Research Article

A proposed gender-based approach for diagnosis of the coronary artery disease

Mofreh A. Hogo[1]

Abstract

The in-time and correct diagnosis of coronary artery disease (CAD) assists cardiologists to take the right decisions. The patient's gender affects the diagnosis, treatment process, and recovery program. The patient's gender affects the structure and performance of the CAD diagnosis system. This paper studies the patient's gender effects on the CAD diagnosis model structure and performance. The work in the paper built two separate and individual models: male and female. The feature set of each model was selected using the features ranking voting (FRV) Algorithm. The memberships of the selected features for each model were computed using the probabilistic clustering technique. We built 38 different classifiers for each model to select the best one with high performance and a simple structure. The results of each selected diagnosis model of each gender were analyzed and compared with related works. The comparison shows that the proposed approach outperforms current models and with a simple structure. The accuracy of the male diagnosis model was 95%, with a sensitivity of 96% and specificity of 100%. The accuracy of the female diagnosis model was 96%, with a sensitivity of 97% and specificity of 96%. The high-performance results prove the success of the proposed gender-based approach for the diagnosis of coronary artery disease.

Keywords CAD · Classification via clustering · Accuracy · Sensitivity · Specificity · Diagnosis model

1 Introduction

Coronary artery disease (CAD) is one of the most critical diseases; it can cause severe heart attacks in patients. Patients and cardiologists should be aware of preventing the occurrence of sudden life-threatening events. A huge quantity of clinical data should be managed so that the cardiovascular risk could be estimated more reliably to stratify patients and foster early intervention [1, 2]. According to [3] CAD occurs when atherosclerotic plaque (hardening of the arteries) builds up in the wall of the arteries that supply the heart; this plaque is primarily made of cholesterol. Plaque accumulation can be accelerated by smoking, high blood pressure, and a high level of cholesterol percentage compared with the normal level and diabetes.

Patients are also at higher risk for plaque development if they are older (greater than 45 years for men and 55 years for women), or if they have a positive family history for early heart artery disease. When a blood clot forms on top of this plaque, the artery becomes blocked, causing a heart attack [3]. The total direct and indirect cost of cardiovascular disease (CVD) and stroke in the USA in 2009 is estimated to be $312.6 billion, compared with $228 billion in 2008 for all cancer and benign neoplasms according to the American Heart Association [4]. CVD costs more than any other diagnostic group. It is important to detect the cardiovascular symptoms precisely because when heart problems are detected, their underlying cause (atherosclerosis) is usually quite advanced and had been progressing for decades.

✉ Mofreh A. Hogo, mofreh_hogo@hotmail.com | [1]Electrical Engineering Department, Faculty of Engineering, Benhā University, Benha, Egypt.

One of the possible solutions to minimize heart problems is to increase the people aware of their respective CAD risks in advance and guide them to take preventive actions accordingly [5]. As in many complex medical problems, early and accurate detection could lead to better decision making.

The gender of the CAD patient is an important issue to consider when building a CAD diagnosis model. Building separate CAD diagnosis models for male and female speeds up detection, decision making, and treatment program. The gender-based analysis affects the development process of the diagnosis model in all of its stages [6].

This paper studies the effects of the patient gender on building the CAD diagnosis model; that leads to speed up the discovery and diagnosis and support the surgeons to make the right decision at the correct time.

The proposed CAD approach aims at the development of CAD diagnosis model achieved with the following characteristics:

A small and simple model with a minimum number of features and high performance around 97% and faster warning based on fewer measurements so there is no need to wait for all measurements. The homogeneity of the dataset in each model when building two separate models one for male gender and another independent separate one for female gender; the training model for each gender will be simple due to homogeneity and small due to the small feature set; therefore, the testing will be faster. The time for learning and testing the model was less than 0.015 s for each model.

The rest of this paper is organized as follows: Sect. 2 reviews the literature. Section 3 introduces the proposed approach. Section 4 describes the experimental design. Section 5 is reserved for the analysis and discussion of the obtained results. Finally, Sect. 6 concludes the paper and proposes future extensions.

2 Literature review

The importance of developing CAD diagnosis systems and CAD decision support systems is introduced and implemented in many research works. CAD diagnosis and decision support systems, designed to assist physicians and other healthcare professionals with decision-making tasks, are called clinical decision support systems (CDSS). These systems mainly link two main components: knowledge and experts. The knowledge component of these systems may come from the experts (via expert systems) and/or it can be extracted from historical data sources via data mining. Data mining offers tools, considering decision making for a particular individual patient rather than a population of patients, resulting in the emphasis on finding out

potentially useful, valid, novel and easily comprehended knowledge from data [7]. According to [3] CAD diagnosis methods include the recording of the electrical activity of the heart, as well as nuclear imaging of the blood flow to different regions of the heart. Using an external camera and ultrasound imaging of the heart muscle with exercise stress testing (stress echocardiography) is also a very accurate technique to detect CAD. The American Heart Association has identified several risk factors; some of them can be modified, treated or controlled, and some cannot; the more risk factors you have, the greater your chance of developing coronary heart disease [8].

Several studies have applied several techniques in collecting datasets from patients while using different data mining algorithms to develop data mining models. The work presented in [9] used a feature creation method to enrich the dataset and applied the information gain and confidence to determine the effectiveness of features on CAD [9]. It reported that chest pain, region RWMA2, and age were the most effective ones with high accuracy. Particle swarm optimization (PSO)-based fuzzy expert system for the diagnosis of coronary artery disease was proposed in [10]. It applied a decision tree (DT) classifier and converted its output into crisp if–then rules and transformed it into the fuzzy rule base. PSO was employed to tune the fuzzy membership functions (MFs). The work presented in [10] introduced a particle swarm optimization (PSO)-based fuzzy expert system to diagnose CAD. It applied a decision tree (DT) classifier and converted its output into crisp if–then rules and transformed it into the fuzzy rule base. PSO was employed to tune the fuzzy membership functions (MFs). A framework for intelligent medical diagnosis using the rough sets with formal concept analysis fuzzy set was developed in [11]. A fuzzy expert system for coronary heart disease diagnosis in Jordan was developed in [12]. In 2013, a newly provided dataset for heart disease, named as Z-Alizadehsani dataset, was published [13]. The collected databank comprises the information of 303 patients. This databank has 55 independent parameters and classifies a person into a normal or CAD class. They utilized several machine learning methods such as SMO, ANN, NB and bagging algorithms to study the Z-Alizadehsani dataset; moreover, they used feature creation and feature selection methods to evaluate the obtained results and concluded that the performances can be improved with utilizing the feature creation and feature selection techniques.

A computer-aided decision making for heart disease detection using a hybrid neural network–genetic algorithm was implemented in [14]. Using decision tree algorithms and data mining approaches for proving that HS-CRP is strongly associated with coronary heart disease (CHD) was implemented in [15]. The use of computational intelligence methods to detect CAD was introduced in

[16]. The benchmarking of feature selection techniques for coronary artery CAD diagnosis was proposed in [17]. A cardiovascular risk prediction method based on text analysis and data mining ensemble system was developed in [18]. A CAD diagnosis model using supervised fuzzy c-means with differential search algorithm was presented in [19]. A fuzzy rule generation for diagnosis of coronary heart disease risk using a subtractive clustering method was presented in [20]. The use of fuzzy c-means clustering employed for predicting heart disease symptoms was presented in [21]. The use of fuzzy classification for obtaining an enhanced risk prediction system for cardiovascular disease in India was introduced in [22]. Data mining with decision trees for the assessment of the risk factors of coronary heart events was developed in [23].

Alizadehsani et al. [24] divided the dataset into the training 90% and test 10% datasets. They used the information gain and SVM methods for feature analysis and feature selection. They employed the SVM methodology in combination with RBF, sigmoid, linear, and polynomial for building classifier.

A model of an association rule discovery with fuzzy decreasing support on syndrome differentiation and medication in coronary heart disease was presented in [24]. The use of cost-sensitive algorithms for a diagnosis of coronary artery disease was implemented in [25], and a comparative study of medical data classification methods based on decision tree and bagging algorithms was presented. According to medical experts, early detection may prevent death due to CAD if the proper medication is given thereafter as presented in [26]. Considering the gender difference of the patient when analyzing and diagnosing the patients' data and the behavior of the CAD and its progress and causes affects the diagnosis and treatment program, as stated in [6]. Gender differences in patients include biological, environmental, behavioral and psychological risk factors. A methodology for the automatic detection of normal and coronary artery disease conditions using heart rate signals is presented in [27] using principal component analysis (PCA) and different analysis algorithms. The work presented in [28] selected the classifier with high performance among different types of classifiers such as logistic regression (LR), classification and regression tree (CART), Multilayer Perceptron (MLP), radial basis function (RBF) and self-organizing feature maps (SOFM); eight predefined attributes were used (age, sex, family history of CAD, smoking status, diabetes mellitus, systemic hypertension, hypercholesterolemia, and body mass index (BMI)), and the major drawback of this work is its low performance. The work presented in [29] applied traditional machine learning algorithms with three types of SVM, the performance enhancement based on data preprocessing of attributes normalization and utilized the genetic algorithm and particle swarm optimization of classifier parameters selection and also introduced a new genetic training, which provided the accuracy of 93.08%.

Ghiasi et al. [30] introduced a tree-based classifier for building a CAD diagnosis approach employing the Z-Alizadehsani dataset. Another work utilized the ANN and GA techniques; Arabasadi et al. [15] classified the Z-Alizadehsani dataset in 2017. The obtained performance was represented by magnitudes of sensitivity, specificity, and accuracy for the GA-ANN which were 97%, 92%, and 93.85%, respectively. On the other hand, the ANN model classified the dataset with 84.62% accuracy, 86% sensitivity, and 83% specificity. The classification capability of the hybrid model of GA-ANN was considerably higher than the developed ANN model. The building of CAD classification system using Naïve Bayes, SVM, SMO, C4.5, and KNN techniques' accuracy levels was also evaluated in a work introduced by Alizadehsani et al. [31]. Alizadehsani et al. [32] proposed a data mining model for CAD classification to detect left circumflex, left anterior descending, and right coronary artery, which improved the classification accuracy. Acharya et al. [33] introduced a comparative study of the performance of three techniques named as discrete wavelet transform (DWT), empirical mode decomposition (EMD), and discrete cosine transform (DCT) in the detection of CAD. Alkeshoush et al. [34] studied the importance of the particle swarm optimization (PSO) algorithm in the diagnosis of heart disease. Steele et al. [35] discovered that machine learning techniques outperform in electronic health records than conventional survival methods for predicting patient mortality in CAD. Johnson et al. [36] employed machine learning techniques for the scoring of CAD characteristics on coronary CT angiograms. In 2019, Alizadehsani et al. [37] conducted a review of machine learning techniques for CAD prediction.

3 Proposed approach

The work in this paper studies and presents the effect of separating the dataset of the CAD patients into two independent datasets for male and female patients and builds two separate diagnosis models for CAD patients based on their gender. The workflow of the proposed approach is comprised of four phases as shown in Fig. 1.

The first phase comprises data segmentation into male and female datasets (segments). The second phase includes the preprocessing for missed values, features discretization, binarization, and features selection processes to extract the most important features in each dataset (using the features ranks voting named FRV [38]. The third phase is the building of 38 different classifiers from seven different classifier categories for female and male datasets
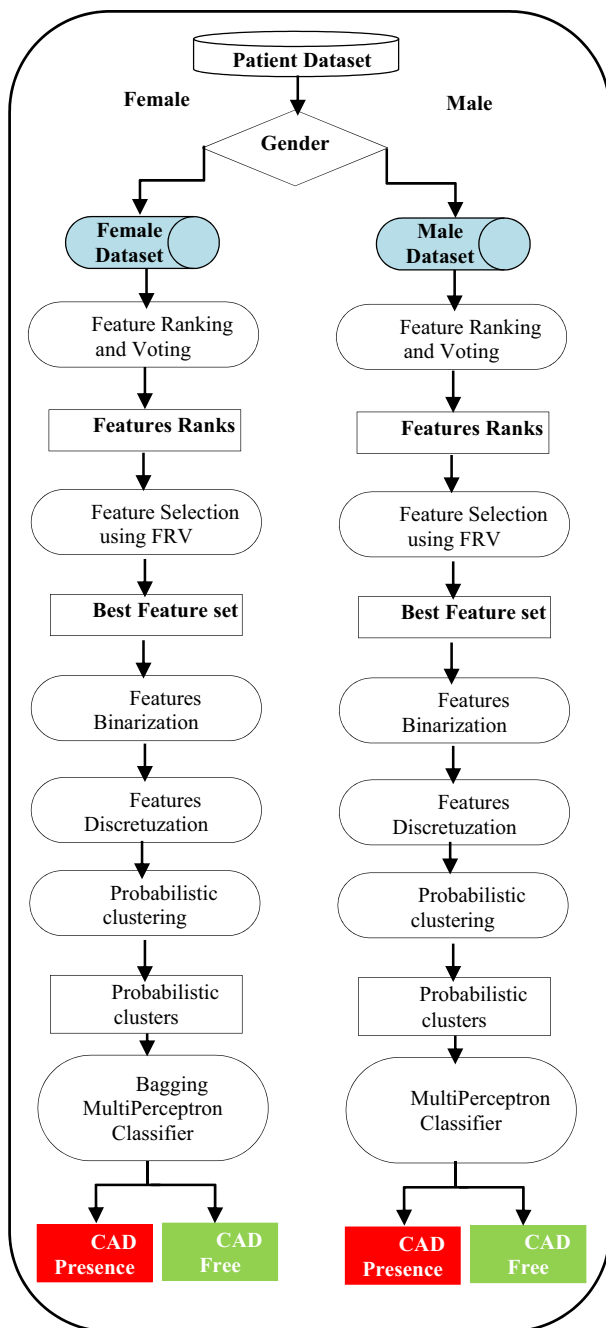
Fig. 1 Overall structure of the proposed approach for building separate diagnosis models for male and female

separately to select the most suitable classifier model for each gender. The last phase applies the classification via clustering technique to increase accuracy, sensitivity, and specificity. The overall structure of the proposed approach of building gender-based diagnosis models for males and females is illustrated in Fig. 1.

More than 50 different data mining methods were applied in implementing the proposed diagnosis system.

The following steps represent the different phases and the structure of the two proposed diagnosis models:

1. The preprocessing methods which solve the problem of missed values, and feature evaluation algorithms (CfsSubsetEval, GainRatioAttributeEval, CorrelationAttributeEval, GainRatioAttributeEval, InfoGain-AttributeEval, OneRAttributeEval, PrincipalComponents, SymmetricalUncertAttributeEval),

2. The features selection algorithm FRV (Feature Rank voting Algorithm), features binarization and discretization processes (NominalToBinary Conversion, attributes discretization, ClusterMembership methods)

3. Classification methods which are grouped into seven groups as:

   - Bayes Net-based classifiers,
   - Functions-based classifiers,
   - Misc,
   - Lazy includes,
   - Meta-based classifiers (optimization techniques).
   - Rule-based classifiers,
   - Tree-based classifiers.

4. The classification via clustering to increase the accuracy of the diagnosis model.

Some stages of the above preprocessing methods such as the feature evaluation methods and classifiers used the Weka software which is implemented [39] and developed in software programs.

4 Experimental design

4.1 Dataset description and preprocessing

The proposed approach in this paper utilized the dataset in [40] with its description presented on the UCI machine learning repository and named as "Heart Disease dataset and Z-Alizadehsani dataset." This dataset is comprised of 270 patients' records, each with 75 attributes.

This dataset is divided into two segments based on the gender of the patient as male data segment (183 records with 74 attributes instead of 75 by excluding the gender feature) and female data segment (87 instances with 74 attributes instead of 75 by excluding the gender feature). The Z-Alizadehsani dataset is comprised of 303 CAD patients' records, each with 56 attributes, categorized into four categories; demographic features set (seven attributes), symptoms and examination feature set (16 attributes), ECG features set (14 attributes), and laboratory feature set (18 attributes), as shown in Tables 1, 2, 3, and 4,

Table 1 Demographic features

ECG features	Range
Rhythm	Sin/AF
Q wave	Yes/No
ST elevation	Yes/No
ST depression	Yes/No
T inversion	Yes/No
LVH (left ventricular hypertrophy)	Yes/No
Poor R progression (poor R wave progression)	Yes/no

Table 2 Symptoms and examination features

Demographic features	Range
Age	30–86
Weight	48–120
Sex	Male, Female
Body mass index (Kg/M2)	18–41
Diabetes mellitus	Yes/No
Hypertension	Yes/No
Current smoker	Yes/No
Ex-smoker	Yes/No
Family history	Yes/No
Obesity	Yes, if MBI > 25, otherwise no
Chronic renal failure	Yes/No
Cerebrovascular accident	Yes/No
Airway disease	Yes/No
Thyroid disease	Yes/No
Congestive heart failure	Yes/No
Dyslipidemia	Yes/no

Table 3 ECG features

Symptoms and examination features	Range
BP (blood pressure)	90–190
PR (pulse rate)	50–110
Edema	Yes/no
Weak peripheral pulses	Yes/no
Lung rales	Yes/no
Systolic murmur	Yes/no
Diastolic murmur	Yes/no
Typical CP (typical chest pain)	Yes/no
Dyspnea	Yes/no
Function class	1/2/3/4
Atypical CP	Yes/No
Non-anginal CP	Yes/no
Exertional CP (exertional chest pain)	Yes/no
LowthAng (low-threshold angina)	Yes/no

Table 4 Laboratory and echo features

Laboratory features	Range
FBS (fasting blood sugar)	62–400
CR (creatine)	0.5–2.2
TG (triglyceride):	37–1050
LDL (low-density lipoprotein)	18–232
HDL (high-density lipoprotein)	15–111
BUN (blood urea nitrogen)	6–52
ESR (erythrocyte sedimentation rate)	1–90
HB (hemoglobin)	8.9–17.6
K (potassium)	3.0–6.6
Na (sodium)	128–156
WBC (white blood cell)	3700–18,000
Lymph (lymphocyte)	7–60
Neut (neutrophil)	32–89
PLT (platelet)	25–742
EF (ejection fraction)	15–60
Region with RWMA (regional wall motion abnormality)	{0,1,2,3,4}
VHD (valvular heart disease)	Normal/mild/moderate/severe

respectively. The total number of attributes is 56 including the class attributes.

The first step in the preprocessing phase is the dataset segmentation into two segments based on the gender of the patient as male data segment (176 records with 55 attributes) and female data segment (127 instances with 55 attributes). The second step in the preprocessing phase is the elimination of the effect of the missed values.

4.2 The features selection

The feature selection process affects the structure and the performance of the data mining models of medical diagnosis systems. Selecting the most important set of attributes and removing noisy or redundant attributes enhance the system performance, simplify the data mining model structure and reduce errors. We applied nine different feature selection algorithms to select the best set of features as shown in Table 7. The outputs of these nine algorithms will be the inputs of the FRV. Features selection/ranking algorithms were implemented in the following two steps.

4.2.1 Using five features ranking algorithms

The first set of experiments was carried out to test the performance of the FRV algorithm for each data segment (male and female) using five features ranking algorithms. As shown in Table 5A–C, each feature ranking algorithm produces different ranks for each feature in the data

Table 5 Attributes ranking using five attributes

Order	Gain ratio	Corr	InfoGain	OneR	Symm
(A) *Attribute ranking using total dataset*					
1	f25	f22	f25	f18	f25
2	f29	f25	f28	f53	f28
3	f28	f28	f54	f52	f54
4	f54	f1	f1	f50	f1
5	f32	f45	f53	f44	f53
6	f33	f7	f7	f3	f29
7	f53	f29	f6	f43	f7
8	f1	f6	f18	f49	f18
9	f37	f18	f29	f51	f6
10	f24	f35	f35	f1	f35
11	f12	f53	f39	f45	f47
12	f21	f3	f45	f2	f32
13	f18	f9	f55	f39	f39
(B) *Attribute ranking for male segment*					
1	f24	f28	f24	f28	f24
2	f53	f24	f1	f54	f53
3	f1	f53	f53	f45	f1
4	f27	f1	f27	f49	f27
5	f34	f27	f34	f9	f28
6	f52	f26	f28	f26	f34
7	f44	f32	f44	f25	f52
8	f28	f31	f6	f53	f44
9	f5	f34	f5	f22	f5
10	f6	f52	f52	f29	f31
11	f54	f44	f33	f30	f26
12	f33	f36	f46	f23	f32
13	f23	f11	f17	f21	f6
(C) *Attribute ranking for the female segment*					
1	f24	f24	f24	f24	f24
2	f27	f27	f27	f27	f27
3	f38	f6	f17	f17	f38
4	f17	f1	f38	f6	f17
5	f40	f40	f53	f38	f40
6	f6	f17	f28	f40	f6
7	f1	f5	f1	f4	f1
8	f53	f52	f40	f28	f53
9	f5	f38	f6	f37	f5
10	f52	f53	f7	f23	f52
11	f28	f28	f36	f54	f28
12	f54	f51	f31	f1	f54
13	f37	f18	f52	f49	f37

Table 6 Ranking order for each segment

Ranks order	Total data	Male dataset	Female dataset
1	f25	f24	f24
2	f28	f1	f27
3	f54	f53	f17
4	f1	f27	f38
5	f53	f34	f53
6	f7	f28	f28
7	f6	f44	f1
8	f18	f6	f40
9	f29	f5	f6
10	f35	f52	f7
11	f39	f33	f36
12	f45	f46	f31
13	f55	f17	f52

which interpret the common and different features' ranking levels as follows:

- The common features with the same ranking levels are f24, f28 at rank order 1 & 6, which are different completely for the total dataset (second column in Table 6) compared with male dataset in the third column and the female dataset in the fourth column.
- The common features with different ranking levels are f1, f53, f27, f6, f52, f17.
- The selected attributes for males only are f34, f44, f5, f33, f46.
- The selected features for females only are f40, f38, f7, f36, f31.

The rest of the results shown in Tables 7 and 8 can be interpreted in the same way. According to these results, we can conclude that the separation of female and male diagnosis models is a crucial and mandatory request.

a. The partitioning of the dataset into female and male data segments divides the problem into two small subproblems and removes the attribute named sex, thus reducing the number of attributes from 56 to 55.

b. An important question is "what is the optimal number of attributes to be selected in the CAD diagnosis system?". Several research works had agreed that 14 attributes are fair enough for CAD. Fourteen or fewer attributes are considered in this work.

c. The experiments through this section are conducted using:

- The total dataset (303 instances and 56 attributes),
- The female data segment (127 instances and 55 attributes), and

segment, which means that the behavior of each ranking algorithm is affected by the gender of the patient.

The different features ranking levels obtained from the InfoGain algorithm for the total data, the male dataset, and the female dataset are presented in Table 6 and represented graphically by the Venn diagram shown in Fig. 2,

Fig. 2 Behavior of information gain ranking algorithm for female and male segments
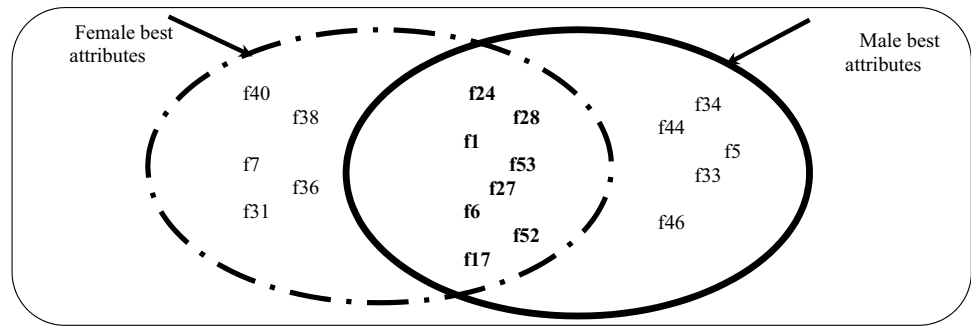
Table 7 Performance comparison for the nine different feature selection algorithms using different datasets

CfsSubset	GainRatio attribute	Correlation attribute	InfoGain attribute	OneR attribute	PCA	ReliefF attribute	Symmetrical uncert attribute
Acc = 84.5%	Acc = 83.8%	Acc = 83.8%	Acc = 83.8%	Acc = 83.8%	Acc = 83.8%	Acc = 83.8%	Acc = 85.8%
Age	Typical chest pain	Typical chest pain	Typical chest pain	Typical chest pain	Typical chest pain	Typical chest pain	Typical chest pain
DM	Nonanginal atypical region RWMA	Atypical	Atypical	Atypical	Atypical	Atypical	Atypical
HTN		Age	Age	Age	Age	Age	Region RWMA
BP		Region RWMA	Region RWMA	Region RWMA	Region RWMA	Region RWMA	Age
Typical chest pain	Q wave	HTN	HTN	HTN	HTN	HTN	EF-TTE
Atypical nonanginal	ST elevation	Nonanginal	Nonanginal	Nonanginal	Nonanginal	Nonanginal	Nonanginal
Q wave	EF-TTE	DM	DM	DM	DM	DM	HTN
T inversion	Age	BP	BP	BP	BP	BP	BP
ESR	Poor R progression	T inversion	T inversion	T inversion	T inversion	T inversion	DM
EF-TTEK	Diastolic murmur	EF-TTE	EF-TTE	EF-TTE	EF-TTE	EF-TTE	T inversion
Region RWMA	CRF	FBS	FBS	FBS	FBS	FBS	K
Cath	Weak peripheral pulse	K	K	K	K	K	Q wave
	BP	ESR	ESR	ESR	ESR	ESR	FBS
	Cath	Cath	Cath	Cath	Cath	Cath	Cath

Table 8 Performance comparison using FRV and CfsSubsetEval feature selection algorithms

Selected features for total dataset using		Selected features for male dataset using CfsSubsetEval	Selected features for male dataset using FRV	Selected features for female dataset using CfsSubsetEval	Selected features for female dataset using FRV
CfsSubsetEval	FRV				
Acc = 84.5%	Acc = 87.13	Acc = 83.5	Acc = 88.63	Acc = 86.6%	Acc = 90.55%
Age	Age	Age	Age	Age	Age
DM	BMI	DM	HTN	DM	BMI
HTN	DM	HTN	FH	HTN	HTN
BP	BP	Typical chest pain	Obesity	Current smoker	Current smoker
Typical chest pain	Weak peripheral pulse	Function class	Typical chest pain	BP	BP
Atypical	Typical chest pain	Nonanginal	Atypical	Diastolic murmur	Typical chest pain
Nonanginal	Atypical	T inversion	Nonanginal	Typical chest pain	Atypical
Q wave	Nonanginal	ESR	ST elevation	Atypical	FBS
T inversion	ST elevation	EF-TTE	T inversion	FBS	TG
ESR	T inversion	Region RWMA	HDL	TG	EF-TTE
EF-TTEK	EF-TTE	Cath	EF-TTE	EF-TTE	Region RWMA
Region RWMA	Region RWMA		Region RWMA	Region RWMA	VHD
Cath	VHD		VHD	Cath	Cath
	Cath		Cath		

- The male data segment (176 instances and 55 attributes).

4.2.2 Using nine features ranking algorithms

The second set of experiments in the feature selection process used nine different features ranking algorithms and was applied on the total dataset before the segmentation into male and female partitions to select the best set of features. The goal of this step to illustrate the importance of building separate diagnosis models based on gender.

The obtained results shown in Table 7 can be interpreted as follows: Each column represents the results obtained by applying one of the feature selection algorithms to the total dataset (before partitioning). For example, the first column represents the 14 features with the highest ranks obtained using CfsSubsetEval algorithms (age, DM, HTN, BP, typical chest pain, atypical, nonanginal, Q wave, T inversion, ESR, EF-TTEK, region RWMA, Cath accuracy). The accuracy of the classifier using this set was 84.5%. The highest performance was 85.8%, obtained using the Symmetrical UncertAttributeEval algorithm, which is considered a very low performance. Table 8 shows the results of applying the FRV algorithm on the total dataset, male dataset, and female dataset. The 14 features with the highest ranks are fed to the backpropagation neural network classifier, and the different performances were measured. The obtained results in Table 8 can be interpreted as follows: Each column represents the results obtained by one of the feature selection algorithms using a specific dataset, and the following examples explain the obtained results:

a. The first column represents the results obtained from CfsSubsetEval algorithm using the total dataset, the accuracy was 84.5%, and the best features were age, DM, HTN, BP, typical chest pain, atypical, nonanginal, Q wave, T inversion, ESR, EF-TTEK, region RWMA, and Cath).
b. The best performance for the female segment was 90.55%, and the number of features was 13 attributes using the FRV algorithm.

The best performance for the male segment was 88.3%, and the number of features was 14 attributes using the FRV algorithm. The obtained results of the female and male segments ensure that the segmentation is necessary. We will proceed with building the CAD model based on the concept of gender difference, and the FRV algorithm provides the best results for both male and female data segments.

4.3 Classifier selection step and results interpretation

This section presents selecting the best classifier for each dataset. Over 38 classifiers were utilized to select the classifier (s) with the highest performance for the female and male models. These classifiers are categorized into seven categories as follows: Bayes (four classifiers), function (four classifiers), lazy (three classifiers), meta (15 classifiers), misc (one classifier), rules (five classifiers), and trees (seven classifiers). Selection of the best model for each gender is carried out using different experiments and is labeled as A, B, C, D, and E.

A. *Male Model (176 instances) with different sets of features as:*

   a. Using the total 55 attributes without applying features selection algorithms.
   b. Using the selected 14 attributes by the FRV algorithm as obtained in Table 8 (age, HTN, FH, obesity, typical chest pain, atypical, nonanginal, ST elevation, T inversion, HDL, EF-TTE, region RWMA, VHD, and Cath).
   c. Using the selected 14 attributes from applying the CfsSubsetEval, as obtained in Table 8 (age, DM, HTN, typical chest pain, function class, nonanginal, T inversion, ESR, EF-TTE, region RWMA, and Cath).

The results of these experiments are shown in Table 9 and can be interpreted as follows: Using the different 38 classifiers for the 55 attributes, the highest performance was 85.8% achieved by the Multilayer Perceptron classifier. Using the selected 14 attributes obtained by FRV, the highest performance was 88.6% achieved by Multilayer Perceptron. Using the selected 14 attributes obtained by CfsSubsetEval, the highest performance was 85.8% achieved by Meta Classifier Attribute Select.

B. *Female model (127 instances) with different sets of features as:*

   d. Using the total 55 attributes.
   e. Using the selected 14 attributes from applying CfsSubsetEval, as obtained in Table 8 (age, DM, HTN, current smoker, BP, diastolic murmur, typical chest pain, atypical, FBS, TG, EF-TTE, region RWMA, and Cath).
   f. Using the selected 13 attributes by the FRV algorithm, as obtained in Table 8 (age, BMI, HTN, current smoker, BP, typical chest pain, atypical, FBS, TG, EF-TTE, region RWMA, VHD, and Cath).

Table 9 Performance of 38 different classifiers using male, female, and total datasets

	A: Male segment			B: Female segment			C: Total	D:Average	E: Average
	1	2	3	4	5	6	7	8	9
	Total of 55	Selected 14 by FRV	Selected 14 by Cfs	Total of 55	Selected 14 by FRV	Selected 14 by Cfs	Male and female using 56	Average male and female using FRV	Average male and female using 14 Cfs
Bayes									
BayesNet	82.4	84.7	83	89.8	91.0	83.4	81.3	87.8	83.23
Naïve Bayes	79.0	85.8	83	85.0	87.4	88.2	79.5	86.6	85.6
Naïve Bayes multinomial text	73.9	73.9	73.7	67.7	67.7	67.7	73.9	70.8	70.7
Naïve Bayes updatable	79.0	85.8	83	85.0	87.4	88.1	79.5	86.6	85.55
Function									
Multilayer perceptron	85.8	88.6	83.6	82.7	90.5	86.4	82.8	89.5	84
SGD Text	73.9	73.9	73.9	67.7	67.7	67.7	71.3	70.8	70.8
SMO	82.4	86.9	85.2	83.5	90.6	84.2	88.1	88.7	84.7
Voted perceptron	73.9	75.0	73.7	67.7	69.2	67	71.3	72.1	70.35
Lazy									
IBk	79.5	78.4	81.2	77.2	82.7	80.3	78.5	80.5	80.77
KStar	69.9	84.1	75	64.6	82.7	83.4	71.3	83.4	79.2
LWL	76.7	80.7	82.4	79.5	79.5	79.5	76.2	80.1	80.95
Meta									
AdaBoostM1	81.3	83.5	80.7	81.9	88.2	88.9	82.5	85.9	84.83
Attribute select	85.2	85.8	85.8	81.9	85.0	89	81.5	85.4	87.4
Bagging	84.7	88.1	83.5	83.5	86.6	81.2	85.1	87.3	82.35
CV parameter selection	73.9	78.3	73.9	67.7	67.7	67.7	71.3	73.0	70.8
Filtered classifier	80.1	85.2	81.5	80.3	81.1	78.7	85.1	83.2	80
Iterative classifier optimizer	80.1	84.1	84.1	81.1	84.3	86.6	85.1	84.2	85.35
LogitBoost	81.8	84.7	83	85.0	85.0	85	86.1	84.8	84
Multischeme	73.9	73.9	73.9	67.7	67.7	67.7	71.3	70.8	70.8
Random committee	82.4	86.9	87.5	85.7	87.4	85.8	82.8	87.2	86.65
Randomizable filtered classifier	66.5	81.3	75.5	82.7	74.0	70	59.4	77.6	72.75
Random Subspace	77.8	84.7	84.1	85.0	87.4	83.4	82.5	86.0	83.75
Stacking	73.9	73.9	67.7	67.7	67.7	67.7	71.3	70.8	67.7
Vote	73.9	73.9	73.9	67.7	67.7	67.7	71.3	70.8	70.8
Weighted instances handler wrapper	73.9	73.9	73.9	67.7	67.7	67.7	71.3	70.8	70.8
Misc									
Input-mapped classifier	73.9	73.9	73.9	67.7	67.7	67.7	71.3	70.8	70.8
Rules									
Decision table	75.7	79.0	84.1	75.6	79.5	80.3	81.5	79.3	82.2
JRip	79.5	81.3	80.7	81.9	86.6	87.4	79.9	83.9	84.05
OneR	71.0	71.6	70.4	70.9	70.9	74	70.6	71.2	72.22

Table 9 (continued)

	A: Male segment			B: Female segment			C: Total	D:Average	E: Average
	1	2	3	4	5	6	7	8	9
	Total of 55	Selected 14 by FRV	Selected 14 by Cfs	Total of 55	Selected 14 by FRV	Selected 14 by Cfs	Male and female using 56	Average male and female using FRV	Average male and female using 14 Cfs
PART	76.1	83.0	83.5	79.5	82.7	87.4	79.5	82.8	85.45
ZeroR	73.9	73.9	73.9	67.7	67.7	67.7	71.3	70.8	70.8
Trees									
Decision stump	66.5	68.2	68.2	79.5	79.5	79.5	76.2	73.9	73.85
Hoeffding Tree	72.7	73.3	72.7	63.8	81.1	82.7	71.3	77.2	77.7
J48	81.3	84.7	83	78.7	85.8	86.6	78.9	85.2	84.8
LMT	85.2	88.1	72.7	82.7	89.8	83.4	86.1	88.9	78.05
Random Forest	84.1	88.1	87	82.7	88.2	88.9	86.1	88.1	87.95
Random Tree	79.5	83.0	77.8	71.7	88.2	77.1	77.9	85.6	77.45
REP Tree	83.0	87.5	82.9	77.2	82.7	82.7	82.5	85.1	82.8

The results of Table 9 show that using the 38 different classifiers for the total 55 attributes, the highest performance was 89.8% achieved by the BayesNet classifier. Using the selected 14 attributes obtained by FRV, the highest performance was 91% achieved by BayesNet. Using the selected 14 attributes obtained by CfsSubsetEval, the highest performance was 89% achieved by Meta Classifier Attribute Select.

C. *Model for male and female using all datasets without partition*

This experiment used all datasets, including both female and male instances before segmentation (303 instances with 56 attributes). The results of these experiments are presented in Table 9; for the 38 different classifiers for the total data (303 instances) with total attributes (56 Attributes), the highest performance was 88.1% achieved by the SMO classifier.

D. *Average performance for female model and male model using CfsSubsetEval*

This step aims at finding the classifier which gives the best accuracy for both male and female segments using the 14 selected attributes by CfsSubsetEval. Table 9 shows that the highest average performance was 86.65% which was achieved by the Meta Random Committee Classifier.

E. *Average performance for female model and male model using FRV*

This step aims at finding out the classifier which gives the best accuracy for both male and female segments

using the best selected 14 attributes by FRV. According to Table 9, the highest average performance was 89.5% which was achieved by the Multilayer Perceptron classifier A.

From these sets of experiments (A through E) shown in Table 9, the highest performance was obtained using the features selected by the FRV for all datasets, and male and female segments. The results obtained by the female segment and male segment were better than those obtained by the total dataset. Accordingly, we selected the following classifiers with the highest performance for the next experiments: Naïve Bayes, Naïve Bayes Updatable, Multilayer Perceptron, SMO, Bagging Multilayer, and Bagging SMO, LMT Tree, Random Forest Tree, REP Tree.

4.4 Classification via clustering

This set of experiments aims at improving the performance of the classifiers through the clustering of the segment dataset into a set of probabilistic clusters using the EM probabilistic clustering algorithm implemented in open-source software.

The first stage is the preprocessing stage including the following steps:

a. Converting nominal attributes to binary attributes.
b. Attributes discretization.
c. Membership clustering: find membership values of each attribute to a set of predefined clusters using the EM clustering on the datasets to cluster them into two main clusters and different subclusters in each cluster.

Converting the nominal attributes to binary ones, except the class attribute, is carried out using the

Table 10 VHD attribute before binarization

No.	Label	Count	Weight
1	Mild	84	84
2	Moderate	21	21
3	N	64	64
4	Severe	7	7

Table 11 VHD attribute after binarization

No	Label	Min	Max	Medium	Std deviation
1	Mild	0	1	0.512	0.502
2	Moderate	0	1	0213	0.047
3	N	0	1	0.409	0.494
4	Severe	0	1	0.031	0.175

Table 12 VHD = mild attribute after discretization

No.	Label	Count	Weight
1	'(-inf-0.25]'	92	92
2	'(0.25–0.5]'	0	0
3	'(0.5–0.75]'	0	0
4	'(0.75-inf)'	84	84

Table 13 VHD = mild attribute after discretization

No.	Label	Count	Weight
1	'(-inf-0.25]'	155	155
2	'(0.25–0.5]'	0	0
3	'(0.5–0.75]'	0	0
4	'(0.75-inf)'	21	21

Table 14 VHD = Severe attribute after discretization

No.	Label	Count	Weight
1	'(-inf-0.25]'	112	112.0
2	'(0.25–0.5]'	0	0
3	'(0.5–0.75]'	0	0
4	'(0.75-inf)'	64	64

Table 15 VHD = Severe attribute after discretization

No.	Label	Count	Weight
1	'(-inf-0.25]'	169	169
2	'(0.25–0.5]'	0	0
3	'(0.5–0.75]'	0	0
4	'(0.75-inf)'	7	7

unsupervised filters (named NominalToBinary), For example, the VHD attribute has four nominal values (mild, moderate, N and severe). These four values are converted to four binary attributes as shown in Tables 10 and 11. Table 10 shows the VHD attribute in the male segment; before binarization, there is one attribute, but after binarization, there are four attributes; each one has minimum, maximum, median, and standard deviation values as shown in Table 11, which increased the number of attributes in the male segment to 17 instead of 14 and in the female segment to 16 instead of 13 before binarization.

Because the attributes have scattered values, attributes discretization defines a set of intervals and groups the features according to those discretized intervals. Four intervals are specified for each attribute. For example, the binary version of attribute VHD became four attributes: VHD = mild, VHD = Moderate, VHD = N and VHD = Severe. The discretized version of VHD = mild attribute is shown in Tables 12, 13, 14, and 15. Another example shows how the numeric attributes are converted to discretized attributes; for example, the age attribute has the values as minimum = 30, maximum = 86, mean = 58.494, and standard deviation = 10.771.

The discretized age attribute is shown in Table 16, where 14 instances have the first interval values from −inf. to 44; the second interval from 44 to 58 has 80

Table 16 Age attribute after discretization

No.	Label	Count	Weight
1	'(-inf-44]'	14	14.0
2	'(44–58]'	80	80.0
3	'(58–72]'	59	59.0
4	'(72-inf)'	23	23.0

instances, the third interval has values from 58 to 72 and represents 59 instances, and the fourth interval from 72 to inf. represents 23 instances. The inf. means maximum and − inf. means the minimum value. The discretized attributes are submitted to the EM cluster algorithm (expectation–maximization algorithm for clustering multidimensional numerical data); more details can be found in [41]. The following example illustrates how it works: The input for this step will be the discretized attributes, and the outputs will be the clusters and subclusters and the membership value for each attribute in these clusters. In the female segment, we applied the ClusterMembership one time, producing two clusters as pcluster0 and pcluster1; each cluster represents one class value. If we apply the ClusterMembership again, it will partition each cluster to subclusters. pcluster0 will be portioned to three subportions as pCluster_0_0,

Table 17 Description of TP, FP, FN, TN

	Positive	Negative	Row summation
Positive	TP	FP	TP + FP
Negative	FN	TN	FN + TN
Column summation	TP + FN	FP + TN	N = TP + FP + FN + TN

pCluster_0_1, and pCluster_0_2, and pcluster1 will be portioned to four subpartitions as pCluster_1_0, pCluster_1_1, pCluster_1_2, and pCluster_1_3.

The final results for these three steps (NominalTo-Binary, Discretization, and ClusterMembership) for the female data segment will be as follows: Instances:127, Attributes: 8: (pCluster_0_0, pCluster_0_1, pCluster_0_2, pCluster_1_0, pCluster_1_1, pCluster_1_2, pCluster_1_3), and Cath. Those attributes will be the input to the classifier Multilayer Perceptron (backpropagation). In the same way, the male data segment is processed.

5 Results analysis and discussion

5.1 Performance metrics

Using tenfold cross-validation technique, the sensitivity, specificity, and accuracy evaluation metrics [9] are used for the performance analysis of male and female diagnosis models, where:

- Accuracy = (TN + TP)/(TN + TP + FN + FP)
- Sensitivity = TP/(TP + FN) (the percentage of actual positives, which are correctly identified)
- Specificity = TN/(TN + FP) (the percentage of negatives which are correctly identified,

Table 17 shows the description of these terms, where TP is true positive, TN true negative, FN false negative, and FP false positive. These terms are obtained from the classification confusion matrix in Table 17, where is positive = summation of the positive column, negative = summation of the negative column, and N = total number of instances.

Table 18 shows the calculation of the different performance metrics: for example, for Naïve Bayes classifier for male data segment with the selected 14 attributes by FRV: TP = 119, FN = 11, FP = 13, and TN = 33; then, positive is 130 (the total number of CAD instances in the male data segment); negative is 46 (the total number of normal instances in the male data segment); accuracy is 0.866; error is 0.14; sensitivity is 0.915; and specificity is 0.75.

Table 18 Performance analysis for the best seven male and female diagnosis models

Models	The male models using the selected 14 attributes										The female models using the selected 13 attributes									
	C	N	C	N	Total		Sens.	Spec	Acc	Err	C	N	C	N	Total		Sens	Spec	Acc	Err
Performance	TP	FN	FP	TN	TP+TN	FP+FN					TP	FN	FP	TN	TP+TN	FP+FN				
Naïve Bayes	119	11	13	33	152	24	0.90	0.75	0.86	0.14	82	4	5	36	118	9	0.94	0.90	0.93	0.07
Naïve Bayes Updatable	119	11	13	33	152	24	0.90	0.75	0.86	0.14	82	4	5	36	118	9	0.94	0.90	0.93	0.07
Multilayer perceptron	130	0	9	37	167	9	0.94	1.00	0.95	0.05	83	3	3	38	121	6	0.97	0.93	0.95	0.05
SMO	129	1	8	38	167	9	0.94	0.97	0.95	0.05	84	2	4	37	121	6	0.95	0.95	0.95	0.05
Bagging multilayer perc.	129	1	8	38	167	9	0.94	0.97	0.95	0.05	84	2	3	38	122	5	0.97	0.95	0.96	0.04
Meta. Bagging.SMO	129	1	9	37	166	10	0.93	0.97	0.94	0.06	84	2	4	37	121	6	0.95	0.95	0.95	0.05
Trees. LMT	130	0	10	36	166	10	0.93	1.00	0.94	0.06	84	2	4	37	121	6	0.95	0.95	0.95	0.05
Trees. Random Forest	129	1	8	38	167	9	0.94	0.97	0.95	0.05	84	2	4	37	121	6	0.95	0.95	0.95	0.05
REP Tree	127	3	15	31	158	18	0.89	0.91	0.90	0.10	82	4	4	37	119	8	0.95	0.90	0.94	0.06

5.2 Performance analysis

The Multilayer Perceptron is the best classifier for the male diagnosis model with the highest performance, where the accuracy = 0.95, the sensitivity = 0.94, specificity = 1, and the error = 0.05. For the female diagnosis model, the accuracy = 0.96, the sensitivity = 0.97, specificity = 0.95, and the error = 0.04. The obtained results for the female data segment were better than those obtained from the male one. To emphasize the role of separating the male and female diagnosis models, using all 55 attributes in developing female and male diagnosis models, Table 19 shows that the best classifier for male diagnosis model was the LMT Tree, where the accuracy = 0.88, the sensitivity = 0.90, specificity = 0.8, and the error = 0.12. For the female diagnosis model, the best classifier was the Bagging Multilayer, where the accuracy = 0.90, the sensitivity = 0.92, specificity = 0.85, and the error = 0.10. As an average performance for male and female classifiers via clustering using all the 55 attributes, the best classifier model for the male diagnosis model was achieved by the Naïve Bayes, where the accuracy = 0.88, the sensitivity = 0.92, specificity = 0.77, and the error = 0.12. Table 20 shows the performance of the male and female diagnosis models using the selected attributes by FRV and classification via clustering using nine different classifiers. For the male diagnosis model, the best classifier was the Multilayer Perceptron, where the accuracy = 0.95, the sensitivity = 0.94, specificity = 1, and the error = 0.05, compared with the accuracy = 0.89, the sensitivity = 0.94, specificity = 0.76, and the error = 0.11. The best classifier for the female diagnosis model was the Bagging Multilayer, where the accuracy = 96%, the sensitivity = 0.97, specificity = 0.95, and the error = 0.04, compared with the normal classifier performance where accuracy = 0.90, sensitivity = 0.92, specificity = 0.85, and the error = 0.10. The performance analysis for the results

obtained from the proposed approach using the two different datasets (Heart Disease dataset and Z-Alizadehsani dataset) is shown in Table 21. The results illustrate the acceptable behavior of the proposed approach which discriminates between the female and male datasets to reduce the diagnosis model and enhance its performance. The performance comparison between the classical approach and our proposed approach using total, female, and male datasets is shown in Fig. 3a–f.

6 Conclusion

The paper proposed an approach for developing diagnosis models for CAD patients, based on their gender. The development of the diagnosis models passes through the following stages: data partitioning, data preprocessing (missing values, binarization, and discretization), features ranking, and features selection algorithms. We built 38 different classifiers to select the most suitable classifier for each data segment. The application of classification via clustering has highly improved the performance of the diagnosis models. The proposed models are found to provide better accuracy data mining models with simple structures than current diagnosis models. The partitioning of the CAD model into separate male and female diagnosis models, each having its characteristics, proved the importance of developing different diagnosis models of patients based on their gender. Using the FRV features selection algorithm was found to be the best among different features selection algorithms. The best classifiers suitable for the two models were based on a Multilayer Perceptron with different functions. The results of the paper were approved by cardiologists to be implemented in the real world. Future investigations will include the development of an interactive system to simplify the diagnosis process for patients and cardiologists. It

Table 19 Performance of classifiers via clustering using 55 attributes

| Classifier	Male classifier via clustering performance using all the 55 attributes				Female classifier via clustering performance using all the 55 attributes				Average performance for male and female classifiers via clustering performance using all the 55 Attributes			
	Sens	Spec	Acc	Err	Sens	Spec	Acc	Err	Sens	Spec	Acc	Err
Naïve Bayes	0.90	0.75	0.86	0.14	0.94	0.80	0.89	0.11	0.92	0.77	0.88	0.12
Naïve Bayes Updatable	0.90	0.75	0.86	0.14	0.93	0.68	0.83	0.17	0.92	0.71	0.85	0.15
Multilayer perceptron	0.90	0.73	0.86	0.14	0.87	0.75	0.83	0.17	0.89	0.74	0.85	0.15
SMO	0.89	0.74	0.85	0.15	0.92	0.81	0.88	0.12	0.90	0.77	0.87	0.13
Bagging multilayer	0.91	0.69	0.85	0.15	0.92	0.85	0.90	0.10	0.91	0.77	0.87	0.13
Bagging SMO	0.89	0.74	0.85	0.15	0.92	0.83	0.89	0.11	0.90	0.78	0.87	0.13
LMT Tree	0.90	0.80	0.88	0.12	0.92	0.81	0.88	0.12	0.91	0.81	0.88	0.12
Random Forest Tree	0.85	0.71	0.82	0.18	0.88	0.81	0.86	0.14	0.86	0.76	0.84	0.16
REP Tree	0.88	0.62	0.81	0.19	0.91	0.74	0.85	0.15	0.90	0.68	0.83	0.17

Table 20 Performance of classifiers and the classifiers via clustering using selected attributes

Classifier	Male using selected 14				Female using selected 13				Female and male avg. selected				Male classifier via clustering using selected 14				Female classifier via clustering using selected 13				Female and male avg. classifier via clustering			
	Sens	Spec	Acc	Err	Sens	Spec	Acc	Err	Sens	Spec	Acc	Err	Sens	Spec	Acc	Err	Sens	Spec	Acc	Err	Sens	Spec	Acc	Err
Naïve Bayes	0.94	0.76	0.89	0.11	0.93	0.81	0.89	0.11	0.93	0.78	0.89	0.11	0.90	0.75	0.86	0.14	0.94	0.90	0.93	0.07	0.92	0.82	0.891	0.11
Naïve Bayes Updatable	0.89	0.74	0.86	0.14	0.93	0.81	0.89	0.11	0.91	0.78	0.87	0.13	0.90	0.75	0.86	0.14	0.94	0.90	0.93	0.07	0.92	0.82	0.891	0.11
Multilayer perceptron	0.94	0.76	0.89	0.11	0.92	0.87	0.91	0.09	0.93	0.81	0.89	0.11	0.94	1.00	0.95	0.05	0.97	0.93	0.95	0.05	0.95	0.96	0.950	0.05
SMO	0.90	0.79	0.89	0.11	0.93	0.85	0.91	0.09	0.91	0.82	0.89	0.11	0.94	0.97	0.95	0.05	0.95	0.95	0.95	0.05	0.95	0.96	0.950	0.05
Bagging Multilayer	0.92	0.77	0.88	0.13	0.93	0.90	0.92	0.08	0.93	0.84	0.90	0.10	0.94	0.97	0.95	0.05	0.97	0.95	0.96	0.04	0.95	0.96	0.954	0.05
Bagging SMO	0.88	0.81	0.87	0.13	0.92	0.87	0.91	0.09	0.90	0.84	0.88	0.12	0.93	0.97	0.94	0.06	0.95	0.95	0.95	0.05	0.94	0.96	0.947	0.05
LMT Tree	0.90	0.78	0.87	0.13	0.93	0.85	0.91	0.09	0.91	0.82	0.88	0.12	0.93	1.00	0.94	0.06	0.95	0.95	0.95	0.05	0.94	0.97	0.947	0.05
Random Forest Tree	0.86	0.85	0.82	0.18	0.90	0.86	0.89	0.11	0.88	0.75	0.87	0.13	0.94	0.97	0.95	0.05	0.95	0.95	0.95	0.05	0.95	0.96	0.950	0.05
REP Tree	0.89	0.74	0.85	0.15	0.86	0.73	0.82	0.18	0.88	0.73	0.84	0.16	0.89	0.91	0.90	0.10	0.95	0.90	0.94	0.06	0.92	0.91	0.914	0.09

Table 21 Performance analysis for the two different datasets

| Result dataset	The results using the classical approach												The results using the proposed approach											
	Male				Female				Total				Male dataset				Female dataset				Total dataset			
Classifier	Multilayer perceptron				Bagging Multilayer				Bagging Multilayer				Multilayer perceptron				Bagging Multilayer				Bagging Multilayer			
Performance metric	Sens	Spec	Acc	Err	Sens	Spec	Acc	Err	Sens	Spec	Acc	Err	Sens	Spec	Acc	Err	Sens	Spec	Acc	Err	Sens	Spec	Acc	Err
Heart Disease dataset	0.72	0.76	0.76	0.24	091	0.7	0.86	0.14	0.80	0.79	0.83	0.19	0.97	0.97	0.97	0.03	0.97	0.91	0.96	0.04	0.9	0.87	0.88	0.12
Z-Alizadehsani dataset	0.94	0.76	0.89	0.11	0.93	0.90	0.92	0.08	0.93	0.84	0.90	0.10	0.94	1.00	0.95	0.05	0.97	0.95	0.96	0.04	0.95	0.96	0.950	0.05
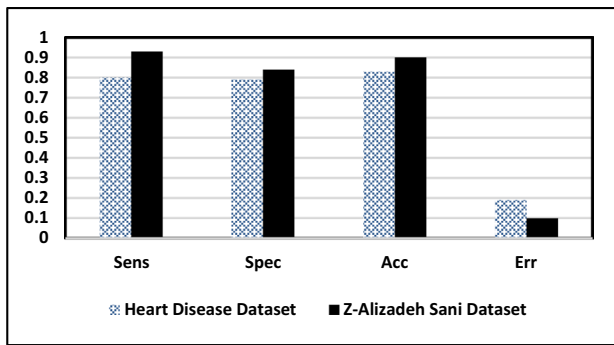
(A) Classical Approach Performance Using Total Dataset
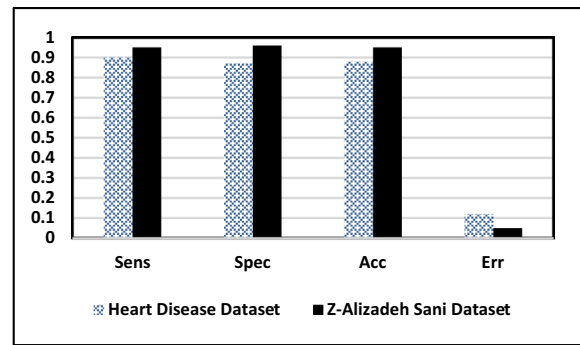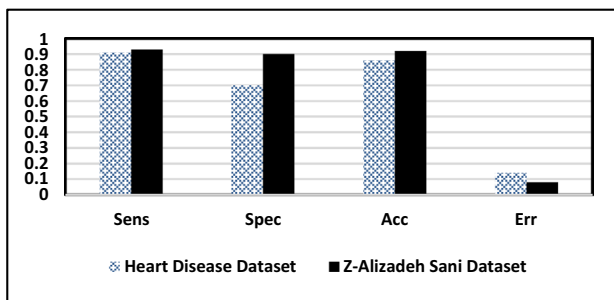

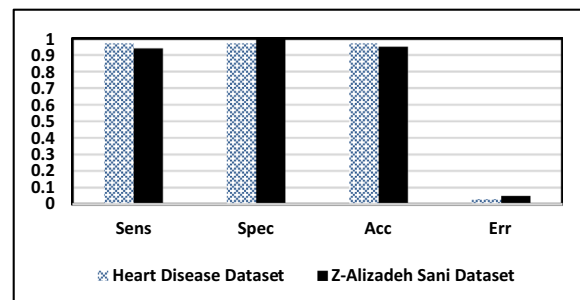**(B)** Proposed Approach Performance for Total Dataset
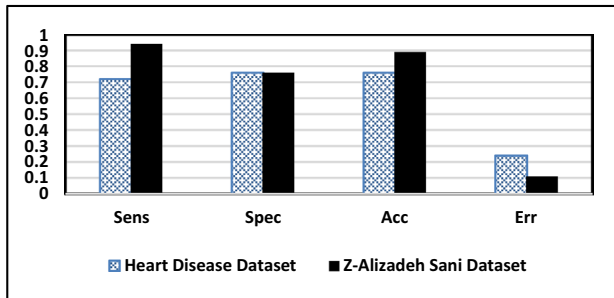
(C) Classical Approach Performance for Female Dataset
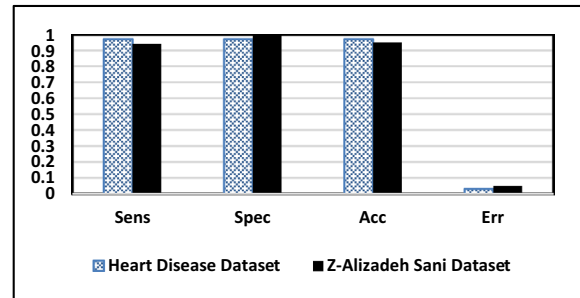
(D) Proposed Approach Performance for Female Dataset

(E) Classical Approach Performance for Male Dataset

(F) Proposed Approach Performance for Male Dataset

Fig. 3 Performance comparison between the classical approach and the proposed approach

encourages the implementation of a warning system for the patients according to their status. The proposed approach will be utilized in the governmental hospital to help the surgeons and cardiologists to take the right decisions to save the patients' life.

Compliance with ethical standards

Competing interests The author declares that he has no competing interests.

References

1. Chang C-D, Wang C-C, Jiang BC (2011) Using data mining techniques for multi-diseases prediction modeling of hypertension and hyperlipidemia by common risk factors. Expert Syst Appl 38(5):5507–5513
2. Ion Titapiccolo J, Ferrario M, Cerutti S, Barbieri C, Mari F, Gatti E, Signorini MG (2013) Artificial intelligence models to stratify cardiovascular risk in incident hemodialysis patients. Expert Syst Appl 40(11):4679–4686

3. Kulick DL, Shie WC (2007) Heart diseases from MediciNet. com. http://www.medicinenet.com/ coronary_artery_bypass_ graft/article.html

4. Benjamin E, Emelia J, Michael JB, Stephanie EC (2015) Heart disease and stroke statistics—2017 update. A Report From the American Heart Association, Circu

5. Koene RJ, Prizment AE, Blaes A, Konety SH (2016) Shared risk factors in cardiovascular disease and cancer. Circulation 133(11):1104–1114

6. Saeidi M, Komasi S, Soroush A, Zakiei A, Shakeri J (2014) Gender differences in patients' beliefs about biological, environmental, behavioral, and psychological risk factors in a cardiac rehabilitation program. Journal of Cardio-Thoracic Medicine 2(4):215–220

7. Di Noia T, Ostuni VC, Pesce F, Binetti G, Naso D, Schena FP, Di Sciascio E (2013) An end stage kidney disease predictor based on an artificial neural networks ensemble. Expert Syst Appl 40(11):4438–4445

8. AHA AHA (2013b) Understand your risk of heart attack. Retrieved 12 April 2013, from http://www.heart.org/HEARTORG/ Conditions/HeartAttack/UnderstandYourRiskofHeartAttack/ Understand-Your-Risk-of-Heart-Attack_UCM_002040_Article.jsp

9. Alizadehsani R, Habibi J, Hosseini MJ, Mashayekhi H, Boghrati R, Ghandeharioun A, Sani ZA (2013) A data mining approach for diagnosis of coronary artery disease. Comput Methods Programs Biomed 111(1):52–61

10. Muthukaruppan S, Er MJ (2012) A hybrid particle swarm optimization based fuzzy expert system for the diagnosis of coronary artery disease. Expert Syst Appl 39(14):11657–11665

11. Tripathy BK, Acharjya DP, Cynthya V (2013) A framework for intelligent medical diagnosis using rough set with formal concept analysis. arXiv preprint arXiv:1301.6011

12. Rajabi M, Hossani S, Dehghani F (2019) A literature review on current approaches and applications of fuzzy expert systems. arXiv preprint arXiv:1909.08794

13. Hedeshi NG, Abadeh MS (2011, December) An expert system working upon an ensemble PSO-based approach for diagnosis of coronary artery disease. In: 2011 18th Iranian conference of biomedical engineering (ICBME). IEEE, pp 249–254

14. Dolatabadi AD, Khadem SEZ, Asl BM (2017) Automated diagnosis of coronary artery disease (CAD) patients using optimized SVM. Comput Methods Programs Biomed 138:117–126

15. Arabasadi Z, Alizadehsani R, Roshanzamir M, Moosaei H, Yarifard AA (2017) Computer-aided decision making for heart disease detection using a hybrid neural network-Genetic algorithm. Comput Methods Programs Biomed 141:19–26

16. Tayefi M, Tajfard M, Saffar S, Hanachi P, Amirabadizadeh AR, Esmaeily H, Ghayour-Mobarhan M (2017) hs-CRP is strongly associated with coronary heart disease (CHD): a data mining approach using decision tree algorithm. Comput Methods Programs Biomed 141:105–109

17. Alizadehsani R, Hosseini MJ, Sani ZA, Ghandeharioun A, Boghrati R (2012, December) Diagnosis of coronary artery disease using cost-sensitive algorithms. In: 2012 IEEE 12th international conference on data mining workshops (ICDMW). IEEE, pp 9–16

18. Setiawan NA, Prabowo DW, Nugroho HA (2014, October) Benchmarking of feature selection techniques for coronary artery disease diagnosis. In: 2014 6th international conference on information technology and electrical engineering (ICITEE). IEEE, pp 1–5

19. Xu S, Shi H, Duan X, Zhu T, Wu P, Liu D (2016, March) Cardiovascular risk prediction method based on text analysis and data mining ensemble system. In: 2016 IEEE international conference on big data analysis (ICBDA). IEEE, pp 1–5

20. Negahbani M, Joulazadeh S, Marateb HR, Mansourian M (2015) Coronary artery disease diagnosis using supervised fuzzy C-means with differential search algorithm-based generalized minkowski metrics. Peertechz J Biomed Eng 1(1):006

21. Muflikhah L, Wahyuningsih Y (2013) Fuzzy rule generation for diagnosis of coronary heart disease risk using subtractive clustering method

22. Chitra R, Seenivasagam V (2013) Heart attack prediction system using Fuzzy C Means classifier. IOSR Journal of Computer Engineering 14:23–31

23. Raju MK, Shafiulilah SK (2015) An Enhanced Risk Prediction System for Cardiovascular Disease in India using Fuzzy Classification. International Journal of Computer Applications 975:8887

24. Alizadehsani R, Zangooei MH, Hosseini MJ, Habibi J, Khosravi A, Roshanzamir M, Nahavandi S (2016) Coronary artery disease detection using computational intelligence methods. Knowl-Based Syst 109:187–197

25. Pal D, Mandana KM, Pal S, Sarkar D, Chakraborty C (2012) Fuzzy expert system approach for coronary artery disease screening using clinical parameters. Knowl-Based Syst 36:162–174

26. Giri D, Rajendra Acharya U, Martis RJ, VinithaSree S, Lim T-C, AhamedVi T, Suri JS (2013) Automated diagnosis of Coronary Artery Disease affected patients using LDA, PCA, ICA, and Discrete Wavelet Transform. Knowl-Based Syst 37:274–282

27. Hogo M (2019) Building data mining models based on a features ranking voting (FRV) algorithm. Egypt Inform J (submitted to)

28. Kurt I, Ture M, Kurum AT (2008) Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease. Expert Syst Appl 34(1):366–374

29. Abdar M, Książek W, Acharya UR, Tan RS, Makarenkov V, Pławiak P (2019) A new machine learning technique for an accurate diagnosis of coronary artery disease. Comput Methods Programs Biomed 179:104992

30. Ghiasi MM, Zendehboudi S, Mohsenipour AA (2020) Decision tree-based diagnosis of coronary artery disease: CART model. Comput Methods Programs Biomed 192:105400

31. Alizadehsani R, Hosseini MJ, Boghrati R, Ghandeharioun A, Khozeimeh F, Sani ZA (2012) Exerting cost-sensitive, and feature creation algorithms for coronary artery disease diagnosis. IGI Global

32. Alizadehsani R, Hosseini MJ, Khosravi A, Khozeimeh F, Roshanzamir M, Sarrafzadegan N et al (2018) Non-invasive detection of coronary artery disease in high-risk patients based on the stenosis prediction of separate coronary arteries. Comput Methods Programs Biomed 162:119–127

33. Acharya UR, Fujita H, Adam M, Lih OS, Sudarshan VK, Hong TJ et al (2017) Automated characterization and classification of coronary artery disease and myocardial infarction by decomposition of ECG signals: a comparative study. Inf. Sci. (NY) 377:17–29

34. Alkeshuosh AH, Moghadam MZ, Mansoori IA, Abdar M (2017) Using PSO algorithm for producing best rules in diagnosis of heart disease. In: 2017 International conference on computer and applications (ICCA), 2017, pp 306–311

35. Steele AJ, Denaxas SC, Shah AD, Hemingway H, Luscombe NM (2018) Machine learning models in electronic health records can outperform conventional survival models for predicting patient mortality in coronary artery disease. PLoS ONE 13(8):e0202344

36. Johnson KM, Johnson HE, Zhao Y, Dowe DA, Staib LH (2019) Scoring of coronary artery disease characteristics on coronary ct angiograms by using machine learning. Radiology 292(2):354–362

37. Alizadehsani R, Abdar M, Roshanzamir M, Khosravi A, Kebria PM, Khozeimeh F et al (2019) Machine learning-based coronary artery disease diagnosis: a comprehensive review. Comput Biol Med 111:103346

38. Witten IH, Frank E, Hall MA, Pal CJ (2016) Data mining: practical machine learning tools and techniques. Morgan Kaufmann, San Francisco

39. Dua D, Graff C (2019) UCI machine learning repository [http://archive.ics.uci.edu/ml]. University of California, School of Information and Computer Science, Irvine, CA

40. Enøe C, Georgiadis MP, Johnson WO (2000) Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. Preventive Veterinary Medicine 45(1–2):61–81

41. Srinivas K, Rao G, Govardhan A (2014) Rough-fuzzy classifier: a system to predict heart disease by blending two different set theories. Arab J Sci Eng 39(4):287–2868

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.