Journal of Computing and Security

January 2017, Volume 4, Number 1 (pp. 13–23)

http://www.jcomsec.org

A GA Approach for Tuning Membership Functions of a Fuzzy Expert System for Heart Disease Prognosis Development Risk

Rana Akhoondi [a] Rahil Hosseini [a,*] Mahdi Mazinani [b]

[a] *Department of Computer Engineering, Shahr-e-Qods branch, Islamic Azad University, Tehran, Iran.*
[b] *Department of Electronic Engineering, Shahr-e-Qods Branch, Islamic Azad University, Tehran, Iran.*

ARTICLE INFO.

Article history:
Received: 09 June 2017
Revised: 05 March 2018
Accepted: 17 March 2018
Published Online: 25 July 2018

Keywords:
FES, Genetic Algorithm, Fuzzy-GA System, HDP.

ABSTRACT

Application of soft computing hybrid models have been concentrated to cope with uncertainty in the medical expert systems, recently. Heart disease is one of the mortal diseases that can be controlled in early stages. In this paper a hybrid Fuzzy-GA model for the Heart Disease Prediction (HDP) problem has been proposed. For this, first a Fuzzy Expert System (FES) using Mamdani model was presented. Then the membership functions parameters of the FES were optimized using the hybrid Fuzzy-Genetic Algorithm (Fuzzy-GA). The reason of selecting fuzzy method was its high potential to address the uncertainty sources in the knowledge of medical experts. Performance of the FES and Fuzzy-GA model were evaluated using a real dataset of 380 patients collected from Parsian Hospital in Karaj, Iran. Accuracy of the designed FES before optimization was 85.52%. After optimization using the hybrid Fuzzy-GA, the accuracy of this system was increased to 92.37%. The proposed hybrid model competes with its counterparts in terms of interpretability and accuracy in prognosis process of the heart disease. This model is promising for early diagnosis of the heart disease and saving more people lives.

1 Introduction

The prognosis and control of disease is much easier and less expensive than their treatment. Controlling the heart disease can prevent sever damage to patients. Heart is an important organ of human body and its disruption may lead to death. Expert systems can assist prognosis of the heart disease and consequently early diagnosis of disease and its survival rates [1, 2]. This paper aims at prediction of the heart disease and improvement of the survival rate of this mortal disease. The prognosis is always simpler, and cheaper than treatment. However several studies have been reported on diagnosis of this disease, the HDP has less been concentrated. The objective is to optimize parameters of the FES using a hybrid fuzzy-Genetic algorithm design a FES for HDP to enhance the accuracy while keeping its interpretability. The role of fuzzy model is to manage the uncertainty in the symptoms and knowledge of experts. It has the capability of computing with words and reasoning in presence of vague and uncertain knowledge of experts [3, 4]. The FES was designed based on Mamdani inference model; the inputs and output of system were considered as fuzzy sets.

A FES suffers from static behavior; *i.e.*, it cannot

* Corresponding author.

Email addresses: rana.akhoondi@gmail.com (R. Akhoondi), rahil.hosseini@qodsiau.ac.ir (R. Hosseini), mahdi.mazinani@qodsiau.ac.ir (M. Mazinani)

https://dx.doi.org/10.22108/jcs.2018.106877.0
ISSN: 2322-4460

learn from the environment. To address this issue, soft computing paradigm suggest to integrate fuzzy models with evolutionary algorithms. Evolutionary computation is simulation of the process of biological evolution of human intelligence and behavioral learning in computer systems for the purpose of optimization. To optimize intelligent models such as neural network and fuzzy systems, the GA has been frequently applied. Between different evolutionary algorithms, GA has been frequently and commonly applied model. This paper focuses on the GA capability for selection and combination of superior genes have been concentrated in the process of evolution and optimization of the FES. For this, GA was combined with the FES to optimize its parameters. Efficiency of a FES can improve through the following processes [5]:

(1) Tuning: It assumes that a fuzzy system including rules and membership functions exist and tries to modify parameters to improve its performance.
(2) Learning: It tries to directly extract the rules, or membership functions from a training dataset.

This study is an attempt toward improving FES dynamic behavior with combining it with the evolutionary nature of the GA during the tuning task. The paper structure is organized as follows: Section 2 presents an overview of related works for prediction of heart disease, in Section 3 the design details of the FES for (HDP) has been explained, followed by the hybrid Fuzzy-GA method for tuning the FES. Experimental results and performance evaluation of the proposed method along with a comparison between the proposed method in this study and related works are presented in Section 4.Then Discussion and conclusion have been drawn in Section 5.

2 Related works and Background

In this section, an overview of related works reported for (HDP) is presented.

A prediction model for heart disease based on hybrid Artificial Neural Network and Gray Wolf Optimization (ANN-GWO) algorithm was designed in [6]. This paper combines two intelligent methods *i.e.*, (GWO) and (ANN) and applied it to the HDP. This algorithm takes advantage of the ANN to search the relation between I/O variables and the GWO to find the optimized weights for the ANN and avoid trapping in the local optima. Efficiency of the hybrid (ANN-GWO) model in comparison with the back-propagation neural network (BPNN) in terms of Root Mean Square Error (RMSE). The results reveals that this model ameliorate the time to convergence of the algorithm and the overall accuracy of the model.

An intelligent hybrid model to assess the risk of the HDP was designed in [7]. The objective of this model was to design a rule set for the HDP through combinations of the Rough Set (RS) theory for addressing uncertainty, and swarm intelligence in the Particle Swarm Optimization (PSO) with learning capability of the ANN and Fuzzy logic (FL). The RS theory was combined with the PSO into the Intelligent Hybrid Quick Reduced PSO (IHQRPSO) algorithm to find a set of optimum features. This feature set was used in the ANN to classify samples into healthy and unhealthy categories. The severity of disease was predicted using the rules of the FL. Accuracy of the IHQRPSO algorithm was 95.25%.

Another system was designed for prediction of heart disease based on hybrid genetic fuzzy model in [8]. The objective of this work was diagnosis of heart disease based on commonly applied soft computing models; *i.e.*, GA and FL. The proposed system takes advantages of the GA and FL to improve accuracy of prediction rates of heart disease in patients. The model accuracy on the UCI heart disease dataset was reported as 86% with the 90% specificity and sensitivity of 80%.

A classification system for healthcare data was presented using Genetic fuzzy method in [9]. This paper integrates fuzzified standard additive model (SAM) with the GA, to address uncertainty and time complexity issues. This method was evaluated using the Wisconsin breast cancer (WBCD) and Cleveland heart disease (CHD) datasets in the UCI repository. Their method was computationally expensive.

A model was proposed for prediction of heart disease based on firefly algorithm in [10]. This paper utilizes rough sets theory and interval type-2 FL system (IT2FLS). Their learning process utilizes the fuzzy C-mean clustering algorithm and adjusting parameter is conducted through chaos theory combined with firefly and the GA.

A system was proposed for classification of Electrocardiogram (ECG) heart beat with a modified algorithm based on Artificial Bee Colony in [11]. It was applied to the ECG signals which was obtained from the MITBIH database. The result was compared to the related classifiers.

A HDP model using a coronary artery disease (CAD) was reported in [12]. This model was optimized through an adaptive Neuro-fuzzy inference system (ANFIS) with combination of a linear discriminant analysis. The prediction rate of the ANFIS-LDA method was 80.2%.

A neural network with back propagation learning was presented to model diagnosis of heart disease in

[13]. The proposed system uses 13 features for heart disease four class prediction model. Their experimental results shows that classification accuracy of 92% on the UCI heart disease dataset.

A classification system was presented for (HDP) using K-Nearest Neighborhood (KNN) and the GA in [14]. In this paper, a hybrid algorithm using combination of the KNN with the GA was proposed. Results shows that this algorithm enhances the diagnosis of the heart disease. The accuracy of this system was equal to 95.73%.

A hybrid PSO and FES was proposed for the diagnosis of CAD in [15] on the Cleveland and Hungarian Heart Disease datasets. The PSO adjusts the parameters of the fuzzy membership functions (MFs). After MFs tuning process, the accuracy of the FES reached to 93.27%. The main advantage of this approach is the capability of interpretation of the FES output compared with related studies.

A FES for CAD prediction was reported in [16]. In this paper, clinical parameters were employed for developing an expert system to predict the CAD (CAD) disease in early stages. The sensitivity of this system was reported 95.85% with the specificity of 83.33%.

A medical diagnosis system was proposed for CD using a GA in [17]. In this paper, a FES was proposed for classifying the patients suffering from cardiovascular diseases. Then the membership functions are genetically tuned. The performance of this system was about 2% more than the other methods.

A prediction model for CAD was presented using data mining tasks in [18]. This paper applies clustering model to define number of clusters. Then, a decision tree and artificial neural network were applied to classify the patients with the CAD. This model error was reported as 0.074. The summery of an analytical review of the related works has been presented in Table 1.

3 Proposed Method

This section presents the design details of the FES for HDP followed by the hybrid Fuzzy-GA approach.

3.1 Design of a FES for HDP

The FES was used for HDP includes the following main components: 1) Inference engine, and 2) knowledge-base; the former performs the reasoning process using and the later includes the fuzzy rules. The details of the FES and its rules for prognosis of the heart disease was reported by the authors in [19]. This section provides a brief review of this system and its improvement are explained.

In this study, the design of the FES was based on our previously reported work in [19]. This study extends the design of this FES for HDP and improves its accuracy through the improved rule extraction method explained in the proposed hybrid Fuzzy-GA method explained in the next sections. The rest of this section briefly explains the design details of the FES. More details are provided in [19]. The rules of the FES was designed using knowledge of the cardiologists. This system includes 6 linguistic variables (5 input variables and one output variable). The MFs of the fuzzy linguistic variable are shown in Table 2.

Gaussian membership function was considered in the FES as shown bellows [19]:

$$f(x,\sigma,c) = e^{\frac{-(x-c)^2}{2\sigma^2}} \quad (1)$$

Where the parameter c and σ^2 are mean and variance, respectively.

3.2 The Process of the Rule Improvement

The rules of the FES can be generated through considering all the combinations of variables and their linguistic terms. However, this way we would end up with a huge number of rules which may deteriorate the overall performance of the FES. One of the methods to extract the most important and effective rules is using the lookup table technique. In this approach, those rules which are compatible with majority of samples are selected. For this, at first each sample is considered as a fuzzy rule and evaluated separately on dataset. Then those fuzzy rules that are fired more by samples in the dataset and covers a greater area of problem state space are selected.

The FES for HDP designed in our previous study [19], has been modified through the rule improvement process using look-up table technique. The improved FES was used for tuning and optimization of the parameters of the MF through the Fuzzy-GA tuning approach. Tuning these parameters can ameliorate the accuracy of FES and improve its interpretability.

3.3 The Proposed Fuzzy-GA Method for Tuning Membership Functions of the FES

In this section, the steps of the hybrid Fuzzy-GA method for tuning membership function parameters in a FES are explained as follows:

(1) **Design the structure of the FIS**: define I/O variables, its granulation and the type of MFs to be used and the fuzzy rules for the FIS.
(2) **Identify the necessary parameters**: For adjusting parameters of the Gaussian MF, the

Table 1. Review of Intelligent Models Applied to (HDP)

Name	Method	Inputs	Outputs	Accuracy (Measure)	Dataset
A prediction system for heart disease [6]	hybrid ANN-GWO algorithm	Clinical parameters	4 classes	0.0036 (RMSE)	UCI (Cleveland)
A mechanism for prediction of heart disease [7]	Combination of rough set with PSO and create the IHQRPSO algorithm	Clinical parameters	3 classes	95.25% (Accuracy)	UCI Machine learning repository
A prediction system for heart disease [8]	Combination of the fuzzy inference system and GA	Clinical parameters	2 classes	86% (Accuracy) 90% (Specificity) 80% (Sensitivity)	UCI Machine learning repository
A classification system for data [9]	fuzzy SAM model with GA	Clinical parameters	5 classes	95.59% (Accuracy) 96% (F- measure)	UCI Machine learning repository
A prediction system for heart disease [10]	RS theory and IT2FL	Rough sets include feature of the variables	The sets that have been reduced some of their features	86% (Accuracy) 78.1% (Sensitivity) 90% (Specificity)	UCI Machine learning repository
A classification system for heart rate [11]	Modified Artificial Bee Colony algorithm	Heart rate	7 clusters relevant to different diseases	99.46% (Sensitivity) 99.37% (Specificity) 99.94% (Positive Productivity)	MITBIH database
A optimization system for diagnosis of heart disease [12]	ANFIS and linear discriminant analysis	Clinical parameters	One crisp number that shows the value of the heart disease risk.	80.2% (Prediction rate)	The dataset includes 8958 records
A prediction system for heart disease [13]	A back propagation neural network	Clinical parameters	4 classes	92% (Accuracy)	UCI (Cleveland)
A classification system for heart disease [14]	Combination of genetic algorithm and KNN	Clinical parameters	8 classes relevant to different heart diseases	95.73% (Accuracy)	UCI Machine learning repository
A system for diagnosis of heart disease [15]	Combination of the fuzzy inference system and PSO algorithm	Clinical parameters	Angiography status	93.27% (Classification Accuracy)	The Cleveland and Hungarian Heart Disease
A FES [16]	Fuzzy system with clinical parameters in inputs	Clinical parameters	The value of the heart disease risk	95.85% (Sensitivity) 83.33% (Specificity)	500 records
A system for diagnosis of heart disease [17]	GA tuning of the FES	Clinical parameters	4 classes that define the values of the risk	74.00% (Accuracy of Test)	The dataset include 904 records
CAD Prediction [18]	Combined Methods of Data Mining	Clinical parameters	3 Classes	0.074 (MSE)	282 records with 58 attributes

Table 2. The Linguistic Terms of Input /Output Variables [19]

Linguistic Variables	**Linguistic Terms**
Blood Pressure	Low, Normal, High
Cholesterol	Normal, High, Very High
Blood Sugar	Normal, High
Heart Rate	Slow, Normal, Fast
Smoking	Non-Smoky ,Smoky
Heart Disease (Output)	Low Risk, High Risk

number of parameters to adjust is computed as follows:

$$\#\text{parameters to adjust} = (\#variables \times \#\text{linguistic terms} \times \#\text{parameters in a Gaussian MF}) \quad (2)$$

In this study, this value is computed as $3*3*2$ (for blood pressure. cholestrol and heart rate variables) $+(3*2*2)$ (for blood sugar. smoking and heart disease variables) $= 30$.

(3) **Build a chromosome:** by sorting the parameters of the MF for each variable. All of the parameters that was specified in previous step, are sorted to create an individual. Each of these parameters are considered as a gene in chromosome. Initialization of parameters are performed using the knowledge of experts. The size of population was considered to be 20.

(4) **Design an objective function**: For this, minimization of the RMSE was considered. This measure compares the response of the FES to the real value.

(5) **Select type of the GA operations and its parameters**: The selection operator choose the superior chromosomes in a population as parents. The crossover operation combines them to create a new offspring with a probability, Pcrossover. The mutation operation selects a bit in a chromosome with a mutation probability, Pm, and changes it to a new possible value. In this study, the P_m and P_c were selected heuristically according to the standard defined in [20, 21]. According to the Table 3, the best performance of the GA for (HDP) was obtained using 0.95 and 0.05 for Pcrossover and Pmutation, respectively.

(6) **Implement the necessary restrictions to build valid individuals**: For each of the genes, upper bound and lower bound are defined according to their valid range of the variables.

(7) **Execute the GA to evolve the FIS**: this process is illustrated in Figure 2. Genetic algorithm was converged after 400 generations.

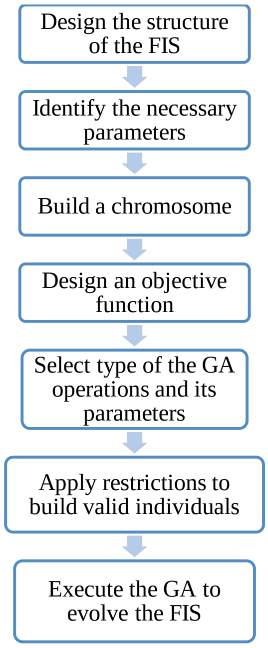

Figure 1. The Flowchart of FGA to Tune FIS.

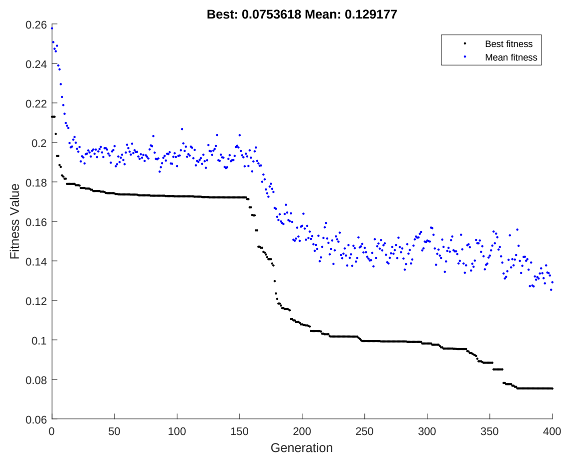

Figure 2. Execute the GA to Evolve the FIS.

The above mentioned steps for tuning of FIS using GA are illustrated in the flowchart presented in Figure 1. The fuzzy rule set after applying this Fuzzy-GA is presented in Table 4.

3.3.1 Representation of Chromosome

Representation of the structure of chromosome is an important stage in design of the Fuzzy-GA. This paper works on the optimization of the MF parameters of the FES. Then MF parameters were considered as gens in a chromosome. The Gaussian membership function includes two parameters, $\sigma.\mu$, that show standard deviation and mean, respectively. In fact, for optimization of database of a FES, the parameters

Table 3. The Different Crossover and Mutation Parameters

(P_m, P_c)	(0.0,1.0)	(0.1,0.9)	(0.2,0.8)	(0.2,0.9)	(0.01,0.99)	(0.05,0.95)	(0.15,0.85)
$P_m \leq 0.2$	0	0.1	0.2	0.2	0.01	0.05	0.15
$P_c \geq 0.8$	1	0.9	0.8	0.9	0.99	0.95	0.85
MSE after the 400 runs of the GA	82.49%	92.36%	91.84%	91.17%	82.07%	93.09%	92.84%

Table 4. The Improved Rule Set of the FES

Rule #	**Diagnostic Fuzzy Rules**
R1	If blood pressure is normal and cholesterol is normal and blood sugar is normal and heart rate is normal and smoking is non-smoker then heart disease is low risk.
R2	If blood pressure is high then heart disease is high risk.
R3	If cholesterol is high then heart disease is high risk.
R4	If blood sugar is high then heart disease is high risk.
R5	If smoking is smoker then heart disease is high risk.

of Gaussian membership functions are required to be optimized. The designed chromosome for proposed Fuzzy-GA approach is illustrated in Figure 3.

3.3.2 The GA Convergence Property

The genetic operators directly affect the convergence property of the GA. The cross-over combines the fittest selected parents. The mutation operation provides variation for the GA. High crossover and low mutation rates can improve convergence of the GA. One of the common problems associated to evolutionary algorithm such as GA is premature convergence. This problem occurs when the algorithm traps in local optima. To address this problem, the following solutions have been considered:

(1) Large population of chromosomes have been considered to ameliorate this problem but it does not guarantee all variations.
(2) The mutation is a useful technique to improve the genetic diversity and avoid trapping in the local optimum solution.

This study applies and controls these parameters to avoid early convergence of the algorithm. To select the appropriate probability of mutation and crossover, the hybrid Fuzzy-GA was executed using different rates for crossover and mutation as shown in Table 3. The best performance of GA was obtained using 0.95 and 0.05 $P_{crossover}$ and $P_{mutation}$, respectively.

4 Performance Evaluation of the Fuzzy-GA Model for HDP

The Fuzzy-GA method for HDP was evaluated and performance of the proposed method was investigated using an ROC curve analysis. Performance evaluation was experimented for the Fuzzy-GA for prognosis of the heart disease on real data of patients.

4.1 Dataset

The dataset includes 380 patients, which was collected from Parsian Hospital in Karaj, Iran. In this dataset 32.1% of samples was related to male and remaining to female. The dataset includes 55.26% of high risk patients and the rest are healthy people. The samples actual labels defined by medical experts were compared to outputs of FES and Fuzzy-GA method. The input features are Gender, Age, Blood pressure, Blood sugar, Cholesterol, Heart rate and Smoking.

4.2 Receiver Operating Characteristic Analysis

One of the common evaluation method for classification models is ROC (Receiver Operating Characteristic) curve analysis. An ROC curve demonstrates the trade-off between system benefits and its cost as the observer changes the decision threshold. The ROC curve plots True positive (TP) on the y-axis against False Positive (FP) on the x-axis. The performance of each classifier is represented by the area under the ROC curve. The area under the ROC curve is a measure of the performance of the classifier [22, 23].

True positive (TP): If the object is positive (abnormal) and is classified as positive.

False negative (FN): If the object is positive and it is classified as negative (normal).

True negative (TN): If an object is negative and is classified as negative.

False positive (FP): If an object is negative and is classified as positive.

According to [22, 23], the ROC curve depicts the trade-off between the classifier's benefits (tp rates) and

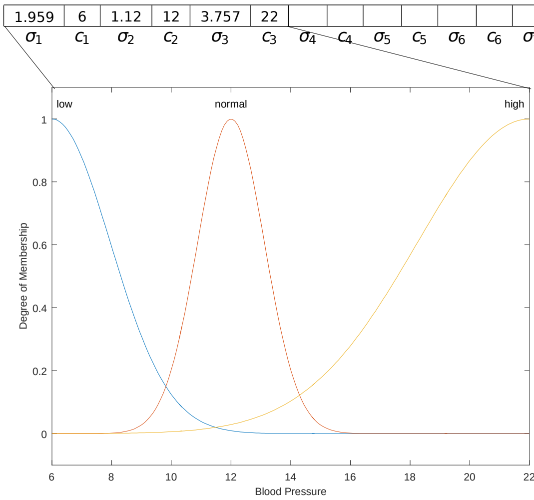

Figure 3. The Chromosome Representation for the Fuzzy-GA Method.

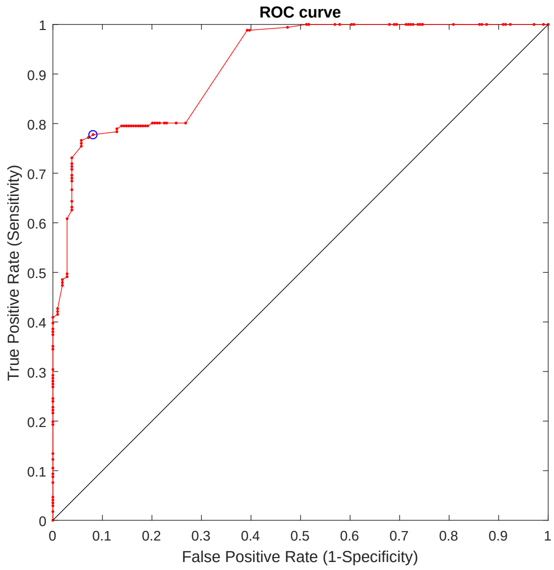

Figure 4. The ROC Curve of FES for HDP.

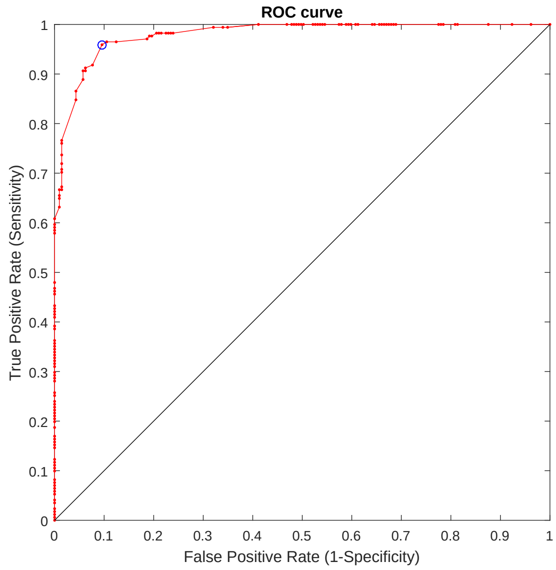

Figure 5. The ROC Curve of Fuzzy-GA system for HDP.

its cost (fp rates). Each point of this curve shows (fp rates, tp rates) a decision based on selected threshold. The area under ROC curve (AUC) has been frequently employed as a performance measure. The performance of fuzzy system and Fuzzy-GA system for HDP were investigated using an ROC curve analysis and related metrics as follows and shown in Figures 4 and 5.

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \tag{3}$$

$$Specificity = \frac{TN}{TN + FN} \tag{4}$$

$$Sensitivity = \frac{TP}{FP + TP} \tag{5}$$

4.3 Performance Evaluation of Fuzzy-GA Method Through ROC Curve Analysis

The performance of 2-class classifier for HDP was demonstrated using the confusion matrix. Confusion matrix shows all the predictions of 2-class classifier for data. A classification system predicts the class label of an object [24, 25]. The confusion matrix of fuzzy system and the Fuzzy-GA system shown in Tables 5 and 6, respectively.

According to Table 6, after applying the Fuzzy-GA approach, the number of the TP and TN circumstances have increased (26 more samples are correctly classified) while number of misclassified samples has decreased from 55 samples in FES to 29 samples.

Table 5. Confusion Matrix of Fuzzy System

		Prediction Class	
Actual Class		**Yes**	**No**
	Yes	133 (TP)	17 (FN)
	No	38 (FP)	192 (TN)

Table 6. Confusion Matrix of Fuzzy-GA System

		Prediction Class	
Actual Class		**Yes**	**No**
	Yes	156 (TP)	14 (FN)
	No	15 (FP)	195 (TN)

Table 7. Comparison of Performance of Fuzzy System and Fuzzy-GA System Using the ROC Curve Criteria.

	Fuzzy System %	**Fuzzy-GA System%**
Accuracy	85.52	92.37
Specificity	91.86	93.3
Sensitivity	77.77	91.23
AUC	91.77	96.23
MSE	82.96	93.09

The ROC analysis results before and after optimization have been compared in Table 7. According to the obtained results in this Table, the Fuzzy-GA system for HDP outperforms the FES. After optimization and tuning of parameters of fuzzy system using Fuzzy-GA algorithm, its accuracy and sensitivity (tp rate) have increased. The AUC (area under the ROC curve) of Fuzzy-GA with 96.23% is greater than the fuzzy system with an accuracy of 91.77%.

4.4 Comparison of Fuzzy-GA Method With Related Works for HDP

Comparison of the proposed methods in this study with the other related works for prediction of heart disease has been presented in Table 8. A Mamdani fuzzy inference model was used which has a high interpretability to interact with the medical experts. The performance of the proposed model is competitive with its counterpart methods for its high accuracy and the interpretability, *i.e.*; lower number of input variables and less fuzzy rules. Also the proposed FES works with input features that are not expensive to achieve for all the people. The Fuzzy-GA method for tuning the parameters of membership functions in FES was initialized using the knowledge of experts. The extensive usage and simple implementation and high interpretability of generated solutions are the reasons for selection of Genetic Algorithm for optimization of FES. The advantages of proposed Fuzzy-GA method for HDP have been summarized as follows:

(1) High interpretability to interact with the medical experts by using fuzzy inference model.
(2) The design of a FES is simple and cheap.
(3) The rules of a FES emulate the experts' knowledge representation model. Therefore, they are very interactive to the clinicians.
(4) Managing uncertainty in symptoms used for HDP using the fuzzy sets.
(5) Dynamic behavior of the GA for optimization of the FL through the tuning process of the FL, using the genetic operators (mutation and crossover).
(6) Evolution of the FES through Fuzzy-GA model and its adaption to the real patients feature space
(7) Higher accuracy (AUC) in comparison with the counterpart methods.

According to the Table 8, the classification system in [11] which uses heart rate (ECG) as input, has higher accuracy than the systems that use clinical parameters as input data. The Fuzzy-GA model proposed in this paper compared to the work in [11], utilizes simple features with lower complexity. Furthermore, the aim of the proposed Fuzzy-GA method was to maintain the trade-off between the FES accuracy and its interpretability. The accuracy of proposed Fuzzy-GA model is competitive while providing higher interpretability in the input feature sets and the process of the design of the FESs.

5 Discussions and Conclusion

In this paper, a Fuzzy-GA system for HDP was proposed. The Fuzzy-GA system has improved accuracy of the FES. The designed system predicts the risk of heart disease development. Performance of the proposed model was investigated applying a dataset including 380 patients collected from Parsian Hospital in Karaj, Iran. The dataset includes actual diagnosis of medical experts. Accuracy of the FES was 85.52% and after tuning with Fuzzy-GA method was reached to 92.37%. The Fuzzy-GA method in comparison with other related works has higher interpretability to interact with the medical experts by using fuzzy inference model with less input parameters. Our future work is to optimize the (HDP) system using other evolutionary algorithms.

Table 8. Comparison of the Accuracy of the Related Works With Proposed Works in This Study.

Name	Method	Accuracy (Measure)
A prediction system for heart disease [6]	hybrid ANN-GWO algorithm	0.0036 (RMSE)
A mechanism for prediction of heart disease [7]	Combination of rough set with PSO and create the IHQRPSO algorithm	95.25% (accuracy)
A prediction system for heart disease [8]	Combination of the fuzzy inference system and GA	86% (Accuracy) 90% (Specificity) 80% (Sensitivity)
A classification system for data [9]	Combined Fuzzy SAM and GA	95.59% (Accuracy) 96% (F- measure)
A prediction system for heart disease [10]	RS based attribute reduction and IT2FLS	86% (Accuracy) 78.1% (Sensitivity) 90% (Specificity)
A classification system for heart rate [11]	Modified artificial bee colony algorithm	99.46% (Sensitivity) 99.37% (Specificity) 99.94% (Positive predictivity)
A optimization system for diagnosis of heart disease [12]	Adaptive network-based fuzzy inference system and linear discriminant analysis	80.2% (Prediction rate)
A prediction system for heart disease [13]	A back propagation neural network	92% (Accuracy)
A classification system for heart disease [14]	Combination of genetic algorithm and KNN	95.73% (Accuracy)
A system for diagnosis of heart disease [15]	Combination of the fuzzy inference system and PSO algorithm	93.27% (Accuracy)
A FES [16]	Fuzzy system with clinical parameters in inputs	95.85% (Sensitivity) 83.33% (Specificity)
A heart disease diagnosis model [17]	GA tuning of the FES	74.00% (Accuracy of Test)
A FES for assessment of the heart disease risk [19]	A FES based on Mamdani model	85.52% (Accuracy) 91.86% (Specificity) 77.77% (Sensitivity) 91.77% (AUC) 82.96% (MSE)
A Fuzzy-GA for HDP (this study)	Combination of the GA and FES to improve accuracy	92.37% (Accuracy) 93.30% (Specificity) 91.23% (Sensitivity) 96.23% (AUC) 93.09% (MSE)

References

[1] R Khezri, R Hosseini, and M Mazinani. A Fuzzy Rule- based Expert System for the Prognosis of the Risk of Development of the Breast Cancer. *International Journal Of Engineering*, (10):1557–1564, 2014.

[2] Razieh Khezri, Rahil Hosseini, and Mahdi Mazinani. A Sugeno Fuzzy Inference System for Breast Cancer Prognosis. In *Proc. of the 14th Fuzzy SystemsConference, Tabriz, Iran*, 2014.

[3] B. M. Gayathri and C. P. Sumathi. Mamdani fuzzy inference system for breast cancer risk detection. In *2015 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, pages 1–6, 2015. doi:10.1109/ICCIC.2015.7435670.

[4] R. Hosseini, S. D. Qanadli, S. Barman, M. Mazinani, T. Ellis, and J. Dehmeshki. An Automatic

Approach for Learning and Tuning Gaussian Interval Type-2 Fuzzy Membership Functions Applied to Lung CAD Classification System. *IEEE Transactions on Fuzzy Systems*, 20(2):224–234, 2012. doi:10.1109/TFUZZ.2011.2172616.

[5] J. Casillas, O. Cordon, M. J. del Jesus, and F. Herrera. Genetic tuning of fuzzy rule deep structures preserving interpretability and its interaction with fuzzy rule set reduction. *IEEE Transactions on Fuzzy Systems*, 13(1):13–29, 2005. ISSN 1063-6706. doi:10.1109/TFUZZ.2004.839670.

[6] Hamza Turabieh. A Hybrid ANN-GWO Algorithm for Prediction of Heart Disease. *American Journal of Operations Research*, 6(2):136–146, 2016.

[7] Sivagowry Shathesh and M Durairaj. An intelligent hybrid mechanism to predict the risk of cardio vascular disease. *Indian Journal of Science and Technology*, 9(4):137–146, 2016.

[8] T Santhanam and EP Ephzibah. Heart disease prediction using hybrid genetic fuzzy model. *Indian Journal of Science and Technology*, 8(9):797–803, 2015.

[9] Thanh Nguyen, Abbas Khosravi, Douglas Creighton, and Saeid Nahavandi. Classification of healthcare data using genetic fuzzy logic system and wavelets. *Expert Systems with Applications*, 42(4):2184 – 2197, 2015. doi:https://doi.org/10.1016/j.eswa.2014.10.027.

[10] Nguyen Cong Long, Phayung Meesad, and Herwig Unger. A highly accurate firefly based algorithm for heart disease prediction. *Expert Systems with Applications*, 42(21):8221 – 8231, 2015. doi:https://doi.org/10.1016/j.eswa.2015.06.024.

[11] Selim Dilmac and Mehmet Korurek. ECG heart beat classification method based on modified ABC algorithm. *Applied Soft Computing*, 36:641 – 655, 2015. doi:https://doi.org/10.1016/j.asoc.2015.07.010.

[12] Jung-Gi Yang, Jae-Kwon Kim, Un-Gu Kang, and Young-Ho Lee. Coronary heart disease optimization system on adaptive-network-based fuzzy inference system and linear discriminant analysis (ANFIS–LDA). *Personal and Ubiquitous Computing*, 18(6):1351–1362, Aug 2014. doi:10.1007/s00779-013-0737-0.

[13] Nabeel Al-Milli. Backpropagation neural network for prediction of heart disease. *Journal of theoretical and applied information Technology*, 56(1):131–135, 2013.

[14] M. Akhil jabbar, B.L. Deekshatulu, and Priti Chandra. Classification of Heart Disease Using K- Nearest Neighbor and Genetic Algorithm. *Procedia Technology*, 10:85 – 94, 2013. doi:https://doi.org/10.1016/j.protcy.2013.12.340.

[15] S. Muthukaruppan and M.J. Er. A hybrid particle swarm optimization based fuzzy expert system for the diagnosis of coronary artery disease. *Expert Systems with Applications*, 39(14):11657 – 11665, 2012. doi:https://doi.org/10.1016/j.eswa.2012.04.036.

[16] Debabrata Pal, K.M. Mandana, Sarbajit Pal, Debranjan Sarkar, and Chandan Chakraborty. Fuzzy expert system approach for coronary artery disease screening using clinical parameters. *Knowledge-Based Systems*, 36:162 – 174, 2012. doi:https://doi.org/10.1016/j.knosys.2012.06.013.

[17] J. Sanz, M. Pagola, H. Bustince, A. Brugos, A. Fernández, and F. Herrera. A case study on medical diagnosis of cardiovascular diseases using a Genetic Algorithm for Tuning Fuzzy Rule-Based Classification Systems with Interval-Valued Fuzzy Sets. In *2011 IEEE Symposium on Advances in Type-2 Fuzzy Logic Systems (T2FUZZ)*, pages 9–15, 2011. doi:10.1109/T2FUZZ.2011.5949553.

[18] M. Shamsollahi, A. badiee, and M. Ghazanfari. Using Combined Descriptive and Predictive Methods of Data Mining for Coronary Artery Disease Prediction: a Case Study Approach. *Journal of AI and Data Mining*, 2018. ISSN 2322-5211. doi:10.22044/jadm.2017.4992.1599.

[19] Rana Akhoondi and Rahil Hosseini. A Fuzzy Expert System for Prognosis of the Risk of Development of Heart Disease. *Journal of Advances in Computer Research*, 7(2):101–114, 2016. ISSN 2345-606X. URL `http://jacr.iausari.ac.ir/article_19844.html`.

[20] A.E. Eiben and J.E. Smith. *Introduction to Evolutionary Computing*. Springer Berlin Heidelberg, 2015. doi:10.1007/978-3-662-44874-8.

[21] T. Back. Selective pressure in evolutionary algorithms: a characterization of selection mechanisms. In *Proceedings of the First IEEE Conference on Evolutionary Computation. IEEE World Congress on Computational Intelligence*, pages 57–62 vol.1, 1994. doi:10.1109/ICEC.1994.350042.

[22] Andrew P. Bradley. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7):1145 – 1159, 1997. doi:https://doi.org/10.1016/S0031-3203(96)00142-2.

[23] Tom Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27 (8):861 – 874, 2006. ISSN 0167-8655. doi:https://doi.org/10.1016/j.patrec.2005.10.010. ROC Analysis in Pattern Recognition.

[24] Jiawei Han, Micheline Kamber, and Jian Pei. 6 - mining frequent patterns, associations, and correlations: Basic concepts and methods. In *Data Mining (Third Edition)*, The Morgan Kaufmann Series in Data Management Systems,

pages 243 – 278. Morgan Kaufmann, Boston, third edition edition, 2012. ISBN 978-0-12-381479-1. doi:https://doi.org/10.1016/B978-0-12-381479-1.00006-X.

[25] Pang-Ning Tan, Michael Steinbach, and Vipin Kumar. *Introduction to Data Mining, (First Edition)*. Addison-Wesley Longman Publishing Co., Inc., 2005. ISBN 0321321367.

Rana Akhoondi received her MSc degree in Artificial Intelligence, from department of Computer Engineering, Shahr-e-Qods Branch, Islamic Azad University, Tehran, Iran. She holds first rank during her BSc. and MSc. studies. Her postgraduate thesis is prognosis of the risk of the heart disease development using hybrid fuzzy models and evolutionary algorithms. Her research of interest is in soft computing, fuzzy modeling and evolutionary optimization. She has some publications in the field of computational intelligence.

Rahil Hosseini is a member of faculty at department of computer engineering at Islamic Azad University in Tehran. In 2001, she received her BSc at Kharazmi University of Tehran followed by her MSc in 2004 in software engineering in Tehran. In 2008, she started working toward the Ph.D, degree with Kingston University London, U.K. From 2008 until 2010, she cooperated in a project involving image analysis techniques to detect and quantify retinal vessel tortuosity in relation to cardiovascular risk factors in collaboration with St. George's University of London. She is a member of cancer study group at Kingston University.

Mahdi Mazinani received the B.Sc. degree (securing the first rank) in electronic engineering and the M.Sc. degree (first rank), both from the Semnan University, Semnan, Iran, in 1992 and 2002, respectively. In 2008, he started toward PhD. degree, which is titled "Computer Aided Detection and quantification of degree of the stenosis of the Coronary Artery Disease using Computed Tomography Angiography".