A noninvasive method for coronary artery diseases diagnosis using a clinically-interpretable fuzzy rule-based system

Hamid Reza Marateb, Sobhan Goudarzi
Department of Biomedical Engineering, Faculty of Engineering, University of Isfahan, Isfahan, Iran

Background: Coronary heart diseases/coronary artery diseases (CHDs/CAD), the most common form of cardiovascular disease (CVD), are a major cause for death and disability in developing/developed countries. CAD risk factors could be detected by physicians to prevent the CAD occurrence in the near future. Invasive coronary angiography, a current diagnosis method, is costly and associated with morbidity and mortality in CAD patients. The aim of this study was to design a computer-based noninvasive CAD diagnosis system with clinically interpretable rules.

Materials and Methods: In this study, the Cleveland CAD dataset from the University of California UCI (Irvine) was used. The interval-scale variables were discretized, with cut points taken from the literature. A fuzzy rule-based system was then formulated based on a neuro-fuzzy classifier (NFC) whose learning procedure was speeded up by the scaled conjugate gradient algorithm. Two feature selection (FS) methods, multiple logistic regression (MLR) and sequential FS, were used to reduce the required attributes. The performance of the NFC (without/with FS) was then assessed in a hold-out validation framework. Further cross-validation was performed on the best classifier.

Results: In this dataset, 16 complete attributes along with the binary CHD diagnosis (gold standard) for 272 subjects (68% male) were analyzed. MLR + NFC showed the best performance. Its overall sensitivity, specificity, accuracy, type I error (a) and statistical power were 79%, 89%, 84%, 0.1 and 79%, respectively. The selected features were “age and ST/heart rate slope categories,” “exercise-induced angina status,” fluoroscopy, and thallium-201 stress scintigraphy results.

Conclusion: The proposed method showed “substantial agreement” with the gold standard. This algorithm is thus, a promising tool for screening CAD patients.

Key words: Classification, clinical prediction rule, coronary artery disease, data mining, fuzzy logic

INTRODUCTION

Coronary heart disease also known as coronary artery disease (CAD) is a chronic disease in which the coronary arteries, responsible for transporting oxygenated blood to heart muscles, get narrowed and are not able to convey enough fresh blood to this blood-pumping organ.[1,2] Narrowing of blood vessels is usually due to atherosclerosis, a common arterial disease in which increased areas of degeneration and cholesterol (CHOL) deposit plaques form on the inner surfaces of the arteries blocking blood flow.[3,4] In case of reduced blood supply, the heart does not receive enough oxygen and nutrition to operate properly resulting in angina pectoris and heart attack. The symptomatic or asymptomatic reduction in coronary artery flow, may occur with exercise or at rest, and may end up with a myocardial infarction, depending on the severity of obstruction and the speed of its development.[5,6] CAD, the most common form of cardiovascular disease (CVD), has the prevalence of 6.9% in men and 6% in women.[7]

Despite the background of increasing health care expenses, CAD has a significant influence on global economics as a principal cause of disability and loss of efficiency.[8] CAD is a major cause of death and disability in developed countries. Although CAD mortality rates have dropped over the past four decades, CAD remains responsible for about one-third of all deaths in people over age 35.[6,7,9]

According to the World Health Organization, more than 60% of the global burden of the CAD occurs in developing countries.[10] CAD is the leading cause of death worldwide: 3.8 million men and 3.4 million women each year.[11] Overall, the prevalence of CAD in Iran (≥20 years old) was reported within 1.1-36.0% in different studies.[12] The age-adjusted (≥30 years old) prevalence of CAD was reported as 21.8% (22.3% in women and 18.8% in men) in Tehran, Iran.[13] The prevalence of CAD among people aged 35-79 years was reported to be 19.4% (21.9% in women and 16.0% in men) in an urban population in

Address for correspondence: Dr. Hamid Reza Marateb, Department of Biomedical Engineering, Faculty of Engineering, University of Isfahan, Hezar Jerib St., Isfahan 81744, Isfahan, Iran. E-mail: h.marateb@eng.ui.ac.ir
Received: 06-09-2014; Revised: 28-10-2014; Accepted: 27-01-2015
It is very difficult to define sharp borders
Marateb, et al.: Coronary artery diseases diagnosis
Journal of Research in Medical Sciences | March 2015 |

A large number of classifiers have been
Marateb, et al.: Coronary artery diseases diagnosis
Journal of Research in Medical Sciences | March 2015 |

Thus,
Marateb, et al.: Coronary artery diseases diagnosis
Journal of Research in Medical Sciences | March 2015 |

Also, fuzzy logic
Marateb, et al.: Coronary artery diseases diagnosis
Journal of Research in Medical Sciences | March 2015 |

Such approaches can also create clinically interpretable rules that take into account
Marateb, et al.: Coronary artery diseases diagnosis
Journal of Research in Medical Sciences | March 2015 |

In this paper, a fuzzy rule-based system was designed to diagnose CAD based on a limited set of features recorded noninvasively in a case-control study. The rest of the paper is organized as follows: Information about the study population, recorded features, and the data mining methodology is given in the next section. Then, the results of the classification and the extracted rules are reported in the “results” section and finally, the discussion about the clinical validity of the proposed system and the comparison with other classifiers will be provided in the “discussion” section.

MATERIALS AND METHODS

Experimental methods

In this work, the CAD dataset from the University of California (UCI, Irvine), which is available online (https://archive.ics.uci.edu/ml/datasets/Heart+Disease), taken from the Cleveland Clinic Foundation datasets, is used for training and testing the proposed system. This dataset has been considered as a benchmark for various computer-aided CAD diagnosis systems. This dataset consisted of 303 records with 76 attributes (features).

The experimental protocol can be found elsewhere in details. However, it is briefly mentioned as follows. A number of 303 consecutive patients referred for coronary angiography at the Cleveland Clinic between May 1981 and September 1984 participated in the experiment. No patient had a history or electrocardiographic evidence of prior myocardial infarction or known valvular or cardiomyopathic diseases. The following attributes were collected: Age, gender, resting blood pressure (trestbps), CHOL, fasting blood sugar (FBS), resting electrocardiographic results (restecg), active smoker type, number of cigarettes per day, number of years as a smoker, family history of CAD (famhist), chest pain type (cp). In addition, patients underwent 3 noninvasive
tests as part of a research protocol. The results of these tests (exercise electrocardiogram, thallium scintigraphy and cardiac fluoroscopy) were not interpreted until after the invasive gold standard test (coronary angiograms) had been read. The CAD status was considered when narrowing of at least one of the coronary arteries was more than 50%, shown by angiography. The definition of attributes, their measurement scales and categories in the raw dataset were listed in Table 1.

**Preprocessing**

In our study, discretization was used on the attribute intervals. In machine learning, discretization refers to the process of converting or partitioning continuous attributes to discretized variables. Discretization is typically used as a preprocessing step for machine learning algorithms because it can significantly impact the performance of classification algorithms used for data analysis. Discretization is a technique to find cut points to partition the range into a small number of intervals by maintaining good class consistency. We used a set of cut points taken from the literature, indicating the diagnosis properties of individual attributes as below to create ordinal-scale variables:

1. "age:" "Young adult" (18-35 years old), "middle-aged adults" (36-55 years old), and "older adults" (older than 55 years old);
2. "CHOL:" "Desirable" (<200 mg/dL), "borderline high" (200-239 mg/dL) and "high" (≥240 mg/dL);
3. "Trestbps:" "Low" (90-100 mmHg), "normal" (100-120 mmHg), "prehypertension" (120-139 mmHg), "stage 1 hypertension" (140-159 mmHg) and "stage 2 hypertension" (≥160 mmHg);
4. "Cigs:" Five categories 0, 1-9, 10-19, 20-39, and ≥40;
5. "Thalrest:" The following categories were extracted ("excel't," "good," "above av," "average," "below av" and "poor") based on resting heart rate (HR) chart considering gender, age and resting HR;
6. "Oldpeak" and "thalach:" The ST depression was divided by the maximum exercise HR (ST/HR slope),

<p>| Table 1: The attributes of the raw Cleveland CAD dataset |</p>
<table>
<thead>
<tr>
<th>Attribute</th>
<th>Measurement scale</th>
<th>Definition</th>
<th>Categories*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Interval</td>
<td>Age in years</td>
<td>–</td>
</tr>
<tr>
<td>Gender</td>
<td>Nominal</td>
<td>Sex</td>
<td>Male/female</td>
</tr>
<tr>
<td>Trestbps</td>
<td>Interval</td>
<td>Resting blood pressure (mmHg)</td>
<td>–</td>
</tr>
<tr>
<td>CHOL</td>
<td>Interval</td>
<td>Serum CHOL (mg/dL)</td>
<td>–</td>
</tr>
<tr>
<td>FBS</td>
<td>Nominal</td>
<td>FBS &gt;120 (mg/dL)</td>
<td>True/false</td>
</tr>
<tr>
<td>Restecg</td>
<td>Nominal</td>
<td>Resting electrocardiographic results</td>
<td>(1) Normal; (2) having ST-T wave abnormality (T wave inversions and/or ST elevation or depression of &gt;0.05 mV); (3) probable or definite left ventricular hypertrophy by Estes’ criteria</td>
</tr>
<tr>
<td>Thalrest</td>
<td>Interval</td>
<td>Resting heart rate (bpm)</td>
<td>–</td>
</tr>
<tr>
<td>Smoke</td>
<td>Nominal</td>
<td>Active smoker type</td>
<td>Yes/no</td>
</tr>
<tr>
<td>Cigs</td>
<td>Interval</td>
<td>Number of cigarettes per day</td>
<td>–</td>
</tr>
<tr>
<td>years</td>
<td>Interval</td>
<td>Number of years as a smoker</td>
<td>–</td>
</tr>
<tr>
<td>Famhist</td>
<td>Nominal</td>
<td>Family history of CAD</td>
<td>Yes/no</td>
</tr>
<tr>
<td>Cp**</td>
<td>Nominal</td>
<td>Chest pain type</td>
<td>(1) Typical angina pectoris; (2) atypical angina; (3) nonanginal pain; (4) no pain</td>
</tr>
<tr>
<td>Tpeakbps</td>
<td>Interval</td>
<td>Peak exercise systolic blood pressure (mmHg)</td>
<td>–</td>
</tr>
<tr>
<td>Tpeakbpd</td>
<td>Interval</td>
<td>Peak exercise diastolic blood pressure (mmHg)</td>
<td>–</td>
</tr>
<tr>
<td>Thalach</td>
<td>Interval</td>
<td>Maximum exercise heart rate achieved (bpm)</td>
<td>–</td>
</tr>
<tr>
<td>Exang</td>
<td>Nominal</td>
<td>Exercise-induced angina</td>
<td>Yes/no</td>
</tr>
<tr>
<td>Oldpeak</td>
<td>Interval</td>
<td>ST depression induced by exercise relative to rest</td>
<td>(1) Upsloping; (2) flat; (3) downsloping</td>
</tr>
<tr>
<td>Slope</td>
<td>Ordinal</td>
<td>The slope of the peak exercise ST segment</td>
<td>–</td>
</tr>
<tr>
<td>Ca</td>
<td>Interval</td>
<td>Number of major vessels (0–3) colored by fluoroscopy</td>
<td>–</td>
</tr>
<tr>
<td>Thal***</td>
<td>Nominal</td>
<td>Thallium-201 stress scintigraphy</td>
<td>(3) Normal; (6) fixed defect; (7) reversible defect</td>
</tr>
<tr>
<td>Num</td>
<td>Nominal</td>
<td>Diagnosis of heart disease (angiographic disease status)</td>
<td>(1) Normal: &lt;50% diameter narrowing; (2) CAD &gt;50% diameter narrowing</td>
</tr>
</tbody>
</table>

*The categories were shown for nominal or ordinal features. **(1) Typical angina pectoris: Pain that occurs in the anterior thorax, neck, shoulders, jaw, or arms is precipitated by exertion and relieved within 20 min by rest. (2) Atypical angina: Pain in one of the above locations and either not precipitated by exertion or not relieved by rest within 20 min. (3) Nonanginal pain: Pain not located in any of the above locations, or if so located not related to exertion, and lasting less than 10 s or longer than 30 min. (4) No pain; ***(1) Normal, (2) Fixed abnormality (defects observed during exercise that persisted at redistribution), and (3) Reversible abnormality (defects present during exercise and significantly corrected during redistribution). CAD = Coronary artery disease; CHOL = Cholesterol; FBS = Fasting blood sugar.
also known as HR adjustment of exercise-induced ST segment depression.[62,63] Then, the cut points of 0.01 and 0.02 μV/beat/min were used to derive three ordinal classes: “low,” “medium” and “high.”

**Fuzzy rule-based system**

Rule generation includes both rule extraction and refinement. One of the methods for extracting rules from the input-out data is neuro-fuzzy rule generation.[83] It combines the powerful description of fuzzy classification techniques with the learning capabilities of Neural Networks. In the neuro-fuzzy classification methods, the feature space is partitioned into multiple fuzzy subspaces that are controlled by fuzzy if-then rules. For determining an optimum fuzzy region, the parameters of the fuzzy rules should be optimized.[86] The classifier consists of the following layers: Fuzzy membership, fuzzification, defuzzification, normalization and output.[67] The K-means clustering method was used to obtain the initial parameters and to formulate the fuzzy if-then rules.[65]

A fuzzy classification rule $R_i$, which describes the relation between the input feature space and the classes, can be defined as follows:

$$R_i: \text{If } x_{p1} = \phi_{il} \text{ and } \ldots \text{ and } x_{pj} = \phi_{ij} \text{ and } R \text{ and } x_{pt} = \phi_{in} \text{ then class is } \text{out}_c,$$

where $x_{pj}$ is the $j$th feature of the $p$th sample, $out_k$ is the $k$th class label, and $\phi_{ij}$ is the fuzzy set of the $j$th feature in the $i$th rule.[67,65] In the neuro-fuzzy method used in this study, scaled conjugate gradient (SCG) algorithm was used to speed up the learning procedure in which the parameters of the Takagi–Sugeno–Kang fuzzy inference system (FIS) were tuned in the training set by shortening the training time per iteration. This method was referred to as neuro-fuzzy classifier (NFC) in this manuscript. In this FIS, “and,” “or,” “implication” from the antecedent to the consequent and “aggregation” of the consequents across the rules operators were product, probabilistic or “probor” (algebraic sum), minimum and maximum. Weighted defuzzification was also used. Note that Probor $(x, y) = x + y - x \times y$.

**Feature selection**

In many classification problems, a lot of candidate attributes are used for problem representation. Many of these are usually irrelevant or redundant.[96] Thus, feature selection (FS) is used to detect relevant features usually leading to an increase in classifier accuracy.[70] In this work, two supervised FS methodologies were used: (1) A statistical FS method, multiple logistic regression (MLR).[71] and (2) a deterministic FS approach, sequential FS (SFS).[72] MLR, known as feature vector machine in machine learning, can be used to select statistically significant features.[73] In our study, the intercept point was not used in the MLR. In SFS, the classifier starts with an empty set and added features until the accuracy was no longer improved by adding more features.

**Validation**

The performance of the classifier was assessed using the “hold-out” method, an approach to out-of-sample evaluation, in which the dataset was randomly split into two equal-size mutually exclusive sets (training and test sets). The classifier was trained on the training set and tested on the test set.[74] The performance measures of the classification are listed in Table 2. The following performance measures were used: Sensitivity (Se), Specificity (Sp), Accuracy (Acc), and Precision (Pr) along with other indices reported in Table 2. Additionally, the McNemar’s (Gillick) statistical test was used[75] to compare the performance of the NFC with different configurations (e.g., with or without FS) on the test set. Classification was performed using Matlab, Statistics Toolbox Release 2011a (The MathWorks, Inc., Natick, Massachusetts, USA) and the NFC toolbox.[67,76] All statistical analyses and calculations were performed using the SPSS statistical package, version 18.0 (SPSS Inc., Chicago, IL, USA).

**RESULTS**

The number of subjects analyzed was 272 (68% male) out of total 303 in the Cleveland CAD dataset. The excluding criterion was the existence of missing value in any attribute. The characteristics of the raw dataset in the CAD and normal groups are shown in Table 3.

The performance of the NFC without FS, and with SFS/MLR on the preprocessed (discretized) Cleveland CAD dataset is shown in Table 4, in the hold-out validation framework. The feature set selected using SFS was (“thal,” “ST/HR slope,” “cp,” “ca,” “CHOL,” “trestbps,” “restecg,” “sex,” and “famhist”) from which the training accuracy increased from 72.1% to 88.2% during SFS procedure. Adding more features did not increase the accuracy. MLR, on the other hand, proposed the following five features

<table>
<thead>
<tr>
<th>Table 2: The reported performance measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity (Se) = $\frac{TP}{TP + FN}$</td>
</tr>
<tr>
<td>Accuracy (Acc) = $\frac{TP + TN}{TP + TN + FP + FN}$</td>
</tr>
<tr>
<td>Precision (Pr) = $\frac{TP}{TP + FP}$</td>
</tr>
<tr>
<td>Recall (Rl) = Power (Se) = $\frac{TP}{TP + FN}$</td>
</tr>
<tr>
<td>F1-score = $\frac{2 \times Pr \times Rl}{Pr + Rl}$</td>
</tr>
<tr>
<td>False alarm = $1 - $Sp$ = $False positive rate = $\alpha$ (type I error)</td>
</tr>
<tr>
<td>$\beta$ = $1 - $Se$ = $False negative rate = $\beta$ (type II error)</td>
</tr>
</tbody>
</table>

| $TP$ = True positive; $FN$ = False negative; $FP$ = False positive; $TN$ = True negative |
Table 3: The attributes of the raw Cleveland dataset for normal and CAD groups, along with their categories (percentage) for nominal/ordinal variables and (minimum-maximum) mean ± SD for interval variables.

<table>
<thead>
<tr>
<th>Attribute*</th>
<th>Normal (n = 151)</th>
<th>CAD (n = 121)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (%)</td>
<td>Male (54.8), female (45.2)</td>
<td>Male (84.0), female (16)</td>
</tr>
<tr>
<td>Cp (%)</td>
<td>1 (10.2), 2 (22.3), 3 (42.7), 4 (24.7)</td>
<td>1 (4.8), 2 (6.4), 3 (13.6), 4 (75.2)</td>
</tr>
<tr>
<td>FBS (%)</td>
<td>False (85.4), true (14.6)</td>
<td>False (84.8), true (15.2)</td>
</tr>
<tr>
<td>Famhist (%)</td>
<td>No (42.0), yes (58.0)</td>
<td>No (32.8), yes (67.2)</td>
</tr>
<tr>
<td>Restecg (%)</td>
<td>1 (56.7), 2 (0.6), 3 (42.7)</td>
<td>1 (39.2), 2 (0.8), 3 (60.0)</td>
</tr>
<tr>
<td>Exang (%)</td>
<td>No (85.4), yes (14.6)</td>
<td>No (44.8), yes (55.2)</td>
</tr>
<tr>
<td>Slope (%)</td>
<td>1 (64.3), 2 (30.6), 3 (5.1)</td>
<td>1 (27.2), 2 (64.8), 3 (8.0)</td>
</tr>
<tr>
<td>Thal (%)</td>
<td>1 (79.5), 2 (3.8), 3 (6.7)</td>
<td>1 (28.2), 2 (6.5), 3 (65.3)</td>
</tr>
</tbody>
</table>

*For the definition of attributes, their units and categories, the reader is referred to Table 1. CAD = Coronary artery disease; CHOL = Cholesterol; FBS = Fasting blood sugar; SD = Standard deviation.

The proposed fuzzy rule-based system (MLR + NFC) was chosen as the best classifier and used for further analysis.

The MLR + NFC method that had been tuned on the training set was shown in Figure 2. The system has two fuzzy rules. Each rule was related to an outcome diagnosis class. A diagnosis example was taken from the test set in the fuzzy rules to clarify the decision-making procedure. For a typical subject whose CAD was confirmed with angiography, the discretized attributes were as the following: Older adult age category, Angina was not induced by exercise, two major vessels colored by fluoroscopy, thallium-201 stress scintigraphy test showed reversible defect and ST/HR slope was high. These attributes were used as the input of the fuzzy system. The range of the output is within (0.9, 2.1) with the cut-off value of 1.5 = (0.9 + 2.1)/2. If the output of the fuzzy system is higher or equal than 1.5, the subject has CAD. The output of the fuzzy system (based on the noninvasive tests) is 1.81, in agreement with that result of the gold standard (angiography). The first and second rules in Figure 2 have been extracted for normal and CAD subjects, respectively.
DISCUSSION

Primary prevention, aiming at preventing heart and blood vessel disease in individuals who have not had a heart attack or symptoms of CAD and have no known clinical evidence of CAD, is highly recommended in comparison with secondary prevention for individuals with known CAD. This requires changing the lifestyle by quitting smoking, bringing down high blood pressure, controlling diabetes, maintaining a healthy body weight, doing regular exercises, following a heart-healthy diet to lower CHOL, LDL, and triglycerides, and to raise HDL, reducing stress and limiting alcohol consumption.\cite{8,78}

It has been suggested to integrate the prediction, prevention and intervention programs of CAD for more effective primary/secondary prevention.\cite{17,79} Invasive coronary angiography is the gold standard for establishing the presence, location, and severity of CAD.\cite{80,81} However, this technique is invasive, costly\cite{81} and associated with a small but definite risk of morbidity (1.5%) and mortality (0.15%).\cite{82,83} Therefore, a convenient, noninvasive alternative method for coronary angiography can provide significant clinical and economic benefits for the public health care system.\cite{84}

A number of noninvasive CAD diagnosis methods have been proposed in the literature. Among which, fluoroscopy and stress thallium scintigraphy, are the two most popular. Unfortunately, the diagnostic accuracy of these methods in comparison with that of the gold standard, coronary angiography, ranges between 35% and 75%.\cite{85,86} Thus, the purpose of our work was to design a computer-aided noninvasive CAD diagnosis system using data mining methods to improve the diagnosis accuracy by combining the results of the noninvasive clinical tests and other attributes recorded from the subjects [Table 1].

A number of computer-aided CAD diagnosis systems have been proposed in the literature\cite{47,50,87-93} whose performance was tested on the Cleveland CAD dataset [Table 6]. Most of them use the black-box mathematical methodology that is not acceptable in medicine where the clinical interpretation of the decision-making procedure is critical. Our proposed fuzzy rule-based system, on the other hand, provided interpretable linguistic terms [Figure 2], which could be regarded as a fuzzy version of the decision tree classifier.\cite{94}

For example, the comparison of rule 1 (normal) and rule 2 (CAD) in Figure 2 shows that the higher the ST/HR slope, and age, the higher the risk of CAD. Also, angina induced by exercise increases the risk of having CAD. Also, the last two categories of thallium-201 stress scintigraphy test (fixed and reversible defects) have higher CAD risk. The above rules are clinically acceptable. However, number of major vessels colored by fluoroscopy had a medium effect on the CAD diagnosis that could be due to the small sample size of the study population. It might be related to the fact that the sensitivity of fluoroscopy could be as low as 35% in some cases.\cite{86} and the system learned it from the training set. Meanwhile, the proposed system (MLR + NFC) only requires five input attributes for the diagnosis that is less than.

Table 5: The overall confusion matrix of the MLR + NFC method*

<table>
<thead>
<tr>
<th>MLR + NFC outcome</th>
<th>Patient with CAD confirmed with angiography</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CAD positive</td>
</tr>
<tr>
<td>Test outcome positive</td>
<td>95 (TP)</td>
</tr>
<tr>
<td>Test outcome negative</td>
<td>26 (FN)</td>
</tr>
</tbody>
</table>

*The classifier was trained on the training set and tested on the whole dataset. “Positive” is related to “CAD diagnosis” while “negative” was used for “normal diagnosis”. TP = True positive; FN = False negative; FP = False positive; MLR = Multiple logistic regression; NFC = Neuro-fuzzy classifier; CAD = Coronary artery disease; TN = True negative

Figure 2: The extracted fuzzy rules from the training set. AGE_CAT: Age category, EX_ANG: Exercise-induced angina, VES_FLU: The number of vessels colored by fluoroscopy, THAL_SCI: Thallium-201 stress scintigraphy category, and ST_HR_S: Heart rate adjustment of exercise-induced ST segment depression category. Fuzzy rules 1 and 2 were related to normal and coronary artery diseases (CAD) classes, respectively. The overall output of the fuzzy system (1.81 in this example), is higher or equal to 1.5 indicating that the subject had CAD that was in agreement with what obtained from the gold standard (angiography). For the description of the input feature categories, refer to the section “preprocessing”
The overall accuracy of the proposed FIS (MLR + NFC) was 84% [Table 5]. Since the number of normal subjects in the dataset was higher than CAD [Table 3], the F_{1}-score was also calculated (~82%) since it is an unbiased accuracy measure in unbalanced datasets. The type I error (α) and the power of the proposed diagnosis test were 0.1 and 76%. Reducing α down to 0.05 and increasing power to 80%, improves the clinical reliability of the proposed system that is the focus of our future activity. The following strategies could be taken in this regard:

1. The selected mixed-type features in MLR could be weighted based on the clinical importance and (or) tuned using particle swarm optimization\[95\] embedded with Generalized Minkowski Metrics.\[96\] Tuning the feature weights, might improve the performance of the designed classifier;

2. Although the sample size of the Cleveland dataset is not low, some important features such as body mass index (BMI) are missing.

We are thus going to design an automated CAD risk assessment program, based on the findings of this study, in collaboration with Isfahan Healthy Heart Program\[97\] whose dataset is quite rich.

In addition to the hold-out validation method, we also used a 10-fold cross-validation in which the original sample was randomly partitioned into 10 equal-size subsamples. Of the 10 subsamples, a single subsample was retained as the validation data for testing the model, and the remaining 9 subsamples were used as training data. The cross-validation process was then repeated 10 times (the folds), with each of the 10 subsamples used exactly once as the validation data. The results from the folds were averaged to estimate a single estimation.\[98\] The overall cross-validation accuracy of the proposed classifier was 83%. Guarding against testing hypotheses suggested by the data (type III errors\[99\]) was done by cross-validation. Comparison with other diagnosis methods designed on the Cleveland CAD dataset, our proposed method ranked in the top first quartile [Table 6].

Although the BMI was not used in our model, it is well-known that BMI is positively correlated with CHOL.\[100-102\] In our dataset, age was positively correlated with hypertension, CHOL, and FBS. This might explain that age in the final model could capture information about CHOL, hypertension, FBS, and BMI as well. The number of cigarettes per day and number of years as a smoker were positively correlated with thallium-201 stress scintigraphy categories. Thus, this attribute could capture the smoking information. Gender was associated with the exercise-induced angina status. Thus, information about the gender could be captured by exercise-induced angina. This is, in fact, the property of MLR that takes into account the interaction between attributes and reports the compact attribute set. Accordingly, most of the input risk factors could be taken into accounts directly/indirectly. Among the features selected, the following ones have been selected in another manuscript:\[103\] “ST/HR slope,” “cp,” “age,” “trestbps.”

**CONCLUSION**

We proposed an interpretable fuzzy rule-based system that could noninvasively predict the CAD based on “age,” “exercise-induced angina status,” number of major vessels colored by fluoroscopy, thallium-201 stress scintigraphy result and ST/HR slope. The proposed computer-aided system was promising in CAD diagnosis and could be implemented as a web-based diagnostic decision support system. However, its performance could be improved by introducing weights to the input attributes taking into account the clinical relevance/priority of the features.

---

**Table 6: Comparison of the proposed system outcome with similar research**

<table>
<thead>
<tr>
<th>Author</th>
<th>Method</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DETRANO et al., 1989</td>
<td>Probability theory (logistic regression)</td>
<td>77.00</td>
</tr>
<tr>
<td>STER and DObNIKAR, 1996</td>
<td>Fisher discriminant analysis</td>
<td>84.20</td>
</tr>
<tr>
<td>STER and DObNIKAR, 1996</td>
<td>LDA</td>
<td>84.50</td>
</tr>
<tr>
<td>STER and DObNIKAR, 1996</td>
<td>Naive Bayes</td>
<td>82.50-83.40</td>
</tr>
<tr>
<td>Jankowski and Kadirkamanathan, 1997</td>
<td>IncNet</td>
<td>90.00</td>
</tr>
<tr>
<td>Cheung, 2001</td>
<td>BNNF</td>
<td>80.96</td>
</tr>
<tr>
<td>Cheung, 2001</td>
<td>C.4.5</td>
<td>81.11</td>
</tr>
<tr>
<td>Cheung, 2001</td>
<td>BNND</td>
<td>81.11</td>
</tr>
<tr>
<td>Cheung, 2001</td>
<td>Naive Bayes</td>
<td>81.48</td>
</tr>
<tr>
<td>Polat et al., 2007</td>
<td>Fuzzy-AIRS-Knn based system</td>
<td>87.00</td>
</tr>
<tr>
<td>Kahramanli and Allahverdi, 2008</td>
<td>Hybrid neural network system</td>
<td>86.80</td>
</tr>
<tr>
<td>Resul, 2009</td>
<td>Neural network ensembles</td>
<td>89.01</td>
</tr>
<tr>
<td>Senthil Kumar, 2011</td>
<td>ANFIS</td>
<td>91.18</td>
</tr>
<tr>
<td>Senthil Kumar, 2012</td>
<td>Fuzzy resolution mechanism</td>
<td>91.83</td>
</tr>
<tr>
<td>Muthukaruppan and Er, 2012</td>
<td>PSO based fuzzy expert system</td>
<td>93.27</td>
</tr>
<tr>
<td>Mahmodabdadi and Saniee Abadeh, 2014</td>
<td>Competitive algorithm based fuzzy expert system</td>
<td>94.92</td>
</tr>
<tr>
<td>Proposed classifier</td>
<td>MLR+NFC</td>
<td>84.00</td>
</tr>
</tbody>
</table>

ANFIS = Adaptive-network-based fuzzy inference system; MLR = Multiple logistic regression; NFC = Neuro-fuzzy classifier; PSO = Particle swarm optimization; AIRS = Artificial immune recognition system; LDA = Linear discriminant analysis.
ACKNOWLEDGMENT

This work was supported by the University of Isfahan.

AUTHOR’S CONTRIBUTION

HRM contributed in the conception and design of the work, drafting and revising the draft, approval of the final version of the manuscript, and agreed for all aspects of the work. SG contributed in the conception of the work, revising the draft, approval of the final version of the manuscript, and agreed for all aspects of the work.

REFERENCES


Source of Support: Nil, Conflict of Interest: None declared.