



A Rule-Based Expert System to Assess Coronary Artery Disease Under Uncertainty

Sohrab Hossain¹ (✉), Dhiman Sarma², Rana Joyti Chakma², Wahidul Alam³,
Mohammed Moshiul Hoque⁴, and Iqbal H. Sarker⁴

¹ Department of Computer Science and Engineering, East Delta University,
Chittagong, Bangladesh

sohrab.h@eastdelta.edu.bd

² Department of Computer Science and Engineering,
Rangamati Science and Technology University, Rangamati, Bangladesh

³ Department of Computer Science and Engineering,
University of Science and Technology Chittagong, Chittagong, Bangladesh

⁴ Department of Computer Science and Engineering, Chittagong
University of Engineering and Technology, Chittagong, Bangladesh

iqbal@cuet.ac.bd

Abstract. The coronary artery disease (CAD) occurs from the narrowing and damaging of major blood vessels or arteries. It has become the most life-threatening disease in the world, especially in the South Asian region. Its detection and treatment involve expensive medical facilities. The early detection of CAD, which is a major challenge, can minimize the patients' suffering and expenses. The major challenge for CAD detection is incorporating numerous factors for detailed analysis. The goal of this study is to propose a new Clinical Decision Support System (CDSS) which may assist doctors in analyzing numerous factors more accurately than the existing CDSSs. In this paper, a Rule-Based Expert System (RBES) is proposed which involves five different Belief Rules, and can predict five different stages of CAD. The final output is produced by combining all BRBs and by using the Evidential Reasoning (ER). Performance evaluation is measured by calculating the success rate, error rate, failure rate and false omission rate. The proposed RBES has higher a success rate and false omission rate than other existing CDSSs.

Keywords: Rule based system · Expert system · Prediction · Uncertainty · Coronary artery disease · Clinical decision support systems · Health analytics

1 Introduction

Coronary artery disease (CAD) is a condition when the coronary arteries become narrow or blocked. It is developed when bad cholesterol and plaques (fatty droplets) deposit inside the wall of arteries. The process is termed as atherosclerosis which means clogging of arteries, and reduces blood flow inside the heart muscle. Blood carries oxygen and essential nutrients to the heart [1]. Lack of sufficient blood supply can cause angina

(chest pain), and lead to a heart attack by injuring heart muscle. The death toll due to heart disease is 16.3 million in America each year which has made it the leading cause of death in the United States. According to the American Heart Association (AHA), one person is suffered from a heart attack in every 40 s. Having zero risk factors of heart disease, any male has 3.6% and any female has less than 1% chance of getting cardiovascular disease in his/her lifetime. Moreover, the chances are 37.5% and 18.3 respectively [2] for having 2 risk factors. In Bangladesh, CAD is responsible for a 17% mortality rate [3]. The regular diagnostic approach of CAD relies on coronary angiogram test [4], echo-cardiogram ram (ECG) [5, 6], nuclear scan test and exercise stress test. ECG and exercise stress do not produce sustainable results for CAD prediction due to their non-invasiveness properties and numerous biases. Moreover, walking on a trade mill in a stress test makes the patient discomfort heart function than normal condition. Nowadays, support vector machine (SVM) [7, 8] and artificial neural network (ANN) [5, 8–17] based Clinical Decision Support Systems (CDSS) [18–20] are developed for CAD prediction. Unfortunately, SVM and ANN have no direct impact on the reasoning process due to their black-box- type modeling approaches. As a result, the degree of significance of individual factors cannot be resolved. So, human judgment and clinical data are both two essential factors for CAD diagnosis. For this purpose, CDSS combines both historical data and doctors' domain specific knowledge. But clinical data, like clinical domain knowledge, signs, and symptom, contain various uncertainties [21–23], and pose challenges for selecting domain knowledge to construct knowledge base. Moreover finding the reasoning under uncertainty requires an excellent computational algorithm. To mitigate the challenges, researches introduced different CDSSs, based on the fuzzy interface system and Bayesian interface system, which also has limitations [10, 24–26]. In this paper, the proposed expert system can predict CAD by five classifications according to the severity. They are as follows:

Class A: (Normal or zero sign of heart disease).

Class B: (Unstable angina) - when new symptoms are introduced beside regular stable angina, and appears frequently (mostly when at rest), last long with more severity, and can lead to a heart attack. It can be treated with oral medications (such as nitroglycerine).

Class C: (Non-ST segment elevation myocardial infarction) - echocardiogram does not indicate the symptom of this type of myocardial infarction (MI) but chemical markers in the blood show the damage of heart muscles. The damage may not be significant and artery blockages are usually partial or temporary.

Class D: (ST-segment elevation myocardial infarction) - this type of MI is occurred quickly due to sudden blockage by blood clogging. It can be detected by ECG and chemical markers in the blood, and causes damage to vast heart muscles.

Class E: (Silent ischemia)- Patient with heart disease can be suffered from a sudden heart attack (called silent ischemia) without any prior or early warning, and diabetic patients are common victims of this type [1].

2 Related Research

Researchers recently worked on machine learning and rule-based systems for different purposes [30–33]. Many researchers developed belief-rule-based interference methodology by using evidential reasoning (RIMER) for CAD diagnosis [18, 27]. The RIMER

process uses belief-rule-base for modeling clinical domain knowledge, and applies an evidential reasoning approach for implementing reasoning. Studies show that RIMER based clinical decision support systems are highly efficient in supporting and interacting with clinical domain knowledge under uncertainty. In [28], Multi-Criteria Decision Making Methods were presented for accessing CAD under uncertainty where presence and absence of CAD is predicted through using symptom and signs of CAD. But these approaches report neither the number of blocked arteries nor the significance of severity of the disease [8, 16, 26, 28, 29]. Weak parameters, like signs and symptoms, are used for predicting CAD as well as for predicting the similar types of diseases like mitral regurgitation, dilated cardiomyopathy, congenital heart disease, hyper-tropic cardiomyopathy, myocardial infarction etc. Some researchers developed the Medical Decision Support System (MDSS) to predict CAD. Other proposer polygenic risk scores (PRS), a nonlinear, for CAD prediction with accuracy an 0.92 under the receiver operating curve (AUC) [8].

Experimental analysis reveals that CAD diagnosis and its severity can be predicted significantly through clinical features along with pathological and demographic features [23, 25, 26, 28]. In this paper, we consider all these parameters, and proposed a cooperative-belief-rule based prototype (CDSS) to assist doctors for CAD analysis under uncertainty.

3 Methodology

3.1 Proposed Rule Based Expert System for CAD

In this paper, five separate BRBs are developed based on five distinct feature sets of patients such as i) patients’ pathological features, ii) patients’ physiological features, iii) patients’ demographic features, iv) patients’ behavioral features, and v) patients’ non-modifiable risk factors. The BRBs are as follows:

$$D_A = f_A(S, P_A) \tag{1}$$

$$D_B = f_B(T, P_B) \tag{2}$$

$$D_C = f_C(X, P_C) \tag{3}$$

$$D_D = f_D(Y, P_D) \tag{4}$$

$$D_E = f_E(Z, P_E) \tag{5}$$

Here, $S = \{a_1, a_2, \dots, a_l\}$, $T = \{b_1, b_2, \dots, b_m\}$, $X = \{c_1, c_2, \dots, c_n\}$, $Y = \{d_1, d_2, \dots, d_o\}$, $Z = \{e_1, e_2, \dots, e_p\}$ represent the demographic, physiological, clinical, behavioral, and Non-modifiable features respectively (where $l, m, n, o,$ and p indicate attributes’ number for factors).

Suppose that P_A, P_B, P_C, P_D and P_E are the corresponding vectors for the five BRBs, and $\omega = [\omega_1, \omega_2, \omega_3, \omega_4, \omega_5]$ represent the weight coefficients to the relative BRB where

f_A, f_B, f_C, f_D and f_E functions are for demographic, physiological, clinical, behavioral, and non-modifiable factors. To calculate the individual matching degree for each rule, the following equation is used:

$$\alpha_{i,j} = \frac{u(A_{i,j+1}) - x_i}{u(A_{i,j+1}) - u(A_{i,j})} \tag{6}$$

Where u is utility value, a_{ij} is individual matching degree, A_{ij} is j^{th} referential value for i^{th} attribute, and x_i is the input for i^{th} antecedent (Fig. 1).

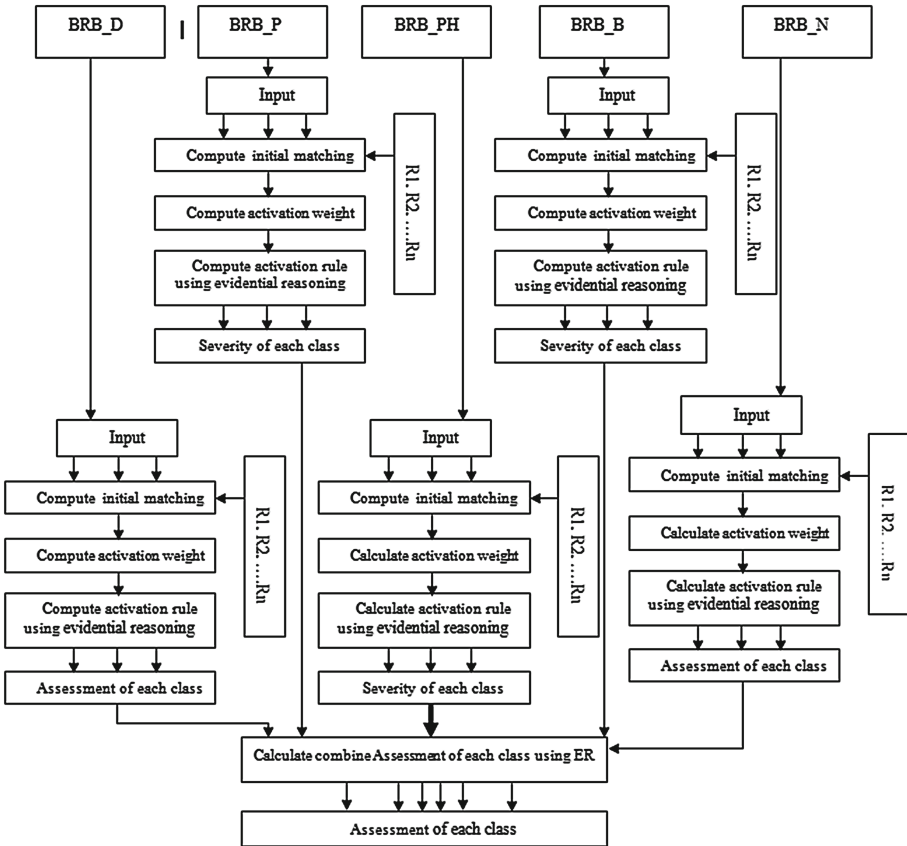


Fig. 1. Rule Based Expert System to assess CAD

To calculate activated weight to each rule the following equation is used:

$$w_k = \frac{\theta_k \alpha_k}{\sum_{i=1}^L \theta_i \alpha_i} \tag{7}$$

Where w_k is the k^{th} rule’s activation weight and a_k is the interrelation between attributes. To calculate a_k , the following equations is used:

$$\alpha_k = \prod_{i=1}^M (\alpha_i^k)^{\bar{\delta}_i} \tag{8}$$

$$\bar{\delta}_i = \frac{\delta_i}{\max_{i=1, \dots, M} (\delta_i)} \tag{9}$$

Where $\bar{\delta}_i$ is the antecedent weight and α_i^k represents individual matching degrees for i^{th} attribute. Five separate BRBs to predict CAD are BRB_P, BRB_PH, BRB_D, BRB_B, and BRB_N. BRB_PH considers physiological factors like blood pressure and stress. BRB_P considers pathological factors like blood sugar level, low density lipoprotein, and triglyceride level. BRB_D considers factors like age and body mass index. BRB_B considers behavior factors like diet, smoking, and physical activities. BRB_N considers non-modifiable risk factors like gender, family history, and residential Area.

3.2 Uncertainties in the Attribute

Attributes like blood pressure, stress, blood sugar, lipoprotein, triglyceride, age, body mass index, unhealthy diet, smoking, family history, and race are categorized into five classes, namely Physiological, Pathological, Demographical, Behavioral, and Non-modifiable risk factors. All the attributes have uncertainties at some level except gender attribute (Table 1).

Table 1. Uncertainties in the attributes

Attributes	Types of uncertainty	Description
Blood Pressure, Stress, Blood sugar, Triglyceride, Lipoprotein	Impression	Information on the attributes is collected through medical instrumentations. The chances are high for storing wrong data in patient profile due to instrumental malfunctioning or operators’ wrong procedural approaches. Moreover, the data for the same patient may vary in different conditions. For example, blood pressure rises after some physical activities, blood sugar level falls after fasting and rises after a meal. So, uncertainty in data may exist due to dataset in different physical conditions or instrumental malfunctions

(continued)

Table 1. (continued)

Attributes	Types of uncertainty	Description
Age	Inconsistency	Old people may face hurdles of remembering their actual age. In underdeveloped countries, the young generation tends to hide actual the age to apply for the Government job
Body mass index	Inconsistency	Patients' weight varies due to clothes, shoes, and after having meals. In this case, wrist measurement is more accurate which are maximum 35 in. for female and 40 in. for male
Unhealthy diet	Vagueness, Incompleteness	Standard calorie demand varies from 1800 to 3000 calories and depends on height, weight, age, and physical activities. The exact amount of food consumption is very hard to measure, and an accurate amount of calorie consumption may be unknown to patients
Smoking	Incompleteness	Patients try to hide their smoking habit as they feel uncomfortable disclosing it in front of their family members
Family history	Vagueness, Incompleteness	Most of the time, parents' heart disease information is not available. Even if some information available, the age, at which they got the disease, cannot be determined
Race	Imprecision, Incompleteness	Usually patients use two or more addresses like present address, mailing address, and permanent address

3.3 Explanation of Antecedent Attributes

Five different types of attributes have been considered in this research. Explanations of the numerical values of each attribute are as follows:

Physiological Factor

Blood Pressure (BP)

The blood pressure which creates heartbeats is known as BP. For BP, several numerical points namely Usual, Elevated, Hypertension Stage 1, Hypertension Stage 2, Hypertension Stage 3 are reflected and shown in Table 2.

Table 2. Numerical values for blood pressure

Terms	Numerical values mm Hg (upper/lower)
Usual (U)	Less than 120/80
Elevated (E)	120/80–129/80
Hypertension Stage 1(H1)	130/89–139/89
Hypertension Stage 2 (H2)	140/90 or higher
Hypertensive Stage 3(H3)	Higher than 180/120

Here, the referential numerical points are presented as in the Eq. (10).

$$PH1 \in \{U, E, H1, H2, H3\} \quad (10)$$

Stress Score (SS)

Intermediate risk of heart problems can be expressed in the SS score. It can be distributed into several referential numerical points, namely regular, mildly irregular, moderately irregular and severely irregular are shown in Table 3.

Table 3. Numerical value for stress score

Terms	Numerical points
Regular (R)	below 4
Insignificantly Irregular (M)	4–8
Significantly Irregular (SI)	9–13
Extremely Irregular (ES)	Higher than 13

The referential numerical values can be presented as in Eq. (11).

$$PH2 \in \{R, M, MI, S\} \quad (11)$$

Pathological Factor

Blood Sugar Level. It is the amount of sugar in the blood. Five referential points are shown in Table 4 and expressed by the Eq. (12).

Table 4. Numerical points for blood sugar

Terms	Amount shown (mg/dL)
Fasting (FA)	Less than 100
Before Meal (BM)	70–130
After meal (1–2 h) (A)	Less than 180
Before Exercise (B)	If taking insulin, at least 100
Bed Time (BT)	100–140

$$P1 \in \{FA, BM, A, B, BT\} \tag{12}$$

Its measures triglycerides amount in blood. Four referential points are described in Table 5 and Eq. (13).

Table 5. Numerical point for triglyceride

Terms	Triglyceride level (mg/dL)
Healthy (H)	Up to 150
Marginal high (BH)	151–200
High (H)	201–500
Extremely high (EH)	above 500

$$P2 \in \{H, BH, H, EH\} \tag{13}$$

Low Density Lipoprotein (LDL)

It contains both lipid and protein, and carries cholesterol to body tissues. Five referential values related to LDL are shown in Table 6 and Eq. (14).

$$P3 \in \{D, NO, BH, H, VH\} \tag{14}$$

Table 6. Numerical point for LDL

Terms	Numerical points (mg/dL)
Good (D)	Less than 100; below 70 if CAD is present
Moderately elevated (NO)	100–140
Near to high (BH)	141–160
High (H)	161–190
Extremely high (VH)	Above 190

Demographic Factor

Age

Older people are more likely to be victims of coronary artery disease, especially, after the age of 65 years. Usually, aged people have higher chance of getting CAD (Table 7).

Table 7. Numerical values for age

Terms	Numerical values (age)
Young (Y)	<35
Middle age (M)	35–49
Old (O)	50–65
Extreme (E)	>65

Four referential values, namely, young (< 35 years), mature (35–49 years), old (50–65), and extremely old (E), have been considered in the following equation from the above table.

$$D1 \in \{Y, M, O, E\} \quad (15)$$

Body Mass Index (BMI)

It indicates the amount of fat ratio. It is applicable for the age range from 18 to 65. It is the ratio of weight to height (Table 8).

Four referential values, namely, healthy weight (18.5–24.9), overweight (25–29.9), obese (30–39.9), and morbidly obese (≥ 40), have been considered in the following equation from the above table.

$$D2 \in \{H, O, OB, MO\} \quad (16)$$

Table 8. Numerical points for BMI

Terms	Numerical points (BMI)
Healthy weight (H)	18.5–24.9
Overweight (O)	25–29.9
Obese (OB)	30–39.9
Morbid Obese (MO)	> = 40

Behavior

Unhealthy Diet

Mediterranean diet can reduce the risk of CAD by 30%. It is mainly plant based food and categorized into four sections shown in Table 10 and expressed by Eq. (17) (Table 9).

Table 9. Numerical values of diet

Terms	Numerical points (Calories/ day)
Low (L)	<1800
Healthy (H)	1800–2200
Moderate (M)	2200–2800
Eating Disorder (ED)	>2800

$$B1 \in \{L, H, M, ED\} \tag{17}$$

Smoking

Smokers or exposers to smoke have a high risk of CAD. Smoking is categorized into four sections, and shown in Table 10 and expressed by Eq. (18).

Table 10. Numerical values for smoking

Terms	Numerical points (Cigarettes/ day)
Non-Smoker (NS)	0
Smoker (S)	1–5
Moderate Smoker (MS)	6–20
Chain Smoker (CS)	>20

$$B2 \in \{NS, S, MS, CS\} \tag{18}$$

Physical Activities

Inactive and less active people are at high risk to develop CAD. Physical activities are categorized into four sections which are shown in Table 11 and expressed by Eq. (19).

Table 11. Numerical values for physical activities

Terms	Numerical points (Minutes/ day)
Inactive (I)	0–10
Less Active (LA)	11–20
Active (A)	21–30
Very Active (VA)	>30

$$B3 \in \{I, LA, A, VA\} \tag{19}$$

Non-modifiable risk Factors

Gender

Male has higher risk of CAD than female. Besides, male suffers from CAD in earlier age than female. But after the age of 70 years, both males and females have similar chances of getting heart disease (Table 12).

Table 12. Numerical values for gender

Terms	Numerical points
Male (M)	0
Female (F)	1
Other (O)	2

$$N1 \in \{M, F, O\} \tag{20}$$

Family History

If parents have histories of heart disease, children have a high risk of developing CAD. The risk is even higher if parents have suffered before early 50 years of age. The numerical points for the family history are represented by 0 (No history of parent’s heart disease), 1 (History of parent’s heart disease), and 2 (History of parent’s heart disease before age of 50), and expressed in Table 13 and by Eq. (21).

$$N2 \in \{L, M, H\} \tag{21}$$

Table 13. Numerical values for family history

Terms	Numerical points
Low (L)	0
Medium (M)	1
High (H)	2

Residential Area

People from mega-cities are more prone to CAD. This is because of a higher rate of diabetes and obesity. On the other hand, people from hill track areas are less likely to develop heart disease. The numerical points for the residential areas are 0(Mega City), (Rural Area), and 2 (Hill track area), and expressed in Table 14 and by Eq. (22).

Table 14. Numerical values for residential area

Terms	Numerical points
Low (L)	0
Medium (M)	1
High (H)	2

$$N3 \in \{L, M, H\} \tag{22}$$

3.4 Rule Base

All attributes from Eqs. (10) to (22) are applied as input variables to predict the CAD class. Sub rules 1 to 20 are expressed in Table 15 for the two Physiological factors from Eqs. (10) and (11).

A sub rule of the CAD can be shown as:

R3: IF blood pressure is usual AND stress score is significantly irregular THEN Overall Prediction is

{(Stage 1, 0.6), (Stage 2, 0.3), (Stage 3, 0.1), (Stage 4, 0.0), (Stage 5, 0.0)}

In the R3, the antecedent attributes are and the consequence attributes are. The rule shows that patient with usual blood pressure and significantly irregular stress score has the probability of developing CAD are (Stage 1 is 60%), (Stage 2 is 30%), (Stage 3 is 10%), (Stage 4 is 0%), (Stage 5 is 0%). The summation of all referential values for R3 is (0.6 + 0.3 + 0.1 + 0.0 + 0.0 =) 1. If the summation of all referential values for a particular rule is 1, we can say that the rule is competed. For some missing attributes or ignorance, the summation may be less than 1 and the rule is incomplete [34].

Table 15. Attributes' sub rule-base for physiological factors

Rule no.	IF		THEN				
	PH1	PH2	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
R1	U	R	1.0	0.0	0.0	0.0	0.0
R2	U	M	0.9	0.1	0.0	0.0	0.0
R3	U	SI	0.6	0.3	0.1	0.0	0.0
R4	U	ES	0.5	0.3	0.2	0.1	0.0
R5	E	R	0.8	0.2	0.0	0.0	0.0
.....
R18	H3	M				0.2	0.8
R19	H3	SI	0.0	0.0	0.0	0.1	0.9
R20	H3	EI	0.0	0.0	0.0	0.0	1.0

3.5 Data Set Description

Dataset was collected from the National Heart Foundation, Bangladesh with proper authorization. The data set description is shown in Table 16.

Table 16. Summary of patients' data

Patient information	Number
Total patients	1100
Age interval in year	30–95
Average age	59
Ratio (Male: Female)	711:389
Class A (Normal)	450
Class B (Unstable angina)	300
Class C (Non-ST segment)	287
Class D (ST segment)	43
Class E (silent Ischemia)	20

4 Result and Discussion

In the binary diagnostic test, a positive or negative diagnosis is made for each patient. When the result of diagnosis is compared to the true condition, we find four possible outcomes: true positive, true negative, false positive, false negative.

4.1 Success Rate

It is the ratio of correctly identified patient' numbers and total patients. Equation (23) is used to calculate the success rate and average success rate.

$$\text{Success Rate} = \frac{\text{Number of correctly identified patients}}{\text{Total patients}} \times 100\% \quad (23)$$

4.2 Error Rate

It is the ratio of incorrectly identified patients' numbers and total patients. Equation (24) is used to calculate the error rate and average error rate.

$$\text{Error Rate} = \frac{\text{Number of patients' incorrectly identified}}{\text{Total number of patients}} * 100\% \quad (24)$$

4.3 Failure Rate

It is the ratio of the number of non-recognized patients to total patients. Equation (25) is used to calculate the failure rate and average failure rate.

$$\text{Failure Rate} = \frac{\text{Number of non recognised patients}}{\text{Total patients}} \times 100\% \quad (25)$$

4.4 False Omission Rate (fOR)

False omission rate is the ratio the of number of patients identified to a class A to the total number of patients belongs to a particular class except for class A. Eq. (26) is used to calculate the false omission rate and average false omission rate.

$$FOR = \frac{\text{Number of patients' identified to a class A}}{\text{Total number of patients belong to a particular class except class A}} \times 100\% \quad (26)$$

Table 17 explains the results obtained by the equation number (23), (24), (25), and (26). Class A is considered as CAD negative patients and the remaining classes are CAD positive patients. It is observed that the success rate of predicting class C type heart disease is the highest (94.08%) among five classes. On the other hand, the success rate of class E prediction is the lowest (50% only). Class E is very hard to predict as most of the time it does show any symptoms.

Table 17. Success rate, failure rate, error rate, false omission rate by expert system

Class	Total No. of Patients (True Condition)	Total No. of Patients (Test Condition)					Total Non-Recognized Patients	False Omission Rate (%)	Error rate (%)	Failure Rate (%)	Success Rate (%)
		A	B	C	D	E					
A	450	405	7	3	0	25	10	–	07.78	10.00	90.00
B	300	15	268	10	0	07	10	05.00	08.33	10.67	89.33
C	287	1	07	270	10	0	6	00.35	03.83	05.92	94.08
D	43	2	03	02	36	0	0	04.65	16.27	16.27	83.72
E	20	05	0	0	0	10	05	25.00	25.00	50.00	50.00
Average Rate (%)								03.54	08.81	11.10	89.90

5 Conclusion

Heart disease is one of the major threats to public health and the reason for the main cause of death worldwide. Although numerous researches are carried out in this area, still there are challenges to diagnose CAD for treatment. In this paper, the proposed expert system results in an average accuracy rate of 89.90% which is the highest among other existing CDSS. The average false omission rate (3.54%) is also the lowest in this system than that of other CDSS. Our test results satisfy one of the main goals of this research. The average failure rate (11.10%) and average error rate (8.81%) also remain as marginal. Class E (silent Ischemia) success rate is the lowest among all classes. The reason is that Class E occurs suddenly without showing any warning signs of heart problems. It was noted that Class E is common to people with diabetes. It requires further research work to investigate whether or not diabetes influences the success rate in Class E type patients. Apart from this, our research concludes that RBES has a higher success rate and false omission rate than other existing CDSS.

References

1. Coronary Artery Disease. <https://my.clevelandclinic.org/health/diseases>. Accessed 21 Nov 2016
2. Benjamin, E.J., et al.: Heart disease and stroke statistics-2017 update: a report from the American Heart Association. *Circulation* **135**, e146–e603 (2017)
3. Islam, A.M., Mohibullah, A., Paul, T.J.B.H.J.: Cardiovascular disease in Bangladesh: a review. *Heart Journal*, **31**, 80–99 (2016)
4. Perez Ruiz De Garibay, A., Kellum, J., Honigschnabel, J., Kreymann, B.: Respiratory and metabolic acidosis correction with the advanced organ Support system. *Intensive Care Medicine Experimental* **7**, 56 (2019)
5. Shadmi, R., Mazo, V., Bregman-Amitai, O., Elnekave, E.: IEEE: Fully-convolutional deep-learning based system for coronary calcium score prediction from non-contrast chest ct. In: 2018 IEEE 15th International Symposium on Biomedical Imaging, pp. 24–28 (2018)

6. Kirsch, J., Buitrago, I., Mohammed, T.L.H., Gao, T.M., Asher, C.R., Novaro, G.M.: Detection of coronary calcium during standard chest computed tomography correlates with multi-detector computed tomography coronary artery calcium score. *Int. J. Cardiovascular Imag.* **28**, 1249–1256 (2012)
7. De Vijver, R.V., et al.: In-field detection of *Alternaria solani* in potato crops using hyperspectral imaging. *Comput. Electron. Agric.* **168**, 105106 (2020)
8. Gola, D., Erdmann, J., Muller-Myhsok, B., Schunkert, H., Konig, I.R.: Polygenic risk scores outperform machine learning methods in predicting coronary artery disease status. *Genetic Epidemiol.* **44**, 125–138 (2020)
9. Weber, M., et al.: Automatic identification of crossovers in cryo-EM images of murine amyloid protein A fibrils with machine learning. *J. Microscopy* **277**, 12–22 (2020)
10. Tang, X.L., Xiao, M.Q., Liang, Y.J., Zhu, H.Z., Li, J.F.: Online updating belief-rule-base using Bayesian estimation. *Knowledge-Based Syst.* **171**, 93–105 (2019)
11. Penson, A., et al.: Development of genome-derived tumor type prediction to inform clinical cancer care. *JAMA Oncol.* **6**, 84–91 (2020)
12. Penafiel, S., et al.: IEEE: Associating risks of getting strokes with data from health checkup records using dempster-shafer theory. In: 2018 20th International Conference on Advanced Communication Technology, pp. 239–246 (2018)
13. Khanna, N.N., et al.: Nonlinear model for the carotid artery disease 10-year risk prediction by fusing conventional cardiovascular factors to carotid ultrasound image phenotypes: A Japanese diabetes cohort study. *Echocardiography-A J. Cardiovascular Ultrasound Allied Techniques* **36**, 345–361 (2019)
14. Hyer, J.M., et al.: Can we improve prediction of adverse surgical outcomes? development of a surgical complexity score using a novel machine learning technique. *J. Am. College Surgeons*, pp. 43–52 (2020)
15. Chicco, D., Jurman, G.: Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. *BMC Med. Inform. Decis. Mak.* **20**, 16 (2020)
16. Apostolopoulos, I.D., Groumpos, P.P., Apostolopoulos, D.I.: A medical decision support system for The Prediction Of The Coronary Artery Disease Using Fuzzy Cognitive Maps. In: Kravets, A., Shcherbakov, M., Kultsova, M., Groumpos, P. (eds.) *Creativity in Intelligent Technologies and Data Science*, vol. 754, pp. 269–283 (2017)
17. Aggarwal, S., Azad, V., Bhattacharyya, S., De, S., Pan, I., Dutta, P.: A Hybrid System Based on FMM and MLP to Diagnose Heart Disease (2017)
18. Zhou, Z.-G., et al.: A cooperative belief rule based decision support system for lymph node metastasis diagnosis in gastric cancer. *Knowledge-Based Syst.* **85**, 62–70 (2015)
19. Zhou, Z.-G., et al.: A bi-level belief rule based decision support system for diagnosis of lymph node metastasis in gastric cancer. *Knowledge-Based Syst.* **54**, 128–136 (2013)
20. Christo, V.R.E., Nehemiah, H.K., Brighty, J., Kannan, A.: Feature selection and instance selection from clinical datasets using co-operative co-evolution and classification using random forest. *IETE Journal of Research*
21. Long, N.C., Meesad, P., Unger, H.: A highly accurate firefly based algorithm for heart disease prediction. *Expert Syst. Appl.* **42**, 8221–8231 (2015)
22. Li, G.L., Zhou, Z.J., Hu, C.H., Chang, L.L., Zhang, H.T., Yu, C.Q.: An optimal safety assessment model for complex systems considering correlation and redundancy. *Int. J. Approximate Reason.* **104**, 38–56 (2019)
23. Hossain, M.S., Ahmed, F., Fatema Tuj, J., Andersson, K.: A belief rule based expert system to assess tuberculosis under uncertainty. *J. Med. Syst.* **41**, 43 (2017)
24. Sparapani, R., Logan, B.R., McCulloch, R.E., Laud, P.W.: Nonparametric competing risks analysis using bayesian additive regression trees. *Stat. Methods Med. Res.* **29**, 57–77 (2020)

25. Onisko, A., Druzdzal, M.J., Austin, R.M.: Application of Bayesian network modeling to pathology informatics. *Diagnostic Cytopathol.* **47**, 41–47 (2019)
26. Gupta, A., et al.: Probabilistic graphical modeling for estimating risk of coronary artery disease: applications of a flexible machine-learning method. *Med. Decis. Mak.* **39**, 1032–1044 (2019)
27. Wang, Y., et al.: Reliability assessment model for industrial control system based on belief rule base. *Int. J. Comput. Commun. Control* **14**, 419–436 (2019)
28. Kishore, A.H.N., Jayanthi, V.E.: Multi criteria decision making methods to predict the prevalence of coronary artery disease. *J. Med. Imag. Health Inform.* **8**, 719–726 (2018)
29. Sofian, H., Ming, J.T.C., Mohamad, S., Noor, N.M.: Calcification detection using deep structured learning in intravascular ultrasound image for coronary artery disease. In: 2018 2nd International Conference on BioSignal Analysis, Processing and Systems (ICBAPS), pp. 47–52 (2018)
30. Sarker, I.H.: Context-aware rule learning from smartphone data: survey, challenges and future directions. *J. Big Data* **6**(1), 1–25 (2019). <https://doi.org/10.1186/s40537-019-0258-4>
31. Sarker, I.H., Colman, A., Han, J.: RecencyMiner: mining recency-based personalized behavior from contextual smartphone data. *J. Big Data* **6**(1), 1–21 (2019). <https://doi.org/10.1186/s40537-019-0211-6>
32. Sarker, I.H., Colman, A., Han, J., Khan, A.I., Abushark, Y.B., Salah, K.: BehavDT: a behavioral decision tree learning to build user-centric context-aware predictive model. *Mobile Netw. Appl.* **25**, 1–11 (2019). <https://doi.org/10.1007/s11036-019-01443-z>
33. Sarker, I.H., Kayes, A.S.M., Watters, P.: Effectiveness analysis of machine learning classification models for predicting personalized context-aware smartphone usage. *J. Big Data* **6**(1), 1–28 (2019). <https://doi.org/10.1186/s40537-019-0219-y>
34. Hossain, S., et al.: A belief rule based expert system to predict student performance under uncertainty. In: 2019 22nd International Conference on Computer and Information Technology (ICCIT), pp. 1–6. IEEE (2019)