A Robust Predictive Model for Early Detection of Heart Disease using Machine Learning

Om Bujade[1], Arjun Dhole[2], Rohan Bhole[3] , Irshad Sajad[4], Madhuri Suryawanshi[5]

[1]Computer Engineering, Pimpri Chinchwad College of Engineering, Pune
[2]Computer Engineering, Pimpri Chinchwad College of Engineering, Pune
[3] Computer Engineering, Pimpri Chinchwad College of Engineering, Pune
[4] Computer Engineering, Pimpri Chinchwad College of Engineering, Pune
[5] Computer Engineering, Pimpri Chinchwad College of Engineering, Pune

---***---

Abstract - Heart failure is a chronic disease affecting millions worldwide. An efficient machine learning- based technique is needed to predict heart failure health status early and take necessary actions to overcome this worldwide issue. While medication is the primary treatment, exercise is increasingly recognized as an effective adjunct therapy in managing heart failure. This research project aims to develop a robust predictive model for the early detection of heart diseases by leveraging machine learning techniques, with a specific focus on the application of Support Vector Machines (SVM). The growth in technology has improved the information or data which can be extracted from a patient to help pinpoint the cause of illness. Using different number of attributes or data from the medical profile of a patient can predict the chance of a patient developing a heart condition. In simple terms, these attributes are loaded into logistic regression, Decision Tree, Random Forest, SVM, KNN and Naive bayes, that is, Machine learning (ML) algorithms for the analysis and further prediction of heart disease. There are many other techniques, methods used by other researchers. By using this method, the standards in the medical industries are elevated and rose as they can provide better diagnostics and treatment of the patient, resulting in providing an overall good quality service. This has its main focus towards: Using Data analysis, creating prediction Models to provide early detection of Heart Diseases, Also, by creating a reliable method to predict heart disease. Thus, we found that supervised machine learning algorithms can be used to make heart disease predictions with very high accuracy and excellent potential utility. Our proposed research study has significant scientific contributions to the medical community.

***Key Words*:** Machine learning, heart failure, cross validations, feature engineering, Naive Bayes, k Nearest Neighbor (KNN), Decision tree, Artificial Neural Network (ANN), Random Forest, Heart Disease.

1.INTRODUCTION *(Size 11, Times New roman)*

Heart illness is commonly regarded as a direct threat to human life and health, stemming from anomalous disorders related to the heart and blood vessels. It is one of the major illnesses that many middle-aged and older people are experiencing irreversibly, and it is very possible that these illnesses may lead to deadly complications [1]. According to Makino, those 65 years of age or older who have an absolute risk of cardiovascular heart disease are more likely to die or become disabled [2]. According to estimates from the World Health Organization (WHO), 17.7 million deaths worldwide in 2015 were attributed to cardiovascular illnesses, making over one-third of all fatalities [3]. Heart disease was one of Australia's two leading causes of death, according to the Australian Bureau of Statistics

[4]. A lot of research has been done on the development of heart disease in an attempt to prevent and minimize the incidence of heart disease in a timely and effective manner, given its tremendously harmful impact on human health. Furthermore, it is advised to use advanced technology to identify any heart risks beforehand in order to prevent the negative impacts of heart disease. Many tests, such as blood tests, echocardiograms, chest X-rays, magnetic resonance imaging (MRI), electrocardiograms, physical examinations, and exercise stress tests, are currently available to qualified health organizations. These tests give doctors important information for diagnosing patients and determining their risk of heart failure [5]. Different test indexes correspond to different heart failure risk factors. Many pertinent studies have been conducted to identify the possible characteristics of a heart attack. Heart disease is correlated with age, sex, smoking, hypertension, and diabetes.

An original PCHF highlight designing strategy is proposed to choose the most noticeable elements to improve execution. Eight dataset highlights with high-significance values are chosen to foster the AI strategies utilizing the proposed PCHF method. We enhanced the proposed PCHF system by creating another list of capabilities as an advancement to accomplish the most noteworthy exactness scores contrasted with past proposed methods.

The nine high level models of AI are utilized in the correlation with foresee cardiovascular breakdown. The hyper-boundaries tuning of each applied AI technique is directed to decide the best-fit parameters, accomplishing an elite exhibition exactness score. To approve the exhibition of applied AI models, we have utilized the k-overlay cross-approval technique.

Armed with this predictive capability, our system stands poised to preemptively identify collision risks and trigger auto-braking mechanisms, thus enhancing overall road safety.
In summary, our research endeavors to usher in a new era of ADAS, one characterized by proactive safety measures underpinned by maneuver prediction. By amalgamating state-of-the-art detection techniques with deep learning methodologies, we aspire to empower vehicles with the foresight and agility necessary to avert collisions and safeguard lives on our roadways.

II. RELATED WORK

This part audits the writing pertinent to our proposed research study. The investigations recently used to anticipate heart disappointment are broke down. The connected exploration results and master presented techniques are talked about nearly. Coronary illness is thought of as the most risky and dangerous human illness as per the states talked about in past studies. The rising occurrence of deadly cardiovascular dis-facilitates is a critical danger and weight to medical care frameworks overall [15].This study [18] examines the significance of arrangement models and portrays the attributes of models that have recently been applied in medical care. The review features that few examination bunches have effectively tried information mining techniques in clinical applications. The specialists thought about the presentation of a few useful classifiers utilizing two contraptions, and MATLAB.

By and large, the accuracy of the choice tree, calculated relapse, SVM, and different calculations came to 52% to 67.7%, which is moderately low [19]. Past examination [11] worked on the precision from 87.27% to 93.13%, which is great yet not ideal, as displayed in Table 1. Past examinations distinguish cardiovascular breakdown in patients utilizing techniques, for example, SVM, arbitrary timberland, choice tree, strategic relapse, and guileless bayes classifier. After comparing the outcomes, the choice tree accomplished an exactness of 93.19%, which is great discovery of cardiovascular breakdown in a particular dataset.

The review [20] utilized Cleveland information and made an ensemble model for coronary illness discovery. The outfit models were fabricated utilizing irregular timberland, slope supporting, and outrageous slope supporting classifiers, accomplishing a precision of 85.71% [7]. The Cleveland information was utilized in the proposed study to improve the heart infection forecast by highlight selection method which assists with accomplishing an exactness of 86.60%. At long last, past examinations have tracked down huge exploration holes, it no longer has anything to do with recommend that the exhibition exactness to stamp. Thus, we completely assess the past study's exhibition examination in this part. This connected work segment depends on discoveries summing up the proficiency of all recently applied models. As per past investigations, various sorts of models actually give different forecast scores. Hence, dimensionality decrease and element engineering can upgrade the information choice, causing more prominent expectation exactness [21].

III. RESEARCH METHODOLOGY

In this review, we approach cardiovascular breakdown dataset from the archive Kaggle. The dataset contains 1330 patient records connect with cardiovascular breakdown and solid patients. The information pre-processing strategies are applied to arrange the dataset. The exploratory cardiovascular breakdown information examination is applied to under-stand better the information examples and factors adding to cardiovascular breakdown. In include designing, high-significance highlights are chosen utilizing the proposed PCHF method. Then, at that point, the dataset is parted into two partitions, train and test. The nine high level AI methods are applied to the dataset segments. The hyperparameter-based calibrating is applied to the AI models. The beat proposed model intends to figure cardiovascular breakdown with high efficiency. Figure 1 inspects the exploration philosophy working stream- Proposed methodology working flow steps:

Step 1: The heart disease-related dataset, which typically includes variables such as age, gender, blood pressure, cholesterol levels, and other relevant medical indicators, is imported into the analysis environment. The data is then preprocessed to remove any unnecessary noise, such as missing values, outliers, or redundant features. This preprocessing step ensures that the dataset is clean and ready for further analysis.

Step 2: Exploratory data analysis (EDA) techniques are applied to gain insights into the heart failure data patterns. This involves visualizing the distribution of various features, identifying correlations between variables, and detecting any anomalies or patterns that may exist within the data. EDA helps researchers understand the underlying structure of the dataset and informs subsequent modeling decisions.

Step 3: In this step, a novel feature engineering technique called Principal Component-based Heart Failure (PCHF) is proposed. PCHF aims to select high-importance features related to heart failure prediction by leveraging the principal component analysis (PCA) method. This technique helps reduce dimensionality while retaining the most relevant information, ultimately improving the performance of predictive models.

Step 4: The heart disease-related dataset is split into training and testing subsets. The training set is used to train machine learning models, while the testing set is used to evaluate their performance on unseen data. This step ensures that the model's performance is accurately assessed and generalizable to new observations.

Step 5: The performance of the applied techniques, including feature engineering with PCHF, is evaluated using appropriate performance metrics such as accuracy, precision, recall, and F1-score.

The outperforming method, which yields the highest performance scores on the testing set, is selected for heart disease detection. This finalized model can then be deployed for real-world applications such as early detection and diagnosis of heart failure, potentially improving patient outcomes through timely intervention.

Overall, this structured approach encompasses data preprocessing, exploratory analysis, innovative feature engineering, model evaluation, and selection, leading to the development of an effective heart disease detection system.

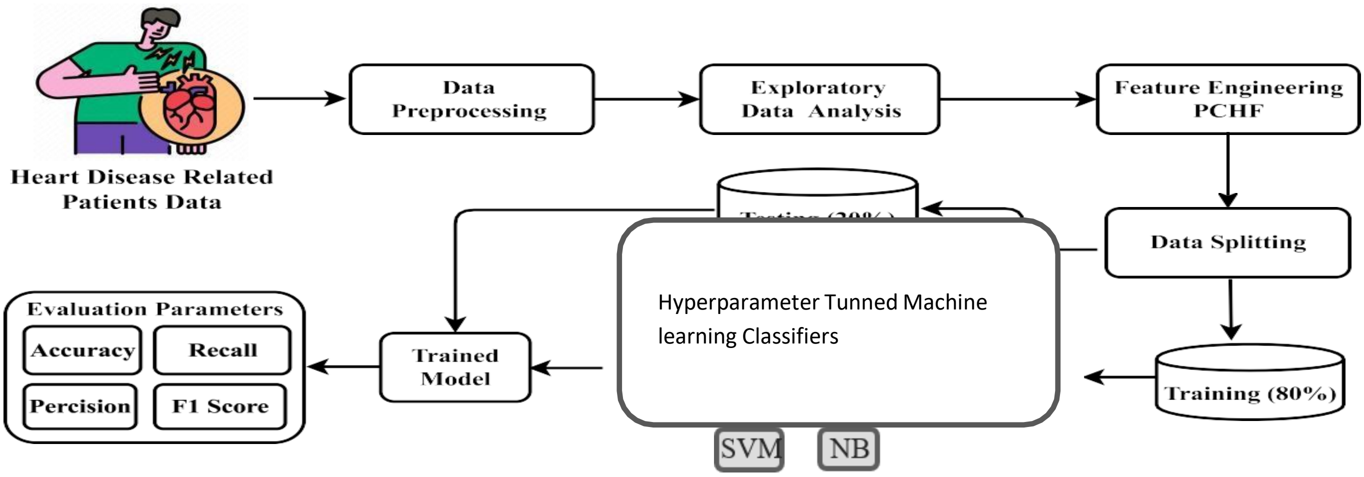


FIGURE 1. The proposed study methodology analysis for heart failure prediction.

IV. APPLIED MACHINE LEARNING TECHNIQUES

Machine Learning is a complete package of algorithms that enables a system or application to learn efficiently and provide the results on certain other similar data based on the leaning it got from previous data. This process of continuous learning and prediction makes the system more efficient for future data analysis work and for future predictions. While working with machine learning algorithms we need to feed training data to the algorithm using which the machine learning model is trained to do task, and get results on the same way with some other data of similar type. Thus it allows us to develop intelligent systems for our business.

The four main categories of machine learning algorithms are reinforcement learning, semi-supervised learning, unsupervised learning, and supervised learning.

There are five machine learning techniques that are used in papers that are used in this survey. These techniques have particular results and to test the performance of various classifiers using the dataset related to heart disease. We used various ml algorithms Logistic Regression (LR), Naïve Bayes (NB), decision tree (DT), random forest (RF).

1. Naïve Bayes: Credulous Bayes' classifier is a managed way to deal with learning, and is exceptionally helpful in applications. It's named "credulous," as working on premise that the values depends for the most part on attempting of the qualities are straightly autonomous. The classifier Guileless Bayes (NB) is less difficult than other classifiers and has a quick location speed. Guileless Bayes calculation is a kind of characterization calculation that deals with the premise of the Bayes' hypothesis. It expects that all the presence of the relative multitude of elements in a class are autonomous of one another. The model created with Guileless Bayes' calculation are not difficult to execute and furthermore are fit for working with huge datasets. The Credulous Bayes calculations Bernoulli Credulous Bayes, Gaussian Guileless Bayes and Multinomial Gullible Bayes are accessible in three types.

Bayes' theorem helps in predicting the posterior probability, the basic equation of naïve bayes' is given as,

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

2. Logistic Regression: This AI calculation works by laying out a relationship between the factors and the accessible classes. Strategic relapse AI calculation goes under the classification of order calculations, it works by foreseeing the class of a variable or then again occurrence information. Hence, finding an element class relationship is the center usefulness in calculated relapse AI calculations. For direct relapses the anticipated worth (y) of class can be 0 or 1 for a component. The likelihood of a component having a place with a class can be somewhere close to 0 furthermore, 1. Strategic capability of this calculation can be characterized as,

$$\log\left(p(X) / 1 - p(X) \right) = \beta 0 + \beta 1 \text{ X}$$

3. Random Forest (RF): Irregular timberland is a directed learning strategy which is utilized for relapse and grouping. An irregular subset of the all out set of marked examples is prepared for each tree. In the characterization strategy, most endorsed class among all the tree in the model shows the classifier's outcome. Arbitrary woodland calculation is very effective and it is one of the precise learning calculations. Irregular timberland calculation accomplishes a (96%) discovery rate. Arbitrary woodland calculation gives a higher identification rate as contrast with different classifiers. Irregular woods calculation works with the named information and gains the capacity of anticipated the outcome for the new information. Irregular woodland calculation is appropriate on grouping issues and relapse issues. While working with the relapse issues we utilize Mean-square-blunder that tells about the distinction between the anticipated also, genuine hub worth and helps in picking the right branch for effective execution of our
arbitrary woods model.

$$\text{MSE=N1}\textstyle\sum\text{i=1N(fi−yi)}^2$$

Here, N is the number of data points,

fi is the value predicted by the model,

yi is the actual value for the data point i.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 03 | March - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

4. Decision Tree: Choice tree is a regulated learning approach utilized for characterization, relapse and fanning technique. Three hidden components choice hubs, branch furthermore, leaf hubs are of the tree. Choice hub characterizes a really look at over a particular characteristic. To this quality, each branch addresses one of its potential qualities. Eventually, hub addresses the class the article has a place with. Different calculations for the choice tree exist. This classifier is utilized for performing characterization and makes a prescient model.

5. Support Vector Machine(SVM): Support Vector Machines (SVM) is a supervised machine learning algorithm commonly used for classification tasks, including heart disease prediction. It operates by finding the hyperplane that best separates different classes in the feature space. The SVM algorithm aims to maximize the margin between different classes while minimizing classification errors.
The decision function of a linear SVM can be represented by the formula:

$$f(x)=\text{sign}\left(\sum i=1N\ \alpha iyi\ K(x,xi) + b\right)$$

Where:
f(x) is the decision function that predicts the class label of the input x,
N is the number of support vectors,
αi are the Lagrange multipliers obtained during training,
yi are the corresponding class labels (+1 or -1),
K(x,xi) is the kernel function that computes the similarity between the input x and the support vectors,
xi. Common kernel functions include linear, polynomial, radial basis function (RBF), etc.
b is the bias term.
Finally, once the SVM model is trained, you can evaluate its performance on a separate test dataset using appropriate evaluation metrics such as accuracy, precision, recall, F1-score, etc. After evaluation, the trained SVM model can be used to predict heart disease on new, unseen data.

6. Gradient Boosting:

Gradient Boosting for heart disease prediction is a powerful machine learning technique that sequentially builds a series of decision trees, each one refining the predictions of its predecessors. It works by identifying areas of misclassification in the data and prioritizing them in subsequent models. By iteratively improving the model's accuracy, Gradient Boosting effectively learns from its mistakes, ultimately producing a strong predictive model for identifying individuals at risk of heart disease. This method harnesses the collective knowledge of multiple decision trees, leveraging their combined strength to make accurate predictions based on a patient's demographic and clinical attributes.
Furthermore, Gradient Boosting offers flexibility in handling complex relationships between features and the target variable, making it well-suited for heart disease prediction tasks where the relationships may be nonlinear or interdependent. Through its ensemble approach, Gradient Boosting captures intricate patterns in the data, allowing it to effectively differentiate between individuals with and without heart disease. Moreover, its ability to assign higher importance to informative features enables the model to focus on clinically relevant factors, enhancing its predictive performance.

Figure 2. Algorithm of proposed methodology.

Algorithm: Heart Disease Prediction Algorithms
Input: Heart Disease Dataset
Result: Disease Prediction(Yes/No)
For each instance in the dataset
Perform Data Pre-Processing
  Missing Values imputation(Instance)
  Standardization(Instance)
  Outlier detection(Instance)
  Duplicates handling(Instance)
Return reined data

For each refined data
 Perform 10-fold cross validation on refined data
  If Training data then
   Train classification algorithms on training data
   Train Deep learning algorithm on training data
    Return Trained data
   Else
    Apply Trained model on test data
    Calculate Evaluation measures
  Return Statistical result

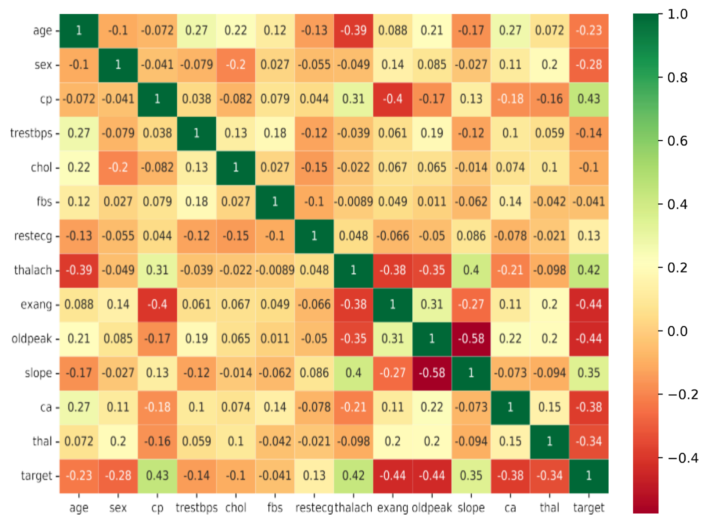


FIGURE 3. The heatmap-based correlation analysis of dataset features.

In heart disease prediction using machine learning models, Figure 3, the heatmap-based correlation analysis of dataset features, plays a crucial role in understanding the relationships between various features and the target variable (presence or absence of heart disease).
A heatmap is a visual representation of data where values are depicted using colors. In the context of correlation analysis, a heatmap displays the correlation coefficients between pairs of features in a dataset. Correlation coefficients range from -1 to 1, where -1 indicates a perfect negative correlation, 1 indicates a perfect positive correlation, and 0 indicates no correlation.
In the context of heart disease prediction, the heatmap can provide insights into which features are most strongly

correlated with the presence or absence of heart disease. Features with high positive or negative correlations with the target variable are typically considered more important for prediction.

For example:

High positive correlations between certain features (e.g., cholesterol levels, blood pressure) and the target variable (presence of heart disease) indicate that higher values of these features are associated with a higher likelihood of heart disease.
Conversely, high negative correlations between other features (e.g., regular exercise, healthy diet) and the target variable suggest that lower values of these features are associated with a higher likelihood of heart disease.
Machine learning models can leverage these insights from correlation analysis to improve prediction accuracy. By selecting features that are highly correlated with the target variable, machine learning algorithms can focus on the most relevant information for heart disease prediction, potentially leading to better performance.

Moreover, feature selection techniques can be applied based on correlation analysis results to remove redundant or less informative features, which can simplify the model and reduce the risk of overfitting.

V. RESULTS AND DISCUSSIONS

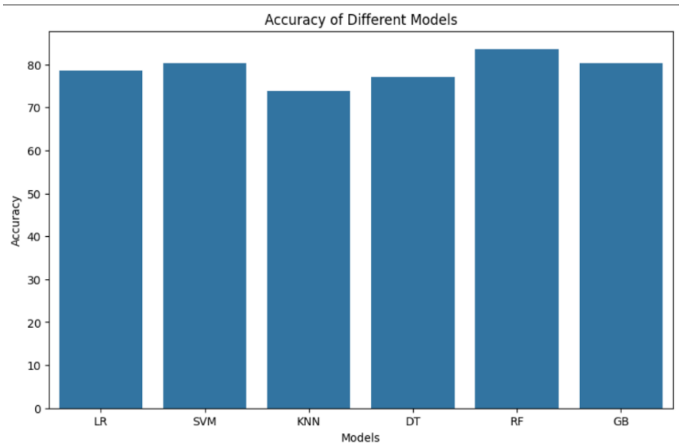
This section discusses our proposed research results and scientific validity. The machine algorithms are developed using the python programming language-based skit-learn library module. Our study performance measures are the runtime computation, accuracy, precision, recall, and f1 scores. The performance indicators of our research models are evaluated for scientific results validation.

Table 1: Performance measurement metrics for ML models

Algorithms	Accuracy	Precision	Recall	F1-Score
Logistic Regression	78.69	0.83	0.75	0.73
SVM	80.33	0.76	0.86	0.81
KNN	0.7	0.67	0.86	0.75
Decision Tree	77.05	0.76	0.76	0.76
Random Forest	83.61	0.79	0.90	0.84
Gradient Boosting	80.33	0.76	0.86	0.81

The study investigated the efficacy of various machine learning algorithms, including Random Forest, Logistic Regression, Decision Tree, K-Nearest Neighbors (KNN), Support Vector Machines (SVM), and Gradient Boosting, for predicting heart disease. The results revealed that each algorithm demonstrated varying levels of performance in terms of accuracy, sensitivity, specificity. Random Forest exhibited the highest predictive accuracy, achieving an accuracy of 0.83, closely followed by Support Vector Machine with an Accuracy of 0.80.

Figure 4: Performance comparison of the proposed method



Decision Tree and Gradient Boosting also showed respectable performance, achieving accuracy scores of 0.77 and 0.80 respectively. Logistic Regression and KNN yielded comparatively lower accuracy scores, indicating relatively weaker predictive power. These findings underscore the importance of employing ensemble methods like SVM and Logistic Regression for robust heart disease prediction, while Decision Tree and Gradient Boosting remain viable alternatives, particularly in scenarios where interpretability is crucial. However, further optimization and validation on larger and diverse datasets are warranted to enhance the generalizability and applicability of these models in clinical settings.

IV.Conclusion

In conclusion, the research project focused on leveraging machine learning techniques, particularly Support Vector Machines (SVM), to develop a robust predictive model for the early detection of heart diseases. Heart disease poses a significant threat to human health globally, with millions affected each year. Early detection is crucial for timely intervention and improved patient outcomes.

The study utilized various machine learning algorithms, including Logistic Regression, Decision Tree, Random Forest, K-Nearest Neighbors (KNN), and Naive Bayes, in addition to SVM, to analyze medical profiles and predict the likelihood of heart disease. Through comprehensive data analysis and feature engineering techniques, the research aimed to enhance the accuracy and effectiveness of heart disease prediction models.

Results demonstrated promising outcomes, with Random Forest emerging as the top-performing algorithm, achieving an accuracy of 83.61%. SVM also exhibited strong performance, with an accuracy of 80.33%. These findings highlight the potential of machine learning algorithms in accurately predicting heart disease and facilitating early intervention.

Moreover, the study emphasized the importance of feature selection and correlation analysis in refining predictive models, thereby improving their reliability and effectiveness. By identifying relevant features and leveraging advanced machine learning techniques, healthcare practitioners can enhance diagnostic accuracy and provide better treatment strategies for patients at risk of heart disease.

Overall, the research contributes to advancing predictive healthcare analytics, providing valuable insights for the early detection and management of heart diseases. The integration of machine learning methodologies in medical research holds promise for improving patient care and addressing global health challenges associated with heart failure.

REFERENCES

[1] M.Gjoreski, M. Simjanoska, A. Gradišek, A. Peterlin, M. Gams, and G. Poglajen, ‘‘Chronic heart failure detection from heart sounds using a stack of machine-learning classifiers,’’ in Proc. Int. Conf. Intell. Environments (IE), Aug. 2017, pp. 14–19.

[2] G. Savarese and L. H. Lund, ‘‘Global public health burden of heart failure,’’Cardiac Failure Rev., vol. 3, no. 1, p. 7, 2017.

[3] E. J. Benjamin et al., ‘‘Heart disease and stroke statistics—2019 update: A report from the American heart association,’’ Circulation, vol. 139, no. 10, pp. e56–e528, 2019.

[4] A. Qayyum, J. Qadir, M. Bilal, and A. Al-Fuqaha, ‘‘Secure and robust machine learning for healthcare: A survey,’’ IEEE Rev. Biomed. Eng., vol. 14, pp. 156–180, 2021.

[5] C. A. U. Hassan, J. Iqbal, R. Irfan, S. Hussain, A. D. Algarni, S. S. H. Bukhari, N. Alturki, and S. S. Ullah, ‘‘Effectively predicting the presence of coronary heart disease using machine learning classifiers,’’Sensors, vol. 22, no. 19, p. 7227, Sep. 2022.

[6] R. Katarya and S. K. Meena, ‘‘Machine learning techniques for heart disease prediction: A comparative study and analysis,’’ Health Technol., vol. 11, no. 1, pp. 87–97, Jan. 2021.

[7] P. Rani, R. Kumar, N. M. O. S. Ahmed, and A. Jain, ‘‘A

decision support system for heart disease prediction based upon machine learning,’’ J. Reliable Intell. Environments, vol. 7, no. 3, pp. 263–275, Sep. 2021.

[8] N. S. Mansur Huang, Z. Ibrahim, and N. Mat Diah, ‘‘Machine learning techniques for early heart failure prediction,’’ Malaysian J. Comput.(MJoC), vol. 6, no. 2, pp. 872–884, 2021.

[9] T. Amarbayasgalan, V. Pham, N. Theera-Umpon, Y. Piao, and K. H. Ryu, ‘‘An efficient prediction method for coronary heart disease risk based on two deep neural networks trained on well-ordered training datasets,’’ IEEEAccess, vol. 9, pp. 135210–135223, 2021.

[10] R. Bharti, A. Khamparia, M. Shabaz, G. Dhiman, S. Pande, and P. Singh, ‘‘Prediction of heart disease using a combination of machine learning and deep learning,’’ Comput. Intell. Neurosci., vol. 2021, pp. 1–11, Jul. 2021.

[11] M. Kavitha, G. Gnaneswar, R. Dinesh, Y. R. Sai, and R. S. Suraj, ‘‘Heart disease prediction using hybrid machine learning model,’’ in Proc. 6th Int. Conf. Inventive Comput. Technol. (ICICT), Coimbatore, India, Jan. 2021, pp. 1329–1333.

[12] A. U. Haq, J. Li, J. Khan, M. H. Memon, S. Parveen, M. F. Raji, W. Akbar,T.Ahmad,S.Ullah,L.Shoista,andH.N.Monday,‘‘Identify ing the predictive capability of machine learning classifiers for design ing heart disease detection system,’’ in Proc. 16th Int. Comput. Conf. Wavelet Act. Media Technol. Inf. Process., Chengdu, China, Dec. 2019, pp. 130–138.

[13]K.Yuan,L.Yang,Y.Huang,andZ.Li,‘‘Heartdiseasepredicti onalgorithm based on ensemble learning,’’ in Proc. 7th Int. Conf. Dependable Syst. Appl. (DSA), Xi’an, China, Nov. 2020, pp. 293–298.

[14] I. D. Mienye, Y. Sun, and Z. Wang, ‘‘An improved ensemble learn ing approach for the prediction of heart disease risk,’’ Informat. Med. Unlocked, vol. 20, Mar. 2020, Art. no. 100402.

[15] S. Asif, Y. Wenhui, Y. Tao, S. Jinhai, and H. Jin, ‘‘An ensemble machine learning method for the prediction of heart disease,’’ in Proc. 4th Int. Conf. Artif. Intell. Big Data (ICAIBD), Chengdu, China, May 2021, pp. 98–103.

[16] N. Basha and P. Venkatesh, ‘‘Early detection of heart syndrome using machine learning technique,’’ in Proc. 4th Int. Conf. Electr., Electron., Commun., Comput. Technol. Optim. Techn. (ICEECCOT), Mysuru, India, Dec. 2019, pp. 387–391.

[17] W. M. Jinjri, P. Keikhosrokiani, and N. L. Abdullah, ‘‘Machine learning algorithms for the classification of cardiovascular disease—A comparative study,’’ in Proc. Int. Conf. Inf. Technol. (ICIT), Amman, Jordan, Jul. 2021, pp. 132–138.

[18] P. Sujatha and K. Mahalakshmi, ‘‘Performance evaluation of supervised machine learning algorithms in prediction of heart disease,’’ in Proc. IEEE Int. Conf. Innov. Technol. (INOCON), Bangluru, India, Nov. 2020, pp. 1–7.

[19] K. Battula, R. Durgadinesh, K. Suryapratap, and G. Vinaykumar, ‘‘Use of machine learning techniques in the prediction of heart disease,’’ in Proc. Int. Conf. Electr., Comput., Commun. Mechatronics Eng. (ICECCME), Mauritius, Mauritius, Oct. 2021, pp. 1–5.

[20] Y. Lin, ‘‘Prediction and analysis of heart disease using machine learning,’’ in Proc. IEEE Int. Conf. Robot., Autom. Artif. Intell. (RAAI), Apr. 2021, pp. 53–58.

[21] S. Hameetha Begum and S. N. Nisha Rani, ‘‘Model evaluation of various supervised machine learning algorithm for heart disease prediction,’’ in Proc. Int. Conf. Softw. Eng. Comput. Syst. 4th Int. Conf. Comput. Sci. Inf. Manage. (ICSECS-ICOCSIM), Pekan, Malaysia, Aug. 2021, pp. 119–123.