ORIGINAL ARTICLE

A hybrid method for heart disease diagnosis utilizing feature selection based ensemble classifier model generation

Jafar Abdollahi[1] · Babak Nouri-Moghaddam[1]

Abstract

Heart disease is one of the most complicated diseases, and it affects a large number of individuals throughout the world. In healthcare, particularly cardiology, early and accurate detection of cardiac disease is critical. The Heart Disease Data Set-UCI repository collects data on heart disease. The search space and complexity of the classification models are increased by this raw dataset, which contains redundant and inconsistent data. We need to eliminate the redundant and unnecessary elements from the data to improve classification accuracy. As a consequence, feature selection approaches might be useful for reducing the cost of diagnosis by identifying the most important qualities. This research developed an ensemble classification model based on a feature selection approach in which selected features play a role in classification. Accordingly, a classification approach was introduced using ensemble learning with a genetic algorithm, feature selection, and biomedical test values to diagnose heart disease. Based on the results, it is deduced that the benefits of using the feature selection method vary depending on the utilized machine learning technique. However, the best-proposed model based on the combination of genetic algorithm and the ensemble learning model has achieved an accuracy of 97.57% on the considered datasets. The suggested diagnosis system achieved better accuracy than previously proposed methods and can easily be implemented in healthcare to identify heart disease.

Graphical abstract
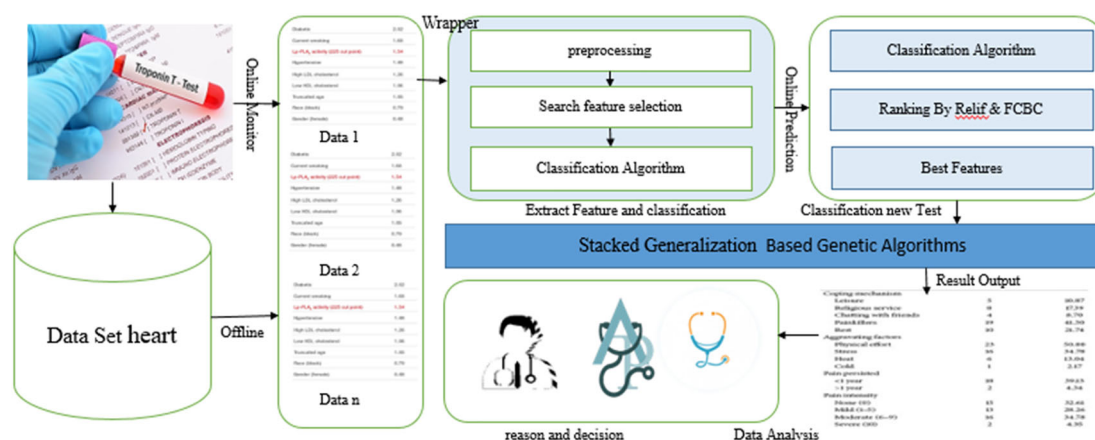
Keywords Feature selection · Hybrid · Stacked-generalization · Genetic algorithm · Heart disease

✉ Babak Nouri-Moghaddam
babaknouriit85@gmail.com

[1] Department of Computer Engineering, Ardabil Branch, Islamic Azad University, Ardabil, Iran

1 Introduction

Many hospitals and health care facilities have sprung up as a result of increased healthcare awareness and technological advancements. However, providing high-quality health care

at a reasonable cost remains a challenge [1]. Chronic disease is one of the most serious public health issues in the world [2], accounting for more than half of all fatalities globally. It also has the greatest non-infectious illness death rate and a high cost of prevention and treatment [3]. Heart disorders also largely afflict those aged 65 and over, and they have surpassed infectious diseases as the leading cause of mortality globally [4]. The considerable rise, complications, and high costs of these illnesses have a negative impact on society and place major financial and physical burdens on the worldwide community. As a result, employing appropriate preventative measures is essential.

Numerous factors are involved in diagnosing heart disease, which complicates a physician's task. To help physicians make quick decisions and minimize errors in diagnosis. Classification systems enable physicians to rapidly examine medical data in considerable detail [1]. These systems are implemented by developing a model that can classify existing records using sample data. Various classification algorithms have been developed and used as classifiers to assist doctors in diagnosing heart disease patients [5].

So, it is vital to identify the root causes responsible for heart diseases. So their remedy could also be planned by using suitable methods. Therefore, it is urgently required to identify factors responsible for heart diseases and develop an effective system for heart disease diagnosis [6]. Traditional methods are ineffective in diagnosing such a disease, and it is necessary to establish a medical diagnostic system based on feature selection approaches to predict and analyze the disease [7].

Feature selection or extraction process is an essential part of pattern recognition and machine learning (ML). The computation cost decreases thanks to the feature selection methods, and the classification performance can also increase. A suitable representation of data from all features is an important problem in machine learning and data mining. Not all original features are always useful for classification or regression tasks. Some features are irrelevant, redundant, or just noise within the distribution of the dataset. This feature can decrease the classification performance. To both increase the classification performance and reduce the computation cost of the classifier, the feature selection process should be used in classification or regression problems [8].

This article proposes an efficient and accurate system to diagnose heart disease, and the system is based on machine learning techniques. The system was developed based on classification algorithms. The feature selection algorithms select the most prominent features to increase the classification accuracy and reduce the execution time of the classification system. Furthermore, cross-validation is a resampling procedure used to evaluate machine learning models and hyper-parameter tuning. The performance measuring metrics are used to assess the performance of the classifiers. The performance of the classifiers has been checked on the selected features as selected by the evaluation metric.

The data set of heart patients available at UCI was used in this study, which included 13 useful features and 270 records. These features and abbreviations are listed in Table 3. Various machine learning algorithms based on genetic algorithms have been used as classifiers for heart prediction. The performance of all algorithms is evaluated on numerous metrics like precision, accuracy, sensitivity, specificity, recall, and F1 score. The experimental results show that the proposed feature selection algorithm (ensemble learning) is feasible with a genetic algorithm for designing a high-level intelligent system to identify heart disease. Among the objectives of this paper are to measure the following:

- Significantly increased forecast accuracy
- Achieve a high level of classification reliability
- Improved accuracy and reduced error compared to single-core models
- Machine learning models for heart disease prediction demonstrated high performance.

The article has been organized as follows: Sect. 2 explains related work, Sect. 3 proposes the method ultimately, Sect. 4 discusses experimental results with evaluation and performance comparison. The last section presents the paper's conclusion and will be introduced in future works.

2 Related work

Due to advances in various measurement techniques, it is likely to have medical data that contains relevant and irrelevant, and redundant features. Irrelevant features have an adverse effect on the description of the target class. The redundant features do not contribute anything but noise towards the description of the target class. Due to the noise contributed by redundant features, target class identification becomes a non-trivial task. Extracting valuable information from these datasets requires an exhaustive search of the sample space [9].

The heart is responsible for blood circulation throughout the body and acts as the body's engine so that heart disease can be fatal. The World Health Organization considers upset one of the foremost vital causes of death worldwide [10]. According to the surveys, 56 million people died in 2012, and the most important cause of the mortality rate was heart disease, controlled by early detection [11].

A large number of studies are being carried out to find efficient methods of medical diagnosis for various diseases. This study uses classification to predict diagnosis efficiently with fewer factors (i.e., attributes) that contribute more to cardiac

disease. Chen et al. developed the breast cancer diagnosis model using the support vector machine (SVM) and a rough set-based feature selection approach [12]. Wang et al. [13] used linear kernel SVM classifiers for heart disease detection and obtained an accuracy of 83.37%. A hybrid neural network method was proposed in [14], and the reported accuracy was 86.8%. In [15], the separability split value, k-nearest neighbor (KNN), and feature space mapping algorithms were used for heart disease detection, and KNN obtained the highest classification accuracy (85.6%).

In [16], a method has been presented to diagnose heart disease using particle swarm optimization and neural network feedforward back-propagation. In [17], a decision tree is used for data mining in heart disease. Researchers attempted to use data mining methods to diagnose heart diseases [18]. Different classification methods, such as neural networks and decision trees, are utilized to predict heart disease and identify its most important factors. The authors tried to diminish the difficulties by using combined methods in this study.

From these effects, it can be seen that feature selection methods can effectively increase the performance of individual classification algorithms in the diagnosis of heart disease. Noisy features and dependency relationships in the heart disease dataset can influence the diagnosis process. Typically, there are numerous records of accompanied syndromes in the original datasets and a large number of redundant symptoms. Consequently, it is necessary to reduce the dimensions of the original feature set by a feature selection method that can remove irrelevant and redundant features.

Over the past few years, many studies have been devoted to evaluating the classification prediction accuracy of the various clustering and classification algorithms applied to heart disease [19], available in the UCI repository. Due to the need to achieve effective analytical techniques for predicting chronic heart disease, many efforts have been made to improve the quality of evidence-based decisions and recommendations in the information environment. One of the most vital functions in health systems is correct medical recommendation supported by predicting the chance of short-run unwellness. It is noteworthy that there is a collection of disease risk prediction models within the medical literature [20].

Researchers have endeavored to search out the most accurate method of ML to explore the relationships in heart disease. Given the need, this article aims to create an intelligent system for predicting and correcting heart disease diagnoses, preventing any unwanted errors, lowering medical costs, and improving treatment quality [21].

Therefore, in the reference [22], the researchers have presented statistical methods for understanding the three medical data sets to produce prediction models by extracting appropriate rules to support the diagnosis process. In this study, methods like decision trees, Naive Bayes (NB), SVM, and a priori algorithms have yielded acceptable results. In

reference [23], a fuzzy system supported by an algorithmic genetic program has been tried to predict the chance of heart disease; the proposed fuzzy decision support system (FDSS) method has a high performance in predicting heart disease. In Reference [24], a call network has been accustomed to diagnosing hearts in clinical settings.

The accuracy of the artificial neural network (ANN) approach, classification and regression tree (CART) algorithm, neural network, and logistic regression has reached 97%, 87.6%, 95.6%, and 72%, respectively. In reference [25], the accuracy of an automated method for early detection of class changes in patients with heart failure using classification algorithms on a data set of 297 patients with evaluation validation approaches reached 97.87 and 67% for the two-, three-, and four-class classification problems, respectively. Numerous researchers have used this dataset to investigate various classification problems with different classification algorithms. Detrano [26] used a logistic regression algorithm and obtained 77.0% classification accuracy.

In 1989, Detrano [26] used an LR algorithm and got 77.0% classification precision. In addition, Edmonds [27] worked on the Cleveland dataset to examine world organic process approaches and ascertained some prognostic performance enhancements once they employed a novel method. However, the performance of his suggested method depends on the features extracted by the algorithm.

In 2010, Gudadhe et al. [28] accomplished an associate's degree field base with a multilayer perception network and support vector machine algorithms. Here, the proposed design obtained an associate's degree precision of 80.41% in classifying the two categories (with or without disease). On the other hand, Doppala et al. [29] achieved an accuracy of 85.40%, employing a combined genetic algorithm (GA) with radial basis function (GA-RBF). Also, experiments performed by the other authors show that the Naive Bayes model has the most effective achievement in terms of accurate prediction (86.112%). The challenger NN model with 86.12% correct predictions and, therefore, the third DT with 84% of the score was correct predictions [30].

Gupta et al. [31] used the Massachusetts Hospital Arrhythmia Database. The performance of the proposed method is compared with previous studies based on the sensitivity (SE) and detection rate (D.R). The proposed plan for the Massachusetts Institute of Technology-Beth Israel Hospital Arrhythmia database (MB Ar DB) Real-time database and RT DB is SE with 99.90%, D.R with 99.81%, and SE with 77.99%, D.R with 99.87%, respectively [31].

In 2020, Verma et al. [32] proposed a newly proposed method of hybrid feature selection technique for evaluating the performance of base learners, and we find that the reduced data subset performed is higher than the whole data set. Also, Osubor et al. [33] used an adaptive fuzzy neural inference system to predict postpartum depression. Thirty-

six data samples were used in model training. This system had a training error of 7.0706e − 005 in period one and an average test error of 3.0185 [33].

2.1 Feature selection

Selecting the appropriate features to achieve the best result in the data classification problem has been one of the most challenging topics in recent decades. Although using more features increases prediction accuracy from the learning theory perspective, practical evidence indicates this is not always true because not all features are essential for detecting the data class label. Some of them are irrelevant to the data label. Feature selection strategies may be divided into three categories: filtering, wrapper, and embedded [34–36].

2.1.1 Filtering methods

Filtering methods measure the accuracy of predictions or classifications based on an indirect criterion, such as the distance criterion, which indicates how well the classes are separated. This method is typically used as a preprocessing step. Instead, the features are selected based on their scores on various statistical tests to relate them to the outcome variable [37–44, 76].

2.1.2 Wrapper methods

Wrapper methods use a search method with a learning model to evaluate a subset of genes in the search phase. Thanks to using a learning model, wrapper methods usually offer better classification performance than filter methods. In contrast, they have several disadvantages, such as high computational overhead and the possibility of overfitting [50, 76].

2.1.3 Embedded methods

These methods perform feature selection in the learning process and are usually assigned to learners. This model also takes advantage of the previous models by using different evaluation criteria in different search stages.

Embedded methods combine filter qualities and wrapper methods. Algorithms do this with internal feature selection methods [40–47]. The various techniques used in this article are described in Table 1.

2.1.4 Differences between filter and wrapper methods

The most important differences between wrapper and filtering methods for selecting features are:

Table 1 Different techniques of the methods studied

Filter methods	Wrapper methods	Embedded methods
Information Gain	Recursive feature elimination:	L1 regularisation (LASSO)
Chi-square test	Sequential feature selection algorithms	Decision tree
Fisher score	Genetic algorithms	
Correlation coefficient		
Variance threshold		

The main differences between filtering and wrapping methods for selecting features:
-Filter methods are much faster than packaging methods because they do not include model training. On the other hand, packaging methods are also very expensive computationally
-Filter methods use statistical methods to evaluate a subset of features, while packaging methods use cross-validation
-Filtering methods may not find the best feature subset in many cases, but packaging methods can always provide the best feature subset
Using the subset of features of wrapping methods makes the model more prone to wear compared to using the subset of features of filter methods

- Filter methods are much faster than wrapper methods because they do not include model training. On the other hand, wrapper methods are also very computationally expensive.
- Filtering methods use statistical methods to evaluate a subset of features, while wrapper methods use cross-validation.
- In many cases, filtering methods may not find the best feature subset, but wrapper methods can always provide the best feature subset.
- In many cases, filtering methods may not find the best feature subset, but wrapper methods can always provide the best feature subset.
- Using a set of wrapper method features makes the model more susceptible to using a subset of filter method features [37–44].

2.2 Classification models

Machine learning algorithms such as Random forest (RF), Gaussian Naïve Bayes (GNB), Decision Tree (DT), Support-Vector Machines (SVM), gradient boosting (GB), K-nearest neighbors (KNN), and logistic regression (LR) were used for classification of people with DM based on data.

The performance of algorithms was evaluated with precision, recall, sensitivity, specificity, f1 score, and accuracy. To train and evaluate training datasets, seven classification algorithms such as RF, GNB, DT, SVM, GB, KNN, and LR were applied. All programming has been done in Python. 3.7 in the Jupyter Notebook [45–48].

Logistic regression (LR) is a machine learning technique for regression and classification issues that assign observations to a distinct set of categories. Provision regression is one of the foremost standard classification algorithms for divided (binary) classification [45–48].

Gaussian Naive classifier (GNB) is a cluster of simple classifiers, which supports possibilities created assumptive the independence of random variables and supported Bayes theorem [45–48].

Decision tree (DT) is a map of the potential results of a series of connected selections or choices. It permits a private person or organization to weigh potential actions regarding prices, opportunities, and edges [45–48].

SVM is classed as a pattern recognition algorithmic rule. The SVM algorithmic rule may be used where there is a necessity to spot patterns or classify objects into specific categories [45–48].

Gradient boosting (GB) classifiers are a bunch of machine learning algorithms that mix several weak learning models to make a robust prophetic model [45–48].

Random Forest (RF) may be a combined learning methodology for regression classification that works on the coaching time and sophistication output (classification) or for predictions of every tree on an individual basis, supported by a structure consisting of an outsized variety of call trees [45–48].

K-nearest neighbors (KNN): the K-nearest neighbors (KNN) algorithmic rule is arguably the only machine learning algorithmic rule. Building the model consists solely of storing the coaching information set. To form a replacement data point prediction, the algorithmic rule finds the nearest information points at intervals within the coaching information set—its "nearest neighbors" [45–48].

Stack generalization (SG) could be a different technique for combining several different classifiers like decision trees, artificial neural networks, support vector machines, etc. (see Fig. 1), which consists of two stages:

- Basic learners at level zero and stacking model learners at level one;
- At level zero, various models are used to learn from the dataset, and the output of each model is used to create a new dataset [45–48].

As a result, due to the speed of data collection, the issue of feature selection has become one of the most important issues.

2.3 Genetic algorithm

GA is one of the primary population-based stochastic algorithms proposed in history. The most common GA operators are selection, crossover, and mutation [56]. These algorithms encode a possible solution to a selected problem in a straightforward chromosome-like arrangement and apply recombination operators to those structures to preserve critical information. Although GAs are often considered performance optimizers, the wide range of genetic algorithms used in them is quite wide [57]. In Fig. 2, the working principle of a simple genetic algorithm is shown.

Since single meta-heuristic algorithms are not enough to solve all the problems, this study proposes a hybrid feature selection method for selecting a combination feature (filter, wrapper) based on the Ensemble approach. Accordingly, the aim is to highlight the identification of the algorithms mentioned in Sect. 2.2 and use them in combination with genetic algorithms to choose the most appropriate prediction method.

To that end, introduce a unique conception, effective correlation, and purpose to a quick filter method that will determine suitable options, likewise as redundancy between appropriate options while not performing pairwise correlation analysis. Among the objectives of this paper are to measure the following:

- Bringing together the best voting machine learning architectures
- Diagnosis of heart disease using a hybrid machine learning model
- Significantly increased forecast accuracy
- Combine several models.
- Achieve a high level of classification reliability.
- Greater accuracy and lower error when compared to single-core models.

Several significant limitations in this study could be addressed in future research. First, the study focused on the prediction heart. And things like that.

- Sample size
- Lack of available and/or reliable data
- Lack of access to hospital data

Therefore, the main contribution of this article is as follows:

- Machine learning models for heart disease prediction demonstrated high performance.
- Comparison of the proposed method results with the most relevant research conducted according to the previous literature.
- We investigate the advantages of ensemble learning methods (proposed model) for diagnosis and prediction.
- Using a hybrid Stacked-genetic approach in the diagnosis of heart disease.

Algorithms 1: Pseudo Code of Ensemble Algorithm

1:    **Input**: Data set $\mathbf{D} = \{(x_1, y_1), (x_2, y_2), \cdots, (x_m, y_m)\}$;
2:         First-level learning algorithms $\mathbf{L_1}, \ldots \mathbf{L_T}$;
3:         Second-level learning algorithm L.
4:    **Process**:
5:    *for $t = 1, \cdots, T$*:
6:       $h_t = L_t(D)$ // Train a first-level individual learner h_t by applying the first-level
7:    *end for*   // learning algorithm £ t to the original data set D
8:    $D' = 0$ // Generate a new data set
9:    *for $i = 1, \cdots, m$*:
10:      *for $t = 1, \cdots, T$*:
11:         $z_{it} = h_t(x_i)$ // Use h_t to classify the training example x_i
12:      *end for*
13:    $D' = D' \cup \{ ((z_{i1}, z_{i2}, \cdots, z_{it}), y_i) \}$
14:    *end for*
15:    $h' = L(D')$ // Train a second-level learner h' by applying the second-level
                       //learning algorithm £ to the new data set D'
16:    **Output**: $H(x) = h'(h_1(x), \ldots, h_T(x))$

Fig. 1 Ensemble algorithm

Algorithms 2: Algorithm GA

1   formulate initial population
2   randomly initialize population
3   repeat
       3.1 evaluate objective function
       3.2 find fitness function
       3.3 apply genetic operators
             3.3.1 reproduction
             3.3.2 crossover
             3.3.3 mutation
4   until stopping criteria

---

3 Methodology

We now describe the datasets we chose, the algorithms used, and the experimental methodology. They have been used for different research purposes in recent years. Such datasets may contain thousands of instances (records), each represented by hundreds or thousands of features (attributes or variables). The large datasets have many features, including valuable data for understanding information and many inappropriate and related features. This decreases learning achievement and computational performance. So, a preprocessing step named "feature selection" is used to reduce the dimensions before using any information extraction techniques such as classification, related rules, clustering, and regression [49]. A feature selection technique was proposed to pick a subset of relevant and non-redundant features in this study to overcome these problems. To this end, the genetic algorithm and the Relief and FCBF are employed to pick the valuable features in the rapid diagnosis of heart diseases. Our work attempts to predict diagnosis efficiently with a reduced num-
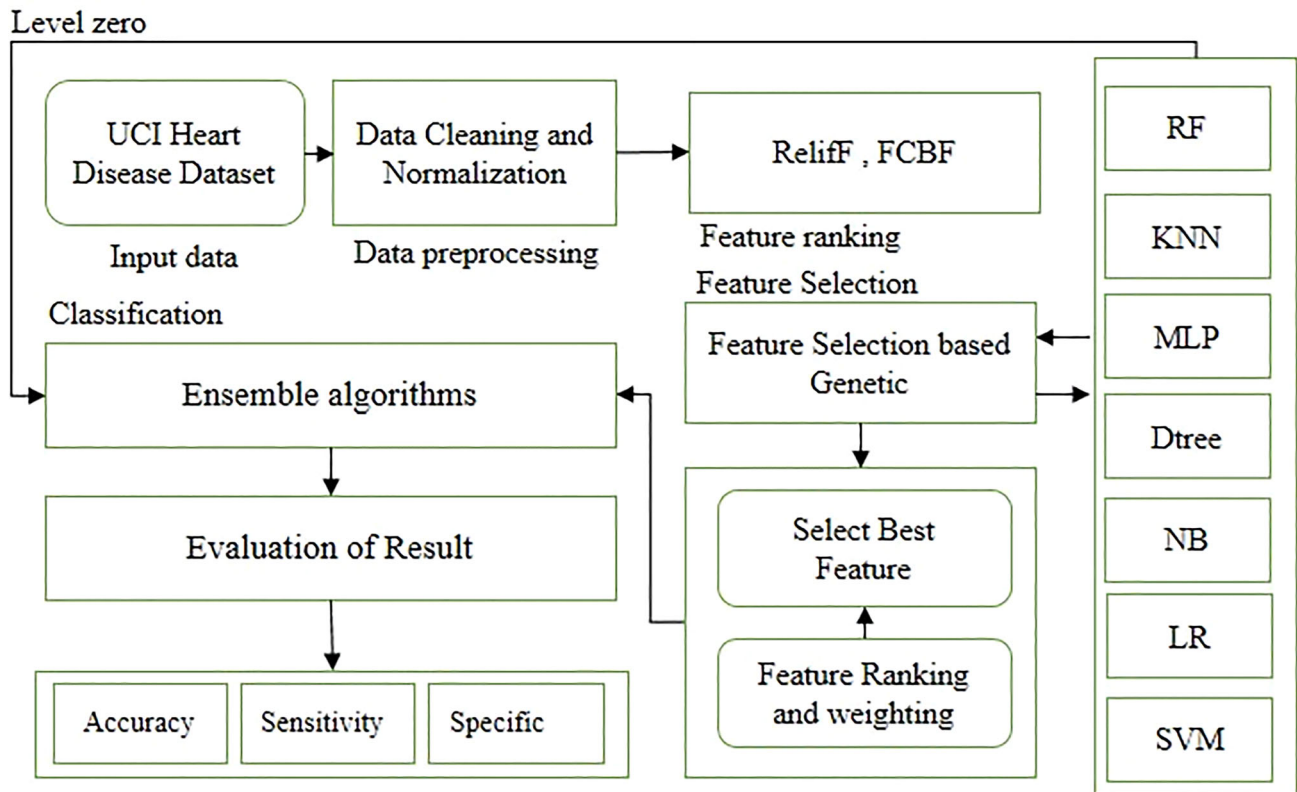
Fig. 3 The structure of feature selection

ber of features (i.e., attributes) that contribute more towards cardiac disease detection using the feature selection technique. The structure is shown in Fig. 3.

This research uses the feature selection-based ensemble learning approach, a machine learning method, to diagnose heart diseases. The structure of the classifier with machine learning-based genetic algorithms is shown in Fig. 4.

Considering that using an algorithm alone to diagnose and predict the disease has not been effective and has not been successful in many scenarios, the ensemble learning algorithm is used to accurately classify heart disease using selected features—the purpose of this.

The task is to improve the accuracy and speed of diagnosis of chronic diseases in the context of the intelligent network by which we want to use ensemble learning approaches and a new meta-learner in stacking learning. The structure of the proposed method is shown in Fig. 5.

The proposed method combines statistical analysis methods, machine learning algorithms, and genetic algorithms, which have the advantages of filtering and wrapper methods to select an optimal subset from the total features space. Also, the genetic algorithm along with Relief and FCBF is used to select the valuable features in the rapid diagnosis of heart disease. The order of execution of the work is shown in Fig. 6.

Initially, filtering algorithms are used to rank data set features. After ranking the features by filter algorithms, we used the subscription criterion ($\cap$) to select the selected features. These features can input the genetic algorithm to choose valuable features from the rated features. Also, the threshold criterion for ranking the features in these methods was 30%. Next, the genetic algorithm selects the best valuable feature compared to the less critical features (k best feature set by the genetic algorithm) for accurate heart disease prediction using group learning.

3.1 Proposed method

Our goal is to reduce and eliminate low-value features with the filter algorithm and select high-value features with the genetic algorithm for the accurate prediction of heart disease using ensemble learning.

In the category of evolutionary algorithms (EA), which generate solutions to optimization problems using methods inspired by natural evolution, including inheritance, mutation, selection, and crossover, genetic algorithms (GA) stand out for the purpose. It is one of the most effective methods to resolve problems that are little known. GA could be a general algorithm, and it works well in any search space. Within
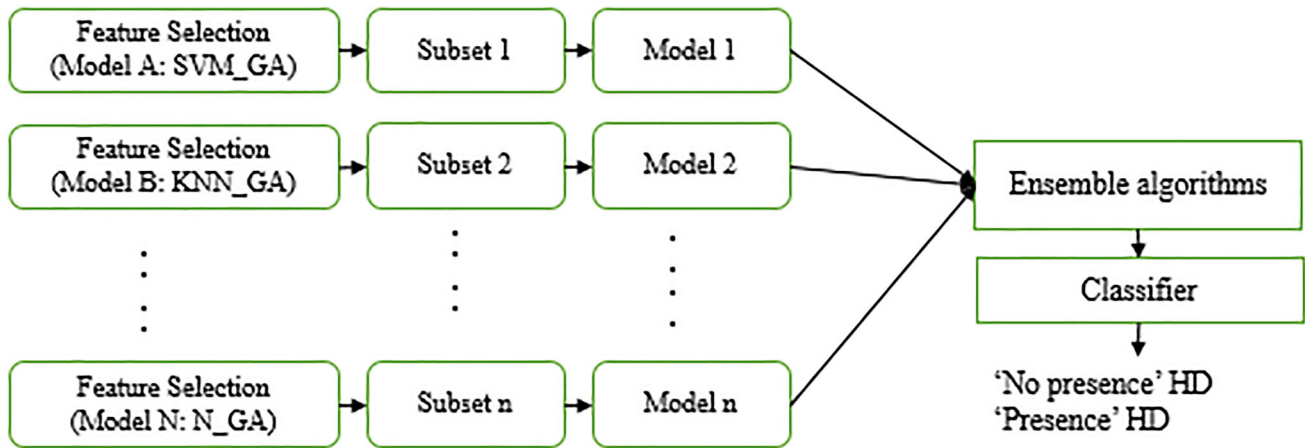
Fig. 4 The structures of the classifier with machine learning-based genetic algorithms

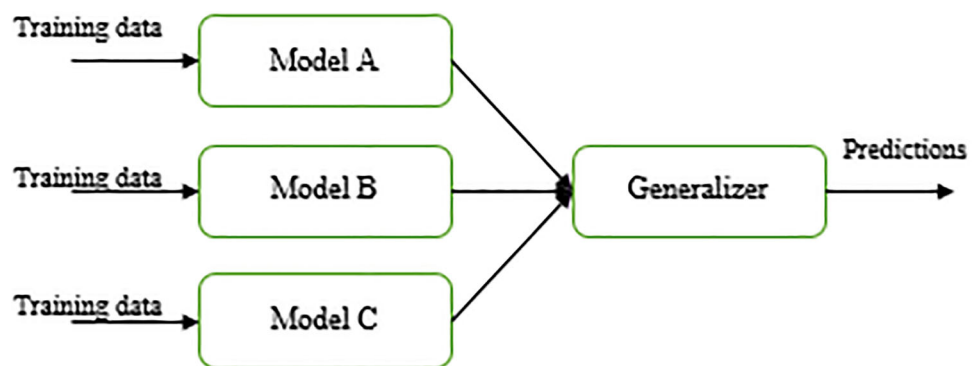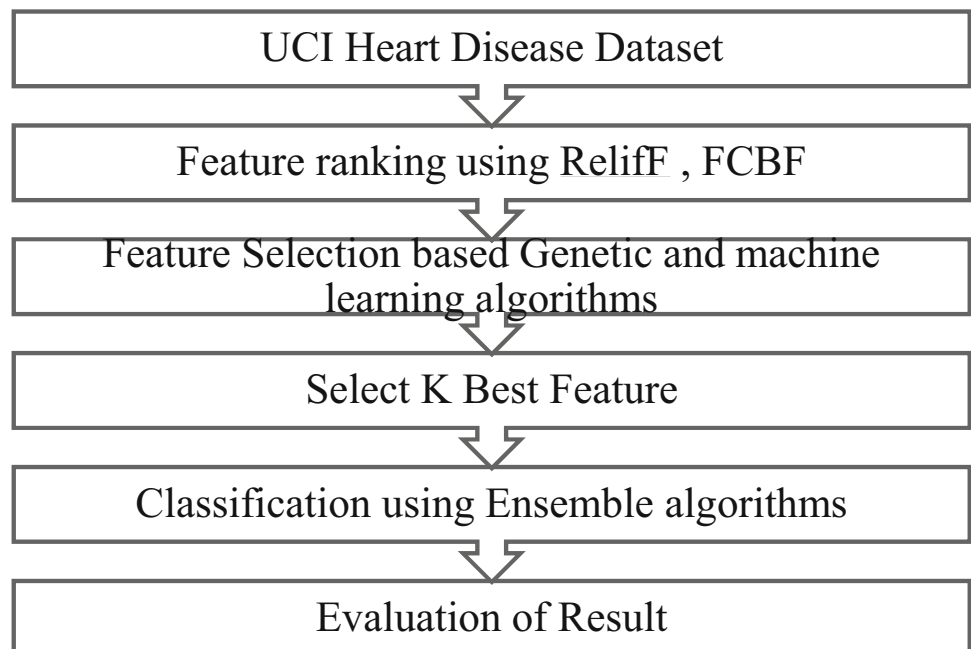**Fig. 5** The structures of ensemble learning

Fig. 6 Flow chart of the implemented method

the AI field, a GA may be a search heuristic that mimics the selection method. This heuristic is often used to generate helpful solutions to problems and optimization.

Initially, 13 attributes were concerned with predicting heart disease. Accordingly, in this study, GA is employed for feature selection, an efficient algorithm for solving large-
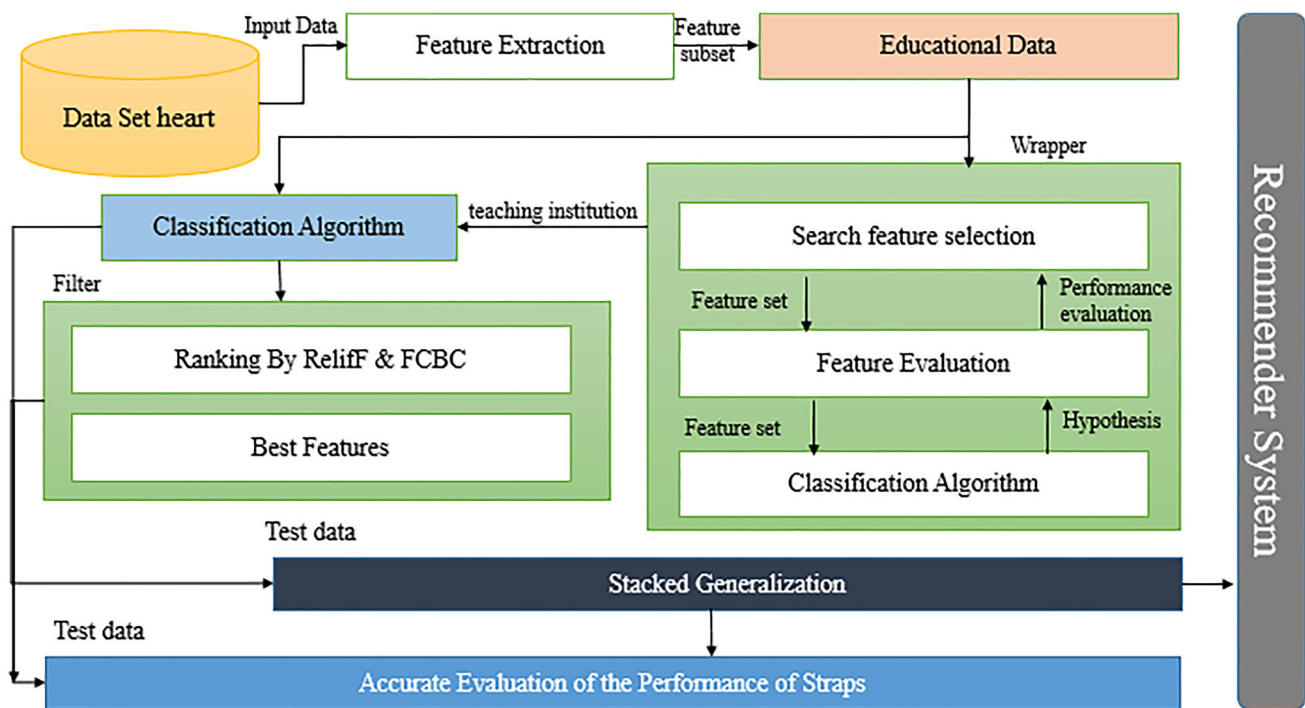
Fig. 7 The flow chart of the proposed method

scale problems, and maybe accustomed to finding an optimal feature subset. In GA, individuals are typically represented by n-bit binary vectors. Each of those individuals would represent a feature subset during a feature selection problem. It is supposed that the standard of every candidate solution will be assessed by employing a fitness function. During this study, classification accuracy was taken into account as a fitness function.

We used the FCBF attribute evaluator with the Ranker search and the ReliefF attribute evaluator with the Ranker search method in the first step. Next, we used the genetic algorithm to select the valuable features in the rapid diagnosis of heart disease.

Subsequently, seven classifiers, including SVM, NB, Dtree, MLP, KNN, RFC, and LR at level zero and AdaBoost algorithm at level one, are used in ensemble learning to predict the diagnosis of patients with the selected feature as obtained. Ensemble learning has advantages like better generalization capability and less computational time than traditional machine learning algorithms. Finally, we aggregate the results of the various models. Figure 7 illustrates the method of multi-model ensemble feature selection. Figure 7 illustrates the method of multi-models ensemble feature selection.

3.2 GA-stacking

The standard GA algorithm was used. The GA method consists of four significant steps: (1) The features were coded as binary genomes, i.e., "1" means selected and "0" means not selected (The phenotypes labeled as "0" represent the removed features, while phenotypes labeled as "1" represent the chosen features)).

Thus, phenotypes labeled as "0" are called "diminished features," while those labeled as "1" are called "highly significant features." On the idea of phenotypes, a bunch of subsets is formed by each genotype. These subsets are used as training sets for the proposed framework. The population of chromosomes was randomly generated; (2) each chromosome was evaluated using the fitness function, and the best features were selected; (3) the chromosomes were modified using crossover and mutation to create a brand-new generation of the population; (4) the new generation continued to (2) again until stopping criteria were met. A flow chart of the GA method is shown in Fig. 8.

The population size was set at 50. The ranking was used because of the selection strategy, where individuals ranking within the best 50% fitness were selected to make offspring. A single-point crossover with a proportion of 0.6 and a single-point mutation with a proportion of 0.033 was applied to develop the following population. The elitist strategy selection was set to 2, i.e., the smallest two individuals of the current generation were included within the next population. The number of generations with identical best fitness values was twenty. The number of generations was set at 100. The quantity of GA runs with different initial conditions (some data splits) was set to 100.

Fig. 8 Overview of genetic algorithm routing with Iterative Steps
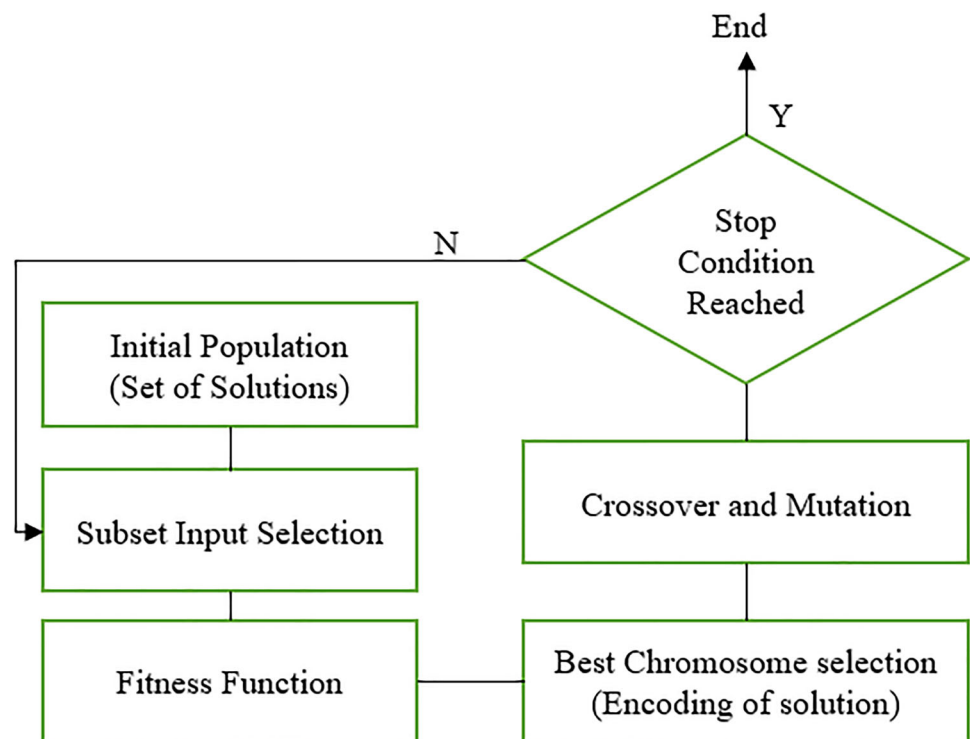
3.2.1 Fitness function

In this study, the linear ranking value was accustomed to evaluating the classifier's fitness. The linear ranking could be a statistical agreement between predicted and actual values. The performance of the classifier depends on how the training set is specified. Due to the limited size of the information, a repeated random sub-sampling strategy was used for cross-validation. Before each GA analysis, the information was randomly split ten times into 70% training and 30% validation sets. The performance on each stage was recorded because of the linear ranking value. The median of those ten linear ranking values was then recorded because of the fitness value of the genome.

3.2.2 Crossover (recombination)

In Crossover, two-parent solutions are taken to supply a toddler. Afterward, the choice (reproduction) process is finished, followed by the event of higher individuals. In this paper, the one-point crossover is employed to perform the mating of two chromosomes. Here, two chromosomes are cut once, exchanging the cuts between the two chromosomes.

3.2.3 Mutation

Once the crossover process is finished, the strings are subjected to the mutation process. "Mutation of touch" encompasses flipping a small amount from 0 to 1 and 1 to 0.

3.2.4 Setting the GA parameters

GAs have other parameters. Structural parameters like the population size and execution parameters like mutation and elite rate have to be set.

Nevertheless, we have used the most frequent values for these parameters, yielding good results. Table 2 shows the importance of the GA parameters used in this experiment.

4 Experimental design

4.1 Dataset

This study, the data set of heart patients available in (http://archive.ics.uci.edu/ml/datasets/statlog+(heart)) has been used, which has 13 useful variables and 270 records. These variables and abbreviations are listed in Table 3. 75% of the data is for training, and 25% of the data is for testing.

The dataset at the University of California, Irvine (UCI) was collected by David Aha [12]. The target of the database is to classify the presence or absence of heart illness given the issues of assorted medical tests given to a patient. The database includes 13 attributes indicating whether or not the patient has a heart illness. This dataset is a benchmark dataset because it factors actual patient data and has been extensively used to test many data processing methods [68–72].

Table 2 Genetic algorithms parameters

Parameter name	Value
Population size	50
Number of generation	100
Probability of crossover	0.6
Probability of mutation	0.033
Type of mutation	Uniform
Type of selection	Rank-Based
Stall generation limit	10
Stall time limit	Infinite
Elite count	2
Number of variables	13
Fitness function	Linear ranking
Encoding	Binary

GAs have other parameters. Structural parameters like the population size, and execution parameters like mutation and elite rate, have to be set. Nevertheless, in this work, we have used the most frequently values for this parameters yielding good results. In this table shows the values of the GA parameters used in this experiment

4.2 Data preprocessing

Data preprocessing is vital to arrange the heart data so that a machine learning model can accept it. Separating the training and testing data sets ensures that the model learns only from training data and tests its performance with the testing data. The data set used was divided into training and test data. The data set used was divided into training and test data. The datasets had irrelevant, unexplained, null, or repeated features. The most important considerations for cleansing were to assign one category for missing values called "null value" and to form rules that consider data consistency. The input file is processed through various steps to enhance the system's performance. This process can be automated using mathematical modeling and statistical knowledge [45–48].

4.3 Evaluation of result

For the study, a Jupyter notebook was used for implementation, and the Python programming language was used for coding. Also, the accuracy, specificity, and sensitivity criteria were utilized to compare the classification efficiency.

According to Table 4, FN and FP stand for the number of false-negative and false-negative samples. Additionally, TN and TP represent the number of true negative or true positive samples. Sensitivity measures the degree of positives that are correctly specified against. In addition, Specificity measures the degree of negatives that are accurately discriminated against. AUC is the associate index to live the efficiency of the classifier. In addition, the F1 score was alive of the accuracy of a binary model. To boot, the efficiency was estimated

with F-measure (F1) to check the similarity and variety of efficiency [45, 46].

This article was implemented using the Python programming language version 3.7 in the Anaconda environment of the Jupyter Notebook platform. The implementation details are shown in Table 5.

Evaluating the accuracy of multi-class algorithms requires complex methods due to the large number of datasets tested, the variety of techniques used, and the characteristics of the data set, which includes both balanced and unbalanced data. The performance of these algorithms is measured through various parameters such as accuracy, sensitivity, and pressure [45–48]. Understanding these metrics allows users to understand how well a developed categorization model analyzes textual data. In the field of multi-class problems, traditionally, only the accuracy obtained from the classification is reported as the primary criterion of evaluation and generality [45–48], which is defined as follows: Reciprocal measurement is also described.

Accuracy: Indicate the quantity of "correct predictions made" by the category, divided by the quantity of "total predictions made" by a similar category

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \qquad (1)$$

Sensitivity: Real positive rate: If the result is positive for the person, in a few percent of cases, the model will be positive, which is calculated from the following formula.

$$Sensitivity = \frac{TP}{TP + FN} \qquad (2)$$

Specificity: If the result is negative for the person, the result will be negative in a few percent of cases. Which is calculated from the following formula [45, 46].

$$Specificity = \frac{TN}{TN + FP} \qquad (3)$$

5 Results and discussion

Feature extraction is one of the most important methods for reducing dimensions in data preprocessing because databases regularly have redundant and irrelevant attributes that negatively affect the efficiency and complexity of classification algorithms. Feature extraction has two main goals: to decrease the number of attributes and improve the classification efficiency due to its inherent nature [50].

We used the ReliefF with the Ranker search method to attribute features in the first step. We left the method's parameters as the default, which compares every instance with its five nearest neighbors. The highest features with ReliefF were thal, sex, and CP, with 0.0821, 0.0793, and 0.0790.

Table 3 Features of the heart-type dataset

ID	Meaning	Type
Age (age in the year)	Patient age	Integer
Sex	Gender	Numerical two values
Chest pain	The location of chest pain	Numeric values
Blood pressure	Blood pressure	Integer
Cholesterol	Cholesterol content	Integer
Blood sugar	Blood sugar	Numerical two values
Electrocardiographic	ECG result (electrocardiographic)	Three-digit numbers (0, 1, and 2)
Heart rate	Heartbeat	Integer
Exercise-induced	Angina result from an exercise test	Numerical two values
Depression	ST depression rate	Real number
Slope	ST Exercise Test Result	Numeric values
Ca	Is there a blockage in the arteries or not?	Numerical two values
Thal	Thalassemia problem	Numeric values
C(Objective variable)	Is the patient at risk for a heart attack?	Two-digit number

As mentioned earlier, the popular and publicly available UCI heart disease dataset is used in this research. The UCI heart disease dataset consists of a total 76 attributes. However, majority of the existing studies have used only a maximum of 14 attributes. Different datasets have been based on the UCI heart disease data. Computational intelligence researchers, however, have mainly used the Cleveland dataset consisting of 14 attributes. The 14 attributes of the Cleveland dataset along with the values and data types are as follow

1. Age: age in years (numeric)
2. Sex: male, female (nominal)
3. Chest pain type (CP):
(a) Typical angina (angina)
(b) Atypical angina (abnang)
(c) Non-anginal pain (notang)
(d) Asymptomatic (asympt) (nominal)
From medical point of view,
(a) Typical angina is the condition in which the past history of the patient shows the usual symptoms and so the possibility of having coronary artery blockages is high
(b) Atypical angina refers to the condition that the patient's symptoms are not detailed and so the probability of blockages is lower
(c) Non-angina pain is the stabbing or knife-like, prolonged, dull, or painful condition that can last for short or long periods of time
(d) Asymptomatic pain shows no symptoms of illness or disease and possibly will not cause or exhibit disease symptoms
4. Trestbps: patient's resting blood pressure in mm Hg at the time of admission to the hospital (numeric)
5. Chol: Serum cholesterol in mg/dl
6. Fbs: Boolean measure indicating whether fasting blood sugar is greater than 120 mg/dl: (1 = True; 0 = false) (nominal)
7. Restecg: electrocardiographic results during rest. Three types of values normal (norm), abnormal (abn): having ST-T wave abnormality, ventricular hypertrophy (hyp) (nominal)
8. Thalach: maximum heart rate attained (numeric)
9. Exang: Boolean measure indicating whether exercise-induced angina has occurred: 1 = yes, 0 = no (nominal)
10. Oldpeak: ST depression brought about by exercise relative to rest (numeric)
11. Slope: the slope of the ST segment for peak exercise. Three types of values upsloping, flat, downsloping (nominal)
12. Ca: number of major vessels (0–3) colored by fluoroscopy (numeric)
13. Thal: the heart status (normal, fixed defect, reversible defect) (nominal)
14. The class attributes: value is either healthy or heart disease (sick type: 1, 2, 3, and 4)

Age, slope, depression, and Ca had significantly lower scores than other features: 0.0188, 0.0157, 0.0118, and 0.0114. We removed these four features and saved the remaining nine into a dataset referred to as the Heart-2 dataset.

Also, we used the FCBF attribute evaluator with the Ranker search method. The most significant features according to FCBF were cp, heart rate, and thal with scores of 0.1728, 0.1701, and 0.1598. This time, the lowest-ranked features were age, slope, depression, and Ca, with 0.0341, 0.0228, 0.0116, and 0.0088. We removed these four features and selected the remaining nine features to create a dataset that we will refer to as the Heart-2 dataset.

Table 6 as represented above, filter strategies perform the feature choice method as a preprocessing step with no induction algorithm. The general characteristics of the coaching information are accustomed to choosing options. The results obtained for the four filters studied (relief, FCBF, and ND) are compared and mentioned. The ultimate aim of this study is to pick out a filter to construct a hybrid technique for feature selection.

Table 4 Confusion matrix

	Actual class	
	Negative	Positives
Predicted class		
Negative	True Positives (TP)	False Positives (FP)
Positives	False Negatives (FN)	True Negatives (TN)

A confusion matrix summarizes the classification performance of a classifier with respect to some test data. It is a two-dimensional matrix, indexed in one dimension by the true class of an object and in the other by the class that the classifier assigns
True Positive:
Interpretation: You predicted positive and it's true
You predicted that a woman is pregnant and she actually is
True Negative:
Interpretation: You predicted negative and it's true
You predicted that a man is not pregnant and he actually is not
False Positive: (Type 1 Error)
Interpretation: You predicted positive and it's false
You predicted that a man is pregnant but he actually is not
False Negative: (Type 2 Error)
Interpretation: You predicted negative and it's false
You predicted that a woman is not pregnant but she actually is
Just Remember, we describe predicted values as Positive and Negative and actual values as True and False

Table 5 Software requirements

Distribution	Anaconda Navigator and Google Colab
Packages	Matplotlib, NumPy, Pandas, Sci-it Learn
Language	Python 3.7
IDE	Jupyter Notebook (google colab)

This article was implemented using the Python programming language version 3.7 in the Anaconda environment of the Jupyter Notebook platform. The implementation details are shown in this table

The following step was to seek out the simplest feature subset using GA-supported seven different classifiers in the Heart-2 dataset. Clear and intuitive clinical interpretation is required to help clinicians in the higher cognitive processes; thus, GA was chosen for its ability to explore the full-feature space and convey the simplest feature subset. SVM, NB, Dtree, MLP, KNN, RFC, and LR were classifiers for the GA. Each classifier selected 7–9 features using the GA. This feature set is employed in our recognition system and its clas-

sification results. Table 7 and Fig. 9 classifiers were adopted to compare the performance of the selection approaches, using Accuracy, sensitivity and specificity as matric.

The result confirms the utility of feature choice for classification and, therefore, the superiority of wrapper strategies. According to studies, filter methods for working with large data sets have little computational cost as opposed to wrapping methods due to some problems caused by use, and therefore, filter methods are a reasonable option.

Ultimately, it proposes analyzing and comparing the accuracy of various data processing classification plans using the ensemble learning method to predict a heart condition. SVM, NB, Dtree, MLP, KNN, RFC, LR at level zero and Adaboost at level one were used. The analyses indicate that the Stacked-Genetic algorithm in diagnosing heart disease using the Adaboost technique outperforms the other methods mentioned above. Also, improved performance in terms of performance metrics provides a better understanding of

Table 7 Results of different wrapper method

	Accuracy	Sensitivity	Specific
SVM-GA	84	82	79
NB-GA	79	81	82
Dtree-GA	94	92	92.87
MLP-GA	92	91	90
KNN-GA	87	86	87
RFC-GA	94	93	92
LR-GA	90	89	91
Stacked-GA	97.57	96	97

In this table, seven classifiers were adopted to compare the performance of the selection approaches, using Accuracy and Receiver Operating Characteristics area (aROC) as matric. Result confirms the utility of feature selection for classification and the superiority of wrapper methods. However, some problems do arise from using wrapper methods and, the evidence is proposed that filters are a reasonable alternative with the limited computational cost for dealing with large datasets
*Note: SVM_GA, Support vector Machine based Genetic Algorithms; NB_GA, Naïve Bayes based Genetic Algorithms; Dtree-GA, Decision Tree Classifier based Genetic Algorithms; MLP_GA, multi layer perceptron based Genetic Algorithms; KNN_GA, nearest neighbor based Genetic Algorithms; RFC_GA, Random Forest Classifier based Genetic Algorithms; LR_GA, Logistic Regression based Genetic Algorithms; SG_GA, Stacked Generalization based Genetic Algorithms

Table 6 Result of filter method

Title	FCBF		Relief	
	Running time	Selected features	Running time	Selected features
Heart	20	8	32	9
Accuracy	0.872		0.9416	

As described above, filter methods carry out the feature selection process as a pre-processing step with no induction algorithm. The general characteristics of the training data are used to select features. The results obtained for the four filters studied (Relief, FCBF) are compared and discussed. The final aim of this study is to select a filter to construct a hybrid method for feature selection
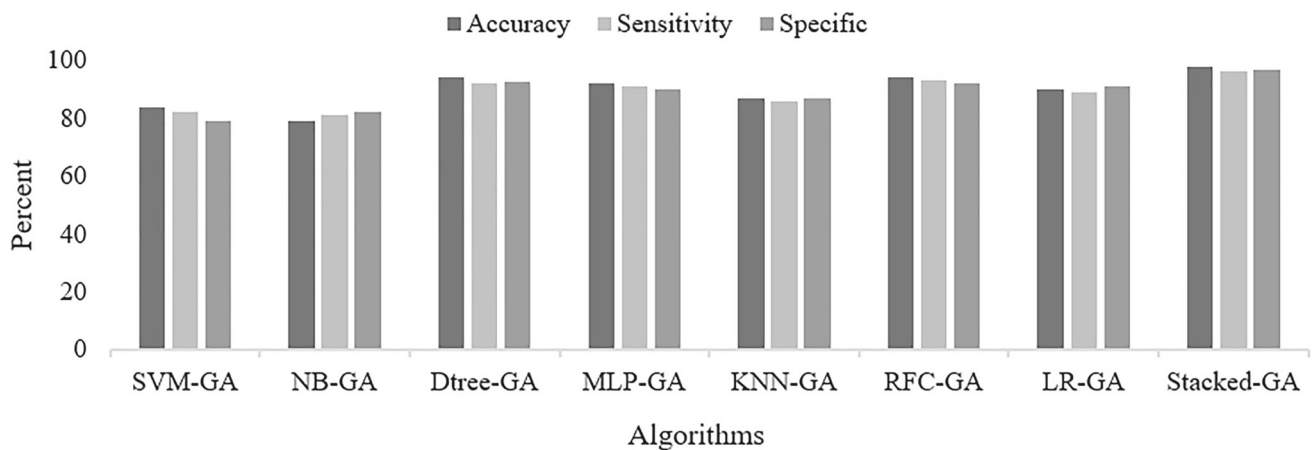
Fig. 9 Results of different wrapper methods

Table 8 Comparison of accuracy rates in holdout and cross-validation approach

Model	Holdout			K-fold = 2			K-fold = 5			K-fold = 10		
	ACC	Sen	Spec	ACC	Sen	Spec	ACC	Sen	Spec	ACC	Sen	Spec
RF	88	87	86	92	91	90	93	92	91	94	93	92
KNN	76	77	73	77	76	77	81	80	80.56	89	85	87
MLP	91	90	92	93	92	91	93	92	91	94	93	92
Dtree	88	88	87	95	94	96	93	92	89	94	91	93
NB	75	74	72	79	78	78.5	78	77	76	86	85	83
LR	74	73	71	75	74	75	81	80	79	87	85	84
SVM	73	71	70	73	72	69	88	85	84	89.05	89	88
Suggest Method (ST-GA)	88	87	87.5	92.56	92	93	96	95	96	97.57	98	97

Diagnosis of detection heart disease was compared in rates among various algorithms

The performance of different machine learning algorithms, such as logistic regression, random forest, and support vector machine, etc. was compared with the accuracy based on evolutionary algorithms. The accuracy of the model predictions showed that Ensemble learning based on a genetic algorithm had the best performance and could lead to 97.57% accuracy

*Note: RF, Random Forest; SVM, support vector machine; ANN, artificial neural network; LR, logistic regression; GB, Gradient Boosting Classifier; DT, Decision Tree Classifier; NB, Naïve Bayes, Accuracy, Sensitivity, Specificity; PPV, positive predictive value; NPV, negative predictive value; TPR, true positive rate; FPR, false positive rate

the ensemble models' accuracy, reliability, and usefulness in favor of improved performance for heart disease prediction. The results of the paper are shown in Table 8 and Fig. 10.

One of the advantages of the filter, a primary component, is that the calculations are apparent, overfitting is prevented, and suitable for specific datasets. However, this method also has disadvantages, including that the unwanted desired subset may be removed from a subset. In general, the biggest distinction between the filtering and wrapper strategies is that the primary one involves a non-repetitive calculation within the information; the latter will adapt itself to machine learning algorithms and be used. It can be concluded that the wrapper results will be better than the filtering method, but the computational cost of this method is high [51–54].

Features selection for the classification model attempts to select a minimally sized subset according to the following

criteria: (1) The classification accuracy should increase; (2) the values for the selected features should be as close as possible to the original class distribution. Figure 9 shows the eight classification models' accuracy, sensitivity, and specificity when applied to the dataset. As we can see from Fig. 9, using feature selection and extraction to the heart-disease dataset had varying results depending heavily on what classification algorithm it was paired with to construct a model. The model built using the Heart-2 dataset by the proposed algorithm had the highest precision of any model we created.

Increasing advances in computer and electronic technology have provided scientists with the opportunity to collect and study data on various phenomena. Data mining and machine learning are vital in analyzing and constructing data-driven diagnostic models [55]. The performance of different ML algorithms, such as LR, RF, SVM, etc., is compared with
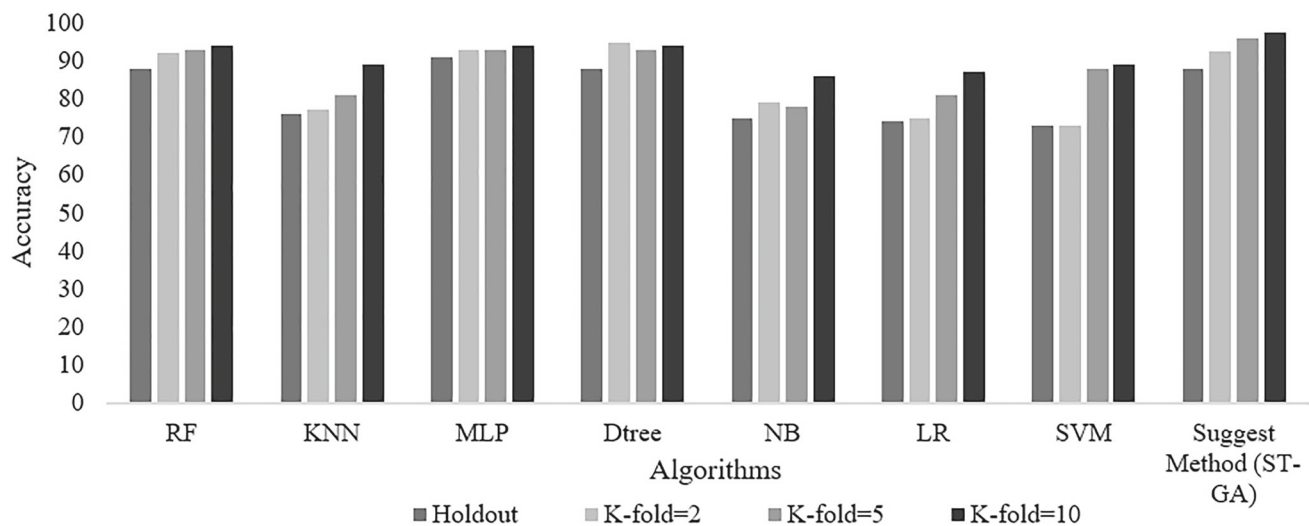
Fig. 10 Comparison of accuracy rates in Cross-Validation and Holdout approach. Diagnosis of detection heart disease was compared in rates among various algorithms

the accuracy based on evolutionary algorithms. The accuracy of the model predictions showed that ensemble learning based on a genetic algorithm had the best performance and could lead to 97.57% accuracy, 96% sensitivity, and 97% specific.

Figure 10 shows the accuracy of the eight classification models when applied to the dataset. Many of the models built on the Heart-2 dataset were performed high. However, the classification model built against these features set using the proposed method achieved the highest recall of any model and had the highest accuracy of any model applied. So experimental results procured for the heart-2 dataset show that the proposed method outperformed the other algorithms in terms of all metrics. We believe that our heart prediction framework will assist doctors in predicting hearts with high accuracy.

6 Conclusion and future work

As mentioned, one of the most common chronic diseases and causes of adult death worldwide is heart disease [50]. According to the Health Announcement, Medical Education Department: 33 to 38% of deaths in the country are due to cardiovascular diseases, and Iran has the highest rate of heart death in the world. Changes in people's lifestyles increase the prevalence of heart disease in Iran. Evidence for lifestyle changes shows that the prevalence of the cardiovascular disease is increasing in Iran.

It is calculable that by 2020, the mortality caused by these diseases can increase to 25 million. Artificial intelligence has many applications in medicine, including forecasting the Spread of COVID-19 [73], diabetes [45], cancer [46], heart [74], and Covid-19 [75]. This article used an ensemble learn-

Table 9 Comparison of our results with those of other studies

Row	Authors	Methods	Accuracy
1	Duch et al. [15]	KNN classifier	85.6
2	Sahan et al. [58]	AWAIS	82.59
3	Kahramanli and Allahverdi [14]	The hybrid neural network method	86.8
4	Helmy and Rasheed [59]	Algebraic Sigmoid	85.24
5	Polat and Gunes [60]	RBF kernel F-score + LS-SVM	83.70
6	Karegowda et al. [61]	GA + Naıve Bayes	85.87
7	Buscema et al. [62]	TWIST algorithm	84.14
8	Tomar and Agarwal [6]	Feature selection-based LSTSVM	85.59
9	Gokulnath et al. [63]	SVM-GA	81.32
10	Arabasadi et al. [64]	NN- Genetic algorithm	89.04
11	Abdullah et al. [65]	RF with RFFS	85.25%
12	Kishor et al. [66]	K-NN, DTs, RF, MLP, NB, L-SVM	92.3%
13	Amen et al. [67]	NB, KNN	82%, 79%

Heart disease is one of the most common chronic diseases and causes of adult death worldwide [33]. Health, Medical Education Department has announced that 33–38 percent of deaths in the country are due to cardiovascular disease, and Iran has the highest rate of heart death in the world. Changes in people's lifestyles increase the prevalence of heart disease in Iran. Evidence for lifestyle changes shows that the prevalence of cardiovascular disease is increasing in Iran. It is estimated that by 2020, the mortality caused by these diseases will increase to 25 million. In this paper, an Ensemble learning approach based on a genetic algorithm was used to select effective features in the immediate diagnosis of heart disease. The results show the high performance of the proposed method in the immediate diagnosis of heart disease with 97.57% accuracy

ing method based on a genetic algorithm to select valuable features in the immediate diagnosis of heart disease. The results show the high efficiency of the suggested approach in diagnosing heart disease with 97.57% accuracy.

As a suggestion for future research, the Naive Bayesian method, decision tree, or support vector regression can be used as a classifier model, and its combination with the new approach to select a combination feature extracted from this study can be used to diagnose and predict metastatic diseases of breast cancer, lung cancer, Covid-19 and various diseases.

Declarations

Conflict of interest The author declares that they have no conflict of interest.

Ethical approval Not required.

References

1. Anbarasi, M., Anupriya, E., Iyengar, N.C.S.N.: Enhanced prediction of heart disease with feature subset selection using genetic algorithm. Int. J. Eng. Sci. Technol. **2**(10), 5370–5376 (2010)
2. Petrie, K.J., Jones, A.S.: Coping with chronic illness. Annu. Rev. Clin. Psychol. (2019). https://doi.org/10.1017/CBO9780511543579.011
3. Zhang, J., Lafta, R.L., Tao, X., Li, Y., Chen, F., Luo, Y., Zhu, X.: Coupling a fast Fourier transformation with a machine learning ensemble model to support recommendations for heart disease patients in a telehealth environment. IEEE Access **5**, 10674–10685 (2017). https://doi.org/10.1109/ACCESS.2017.2706318
4. Nourmohammadi-Khiarak, J., Feizi-Derakhshi, M.R., Behrouzi, K., Mazaheri, S., Zamani-Harghalani, Y., Tayebi, R.M.: A new hybrid method for heart disease diagnosis utilizing optimization algorithm in feature selection. Health Technol. **10**(3), 667–678 (2020). https://doi.org/10.1007/s12553-019-00396-3
5. Liu, X., Wang, X., Su, Q., Zhang, M., Zhu, Y., Wang, Q., Wang, Q.: A hybrid classification system for heart disease diagnosis based on the RFRS method. Comput. Math. Methods Med. (2017). https://doi.org/10.1155/2017/8272091
6. Tomar, D., Agarwal, S.: Feature selection based least square twin support vector machine for diagnosis of heart disease. Int. J. Bio-Sci. Bio-Technol. **6**(2), 69–82 (2014)
7. Karayılan, T., Kılıç, Ö.: Prediction of heart disease using neural network. In: Computer Science and Engineering (UBMK), 2017 international conference on (pp. 719–723). IEEE (2017). https://doi.org/10.1109/UBMK.2017.8093512
8. Polat, K., Güneş, S.: A new feature selection method on classification of medical datasets: Kernel F-score feature selection. Expert Syst. Appl. **36**(7), 10367–10373 (2009). https://doi.org/10.1016/j.eswa.2009.01.041
9. Shilaskar, S., Ghatol, A.: Feature selection for medical diagnosis: Evaluation for cardiovascular diseases. Expert Syst. Appl. **40**(10), 4146–4153 (2013). https://doi.org/10.1016/j.eswa.2013.01.032
10. Shah, S.M.S., Batool, S., Khan, I., Ashraf, M.U., Abbas, S.H., Hussain, S.A.: Feature extraction through parallel probabilistic principal component analysis for heart disease diagnosis. Phys. A Stat. Mech. Appl. **482**, 796–807 (2017). https://doi.org/10.1016/j.physa.2017.04.113
11. Tripoliti, E.E., Papadopoulos, T.G., Karanasiou, G.S., Naka, K.K., Fotiadis, D.I.: Heart failure: diagnosis, severity estimation and prediction of adverse events through machine learning techniques. Comput. Struct. Biotechnol. J. **15**, 26–47 (2017). https://doi.org/10.1016/j.csbj.2016.11.001
12. Chen, H.L., Yang, B., Liu, J., Liu, D.Y.: A support vector machine classifier with rough set-based feature selection for breast cancer diagnosis. Expert Syst. Appl. **38**(7), 9014–9022 (2011). https://doi.org/10.1016/j.eswa.2011.01.120
13. Wang, S.-J., Mathew, A., Chen, Y., Xi, L.-F., Ma, L., Lee, J.: Empirical analysis of support vector machine ensemble classifiers. Expert Syst. Appl. **36**(3), 6466–6476 (2009). https://doi.org/10.1016/j.eswa.2008.07.041
14. Kahramanli, H., Allahverdi, N.: Design of a hybrid system for the diabetes and heart diseases. Expert Syst. Appl. **35**(1–2), 82–89 (2008). https://doi.org/10.1016/j.eswa.2007.06.004
15. Duch, W., Adamczak, R., Grabczewski, K.: A new methodology of extraction, optimization and application of crisp and fuzzy logical rules. IEEE Trans. Neural Netw. **12**(2), 277–306 (2001). https://doi.org/10.1109/72.914524
16. Feshki, M.G., Shijani, O.S.: Improving the heart disease diagnosis by evolutionary algorithm of PSO and Feed Forward Neural Network. In: Paper presented at the artificial intelligence and robotics (IRANOPEN) (2016). https://doi.org/10.1109/RIOS.2016.7529489
17. Taye, M., Tajfard, M., Saffar, S., Hanachi, P., Amirabadizadeh, A.R., Esmaeily, H., et al.: Hs-CRP is strongly associated with coronary heart disease (CHD): a data mining approach using decision tree algorithm. Comput. Methods Prog. Biomed. **141**, 105–109 (2017). https://doi.org/10.1016/j.cmpb.2017.02.001
18. Kangwanariyakul, Y., Naenna, T., Nantasenamat, C., Tantimongcolwat, T.: Data mining of magnetocardiograms for prediction of ischemic heart disease. EXCLI J **9**, 82 (2010)
19. Cleveland Heart Disease Dataset: http://archive.ics.uci.edu/ml/datasets/Heart+Disease (2017). Accessed 3 Feb 2017
20. Pouriyeh, S., Vahid, S., Sannino, G., De Pietro, G., Arabnia, H., Gutierrez, J.: A comprehensive investigation and comparison of machine learning techniques in the domain of heart disease. In: Computers and communications (ISCC), 2017 IEEE Symposium on (pp. 204–207). IEEE (2017). https://doi.org/10.1109/ISCC.2017.8024530
21. Singh, N., Singh, P.: A stacked generalization approach for diagnosis and prediction of type 2 diabetes mellitus. In: computational intelligence in data mining (pp. 559–570). Springer, Singapore (2020). https://doi.org/10.1007/978-981-13-8676-3_47
22. Babič, F., Olejár, J., Vantová, Z., Paralič, J.: Predictive and descriptive analysis for heart disease diagnosis (2021). https://doi.org/10.15439/2017F219
23. Paul, A.K., Shill, P.C., Rabin, M.R.I., Akhand, M.A.H.: Genetic algorithm based fuzzy decision support system for the diagnosis of heart disease. In: Informatics, electronics and vision (ICIEV), 2016 5th international conference on (pp. 145–150), IEEE (2016). https://doi.org/10.1109/ICIEV.2016.7759984
24. Safdar, S., Zafar, S., Zafar, N., Khan, N.F.: Machine learning-based decision support systems (DSS) for heart disease diagnosis: a review. Artif. Intell. Rev. (2017). https://doi.org/10.1007/s10462-017-9552-8

25. Tripoli, E.E., Papadopoulos, T.G., Karanasiou, G.S., Kalatzis, F.G., Bechlioulis, A., Goletsis, Y., Fotiadis, D.I.: Estimation of New York Heart Association class in heart failure patients based on machine learning techniques. In: Biomedical & Health Informatics (BHI), 2017 IEEE EMBS International Conference on (pp. 421–424). IEEE (2017). https://doi.org/10.1109/BHI.2017.7897295

26. Detrano, R., Janosi, A., Steinbrunn, W., Pfisterer, M., Schmid, J.-J., Sandhu, S., Guppy, K.H., Lee, S., Froelicher, V.: International application of a new probability algorithm for the diagnosis of coronary artery disease. Am. J. Cardiol. **64**(5), 304–310 (1989). https://doi.org/10.1016/0002-9149(89)90524-9

27. Edmonds, B.: Using localised 'gossip' to structure distributed learning (2005)

28. Gudadhe, M., Wankhade, K., Dongre, S.: Decision support system for heart disease based on support vector machine and artificial neural network. In: Computer and communication technology (ICCCT), 2010 international, 2010, pp. 741–745 (2010). https://doi.org/10.1109/ICCCT.2010.5640377

29. Doppala, B.P., Bhattacharyya, D., Chakkravarthy, M., Kim, T.H.: A hybrid machine learning approach to identify coronary diseases using feature selection mechanism on heart disease dataset. Distrib. Parall. Databases (2021). https://doi.org/10.1007/s10619-021-07329-y

30. Palaniappan, S., Awang, R.: Intelligent heart disease prediction system using data mining techniques. In: IEEE/ACS international conference on computer systems and applications, IEEE, pp. 108–115 (2008). https://doi.org/10.1109/AICCSA.2008.4493524

31. Gupta, V., Mittal, M., Mittal, V., Gupta, A.: ECG signal analysis using CWT, spectrogram and autoregressive technique. Iran J. Comput. Sci. (2021). https://doi.org/10.1007/s42044-021-00080-8

32. Verma, A.K., Pal, S., Tiwari, B.B.: Skin disease prediction using ensemble methods and a new hybrid feature selection technique. Iran J. Comput. Sci. **3**(4), 207–216 (2020). https://doi.org/10.1007/s42044-020-00058-y

33. Osubor, V.I., Egwali, A.O.: A neuro-fuzzy approach for the diagnosis of postpartum depression disorder. Iran J. Comput. Sci. **1**(4), 217–225 (2018). https://doi.org/10.1007/s42044-018-0021-6

34. Schiezaro, M., Pedrini, H.: Data feature selection based on Artificial Bee Colony algorithm. EURASIP J. Image Video Process. **2013**(1), 47 (2013). https://doi.org/10.1186/1687-5281-2013-47

35. Chandrashekar, G., Sahin, F.: A survey on feature selection methods. Comput. Electr. Eng. **40**(1), 16–28 (2014). https://doi.org/10.1016/j.compeleceng.2013.11.024

36. Liu, H., Yu, L.: Toward integrating feature selection algorithms for classification and clustering. IEEE Trans. Knowl. Data Eng. **17**(4), 491–502 (2005). https://doi.org/10.1109/TKDE.2005.66

37. Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. J. Mach. Learn. Res. **3**, 1157–1182 (2003)

38. Tuv, E., Borisov, A., Runger, G., Torkkola, K.: Feature selection with ensembles, artificial variables, and redundancy elimination. J. Mach. Learn. Res. **10**, 1341–1366 (2009)

39. Vergara, J.R., Estévez, P.A.: A review of feature selection methods based on mutual information. Neural Comput. Appl. **24**(1), 175–186 (2014). https://doi.org/10.1007/s00521-013-1368-0

40. Jović, A., Brkić, K., Bogunović, N.: A review of feature selection methods with applications. In: 2015 38th international convention on information and communication technology, electronics and microelectronics (MIPRO) (pp. 1200–1205), IEEE (2015). https://doi.org/10.1109/MIPRO.2015.7160458

41. Bolón-Canedo, V., Sánchez-Maroño, N., Alonso-Betanzos, A.: A review of feature selection methods on synthetic data. Knowl. Inf. Syst. **34**(3), 483–519 (2013). https://doi.org/10.1007/s10115-012-0487-8

42. Remeseiro, B., Bolon-Canedo, V.: A review of feature selection methods in medical applications. Comput. Biol. Med. **112**, 103375 (2019). https://doi.org/10.1016/j.compbiomed.2019.103375

43. Shah, F.P., Patel, V.: A review on feature selection and feature extraction for text classification. In: 2016 international conference on wireless communications, signal processing and networking (WiSPNET) (pp. 2264–2268), IEEE (2016). https://doi.org/10.1109/WiSPNET.2016.7566545

44. Mwadulo, M.W.: A review on feature selection methods for classification tasks (2016)

45. Abdollahi, J., Moghaddam, B.N., Parvar, M.E.: Improving diabetes diagnosis in smart health using genetic-based ensemble learning algorithm. Approach to IoT infrastructure. Future Gen. Distrib. Syst. J. **1**, 23–30 (2019)

46. Abdollahi, J., Keshandehghan, A., Gardaneh, M., Panahi, Y., Gardaneh, M.: Accurate detection of breast cancer metastasis using a hybrid model of artificial intelligence algorithm. Arch. Breast Cancer (2020). https://doi.org/10.32768/abc.20207122-28

47. Abdollahi, J., Nouri-Moghaddam, B.: Hybrid stacked ensemble combined with genetic algorithms for diabetes prediction. Iran J. Comput. Sci. (2022). https://doi.org/10.1007/s42044-022-00100-1

48. Jafar, A., Firouz, A., Alireza, M., Paniz, A., Ghasem, F.-A.: Using Stacking methods based Genetic Algorithm to predict the time between symptom onset and hospital arrival in stroke patients and its related factors. JBE. **8**(1), 8–23 (2022)

49. Sutha, K., Tamilselvi, J.J.: A review of feature selection algorithms for data mining techniques. Int. J. Comput. Sci. Eng. **7**(6), 63 (2015)

50. Nouri-Moghaddam, B., Ghazanfari, M., Fathian, M.: A novel multi-objective forest optimization algorithm for wrapper feature selection. Expert Syst. Appl. (2021). https://doi.org/10.1016/j.eswa.2021.114737

51. Bashir, S., Khan, Z.S., Khan, F.H., Anjum, A., Bashir, K.: Improving heart disease prediction using feature selection approaches. In: 2019 16th International Bhurban conference on applied sciences and technology (IBCAST) (pp. 619–623). IEEE (2019). https://doi.org/10.1109/IBCAST.2019.8667106

52. Qin, C.J., Guan, Q., Wang, X.P.: Application of ensemble algorithm integrating multiple criteria feature selection in coronary heart disease detection. Biomed. Eng. Appl. Basis Commun. **29**(6), 1750043 (2017). https://doi.org/10.4015/S1016237217500430

53. Kavitha, R., Kannan, E.: An efficient framework for heart disease classification using feature extraction and feature selection technique in data mining. In: 2016 international conference on emerging trends in engineering, technology and science (icetets) (pp. 1–5), IEEE (2016). https://doi.org/10.1109/ICETETS.2016.7603000

54. Al-Tashi, Q., Rais, H., Jadid, S.: Feature selection method based on grey wolf optimization for coronary artery disease classification. In International conference of reliable information and communication technology (pp. 257–266). Springer, Cham (2018). https://doi.org/10.1007/978-3-319-99007-1_25

55. Nouri-Moghaddam, B., Ghazanfari, M., Fathian, M.: A novel filter-wrapper hybrid gene selection approach for microarray data based on multi-objective forest optimization algorithm. Dec. Sci. Lett. **9**(3), 271–290 (2020). https://doi.org/10.5267/j.dsl.2020.5.006

56. Mirjalili, S.: Genetic algorithm. In: Evolutionary algorithms and neural networks (pp. 43–55). Springer, Cham (2019). https://doi.org/10.1007/978-3-319-93025-1_4

57. Mathew, T.V.: Genetic algorithm. In: Report submitted at IIT Bombay (2012)

58. Sahan, S., Polat, K., Kodaz, H., Gunes, S.: The medical applications of attribute weighted artificial immune system (AWAIS): diagnosis of heart and diabetes diseases. Artif. Immune Syst. **3627**, 456–468 (2005). https://doi.org/10.1007/11536444_35

59. Helmy, T., Rasheed, Z.: Multi-category bioinformatics dataset classification using extreme learning machine. In: Proceedings of the IEEE congress on evolutionary computation (CEC '09), pp. 3234–3240, Trondheim, Norway

60. Polat, K., Gunes, S.: A new feature selection method on classification of medical datasets: kernel F-score feature selection. Expert Syst. Appl. **36**(7), 10367–10373 (2009). https://doi.org/10.1016/j.eswa.2009.01.041

61. Karegowda, G., Manjunath, A.S., Jayaram, M.A.: Feature subset selection problem using wrapper approach in supervised learning. Int. J. Comput. Appl. **1**(7), 13–17 (2010)

62. Buscema, M., Breda, M., Lodwick, W.: Training with Input Selection and Testing (TWIST) algorithm: a significant advance in pattern recognition performance of machine learning. J. Intell. Learn. Syst. Appl. **5**(1), 29–38 (2013)

63. Gokulnath, C.B., Shantharajah, S.P.: An optimized feature selection based on genetic approach and support vector machine for heart disease. Cluster Comput. **22**(6), 14777–14787 (2019)

64. Arabasadi, Z., Alizadehsani, R., Roshanzamir, M., Moosaei, H., Yarifard, A.A.: Computer-aided decision making for heart disease detection using hybrid neural network-Genetic algorithm. Comput. Methods Progr. Biomed. **141**, 19–26 (2017). https://doi.org/10.1016/j.cmpb.2017.01.004

65. Abdullah, A.A., Alhadi, N.A., Khairunizam, W.: Diagnosis of heart disease using machine learning methods. In: Intelligent manufacturing and mechatronics (pp. 77–89). Springer, Singapore (2021). https://doi.org/10.1007/978-981-16-0866-7_6

66. Kishor, A., Jeberson, W.: Diagnosis of heart disease using internet of things and machine learning algorithms. In: Proceedings of second international conference on computing, communications, and cyber-security (pp. 691–702). Springer, Singapore (2021). https://doi.org/10.1007/978-981-16-0733-2_49

67. Amen, K., Zohdy, M., Mahmoud, M.: Towards comparing machine learning models to foresee the stages for heart (2022)

68. Nishi, M., Ahmadi, H., Manaf, A.A., Rashid, T.A., Samad, S., Shahmoradi, L., Akbari, E., et al.: Coronary heart disease diagnosis through self-organizing map and fuzzy support vector machine with incremental updates. Int. J. Fuzzy Syst. **22**(4), 1376–1388 (2020). https://doi.org/10.1007/s40815-020-00828-7

69. Tama, B.A., Im, S., Lee, S.: Improving an intelligent detection system for coronary heart disease using a two-tier classifier ensemble. BioMed Res. Int. (2020). https://doi.org/10.1155/2020/9816142

70. Karadeniz, T., Tokdemir, G., Maraş, H.H.: Ensemble methods for heart disease prediction. New Gener. Comput. (2021). https://doi.org/10.1007/s00354-021-00124-4

71. Porto, R., Molina, J.M., Berlanga, A., Patricio, M.A.: Minimum relevant features to obtain explainable systems for predicting cardiovascular disease using the starlog data set. Appl. Sci. **11**(3), 1285 (2021). https://doi.org/10.3390/app11031285

72. Tougui, I., Jilbab, A., El Mhamdi, J.: Heart disease classification using data mining tools and machine learning techniques. Health Technol. **10**, 1137–1144 (2020). https://doi.org/10.1007/s12553-020-00438-1

73. Abdollahi, J., Irani, A.J., Nouri-Moghaddam, B.: Modeling and forecasting Spread of COVID-19 epidemic in Iran until Sep 22, 2021, based on deep learning. arXiv:2103.08178 (2021). https://doi.org/10.48550/arXiv.2103.08178

74. Abdollahi, J., Nouri-Moghaddam, B., Ghazanfari, M.: Deep neural network based ensemble learning Algorithms for the healthcare system (diagnosis of chronic diseases). arXiv:2103.08182 (2021). https://doi.org/10.48550/arXiv.2103.08182

75. Abdollahi, J.: A review of Deep learning methods in the study, prediction and management of COVID-19 (2022)

76. Nouri-Moghaddam, B., Ghazanfari, M., Fathian, M.: A novel bio-inspired hybrid multi-filter wrapper gene selection method with ensemble classifier for microarray data. Neural Comput. Appl. (2021). https://doi.org/10.1007/s00521-021-06459-9