

Article

Ensemble Heuristic–Metaheuristic Feature Fusion Learning for Heart Disease Diagnosis Using Tabular Data

Mohammad Shokouhifar ^{1,*}, Mohamad Hasanvand ², Elaheh Moharamkhani ^{3,4} and Frank Werner ^{5,*}¹ Department of Electrical and Computer Engineering, Shahid Beheshti University, Tehran 1983969411, Iran² Department of Computer Engineering, University of Mohaghegh Ardabili, Ardabil 5619911367, Iran; mohamadhasanvand5691@gmail.com³ Department of Computer Engineering, Institute of Higher Education Saeb, Abhar 3697345619, Iran; elaheh.moharamkhani@uoh.edu.iq⁴ IRO, Computer Science Department, University of Halabja, Halabja 46018, Iraq⁵ Faculty of Mathematics, Otto-von-Guericke University, 39016 Magdeburg, Germany

* Correspondence: m_shokouhifar@sbu.ac.ir (M.S.); frank.werner@ovgu.de (F.W.)

Abstract: Heart disease is a global health concern of paramount importance, causing a significant number of fatalities and disabilities. Precise and timely diagnosis of heart disease is pivotal in preventing adverse outcomes and improving patient well-being, thereby creating a growing demand for intelligent approaches to predict heart disease effectively. This paper introduces an ensemble heuristic–metaheuristic feature fusion learning (EHMFFL) algorithm for heart disease diagnosis using tabular data. Within the EHMFFL algorithm, a diverse ensemble learning model is crafted, featuring different feature subsets for each heterogeneous base learner, including support vector machine, K-nearest neighbors, logistic regression, random forest, naive bayes, decision tree, and XGBoost techniques. The primary objective is to identify the most pertinent features for each base learner, leveraging a combined heuristic–metaheuristic approach that integrates the heuristic knowledge of the Pearson correlation coefficient with the metaheuristic-driven grey wolf optimizer. The second objective is to aggregate the decision outcomes of the various base learners through ensemble learning. The performance of the EHMFFL algorithm is rigorously assessed using the Cleveland and Statlog datasets, yielding remarkable results with an accuracy of 91.8% and 88.9%, respectively, surpassing state-of-the-art techniques in heart disease diagnosis. These findings underscore the potential of the EHMFFL algorithm in enhancing diagnostic accuracy for heart disease and providing valuable support to clinicians in making more informed decisions regarding patient care.

Keywords: heart disease diagnosis; ensemble learning; feature selection; heuristics; metaheuristics; Pearson correlation coefficient (PCC); grey wolf optimizer (GWO)



Citation: Shokouhifar, M.; Hasanvand, M.; Moharamkhani, E.; Werner, F. Ensemble Heuristic–Metaheuristic Feature Fusion Learning for Heart Disease Diagnosis Using Tabular Data. *Algorithms* **2024**, *17*, 34. <https://doi.org/10.3390/a17010034>

Academic Editors: Roberto Montemanni and Günther Raidl

Received: 27 October 2023
Revised: 18 December 2023
Accepted: 12 January 2024
Published: 14 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Currently, a person’s workload has significantly increased as a result of more work. There is a great likelihood that the person will get heart disease as a result of this terrible situation, which cannot be avoided [1–3]. Heart diseases are brought on by a reduction in the amount of blood circulating to the brain, heart, lungs, and other vital organs. The most prevalent and least serious kind of cardiovascular illness is congestive heart failure. Blood is transported to the heart by blood veins in the human anatomy. Defective heart valves, which can cause heart failure, are one of the additional causes of heart disease. Anesthesia may also be present together with upper abdominal muscle pain, which is a characteristic indication of heart illness. It is advised to reduce blood pressure, lower cholesterol, and exercise frequently to reduce the risk of heart disease. Angina pectoris, dilated cardiomyopathy, stroke, and congestive heart failure are among the conditions most closely associated with heart disease. As a result, it is important to keep an eye on indicators of cardiovascular disease and speak with medical professionals [4–6].

Cardiovascular diseases are one of the most prevalent causes of global mortality, and their diagnosis and prediction have consistently posed substantial challenges due to their dynamic nature. Risk factors contributing to the elevated risk of heart disease encompass age, gender, smoking habits, family medical history, cholesterol levels, poor dietary choices, high blood pressure, obesity, physical inactivity, and alcohol consumption. Additionally, hereditary factors like high blood pressure and diabetes heighten susceptibility to heart disease. Certain risk factors can be influenced by individual choices. In conjunction with the aforementioned risk factors, lifestyle decisions, such as dietary patterns, sedentary behavior, and obesity, are recognized as significant contributors [7–9]. Heart conditions manifest in various forms, including myocarditis, angina pectoris, congestive heart failure, cardiomyopathy, congenital heart disease, and coronary heart disease. Manual calculations to assess the likelihood of heart disease based on these risk factors are intricate, necessitating the adoption of computer-assisted techniques for efficient and accurate evaluation [10].

Machine learning is effective for a wide range of problems. Utilizing the values of independent variables to predict the value of a dependent variable is one use for this technique. Since the healthcare industry has huge data resources that are challenging to manage manually, it is an application area for data mining. Even in wealthy nations, heart disease has been found to be one of the leading causes of death. The hazards are either not recognized or are not recognized until much later, which is one of the causes of fatalities from heart disease. Machine learning techniques, on the other hand, can be helpful in overcoming this issue and enabling early risk prediction [11].

In this study, we introduce an advanced method for detecting and predicting heart disease patients using ensemble learning, feature selection, and heuristic–metaheuristic optimization. The presented method has two stages. In the first stage, we utilize a combined heuristic–metaheuristic feature selection algorithm based on the Pearson correlation coefficient (PCC) and the grey wolf optimizer (GWO), called the PCC–GWO, to increase the accuracy and performance of each machine learning model. In the second stage, a heterogeneous ensemble learning model is applied to generate the final outputs based on the aggregation of the opinion of the different base learners. As a result, the following significantly contribute to this evolved diagnosis model of heart disease:

- The introduction of an advanced ensemble heuristic–metaheuristic feature fusion learning (EHMFFL) algorithm as a robust model in predicting heart diseases;
- The construction of a heterogeneous ensemble learning model for heart disease diagnosis comprising seven base learners: support vector machine (SVM), K-nearest neighbors (KNNs), logistic regression (LR), random forest (RF), naive bayes (NB), decision tree (DT), and eXtreme Gradient Boosting (XGBoost) techniques;
- The presentation of a combined heuristic–metaheuristic algorithm (called PCC–GWO) to select an optimal feature subset for each machine learning model, separately. In the PCC–GWO model, at first, the PCC is used to calculate an importance score for each feature. Then, these scores are used as heuristic knowledge to guide the search process of the GWO for obtaining the best achievable feature subset;
- The analysis of the relationships between different variables within the cardiovascular datasets using a correlation heat map (CHM), and the evaluation of the performance of the EHMFFL algorithm using different measures: accuracy, precision, recall, F1 score, specificity, and the receiver operating characteristic (ROC);
- The successful development of the EHMFFL algorithm in MATLAB R2022b for heart disease prediction using the Cleveland and Statlog datasets, respectively.

The rest of this paper is organized as follows: In the Section 2, we examine related works. The Section 3 provides details on the two datasets used in this paper. The proposed EHMFFL algorithm is introduced in the Section 4. The results are provided and assessed in the Section 5 and, finally, concluding remarks are presented in the Section 6.

2. Literature Review

In this section, we delve into the realm of machine learning, ensemble learning, and deep learning techniques. Machine learning methods for classification are widely adopted across various industries, and researchers are continually working on advancing their categorization capabilities. One such approach is ensemble learning, which can be either homogeneous or heterogeneous. Early techniques, such as bootstrap aggregating (bagging) [12] and boosting [13], exemplify the power of ensemble learning, often leading to improved classification performance when implemented. In addition to these, various strategies have been explored by researchers, including methods like majority voting, to effectively combine multiple classifiers or partitions for enhanced results.

2.1. Machine Learning Approaches

Miao and Miao [14] underscored the critical significance of early detection and diagnosis of coronary heart disease (CHD), a leading global cause of mortality. To facilitate the training and evaluation of diverse deep neural network (DNN) architectures, including convolutional neural networks and recurrent neural networks, they curated a comprehensive dataset comprising 303 patients and 14 clinical attributes, encompassing factors like age, gender, and cholesterol levels. Their results demonstrated that the proposed DNN models outperformed established methods like logistic regression and decision trees, showcasing high accuracy in CHD detection. Furthermore, feature importance analysis revealed that age, maximum heart rate, and ST segment depression, were the three most critical variables for predicting CHD.

Vijayashree and Sultana [15] introduced a machine learning framework designed for feature selection in heart disease classification, leveraging an enhanced particle swarm optimization (PSO) algorithm in conjunction with an SVM classifier. The innovative PSO algorithm, crafted with the unique blend of a hybrid mutation operator, velocity clamping, and adaptive inertia weight, aimed to overcome the limitations of conventional PSO methods. Evaluating the framework using the Cleveland heart disease dataset, the results showcased its superiority over alternative feature selection techniques. Notably, the framework exhibited a high degree of accuracy in classifying heart disease, underscoring its potential for improving the accuracy and effectiveness of heart disease diagnosis.

Waigi et al. [16] presented a study focused on predicting the risk of heart disease by employing advanced machine learning techniques. The research explores innovative approaches to risk assessment in cardiovascular health, utilizing a diverse range of machine learning algorithms. By leveraging extensive data and applying advanced analytics, the study aims to enhance the accuracy and effectiveness of heart disease risk prediction. This work contributes to the field of cardiovascular medicine and underscores the potential of machine learning in improving heart disease risk assessment and patient care.

Tuli et al. [17] presented HealthFog, a smart healthcare system that used ensemble deep learning techniques for the autonomous diagnosis of cardiac illnesses in an integrated Internet of Things (IoT) and fog computing environment. The system was able to effectively diagnose heart illnesses by processing real-time data from numerous sensors and devices, including blood pressure monitors and electrocardiogram (ECG) devices. The HealthFog system's patient monitoring module, data preprocessing module, feature extraction and selection module, and classification module, were all covered in the authors' full architecture presentation. The findings demonstrated that the HealthFog system performed better than other current systems in terms of precision and timeliness.

Jindal et al. [18] focused on heart disease prediction through the application of numerous algorithms, including KNNs, LR, and RF. Their research explores the utilization of these algorithms to enhance the accuracy of heart disease risk assessment and prediction. By leveraging advanced data analytics and machine learning techniques, the study aims to contribute to the field of cardiovascular medicine and improve the effectiveness of heart disease prediction, potentially leading to better patient care.

Sarra et al. [19] reported on a study that used machine learning and statistical analysis to increase the precision of heart disease prediction. They chose the most important candidate features from a list of candidate features using the two statistical models. On the basis of the chosen features, they then applied a support vector machine to create prediction models. According to the findings, the two statistical models and the SVM combination had the highest level of success in predicting heart disease.

Aliyar Vellameeran and Brindha [20] introduced a new type of deep belief network (DBN) for diagnosing heart disease utilizing IoT wearable medical devices that was supported with optimal feature selection. The main objective of the study was to train the DBN model by analyzing and selecting the most important features from a big dataset. The proposed method was evaluated using actual data gathered from wearable medical devices connected to the Internet of Things, and it has shown promising results in correctly identifying heart disease.

2.2. Ensemble Learning Approaches

In the case of ensemble learning models, Latha and Jeeva [21] examined the effectiveness of several machine learning techniques, including support vector machines, decision trees, and random forests. They contrasted the distinct methods with an ensemble method that brought these models together. The results showed that the ensemble method outperformed the individual algorithms in terms of prediction accuracy, sensitivity, and specificity. The study also emphasized the importance of feature selection in raising the model's accuracy.

Ali et al. [22] have innovated a smart healthcare monitoring system designed to integrate multiple clinical data sources for accurate heart disease prediction. This system employs a combination of deep learning models, outperforming traditional methods in terms of accuracy. A standout feature of this system is its real-time patient data monitoring capability, facilitating timely intervention and heart disease prevention. By incorporating ECG readings, blood pressure, body temperature, and other pertinent clinical factors, the system provides precise cardiac illness prognosis.

Shorewala [23] delved into the realm of coronary heart disease early detection, with a specific focus on harnessing the potential of ensemble methods. They pinpointed the most effective approach for early disease detection by rigorously analyzing a spectrum of models and algorithms, including DT, RF, SVM, KNNs, and artificial neural networks (ANNs). The results underscore the superiority of ensemble approaches, which seamlessly integrate multiple algorithms, yielding the highest accuracy in disease prediction. This research highlights the significance of ensemble techniques in enhancing early detection capabilities for coronary heart disease.

Ghasemi Darehnaei et al. [24] introduced an approach known as swarm intelligence ensemble deep transfer learning (SI-EDTL), designed for the task of multiple vehicle detection in images captured by unmanned aerial vehicles (UAVs). This method combines the power of swarm intelligence algorithms and deep transfer learning to enhance the accuracy of vehicle detection in UAV imagery. The research demonstrated the effectiveness of SI-EDTL, offering a solution for the challenging task of detecting multiple vehicles in aerial images, which has significant applications in fields such as surveillance and autonomous navigation.

Shokouhifar et al. [25] have presented a novel approach for accurately measuring arm volume in patients with lymphedema. This method utilized a three-stage ensemble deep learning framework empowered by swarm intelligence techniques. By combining the power of deep learning and swarm intelligence, the research aimed to enhance the precision of arm volume measurement, which is crucial in the diagnosis and management of lymphedema. The proposed model demonstrated promising results, showcasing its potential to improve healthcare outcomes for individuals with lymphedema by providing more accurate and reliable measurements of arm volume.

2.3. Feature Selection Algorithms

There are also various feature selection techniques applied for the enhancement of prediction accuracy in heart diseases. For example, Nagarajan et al. [26] introduce a feature selection and classification model tailored for the prediction of heart disease. The research explores advanced techniques for selecting relevant features and enhancing the accuracy of heart disease prediction. Their results showed that this technique can efficiently improve the effectiveness of early detection and risk assessment for heart disease, potentially benefiting both patients and healthcare providers.

Al-Yarimi et al. [27] presented a heart disease prediction model using supervised learning techniques. The focus of their study was on feature optimization, where they employ discrete weights to enhance the accuracy of heart disease prediction models. By selecting and assigning weights to relevant features, the research aims to improve the efficiency and precision of predictive models in diagnosing heart disease.

Ahmad et al. [28] conducted a comparative investigation on the optimal medical diagnosis of human heart disease using machine learning techniques. They specifically examined the impact of sequential feature selection, comparing its inclusion with conventional machine learning approaches that do not employ this feature selection method. The research aimed to enhance the efficiency and accuracy of heart disease diagnosis through the identification of the most relevant features. They provided some insights into the utility of sequential feature selection in improving the performance of machine learning-based heart disease diagnostic models.

Pathan et al. [29] proposed an analysis to assess the influence of feature selection on the accuracy of heart disease prediction. The study specifically focused on understanding how different feature selection techniques could enhance or affect the accuracy of predictive models for heart disease. By investigating the impact of feature selection, the research aimed to optimize the heart disease prediction model.

Zhang et al. [30] developed a heart disease prediction model that combines feature selection methods with deep neural networks. The research focused on optimizing the feature selection process to enhance the accuracy of heart disease prediction. By utilizing deep neural networks, they achieved more efficient results for diagnosing heart disease, which resulted in the development of diagnostic tools for heart disease diagnosis.

2.4. Our Contributions Compared with the Literature

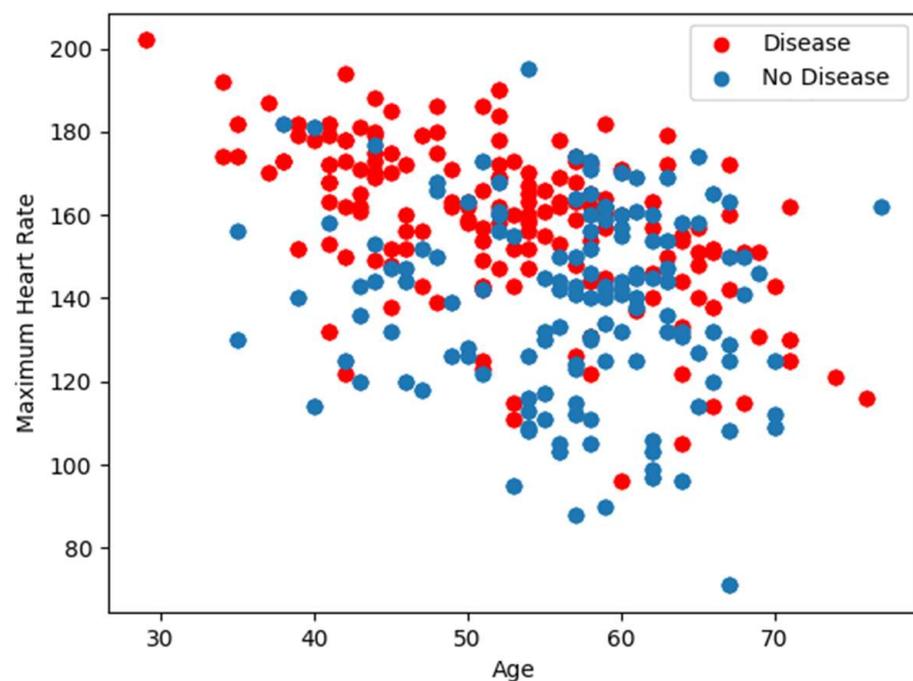
This paper addresses a significant gap in the existing literature by introducing an innovative EHMFFL algorithm for heart disease diagnosis using tabular data. The EHMFFL approach stands out by seamlessly integrating ensemble learning and feature fusion into a comprehensive framework. The EHMFFL algorithm not only leverages an ensemble of base learners, including SVM, KNNs, LR, RF, NB, DT, and XGBoost techniques, but it also combines the advantages of heuristic–metaheuristic approaches for the selection of a specific feature subset for each base learner within the ensemble learning model. The hybridization of the PCC as a heuristic knowledge source with the metaheuristic-driven GWO sets our combined PCC–GWO feature selection algorithm apart.

3. Data Gathering

In our analysis, we utilize two well-established datasets on cardiac illnesses sourced from the University of California’s Irvine Machine Learning Repository, specifically the Cleveland and Statlog datasets [31,32]. Table 1 details the attributes common to both datasets, with the final attribute serving as an indicator of a person’s heart disease status. To gain deeper insights into the feature distribution, we present Figures 1 and 2, which illustrate the relationship between the maximum heart rate and age, as well as the distribution of the remaining 12 features, respectively.

Table 1. Description of the attributes in the datasets.

Feature	Description	Type	Values
Age	Age of the patients	Numeric	Years
Sex	Gender of the patients	Categorical	M, F
Ca	Number of major vessels	Categorical	0–4
Chol	Serum cholesterol	Numeric	mg/dL
Exang	Exercise induced angina	Categorical	Yes = 1, No = 0
Cp	Chest pain type	Categorical	Male = 1, Female = 0
Oldpeak	ST depression induced by exercise relative to rest	Numeric	0–6.2
Fbs	Fasting blood sugar	Categorical	mg/dL
Restecg	Resting electrocardiographic	Categorical	0, 1, 2
Thal	Normal; Fixed defect; Reversible defect	Categorical	0, 1, 2, 3
Thalach	Maximum heart rate achieved	Numeric	71–202
Slope	Slope of the peak exercise ST segment	Categorical	0, 1, 2
Trestbps	Resting blood pressure	Numeric	94–200
Num	Heart disease status	Categorical	Yes/No

**Figure 1.** Maximum heart rate versus age.

The Cleveland dataset comprises medical records from individuals who underwent heart disease evaluations at the Cleveland Clinic Foundation in the late 1980s, containing 303 instances, each representing a patient, and encompassing 13 features, including critical factors like age, gender, blood pressure, cholesterol level, chest pain presence, and results from various medical tests. This dataset has played a pivotal role in the development and testing of machine learning algorithms aimed at predicting cardiac disease.

The Statlog dataset, part of a dataset collection, consists of 270 instances (patients) with 13 attributes, including age, gender, blood pressure, cholesterol, fasting blood sugar, and various electrocardiography (ECG) and exercise stress test readings. Originally sourced from the Cleveland Clinic Foundation, this dataset has been widely employed in studies related to machine learning algorithms for medical diagnosis. The primary objective is to enable physicians to make more informed treatment decisions by accurately identifying patients based on their feature values.

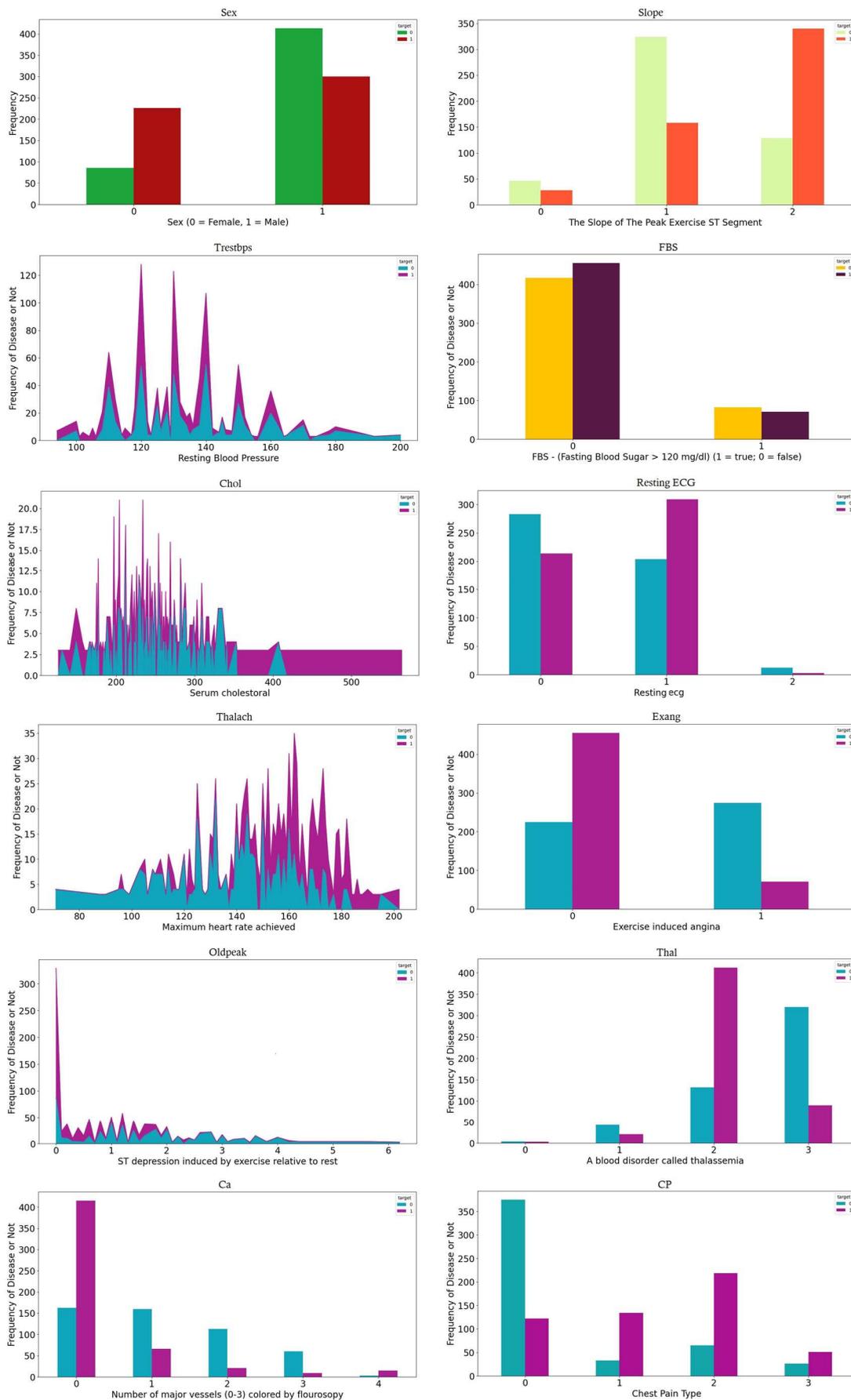


Figure 2. Distribution of all the features.

4. Proposed EHMFFL Algorithm

The proposed EHMFFL algorithm represents a heterogeneous ensemble learning framework, featuring seven base learners, namely SVM, KNNs, LR, RF, NB, DT, and XGBoost techniques. To optimize the performance of each machine learning model, a combined heuristic–metaheuristic algorithm known as the PCC–GWO is performed on each base learner, separately. Initially, the PCC method is employed to calculate the feature importance scores, serving as critical heuristic knowledge for guiding the GWO in selecting the most effective features for the heart disease diagnosis. Subsequently, the tuned machine learning models (SVM, KNNs, LR, RF, NB, DT, and XGBoost) are employed to create the final ensemble learning model. The subsequent sections provide a detailed account of the feature selection process using the PCC–GWO algorithm and the comprehensive classification process with the tuned EHMFFL model.

4.1. Feature Selection Using PCC–GWO

Feature selection is a crucial step in machine learning, particularly when dealing with datasets with high dimensionality. Its primary objective is to streamline the dataset by reducing its dimensionality, thereby identifying the most relevant features that contribute significantly to predictive accuracy, while discarding irrelevant or noisy attributes. This process not only enhances computational efficiency, but also minimizes redundancy among the selected features. Feature selection is essential in various domains, including text categorization, data mining, pattern recognition, and signal processing [33], where it aids in improving model performance by focusing on the most informative attributes and discarding superfluous ones.

Feature selection poses a challenging problem, acknowledged as non-deterministic polynomial hard (NP-hard) [34], making exact (exhaustive) search methods impractical due to their computational complexity and time requirements. Therefore, heuristic and metaheuristic algorithms and their hybridizations become essential in this context [35]. When crafting a metaheuristic algorithm for an NP-hard problem, a delicate balance between exploration and exploitation must be carefully maintained to optimize search algorithms [36]. The GWO is recognized in the literature for its adeptness in striking the right equilibrium between exploration and exploitation. Simultaneously, the PCC stands out as a swift heuristic method for identifying and eliminating highly correlated features [37]. Hence, we have chosen to employ PCC and GWO as the heuristic and metaheuristic components of our integrated PCC–GWO feature selection algorithm. This strategy aims to harness the advantages of both methods concurrently, combining the speed of heuristic-based PCC with the precision of metaheuristic-driven GWO to enhance the feature selection process.

Algorithm 1 outlines the PCC–GWO feature selection approach, offering a hybrid method for selecting an optimal feature subset for each base learner within the ensemble learning model. Initially, the algorithm employs the PCC method to compute an importance score for each feature. Subsequently, these scores serve as heuristic knowledge to guide the GWO during the search process. To achieve this, the importance scores are normalized within the range of [0, 1], and a roulette wheel selection method is utilized to choose features for each grey wolf within the initial population generation procedure. The subsequent sections delve into the specifics of the PCC–GWO algorithm, encompassing both the PCC and GWO phases, facilitating a comprehensive understanding of the feature selection process.

Algorithm 1. Feature selection using PCC–GWO algorithm.**Input:**

Full heart disease dataset

Output:

Optimal Feature Subset for Machine Learning Model

Heuristic Feature Selection: Calculation of Importance Scores using PCC:

1. **For** ($i = 1$: Number of Features)
2. Calculation of the correlation of feature i with the class: CC_i
3. Calculation of the correlation of feature i in relation to the other features: CF_i
4. Calculation of the PCC importance score of feature i : $IS_i = CC_i/CF_i$
5. **End For**

Metaheuristic Feature Selection: Final Feature Subset Selection using GWO:

1. $t = 0$ (Initial Population)
2. **For** ($s = 1$: PopSize)
3. **for** ($i = 1$: Number of Features)
4. Calculation of the probability of feature i in solution s using Equation (3)
5. Deciding to select or decline feature i using roulette wheel selection
6. **end for**
7. Calculation of the fitness of each grey wolf s using Equation (4)
8. **End For**
9. Considering the best solution as alpha wolf: X_α
10. Considering the second best solution as beta wolf: X_β
11. Considering the third best solution as delta wolf: X_δ
12. **For** ($t = 1$: MaxIter)

% Population Updating

13. **for** each grey wolf s
14. Updating a , A_i , and C_i , r_{A_i} , and r_{C_i} .
15. Calculation of updating factor towards alpha grey wolf using Equation (13)
16. Calculation of updating factor towards beta grey wolf using Equation (14)
17. Calculation of updating factor towards delta grey wolf using Equation (15)
18. **if** ($|A_i| \geq 1$)
19. Updating the wolf s using search for prey by Equation (16)
20. **elseif** ($|A_i| < 1$)
21. Updating the wolf s using attacking prey by Equation (16)
22. **end if**
23. **end for**

% Fitness Evaluation

24. **for** ($s = 1$: PopSize)
25. Calculation of the fitness of each grey wolf s using Equation (4)
26. **End for**
27. Updating the best solution as alpha wolf: X_α
28. Updating the second best solution as beta wolf: X_β
29. Updating the third best solution as delta wolf: X_δ
30. **End For**

Return X_α as the optimized feature subset

4.1.1. Calculating the Importance Score of Features Using PCC

The PCC is a measure of the degree and direction of a relationship between two variables [38]. The PCC values vary from -1 to $+1$. A value of zero shows that there is no correlation between the two variables, while values near -1 or $+1$ suggest that there is a strong association between the two variables. The PCC is determined by:

$$r_{x,y} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}, \quad (1)$$

where \bar{x} and \bar{y} are the means of the two variables x and y , respectively. Moreover, x_i denotes the i -th value of the variable x , and y_i denotes the i -th value of the variable y .

By computing the correlation coefficient between each feature and the target variable, the method identifies the most informative features for an accurate classification. Then, by considering the correlation of each feature with respect to all the other features in the dataset, the method identifies redundant or highly correlated features that may not provide much additional information. The selection status of each feature is then determined based on a threshold value derived from its correlation coefficients. Finally, the GWO algorithm is used to repeat the selection process multiple times, and the feature subset with the highest fitness value is selected as the final solution. This method provides an effective way to identify and select the most valuable features in high-dimensional datasets, leading to improved predictive accuracy and better performance of machine learning models. The overall operation of PCC can be summarized as follows:

- (1) The correlation coefficient of each feature i with the class is computed as CC_i ;
- (2) The correlation coefficient of each feature i in relation to the other features is calculated as CF_i ;
- (3) The importance score of each feature i can be calculated as $IS_i = CC_i/CF_i$.

Concerning the PCC, if the value of IS_i is greater than a specific threshold TH ($IS_i > TH$), the feature i is selected; otherwise, it is not chosen. However, in the proposed combined PCC–GWO algorithm, the importance scores obtained by the PCC are used to guide the search process of the GWO for achieving a better level of convergence.

4.1.2. Feature Subset Selection Using GWO

The GWO was originally introduced by Mirjalili et al. [39]. It is based on the hunting behavior and social order of grey wolves found in nature. The social hierarchy of grey wolves is described by four types of wolves, which are the following:

- Alpha (α): the finest solution;
- Beta (β): the second best solution;
- Delta (δ): the third best solution;
- Omega (ω): the rest of the grey wolves.

Similar to other metaheuristic algorithms, the GWO initiates its search procedure by creating an initial population of viable solutions. Subsequently, it undergoes iterative phases, comprising a fitness assessment and population adaptation, until it fulfills a predefined stopping condition, such as reaching a specific number of iterations.

Representation of Feasible Solutions: The encoding of a feasible solution X (i.e., a grey wolf) is depicted in Figure 3. If the quantity of the i -th variable is equal to 1, the feature i is selected by the grey wolf; otherwise, it is not picked. Consequently, a value of 1 is used to represent the feature subset’s scope, which is expressed as follows:

$$X(i) = \begin{cases} 1 & \text{if feature } i \text{ is selected} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

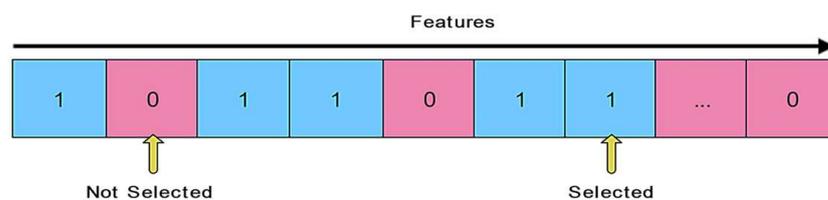


Figure 3. Representation of a feasible solution.

Initial Population Generation: As mentioned above, the original GWO algorithm starts its search process with a random population of grey wolves. However, in the proposed combined PCC–GWO algorithm, the importance scores of the features obtained by the PCC

are utilized to generate a set of near-optimal initial solutions for the GWO. To achieve this purpose, at first the normalized importance score for each feature i is calculated and, then, the probability of feature i to be selected in each solution (grey wolf) s can be expressed using the roulette wheel selection method, as follows:

$$NIS_i = \frac{IS_i - \min(ISs)}{\max(ISs) - \min(ISs)}. \quad (3)$$

Fitness Evaluation: The original dataset is separated into train and test datasets. The train dataset is considered for the optimization procedure via the GWO by means of K-fold cross-validation. However, the test dataset is unseen for the final evaluation of the generalizability of the trained model. The following is the fitness function of the GWO to assign the quality of each solution, which aims to be maximized:

$$\text{maximize Fitness} = \left(\mu \times \text{accuracy} + (1 - \mu) \times \frac{\text{Number of all features}}{\text{Number of selected features}} \right), \quad (4)$$

where *accuracy* is the total accuracy of the base learner using the validation dataset, and μ is a parameter ($0 < \mu < 1$) that determines the relative importance of accuracy and the number of selected features on the fitness value. The higher μ , the higher impact of *accuracy* on the fitness value. We consider $\mu = 0.99$ to ensure that high-accuracy solutions are achieved, while the number of features in the second rank is minimized.

Population Updating: At every iteration of the GWO, after the fitness evaluation of all the wolves, the first three best wolves, α , β , and δ , are in charge of leading the optimizer's hunting process, while ω simply obeys and follows them. Encircling, hunting, and attacking are the three well-organized steps that the GWO does during the optimization process. The following equations were used to determine the encircling process:

$$\vec{D} = \left| \vec{C} \cdot \vec{X}_p(t) - \vec{X}(t) \right|, \quad (5)$$

$$\vec{X}(t+1) = \vec{X}_p(t) + \vec{A} \cdot \vec{D}, \quad (6)$$

where t indicates the number of iterations, X represents the location vector of the wolf, and X_p represents the location vector of the prey. Moreover, A and C represent the vector coefficients expressed as follows:

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a}, \quad (7)$$

$$\vec{C} = 2 \cdot \vec{r}_2, \quad (8)$$

Where $[0, 1]$ is a random range for the vectors r_1 and r_2 , and the elements within the vector a start at 2 and fall linearly to 0 during the execution of the algorithm, as follows:

$$\vec{a} = 2 - t \cdot \frac{2}{MaxIter}, \quad (9)$$

where *MaxIter* denotes the maximum number of iterations.

The GWO keeps the top three solutions (α , β , and δ) obtained so far and compels ω to modify their placements in order to follow them. As a result, a series of equations that run for each search candidate is used to simulate the GWO hunting process. To achieve this, at first, the parameters of D for alpha, beta, and delta wolves are expressed as follows:

$$\vec{D}_\alpha = \left| \vec{C}_1 \cdot \vec{X}_\alpha - \vec{X} \right|, \quad (10)$$

$$\vec{D}_\beta = \left| \vec{C}_2 \cdot \vec{X}_\beta - \vec{X} \right|, \tag{11}$$

$$\vec{D}_\delta = \left| \vec{C}_3 \cdot \vec{X}_\delta - \vec{X} \right|, \tag{12}$$

Then, the moving vectors of the grey wolf X towards the alpha, beta, and delta wolves can be calculated as Equations (12)–(14), respectively. Finally, the movement of the grey wolf X is obtained through the aggregation of the three moving vectors according to Equation (15).

$$\vec{X}_1 = \vec{X}_\alpha - A_1 \cdot (\vec{D}_\alpha), \tag{13}$$

$$\vec{X}_2 = \vec{X}_\beta - A_2 \cdot (\vec{D}_\beta), \tag{14}$$

$$\vec{X}_3 = \vec{X}_\delta - A_3 \cdot (\vec{D}_\delta), \tag{15}$$

$$\vec{X}(t + 1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3}. \tag{16}$$

4.2. Ensemble Learning Model

Ensemble learning is a technique for improving the performance of a classifier. It is an efficient classification strategy that combines a weak classifier with a strong classifier to improve the effectiveness of the weak learner [40]. The proposed EHMFFL algorithm utilizes the ensemble technique to improve the accuracy of the SVM, KNNs, LR, RF, NB, DT, and XGBoost base learners for diagnosing heart disease. When compared to a single classification, the goal of integrating numerous learning models is to achieve better performance with more robustness. Figure 4 illustrates how ensemble learning is used to improve heart disease diagnosis using these seven base learners.

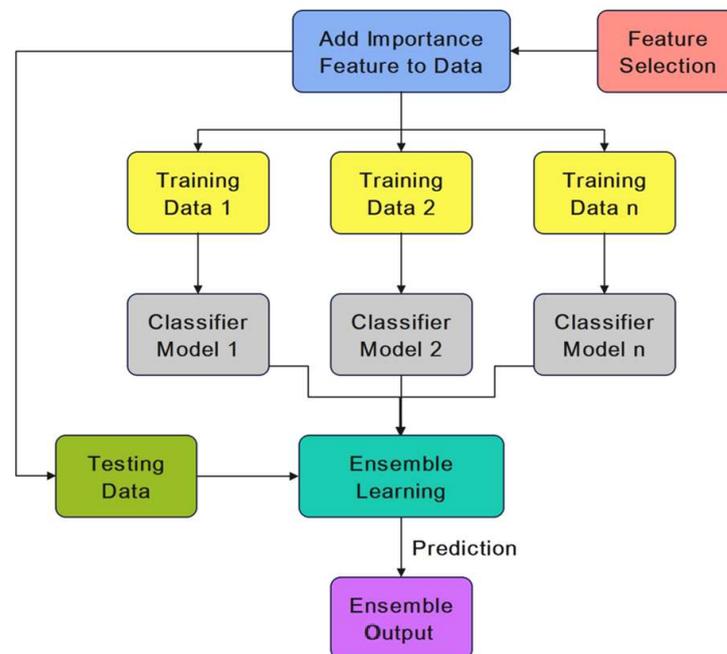


Figure 4. The proposed ensemble learning model.

Finally, using a weighted averaging method, we predict heart disease using each dataset. The weights of the different base learners are adjusted so that each learner with a higher accuracy has a higher weight in the ensemble learning model. The algorithm involves separately predicting each class and, then, using a weighted function to combine the outcomes. In contrast to hard voting with an equal chance for each base learner,

each prediction receives a weight, and the final results are combined by computing the weighted average. More specifically, the weight of base learner b is proportional to its normalized accuracy using the validation dataset against all the base learners within the ensemble model.

5. Evaluation and Findings

This section offers a comprehensive view of the performance metrics and results obtained in our study. All simulations were meticulously conducted on a PC, featuring an Intel i7 CPU with 2.6 GHz and 16 GB of RAM, and executed on MATLAB R2022b within the Windows 10 environment. Table 2 provides a snapshot of the parameter set applied to the GWO algorithm, facilitating a clearer understanding of the experimental setup. In the following, we evaluate the performance of the proposed EHMFFL algorithm against the seven base learners, as well as the state-of-the-art techniques.

Table 2. Parameter settings for the GWO.

Parameter	Value
Number of grey wolves (PopSize)	30
Number of iterations (MaxIter)	100
Search domain	{0, 1}
Solution dimension	No. Features

5.1. Performance Metrics

In this paper, each dataset was split into 80% and 20% to train and test the datasets. The train dataset (using K-fold cross-validation with $K = 10$) was applied to optimize the model, while the test dataset was used to assess the generalizability of the tuned model on new unseen data samples. Considering the true positive (TP), true negative (TN), false positive (FP), and false negative (FN), we utilized different performance measures to evaluate the performance of the different techniques:

- True positive (TP): the number of correctly identified positive instances inside the desired class;
- True negative (TN): the number of correctly identified negative instances outside the desired class;
- False positive (FP): the number of incorrectly predicted positive samples when the actual target was negative;
- False negative (FN): the number of incorrectly predicted negative samples when the actual target was positive.

Accuracy: Occurs when the proportion of occurrences correctly classified by the classification learner equals the proportion of correctly predicted samples to the total number of examples, which can be calculated as follow:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}. \quad (17)$$

Precision: It is one of the performance indicators that will be used to determine how many correct positive forecasts were made. So, precision measures the minority class's accuracy; then, the ratio of correctly predicted positive instances divided by the total number of positive cases predicted is utilized to compute it, using:

$$\text{Precision} = \frac{TP}{TP + FP}. \quad (18)$$

Recall: It is a measurement that quantifies the proportion of actual positive predictions correctly identified out of all potential positive predictions. Unlike precision, which considers the correctly predicted positives relative to all positive predictions, recall focuses on the

positives that were overlooked. Essentially, it signifies the extent to which the positive class is comprehensively captured, which is calculated as follows:

$$\text{Recall} = \frac{TP}{TP + FN}. \quad (19)$$

F1 score: In an ideal classifier, we aim for both accuracy and recall to be maximized, equating to values of one. This optimal scenario indicates that both the FP and FN are reduced to zero, highlighting the classifier's ability to make accurate and comprehensive predictions; ultimately, minimizing errors in both positive and negative classifications. As a result, we need a statistic that takes precision and recall into account. The F1 score is a precision and recall-based measure that is defined as follows:

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (20)$$

Specificity: It is the proportion of true negative samples to all actual negative samples, which indicates the ratio of the projected presence to the total samples with heart disease presence. The specificity is expressed as follows:

$$\text{Specificity} = \frac{TN}{TN + FP}. \quad (21)$$

5.2. Experimental Findings

As mentioned above, 80% of each dataset were used for the training of the proposed model, while the remaining 20% of the data samples were kept unseen for the validation of the tuned model. More specifically, 61 and 54 data samples were used to test the proposed model and compare it with the other techniques using the Cleveland and Statlog datasets, respectively. The obtained confusion matrix by the proposed EHMFFL algorithm using both datasets can be seen in Figure 5.

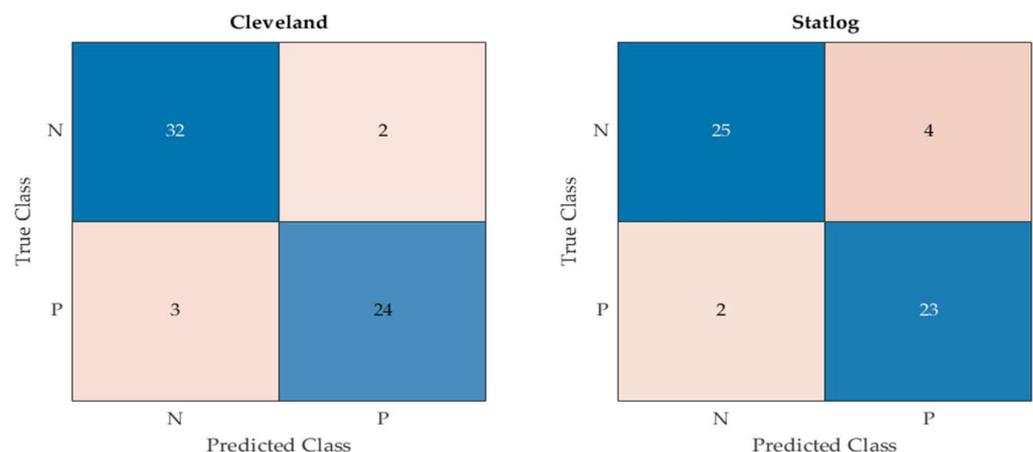


Figure 5. Confusion matrix of the proposed EHMFFL algorithm using the Cleveland and Statlog datasets.

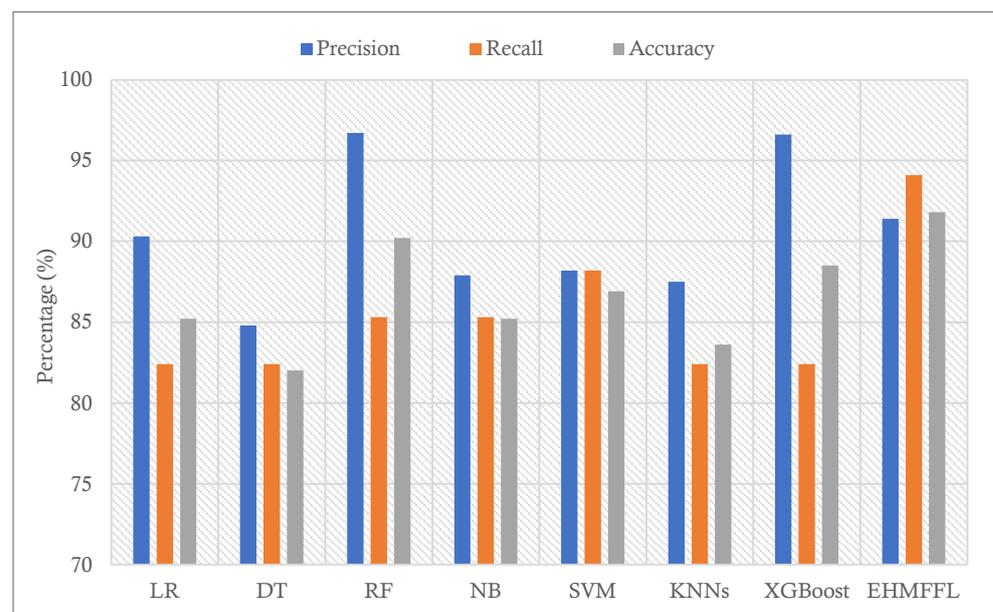
To find the effectiveness of the proposed ensemble EHMFFL algorithm against the base learners, a comparison of various performance measures using the test data samples from the Cleveland and Statlog datasets is provided in Tables 3 and 4, respectively. While some algorithms may display higher performance than the EHMFFL algorithm on a measure, the proposed method outperforms all the techniques on average for both datasets. Figures 5 and 6 show the accuracy of the EHMFFL algorithm using various methods. The EHMFFL algorithm surpasses all other methods, as illustrated in Figures 6 and 7.

Table 3. Comparison of the EHMFFL algorithm with existing methods using the Cleveland dataset.

Algorithms	Accuracy	Precision	Recall	Specificity	F1 Score
LR	85.2	90.3	82.4	88.9	86.2
DT	82	84.8	82.4	81.5	83.6
RF	90.2	96.7	85.3	96.3	90.6
NB	85.2	87.9	85.3	85.2	86.6
SVM	86.9	88.2	88.2	85.2	88.2
KNNs	83.6	87.5	82.4	85.2	84.8
XGBoost	88.5	96.6	82.4	96.3	88.9
EHMFFL (Proposed)	91.8	91.4	94.1	88.9	92.8

Table 4. Comparison of the EHMFFL algorithm with existing methods using the Statlog dataset.

Algorithms	Accuracy	Precision	Recall	Specificity	F1 Score
LR	79.6	80	82.8	76	81.4
DT	81.5	85.2	79.3	84	82.1
RF	84.4	86.2	85.5	76	87.4
NB	77.8	79.3	79.3	76	79.3
SVM	83.3	81.3	89.7	76	85.2
KNNs	80.8	84.5	78.9	83	81.6
XGBoost	85.2	88.9	82.8	88	85.7
EHMFFL (Proposed)	88.9	92.6	86.2	92	89.3

**Figure 6.** Comparison of the results of different methods using the Cleveland dataset in terms of precision, recall, and accuracy.

According to the different performance metrics for various classification techniques using the Cleveland dataset, the EHMFFL algorithm outperforms all the base learners with an accuracy of 91.8%, a precision of 91.4%, a recall of 94.1%, an F1 score of 92.8%, and a specificity of 88.9. This shows that the EHMFFL algorithm is the most effective and efficient method for the supplied dataset. Other algorithms also perform well in some cases, with accuracies ranging from 82% to 90.2%. When comparing the other algorithms, RF is the best base learner with an accuracy of 90.2% and, then, XGBoost and SVM with accuracies of 88.5% and 86.9% are in the next in order. Also, based on the results in Table 4, the EHMFFL algorithm exceeds all other algorithms with an accuracy score of 88.9%. After the EHMFFL algorithm, XGBoost, RF, and SVM, obtained better results than the other base learners with

accuracy scores of 85.2%, 84.4%, and 83.3%, indicating that these three methods are the most accurate base learners, the same as observed for the Cleveland dataset. The results show that the EHMFFL algorithm again shines out in terms of all the performance metrics, on average.

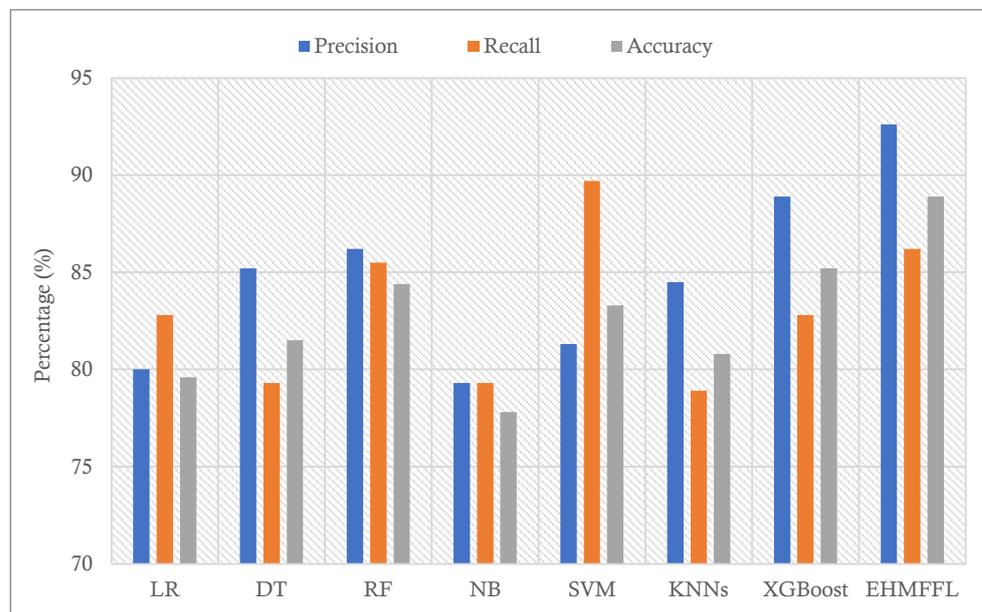


Figure 7. Comparison of the results of different methods using the Statlog dataset in terms of precision, recall, and accuracy.

5.2.1. Running Time Analysis

To assess the trade-off between enhanced performance and time-consuming cost in the proposed EHMFFL algorithm, we conducted a comprehensive analysis comparing its running time with the seven base learners and an ensemble model utilizing all the original features for each base learner, named “Ensemble” in Table 5. Our findings, detailed in Table 5, delineate the offline training/tuning phase and the online test phase time comparisons for both datasets. The offline time for each base learner signifies its individual training time, while for the Ensemble model, it represents the total training duration of the seven base learners within the model. Conversely, in our EHMFFL model, this duration encapsulates the time for feature selection via the PCC–GWO algorithm and training the base learners with the reduced feature subsets.

Table 5. Comparison of the running time (in seconds) of the different techniques.

Algorithms	Offline Training/Tuning Phase		Online Test Phase	
	Cleveland Dataset	Statlog Dataset	Cleveland Dataset	Statlog Dataset
LR	0.01	0.01	0.02	0.02
DT	0.13	0.12	0.03	0.03
RF	0.74	0.7	0.08	0.07
NB	0.12	0.14	0.04	0.04
SVM	0.21	0.23	0.06	0.07
KNNs	0.08	0.07	0.03	0.02
XGBoost	0.86	0.83	0.1	0.09
Ensemble	2.3	2.2	0.17	0.16
EHMFFL (Proposed)	142	137	0.13	0.13

Analysis of Running Time in Offline Training/Tuning Phase: Table 5 indicates that the Ensemble model, leveraging all the original features for each base learner, approximately accumulates the total time consumption of the training phase for individual base learners, given their sequential training approach. However, the huge increase in the running time of the EHMFFL algorithm, compared to the Ensemble model, primarily stems from the implementation of the time-intensive PCC–GWO feature selection algorithm for selecting an appropriate feature subset for each base learner. Notably, these processes are confined to the offline model’s tuning and do not affect the test phase time.

Analysis of Running Time in Online Test Phase: In the online test phase, both the Ensemble and EHMFFL models exhibit slightly higher time consumption than the single base learners. As the test phase operations occur in parallel, the cumulative time of the base learners is not entirely additive for these models. While the Ensemble model shows around a 70% increase compared to the most time-consuming base learner (i.e., XGBoost), the EHMFFL model, owing to the PCC–GWO feature selection, presents reduced response times for the base learners compared to the Ensemble model. As a result, the online time of the EHMFFL algorithm exhibits a 24% and 19% reduction when using the Cleveland and Statlog datasets, respectively.

Despite the significant increase in the offline running time of the EHMFFL model caused by the combined heuristic–metaheuristic PCC–GWO feature selection algorithm, the implementation of PCC–GWO yields a reduction in the online processing time when compared to the Ensemble model encompassing all the original features.

5.2.2. Analysis of the Correlation Heat Map (CHM)

This section presents the CHM, illustrating the relationships between different variables within the cardiovascular data for both the Cleveland and Statlog datasets. Figures 8 and 9 display these CHMs, where each column signifies a specific variable, and each row visualizes its correlations with other variables. The numerical values within the tables convey the strength and direction of these correlations, which can span from -1 , indicating a perfect inverse correlation, to 1 , representing a perfect positive correlation. This visual representation offers valuable insights into the interplay among the dataset variables and their potential impacts on heart disease prediction.

In Figure 8, the CHM of the Cleveland dataset illustrates the comparisons among various variables. These variables include age, gender, blood pressure (trestbps), cholesterol level (chol), fasting blood sugar (fbs), electrocardiogram results (restecg), maximum heart rate achieved (thalach), exercise-induced angina (exang), ST depression induced by exercise relative to rest (oldpeak), the number of major vessels colored by fluoroscopy (ca), type of chest pain (cp), and slope of the peak exercise ST segment (slope). Additionally, the presence of two types of thalassemia, indicated as thal and thal2, are considered. The values within the matrix fall within the -1 to 1 range, where positive values signify a positive correlation, negative values indicate a negative correlation, and a value of 0 denotes no discernible correlation between the variables. Analysis of the CHM for the Cleveland dataset reveals the following insights:

- The first section of the matrix compares age, sex, and blood pressure. The correlation between age and blood pressure is weakly positive (0.28), while the correlation between sex and blood pressure is weakly negative (-0.098);
- The second section compares cholesterol and blood sugar. Cholesterol and blood sugar have a weakly negative correlation (-0.057);
- The third section compares restecg, thalach, exang, and oldpeak. The resting electrocardiogram results (restecg) and exercise-induced angina (exang) have a weakly positive correlation (0.14), while the maximum heart rate achieved during exercise (thalach) has a weakly negative correlation (-0.044) with ST depression induced by exercise relative to rest (oldpeak);

- The fourth section compares the number of major vessels colored by fluoroscopy (ca) with the other variables. There is a weakly positive correlation between ca and age (0.12), and a weakly positive correlation between ca and cholesterol (0.097);
- The fifth section compares the different types of chest pain (cp) and their correlations with the other variables. Chest pain type 0 (cp_0) has a weakly positive correlation with ca (0.14), while chest pain type 1 (cp_1) has a weakly negative correlation with thal2 (−0.15). Chest pain type 2 (cp_2) has a weakly positive correlation with fbs (0.084), and chest pain type 3 (cp_3) has a weakly positive correlation with age (0.048);
- The final section of the matrix compares the slope of the peak exercise ST segment (slope) and the two types of thalassemia (thal and thal2). There is a weakly positive correlation between slope and thal2 (0.18), and a weakly negative correlation between slope and thal (−0.42).

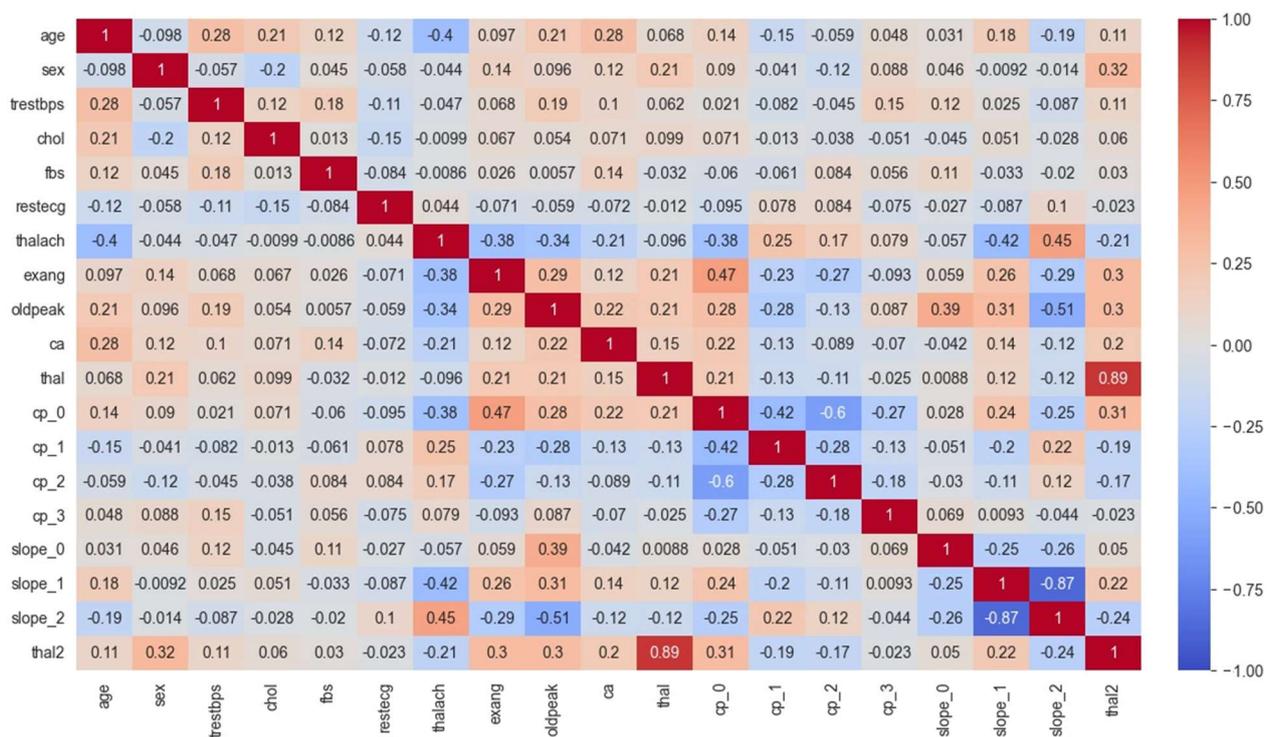


Figure 8. CHM analysis for the Cleveland dataset.

Also, according to the results of the CHM analysis for the Statlog dataset in Figure 9, it can be understood that:

- The values in the matrix represent the correlations between each pair of variables. A positive value indicates a positive correlation (as one variable increases, so does the other), while a negative value indicates a negative correlation (as one variable increases, the other decreases);
- For example, we can see that age is highly negatively correlated with itself (correlation coefficient of −1.00), since it is impossible for someone’s age to be negatively correlated with their own age. Sex is negatively correlated with BP and positively correlated with cholesterol levels. We can also see that the ST depression is positively correlated with exercise-induced angina, thallium stress test results, and chest pain types 3 and 4;
- Some notable correlations include a positive correlation between age and BP ($r = 0.27$), a negative correlation between age and max HR ($r = -0.4$), and a positive correlation between chest pain type 3 and ST depression ($r = 0.35$). There also appear to be some negative correlations between certain variables, such as sex and chest pain type 3 ($r = -0.26$) and slope of ST 3 and thal2 ($r = -0.24$).

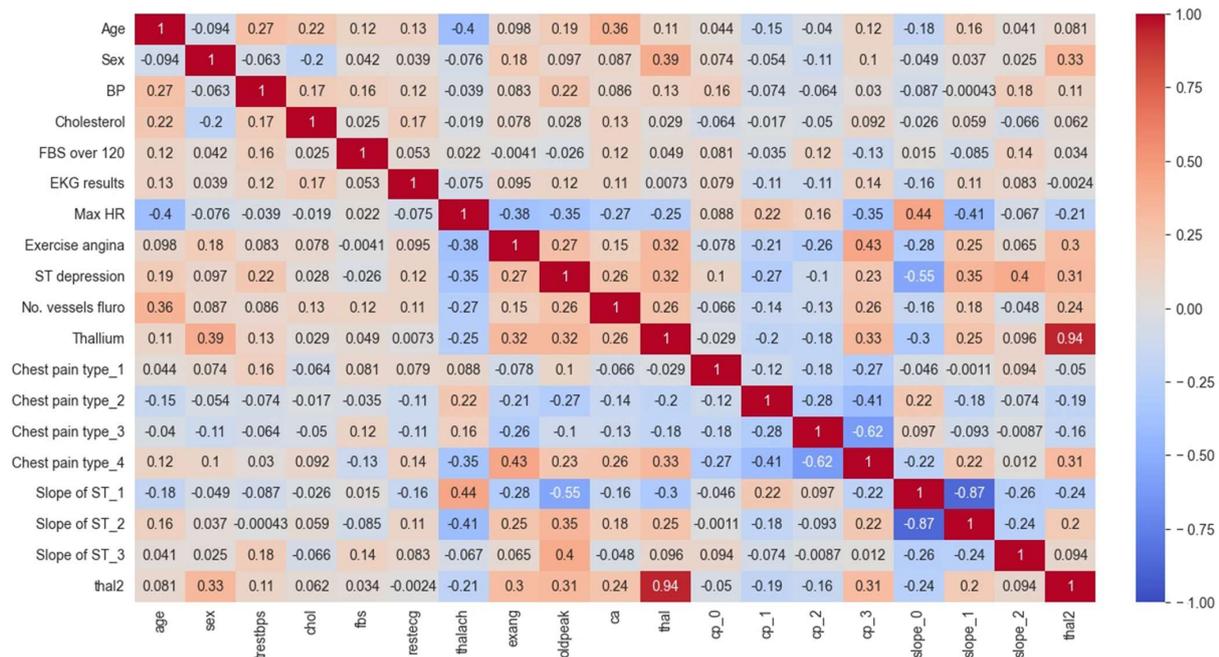


Figure 9. CHM analysis for the Statlog dataset.

5.2.3. Analysis of the Receiver Operating Characteristic (ROC)

In Figures 10 and 11, we present the ROC curves, which illustrate the performance of various heart disease prediction models, encompassing the seven base learners within our ensemble learning model, the EHMFFL algorithm as a whole, and random classification. These curves showcase the trade-off between sensitivity and specificity at different decision thresholds. To gauge the diagnostic value of these tests, we calculate the area under the ROC curve (AUC), where a larger AUC signifies a more effective test. According to the obtained results, the EHMFFL algorithm outperforms all the other compared techniques, achieving AUC values of 0.95 for the Cleveland dataset and 0.88 for the Statlog dataset, underscoring its superior predictive capability in heart disease diagnosis.

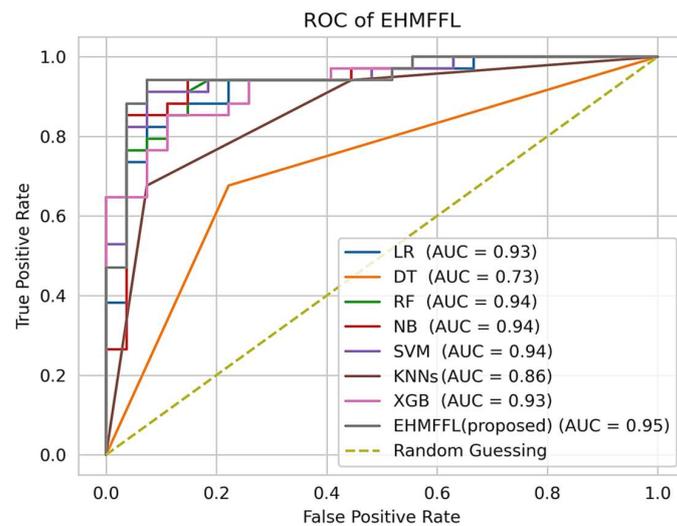


Figure 10. Analyzing the ROC curves of the different techniques using the Cleveland dataset.

5.3. Comparison with Existing Techniques

In this section, we conduct a comparative analysis of the EHMFFL algorithm against three existing heart disease diagnosis methods. Since the EHMFFL algorithm utilizes a

heuristic–metaheuristic-driven feature selection algorithm (i.e., PCC–GWO) embedded in an ensemble learning model, we considered the machine learning approach based on LR by Jindal et al. (2021) [18] (referred to as ML), an ensemble of different machine learning models, including DT, RF, SVM, KNNs, and ANN, developed by Shorewala (2021) [23] (referred to as EL), and a sequential heuristic feature selection method by Ahmad et al. (2022) [28] (referred to as FS). We assess the precision, recall, and accuracy of the different methods, as depicted in Figures 12 and 13, for the Cleveland and Statlog datasets, respectively. The results in Figure 12 indicate that the EL model outperforms the EHMFFL algorithm in terms of recall when using the Cleveland dataset. A similar trend is observed in Figure 13 for the Statlog dataset, where the FS method excels in achieving the highest recall among all the techniques. However, when considering overall performance and emphasizing accuracy as the main metric, the proposed EHMFFL model demonstrates its superiority over all the compared techniques for both datasets, underscoring its effectiveness in heart disease diagnosis.

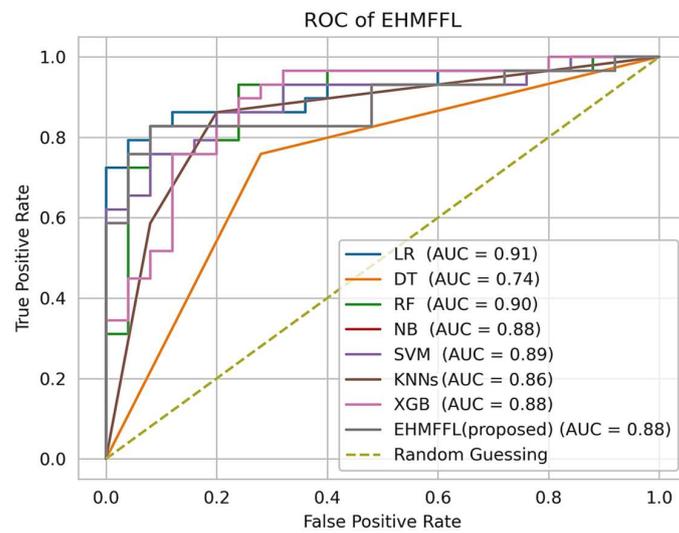


Figure 11. Analyzing the ROC curves of the different techniques using the Statlog dataset.

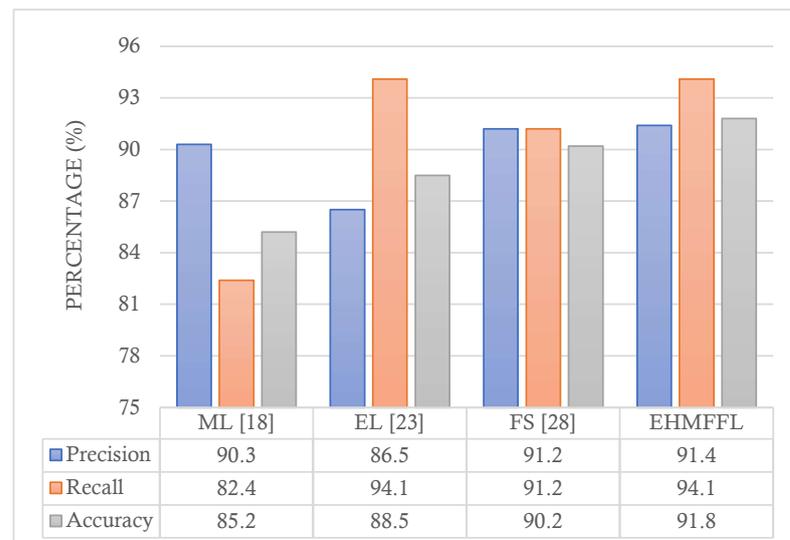


Figure 12. Comparison of the results of the EHMFFL algorithm with the existing techniques using the Cleveland dataset.

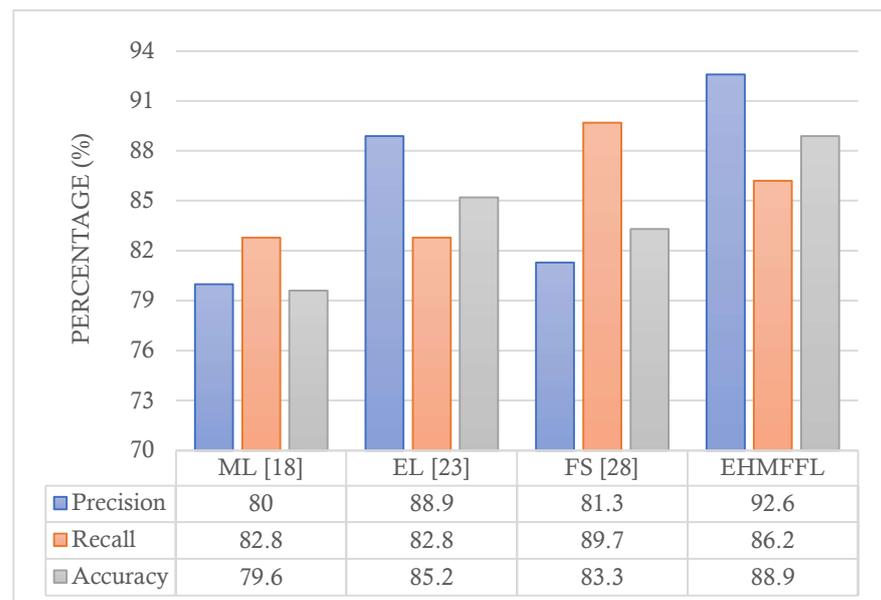


Figure 13. Comparison of the results of the EHMFFL algorithm with the existing techniques using the Statlog dataset.

6. Conclusions

In this study, we have introduced an ensemble heuristic–metaheuristic feature fusion learning (EHMFFL) algorithm, as a powerful tool for heart disease diagnosis using tabular data. The EHMFFL algorithm’s first phase employed a hybrid feature selection approach, combining heuristic-based PCC and metaheuristic-driven GWO techniques through the innovative PCC–GWO method. This approach effectively selects the essential features for each machine learning model, facilitating the construction of a robust predictive model through ensemble learning. We evaluated the performance of the EHMFFL algorithm using the Cleveland and Statlog datasets. With an accuracy rate of 91.8% for the Cleveland dataset and 88.9% for the Statlog dataset, our method outperformed the base learners and state-of-the-art approaches. These outcomes highlight the potential of our strategy to elevate cardiac disease prediction accuracy and support healthcare professionals in making more informed patient care decisions.

In implementing the EHMFFL algorithm, the choice of classical machine learning methods within the ensemble model was driven by their inherent benefits related to the utilized datasets in this paper. Classical models offer superior interpretability crucial for comprehending feature influences, especially conspicuous with straightforward attributes. Their efficiency in handling smaller datasets, computational simplicity, and lower risk of overfitting in scenarios with limited samples, outweigh the potential gains from deep learning techniques, particularly considering the simplicity of the provided features. However, if we encounter larger datasets with intricate patterns, the utilization of deep learning techniques becomes imperative. Looking ahead, expanding the application of the EHMFFL algorithm to larger-scale datasets with complex features and more extensive sample sizes stands as a key future endeavor. An appealing prospect is to use deep neural networks integrated with metaheuristic-driven ensemble learning techniques, employing deep networks as base learners within the ensemble model. This amalgamation holds the promise of creating resilient systems adept at deciphering intricate heart disease data streams. Additionally, the integration of real-time patient data streams and the development of a user-friendly interface remain pivotal avenues, promising transformative healthcare tools for timely disease detection and proactive intervention. Furthermore, we plan to explore further enhancements to the proposed algorithm by delving into more advanced feature selection methods and metaheuristic algorithms, continually striving to optimize diagnostic accuracy and efficiency in healthcare.

Author Contributions: Conceptualization, M.H. and E.M.; Data curation, M.H. and E.M.; Formal analysis, M.S. and F.W.; Investigation, M.H. and E.M.; Methodology, M.H., E.M. and M.S.; Resources, M.H. and E.M.; Software, M.H. and E.M.; Supervision, M.S. and F.W.; Validation, M.S. and F.W.; Writing—original draft, M.H. and E.M.; Writing—review and editing, M.S. and F.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data used in the study are available from the authors and can be shared upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Das, S.; Sharma, R.; Gourisaria, M.K.; Rautaray, S.S.; Pandey, M. Heart disease detection using core machine learning and deep learning techniques: A comparative study. *Int. J. Emerg. Technol.* **2020**, *11*, 531–538.
2. Hasan, T.T.; Jasim, M.H.; Hashim, I.A. FPGA design and hardware implementation of heart disease diagnosis system based on NVG-RAM classifier. In Proceedings of the 2018 3rd Scientific Conference of Electrical Engineering (SCEE), Baghdad, Iraq, 19–20 December 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 33–38.
3. Rahman, A.U.; Saeed, M.; Mohammed, M.A.; Jaber, M.M.; Garcia-Zapirain, B. A novel fuzzy parameterized fuzzy hypersoft set and riesz summability approach based decision support system for diagnosis of heart diseases. *Diagnostics* **2022**, *12*, 1546. [[CrossRef](#)] [[PubMed](#)]
4. Javid, I.; Alsaedi AK, Z.; Ghazali, R. Enhanced accuracy of heart disease prediction using machine learning and recurrent neural networks ensemble majority voting method. *Int. J. Adv. Comput. Sci. Appl.* **2020**, *11*. [[CrossRef](#)]
5. Muhsen, D.K.; Khairi TW, A.; Alhamza NI, A. Machine learning system using modified random forest algorithm. In Proceedings of the Intelligent Systems and Networks (ICISN 2021), Hanoi, Vietnam, 19 March 2021; Springer: Singapore; pp. 508–515.
6. Mastoi QU, A.; Wah, T.Y.; Mohammed, M.A.; Iqbal, U.; Kadry, S.; Majumdar, A.; Thinnukool, O. Novel DERMA fusion technique for ECG heartbeat classification. *Life* **2022**, *12*, 842. [[CrossRef](#)] [[PubMed](#)]
7. Nahar, J.; Imam, T.; Tickle, K.S.; Chen YP, P. Computational intelligence for heart disease diagnosis: A medical knowledge driven approach. *Expert Syst. Appl.* **2013**, *40*, 96–104. [[CrossRef](#)]
8. Lee, H.G.; Noh, K.Y.; Ryu, K.H. Mining biosignal data: Coronary artery disease diagnosis using linear and nonlinear features of HRV. In Proceedings of the Emerging Technologies in Knowledge Discovery and Data Mining: PAKDD 2007 International Workshops, Nanjing, China, 22–25 May 2007; Revised Selected Papers 11. Springer: Berlin/Heidelberg, Germany, 2007; pp. 218–228.
9. Sudhakar, K.; Manimekalai, D.M. Study of heart disease prediction using data mining. *Int. J. Adv. Res. Comput. Sci. Softw. Eng.* **2014**, *4*, 1157–1160.
10. Khazaee, A. Heart beat classification using particle swarm optimization. *Int. J. Intell. Syst. Appl.* **2013**, *5*, 25. [[CrossRef](#)]
11. Xing, Y.; Wang, J.; Zhao, Z. Combination data mining methods with new medical data to predicting outcome of coronary heart disease. In Proceedings of the 2007 International Conference on Convergence Information Technology (ICCIT 2007), Gwangju, Republic of Korea, 21–23 November 2007; IEEE: Piscataway, NJ, USA, 2007; pp. 868–872.
12. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [[CrossRef](#)]
13. Schapire, R.E.; Singer, Y. Improved boosting algorithms using confidence-rated predictions. In Proceedings of the 11th Annual Conference on Computational Learning Theory, Madison, WI, USA, 24–26 July 1998; pp. 80–91.
14. Miao, K.H.; Miao, J.H. Coronary heart disease diagnosis using deep neural networks. *Int. J. Adv. Comput. Sci. Appl.* **2018**, *9*, 1–8. [[CrossRef](#)]
15. Vijayashree, J.; Sultana, H.P. A machine learning framework for feature selection in heart disease classification using improved particle swarm optimization with support vector machine classifier. *Program. Comput. Softw.* **2018**, *44*, 388–397. [[CrossRef](#)]
16. Waigi, D.; Choudhary, D.S.; Fulzele, D.P.; Mishra, D. Predicting the risk of heart disease using advanced machine learning approach. *Eur. J. Mol. Clin. Med* **2020**, *7*, 1638–1645.
17. Tuli, S.; Basumatary, N.; Gill, S.S.; Kahani, M.; Arya, R.C.; Wander, G.S.; Buyya, R. HealthFog: An ensemble deep learning based Smart Healthcare System for Automatic Diagnosis of Heart Diseases in integrated IoT and fog computing environments. *Future Gener. Comput. Syst.* **2020**, *104*, 187–200. [[CrossRef](#)]
18. Jindal, H.; Agrawal, S.; Khera, R.; Jain, R.; Nagrath, P. Heart disease prediction using machine learning algorithms. *IOP Conf. Ser. Mater. Sci. Eng.* **2021**, *1022*, 012072. [[CrossRef](#)]
19. Sarra, R.R.; Dinar, A.M.; Mohammed, M.A.; Abdulkareem, K.H. Enhanced heart disease prediction based on machine learning and χ^2 statistical optimal feature selection model. *Designs* **2022**, *6*, 87. [[CrossRef](#)]
20. Aliyar Vellameeran, F.; Brindha, T. A new variant of deep belief network assisted with optimal feature selection for heart disease diagnosis using IoT wearable medical devices. *Comput. Methods Biomech. Biomed. Eng.* **2022**, *25*, 387–411. [[CrossRef](#)] [[PubMed](#)]
21. Latha CB, C.; Jeeva, S.C. Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques. *Inform. Med. Unlocked* **2019**, *16*, 100203. [[CrossRef](#)]

22. Ali, F.; El-Sappagh, S.; Islam, S.R.; Kwak, D.; Ali, A.; Imran, M.; Kwak, K.S. A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. *Inf. Fusion* **2020**, *63*, 208–222. [[CrossRef](#)]
23. Shorewala, V. Early detection of coronary heart disease using ensemble techniques. *Inform. Med. Unlocked* **2021**, *26*, 100655. [[CrossRef](#)]
24. Ghasemi Darehnaei, Z.; Shokouhifar, M.; Yazdanjouei, H.; Rastegar Fatemi SM, J. SI-EDTL: Swarm intelligence ensemble deep transfer learning for multiple vehicle detection in UAV images. *Concurr. Comput. Pract. Exp.* **2022**, *34*, e6726. [[CrossRef](#)]
25. Shokouhifar, A.; Shokouhifar, M.; Sabbaghian, M.; Soltanian-Zadeh, H. Swarm intelligence empowered three-stage ensemble deep learning for arm volume measurement in patients with lymphedema. *Biomed. Signal Process. Control.* **2023**, *85*, 105027. [[CrossRef](#)]
26. Nagarajan, S.M.; Muthukumaran, V.; Murugesan, R.; Joseph, R.B.; Meram, M.; Prathik, A. Innovative feature selection and classification model for heart disease prediction. *J. Reliab. Intell. Environ.* **2022**, *8*, 333–343. [[CrossRef](#)]
27. Al-Yarimi FA, M.; Munassar NM, A.; Bamashmos MH, M.; Ali MY, S. Feature optimization by discrete weights for heart disease prediction using supervised learning. *Soft Comput.* **2021**, *25*, 1821–1831. [[CrossRef](#)]
28. Ahmad, G.N.; Ullah, S.; Algethami, A.; Fatima, H.; Akhter SM, H. Comparative study of optimum medical diagnosis of human heart disease using machine learning technique with and without sequential feature selection. *IEEE Access* **2022**, *10*, 23808–23828. [[CrossRef](#)]
29. Pathan, M.S.; Nag, A.; Pathan, M.M.; Dev, S. Analyzing the impact of feature selection on the accuracy of heart disease prediction. *Healthc. Anal.* **2022**, *2*, 100060. [[CrossRef](#)]
30. Zhang, D.; Chen, Y.; Chen, Y.; Ye, S.; Cai, W.; Jiang, J.; Xu, Y.; Zheng, G.; Chen, M. Heart disease prediction based on the embedded feature selection method and deep neural network. *J. Healthc. Eng.* **2021**, *2021*, 6260022. [[CrossRef](#)] [[PubMed](#)]
31. Heart Disease. UCI Machine Learning Repository. Available online: <https://doi.org/10.24432/C52P4X> (accessed on 1 August 1989).
32. Statlog (Heart). UCI Machine Learning Repository. Available online: <https://doi.org/10.24432/C57303> (accessed on 13 February 1993).
33. Jensen, R. Combining Rough and Fuzzy Sets for Feature Selection. Ph.D. Thesis, University of Edinburgh, Edinburgh, UK, 2005.
34. Seyyedabbasi, A. Binary Sand Cat Swarm Optimization Algorithm for Wrapper Feature Selection on Biological Data. *Biomimetics* **2023**, *8*, 310. [[CrossRef](#)] [[PubMed](#)]
35. Shokouhifar, M.; Sohrabi, M.; Rabbani, M.; Molana SM, H.; Werner, F. Sustainable Phosphorus Fertilizer Supply Chain Management to Improve Crop Yield and P Use Efficiency Using an Ensemble Heuristic–Metaheuristic Optimization Algorithm. *Agronomy* **2023**, *13*, 565. [[CrossRef](#)]
36. Sohrabi, M.; Zandieh, M.; Shokouhifar, M. Sustainable inventory management in blood banks considering health equity using a combined metaheuristic-based robust fuzzy stochastic programming. *Socio-Econ. Plan. Sci.* **2023**, *86*, 101462. [[CrossRef](#)]
37. Xie, W.; Li, W.; Zhang, S.; Wang, L.; Yang, J.; Zhao, D. A novel biomarker selection method combining graph neural network and gene relationships applied to microarray data. *BMC Bioinform.* **2022**, *23*, 303. [[CrossRef](#)]
38. Pearson, K. Contributions to the mathematical theory of evolution. *Philos. Trans. R. Soc. Lond. A* **1894**, *185*, 71–110.
39. Mirjalili, S.; Mirjalili, S.M.; Lewis, A. Grey wolf optimizer. *Adv. Eng. Softw.* **2014**, *69*, 46–61. [[CrossRef](#)]
40. Grover, P.; Chaturvedi, K.; Zi, X.; Saxena, A.; Prakash, S.; Jan, T.; Prasad, M. Ensemble Transfer Learning for Distinguishing Cognitively Normal and Mild Cognitive Impairment Patients Using MRI. *Algorithms* **2023**, *16*, 377. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.