Accuracy based-stacked ensemble learning model for the prediction of coronary heart disease

Santosini Bhutia, Bichitrananda Patra, Mitrabinda Ray
Department of Computer Science and Engineering, Siksha 'O' Anusandhan University, Odisha, India

Article Info

Article history:

Received Oct 17, 2023
Revised Apr 1, 2024
Accepted Apr 17, 2024

Keywords:

Heart disease
Machine learning
Ensemble learning
Logistic regression-recursive feature elimination feature selection
Grid search
Random search
5-fold cross-validation

ABSTRACT

Coronary heart disease (CHD) is the primary cause of silent and noncommunicable deaths. Early detection is essential for slowing the progression of death and saving lives. Medical researchers use machine learning techniques to predict CHD. This article proposes an accuracy based-stacked ensemble learning (AB-SEL) model to predict CHD while minimizing computational time (CT). The dataset undergoes the logistic regression recursive feature elimination (LR-RFE) method to identify the important features. The three strong classifiers, logistic regression (LR), random forest (RF), and AdaBoost, are chosen to build ensemble machine-learning models, including techniques like bagging, majority voting, and stacking, for the Cleveland dataset accessible from Kaggle. Data scaling was done using the normal scalar method, and hyperparameter optimization was done using random search and grid search. Effectiveness is measured by accuracy, precision, recall, F1 score, and CT is validated through 5-fold cross-validation. The suggested approach achieved a 90.16% accuracy rate, required only 0.2 seconds of CT, and yielded an area under the curve (AUC) of 0.892.

Corresponding Author:

Santosini Bhutia
Department of Computer Science and Engineering, Siksha 'O' Anusandhan University
J-15, Khandagiri Marg, Dharam Vihar, Jagamara, Bhubaneswar, Odisha 751030, India
Email: santosini.bhutia@gmail.com

1. INTRODUCTION

Coronary heart disease (CHD) is a major chronic disease. It includes heart and blood vessel disorders. The primary function of the heart is to pump blood promptly to keep humans alive [1]. If the heart fails, essential organs such as the brain may cease to function, leading to a potential shutdown, and the person may succumb. Poor diet, heavy alcohol consumption, patient sex, and age are all risk factors [2]. These factors can negatively impact blood pressure, blood sugar levels, and blood lipid profiles and cause obesity. These aspects lead to a rise in heart disorders and other complexities [3]. The World Health Organisation (WHO) claims that these coronary heart disorders are the leading global cause of death. CHD caused 17.9 million deaths in 2019, 38% of all fatalities. Since these numbers are expected to approach 21 million in the coming decade, this should be addressed seriously [4]. Electronic devices, like electrocardiograms and computed tomography scans, are used to detect CHD. But these devices are very expensive. In poorer countries, there may not be enough doctors to diagnose CHD patients. Diagnostic tests can be erroneous due to infrastructural issues, causing more problems and patient deaths [5]. It also takes more computational time (CT) for assessments. Therefore, it is crucial to have a reliable and precise machine learning (ML) model for predicting CHD at an early stage to prevent death. However, leveraging ML techniques to predict CHD is a critical challenge, and selecting the optimum features from the dataset is another challenge in developing ML models.

Ensemble learning methods in ML are being used in several medical sectors, and feature selection is crucial to their performance [6], [7]. The feature selection approach eliminates certain unnecessary features that may hinder algorithm performance [8], [9]. Choosing an appropriate feature subset and classification algorithm improves the model's performance. Ensemble techniques like bagging, boosting, majority voting, and stacking are reported by Latha and Jeeva [10] using the Cleveland heart disease dataset. These ensemble models were used to enhance the classification accuracy of weak classifiers. The study found that majority voting ensembles of weak classifiers improved accuracy by 7.26%, whereas feature selection approaches boosted performance even more. Uddin and Halder [11] proposed an ensemble-based multilayer dynamic system (MLDS) for better cardiovascular disease prediction. Feature selection methods like the correlation, gain ratio, information gain, lasso, and extra trees classifiers were used to select relevant features. The ensemble model was subsequently created by using random forest (RF), naïve Bayes (NB), and gradient boosting. After dividing 70,000 instances into 50:50, 60:40, 70:30, 80:20, and 87.5:12.5, they got better results. Furthermore, the area under the curve (AUC) curve measures the probability of accurate classification. Gao *et al*. [12] improved heart disease prediction with ensemble learning methods, including bagging and boosting. To find important features, they used principal component analysis (PCA) and linear discriminant analysis (LDA). The authors also compared these methods with five traditional ML methods using various performance metrics. The bagging ensemble learning approach, using the decision tree (DT) classifier and PCA feature selection, has 98.6% accuracy.

Rahim *et al*. [13] introduced the machine learning-based cardiovascular disease diagnosis (MaLCaDD) system. They started by replacing missing data with the mean replacement technique and imbalanced data with the synthetic minority over-sampling technique. They selected the most relevant features using SelectKBest and built an ensemble model using logistic regression (LR) and k-nearest neighbor (KNN). On Cleveland, Framingham, and heart disease, the system achieved 95.5%, 99.1%, and 98.0% accuracy. Mienye *et al*. [14] randomly partitioned the dataset and used classification and regression tree (CART) to model each subset. Subsequently, they developed a homogeneous ensemble model from these multiple CART models by applying an accuracy-based weighted aging (AB-WAE) classifier. This model achieved 93% and 91% classification accuracy on the Cleveland and Framingham datasets, respectively.

According to the literature, many researchers use ensemble approaches to improve classification accuracy over single classifiers. Additionally, feature selection before classification also improves model efficiency. However, they did not provide the CT for model fitting. The literature review does not explain hyperparameter tuning's effects on models. Based on model performance and feature interactions, LR with recursive feature elimination (RFE) removes features repeatedly. LR-RFE is model-driven and identifies LR-relevant features, making feature selection more personalized and optimized than SelectKBest or PCA.

In this study, the LR-RFE feature selection technique is used to select the significant features for predicting CHD. Thereafter, numerous ML algorithms are explored to select the three high-accuracy learners to build ensemble learning models such as majority voting, bagging, and stacking. Comparing the performance of these ensemble models, a stacked ensemble model named the accuracy based-stacked ensemble learning (AB-SEL) model is proposed, which performs better for predicting CHD with the least CT.

The following is a summary of the contributions that this article makes:

– To explore six traditional ML methods: LR, KNN, RF, DT, NB, and support vector machine (SVM), and four boosting methods: AdaBoost, GBoost, XGBoost, and LightGBM. To improve model accuracy, the LR-RFE feature selection technique is used to select a significant feature subset. The hyperparameters are optimized using both grid search and random search approaches.
– Three high-accuracy strong learners-LR, RF, and AdaBoost-are chosen to construct bagging, majority voting, and stacking ensemble models. In clinical applications, ensemble models are evaluated for efficiency using measures including accuracy, precision, recall, F1 score, CT, and AUC to ascertain the most suitable one.
– All models are constructed using the publicly accessible Cleveland dataset, which is available on Kaggle. Python-based Jupyter Notebook is employed for all data processing and computational tasks.

The remainder of the paper is structured as follows: section 2 provides a detailed description of the methodology, highlighting the proposed approach for predicting CHD. It also covers the comprehensive study procedure employed to ensure the reliability and accuracy of the predictions. Section 3 presents the finding of the study, discussion and how they compare with existing literature. Finally, section 4 concludes the paper with a summary, offering final remarks and suggesting potential directions for future research in this field.

2. METHODOLOGY

Let us consider the dataset as $DS = \{(x_i, y_i), i = 1,2,3, \dots n\}$, where x_i is the independent variable and y_i dependent variable used for the prediction of CHD.

$x_i = [x_{i1}, x_{i2}, x_{i3}, \ldots, x_{ip}]$ and $y_i \in \{0, 1\}$

The conceptual framework for the prediction of CHD is shown in Figure 1. This methodology includes feature selection, classifier modeling with hyperparameter optimization, validation, and performance analysis. The first step identifies a set of features that are most relevant to detecting CHD. The LR-RFE method is used to reduce the weakest features until the desired number is reached. The next phase builds the model. Based on their learning type, six common standard ML methods-LR, KNN, RF, DT, NB, and SVM-and four boosting algorithms-AdaBoost, GBoost, XGBoost, and LightGBM-are chosen to explore. The aforementioned models are optimized using grid search and random search with 5-fold cross-validation. The classification algorithms mentioned above form the basis for the classification analysis and effectiveness comparisons outlined in section 3. The three high-accuracy strong learners-LR, RF, and AdaBoost-are chosen to build bagging, majority voting, and stacking ensemble models. Finally, an assessment is carried out to measure the accuracy, precision, recall, F1 score, CT, and AUC of these ensemble models.

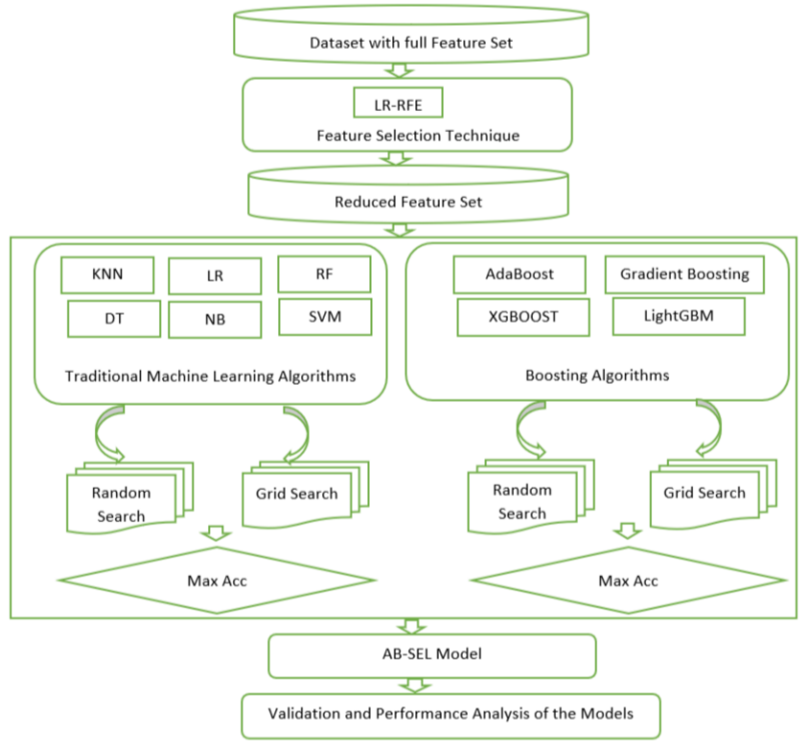


Figure 1. Framework for prediction of CHD

2.1. Dataset

This study uses Kaggle's Cleveland dataset to analyze CHD. Table 1 shows 303 instances and 14 features or attributes. It has 165 CHD patients, and the remainder are healthy. The dataset has been preprocessed using the standard scaler method (1) for ML algorithms. Table 2 lists age, gender, cholesterol, blood pressure, alcohol use, diabetes, and other health factors in the dataset. The discussion suggests that 5-fold cross-validation has produced the most accurate and insightful outcomes for the same dataset.

$$\text{Standardization: } x' = \frac{x-\bar{x}}{\sigma} \tag{1}$$

where $\bar{x}$(mean)=$(\frac{1}{N})\sum_{i=1}^{N} x_i$ and σ (standard deviation)=$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}$

Table 1. Overview of the dataset

Dataset	Instances	Attributes
Cleveland	303	14

Table 2. Detailed attribute description of the dataset

Attribute code	Attribute name	Attribute description
age	Age	Age in years
sex	Sex	0-Female 1-Male
cp	Chest pain types	1-Typical Angina 2-Atypical Angina 3-Non-Angina Pain 4-Asymptomatic Pain
trestbps	Resting blood pressure	Resting blood pressure (in mm Hg)
chol	Cholesterol	Serum cholesterol (in mm/dl)
fbs	Fasting blood sugar	Fasting blood sugar >120 mg/dl 0-False 1-True
restecg	Resting electrocardiographic	Resting electrocardiographic results Value 0-Normal Value 1-Having ST-T wave abnormality Value 2-Showing probable or defined left ventricular hypertrophy
thalach	Maximum heart rate	Maximum heart rate attained at the time of thallium test
exang	Exercise-induced angina	0-No 1-Yes
oldpeak	ST depression	ST depression induced by exercise relative to rest
slope	ST slope	The slope of the peak exercise ST segment 1-Up sloping 2-Flat 3-Down sloping
ca	Number of major vessels	Number of major vessels (0-3) colored by fluoroscopy
thal	Thallium heart test	Thalassemia value 1-Normal 2-Fixed defect 3-Reversible defect
target	Heart disease	0-Patient not suffering from heart disease 1-Patient suffering from heart disease

2.2. Feature selection

Feature selection reduces noisy, insignificant, and redundant features, improving prediction model efficiency [15]. RFE is a wrapper feature selection method [16]. The wrapper method generates a feature subset with the highest classifier accuracy. When fewer features are employed, wrapper methods provide improved accuracy. Haoran Wu suggested an LR-RFE method for eliminating features and enhancing the efficiency of the model in another work [17]. For this reason, we have chosen LR-RFE in our study. LR will be used to score and rank the features. The LR-RFE algorithm reduces the features recursively until the target number is reached. This approach includes the following steps: i) LR is used to fit the model; ii) rank the most important features; iii) eliminate the least important features; and iv) refit the model until the required features are found.

2.3. Base models for ensemble learning

2.3.1. Logistic regression

The most popular simple method for classification in supervised ML is LR [18]. When used on a categorical dependent variable, the result can be discrete or binary. The LR model uses a sigmoid function (2) as a cost function instead of a linear function. The sigmoid function converts a predicted actual value into a probabilistic value between '0' and '1'. Sigmoid function:

$$\sigma(x) = \frac{1}{1+e^{-x}} \tag{2}$$

Here, σ(x) is the probability evaluation function whose value lies between 0 and 1, x is the input of the probability function, and e is Euler's number a mathematical constant. The presented data set shows that 0 implies no heart risks and 1 implies a high risk of CHD.

2.3.2. Random forest

The RF is a collection of DT that takes into account multiple DT before generating a result [19]. This strategy is based on the assumption that more trees will converge on the correct conclusion. In binary classification situations, the RF performs remarkably well. The complete data set has been split into sub-datasets, with each sub-dataset trained in n DT. Each DT is trained separately and makes predictions for its respective sub-dataset. The final prediction relies heavily on the predictive outcomes of these subgroups. In (3) outlines the probability for each subset belonging to a predictive class.

$$prob\left(\frac{C}{F}\right) = P_1 + P_2 + \cdots + P_n \sum_{i=1}^{n}(P_i(\frac{C}{F})) \quad (3)$$

where, C is class, F is features, n is number of sub-datasets, and P_1 ... P_n is probability of each feature and class.

2.3.3. AdaBoost

AdaBoost is a type of estimator that produces a series of weak classifiers, with the algorithm determining the best classifier at each iteration based on the most recent sample weights. Initially, all the data points are assigned equal weights. In the (k+1)th iteration, the samples that were previously misclassified receive increased importance, whereas the samples that were correctly classified are assigned reduced importance. Subsequently, all data points with higher weights are given priority. In each iteration, a stage weight is determined based on the error rate at that particular iteration. The ensemble formed by these weighted classifiers as a whole has a higher probability of being correctly classified than any of the individual classifiers [20].

2.4. Ensemble methods

The ensemble approach is used to enhance classifier accuracy by integrating weak and strong learners. Integrating numerous classifiers aims to improve performance, resulting in better outcomes than using single classifiers. This is a remarkable ensemble learning technique for enhancing the accuracy of multiple algorithms in predicting CHD.

2.4.1. Bagging

A model is trained for each subset of a replacement training set with multiple subsets using bagging classifiers [21]. The final classification outcome is determined by selecting the majority based on the average of the predicted outcomes of the sub-models. In this study, bagging ensemble models are constructed using base classifiers along with grid search and random search to determine the optimum hyperparameter, and their performance is validated using 5-fold cross-validation. These models are more accurate, but their CT is longer.

2.4.2. Voting

The majority voting classifier is a type of meta-classifier that combines predictions from multiple individual classifiers to make a final decision based on the majority vote of the models [22]. In this study, LR, RF, and AdaBoost are the chosen classifiers to form the ensemble model. The ensemble model employs 'hard voting', where each classifier contributes a single vote to the final prediction, and hyperparameter optimization techniques are applied to fine-tune the performance of each classifier, ensuring the most accurate ensemble output.

2.4.3. Stacking

Stacking is a type of ensemble learning in which final predictions are made using multiple layers of models [23]. The stacking classifier was built in this work using RF and AdaBoost as the basis classifiers and LR as the meta classifier. The model was trained using 5-fold cross-validation along with grid search and random search to identify the optimum hyperparameters. The stacking classifier was trained again after selecting the optimum hyperparameters, and its performance was evaluated. The above-mentioned ensemble models are implemented for the prediction of CHD. The effectiveness of these ensemble models is evaluated with regard to accuracy, precision, recall, F1 score, and CT, and this evaluation is detailed in section 4. The following is the pseudocode for the AB-SEL model.

Procedure for AB-SEL model:
Input: Coronary heart disease dataset $DS = \{(x_i, y_i), i = 1,2,3, \ldots n\}$ in CSV format
Output: Performance of the AB-SEL model
Start
Step 1: Import libraries
Step 2: Input the CHD dataset
Step 3: Preprocessing is cleaning and taking care of missing values
Step 4: For all features 1 to n
    Fit the features using the LR-RFE approach
Step 5: Split the dataset in an 80:20 ratio between the training and test sets
Step 6: Scale the feature using standard scalar as (1)
Step 7: Train various classification models on the training set along with hyperparameter optimization and
    select the top three classifiers based on the accuracy scores
Step 8: Apply the stacking algorithm to implement AB-SEL model
Step 9: Stop
End

3. RESULTS AND DISCUSSION

This study investigated the effectiveness of traditional ML, boosting, and ensemble models for predicting CHD using the Cleveland dataset publicly available on the Kaggle database. The performance of all the models is measured in terms of accuracy and CT, whether the reference literature has not addressed the CT. Firstly, the accuracy of the models is assessed by considering all features in the dataset. Subsequently, the LR-RFE approach is applied to eliminate the irrelevant features. Using these feature subsets, all models are optimized through grid search and random search approaches. We found that by using these relevant features, the performance of the models improved, as provided in Tables 3 and 4. It is also observed that grid search generally takes longer than random search across all algorithms. The standard scaler scaling approach is implemented to prevent algorithms from exhibiting bias toward higher values. The hyperparameters employed to achieve optimal results are detailed in Table 5. The 5-fold cross-validation technique is employed to prevent problems related to overfitting. Figures 2(a) to 2(d) show the accuracies and CTs of all models based on the relevant features selected by the LR-RFE feature selection technique. Upon observation, we identified LR (accuracy=88.52%, CT=0.1s), RF (accuracy=88.52%, CT=6.4s), and AdaBoost (accuracy=88.52%, CT=2.7s) as the top three high-accuracy strong learners.

Table 3. Effectiveness of traditional ML models

Traditional ML algorithms	Accuracy with full feature set	LR-RFE with random search		LR-RFE with grid search	
	Accuracy (%)	Accuracy (%)	CT (s)	Accuracy (%)	CT (s)
LR	88.52	88.52	0.1	86.89	0.3
KNN	63.93	83.60	0.1	83.60	0.6
RF	75.41	88.52	6.4	88.52	27.6
DT	77.05	80.32	0.1	85.24	0.7
NB	85.24	86.88	0.1	86.88	2.2
SVM	55.73	81.96	0.1	85.25	5.5

Table 4. Effectiveness of boosting models

Ensemble learning algorithms	Accuracy with full feature set	LR-RFE with random search		LR-RFE with grid search	
	Accuracy (%)	Accuracy (%)	CT (s)	Accuracy (%)	CT (m)
AdaBoost	86.88	88.52	2.7	85.24	5.2
GBoosting	81.96	83.60	3.1	81.96	5.8
XGBoost	83.60	85.24	3.8	85.24	6.3
LightGBM	85.24	81.97	1.2	83.60	1.7

Table 5. The hyperparameters used to achieve optimum results

Algorithms	Parameters of random search	Parameters of grid search
LR	C: 0.23 Penalty: l2 Solver: 'saga'	C: 0.1 Penalty: l2 Solver: 'saga'
KNN	No of neighbor: 9	No of neighbor: 9
RF	n_estimators: 300 max_depth: 3 criterion: gini	n_estimators: 300 max_depth: 3 criterion: gini
DT	Min samples leaf: 7 max_depth: None criterion: gini	Min samples leaf: 7 max_depth: None criterion: gini
NB	var_smoothing: 0.001	var_smoothing: 0.001
SVM	kernel: 'rbf' gamma: 0.01 C: 1.0	kernel: 'sigmoid' gamma: 0.01 C: 1.0
AdaBoost	n_estimators: 40 learning_rate: 1.0	n_estimators: 20 learning_rate: 0.4
GBoosting	n_estimators: 80 max_depth: 6 learning_rate: 0.275	n_estimators: 80 max_depth:3 learing_rate:0.1
XGBoost	n_estimators: 200 max_depth: 5 learning_rate: 0.125	n_estimators: 80 max_depth: 3 learning_rate: 0.1
LightGBM	n_estimators: 80 max_depth: 6 learning_rate: 0.25	n_estimators: 80 max_depth: 3 learning_rate: 0.1

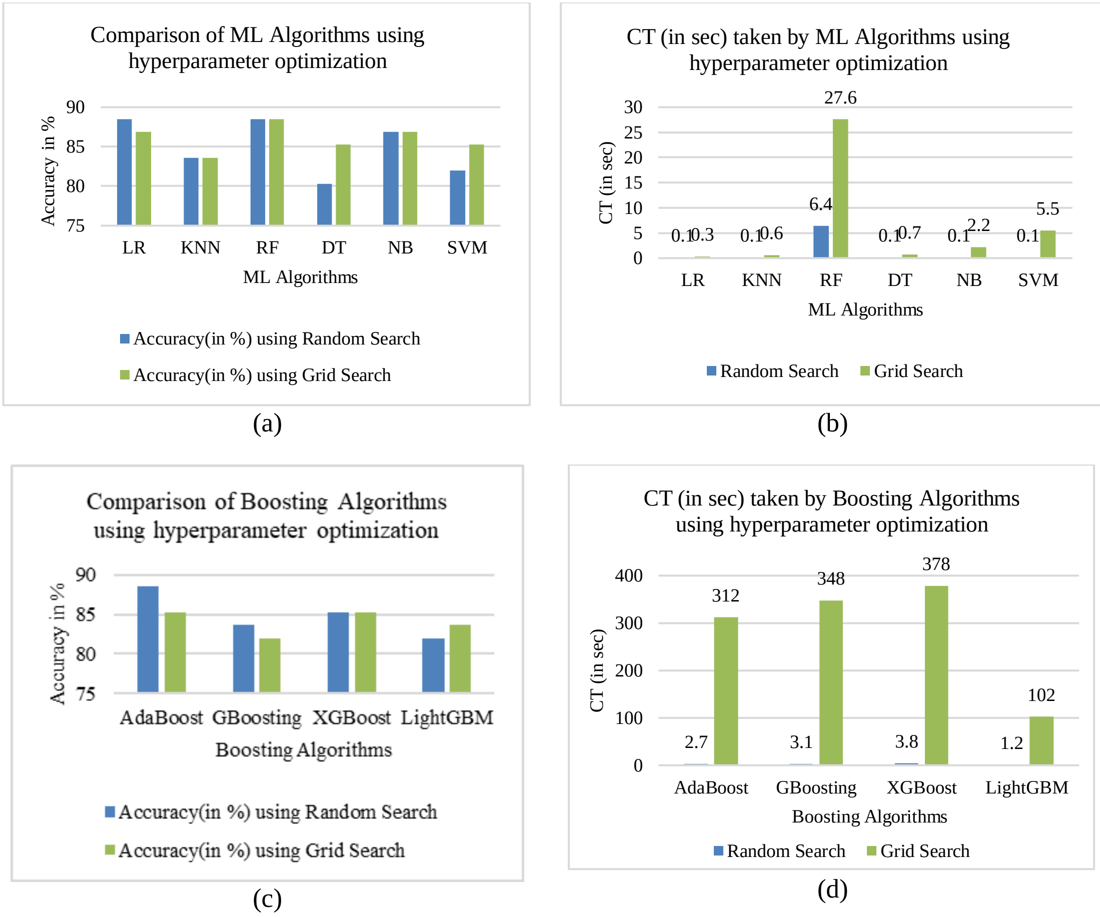


Figure 2. Effectiveness of all models performance of (a) ML algorithms in terms of accuracy, (b) ML algorithms in terms of CT, (c) boosting algorithms in terms of accuracy, and (d) boosting algorithms in terms of CT

The strong learners identified-LR, RF, and AdaBoostare employed to implement three ensemble models, namely bagging, majority voting, and stacking. The effectiveness of these ensemble models is evaluated in terms of accuracy, precision, recall, F1 score, and CT using both random search and grid search approaches, as depicted in Tables 6 and 7, respectively. Figures 3 and 4 present the comparative analysis of these ensemble models, and it is observed that our proposed stacked ensemble model, named the AB-SEL model with random search hyperparameter optimization, performs better than majority voting and bagging ensemble models. This model achieves an accuracy rate of 90.16%, precision of 92%, recall of 85.19%, and an F1-score of 88.46%. The CT taken by these ensemble models using both random search and grid search hyperparameter optimization is compared and shown in Figure 5. The CT taken by the AB-SEL model is 0.2 seconds, with a 0.892 AUC. A higher AUC value indicates better performance [24]. Figure 6 depicts the receiver operating characteristic (ROC) curve for the AB-SEL model.

Table 6. Performance of ensemble models using random search

Ensemble of RF, LR, AdaBoost	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	CT
Stacking	90.16	92	85.19	88.46	0.2s
Majority Voting	88.52	88.57	91.18	89.85	1.6s
Bagging	88.52	80	88.89	84.21	4.3s

Table 7. Performance of ensemble models using grid search

Ensemble of RF, LR, AdaBoost	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	CT
Stacking	90.16	92	85.19	88.46	22.9s
Majority Voting	88.52	88.57	91.18	89.85	3.5s
Bagging	88.52	80	88.89	84.21	1.8m

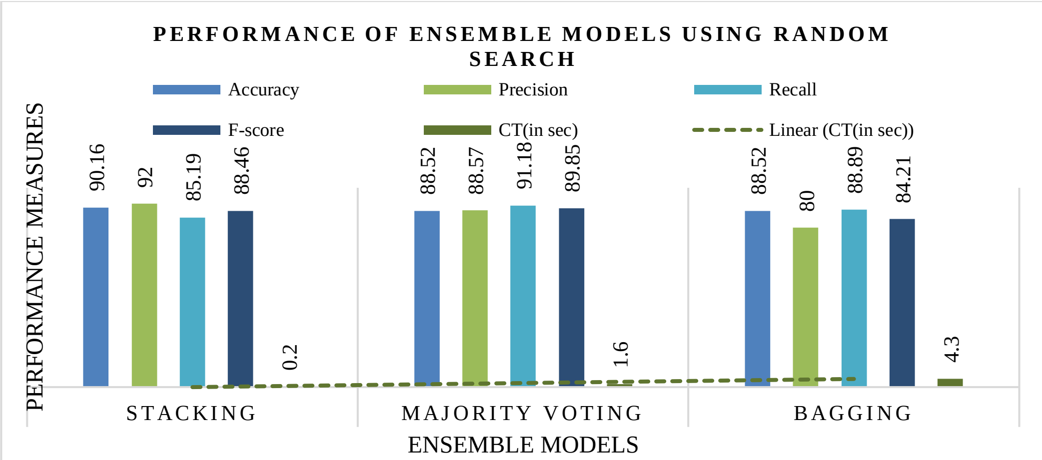


Figure 3. Comparative results graph representation for ensemble models using Random Search

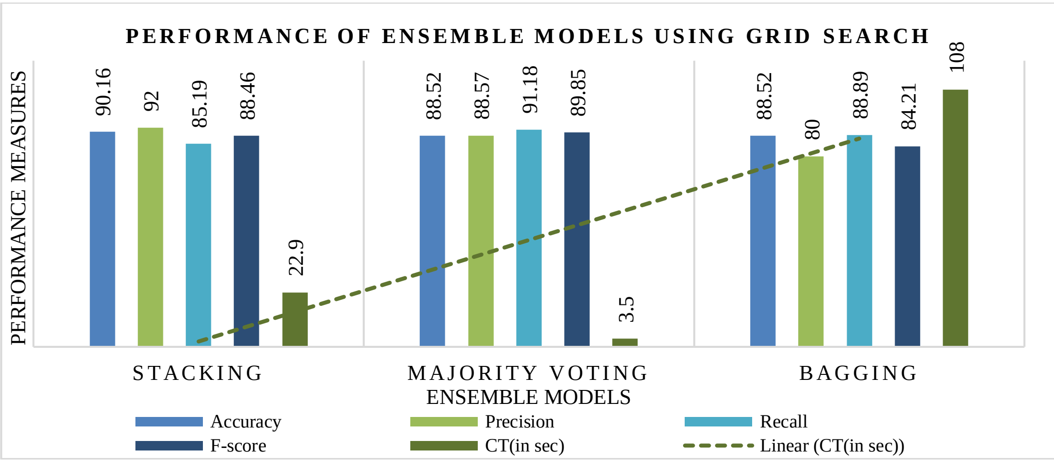
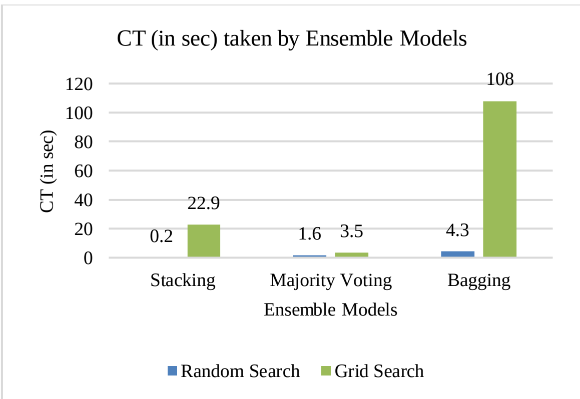


Figure 4. Comparative results graph representation for ensemble models using Grid Search



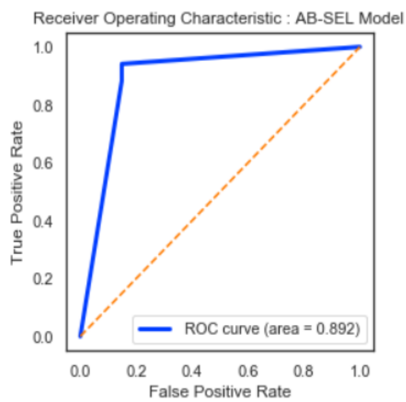


Figure 5. Comparison of CT taken by the ensemble models Figure 6. ROC curve of AB-SEL model

Our study conducted a comprehensive analysis of existing research in this domain, and the findings were systematically organized in Table 8. It shows an accuracy of 85.48% by Latha and Jeeva [10], 88.88% by Raza [22], 87% by Khanna *et al*. [25], 80.14% by Miao *et al.* [26], 85.86% by Tama *et al*. [27], and 87.37% by Mehanović *et al*. [28] using different ensemble learning techniques. The comparative analysis revealed that our proposed AB-SEL model achieved a significantly enhanced accuracy of 90.16%, except for the bagging-quantum support vector classifier (QSVC) model by Abdulsalam *et al*. [29]. The Bagging-QSVC model did not supply the crucial CT for model fitting. Notably, our proposed model demonstrated efficient fitting, completing the process in a mere 0.2 s.

Table 8. Comparison of the proposed models with existing models

Authors	Methods	Feature selection	Accuracy (%)
Proposed AB-SEL Model	Stacked with RF, LR, AdaBoost	LR-RFE	90.16
Latha and Jeeva [10]	Majority vote with NB, BN, MP, and RF	Brute Force	85.48
Raza [22]	Majority voting		88.88
Khanna *et al*. [25]	SVM		87
Miao *et al*. [26]	Adaptive boosting		80.14
Abdulsalam *et al*. [29]	Bagging-QSVC	RFE	90.16
Tama *et al*. [27]	Stacking of RF, GBM, and XGBoost	CFS-PSO	85.86
Mehanović *et al*. [28]	Majority voting		87.37

4. CONCLUSION

In this study, we integrated LR, RF, and AdaBoost to analyze the performance of bagging, majority voting, and stacking ensemble models for the Cleveland dataset available on Kaggle. The standard scalar method is used for the preprocessing of the dataset. LR-RFE is used to select a relevant feature subset, which improves the prediction performance and reduces the overall CT. Grid search and random search are two searching approaches used to select optimal hyperparameters, with random search taking less CT than grid search. According to the findings of this research, the AB-SEL model suggested in the study attains an improved accuracy of 90.16%, requires only 0.2 seconds of CT, and achieves an AUC of 0.892. The suggested intelligent model shows great potential, but it has certain restrictions because of the stacking ensemble model, which allows any models to be utilized as base models and meta-models. For this reason, in the future, with the help of the additional datasets, we will be able to draw more trustworthy conclusions. We may also use metaheuristic methods and nature-inspired algorithms to optimize the classifier parameters and determine the accuracy of the algorithms. Such a type of intelligent model can work with medical professionals to offer a different perspective and actively assist people in detecting CHD.

REFERENCES

[1] A. Kadam, S. Patil, P. Pethkar, R. Shikare, and S. Sarnayak, "A cardiovascular disease prediction system using machine learning," *Journal of Pharmaceutical Negative Results*, pp. 7216–7225, 2023, doi: 10.47750/pnr.2022.13.S09.849.
[2] R. Atat, L. Liu, J. Wu, G. Li, C. Ye, and Y. Yang, "Big data meet cyber-physical systems: a panoramic survey," *IEEE Access*, vol. 6, pp. 73603–73636, 2018, doi: 10.1109/ACCESS.2018.2878681.
[3] D. GhoshRoy, P. A. Alvi, and J. M. R. S. Tavares, "Detection of cardiovascular disease using ensemble feature engineering with decision tree," *International Journal of Ambient Computing and Intelligence*, vol. 13, no. 1, 2022, doi: 10.4018/IJACI.300795.
[4] A. S. Kumar and N. Sinha, "Cardiovascular disease in India: A 360 degree overview," *Medical Journal Armed Forces India*, vol. 76, no. 1, pp. 1–3, 2020, doi: 10.1016/j.mjafi.2019.12.005.
[5] V. Shorewala, "Early detection of coronary heart disease using ensemble techniques," *Informatics in Medicine Unlocked*, vol. 26, 2021, doi: 10.1016/j.imu.2021.100655.
[6] P. Ghosh *et al.*, "Efficient prediction of cardiovascular disease using machine learning algorithms with relief and lasso feature selection techniques," *IEEE Access*, vol. 9, pp. 19304–19326, 2021, doi: 10.1109/ACCESS.2021.3053759.
[7] S. Bhutia, B. Patra, and M. Ray, "COVID-19 epidemic: analysis and prediction," *IAES International Journal of Artificial Intelligence*, vol. 11, no. 2, pp. 736–745, 2022, doi: 10.11591/ijai.v11.i2.pp736-745.
[8] B. Patra, S. Bhutia, and N. Panda, "Machine learning techniques for cancer risk prediction," *TEST Engineering & Management*, vol. 83, 2020, pp. 7414–7420, 2020.
[9] S. Bhutia, B. Patra, and M. Ray, "A hybrid approach for cancer classification based on squirrel search," *Journal of Information and Optimization Sciences*, vol. 43, no. 5, pp. 905–914, 2022, doi: 10.1080/02522667.2022.2091095.
[10] C. B. C. Latha and S. C. Jeeva, "Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques," *Informatics in Medicine Unlocked*, vol. 16, 2019, doi: 10.1016/j.imu.2019.100203.
[11] M. N. Uddin and R. K. Halder, "An ensemble method based multilayer dynamic system to predict cardiovascular disease using machine learning approach," *Informatics in Medicine Unlocked*, vol. 24, 2021, doi: 10.1016/j.imu.2021.100584.
[12] X. Y. Gao, A. Amin Ali, H. S. Hassan, and E. M. Anwar, "Improving the accuracy for analyzing heart diseases prediction based on the ensemble method," *Complexity*, vol. 2021, 2021, doi: 10.1155/2021/6663455.
[13] A. Rahim, Y. Rasheed, F. Azam, M. W. Anwar, M. A. Rahim, and A. W. Muzaffar, "An integrated machine learning framework for effective prediction of cardiovascular diseases," *IEEE Access*, vol. 9, pp. 106575–106588, 2021, doi: 10.1109/ACCESS.2021.3098688.
[14] I. D. Mienye, Y. Sun, and Z. Wang, "An improved ensemble learning approach for the prediction of heart disease risk," *Informatics in Medicine Unlocked*, vol. 20, 2020, doi: 10.1016/j.imu.2020.100402.
[15] B. Patra, L. Jena, S. Bhutia, and S. Nayak, "Evolutionary hybrid feature selection for cancer diagnosis," *Smart Innovation, Systems and Technologies*, vol. 153, pp. 279–287, 2021, doi: 10.1007/978-981-15-6202-0_28.

[16] B. Patra, S. Bhutia, T. Pandey, and L. Jena, "An innovative IoT-based breast cancer monitoring system with the aid of machine learning approach," *The Internet of Medical Things: Enabling Technologies and Emerging Applications*, pp. 155–180, 2022, doi: 10.1049/pbhe034e_ch9.
[17] H. Wu, "A deep learning-based hybrid feature selection approach for cancer diagnosis," *Journal of Physics: Conference Series*, vol. 1848, no. 1, 2021, doi: 10.1088/1742-6596/1848/1/012019.
[18] I. Kamkar, S. K. Gupta, D. Phung, and S. Venkatesh, "Stable feature selection for clinical prediction: Exploiting ICD tree structure using Tree-Lasso," *Journal of Biomedical Informatics*, vol. 53, pp. 277–290, 2015, doi: 10.1016/j.jbi.2014.11.013.
[19] Y. Luo *et al.*, "Predicting congenital heart defects: A comparison of three data mining methods," *PLoS ONE*, vol. 12, no. 5, 2017, doi: 10.1371/journal.pone.0177811.
[20] T. R. Mahesh *et al.*, "AdaBoost ensemble methods using k-fold cross validation for survivability with the early detection of heart disease," *Computational Intelligence and Neuroscience*, vol. 2022, 2022, doi: 10.1155/2022/9005278.
[21] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996, doi: 10.1023/A:1018054314350.
[22] K. Raza, "Improving the prediction accuracy of heart disease with ensemble learning and majority voting rule," *U-Healthcare Monitoring Systems: Volume 1: Design and Applications*, pp. 179–196, 2018, doi: 10.1016/B978-0-12-815370-3.00008-6.
[23] J. Wang *et al.*, "A stacking-based model for non-invasive detection of coronary heart disease," *IEEE Access*, vol. 8, pp. 37124–37133, 2020, doi: 10.1109/ACCESS.2020.2975377.
[24] F. Teng, Z. Ma, J. Chen, M. Xiao, and L. Huang, "Automatic medical code assignment via deep learning approach for intelligent healthcare," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 9, pp. 2506–2515, 2020, doi: 10.1109/JBHI.2020.2996937.
[25] D. Khanna, R. Sahu, V. Baths, and B. Deshpande, "Comparative study of classification techniques (SVM, logistic regression and neural networks) to predict the prevalence of heart disease," *International Journal of Machine Learning and Computing*, vol. 5, no. 5, pp. 414–419, 2015, doi: 10.7763/ijmlc.2015.v5.544.
[26] K. H. Miao, J. H. Miao, and G. J. Miao, "Diagnosing coronary heart disease using ensemble machine learning," *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 10, 2016, doi: 10.14569/ijacsa.2016.071004.
[27] B. A. Tama, S. Im, and S. Lee, "Improving an intelligent detection system for coronary heart disease using a two-tier classifier ensemble," *BioMed Research International*, vol. 2020, 2020, doi: 10.1155/2020/9816142.
[28] D. Mehanović, Z. Mašetić, and D. Kečo, "Prediction of heart diseases using majority voting ensemble method," *IFMBE Proceedings*, vol. 73, pp. 491–498, 2020, doi: 10.1007/978-3-030-17971-7_73.
[29] G. Abdulsalam, S. Meshoul, and H. Shaiba, "Explainable heart disease prediction using ensemble-quantum machine learning approach," *Intelligent Automation and Soft Computing*, vol. 36, no. 1, pp. 761–779, 2023, doi: 10.32604/iasc.2023.032262.

BIOGRAPHIES OF AUTHORS

Santosini Bhutia is currently working as a Research Scholar at Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, Odisha, India. She received her Bachelor of Technology (B.Tech.) in Information Technology, Master of Technology (M.Tech.) in Computer Science and Engineering, and pursuing her Ph.D. in Computer Science and Engineering. She can be contacted at email: santosini.bhutia@gmail.com.

Bichitrananda Patra is currently working as a Professor in the Department of Computer Application at the Institute of Technical Education and Research, Siksha 'O' Anusandhan (Deemed to be University). Bhubaneswar, Odisha, India. He received his M.Tech. in Computer Science from Utkal University, Bhubaneswar, and Ph.D. from Berhampur University, Orissa, India. He has published more research papers in international and national journals, conferences, and book chapters in different books and also has membership in different professional bodies like ISTE and CSI. He can be contacted at email: bichitranandapatra@soa.ac.in.

Mitrabinda Ray is currently working as an Associate Prof. in the Department of Computer Science and Engineering at Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, Odisha, India. She holds her Master of Technology (M. Tech.) and Ph.D. in Computer Science and Engineering from the National Institute of Technology, Rourkela, Odisha, India. She has more than 15 years of teaching and research experience. She has published more than 25 international journals. Her research areas of interest include software testing and software reliability analysis. She can be contacted at email: mitrabindaray@soa.ac.in.