Received: 23 October 2019 | Revised: 16 February 2020 | Accepted: 18 April 2020

DOI: 10.1111/exsy.12573

SPECIAL ISSUE PAPER

Expert Systems WILEY

Hybrid genetic-discretized algorithm to handle data uncertainty in diagnosing stenosis of coronary arteries

Roohallah Alizadehsani[1] | Mohamad Roshanzamir[2] | Moloud Abdar[1] | Adham Beykikhoshk[3] | Abbas Khosravi[1] | Saeid Nahavandi[1] | Pawel Plawiak[4,5] | Ru San Tan[6] | U Rajendra Acharya[7,8,9]

[1]Institute for Intelligent Systems Research and Innovation (IISRI), Deakin University, Geelong, Victoria, Australia

[2]Department of Engineering, Fasa Branch, Islamic Azad University, Fasa, Fars, Iran

[3]Centre for Pattern Recognition and Data Analytics, Deakin University, Geelong, Victoria, Australia

[4]Department of Information and Communications Technology, Faculty of Computer Science and Telecommunications, Cracow University of Technology, Warszawska 24 st., F-3, 31-155, Krakow, Poland

[5]Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, Bałtycka 5, 44-100, Gliwice, Poland

[6]National Heart Centre, Singapore, Singapore

[7]Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore

[8]Department of Biomedical Engineering, School of Science and Technology, Singapore University of Social Sciences, Singapore

[9]Department of Bioinformatics and Medical Engineering, Asia University, Taiwan

Correspondence
Roohallah Alizadehsani, Institute for Intelligent Systems Research and Innovation (IISRI), Deakin University, Geelong, Victoria, Australia.
Email: ralizadehsani@deakin.edu.au

Abstract

Coronary artery disease (CAD) is the leading cause of morbidity and death worldwide. Invasive coronary angiography is the most accurate technique for diagnosing CAD, but is invasive and costly. Hence, analytical methods such as machine learning and data mining techniques are becoming increasingly more popular. Although physicians need to know which arteries are stenotic, most of the researchers focus only on CAD detection and few studies have investigated stenosis of the right coronary artery (RCA), left circumflex (LCX) artery and left anterior descending (LAD) artery separately. Meanwhile, most of the datasets in this field are noisy (data uncertainty). However, to the best of our knowledge, there is no study conducted to address this important problem. This study uses the extension of the Z-Alizadeh Sani dataset, containing 303 records with 54 features. A new feature selection algorithm is proposed in this work. Meanwhile, by discretization of data, we also handle the uncertainty in CAD prediction. To the best of our knowledge, this is the first study attempted to handle uncertainty in CAD prediction. Finally, the genetic algorithm (GA) is used to determine the hyper-parameters of the support vector machine (SVM) kernels. We have achieved high accuracy for the stenosis diagnosis of each main coronary artery. The results of this study can aid the clinicians to validate their manual stenosis diagnosis of RCA, LCX and LAD coronary arteries.

KEYWORDS

coronary artery disease, discretization, feature selection, machine learning, uncertainty

1 | INTRODUCTION

Coronary artery disease (CAD) is a leading cause of morbidity and mortality worldwide. Accurate diagnosis and early institution of secondary prevention measures can avert complications, deaths and also can reduce the treatment costs (Turk-Adawi *et al*., 2017). Invasive coronary angiography (ICA) is the reference standard for CAD diagnosis. It depicts intraluminal coronary anatomy with high accuracy and spatial resolution which is necessary to make correct decisions (Hamon, Morello, Riddell, & Hamon, 2007). But this procedure requires expert operators, has inherent procedural risks and high costs. So, it is better not to use it as a screening tool for all patients but use it only for suspicious cases. For clinical assessment and non-invasive tests electrocardiography (ECG) and echocardiography are commonly used instead, but these are not as sensitive or specific as ICA. Researchers try to improve the sensitivity and specificity of non-invasive methods for CAD diagnoses using machine learning and data mining algorithms.

Machine learning can process multiple parallel clinical and/or diagnostic tests. Various algorithms such as support vector machine (SVM) (Alizadehsani *et al*., 2016; Alizadehsani *et al*., 2018; Kadi, Idri, & Fernandez-Aleman, 2017; Mustaqeem, Anwar, Khan, & Majid, 2017; Pławiak, 2018b; Sharma & Rajendra Acharya, 2019), C4.5 (Abdar, 2015; Abdar, Kalhori, Sutikno, Subroto, & Arji, 2015; Alizadehsani, Habibi, Hosseini, *et al*., 2012), particle swarm optimization (PSO) (Zomorodi-moghadam *et al*., 2019), neural networks (Süt & Şenocak, 2007; Tan *et al*., 2018) and deep learning (Butun, Yildirim, Talo, Tan, & Rajendra Acharya, 2020; Yıldırım, Pławiak, Tan, & Acharya, 2018) have been reported to improve the accuracy of non-invasive triage of CAD. Babagolu *et al*. (Babaoğlu, Fındık, & Bayrak, 2010) employed SVM on exercise stress ECG data and achieved 81.46% accuracy. Kara *et al*. (Kara & Dirgenali, 2007) used neural network on echo-Doppler signal to diagnose CAD.

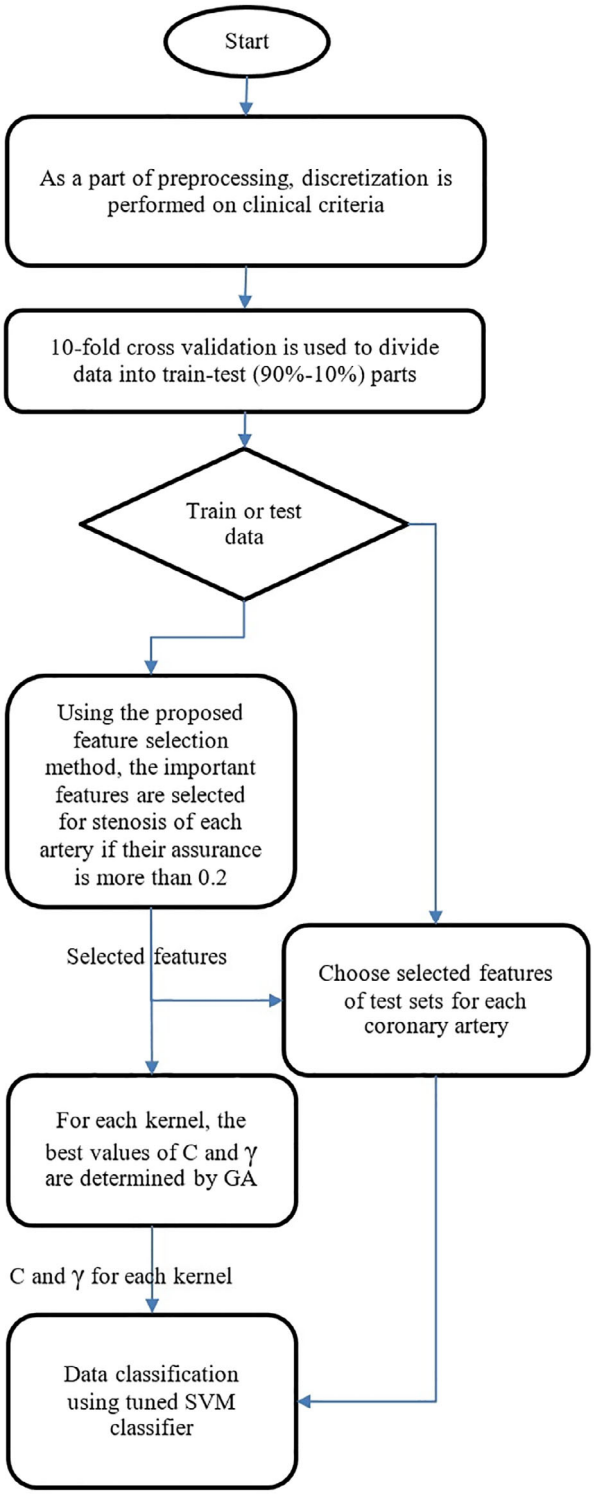

FIGURE 1 The flowchart of the proposed algorithm

Raghavendra *et al*. (Raghavendra *et al*., 2018) used linear discriminant analysis (LDA), and SVM (Polynomial kernel) on ultrasound images and achieved accuracy rate of more than 94% for CAD diagnosis. Polat *et al*. (Polat, Şahan, & Güneş, 2007) used artificial immune recognition system (AIRS) for CAD diagnosis and attained 87% accuracy. In (Acharya *et al*., 2017), the performance of an ECG-based system for CAD detection was tested using discrete cosine transform, empirical mode decomposition and discrete wavelet transform. Then, K-nearest neighbour (KNN) was used for classification, yielded 98.5% accuracy.

In patients who have undergone ICA, machine learning may also add incremental diagnostic value. Wan *et al*. (Wan, Feng, Tong, Li, & Qin, 2018) developed an automated method that could grade the severity of coronary artery stenosis on ICA. Tested on 143 patients with 267 stenosed segments, the algorithm achieved 93.93% accuracy for CAD diagnosis. In (Banchhor *et al*., 2017), principle component analysis (PCA) was used for dimensionality reduction, after which SVM classifier was applied for classification. The accuracy of the designed system was 91.28%. Cost sensitive algorithm is another algorithms used for this purpose which focuses more on improving the sensitivity of algorithm than its accuracy (Alizadehsani, Hosseini, Sani, Ghandeharioun, & Boghrati, 2012; Roohallah *et al*., 2012). In another study, combination of genetic algorithms and artificial neural networks yielded good results for the CAD diagnosis (Arabasadi, Alizadehsani, Roshanzamir, Moosaei, & Yarifard, 2017).

The detection of CAD can either be patient- or vessel-specific. Knowledge about CAD in individual major coronary arterial branches such as left anterior descending (LAD), left circumflex (LCX) and right coronary arteries (RCA) are clinically relevant. For instance, a LAD lesion connotes worse prognosis (as it subtends a larger area of arterial distribution than a LCX or RCA lesion) and may obligate earlier consideration of intervention (Eghbali-Babadi, Khosravi, Feizi, & Sarrafzadegan, 2017). The literature on the use of machine learning methods to diagnose vessel-specific stenosis is rare, as the complexity of diagnosing individual coronary artery stenosis is more challenging than CAD diagnosis (Alizadehsani *et al*., 2018; Alizadehsani, Roshanzamir, *et al*., 2019b). Babaoglu *et al*. (Babaoglu, Baykan, Aygul, Ozdemir, & Bayrak, 2009) used a neural network on exercise stress ECG data and achieved accuracy rates of 73, 64.85, and 69.39% for diagnosis of ICA-validated LAD, LCX and RCA stenosis, respectively. Various decision support rules have been used for the diagnosis of stenosis of individual arteries (Ordonez *et al*., 2001; Soni, Ansari,

TABLE 1 Feature discretization of continuous features for Z-Alizadeh Sani dataset

Feature	Low range	Medium range	High range
Cr^D	0.7 > Cr	1.5 ≥ Cr≥0.7	1.5 < Cr
FBS^D	70 > FBS	105 ≥ FBS≥70	105 < FBS
LDL^D		130 ≥ LDL	130 < LDL
HDL^D	35 > HDL	35 ≤ HDL	—
BUN^D	7 > BUN	20 ≥ BUN≥7	20 < BUN
ESR^D		If the patient is female & age ≥ (ESR-5)*2 or if the patient is male & Age ≥ 2*ESR	If the patient is female & age < (ESR-5)*2 or if the patient is male & age < 2* ESR
Hb^D	If male & 14 > Hb or if female & 12.5 > Hb	If the patient is female & 15 ≥ Hb≥12.5 or if the patient is male & 17 ≥ Hb≥14	If the patient is female & 15 < Hb or if the patient is male & 17 < Hb
K^D	3.8 > K	5.6 ≥ K ≥ 3.8	5.6 < K
Na^D	136 > Na	146 ≥ Na≥136	146 < Na
WBC^D	4,000 > WBC	11,000 ≥ WBC≥4,000	11,000 < WBC
PLT^D	150 > PLT	450 ≥ PLT≥150	450 < PLT
EF^D	50 ≥ EF	50 < EF	
Regional Wall motion $Abnormalities^D$	—	Regional Wall motion abnormalities = 0	Regional Wall motion abnormalities≠0
Age^D		If the patient is female & 55 ≥ age or if the patient is male & 45 ≥ age	If the patient is female & 55 < age or if the patient is male & 45 < age
BP^D	90 > BP	140 ≥ BP≥90	140 < BP
PR^D	60 > PR	100 ≥ PR≥60	100 < PR
Neut ^{D}		65 ≥ Neut	65 < Neut
TG^D		200 ≥ TG	200 < TG
Function $Class^d$		1	2, 3, 4

Sharma, & Soni, 2011; Srinivas, Rani, & Govrdhan, 2010). The performance of few classifiers were compared to determine the association between work-related features and CAD (Nasarian *et al*., 2020). A comprehensive review in this field can be found in (Alizadehsani, Abdar, *et al*., 2019).

In this work, we proposed a novel algorithm to diagnose stenosis in individual coronary arteries separately. The present study used the Z-Alizadeh Sani dataset (described in Table S1) (Alizadehsani, Habibi, Hosseini, *et al*., 2013), which contains granular annotation of clinical variables and individual coronary artery status. The continuous features of dataset are first discretized to reduce the noise. After that, our proposed feature selection approach is applied. Finally, the best values of two SVM parameters (C and γ) are identified using a genetic algorithm.

The structure of this paper is organized as follows: Section 2 introduces the medical dataset used in this study. In Section 3, we describe the data mining methods. In Section 4, experimental results are explained and finally, we conclude the paper in Section 5.

2 | MEDICAL DATASET

There are various datasets available in this field and the complete list can be found in (Alizadehsani, Roshanzamir, *et al*., 2019a). In this work, Z-Alizadeh Sani dataset is used (Alizadehsani, Habibi, Hosseini, *et al*., 2013) as it investigates LAD, LCX and RCA stenosis separately. The cohort comprised of 303 patients with 54 features grouped into: laboratory data, patient's medical history, ECG, physical examinations and echocardiogram. All patients underwent ICA. The coronary artery stenosis in a segment is defined as segment diameter having less than 50% of the diameter

TABLE 2 Weight of features based on assurance for LAD

Feature	Number	Assurance
ST elevation	14	1
Q wave	16	0.9375
Regional Wall motion Abnormalities^{D}	86	0.872093
Edema	12	0.833333
Ex-smoker	10	0.8
Typical CP	164	0.792683
T inversion	90	0.744444
ST depression	71	0.71831
EF^{D}	197	0.71066
FBS^{D}	84	0.690476
DM	90	0.677778
Neut^{D}	89	0.674157
Current smoker	63	0.666667
BP^{D}	48	0.666667
HTN	179	0.659218
Age^{D}	238	0.638655
LDL^{D}	61	0.622951
TG^{D}	62	0.612903
Systolic murmur	41	0.609756
FH	48	0.604167
Week peripheral pulse	5	0.6
LVH	20	0.6
HDL^{D}	87	0.597701
HB^{D}	157	0.592357
Obesity	211	0.563981
DLP	112	0.553571
Dyspnea	134	0.537313
PLT^{D}	12	0.5
Atypical CP	93	0.322581
Lung rales	9	0.222222

of an adjacent normal segment [1]. When any coronary stenosis is present, the affected artery is labelled 1; and 0 when it is absent. Among 303 patients, 177 LAD, 119 LCX and 114 RCA arteries had demonstrable stenosis.

3 | METHOD

In this section, we discussed our approach to predict the diagnosis of coronary artery stenosis.

TABLE 3 Weight of features based on assurance for LCX

Feature	Number	Assurance
Ex-smoker	10	0.7
CR^{D}	22	0.681818
Airway disease	11	0.636364
Poor R progression	9	0.555556
Typical CP	164	0.52439
BP^{D}	48	0.520833
DM	90	0.511111
Edema	12	0.5
PLT^{D}	12	0.5
Regional Wall motion Abnormalities^{D}	86	0.5
Q wave	16	0.5
FBS^{D}	84	0.464286
HTN	179	0.463687
Lung rales	11	0.454545
EF^{D}	197	0.451777
TG^{D}	62	0.451613
Function Class^{D}	92	0.445652
Age^{D}	238	0.445378
T inversion	90	0.444444
Current smoker	63	0.444444
WBC^{D}	27	0.444444
ESR^{D}	46	0.434783
ST elevation	14	0.428571
Neut^{D}	89	0.426966
ST depression	71	0.422535
HB^{D}	157	0.422535
LDL^{D}	62	0.403226
Week peripheral pulse	5	0.4
LVH	20	0.4
FH	48	0.395833
DLP	112	0.383929
Obesity	211	0.379147
Dyspnea	134	0.365672
K^{D}	37	0.351351
HDL^{D}	87	0.344828
Na^{D}	34	0.323529
Systolic murmur	41	0.317073
Atypical CP	93	0.268817
Diastolic murmur	9	0.222222
Thyroid disease	7	0.21

3.1 | Feature analyzing

We used the information gain (Alizadehsani, Habibi, Hosseini, *et al*., 2013), wrapper (Monirul Kabir, Monirul Islam, & Murase, 2010) and embedding (Chandrashekar & Sahin, 2014) methods to rank the features in our dataset based on the influence of features on LAD, LCX and RCA stenosis. The more weight of feature causes more influence on the stenosis of arteries. The selection of each feature, which forms the basis for further

TABLE 4 Weight of features based on assurance for RCA

Feature	Number	Assurance
Poor R progression	9	0.89
Q wave	16	0.63
Weak peripheral pulse	5	0.6
CVA	5	0.6
DM	90	0.58
WBC^D	27	0.52
$Neut^D$	89	0.52
FBS^D	84	0.52
Typical CP	164	0.51
ST elevation	14	0.5
Ex-smoker	10	0.5
PLT^D	12	0.5
BP^D	48	0.46
ESR^D	46	0.46
Airway disease	11	0.45
Function $Class^D$	92	0.45
HTN	179	0.44
Age^D	238	0.44
EF^D	197	0.43
Regional Wall motion $Abnormalities^D$	86	0.43
T inversion	90	0.42
Edema	12	0.42
ST depression	71	0.41
Current smoker	63	0.41
CR^D	22	0.41
FH	48	0.4
DLP	112	0.4
TG^D	62	0.4
LDL^D	62	0.4
HB^D	157	0.4
Obesity	211	0.37
Lung rales	11	0.36
K^D	37	0.35
LVH	20	0.35
HDL^D	87	0.34
Systolic murmur	41	0.34
CRF	6	0.33
Dyspnea	134	0.32
Na^D	34	0.32
Thyroid disease	7	0.29

TABLE 5 The best C and γ values for each kernel in each dataset

	Selected features using assurance											
	LAD				LCX				RCA			
	RBF	Polynomial	Sigmoid	Line7ar	RBF	Polynomial	Sigmoid	Linear	RBF	Polynomial	Sigmoid	Linear
C	4.e+0	5.e-1	4.e+0	1.25e-1	1.6e+1	2.e+0	4.e+0	3.125e-2	1.6e+1	2.e+0	1.024e+3	1.25e-1
γ	1.5625e-2	3.125e-2	3.90625e-3	3.05176e-05	1.953125e-3	6.10352e-05	3.90625e-3	9.76563e-4	3.90625e-3	1.953125e-3	2.44141e-4	4.88281e-4
	Selected features using information gain											
	LAD				LCX				RCA			
	RBF	Polynomial	Sigmoid	Linear	RBF	Polynomial	Sigmoid	Linear	RBF	Polynomial	Sigmoid	Linear
C	6.4e+1	2.e+0	2.56e+2	5.e-1	4.096e+3	4.096e+3	2.56e+2	2.56e+2	8.192e+3	4.096e+3	5.12e+2	2.048e+3
γ	1.953125e-3	1.e+0	9.76563e-4	3.125e-2	6.10352e-05	3.90625e-3	7.8125e-3	5.e-1	1.2207e-4	3.90625e-3	1.953125e-3	5.e-1

analyses, is explained in Section 3.3. In all feature selection methods used in the diagnosis of each coronary artery stenosis, features with weights greater than 0.2 are selected and subsequently classified.

3.2 | SVM classifier

It is a supervised machine learning technique used for classification and regression (Alizadehsani *et al*., 2016). It uses an optimal hyperplane to maximize the margins between data points of different classes in a linear space. Different kernels namely: (a) sigmoid, (b) linear, (c) polynomial and radial basis function (RBF) (Caruana & Niculescu-Mizil, 2006) are used. Choosing the appropriate type of kernel as well as the best values of regularization parameter (C) and gamma (γ) are important to achieve the highest performance. The γ: lower values of γ imply higher influence and higher values represent lower influence (Ruta & Gabrys, 2000). The parameter ***C*** trades off simplicity of decision surface against training misclassification. The higher value of ***C*** tend to classify all training examples correctly and reduces the generalizability of the model to the test data. The lower value for C results in a less complex model with smoother decision surface.

3.3 | Assurance feature selection

In this study, a novel feature selection method named 'assurance feature selection' is proposed. This method is defined for a binomial feature *F* having the values 0 or 1. The assurance of feature *F* for a binomial class label *D* shown by $A_{F,D}$ is calculated according to Equation (1).

$$A_{F,D} = \frac{\text{Count}(F\&D)}{\text{Count}(F)} \times 100, \tag{1}$$

TABLE 6 The best performance using optimum values of C and γ for various SVM kernels by wrapper and embedded feature selection methods

			Naive Bayes	Neural network	Random forest	SVM with polynomial kernel	SVM with RBF kernel	SVM with linear kernel	SVM with sigmoid kernel
Selected features using wrapper	LAD	Accuracy (%)	77.56	81.11	81.02	79.04	82.04	79.82	82.08
		Sensitivity (%)	84.29	87.22	87.02	85.84	88.21	81.84	84.64
		Specificity (%)	70.98	69.04	70.04	71.04	75.12	71.02	77.01
	LCX								
		Accuracy (%)	75.01	79.01	78.01	78.11	79.04	79.03	81.41
		Sensitivity (%)	78.02	86.02	87.14	84.94	84.34	83.19	83.99
		Specificity (%)	72.33	69.00	68.04	69.02	73.22	72.44	76.04
	RCA								
		Accuracy (%)	77.02	78.92	79.00	78.02	79.24	79.21	77.59
		Sensitivity (%)	79.16	86.04	81.02	82.04	85.00	82.33	81.02
		Specificity (%)	75.02	69.84	69.02	71.11	69.98	71.88	72.04
Selected features using embedded	LAD								
		Accuracy (%)	77.09	82.02	81.04	79.31	83.01	82.65	83.11
		Sensitivity (%)	85.03	89.34	88.21	88.96	88.32	87.06	86.87
		Specificity (%)	68.02	69.91	71.02	71.22	75.41	73.19	73.96
	LCX								
		Accuracy (%)	76.04	78.91	79.52	79.02	79.92	79.54	81.34
		Sensitivity (%)	79.94	88.43	89.21	87.94	81.12	83.12	86.01
		Specificity (%)	74.01	69.52	70.05	70.56	74.38	71.12	77.21
	RCA								
		Accuracy (%)	77.04	79.34	79.94	77.46	79.12	78.92	77.97
		Sensitivity (%)	78.80	88.11	84.12	82.84	86.31	83.24	81.30
		Specificity (%)	75.11	69.12	69.43	71.41	69.32	71.69	72.92

TABLE 7 The best performance using optimum values of C and γ for various SVM kernels by information gain and assurance feature selection methods

			Naive Bayes	Neural network	Random forest	SVM with polynomial kernel	SVM with RBF kernel	SVM with linear kernel	SVM with sigmoid kernel
Selected features using information gain	LAD	Accuracy (%)	77.86	81.19	81.85	79.19	82.51	82.31	83.17
		Sensitivity (%)	84.92	89.27	89.29	85.57	88.70	88.14	89.27
		Specificity (%)	70.67	69.84	71.43	72.22	73.81	74.60	74.60
	LCX								
		Accuracy (%)	76.56	79.92	79.22	79.12	81.84	80.86	81.51
		Sensitivity (%)	80.16	89.83	88.40	87.57	86.44	85.31	86.44
		Specificity (%)	75.71	69.84	69.05	70.63	77.78	74.60	76.98
	RCA								
		Accuracy (%)	77.56	79.87	79.68	77.27	80.86	80.52	78.51
		Sensitivity (%)	78.53	88.14	83.01	82.44	85.88	88.14	83.70
		Specificity (%)	76.19	68.25	70.63	70.63	73.81	72.22	73.81
Selected features using assurance (Proposed)	LAD								
		Accuracy (%)	78.56	83.83	84.82	80.12	**86.64**	84.49	85.15
		Sensitivity (%)	79.57	89.83	90.96	86.44	92.96	90.41	90.96
		Specificity (%)	76.84	75.40	76.19	72.63	79.37	75.37	76.98
	LCX								
		Accuracy (%)	78.22	81.82	81.85	79.16	**83.47**	83.12	82.19
		Sensitivity (%)	78.53	93.62	89.83	84.44	90.96	90.40	89.83
		Specificity (%)	77.78	65.08	70.63	75.12	75.22	73.02	69.05
	RCA								
		Accuracy (%)	77.95	79.89	80.58	78.26	**82.85**	80.71	81.10
		Sensitivity (%)	78.66	84.27	83.01	79.66	87.01	88.74	83.44
		Specificity (%)	76.98	71.43	75.40	76.98	74.60	72.22	76.12

Note: The best accuracy achieved for diagnosing stenosis of LAD, LCX, and RCA arteries are bold in this table.

where $D \in \{LAD, LCX, RCA\}$, $F \in Z - Alizade\ Sani\ dataset\ features$, Count($F\&D$) shows the number of patients with feature F equals to 1 having stenosis D, and Count (F) shows the number of patients with feature F equals to 1. After the calculation of assurance of each feature, features with high assurance values are selected for classification.

3.4 | Proposed method

Figure 1 shows our proposed method used to handle uncertainty in CAD prediction. We first discretized the features based on clinical criteria [23]. Then, data is divided according to 10-fold cross validation. Based on the training results, high assurance features are selected. A feature is selected, if its assurance is more than 0.2. This value is selected based on what yielding the best performance for the validation data. This is done to separate LAD, LCX and RCA clearly. Finally, optimal values for the hyper-parameters (C and γ) of SVM are derived using genetic algorithm. Accordingly, high assurance features are applied to the test data and then using the best values calculated for C and γ, the SVM classifies the test data. It is described in detail in the following sections.

3.4.1 | Discretization

Z-Alizadeh Sani dataset comprises of 54 features: 35 discretized, and 19 continuous. We discretized the continuous features based on the clinical criteria according to (Mann, Zipes, Libby, & Bonow, 2014). The discretized values are presented in Table 1. This discretization reduces uncertainty presented in the imprecise data, thereby enhancing the performance of classification algorithms.

3.4.2 | Hyper-parameter optimization of SVM

The best value for C and γ are selected using genetic algorithm (GA) (Haupt & Haupt,2004). The GA is an efficient and reliable meta-heuristic search algorithm for optimal solutions. The method mimics natural evolutionary processes based on bio-inspired operators such as selection, crossover and mutation. In this problem, we are looking for optimal values for C and γ. From a population of 1,000 initial values for C and γ, tournament selection (Haupt & Haupt,2004) is implemented for crossover and mutation, producing the next generation. To update these two variables, a random number of 0≤α≤1 is generated. Then if parents are x and y, their corresponding offspring's are (Zhou, Jin, Zhang, Sendhoff, & Tsang,2006):

$$\text{Offspring}(1) = \alpha.x + (1-\alpha).y \tag{2}$$

$$\text{Offspring}(2) = (1-\alpha).x + \alpha.y. \tag{3}$$

The crossover and mutation rates are 1 and 0.01, respectively. The algorithm terminates after 100,000 generations.

4 | EXPERIMENTAL RESULTS AND DISCUSSION

In this section, the effect of features for the diagnosis of LAD, LCX and RCA stenosis are discussed. Then, classification results are presented in terms of accuracy, sensitivity and specificity.

4.1 | Analysis of arteries' stenosis

To analyze the arteries' stenosis, the assurance (the proposed feature selection method) of the features in classifying the arteries are used. The effect of features on the diagnosis of LAD stenosis is shown in Table 2 based on their assurance. In this table, the assurance of features and the number of patients corresponding those features is equal to 1.

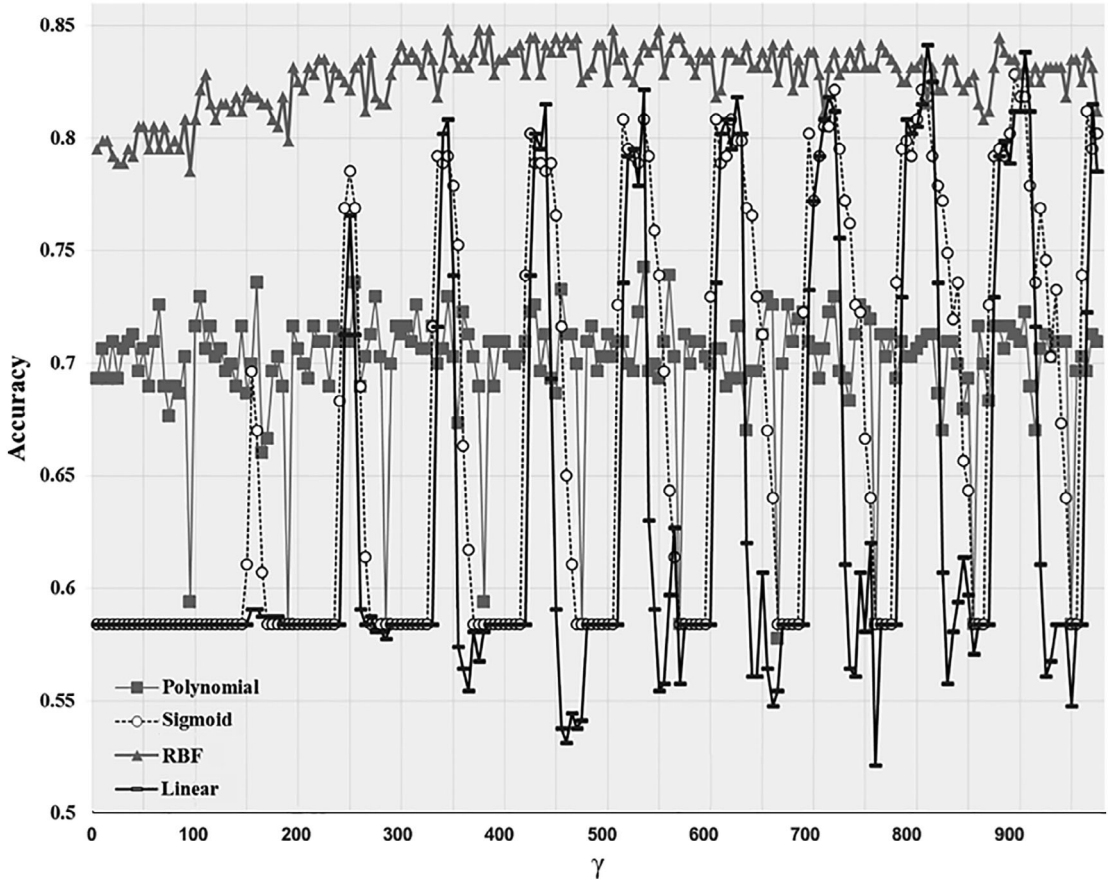

FIGURE 2 Graph of accuracy versus γ of various kernel functions for LAD artery with C = 10

Table 2 demonstrates that the highest assurance values related to ST Elevation, Q Wave, Regional Wall Motion Abnormalities, Edema, Ex-Smoker, Typical chest pain (CP), T inversion, ST Depression, ejection fraction (EF)^{D}, fasting blood sugar (FBS)^{D}, diabetes mellitus (DM), neutrophil count (Neut)^{D}, Current Smoker, blood pressure (BP)^{D} and hypertension (HTN), respectively. The effect of features on the diagnosis of LCX stenosis including their assurance and the number of patients who had those features (value = 1) are shown in Table 3. Ex-Smoker, serum creatinine (Cr)^{D}, Airway Disease, Poor R Progression, Typical CP, BP^{D}, DM, Edema, platelet count (PLT)^{D}, Regional Wall Motion Abnormalities^{D}, Q wave, FBS^{D}, HTN, Lung Rales, EF^{D}, serum triglycerides (TG)^{D}, Function Class^{D}, Age^{D}, T inversion, Current Smoker and white blood cell count (WBC)^{D} are shown in the table in descending order of their assurance values. The effect of features on the diagnosis of RCA stenosis including their assurance and the number of patients who had those features (value = 1) are shown in Table 4.

Poor R Progression, Q Wave, Weak Peripheral Pulse, cerebral vascular accident (CVA), DM, WBC^{D}, Neut^{D}, FBS^{D}, Typical CP, ST Elevation, Ex-Smoker, PLT^{D}, BP^{D}, erythrocyte sedimentation rate (ESR)^{D}, Airway Disease, Function Class^{D}, HTN, Age^{D}, EF^{D}, Regional Wall Motion Abnormalities^{D}, T inversion, Edema, ST Depression, Current Smoker and Cr^{D} have the highest impact on RCA stenosis in terms of assurance.

A comparison of Tables 2–4 shows that the weight of the features in Table 2 is higher than other tables, demonstrating that LAD stenosis can be diagnosed with better accuracy rate than the other two arteries (LCX and RCA).

4.2 | Classification results

The two SVM tuning parameters namely regularization parameter (C) and gamma (γ) are used to obtain the highest classification performance. Table 5 shows the best values of (C, γ) for each kernel and dataset. The accuracy of the method obtained for different values of γ is presented in Figures 2–4 for various kernel functions with C = 10. The RBF kernel does not show much variance for different values of γ. The polynomial kernel is ranked second highest which shows slightly higher variation than RBF, and its fluctuations are less than linear and sigmoid kernels. The linear and sigmoid kernel variances are significantly more than RBF and polynomial kernels. Hence, finding the optimal values for linear and sigmoid kernels are more important than RBF and polynomial kernels for this dataset.

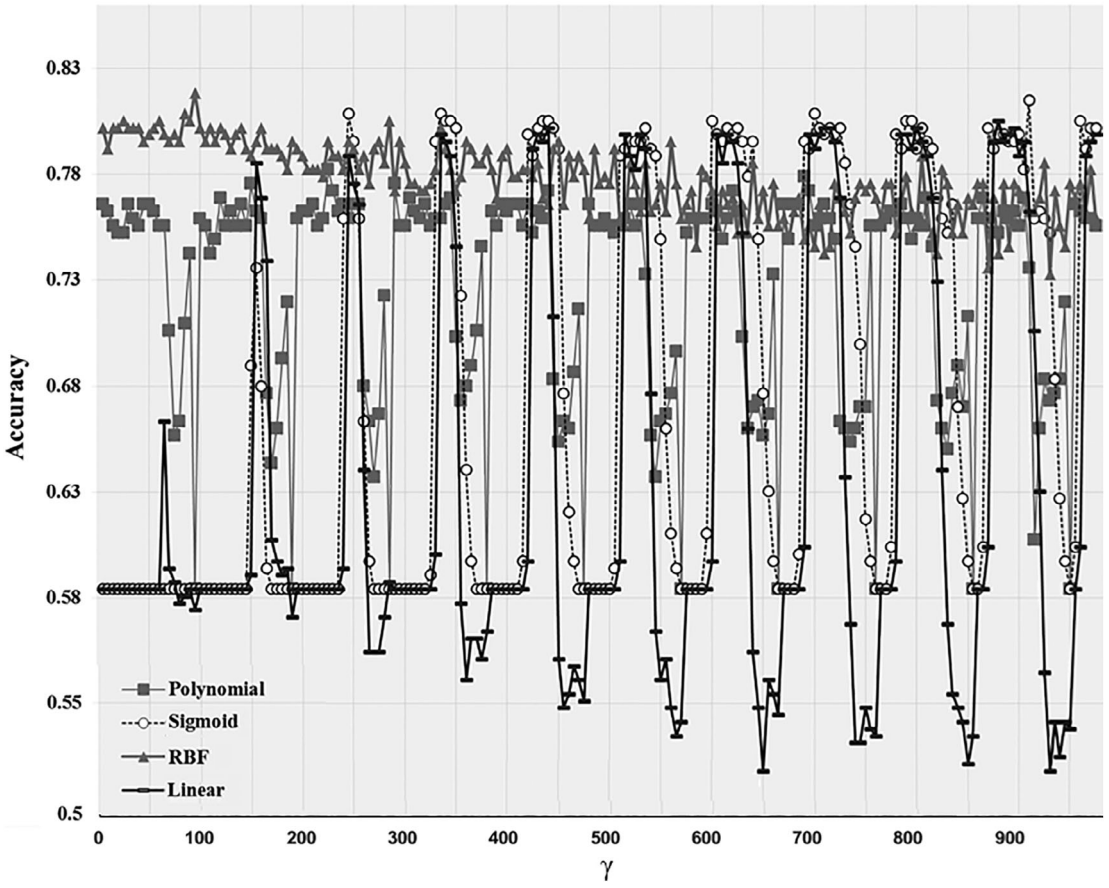

FIGURE 3 Graph of accuracy versus γ of various kernel functions for LCX artery with C = 10

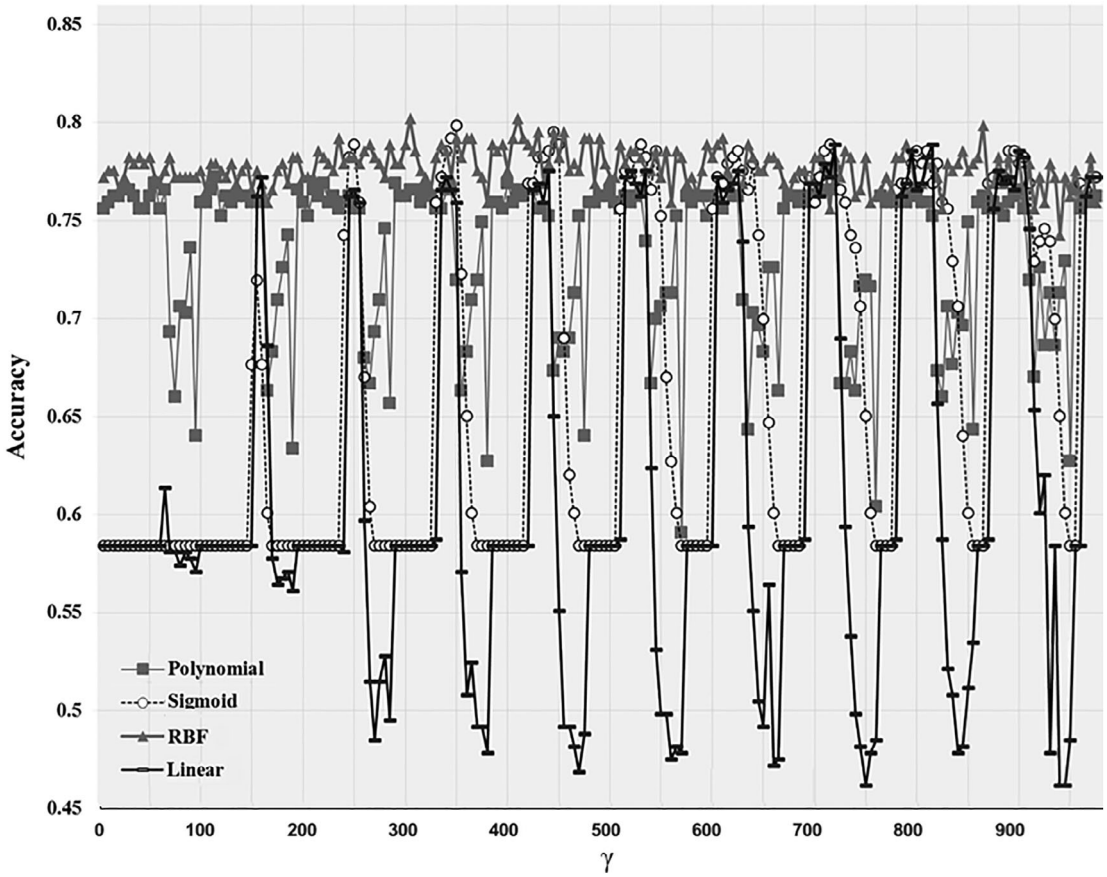

FIGURE 4 Graph of accuracy versus γ of various kernel functions for RCA artery with C = 10

To study the effect of assurance feature selection and discretization, we compared the experimental results with *three* other feature selection algorithms: (a) embedded, (b) wrapper and (c) information gain. The algorithms are compared in Tables 6 and 7. With information gain s feature selection method, the highest accuracy of 83.1% is obtained for the diagnosis of LAD stenosis with sigmoid kernel as indicated in the first row of Table 7. Using RBF kernel, the highest accuracies of 81.84 and 80.86% are achieved for the diagnosis of LCX and RCA, respectively.

It can be noted from Figure 5 that, the best results are obtained for the proposed algorithm, followed by information gain, embedding and wrapper methods. The second row in Table 7 shows the results using assurance feature selection method followed by SVM algorithm (RBF kernel function). The best accuracy for the stenosis diagnosis of LAD, LCX and RCA is 86.64, 83.47 and 82.85%, respectively. Our proposed method achieved the highest accuracy as compared to other reported methods when other feature selection algorithms such as wrapper, embedded and information gain are applied.

4.3 | Discussion

As far as we know, there is no paper that considered the uncertainty in CAD data for vessel-specific diagnosis of CAD. In Table 8, we compared the works conducted in the automated diagnosis of major coronary arteries, separately. It is clear from this discussion table that our proposed method is the best algorithm for diagnosing each main coronary artery stenosis.

Two types of uncertainties exists: (a) uncertainty in the data and (b) uncertainty in the model (Gal & Ghahramani, 2016; Kendall & Gal, 2017). Investigating uncertainty in the models while each vessel is considered separated was done by (Alizadehsani, Roshanzamir, *et al*., 2019b). Distance of sample from hyperplane alongside with accuracy rate of the classifier was used to handle the model uncertainty. As far as we know there is no research investigating the uncertainty in the data. The advantage of our proposed method is its robustness against uncertainty in the data. The uncertainty in the clinical data is a common problem. For example, the error caused due to usage of instruments or human mistakes can easily generate uncertainty in the extracted data. Some of them may be vital in the disease diagnosis process. Unfortunately, this important problem is ignored in most of those studies which use machine learning and data mining algorithms for diagnosing different diseases and cancers. Hence, solving this issue can improve the reliability of machine learning and data mining methods in disease diagnosis. However, in our proposed method, we cannot handle uncertainty in both data and models. This is

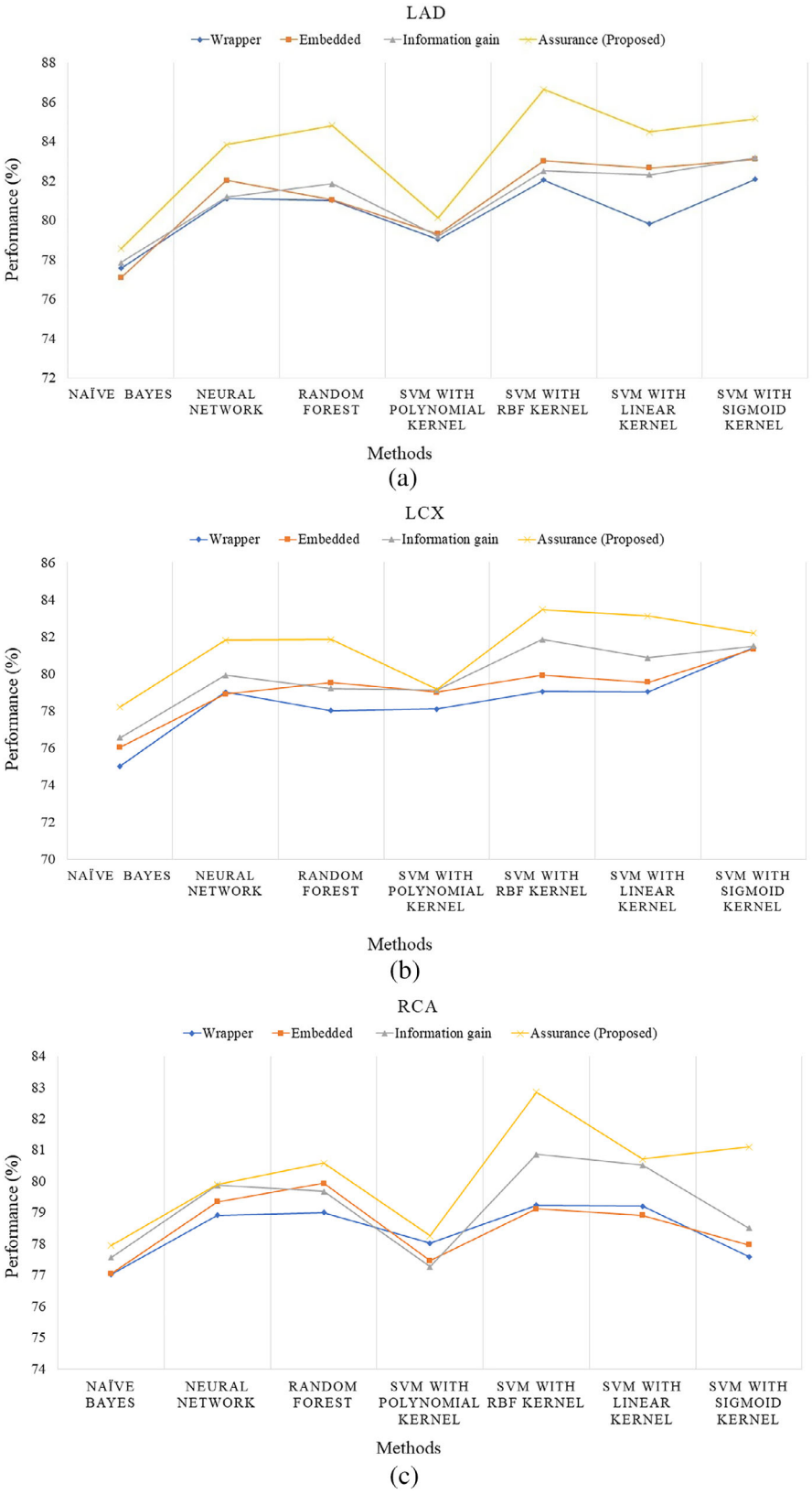

FIGURE 5 Graph of performance (accuracy %) versus different classifiers for different stenosis: (a) LAD, (b) LCX and (c) RCA

one of the weaknesses of our method that we can overcome it in our future works using evolutionary algorithms (Alkeshuosh, Moghadam, Mansoori, & Abdar, 2017; Hassoon, Kouhi, Zomorodi-Moghadam, & Abdar, 2017; Książek, Abdar, Acharya, & Pławiak, 2019; Pławiak, 2018a; Pławiak & Acharya, 2019). Meanwhile, it is obvious that handling uncertainty is computationally intensive. Having sound primary knowledge in the area of stenosis diagnosis for the discretization of features is an important weakness of our proposed method.

TABLE 8 Summary of works conducted on automated detection of stenosis

References	Partition type	Best classifier	Accuracy (%)		
			LAD	LCX	RCA
Babaoğlu *et al*. (2010)	10-fold	Artificial neural network	73.0	64.8	69.4
Alizadehsani, Habibi, Alizadeh Sani *et al*. (2013)	10-fold	Bagging	79.5	65.1	68.0
Alizadehsani, Habibi, Bahadorian et al. (2012)	10-fold	Decision tree	74.2	63.8	68.3
Garcia *et al*. (2001)	NA	Expert system	72.0	74.0	77.0
Our method	10-fold	SVM	**86.6**	**83.5**	**82.9**

5 | CONCLUSION

In this work, a novel method for automated stenosis diagnosis of LAD, LCX and RCA coronary arteries is proposed. We have achieved an accuracy of 86.64, 83.47 and 82.85% for the stenosis diagnosis of LAD, LCX and RCA, respectively. As far as we know, this paper is the first one that investigates the data uncertainty in the stenosis diagnosis of LAD, LCX and RCA. We have obtained the highest accuracy for the diagnosis of LAD stenosis as compared to LCX and RCA. Moreover, SVM with RBF kernel function is the best as compared to other classification methods. The results show that discretization together with assurance feature selection can improve the efficiency of classification algorithms significantly. In future, we intend to use new methods to handle the data uncertainty and improve the performance of stenosis diagnosis. Also, different types of evolutionary algorithms may be used instead of genetic algorithm to develop a hybrid feature selection algorithm (Alizadehsani, Abdar, *et al*., 2019; Alizadehsani, Habibi, Hosseini, *et al*., 2013; Alizadehsani, Hosseini, *et al*., 2012; Alizadehsani, Roshanzamir, *et al*., 2019b; Alizadehsani, Roshanzamir, *et al*., 2019a; Arabasadi *et al*., 2017; Roohallah *et al*., 2012; Zomorodi-moghadam *et al*., 2019).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Roohallah Alizadehsani https://orcid.org/0000-0002-3069-7932

REFERENCES

Abdar, M. (2015). Using decision trees in data mining for predicting factors influencing of heart disease. *Carpathian Journal of Electronic and Computer Engineering*, *8*(2), 31–36.

Abdar, M., Kalhori, S. R. N., Sutikno, T., Subroto, I. M. I., & Arji, G. (2015). Comparing performance of data mining algorithms in prediction heart diseases. *International Journal of Electrical and Computer Engineering (IJECE)*, *5*(6), 1569–1576.

Acharya, U. R., Fujita, H., Adam, M., Lih, O. S., Sudarshan, V. K., Hong, T. J., … San, T. R. (2017). Automated characterization and classification of coronary artery disease and myocardial infarction by decomposition of ECG signals: A comparative study. *Information Sciences*, *377*, 17–29. https://doi.org/10.1016/j.ins.2016.10.013

Aimin, Z., Yaochu, J., Qingfu, Z., Sendhoff, B., & Tsang, E. (2006, July). *Combining model-based and genetics-based offspring generation for multi-objective optimization using a convergence criterion*. Paper presented at the 2006 IEEE International Conference on Evolutionary Computation, Vancouver, Canada.

Alizadehsani, R., Abdar, M., Roshanzamir, M., Khosravi, A., Kebria, P. M., Khozeimeh, F., … Acharya, U. R. (2019). Machine learning-based coronary artery disease diagnosis: A comprehensive review. *Computers in Biology and Medicine*, *111*, 103346. https://doi.org/10.1016/j.compbiomed.2019.103346

Alizadehsani, R., Habibi, J., Alizadeh Sani, Z., Mashayekhi, H., Boghrati, R., Ghandeharioun, A., … Alizadeh-Sani, F. (2013). Diagnosing coronary artery disease via data mining algorithms by considering laboratory and echocardiography features. *Research in Cardiovascular Medicine*, *2*(3), 133–139. https://doi.org/10.5812/cardiovascmed.10888

Alizadehsani, R., Habibi, J., Bahadorian, B., Mashayekhi, H., Ghandeharioun, A., Boghrati, R., & Sani, Z. A. (2012). Diagnosis of coronary arteries stenosis using data mining. *Journal of Medical Signals and Sensors*, *2*(3), 153–159.

Alizadehsani, R., Habibi, J., Hosseini, M. J., Boghrati, R., Ghandeharioun, A., Bahadorian, B., & Sani, Z. A. (2012). Diagnosis of coronary artery disease using data mining techniques based on symptoms and ECG features. *European Journal of Scientific Research*, *82*(4), 542–553.

Alizadehsani, R., Habibi, J., Hosseini, M. J., Mashayekhi, H., Boghrati, R., Ghandeharioun, A., … Sani, Z. A. (2013). A data mining approach for diagnosis of coronary artery disease. *Computer Methods and Programs in Biomedicine*, *111*(1), 52–61. https://doi.org/10.1016/j.cmpb.2013.03.004

Alizadehsani, R., Hosseini, M. J., Khosravi, A., Khozeimeh, F., Roshanzamir, M., Sarrafzadegan, N., & Nahavandi, S. (2018). Non-invasive detection of coronary artery disease in high-risk patients based on the stenosis prediction of separate coronary arteries. *Computer Methods and Programs in Biomedicine*, *162*, 119–127. https://doi.org/10.1016/j.cmpb.2018.05.009

Alizadehsani, R., Hosseini, M. J., Sani, Z. A., Ghandeharioun, A., & Boghrati, R. (2012, December). *Diagnosis of coronary artery disease using cost-sensitive algorithms*. Paper presented at the 2012 IEEE 12th International Conference on Data Mining Workshops, Brussels, Belgium.

Alizadehsani, R., Roshanzamir, M., Abdar, M., Beykikhoshk, A., Khosravi, A., Panahiazar, M., … Sarrafzadegan, N. (2019a). A database for using machine learning and data mining techniques for coronary artery disease diagnosis. *Scientific Data*, *6*(1), 227. https://doi.org/10.1038/s41597-019-0206-3

Alizadehsani, R., Roshanzamir, M., Abdar, M., Beykikhoshk, A., Zangooei, M. H., Khosravi, A., … Acharya, U. R. (2019b). Model uncertainty quantification for diagnosis of each main coronary artery stenosis. *Soft Computing*, *2*, 1–12. https://doi.org/10.1007/s00500-019-04531-0

Alizadehsani, R., Zangooei, M. H., Hosseini, M. J., Habibi, J., Khosravi, A., Roshanzamir, M., … Nahavandi, S. (2016). Coronary artery disease detection using computational intelligence methods. *Knowledge-Based Systems*, *109*, 187–197. https://doi.org/10.1016/j.knosys.2016.07.004

Alkeshuosh, A. H., Moghadam, M. Z., Mansoori, I. A., & Abdar, M. (2017). *Using PSO algorithm for producing best rules in diagnosis of heart disease*. Paper presented at the International Conference on Computer and Applications (ICCA), Doha, UAE.

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

Babaoglu, I., Baykan, O. K., Aygul, N., Ozdemir, K., & Bayrak, M. (2009). Assessment of exercise stress testing with artificial neural network in determining coronary artery disease and predicting lesion localization. *Expert Systems with Applications*, *36*(2), Part 1, 2562–2566. https://doi.org/10.1016/j.eswa.2007.11.013

Babaoğlu, I., Fındık, O., & Bayrak, M. (2010). Effects of principle component analysis on assessment of coronary artery diseases using support vector machine. *Expert Systems with Applications*, *37*(3), 2182–2185. https://doi.org/10.1016/j.eswa.2009.07.055

Banchhor, S. K., Londhe, N. D., Araki, T., Saba, L., Radeva, P., Laird, J. R., & Suri, J. S. (2017). Wall-based measurement features provides an improved IVUS coronary artery risk assessment when fused with plaque texture-based features during machine learning paradigm. *Computers in Biology and Medicine*, *91*, 198–212. https://doi.org/10.1016/j.compbiomed.2017.10.019

Butun, E., Yildirim, O., Talo, M., Tan, R.-S., & Rajendra Acharya, U. (2020). 1D-CADCapsNet: One dimensional deep capsule networks for coronary artery disease detection using ECG signals. *Physica Medica*, *70*, 39–48. https://doi.org/10.1016/j.ejmp.2020.01.007

Caruana, R., & Niculescu-Mizil, A. (2006, June). *An empirical comparison of supervised learning algorithms*. Paper presented at the Proceedings of the 23rd International Conference on Machine learning, Pittsburgh, PA.

Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. *Computers & Electrical Engineering*, *40*(1), 16–28. https://doi.org/10.1016/j.compeleceng.2013.11.024

Eghbali-Babadi, M., Khosravi, A., Feizi, A., & Sarrafzadegan, N. (2017). Design and implementation of a combined observational and interventional study: Trends of prevalence, awareness, treatment and control hypertension and the effect of expanded chronic care model on control, treatment and self-care. *ARYA Atherosclerosis*, *13*(5), 211–220.

Gal, Y., & Ghahramani, Z. (2016, June). *Dropout as a Bayesian approximation: Representing model uncertainty in deep learning*. Paper presented at the Proceedings of the 33rd International Conference on Machine Learning. Retrieved from http://proceedings.mlr.press

Garcia, E. V., Cooke, C. D., Folks, R. D., Santana, C. A., Krawczynska, E. G., De Braal, L., & Ezquerra, N. F. (2001). Diagnostic performance of an expert system for the interpretation of myocardial perfusion SPECT studies. *Journal of Nuclear Medicine*, *42*(8), 1185–1191.

Hamon, M., Morello, R., Riddell, J. W., & Hamon, M. (2007). Coronary arteries: Diagnostic performance of 16-versus 64-section spiral CT compared with invasive coronary angiography—Meta-analysis. *Radiology*, *245*(3), 720–731. https://doi.org/10.1148/radiol.2453061899

Hassoon, M., Kouhi, M. S., Zomorodi-Moghadam, M., & Abdar, M. (2017, September). *Rule optimization of boosted C5.0 classification using genetic algorithm for liver disease prediction*. Paper presented at the 2017 International Conference on Computer and Applications (ICCA), Doha, UAE.

Haupt, R. L., & Haupt, S. E. (2004). *Practical genetic algorithms*, 2nd ed. Hoboken, NJ: John Wiley & Sons.

Kadi, I., Idri, A., & Fernandez-Aleman, J. L. (2017). Knowledge discovery in cardiology: A systematic literature review. *International Journal of Medical Informatics*, *97*, 12–32. https://doi.org/10.1016/j.ijmedinf.2016.09.005

Kara, S., & Dirgenali, F. (2007). A system to diagnose atherosclerosis via wavelet transforms, principal component analysis and artificial neural networks. *Expert Systems with Applications*, *32*(2), 632–640. https://doi.org/10.1016/j.eswa.2006.01.043

Kendall, A., & Gal, Y. (2017). *What uncertainties do we need in Bayesian deep learning for computer vision?* Paper presented at the Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA.

Książek, W., Abdar, M., Acharya, U. R., & Pławiak, P. (2019). A novel machine learning approach for early detection of hepatocellular carcinoma patients. *Cognitive Systems Research*, *54*, 116–127. https://doi.org/10.1016/j.cogsys.2018.12.001

Mann, D. L., Zipes, D. P., Libby, P., & Bonow, R. O. (2014). *Braunwald's heart disease: A textbook of cardiovascular medicine*, Amsterdam, Netherlands: Elsevier Health Sciences.

Monirul Kabir, M., Monirul Islam, M., & Murase, K. (2010). A new wrapper feature selection approach using neural network. *Neurocomputing*, *73*(16), 3273–3283. https://doi.org/10.1016/j.neucom.2010.04.003

Mustaqeem, A., Anwar, S. M., Khan, A. R., & Majid, M. (2017). A statistical analysis based recommender model for heart disease patients. *International Journal of Medical Informatics*, *108*, 134–145. https://doi.org/10.1016/j.ijmedinf.2017.10.008

Nasarian, E., Abdar, M., Fahami, M. A., Alizadehsani, R., Hussain, S., Basiri, M. E., … Sarrafzadegan, N. (2020). Association between work-related features and coronary artery disease: A heterogeneous hybrid feature selection integrated with balancing approach. *Pattern Recognition Letters*, *133*, 33–40. https://doi.org/10.1016/j.patrec.2020.02.010

Ordonez, C., Omiecinski, E., Braal, L. d., Santana, C. A., Ezquerra, N., Taboada, J. A., … Garcia, E. V. (2001, November–December). *Mining constrained association rules to predict heart disease*. Paper presented at the Proceedings IEEE International Conference on Data Mining, San Jose, CA.

Pławiak, P. (2018a). Novel genetic ensembles of classifiers applied to myocardium dysfunction recognition based on ECG signals. *Swarm and Evolutionary Computation*, *39*, 192–208. https://doi.org/10.1016/j.swevo.2017.10.002

Pławiak, P. (2018b). Novel methodology of cardiac health recognition based on ECG signals and evolutionary-neural system. *Expert Systems with Applications*, *92*, 334–349. https://doi.org/10.1016/j.eswa.2017.09.022

Pławiak, P., & Acharya, U. R. (2019). Novel deep genetic ensemble of classifiers for arrhythmia detection using ECG signals. *Neural Computing and Applications*, *2*, 1–25. https://doi.org/10.1007/s00521-018-03980-2

Polat, K., Şahan, S., & Güneş, S. (2007). Automatic detection of heart disease using an artificial immune recognition system (AIRS) with fuzzy resource allocation mechanism and k-nn (nearest neighbour) based weighting preprocessing. *Expert Systems with Applications*, *32*(2), 625–631. https://doi.org/10.1016/j.eswa.2006.01.027

Raghavendra, U., Fujita, H., Gudigar, A., Shetty, R., Nayak, K., Pai, U., … Acharya, U. R. (2018). Automated technique for coronary artery disease characterization and classification using DD-DTDWT in ultrasound images. *Biomedical Signal Processing and Control*, *40*, 324–334. https://doi.org/10.1016/j.bspc.2017.09.030

Roohallah, A., Mohammad Javad, H., Reihane, B., Asma, G., Fahime, K., & Zahra Alizadeh, S. (2012). Exerting cost-sensitive and feature creation algorithms for coronary artery disease diagnosis. *International Journal of Knowledge Discovery in Bioinformatics (IJKDB)*, *3*(1), 59–79. https://doi.org/10.4018/jkdb.2012010104

Ruta, D., & Gabrys, B. (2000). An overview of classifier fusion methods. *Computing and Information Systems*, *7*(1), 1–10.

Sharma, M., & Rajendra Acharya, U. (2019). A new method to identify coronary artery disease with ECG signals and time-Frequency concentrated antisymmetric biorthogonal wavelet filter bank. *Pattern Recognition Letters*, *125*, 235–240. https://doi.org/10.1016/j.patrec.2019.04.014

Soni, J., Ansari, U., Sharma, D., & Soni, S. (2011). Predictive data mining for medical diagnosis: An overview of heart disease prediction. *International Journal of Computer Applications*, *17*(8), 43–48.

Srinivas, K., Rani, B. K., & Govrdhan, A. (2010). Applications of data mining techniques in healthcare and prediction of heart attacks. *International Journal of Advanced Trends in Computer Science and Engineering*, *2*(02), 250–255.

Süt, N., & Şenocak, M. (2007). Assessment of the performances of multilayer perceptron neural networks in comparison with recurrent neural networks and two statistical methods for diagnosing coronary artery disease. *Expert Systems*, *24*(3), 131–142. https://doi.org/10.1111/j.1468-0394.2007.00425.x

Tan, J. H., Hagiwara, Y., Pang, W., Lim, I., Oh, S. L., Adam, M., … Acharya, U. R. (2018). Application of stacked convolutional and long short-term memory network for accurate identification of CAD ECG signals. *Computers in Biology and Medicine*, *94*, 19–26. https://doi.org/10.1016/j.compbiomed.2017.12.023

Turk-Adawi, K., Sarrafzadegan, N., Fadhil, I., Taubert, K., Sadeghi, M., Wenger, N. K., … Grace, S. L. (2017). Cardiovascular disease in the Eastern Mediterranean region: Epidemiology and risk factor burden. *Nature Reviews Cardiology*, *15*, 106–119. https://doi.org/10.1038/nrcardio.2017.138

Wan, T., Feng, H., Tong, C., Li, D., & Qin, Z. (2018). Automated identification and grading of coronary artery stenoses with X-ray angiography. *Computer Methods and Programs in Biomedicine*, *167*, 13–22. https://doi.org/10.1016/j.cmpb.2018.10.013

Yıldırım, Ö., Pławiak, P., Tan, R.-S., & Acharya, U. R. (2018). Arrhythmia detection using deep convolutional neural network with long duration ECG signals. *Computers in Biology and Medicine*, *102*, 411–420. https://doi.org/10.1016/j.compbiomed.2018.09.009

Zomorodi-moghadam, M., Abdar, M., Davarzani, Z., Zhou, X., Pławiak, P., & Acharya, U. R. (2019). Hybrid particle swarm optimization for rule discovery in the diagnosis of coronary artery disease. *Expert Systems*, 2, e12485, https://doi.org/10.1111/exsy.12485

AUTHOR BIOGRAPHIES

Roohallah Alizadehsani obtained a Bachelor of Science degree in computer engineering-software from Sharif University of Technology and then received a Master in Science in computer engineering software from Sharif University of Technology in 2012. He has worked in Software engineering Lab under supervision of Dr. Jafar Habibi since 2011. Roohallah's research interests include mainly in the areas of data mining, machine learning, bioinformatics, Heart Disease, Skin Disease, Diabetes disease, Hepatitis disease, and Cancer Disease. He is currently a Ph.D. student at Deakin University in Australia under supervision of Dr. Abbas Khosravi.

Mohamad Roshanzamir is a PhD student and has some papers in the field of deep learning, machine learning, and bioinformatics.

Moloud Abdar received the bachelor's degree in computer engineering from Damghan University, Iran, in 2015, and the master's degree in computer science and engineering from the University of Aizu, Aizu, Japan, in 2018. He is currently pursuing the Ph.D. degree. He has written several articles in the fields of data mining, machine learning, and user modeling in some refereed international journals and conferences. He is also very active in several international conferences, including the TPC in the IEEE AINA 2018, the IEEE AINA 2019, and the IEEE AINA 2020 and some referred international journals, such as the IEEE ACCESS, the Applied Soft Computing, the Future Generation Computer Systems (outstanding reviewer in October 2017), the Neurocomputing (Outstanding Reviewer in January 2017), and the Neural Computing and Applications as a Reviewer. His research interests include data mining, machine learning, ensemble learning, evolutionary algorithms, and user modeling. He was a recipient of the Fonds de Recherche du Quebec—Nature et Technologies Award (ranked 5th among 20 candidates in the second round of selection process), in 2019.

Adham Beykikhoshk has some papers in the field of deep learning, machine learning, and bioinformatics.

Abbas Khosravi (M'10) received the B.Sc. degree in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 2002 and the M.Sc. degree (Hons.) in electrical engineering from the Amirkabir University of Technology, Tehran, in 2005. He joined the eXiT Group as a Research Academic with the University of Girona, Girona, Spain, in 2006, researching in the area of artificial intelligence (AI) applications. He is currently a Senior Research Fellow with the Institute for Intelligent Systems Research and Innovation, Deakin University (Waurn Ponds Campus), Waurn Ponds, VIC, Australia. His current research interests include the development and application of AI techniques for (meta) modeling, analysis, control, and optimization of operations within complex systems.

Saeid Nahavandi (M'92–SM'07) received the Ph.D. degree from Durham University, Durham, U.K., in 1991. He is an Alfred Deakin Professor, the Pro Vice-Chancellor (Defence Technologies), the Chair of Engineering, and the Director for the Institute for Intelligent Systems Research and Innovation, Deakin University (Waurn Ponds Campus), Waurn Ponds, VIC, Australia. He has published over 600 papers in various international journals and conferences. His current research interests include modeling of complex systems, robotics, and haptics. Dr. Nahavandi is the Co-Editor-in-Chief of the IEEE Systems Journal, an Associate Editor of the IEEE/ASME Transactions on Mechatronics and the IEEE

Transactions on Systems, Man and Cybernetics: Systems, and an Editorial Board Member of IEEE Access. He is a fellow of Engineers Australia and the Institution of Engineering and Technology.

Paweł Pławiak, B.Eng., M.Sc., Ph.D., D.Sc. was born in Ostrowiec, Poland, in 1984. He obtained his B.Eng. and M.Sc. degree in Electronics and Telecommunications and his Ph.D. degree with honors in Biocybernetics and Biomedical Engineering at the AGH University of Science and Technology, Cracow, Poland, in 2012 and 2016, respectively. He obtained his D.Sc. degree (habilitation) in Computer Science at The Silesian Technical University, Gliwice, Poland, in 2020. He is the Head of the Department of Information and Communications Technology and an Associate Professor in Cracow University of Technology, Cracow, and Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, Gliwice Poland. He has published more than 20 papers in refereed international SCI-IF journals. He is a reviewer of many prestigious and reputed journals. His research interests include machine learning and computational intelligence (e.g., artificial neural networks, genetic algorithms, fuzzy systems, support vector machines, k-nearest neighbors, and hybrid systems), ensemble learning, deep learning, evolutionary computation, classification, pattern recognition, signal processing and analysis, data analysis and data mining, sensor techniques, medicine, biocybernetics, and biomedical engineering.

Ru San Tan is currently a cardiologist with the National Heart Center Singapore and holds concurrent appointments as a Deputy Director with the National Heart Research Institute of Singapore, and an Adjunct Associate Professor with the Duke-National University Singapore Graduate Medical School Singapore. He specializes in noninvasive cardiovascular imaging (echocardiography and magnetic resonance imaging), and leads a team of scientists involved in bioengineering and cardiovascular imaging research. He is an experienced clinical trialist and site principal investigator of several multicentre cardiology clinical trials.

U. Rajendra Acharya received the Ph.D. degree from the National Institute of Technology Karnataka, Surathkal, India, and the D.Eng. degree from Chiba University, Japan. He is currently a Senior Faculty Member with Ngee Ann Polytechnic, Singapore. He is also an Adjunct Professor with Taylor's University, Malaysia, an Adjunct Faculty with the Singapore Institute of Technology, University of Glasgow, Singapore, and an Associate Faculty with the Singapore University of Social Sciences, Singapore. He has published more than 400 articles, refereed international SCI-IF journals (345), international conference proceedings (42), books (17) with more than 24,500 citations in Google Scholar (with H-index of 93), and Research Gate RG Score of 47.66. He is ranked in the top 1% of the Highly Cited Researchers for the last three consecutive years (2016, 2017, and 2018) in computer science according to the Essential Science Indicators of Thomson. He has worked on various funded projects, with grants worth more than two million SGD. He has three patents. His research interests include biomedical signal processing, biomedical imaging, data mining, visualization and biophysics for better healthcare design, delivery, and therapy. He has served as a Guest Editor for many journals. He is also an Editorial Board Member of many journals.

How to cite this article: Alizadehsani R, Roshanzamir M, Abdar M, et al. Hybrid genetic-discretized algorithm to handle data uncertainty in diagnosing stenosis of coronary arteries. *Expert Systems*. 2020;1–17. https://doi.org/10.1111/exsy.12573