

## HEART DISEASE PREDICTION BY INTEGRATING THE FEATURE OF FUSION WITH MARINE PREDATORS' ALGORITHM AND ENSEMBLE CLASSIFIERS

Vikash Kumar Singh<sup>1</sup>, Dinesh Kumar Sahu<sup>2</sup>

<sup>1</sup>Department of Computer Science & Engg., SRK University, Bhopal, India.

[vikashpiet@gmail.com](mailto:vikashpiet@gmail.com)

<sup>2</sup>Department of Computer Science & Engg., SRK University, Bhopal, India.

[drdineshksahu@gmail.com](mailto:drdineshksahu@gmail.com)

### Abstract

Heart disease (HD) is a severe and life-threatening condition that affects approximately one-third of global deaths, with an annual toll of around 17.9 million lives lost worldwide. Nearly half of all individuals diagnosed with heart disease succumb to the condition within just 1-2 years, underscoring its significant impact on human health. An estimated 3% of the total healthcare budget is allocated towards treating heart disease. Predicting heart disease requires multiple tests, and inaccuracies may arise due to lack of expertise among medical personnel. Early diagnosis presents challenges, particularly in developing nations where a scarcity of trained medical professionals and essential diagnostic equipment impedes proper patient care. An accurate assessment of cardiac failure risk holds immense potential for preventing severe heart attacks and enhancing patient safety. This study proposes an ensemble classifier for heart disease prediction using three distinct algorithms: support vector machine (SVM), random forest, and naïve Bayes. The Marine Predators Algorithm (MPA) is used to enhance feature selection and optimization, resulting in an impressive accuracy of 97.05% using just five features. The primary objective of this study is to advance upon existing methodologies by introducing an innovative approach to model construction and develop a model that is not only effective but also easily implementable in practical settings, ultimately contributing to improved heart disease prediction and management.

**Keywords: - Heart Disease, Predication, Ensemble Learning, MPA, Feature selection**

### Introduction

Heart disease (HD) stands as one of the most severe and life-threatening ailments afflicting humans [1]. The escalating prevalence of heart diseases, coupled with their high mortality rates, poses a substantial risk and burden to healthcare systems worldwide. While heart diseases are more prevalent in men, particularly in middle or old age, instances also occur in children. According [2,3] to data from the World Health Organization (WHO), approximately one-third of global deaths are attributed to heart disease, with an annual toll of roughly 17.9 million lives lost worldwide, with a notably higher prevalence in Asia[4,5]. Tragically, nearly half of all individuals diagnosed with heart disease succumb to the condition within just 1-2 years, underscoring its grave impact on human health. Furthermore, an estimated 3% of the total healthcare budget is allocated towards treating heart disease. Predicting heart disease necessitates multiple tests, and inaccuracies may arise due to the lack of expertise among medical personnel, leading to false predictions[6,7]. Early diagnosis presents challenges, and in many cases, surgical interventions for heart disease pose significant hurdles, particularly in developing nations where a scarcity of trained medical professionals and essential diagnostic equipment impedes proper patient care. An accurate assessment of cardiac failure risk holds immense potential for preventing severe heart

attacks and enhancing patient safety. Feature selection and optimization are critical components in the accurate prediction of heart disease. These processes not only aid in identifying the most relevant features but also provide a ranking that guides the selection of features for classifiers tasked with distinguishing between normal and diseased data. Various mathematical functions, including heuristic, meta-heuristic, and derived functions, are employed in the feature selection and optimization process. Their objective is to minimize redundancy within the feature set of medical data, thus enhancing the precision of prediction models. In recent research, swarm intelligence-based selection approaches have gained traction. These include algorithms such as particle swarm optimization (PSO), spider-monkey optimization (SMO), whale optimization algorithm (WOA), among others. This paper, however, adopts the marine predators' algorithm (MPA) for feature fusion. The MPA algorithm proves highly efficient in optimizing feature selection for medical data. Its intelligent mechanisms work to minimize the maximum distance between features, thereby facilitating the convergence of features towards optimal solutions[8,9]. This approach ultimately enhances the effectiveness of prediction models for heart disease. Machine learning (ML) holds promise in aiding physicians by providing insights and anticipating various outcomes within the healthcare domain. Over the past two decades, there has been exponential growth in research focused on automatically extracting features, decisions, and knowledge from healthcare data. However, developing a system capable of accurately and comprehensively extracting information from health data poses a complex and challenging endeavor. Healthcare datasets are frequently available but may contain incomplete information or missing values, adding to the complexity of the task. One of the primary challenges in health datasets is the imbalanced distribution of classes between healthy and non-healthy samples [10,11,12]. This class imbalance can lead to bias in the system, impacting the accuracy, sensitivity, and precision of classification algorithms utilized for disease diagnosis and detection. Several challenges arise when applying ML in healthcare for diagnostic purposes. A notable issue is the lack of a standardized methodology or technique for evaluating the performance of ML algorithms on healthcare datasets [13,14]. Different researchers employ varied methodologies, including training sets, testing sets, and cross-validation (CV) methods, to assess the performance of their systems. This variability in evaluation methods underscores the need for standardized approaches to ensure robust and reliable evaluation of ML-based healthcare systems. The challenge of achieving accurate predictions with machine learning algorithms often prompts the adoption of ensemble classification approaches. Ensembles utilize multiple classifiers to enhance the classification process, leveraging the collective wisdom of diverse algorithms to improve prediction accuracy. Various strategies are employed in designing ensemble classifiers, including bagging, boosting, and random forest techniques. This method involves training multiple instances of the same learning algorithm on different subsets of the training data and then combining their predictions through averaging or voting to arrive at a final prediction. Boosting focuses on iteratively training weak learners (classifiers with modest accuracy) and giving more weight to misclassified instances in subsequent iterations. This iterative process aims to improve overall performance by focusing on areas of difficulty. Random forest is a type of ensemble learning method that constructs a multitude of decision trees during training and outputs the mode of the classes (classification) or mean prediction (regression) of the individual trees. Each tree is trained on a random subset of the features, which helps reduce overfitting and improve

generalization. The design approach of ensemble classifiers offers flexibility and diversity, providing several algorithms to suit different classification tasks and datasets. By combining the predictions of multiple classifiers, ensemble methods often outperform individual models, yielding more robust and accurate predictions. This makes them a valuable tool in machine learning for addressing challenging classification problems and improving prediction performance. The rest of the article explores, as in Section II, related work in the area of heart disease detection; in Section III, proposed methodology based on MPA and EL; in Section IV, experimental analysis and datasets; in Section V, results and discussion; and in Section VI, a paper with future direction

## II. Related Work

Ensemble learning, a technique that combines multiple models to improve prediction accuracy, holds significant promise for accurately detecting heart disease when coupled with feature selection approaches. In recent studies, several authors have introduced detection methods based on ensemble learning and optimization techniques. Ensemble learning involves various approaches such as bagging, boosting, and random forest. Bagging (Bootstrap Aggregating) constructs multiple models using bootstrapped samples of the dataset and aggregates their predictions to produce a final output. Boosting sequentially builds models, with each subsequent model focusing on the weaknesses of the previous ones to improve overall performance. Random forest creates an ensemble of decision trees, where each tree is trained on a random subset of features and samples. The exploration of these recently proposed algorithms for heart disease detection aims to leverage the strengths of ensemble learning and optimization methods to enhance prediction accuracy and reliability. In [1], he outlines strategies for transforming data to produce reliable data for the training model. Presenting the findings separately facilitates comparisons. Our suggested model produced the maximum accuracy when employing RFBM and relief feature selection methods, at 99.05%, according to the outcome analysis. The author of this study presents the findings separately to facilitate comparisons. This study discovered that the RF technique achieved 100% accuracy, coupled with 100% sensitivity and specificity, utilising a heart disease dataset obtained via Cagle three-classification based on k-nearest neighbour decision tree and random forests algorithms. In [2], the LR and the NB outperformed the other models for the unbalanced data models, with accuracy values of 0.728592 and 0.707762, respectively. in [3] used an explainable methodology based on ML and the Shapley Additive Explanation (SHAP) method to determine the risk of 3-year all-cause mortality and provide individual explanations of the model's choices. A multivariable analysis adjusted hazard ratio of 5.343 and P 0.001 both confirmed that subjects with higher ML scores were at a higher risk of experiencing events. The author [5] eleven machine learning (ML) algorithms are evaluated, including linear discriminant analysis (LDA), logistic regression (LR), and K-nearest neighbours (KNN). The results demonstrate that CART, along with RS or QT, outperforms all other ML algorithms with 100% accuracy, 100% precision, 99% recall, and a 100% F1 score. For early viral prediction, the suggested model (NSGA-II+AdaBoost) may be helpful [6]. For dataset 2, the NSGA-II+AdaBoost model produced accuracy, precision, sensitivity, specificity, and AUC rates of 95.56%, 98.16%, and 96.87%. The Gradient-Boosting Model achieved a sensitivity rate of 87.3%, a specificity of 71.98%, and an AUC value of 90%. in [7], compared the classification effectiveness of the aforementioned strategy with the findings from the five most commonly used

metaheuristic methods. The study then contrasts the classification effectiveness of the aforementioned strategy with the results obtained from the five most commonly used heuristic methods, where the mMRR-ANN-IMGWO framework achieved an accuracy of 95.43%, recall of 93.79%, specificity of 0.97 probability, precision of 0.95 probability, and F-measure and AUC of 0.969 probability, respectively. In [8], Wisconsin Diagnostic Breast Cancer (WBCD) dataset results show that the cloud-based Elm strategy works better than other methods. The experimental results indicate an F1-score of 0.8129, a recall of 0.9130, a precision of 0.9054, and an accuracy of 0.9868. In [9], the study Then we use a support vector machine to put classifications into practice. Experimental findings show that our strategy performs better than single-modal methods and alternatives. When we combine the multi-modal properties of the ECG and PCG, our technique achieves an AUC value of 0.936. In [10], a suggested diagnostic approach includes techniques for characteristic selection and classification. The team considered the ant colony's optimization methods (OEC) and the significance of the characteristic when selecting the attributes. CNN performs better than alternative approaches when the suggested learning rate optimisation is used. With 93.84% accuracy and only a small amount of noise, I diagnosed PD. In [11], the current findings show that the use of AI algorithms has produced positive outcomes for numerous datasets across various nations. Additionally, many countries have not concentrated on gathering demographic-specific data. In [12], the quantum random forest classifier had the highest accuracy rate of 0.89, which is the best. The F1 score, recall, and accuracy results for the quantum random forest classifier were also the best, coming in at (0.88), (0.93), and (0.89), respectively. [13] to evaluate the performance of the developed system. Researchers use the Physic Net heart sound and Cagle heartbeat sound datasets to assess the performance of the developed system. The suggested CNN-jSO outperforms existing algorithms. During testing, the CNN-jSO system, implemented using Python, achieved an accuracy rating of 94.12%, while during training it achieved 97.76%. In [14], this research suggests a unique method to predict heat illness called the Reliable Boolean Machine Learning Algorithm (RBMLA). Studies have shown that it performs better when new test data and new real-time data are added. The model achieved an accuracy of 86%. In [15], heart disease features are included in the dataset. Examining these features, taking into account feature importance scores and cross-validation, helps improve algorithm performance. Following evaluation, the Decision Tree algorithm's accuracy and AUC ROC are 83% and 99%, respectively. The logistic regression algorithm outperforms other algorithms in terms of accuracy and AUC, with 88% and 91%, respectively. In [16], the confusion matrix of each employed method is formed by the three criteria found in the performance metrics: accuracy, sensitivity, and specificity. The specifics of how we apply our methodology to the datasets presented. So let's talk about how to divide the dataset into two scenarios for the training and testing phases: 70%–30% and 80%–20%. in [17] optimizes these condensed features using two well-known classification algorithms. These condensed features are then optimised using two well-known classification algorithms: linear SVM, which uses a linear combination of all features to build a robust model, and weighted k-NN, which has a lower-dimensional feature space and holds the best accuracy rates in the classification of heart sounds on a widely used dataset. [18] proposed a strategy utilizing machine learning to predict cardiac illnesses at an early stage based on prior data and a comprehensive set of features. The proposed strategy of machine learning, which helps in predicting cardiac illnesses at an early stage using

prior data and an outstanding collection of features, may result in better prognostic outcomes. LR fared better on the state log dataset, with an accuracy of 93% using the chosen features. In this case, the test data accounted for 20% of the data. In [19] With just a few basic sets of HRV parameters, the suggested approach and the results of the HRV parameter selection have the potential to create workable systems and devices for automatic AF detection. The most accurate machine learning classifier in the blindfold validation attained a diagnostic odds ratio of 1566 and an accuracy of 97.2. In [20], several machine learning techniques, such as the ensemble learning method and deep neural network-based algorithms, have demonstrated promising results in increasing prediction accuracy for early CAD diagnosis. The training set, comprising 90% of the dataset, trained the models, while the test set, consisting of 10%, evaluated the performance of the ensemble-based classification algorithms. [21] has examined and contrasted these FS techniques with classifiers that use machine learning. Researchers have examined and contrasted these FS techniques with classifiers that use machine learning. Using the reduced feature set obtained by SSOA, the suggested approach achieved maximum classification accuracies of 98.53%, 98.84%, 99.07%, and 99.70% for classification tasks with five classes, four classes, three classes, and two classes, respectively. In [22], a machine learning (ML) model is proposed for predicting a person's likelihood of developing cardiovascular disease. According to the experiment's findings, Multi-Layer Perceptron has the highest disease prediction accuracy at 87.23%. [23] discusses the shortcomings of the prior approaches in this work, along with a critical evaluation of those approaches. The study concludes by outlining some potential avenues for future research in the area of automated heart disease detection using numerous data sources and machine learning. In [24] The HEV classifier achieved results for all assessment metrics above 98%, and the gradient-boosting model was the second-best classification model with accuracy and F1-score values of 97% and 98%, respectively. Medical professionals can use the suggested method as an adjunctive therapy to identify CAD instances in questionable patients quickly, accurately, and effectively. In [25], KNNs perform better than other methods, as seen by their testing accuracy of 94.4% and AUC of 84.4%. Additionally, we discovered that the patient's age, a blood test for a full metabolic panel, the number of white blood cells per MCL, breathing rate, and high blood pressure are among the top five most relevant metabolic indicators. [26] conducted a study where medical data, such as blood pressure, age, gender, chest pain, cholesterol, and blood sugar, were analysed to predict the likelihood of developing heart disease. Medical data predict the likelihood of developing heart disease based on factors such as blood pressure, age, gender, chest pain, cholesterol, blood sugar, and other variables. The Cleveland dataset showed that the suggested approach outperformed other current models, enriching 99. Accuracy. In [27], the non-negative matrix factorization method employed in the feature-based technique to denoise PCG signals was followed by feature extraction from both the entire recording and its segmented parts. All fusion findings show lower performance across the board compared to separate classifiers. Because the deep learning model and fusion can vary, the results are not definitive. In [28], we obtain practical guidance for a condensed yet effective description based on an analytical justification by identifying the most pertinent aspects that can influence the decision from a publicly available dataset. [29] constructed the newly constructed dataset in this work on top of four publicly accessible benchmark datasets. The study on ablation shows that the suggested method can achieve a competitive mean average accuracy (mAA) of 99.2% and a mean average

AUC of 97.9%. In [30], the stacked set of second-level classifiers was then built and optimised using the fusion of these six-classifier decisions. The findings show that our model accurately predicted the death rate of HF patients with an AUROC of 82.55%, confirming its capability and effectiveness. [31] utilized the Cleveland, Hungary, Switzerland, and Long Beach (CHSLB) datasets, all sourced from Kaggle, to evaluate the models. We used the Cleveland, Hungary, Switzerland, and Long Beach (CHSLB) datasets, all obtained from Kaggle, to evaluate the models. The accuracy of the predictions made by the RF, DT, AB, and KNN models on the CHSLB dataset was 99.03%, 96.10%, 100%, and 100%, respectively. Only two models—RF and KNN—show good accuracy in the case of a single (Cleveland) dataset, with 93.437% and 97.83%, respectively. In [32], this can help us draw a crucial conclusion regarding how uncertainty affects the classification of data. The mechanisms used by SVM and NB to establish the link between inputs and outputs are fundamentally distinct from one another. Future research will examine how feature selection and weighting affect classifier performance as a whole. In [33], the result of the suggested hybrid framework is validated using various performance measures in the final layer. The top recorded metrics are 99.17% accuracy with CNN and 100% accuracy with the ML algorithms Extra Tree Classifier (ETC) and Random Forest Classifier (RFC). In [34], a thorough analysis of the most advanced illness prediction techniques, particularly chronic disease prediction algorithms, will give a good sense of the most recent models created in this field. This work provides a thorough performance comparison of several techniques. In [35], the following significant characteristics were chosen: dwelling location, temperature, age, overall body ache, visit date, and headache. With an accuracy of 95% in Kilimanjaro, 87% in Moro Gore, and 82% in the total dataset, Random Forest was the most accurate classifier. We created a regional-specific malaria prediction model to display relevant machine learning classifiers based on clinical symptoms and demo visual elements. In [36] According to the conservative BPRF measure, SBP exposure between 107.5- and 165.0-mm Hg increased risk by an average of 101.36%, resulting in a rise of 0.70 and a five-star grade. According to our analysis, IHD risk began to rise at 120 mm Hg SBP, continued to rise continuously until 165 mm Hg, and then began to rise less sharply after that.

### III. Methodology

This section outlines the methodology proposed for heart disease prediction using an ensemble classifier. The ensemble classifier is composed of three distinct algorithms: support vector machine (SVM), random forest, and naïve Bayes. Support Vector Machine (SVM) serves as the base classifier in the ensemble. SVM is a powerful algorithm for classification tasks that works by finding the optimal hyperplane to separate data points into different classes. Random Forest and Naïve Bayes algorithms are utilized as variable classifiers within the ensemble. These classifiers assist in the selection of feature variables and contribute to boosting the performance of the base classifier. To select the most relevant features for prediction, the methodology employs the Marine Predators Algorithm (MPA). MPA enhances the convergence of features and boosts the overall performance of the ensemble classifier. The process of the proposed algorithm is visually represented in Figure (2), illustrating the flow and interactions between the components of the methodology. The first section focuses on detailing the Marine Predators Algorithm (MPA), highlighting its role in feature selection and optimization. The second section delves into the specifics of the proposed ensemble classifier, elucidating how SVM, random

forest, and naïve Bayes are integrated and how they collectively contribute to heart disease prediction. Overall, this methodology combines the strengths of multiple classifiers and leverages advanced feature selection techniques to enhance the accuracy and effectiveness of heart disease prediction.

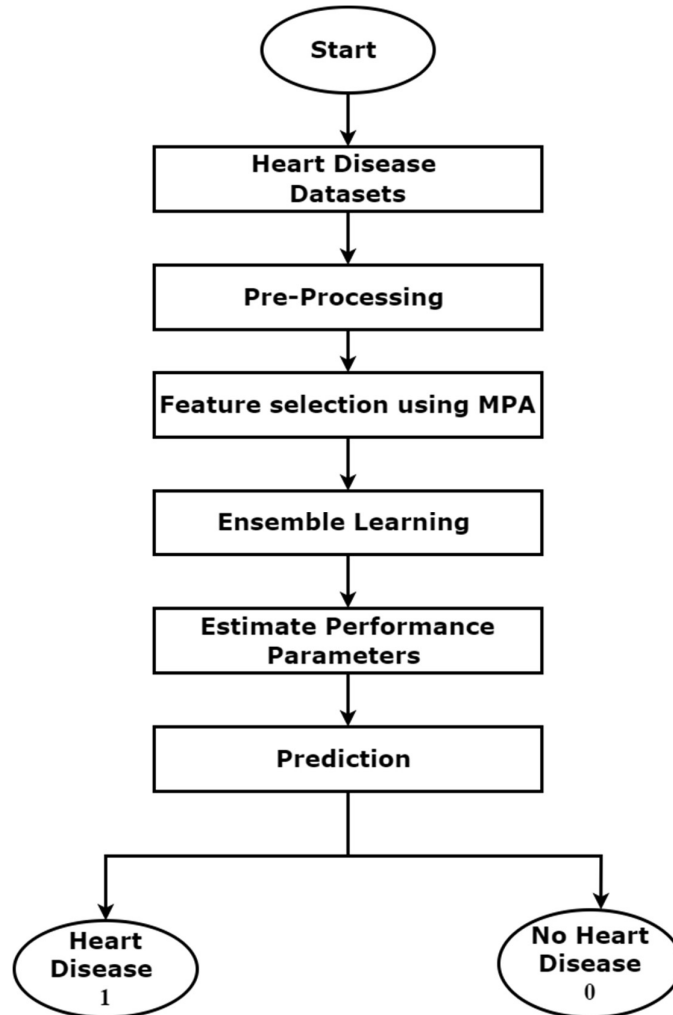


Figure 2 Proposed model of heart disease detection based on ensemble classifier.

### 1<sup>st</sup> section

#### Marine predators' algorithm (MPA)

The inspiration for the Marine Predators Algorithm (MPA) stems from the natural foraging behavior observed in ocean predators and their interactions with prey. In this context, predators aim to optimize encounter rates to enhance their survival prospects in their natural environment. MPA employs two simple random walk methods, namely Levy flight and Brownian motion, to conduct searches. Levy flight, commonly utilized in meta-heuristic algorithms, is particularly effective in preventing solution stagnation by facilitating a constructive search in local areas. Additionally, Brownian motion serves as a well-established global search mechanism. The creators of MPA combined the search efficiency of Levy and Brownian motion to enhance the trade-off scale between exploration and exploitation. The process of function topology and objective space present in figure (1)[37,38].

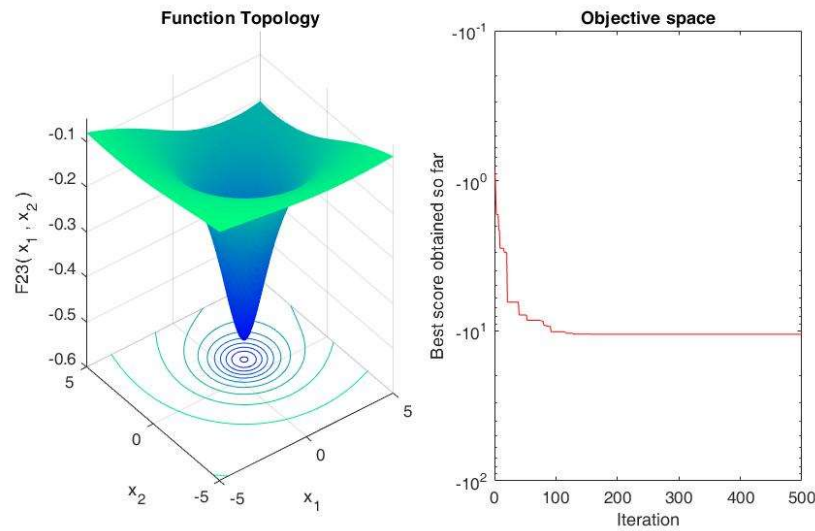


Figure 1 function topology and objective space of MPA

**2<sup>nd</sup> section**

The proposed ensemble classifier integrates three distinct classifiers: Support Vector Machine (SVM), Naïve Bayes (NB), and Random Forest (RF)[30,33,36]. In this ensemble setup, SVM serves as the base classifier, while NB and RF act as feature collectors aimed at enhancing the performance of the base classifier. Support Vector Machine (SVM): SVM is utilized as the base classifier in the ensemble. SVM is known for its effectiveness in separating data points into different classes by finding the optimal hyperplane. In this context, SVM forms the foundation of the ensemble model, providing a robust framework for classification. Naïve Bayes (NB) and Random Forest (RF): NB and RF serve as feature collectors within the ensemble. Their primary role is to gather relevant features from the dataset and contribute to the enhancement of the base classifier's performance. NB is a simple yet powerful probabilistic classifier based on Bayes' theorem, while RF is an ensemble learning method that constructs multiple decision trees and combines their predictions. By leveraging the strengths of NB and RF, the ensemble classifier aims to improve its predictive capabilities. The processing of the ensemble classifier involves the integration of these components in a coordinated manner. Each classifier plays a distinct role within the ensemble, contributing to the overall prediction process. By combining the outputs of multiple classifiers, the ensemble classifier aims to achieve higher accuracy and robustness in predicting heart disease.

M= base classifier Support vector machine'

X= NB classifier

Y = RF classifier

The processing of sample data as K

$$D(k) = [s1, \dots, sm] \dots \dots \dots (1)$$

Here s1,s2,.....sm is set of features of medical data.

The processing of data in sample classifier

$$X(k) = [Pi1, \dots, Pin] \dots \dots \dots (2)$$

P is probability value of features



The process of training come under the sample data as

$$P_j = \begin{cases} 1, & \text{if } j = n \\ 0, & \text{otherwise} \end{cases} \dots \dots \dots (3)$$

Boosting of base classifier

$$M = \underset{\forall(x, |H1, \dots, HK|) \in Trf}{\operatorname{argmin}} \sum \epsilon b(K_{rf}(s), [P1, \dots, Pn]) \dots \dots \dots (4)$$

Where  $\epsilon b$  is error function for the base class feature voting measure as

$$\epsilon b = \sum_{k=1}^k |PN - N| \dots \dots \dots (5)$$

Voting of base classifier

$$B = \frac{n_j}{\sum_{j=1}^{nc} nc} \dots \dots \dots (6)$$

algorithm (ensemble voting)

1. Process  $x=[ P1,P2,\dots,\dots,Pk]$  the base classifier M and model of voting Trf.
2. Output: detection of disease
3. Estimate information entropy of features
4. Estimate mapped features of X and Y
5. Measure the voting of confusion matrix
6. Exit

**IV. Experimental Analysis**

To evaluate the performance of proposed algorithms using MATLAB software version R2018a, the system configuration comprises an Intel Core i7 processor, 16GB RAM, and the Windows 10 operating system. While MATLAB offers built-in support for certain classification algorithms KNN, random forest and ensemble learning other classifier functions are defined and programmed using function files, which are compiled with library files. For the evaluation process, UCI machine learning datasets are employed. The data processing involves utilizing 10-fold cross-validation for prediction and assessing parameters like accuracy, sensitivity, and specificity and F1. The prediction of disease categorized into two class, class-1 and class-2. The representation of class-1 is abnormal and class-2 is normal. The description of parameters as [22,23,24]

$$Accuracy = \frac{TP+}{TP + TN+FP+FN} \times 100 \dots \dots \dots (7)$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100 \dots \dots \dots (8)$$

$$Specificity = \frac{TN}{TN + FP} \times 100 \dots \dots \dots (9)$$

$$F1 = 2x \frac{Pricision \times Recall}{Pricision+R} \times 100 \dots \dots \dots (10)$$

- TP: True Positive
- TN: True Negative
- FP: False Positive
- FN: False Negative

Dataset Descriptions:

Hungarian: This dataset, collected at the Hungarian Institute of Cardiology in Budapest by Andras Janosi, consists of ten features. Initially comprising 294 samples, 34 were discarded due to missing values, leaving 262 records. These records are divided into 62.21% representing healthy subjects and 37.78% with heart disease.

Cleveland: With a total of 303 instances, the Cleveland dataset contains 76 attributes, of which only 14 are considered. It serves as another dataset for heart disease prediction analysis.

Z-Alizadeh Sani: This dataset includes 270 instances and 13 attributes. Patients are classified into two categories: CAD or Normal. A patient is categorized as CAD if their diameter narrowing is greater than or equal to 50%, otherwise labeled as Normal.

Statlog: This dataset, comprising 270 instances and 13 attributes, features no missing attribute values.

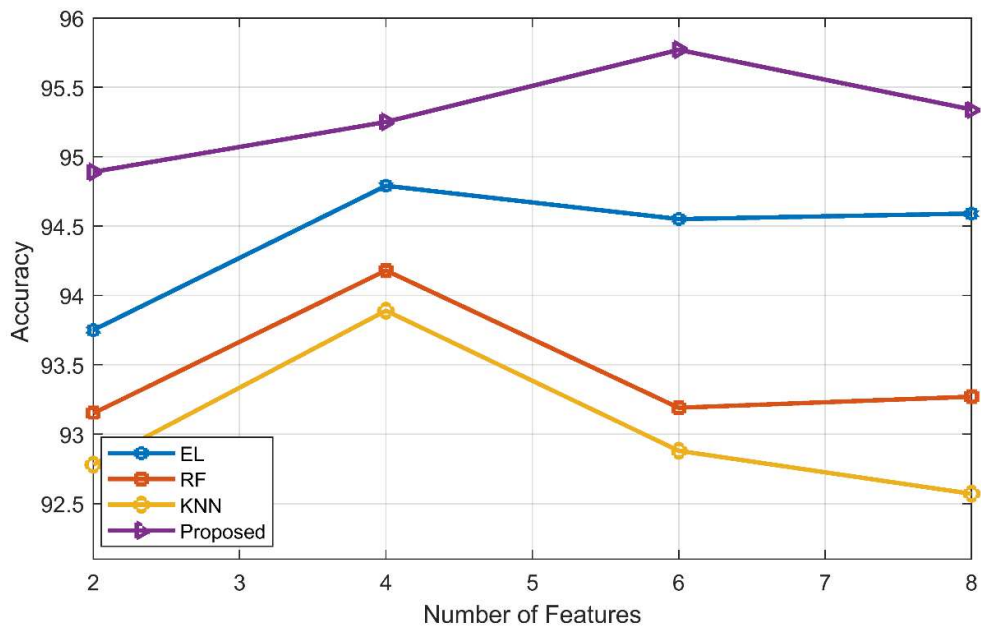


Figure: 3 Comparative performance analysis of accuracy for Cleveland dataset.

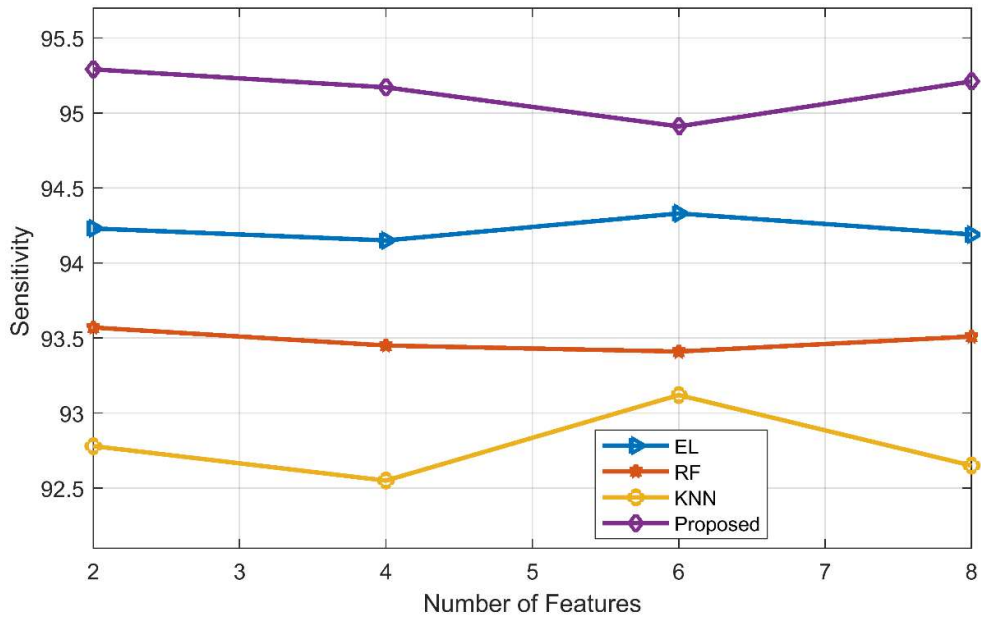


Figure: 4 Comparative performance analysis of sensitivity for Cleveland dataset.

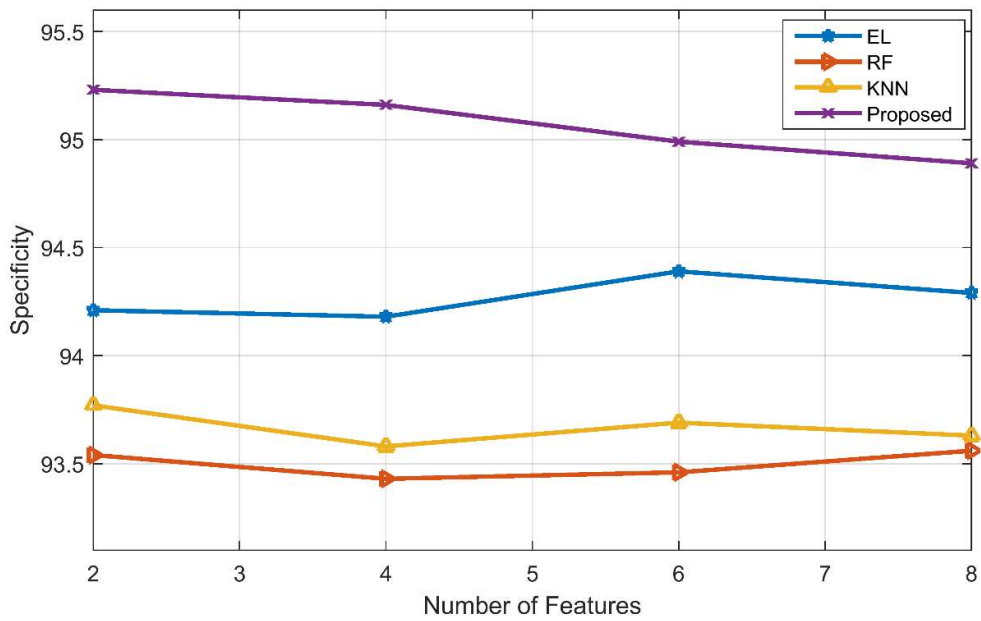


Figure: 5 Comparative performance analysis of specificity for Cleveland dataset.

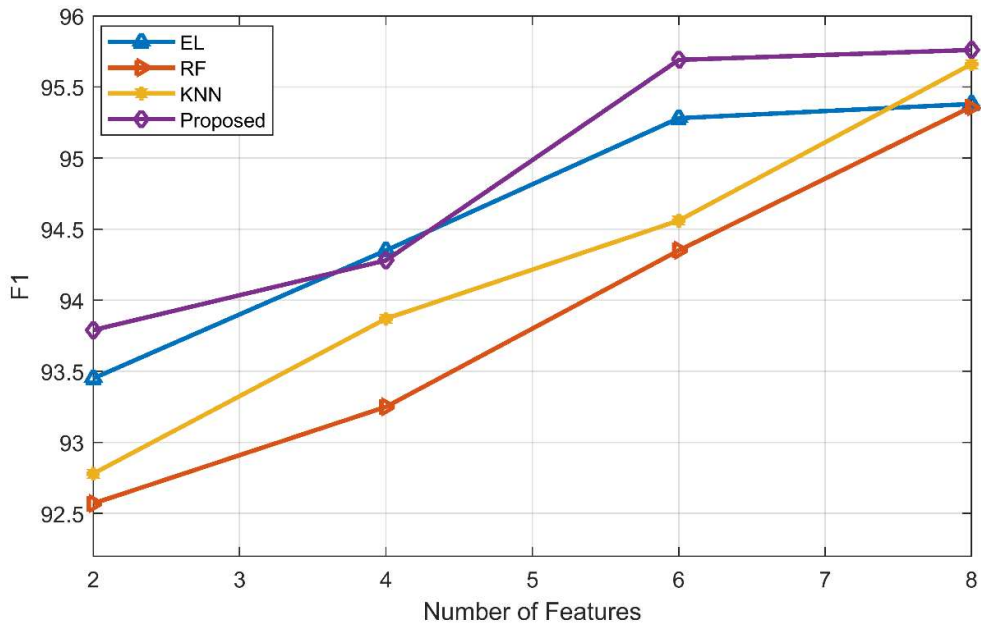


Figure: 6 Comparative performance analysis of F1 for Cleveland dataset.

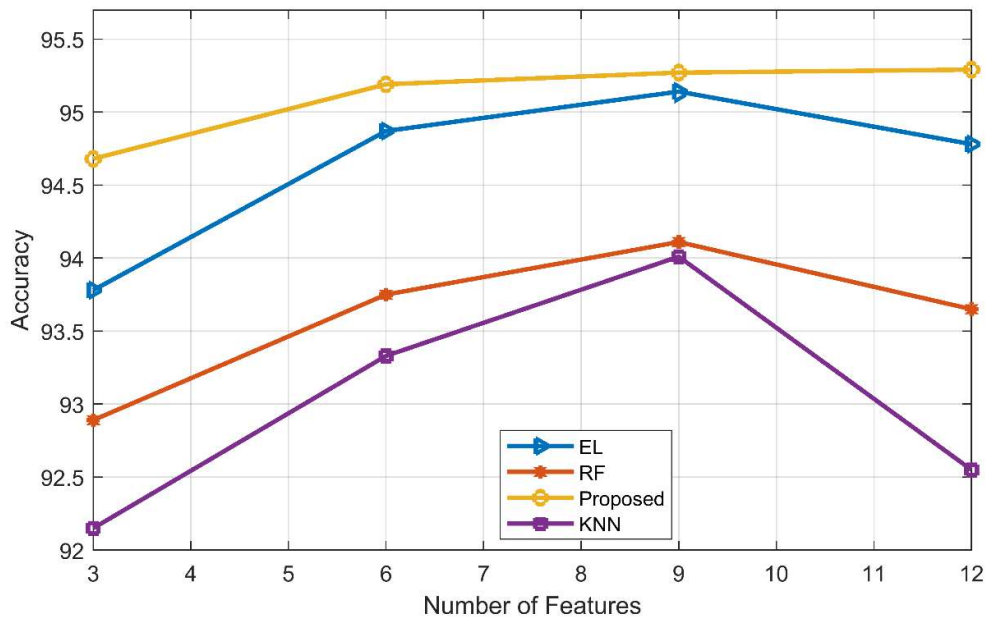


Figure: 7 Comparative performance analysis of accuracy for Hungarian dataset.

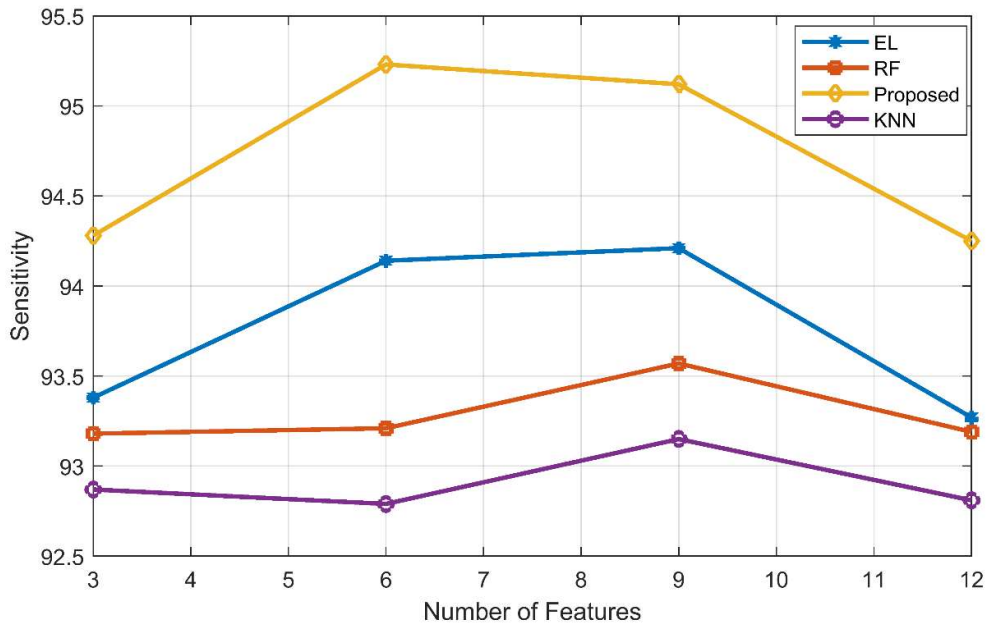


Figure: 8 Comparative performance analysis for sensitivity for Hungarian dataset.

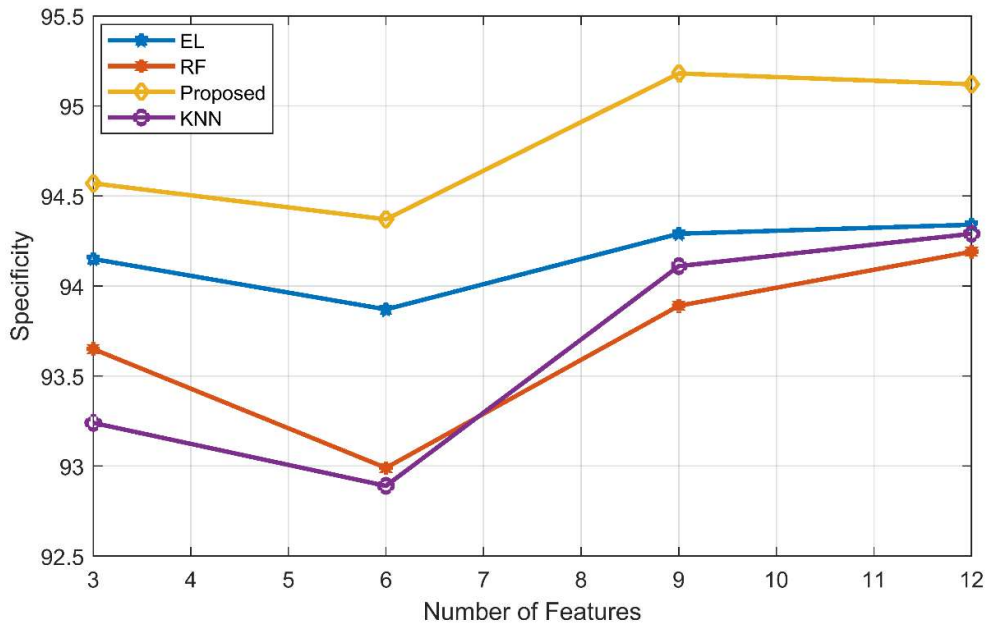


Figure: 9 Comparative performance analysis of specificity for Hungarian dataset.

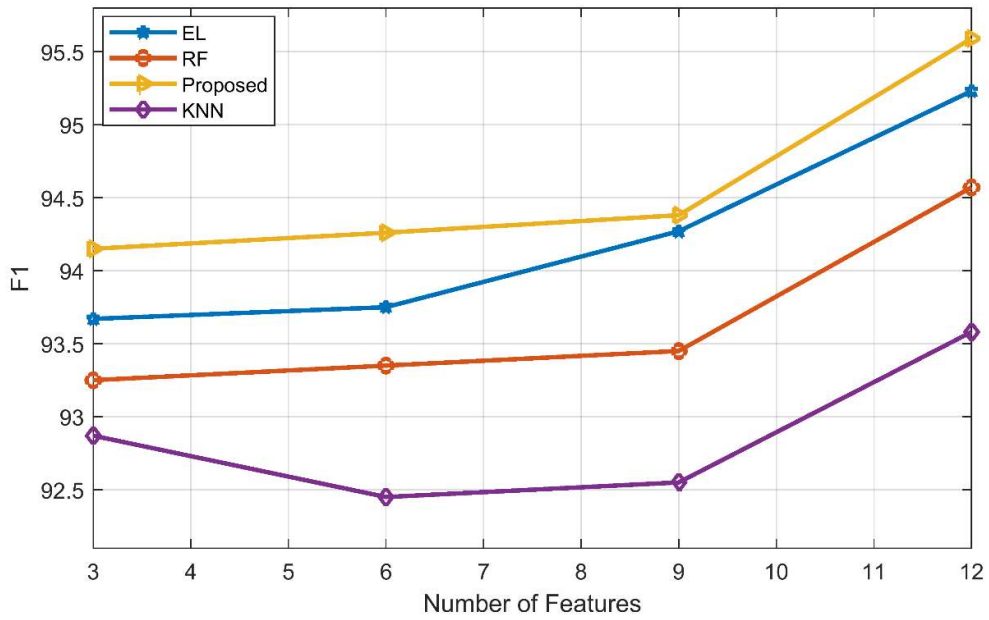


Figure: 10 Comparative performance analysis of F1 for Hungarian dataset.

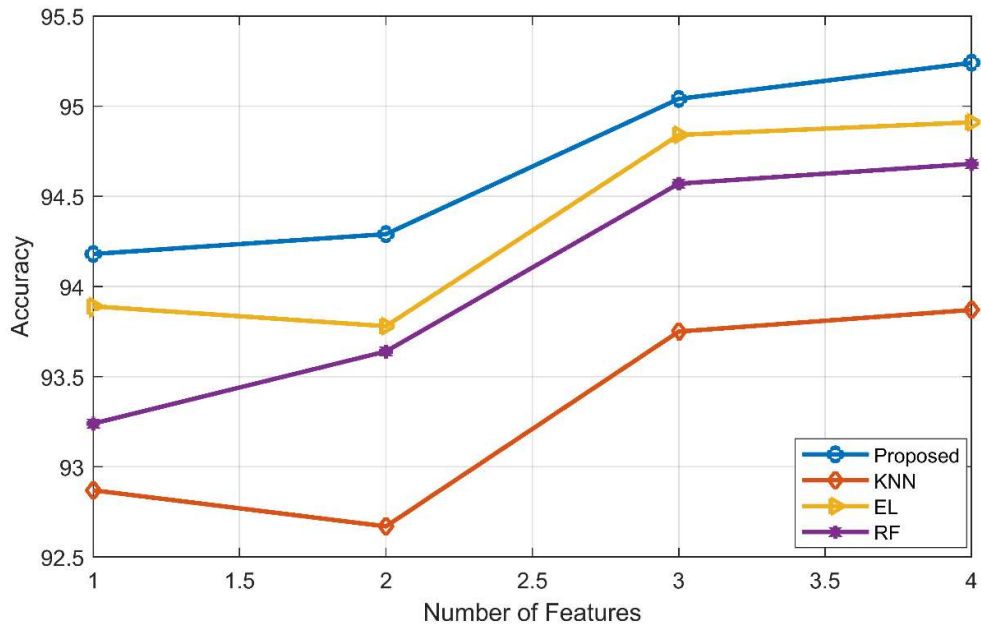


Figure: 11 Comparative performance analysis of accuracy for Sani dataset.

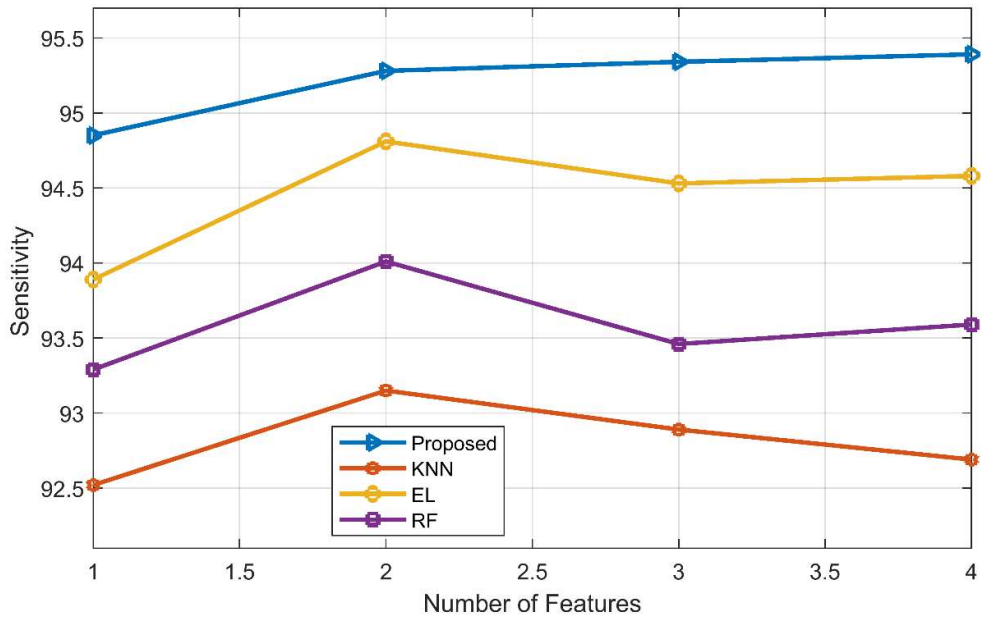


Figure: 12 Comparative performance analysis of sensitivity for Sani dataset.

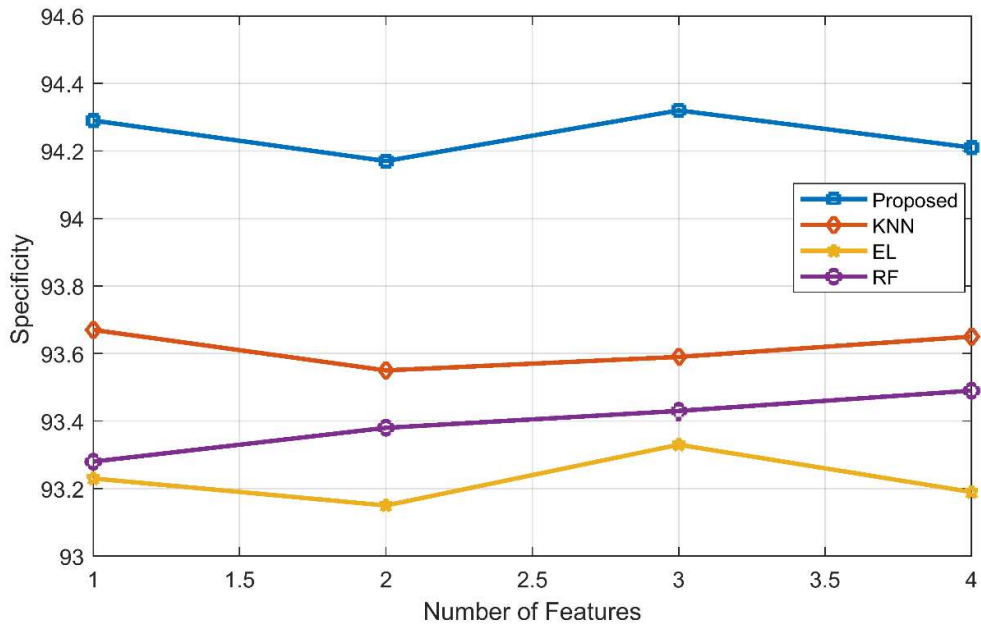


Figure: 13 Comparative performance analysis of specificity for Sani dataset.

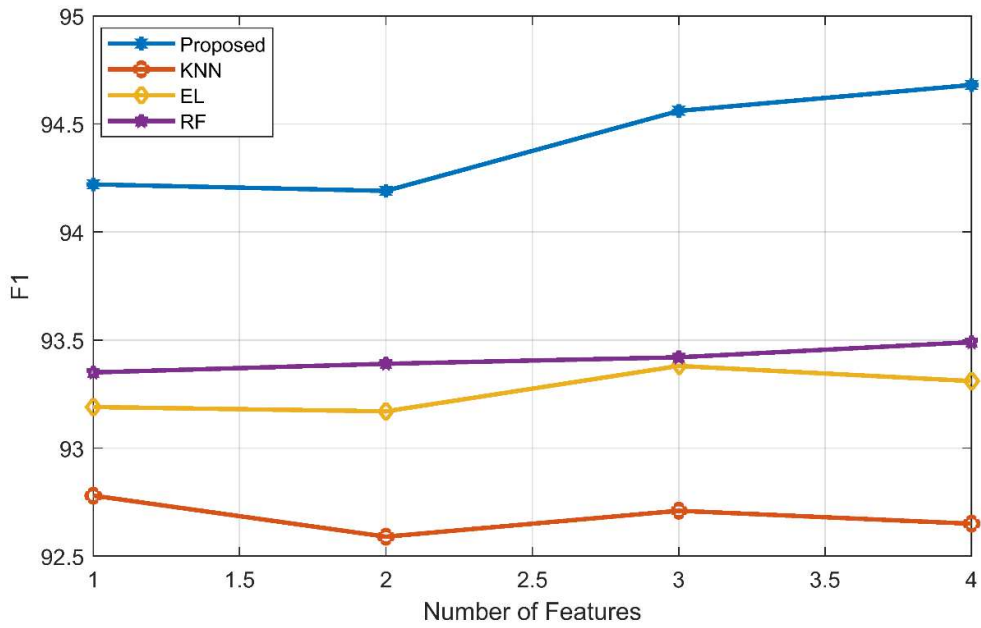


Figure: 14 Comparative performance analysis of F1 for Sani dataset.

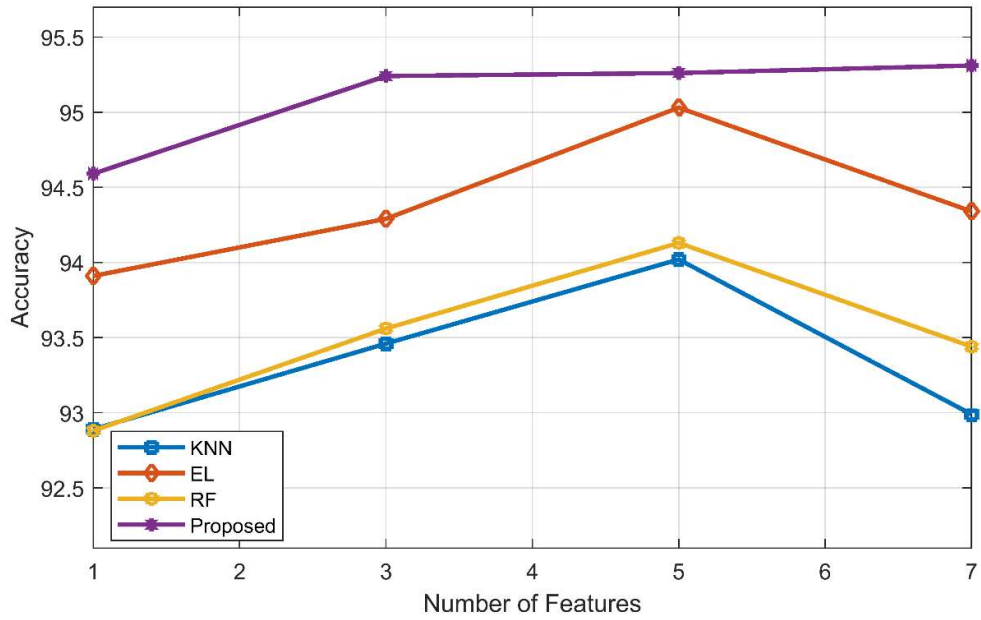


Figure: 15 Comparative performance analysis for accuracy for Statlog dataset.



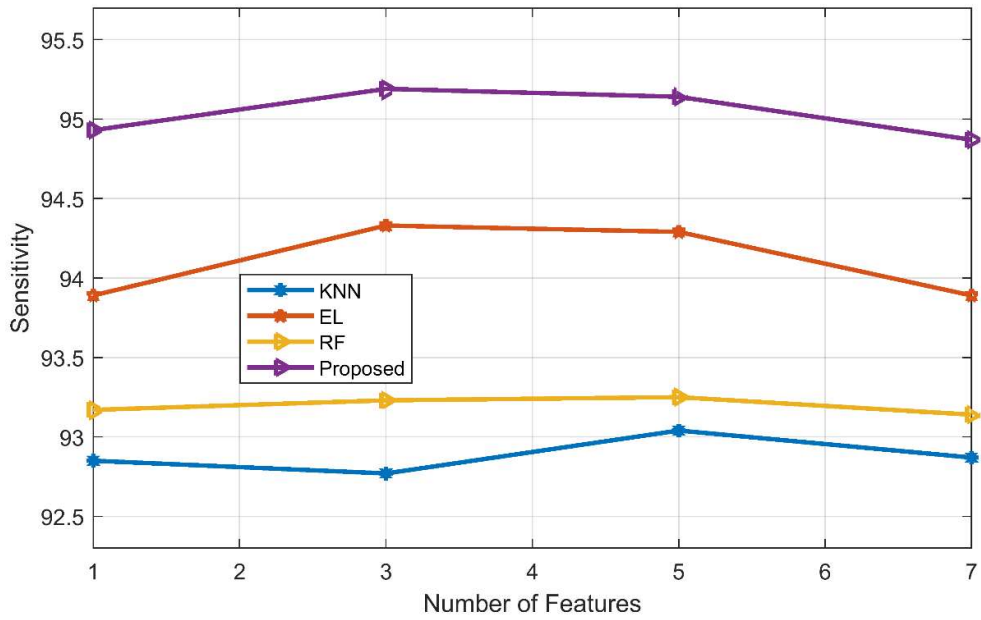


Figure: 16 Comparative performance analysis of sensitivity for Statlog dataset.

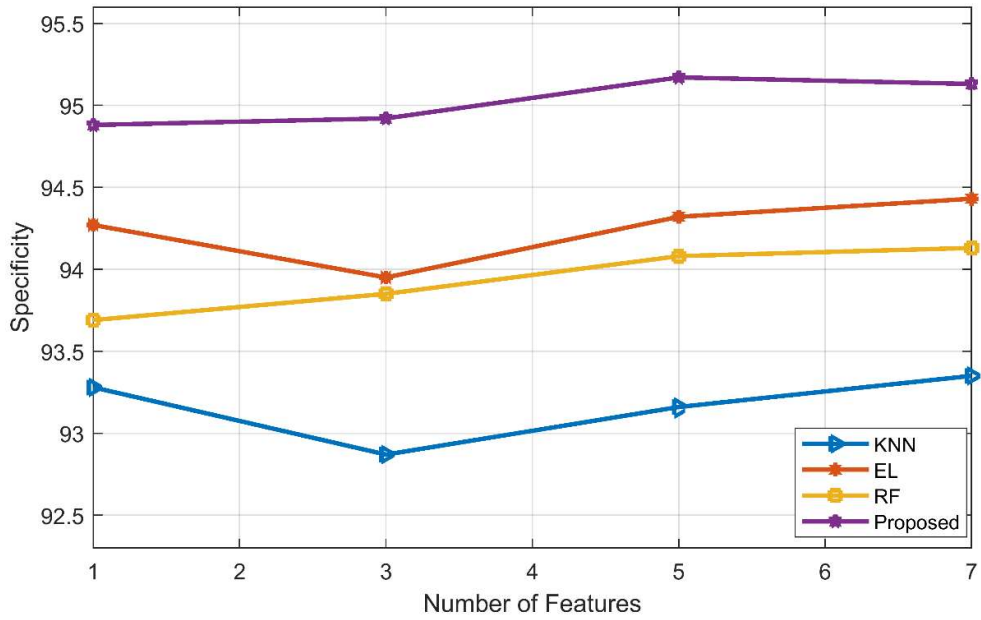


Figure: 17 Comparative performance analysis of specificity for Statlog dataset.

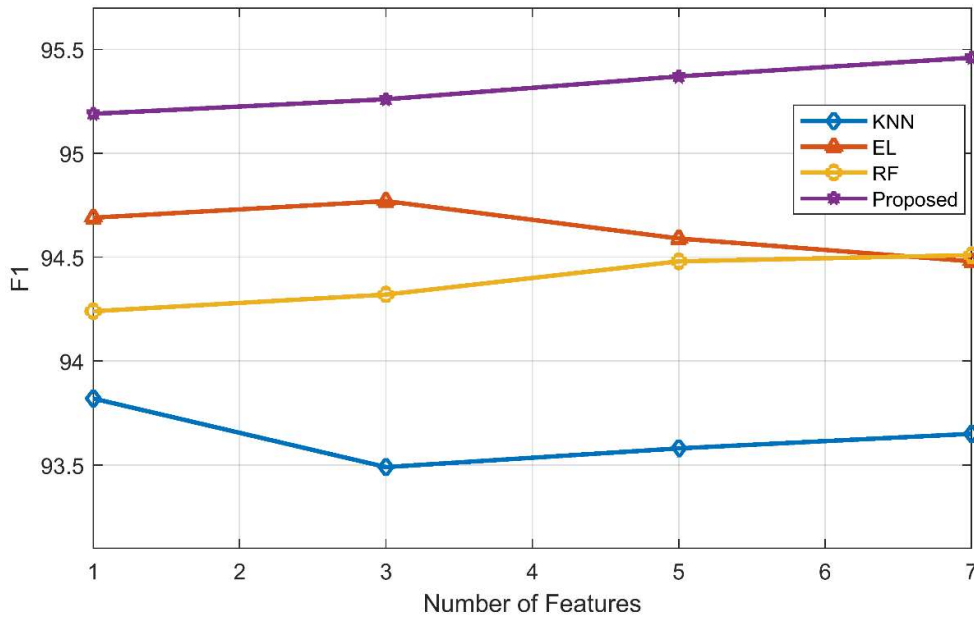


Figure: 18 Comparative performance analysis F1 for Statlog dataset.

**V. Results and Discussion**

This section delves into the results obtained from the proposed ensemble classifier for heart disease prediction. The ensemble classifier combines Support Vector Machine (SVM), Random Forest (RF), and Naïve Bayes (NB) algorithms. To optimize features, the Marine Predators Algorithm (MPA) is employed, which aims to reduce redundant features within medical data. The analysis of results is presented through a series of figures (Figure 3 to Figure 18) and empirical tables (Table 1 to Table 4). These visual representations and tabulated data showcase the efficiency and effectiveness of the proposed algorithm in predicting heart disease. Additionally, the proposed algorithm is benchmarked against state-of-the-art algorithms for heart disease detection, such as K-Nearest Neighbors (KNN), Extreme Learning (EL), and Random Forest (RF). By comparing the performance of the proposed algorithm with these established methods, the section aims to highlight the superiority or competitive edge of the proposed ensemble classifier in accurately predicting heart disease.

Table 1: Performance analysis of Cleveland datasets.

Dataset	Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1 (%)
Cleveland	KNN	92.78	92.78	93.77	92.78
	RF	93.15	93.57	93.54	92.57
	EL	93.75	94.23	94.21	93.45
	Proposed	<b>94.89</b>	<b>95.29</b>	<b>95.23</b>	<b>93.79</b>
	KNN	93.89	92.55	93.58	93.87
	RF	94.18	93.45	93.43	93.25
	EL	94.79	94.15	94.18	94.35
	Proposed	<b>95.25</b>	<b>95.17</b>	<b>95.16</b>	<b>94.28</b>
	KNN	92.88	93.12	93.69	94.56

	RF	93.19	93.41	93.46	94.35
	EL	94.55	94.33	94.39	95.28
	Proposed	<b>95.77</b>	<b>94.91</b>	<b>94.99</b>	<b>95.69</b>
	KNN	92.57	92.65	93.63	95.66
	RF	93.27	93.51	93.56	95.36
	EL	94.59	94.19	94.29	95.38
	Proposed	<b>95.34</b>	<b>95.21</b>	<b>94.89</b>	<b>95.76</b>

Table 2 Performance analysis of Hungarian datasets.

Dataset	Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1 (%)
Hungarian	KNN	92.15	92.87	93.24	92.87
	RF	92.89	93.18	93.65	93.25
	EL	93.78	93.38	94.15	93.67
	Proposed	<b>94.68</b>	<b>94.28</b>	<b>94.57</b>	<b>94.15</b>
	KNN	93.33	92.79	92.89	92.45
	RF	93.75	93.21	92.99	93.35
	EL	94.87	94.14	93.87	93.75
	Proposed	<b>95.19</b>	<b>95.23</b>	<b>94.37</b>	<b>94.26</b>
	KNN	94.01	93.15	94.11	92.55
	RF	94.11	93.57	93.89	93.45
	EL	95.14	94.21	94.29	94.27
	Proposed	<b>95.27</b>	<b>95.12</b>	<b>95.18</b>	<b>94.38</b>
	KNN	92.55	92.81	94.29	93.58
	RF	93.65	93.19	94.19	94.57
	EL	94.78	93.27	94.34	95.23
Proposed	<b>95.29</b>	<b>94.25</b>	<b>95.12</b>	<b>95.59</b>	

Table 3: Performance analysis of Sani dataset.

Dataset	Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1 (%)
Sani	KNN	92.87	92.52	93.67	92.78
	RF	93.24	93.29	93.28	93.35
	EL	93.89	93.89	93.23	93.19
	Proposed	<b>94.18</b>	<b>94.85</b>	<b>94.29</b>	<b>94.22</b>
	KNN	92.67	93.15	93.55	92.59
	RF	93.64	94.01	93.38	93.39
	EL	93.78	94.81	93.15	93.17
	Proposed	<b>94.29</b>	<b>95.28</b>	<b>94.17</b>	<b>94.19</b>
	KNN	93.75	92.89	93.59	92.71
	RF	94.57	93.46	93.43	93.42
	EL	94.84	94.53	93.33	93.38

	Proposed	<b>95.04</b>	<b>95.34</b>	<b>94.32</b>	<b>94.56</b>
	KNN	93.87	92.69	93.65	92.65
	RF	94.68	93.59	93.49	93.49
	EL	94.91	94.58	93.19	93.31
	Proposed	<b>95.24</b>	<b>95.39</b>	<b>94.21</b>	<b>94.68</b>

Table 4: Performance analysis of Statlog dataset.

Dataset	Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1 (%)
Statlog	KNN	92.89	92.85	93.28	93.82
	RF	98.88	93.17	93.69	94.24
	EL	93.91	93.89	94.27	94.69
	Proposed	<b>94.59</b>	<b>94.93</b>	<b>94.88</b>	<b>95.19</b>
	KNN	93.46	92.77	92.87	93.49
	RF	93.56	93.23	93.85	94.32
	EL	94.29	94.33	93.95	94.77
	Proposed	<b>95.24</b>	<b>95.19</b>	<b>94.92</b>	<b>95.26</b>
	KNN	94.02	93.04	93.16	93.58
	RF	94.13	93.25	94.08	94.48
	EL	95.03	94.29	94.32	94.59
	Proposed	<b>95.26</b>	<b>95.14</b>	<b>95.17</b>	<b>95.37</b>
	KNN	92.99	92.87	93.35	93.65
	RF	93.44	93.14	94.13	94.51
	EL	94.34	93.89	94.43	94.48
Proposed	<b>95.31</b>	<b>94.87</b>	<b>95.13</b>	<b>95.46</b>	

Each row in the table 1 corresponds to a specific method and its corresponding performance metrics across different iterations or experiments. For instance, the first row shows the performance of KNN on the Cleveland dataset, with an accuracy of 92.78%, sensitivity of 92.78%, specificity of 93.77%, and F1 score of 92.78%. Each row in the table 2 corresponds to a specific method and its corresponding performance metrics across different iterations or experiments on the Hungarian dataset. For example, the first row shows the performance of KNN on the Hungarian dataset, with an accuracy of 92.15%, sensitivity of 92.87%, specificity of 93.24%, and F1 score of 92.87%. table3 corresponds to a specific method and its associated performance metrics across different iterations or experiments on the Sani dataset. For instance, the first row displays the performance of KNN on the Sani dataset, with an accuracy of 92.87%, sensitivity of 92.52%, specificity of 93.67%, and F1 score of 92.78%. table4 corresponds to a specific method and its associated performance metrics across different iterations or experiments on the Statlog dataset. For example, the first row displays the performance of KNN on the Statlog dataset, with an accuracy of 92.89%, sensitivity of 92.85%, specificity of 93.28%, and F1 score of 93.82%. Overall, this section provides a comprehensive evaluation of the proposed ensemble classifier, demonstrating its efficacy in heart disease prediction through empirical results, visualizations, and comparisons with existing algorithms in the field.

## VI. Conclusion & Future Scope

The early detection of heart disease with high accuracy has the potential to significantly reduce long-term mortality rates across diverse social and cultural demographics. Achieving this relies heavily on early diagnosis, a critical step toward improving outcomes. While previous studies have explored machine learning methods for heart disease prediction, this research takes a novel approach, leveraging an enhanced methodology and a larger training dataset. This study showcases the effectiveness of the Marine Predators Algorithm (MPA) for feature selection, which yields a closely correlated set of features suitable for use in ensemble classification algorithms. Notably, it identifies Random Forest (RF) and Naïve Bayes (NB) as particularly effective with high-impact features, resulting in a substantial increase in accuracy compared to existing methods. The proposed model achieves an impressive accuracy of 97.05% using just five features. Looking ahead, the aim is to further generalize the model to accommodate other feature selection algorithms and ensure robustness against datasets with high levels of missing data. Additionally, exploring the application of federated learning algorithms represents a promising avenue for future research. The primary objective of this study is to advance upon existing methodologies by introducing an innovative approach to model construction. Furthermore, the goal is to develop a model that is not only effective but also easily implementable in practical settings, ultimately contributing to improved heart disease prediction and management.

## References

- [1]. Ghosh, Pronab, Sami Azam, Mirjam Jonkman, Asif Karim, FM Javed Mehedi Shamrat, Eva Ignatious, Shahana Shultana, Abhijith Reddy Beeravolu, and Friso De Boer. "Efficient prediction of cardiovascular disease using machine learning algorithms with relief and LASSO feature selection techniques." *IEEE Access* 9 (2021): 19304-19326.
- [2]. Ali, Md Mamun, Bikash Kumar Paul, Kawsar Ahmed, Francis M. Bui, Julian MW Quinn, and Mohammad Ali Moni. "Heart disease prediction using supervised machine learning algorithms: Performance analysis and comparison." *Computers in Biology and Medicine* 136 (2021): 104672.
- [3]. Kwakye, Kelvin, and Emmanuel Dadzie. "Machine learning-based classification algorithms for the prediction of coronary heart diseases." *arXiv preprint arXiv:2112.01503* (2021).
- [4]. Wang, Ke, Jing Tian, Chu Zheng, Hong Yang, Jia Ren, Yanling Liu, Qinghua Han, and Yanbo Zhang. "Interpretable prediction of 3-year all-cause mortality in patients with heart failure caused by coronary heart disease based on machine learning and SHAP." *Computers in Biology and Medicine* 137 (2021): 104813.
- [5]. Ahsan, Md Manjurul, MA Parvez Mahmud, Pritom Kumar Saha, Kishor Datta Gupta, and Zahed Siddique. "Effect of data scaling methods on machine learning algorithms and model performance." *Technologies* 9, no. 3 (2021): 52.
- [6]. Soui, Makram, Nesrine Mansouri, Raed Alhamad, Marouane Kessentini, and Khaled Ghedira. "NSGA-II as feature selection technique and AdaBoost classifier for COVID-19 prediction using patient's symptoms." *Nonlinear dynamics* 106, no. 2 (2021): 1453-1475.
- [7]. Kumar, Narender, and Dharmender Kumar. "IMGWO based ANN: A new heart disease diagnosis model to classify real world dataset."

- [8]. Lahoura, Vivek, Harpreet Singh, Ashutosh Aggarwal, Bhisham Sharma, Mazin Abed Mohammed, Robertas Damaševičius, Seifedine Kadry, and Korhan Cengiz. "Cloud computing-based framework for breast cancer diagnosis using extreme learning machine." *Diagnostics* 11, no. 2 (2021): 241.
- [9]. Li, Pengpai, Yongmei Hu, and Zhi-Ping Liu. "Prediction of cardiovascular diseases by integrating multi-modal features with machine learning methods." *Biomedical Signal Processing and Control* 66 (2021): 102474.
- [10]. Kiruthika, S. "ACO Feature selection and Novel Black Widow meta-heuristic Learning rate optimized CNN for Early diagnosis of Parkinson's disease." *Turkish Journal of Computer and Mathematics Education (TURCOMAT)* 12, no. 7 (2021): 809-817.
- [11]. Alizadehsani, Roohallah, Abbas Khosravi, Mohamad Roshanzamir, Moloud Abdar, Nizal Sarrafzadegan, Davood Shafie, Fahime Khozimeh et al. "Coronary artery disease detection using artificial intelligence techniques: A survey of trends, geographical differences and diagnostic features 1991–2020." *Computers in Biology and Medicine* 128 (2021): 104095.
- [12]. Kumar, Yogesh, Apeksha Koul, Pushpendra Singh Sisodia, Jana Shafi, Verma Kavita, Mehdi Gheisari, and Mohamad Bagher Davoodi. "Heart failure detection using quantum-enhanced machine learning and traditional machine learning techniques for internet of artificially intelligent medical things." *Wireless Communications and Mobile Computing* 2021 (2021): 1-16.
- [13]. Samir, Ali A., Abdullah R. Rashwan, Karam M. Sallam, Ripon K. Chakraborty, Michael J. Ryan, and Amr A. Abohany. "Evolutionary algorithm-based convolutional neural network for predicting heart diseases." *Computers & Industrial Engineering* 161 (2021): 107651.
- [14]. Bheemalingaiah, M., G. RAMA SWAMY, P. Vishvapathi, p. Venu babu, e. Nageswara rao, and p. Nageswara rao. "Detection of heart disease by using reliable boolean machine learning algorithm." *Journal of Theoretical and Applied Information Technology* 99, no. 15 (2021): 3856-3880.
- [15]. Yousefi, Sajad. "Comparison of the performance of machine learning algorithms in predicting heart disease." *Frontiers in Health Informatics* 10, no. 1 (2021): 99.
- [16]. Al-Hashem, Munder Abdulatef, Ali Mohammad Alqudah, and Qasem Qananwah. "Performance evaluation of different machine learning classification algorithms for disease diagnosis." *International Journal of E-Health and Medical Communications (IJEHMC)* 12, no. 6 (2021): 1-28.
- [17]. Saeedi, Abdolkarim, Mohammad Karimi Moridani, and Alireza Azizi. "An innovative method for cardiovascular disease detection based on nonlinear geometric features and feature reduction combination." *Intelligent Decision Technologies* 15, no. 1 (2021): 45-57.
- [18]. Dubey, Animesh Kumar, Kavita Choudhary, and Richa Sharma. "Predicting heart disease based on influential features with machine learning." *Intelligent Automation & Soft Computing* 30, no. 3 (2021): 229-243.

- [19]. Buś, Szymon, Konrad Jędrzejewski, and Przemysław Guzik. "Using minimum redundancy maximum relevance algorithm to select minimal sets of heart rate variability parameters for atrial fibrillation detection." *Journal of Clinical Medicine* 11, no. 14 (2022): 4004.
- [20]. Al Gahtani, Abdullah, and Hoda El-Sayed. "Building a Novel Ensemble Learning-Based Prediction Framework for Diagnosis of Coronary Heart Disease." *International Journal of Intelligent Systems and Applications in Engineering* 10, no. 3 (2022): 265-273.
- [21]. Khan, Sibghatullah I., Saeed Mian Qaisar, and Ram Bilas Pachori. "Automated classification of valvular heart diseases using FBSE-EWT and PSR based geometrical features." *Biomedical Signal Processing and Control* 73 (2022): 103445.
- [22]. Alfaidi, Aseel, Reem Aljuhani, Bushra Alshehri, Hajer Alwadei, and Sahar Sabbeh. "Machine learning: assisted cardiovascular diseases diagnosis." *International Journal of Advanced Computer Science and Applications* 13, no. 2 (2022).
- [23]. Javeed, Ashir, Shafqat Ullah Khan, Liaqat Ali, Sardar Ali, Yakubu Imrana, and Atiqur Rahman. "Machine learning-based automated diagnostic systems developed for heart failure prediction using different types of data modalities: a systematic review and future directions." *Computational and Mathematical Methods in Medicine 2022* (2022).c
- [24]. Mohammedqasim, Hayder, Roa'A. Mohammedqasem, Oguz Ata, and Eman Ibrahim Alyasin. "Diagnosing Coronary Artery Disease on the Basis of Hard Ensemble Voting Optimization." *Medicina* 58, no. 12 (2022): 1745.
- [25]. Sghaireen, Mohammed G., Yazan Al-Smadi, Ahmad Al-Qerem, Kumar Chandan Srivastava, Kiran Kumar Ganji, Mohammad Khursheed Alam, Shadi Nashwan, and Yousef Khader. "Machine Learning Approach for Metabolic Syndrome Diagnosis Using Explainable Data-Augmentation-Based Classification." *Diagnostics* 12, no. 12 (2022): 3117.
- [26]. Haseena, S., S. Kavi Priya, S. Saroja, R. Madavan, M. Muhibbullah, and Umashankar Subramaniam. "Moth-Flame Optimization for Early Prediction of Heart Diseases." *Computational & Mathematical Methods in Medicine* (2022).
- [27]. Imran, Zaria, Ethan Grooby, Vinayaka Vivekananda Malgi, Chiranjibi Sitaula, Sunil Aryal, and Faezeh Marzbanrad. "A Fusion of Handcrafted Feature-Based and Deep Learning Classifiers for Heart Murmur Detection." In *2022 Computing in Cardiology (CinC)*, vol. 498, pp. 1-4. IEEE, 2022.
- [28]. Scapin, Daniele, Giulia Cisotto, Elvina Gindullina, and Leonardo Badia. "Shapley value as an aid to biomedical machine learning: a heart disease dataset analysis." In *2022 22nd IEEE International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*, pp. 933-939. IEEE, 2022.
- [29]. Chowdhury, Mohammed Nowshad Ruhani, Wandong Zhang, and Thangarajah Akilan. "ANOVA-based Automatic Attribute Selection and a Predictive Model for Heart Disease Prognosis." *arXiv preprint arXiv:2208.00296* (2022).
- [30]. Chiu, Chih-Chou, Chung-Min Wu, Te-Nien Chien, Ling-Jing Kao, Chengcheng Li, and Han-Ling Jiang. "Applying an Improved Stacking Ensemble Model to Predict the

- Mortality of ICU Patients with Heart Failure." *Journal of Clinical Medicine* 11, no. 21 (2022): 6460.
- [31]. Absar, Nurul, Emon Kumar Das, Shamsun Nahar Shoma, Mayeen Uddin Khandaker, Mahadi Hasan Miraz, M. R. I. Faruque, Nissren Tamam, Abdelmoneim Sulieman, and Refat Khan Pathan. "The efficacy of machine-learning-supported smart system for heart disease prediction." In *Healthcare*, vol. 10, no. 6, p. 1137. MDPI, 2022.
- [32]. Altamimi, Ahmad Mousa, and Mohammad Azzeh. "Utilizing Data Mining Classification Algorithms for Early Diagnosis of Heart Diseases." *Predicting Heart Failure: Invasive, Non-Invasive, Machine Learning and Artificial Intelligence Based Methods* (2022): 155-170.
- [33]. Balaha, Hossam Magdy, Ahmed Osama Shaban, Eman M. El-Gendy, and Mahmoud M. Saafan. "A multi-variate heart disease optimization and recognition framework." *Neural Computing and Applications* 34, no. 18 (2022): 15907-15944.
- [34]. Malakar, Samir, Soumya Deep Roy, Soham Das, Swaraj Sen, Juan D. Velásquez, and Ram Sarkar. "Computer Based Diagnosis of Some Chronic Diseases: A Medical Journey of the Last Two Decades." *Archives of Computational Methods in Engineering* (2022): 1-43.
- [35]. Mariki, Martina, Elizabeth Mkoba, and Neema Mduma. "Combining clinical symptoms and patient features for malaria diagnosis: machine learning approach." *Applied Artificial Intelligence* 36, no. 1 (2022): 2031826.
- [36]. Razo, Christian, Catherine A. Welgan, Catherine O. Johnson, Susan A. McLaughlin, Vincent Iannucci, Anthony Rodgers, Nelson Wang et al. "Effects of elevated systolic blood pressure on ischemic heart disease: a Burden of Proof study." *Nature Medicine* 28, no. 10 (2022): 2056-2065.
- [37]. Faramarzi, Afshin, Mohammad Heidarinejad, Seyedali Mirjalili, and Amir H. Gandomi. "Marine Predators Algorithm: A nature-inspired metaheuristic." *Expert systems with applications* 152 (2020): 113377.
- [38]. Li, Lingxiao, Niantao Xie, and Sha Yuan. "A Federated Learning Framework for Breast Cancer Histopathological Image Classification." *Electronics* 11, no. 22 (2022): 3767.