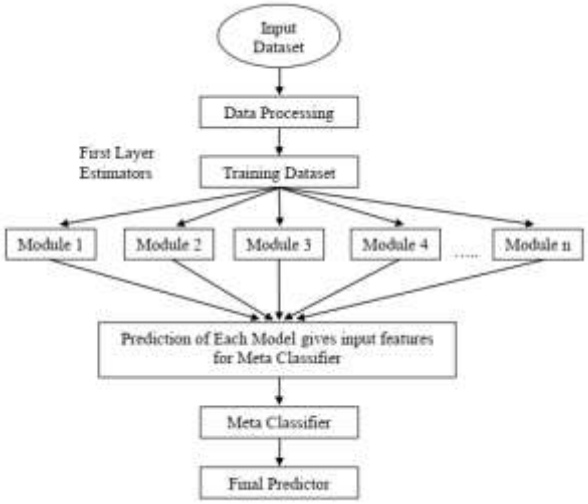


Fig. 5. Proposed Methodology

4. Model Evaluation

4.1 True Positive

A true positive is an outcome where the model correctly predicts the positive class.

4.2 False Positive

A false positive is an outcome where the model incorrectly predicts the positive class.

4.3 True Negative

A true negative is an outcome where the model correctly predicts the positive class.

4.4 False Negative

A test result that indicates that a person does not have a specific disease or condition when the person does have the disease or condition.

4.5 Sensitivity

Recall is the summation of all correctly identified positive values which is divided by the total number of true positive and false negative values. Moreover, recall is a quantitative measure which represent completeness of the result. True Positive Rate measures means positive factors which are identified correctly. High Recall specifies the

correctly diagnosed cases. Recall is calculated by dividing the total number of true positive samples by the total number of true positive and false negative samples. The recall measure is used to assess the model's ability to identify positive samples. There are more positive samples when the recall is higher. The occurrence of the event is often referred to as the event's sensitivity.

$$\text{Sensitivity} = TP/(TP+FN) \tag{5}$$

4.6 Specificity

Specificity itself can be described as the model's ability to predict a true negative of each category available

$$\text{Sensitivity} = TN/(FP+TN) \tag{6}$$

4.7 Precision

Precision measures the number of positive class predictions which actually belong to the positive class. Precision is calculated as the number of True Positives divided by the sum of correctly identified cases. In mathematics, precision is calculated by dividing the total number of true positive samples by the total number of true positive and false positive samples. High precision results in fewer false positives.

$$\text{Precision} = TP/(TP+FP) \tag{7}$$

4.8 F1-score

Precision measures the number of positive class predictions which actually belong to the positive class. Precision is

$$\text{F1-score} = (2*\text{Recall}*\text{Precision})/(\text{Recall}+\text{Precision}) \tag{8}$$

4.9 Accuracy

Accuracy indicates how comfortable the model is with detecting the positive and negative class. The accuracy of a machine learning algorithm is one way to measure how often it is successful in classifying a data point correctly. Accuracy refers to the percentage of data points that were correctly predicted out of all the data points. Accuracy is calculated by dividing the total number of true positives and true negatives samples by the total number of true positive, true negative, false positive, and false negative samples. Data that the algorithm correctly detects as true or untrue are "true positive" or "true negative" data points. A false positive or false negative, on the other hand, is a data point that the algorithm incorrectly classified

$$\text{Accuracy} = (TP+TN)/(TP+TN+FP+FN) \tag{12}$$

4.10 LogLoss

Logarithmic Loss (i.e. Logloss) is an error calculation used to determine how close the predictions of a model are to the actual values.

4.11 Coefficient of Determination

Coefficient of Determination, also popularly known as R square value is a regression error metric to evaluate the accuracy and efficiency of a model on the data values that it would be applied to. R square values describe the performance of the model. Coefficient of Determination also popularly known as R square value is a regression error metric to evaluate the accuracy and efficiency of a model on the data values that it would be applied to. R square values describe the performance of the model. It describes the variation in the response or target variable predicted by the data model's independent variables. Thus, in simple words we can say that, the R square value helps

determine how well the model is blend and how well the output value is explained by the determining(independent) variables of the dataset. The value of R square ranges between [0,1].

$$\text{Coefficient of Determination} = 1 - (SS\_res)/(SS\_tot) \quad (14)$$

Here, SSres represents the sum of squares of the residual errors of the data model. SStot represents the total sum of the errors. Higher is the value of coefficient of determination, better is the model and the results.

4.12 Root Mean Squared Error:

Root Mean Square Error (RMSE) is a standard way to measure the error of a model in predicting quantitative data. Formally it is defined as follows:

$$RMSE = \sqrt{(\sum\_(i=1)^n (\hat{y}\_i - y\_i)^2/n)} \quad (15)$$

where ŷ1, ŷ2.. are the predicted values, y1, y2..are the observed values and n is the number of observations.

4.13 Computation Time

The amount of time required to complete the simulation is termed as the computation time.

5. Experimental Results

5.1. Experimental Setup

The above-mentioned procedure was performed on a 64-bit machine with a 6th Gen Intel i5 CPU (8 GB DDR3+1 TB Hard drive+20 GB SSD). Python was chosen as the programming language on the Google Colaboratory.

5.2. The Result of Applying Data Pre-Processing Methods

In this research work, different data-processing techniques were used and the best has been chosen from them. Figure 1 shows the comparative analysis of the accuracy when the respective pre-processing methods are applied. It is seen that the highest accuracy 94.118% is obtained when Minmax and PCA both are applied in compared to methods where only minmax or z score is applied.

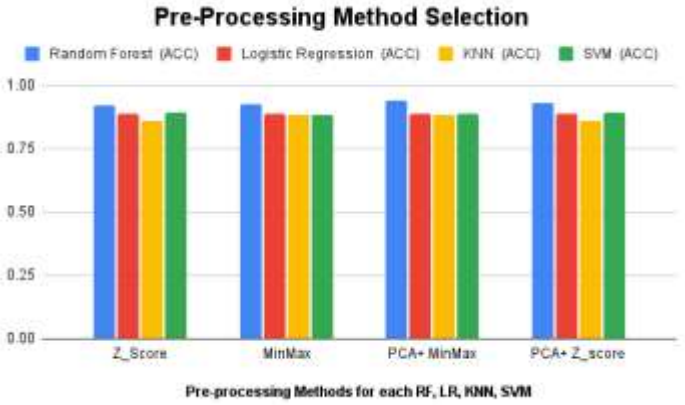


Fig. 6. Result of Applying Pre-Processing Methods on the Features

5.3. Performance Metrics

In this research work, the confusion matrix is shown in Figure 7. The evaluation metrics of the ensemble stacking method using different meta classifiers are mentioned in Table II.

Table 2. Evaluation Metrics

Algorithms	Sensitivity	Specificity	Precision	F1-score	Logloss	Rot Mean Square	Coefficient of Determination
RF	**0.972973**	**0.91339**	**0.907563**	**0.94**	**2.032**	**0.2425356**	**0.9411764**
LR	0.900901	0.87402	0.862069	0.88	3.918	0.336816	0.8865546
KNN	0.900901	0.86614	0.854701	0.88	4.063	0.342997	0.8823529
SVM	0.927928	0.85039	0.844262	0.88	3.918	0.336816	0.8865546

In this research work, the evaluation metrics of the ensemble stacking method using different meta classifiers are mentioned in Table II. Here, it is seen that the Random Forest algorithm gives best results, although its computation time is larger.

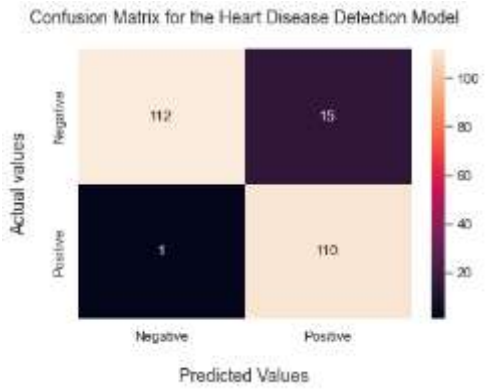


Fig. 7. Confusion Matrix of the Proposed Model

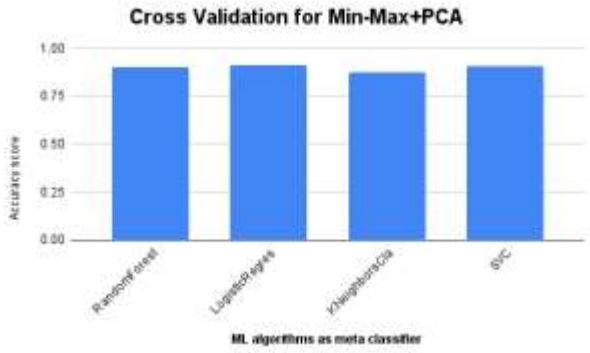


Fig. 8 Comparison of Accuracy from Cross-validation Method using Minmax & PCA as Data Pre-Processing Methods for different meta learners.

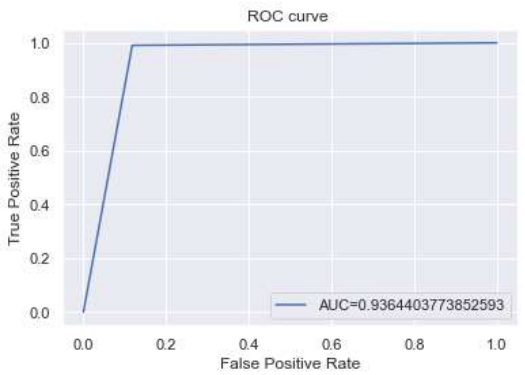


Fig. 9. ROC curve of the Proposed Methodology

In this Fig. 7, the accuracy obtained via cross validation for all the four meta learners are shown for the pre-processed data using MinMax method and PCA method. It is seen that, the accuracy of Random Forest is almost near to that of obtained earlier. In this figure, the area under curve is shown for the Random Forest as meta-classifier when the data is pre-processed with both MinMax and PCA techniques.

Table 3. Comparison of accuracy of our proposed model with other references.

Ref.	Model Used	Meta Learner	Accuracy
[5]	Stacking Classifer	Logistic Regression	89.86%
[6]	Stacking Classifer	Multi Layer Perceptron	92%
[7]	Stacking Classifer	Multilayer Perceptron, Gradient Boosting Machine	91.89%
[8]	Stacking Classifer	Majority Voting	75.1%
[9]	Stacking Classifier	SVM, Random Forest	88.71%
Proposed	Stacking Classifer	Random Forest	94.118%

In the references mentioned in Table III, different base learners were used as SVM, KNN, LR, RF, ET, GBDT, XGBoost, LightGBM, CatBoost, MLP and GBM. Our Proposed Model outperformed all of them with a classification accuracy of 94.118%.

6. Conclusion

The ensemble stacking machine learning method was carried out in this research work on the Statlog heart dataset taken from the UCI Machine Learning Repository. The data was pre-processed using three different methods, and it was found that the combination of the PCA feature extraction method and minmax normalization method gave the highest accuracy where random forest was used as meta learner. In the stacking ensemble method, a total of 10 machine-learning algorithms were employed as base learners, whereas four algorithms were used as meta-learners, namely RF, LR, SVM and KNN. The result was cross-validated where it has 90.08% accuracy.

References

[1] “World Health Organization.” https://www.who.int/newsroom/fact-sheets/detail/cardiovascular-diseases-(cvds).
[2] “Louis Pressbook.”https://louis.pressbooks.pub/medicalterminology/chapter/heartphysiology/.
[3] R. G. Franklin and B. Muthukumar, “Survey of heart diseaseprediction and identification using machine learning approaches,”Proc. 3rd Int. Conf. Intell. Sustain. Syst. ICISS 2020, pp. 553–557, 2020, doi: 10.1109/ICISS49785.2020.9316119.
[4] C. J. Harrison and C. J. Sidey-Gibbons, “Machine learning in medicine: a practical introduction to natural language processing,” BMC Med. Res. Methodol., vol. 21, no. 1, pp. 1–18, 2021, doi:10.1186/s12874-021-01347-1.
[5] I. D. Mienye and Y. Sun, “A Survey of Ensemble Learning:Concepts, Algorithms, Applications, and Prospects,” IEEE Access, vol. 10, no. September, pp. 99129–99149, 2022, doi:10.1109/ACCESS.2022.3207287.
[6] D. C. Yadav and S. Pal, “Thyroid prediction using ensemble data mining techniques,” Int. J. Inf. Technol., vol. 14, no. 3, pp. 1273–1283,

2022, doi: 10.1007/s41870-019-00395-7.
[7] "UCI Machine Learning Repository Statlog Heart dataset."https://archive.ics.uci.edu/ml/datasets/statlog+(heart) (accessed Apr. 10, 2023).
[8] J. Liu, X. Dong, H. Zhao, and Y. Tian, "Predictive Classifier for Cardiovascular Disease Based on Stacking Model Fusion," Processes, vol. 10, no. 4, 2022, doi: 10.3390/pr10040749.
[9] S. Mohapatra et al., "A stacking classifiers model for detecting heart irregularities and predicting Cardiovascular Disease," Healthc. Anal., vol. 3, no. November 2022, p. 100133, 2023, doi: 10.1016/j.health.2022.100133.
[10] H. P. Jia, J. L. Zhao, Jun-Liu, Min-Zhang, and W. X. Sun, "Accurate heart disease prediction via improved stacking integration algorithm," J. Imaging Sci. Technol., vol. 65, no. 3, pp. 1–9, 2021, doi: 10.2352/J.ImagingSci.Technol.2021.65.3.030408.
[11] K. Budholiya, S. K. Shrivastava, and V. Sharma, "An optimized XGBoost based diagnostic system for effective prediction of heart disease," J. King Saud Univ. - Comput. Inf. Sci., vol. 34, no. 7, pp. 4514–4523, 2022, doi: 10.1016/j.jksuci.2020.10.013.
[12] S. Mohan, C. Thirumalai, and G. Srivastava, "Effective heart disease prediction using hybrid machine learning techniques," IEEE Access, vol. 7, pp. 81542–81554, 2019, doi: 10.1109/ACCESS.2019.2923707.
[13] A. Menshawi, M. M. Hassan, N. Allheeib, and G. Fortino, "A Hybrid Generic Framework for Heart Problem Diagnosis Based on a Machine Learning Paradigm," Sensors, vol. 23, no. 3, pp. 1–17, 2023, doi: 10.3390/s23031392.
[14] A. U. L. Haq et al., "Identifying the Predictive Capability of Machine Learning Classifiers for Designing Heart Disease Detection System," 2019 16th Int. Comput. Conf. Wavelet Act. Media Technol. Inf. Process. ICCWAMTIP 2019, no. December, pp. 130–138, 2019.
[15] A. Alqahtani, S. Alsubai, M. Sha, L. Vilcekova, and T. Javed, "Cardiovascular Disease Detection using Ensemble Learning," Comput. Intell. Neurosci., vol. 2022, 2022, doi:10.1155/2022/5267498.
[16] P. Kunwar and P. Choudhary, "A stacked ensemble model for automatic stroke prediction using only raw electrocardiogram," Intell. Syst. with Appl., vol. 17, no. December 2022, p. 200165, 2023, doi: 10.1016/j.iswa.2022.200165.
[17] H. Benhar, A. Idri, and J. L Fernández-Alemán, Data preprocessing for heart disease classification: A systematic literature review., vol. 195. 2020.
[18] S. Mohapatra, S. Maneesha, P. K. Patra, and S. Mohanty, "Heart Diseases Prediction based on Stacking Classifiers Model," Procedia Comput. Sci., vol. 218, no. 2022, pp. 1621–1630, 2023, doi: 10.1016/j.procs.2023.01.140.
[19] V. Shorewala, "Early detection of coronary heart disease using ensemble techniques," Informatics Med. Unlocked, vol. 26, p. 100655, 2021, doi: 10.1016/j.imu.2021.100655.
[20] A. Khan, A. Khan, M. M. Khan, K. Farid, M. M. Alam, and M. B.M. Su'ud, "Cardiovascular and Diabetes Diseases Classification Using Ensemble Stacking Classifiers with SVM as a Meta Classifier," Diagnostics, vol. 12, no. 11, pp. 1–15, 2022, doi:10.3390/diagnostics12112595.