International Journal of Engineering & Technology, 8 (1.7) (2019) 119-125

SPC

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Fuzzy Logic System for Diagnosing Coronary Heart Disease

Mustafa Salah Mahdi[1], Mohd Faisal Ibrahim[2]*, Shakir Mahmood Mahdi[3], Praveen Singam[4], Aqilah Baseri Huddin[2]

[1]Electrical & Electronic Programme,
[2]Centre for Integrated Systems Engineering and Advanced Technologies (INTEGRA),
Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, MALAYSIA.
[3]Information Technology Center (ITC), University of Technology Baghdad, IRAQ.
[4]KPJ Kajang Specialist Hospital, Jalan Cheras,43000 Kajang, Selangor, MALAYSIA.
**Corresponding author E mail: faisal.ibrahim@ukm.edu.my*

Abstract

The diagnosis of coronary heart disease is a non-trivial task that requires a careful and time-consuming examination. Hence, the application of a computer-aided diagnostic (CAD) system to assess the condition of a person of having coronary heart disease is greatly beneficial. Although the usage of CADs related to coronary heart disease diagnosis is widely implemented, the inputs medical data required is still a significant challenge to encourage rapid diagnosis among public. This paper presents a fuzzy logic system for diagnosing coronary heart disease. In this study, the fuzzy logic based CAD system has been developed with five input variables that requires minimal medical procedure to obtain to determine the presence of coronary heart disease. Experiments to assess the accuracy of the system and comparison with a previous work of rough-fuzzy classifier were carried out using an open source coronary heart disease dataset. The results show that the proposed work is able to achieve significant accuracy of 72.6%, an improvement of 30% from the previous work.

Keywords: *fuzzy logic; expert system; coronary heart disease (CHD); computer-aided diagnostic system (CAD); health care;*

1. Introduction

Recent health information by World Health Organisation (WHO) indicates that cardiovascular disease including coronary heart disease is the number one cause of death globally. Among the country with high death toll due to coronary heart disease is the United State of America with 43.8% of deaths recorded in 2015 [1]. The statistics also show that some developing countries such as Malaysia has higher mortality rate with 35% deaths due to the cardiovascular disease [2]. It is also reported that cardiovascular disease represents one of the main causes of adult deaths in the other developed countries [3].

The alarming statistics require immediate actions in reducing the mortality rate due to the complication of such disease. One way to boost the effort is to use computer technology to diagnose the disease. Medical diagnosis can be referred as the procedure of defining which disease or condition explains a patient's symptoms and signs [4]. Such technological diagnosis approach is known as computer-aided diagnostic (CAD) systems. CAD systems can help medical practitioners and patients in many ways. The systems can helps in identifying complex relationships between patient's symptoms and attributes to measure the level of severity of a disease. CAD systems can be used to deal with ambiguous information to classify accurately the health condition of a patient [5].

Many studies from the past have shown the effectiveness of the CAD systems in diagnosing disease despite of continuing technological improvement. The adoption of artificial intelligence technique in CAD systems has been recognised as an important milestone in increasing the accuracy of diagnostic results. One class of such intelligent CAD systems can be referred as Fuzzy logic-based diagnostic system. Fuzzy logic is an artificial intelligence technique pioneered by Professor Lotfi A. Zadeh in the mid-1960 to imitate the capability of human thinking in making decision [6]. The technique is able to deal with problems that involve uncertain, imprecise and ambiguous data. Among previous works on fuzzy logic based CAD system include [7] for Thalassemia disease, [8] (2017) for liver disease and [9] for thypoid fever.

For heart disease diagnosis, there are various fuzzy logic based CAD systems that have been developed [3, 10-14]. One interesting feature that worth to be highlighted when designing the system is the selection and the number of factors or input variables related to heart disease. The database of [15] shows an instance of more than 50 heart disease factors where the complexity and relationship between factors are indirect. From the past works, it is stated that at least four (4) input variables to eleven (11) input variables are required and some of them require particular medical procedure to obtain the input data.

This paper presents a fuzzy logic system for diagnosing coronary heart disease with the aim to encourage rapid diagnosis among public as well as learning tool for junior medical practitioners. In this study, we develop a fuzzy logic system with five (5) input variables represent factors to determine the presence of coronary heart disease. The selection of the 5 variables are carefully taken by getting the expert view and recommendation. The criteria of the selection is to consider the practicality of a person to take the reading of the input variables regularly. Our approach differs from other works in that the required input variables can be taken with little medical procedure but preserves significant diagnostic results. The rules of fuzzy logic inference engine are carefully designed by a medical practitioner to demonstrate the capability of the fuzzy logic system to gather and transfer expert knowledge into useful machine information. We have ana-

lysed the accuracy of our work by using an open dataset of real coronary heart disease patients obtained from the University of California, Irvine (UCI) Machine Learning Repository [15]. The results show that the proposed work is able to achieve the accuracy and specificity of 72.6% and 77.6%, respectively. Finally, our fuzzy logic system for diagnosing coronary heart disease has been developed in the form of Java software.
This paper is organised as follows. The next section describes the research method involving the proposed fuzzy logic design and implementation for coronary heart disease diagnosis. Section 3 provides the experimental results and discussion. Finally, the conclusion of this work is presented in section 4.

2. Research method

This section presents all necessary information regarding research process, dataset and the fuzzy logic design for coronary heart disease diagnostic system.

2.1. Research process

Fig. 1 presents the research process taken to design the fuzzy logic system for diagnosing coronary heart disease. The process is divided into four phases to achieve the objectives of the research. The phases are 1) Data collection phase; 2) Fuzzy logic design phase; 3) System implementation phase; and 4) System evaluation phase.

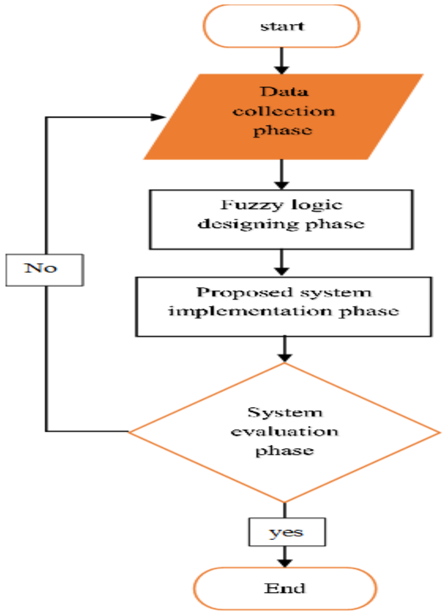

Fig. 1: Research Process

2.2. Data collection and analysis

An open source medical data from the University of California, Irvine (UCI) Machine Learning Repository, Heart Disease Dataset [4] are taken as the main source for experimental procedures and performance benchmarking of the proposed work. The data has three sets namely, Cleveland, Hungarian, and Switzerland sets. We select the Hungarian data set for the proposed system which contains 294 cases of coronary heart disease (CHD). In this study, we use only five (5) input variables from the dataset after thorough screening with medical experts. The input variables are age, blood pressure, cholesterol level, heart rate and gender. Meanwhile, the output from the diagnostic refers to the presence of CHD ranging from 0 (no presence) to 4 (the highest risk). The increasing value shows the increment of the coronary heart disease risk.

2.3. Fuzzy logic design and implementation

Fuzzy logic is adopted in this study to predict the health condition of CHD patients based on the selected symptoms. In order to develop the fuzzy logic system, four stages with seven steps are taken as summarised in Table 1.

Table 1: Stages and steps to develop a fuzzy logic system

Stages	Steps
1. Fuzzification	i. Define linguistic variables and range. ii. Construct the membership functions for the input and output variables. iii. Convert crisp data into fuzzy data sets using membership functions.
2. Fuzzy rule base	iv. Construct knowledge base of rules from expert input.
3. Fuzzy inference engine	v. Evaluate rules in the rule base. vi. Combine results from each rule.
4. Defuzzification	vii. Convert output data into non-fuzzy values.

2.3.1. Fuzzification

In the fuzzification stage, linguistic variables, linguistic values, and their range are defined. The linguistic variables consist of two parameters which are input and output. The input variables are limited to five (5) variables that are the age of patients, blood pressure, cholesterol, gender, and heart rate while the output is the level of coronary heart disease (CHD). For each of the linguistic variables, the linguistic values and range are shown in Table 2. The range of linguistic variables are classified based on the data analysed from the reference dataset as explained in section 2.2 and advice from the medical experts.

Table 2: Linguistic variables of the proposed CHD fuzzy logic diagnostic system.

Variable	Linguistic Variable	Linguistic Value	Range
Input	Age	Young	<38
		Middle	35-50
		Old	>45
	Blood Pressure	Low	<125
		Medium	120-145
		High	140-165
		Very High	>160
	Cholesterol	Low	<197
		Medium	188-250
		High	217-307
		Very High	>281
	Gender	Male	1
		Female	0
	Heart Rate	Low	<140
		Medium	110-195
		High	>150
Output	Result	Healthy	0-1.6
		Middle	1.5-2.6
		Sick	2.5-4

When all information of the linguistic variables has been gathered, the numeric crisp variables are converted into their equivalent fuzzy sets through membership functions. The membership functions can be designed in several types such as triangular, trapezoidal, bell shaped or Gaussian. Nevertheless, only one form of membership functions is selected this work i.e. the trapezoidal membership function. The trapezoidal membership function can be constructed by referring to equation (1):

$$\mu_n(x) = \begin{cases} 0 \quad for \quad x < n_1, \\ \dfrac{x - n_1}{n_2 - n_1} \quad for \quad n_1 \leq x \leq n_2, \\ 1 \quad for \quad n_2 \leq x \leq n_3, \\ \dfrac{n_4 - x}{n_4 - n_3} \quad for \quad n_3 \leq x \leq n_4, \\ 0 \quad for \quad x > n_4. \end{cases} \tag{1}$$

where $n_1 - n_4$ refer to the points in the universal of discourse of a variable. Based on the advice of medical experts, the fuzzy sets for the input and output linguistic variables have been constructed by using equation (1) as in Fig. 2 to Fig. 6.

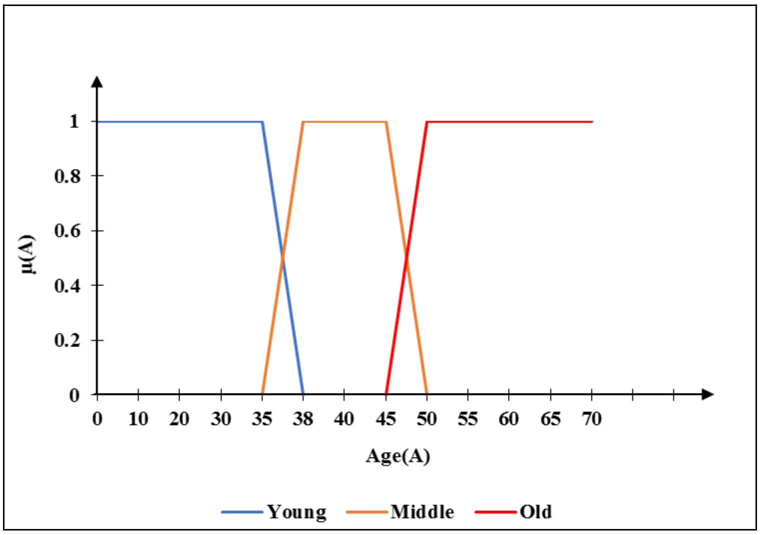

Fig. 2: Membership function of age

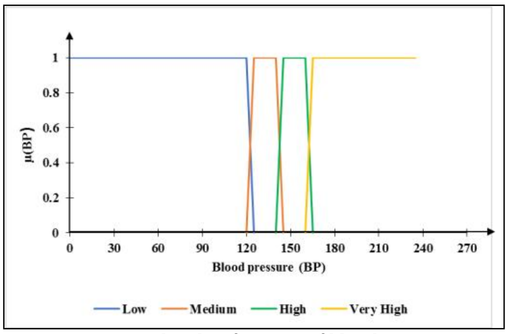

Fig. 3: Membership function of blood pressure

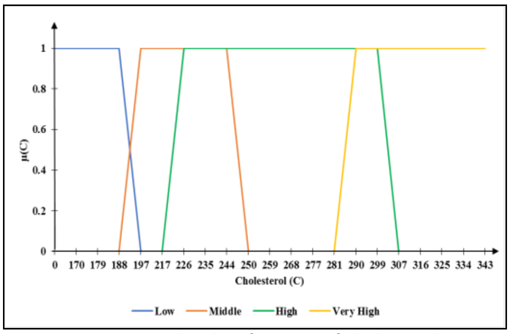

Fig. 4: Membership function of cholesterol

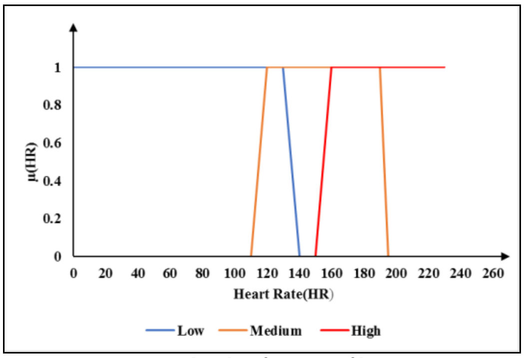

Fig. 5: Membership function of heart rate

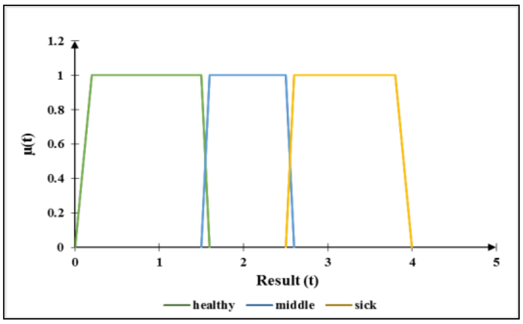

Fig. 6: Membership function of level of CHD (output)

2.3.2. Fuzzy rule base

During this stage, knowledge base of rules is constructed. In a fuzzy logic design, a rule base is created to infer the decision making and logic process of human being. A fuzzy rule is a simple IF-THEN rule with a condition and a consequence. The fuzzy rules will affect the quality of the result in a fuzzy system. In this study, the total of 288 feasible rules are obtained and verified by medical experts to cater all possible cases depicted by the developed input-output membership functions. Table 3 represents the sample of the rules.

Table 3: Sample of fuzzy rules for CHD diagnostic system

	Fuzzy IF-Then Rules
1.	IF A IS young AND BP IS low AND C IS low AND HR IS low AND G IS male THEN t IS healthy
2.	IF A IS young AND BP IS low AND C IS low AND HR IS low AND G IS female THEN t IS healthy
3.	IF A IS young AND BP IS low AND C IS low AND HR IS medium AND G IS male THEN t IS healthy
4.	IF A IS young AND BP IS low AND C IS low AND HR IS medium AND G IS female THEN t IS healthy
5.	IF A IS young AND BP IS low AND C IS low AND HR IS high AND G IS male THEN t IS middle
6.	IF A IS young AND BP IS low AND C IS low AND HR IS high AND G IS female THEN t IS middle
7.	IF A IS young AND BP IS low AND C IS medium AND HR IS low AND G IS male THEN t IS healthy

where (A) represents the age, (BP) represents the blood pressure, (C) represents the cholesterol, (HR) represents the heart rate, (G) represents the gender and (t) represents the level of CHD.

2.3.3. Fuzzy inference engine

During this stage, the step of evaluation of rules is performed by fuzzy set operation. Then, in order to make a decision in a particular approach, the rules need to be tied together using aggregation method. Aggregation method is a process of combining all the individual fuzzy input and infer it into single fuzzy output. In this study, the standard Mamdani approach is selected.

2.3.4. Defuzzification

Defuzzification is the process of generating and producing a final output in the form of crisp value from the inferred single output fuzzy set. Centroid method is used as the deffuzification method in this study. It is used to regulate a point that indicates the center of gravity (COG) of the fuzzy set. The formula that has been used to calculate the COG is defined in equation (2).

$$t^* = \frac{\int \mu_A(t) * t \, dt}{\int \mu_A(t) \, dt} \tag{2}$$

where t^* is the COG point represents the output of the system in the form of a crisp numerical value, $\mu A(t)$ is the membership value of point t in the output fuzzy set.

3. Results and discussion

Fig. 7 shows the diagram of the proposed fuzzy logic system design.

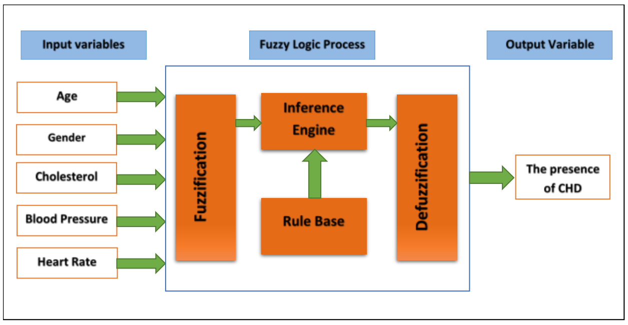

Fig. 7: Shows the diagram of the proposed fuzzy logic system design.

The design of the proposed system is deployed by using NetBeans IDE 8.2 in Java Programming language. Fig. 8 shows the interface for the proposed CHD Diagnostic System. Firstly, once the user runs the system, a window will be pop up. This window displays five empty text boxes for each of the input variables where the user needs to fill in the details accordingly. Then, there are a green and a blue buttons. The green button is used to perform diagnosis using the proposed fuzzy logic CHD diagnostic system, while the blue button is to reset the system. By clicking on the diagnose button (green), the result of CHD level will be displayed. Then, the user needs to refer to the table provided in order to know status of the result. By clicking on the reset button (blue), the five text boxes will reset then the user can submit another case. The significance of the CHD diagnostic system designed in this study is its ability to produce rapid result and can be used by the public and junior medical practitioners.

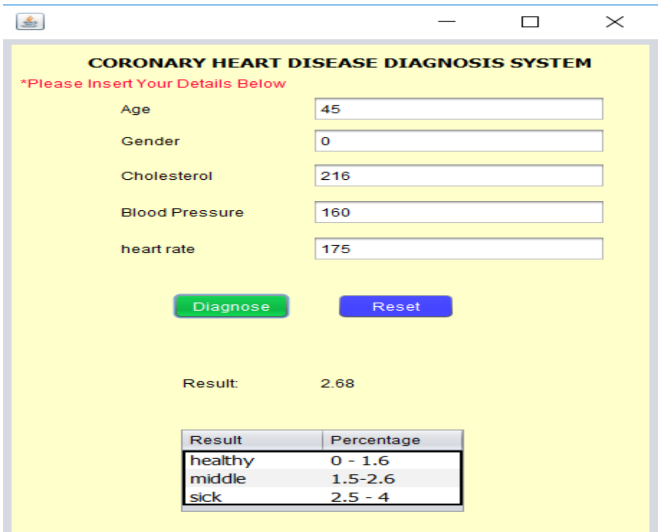

Fig. 8: Interface of CHD Diagnosis System

The evaluation of the system is conducted by measuring the performance of the proposed work and comparing it with the performance of an existing work of rough-fuzzy classifier defined by [14]. The measurement of the performance is conducted by using the confusion matrix. The confusion matrix is a table that applied to define the performance of a classification class or a classifier on a set of trial data for which the true values are identified.

The performance measurement applied to a medical diagnostic system are accuracy, sensitivity, and specificity. The occurrence of a condition is stated as a positive situation while the non-occurrence is stated as a negative situation. Accuracy is defined as the capability of the system to give an accurate diagnosis. Sensitivity is the capability of the system to classify the existence of a target class accurately. Meanwhile, specificity is the capability of the system to separate the target class. Hence, accuracy, sensitivity, and specificity are determined by equations (3) to (11):

$$Accuracy = \frac{TP_A + TP_B + TP_C}{N} \tag{3}$$

$$Sensitivity(A) = \frac{TP_A}{TP_A + FN_A} \tag{4}$$

$$Sensitivity(B) = \frac{TP_B}{TP_B + FN_B} \tag{5}$$

$$Sensitivity(C) = \frac{TP_C}{TP_C + FN_C} \tag{6}$$

$$Overall_{Sensitivity} = \frac{Sensitivity(A) + Sensitivity(B) + Sensitivity(C)}{3} \tag{7}$$

$$Specificity(A) = \frac{TN_A}{TN_A + FP_A} \tag{8}$$

$$Specificity(B) = \frac{TN_B}{TN_B + FP_B} \tag{9}$$

$$Specificity(C) = \frac{TN_C}{TN_C + FP_C} \tag{10}$$

$$Ocerall_{Specificity} = \frac{Specificity(A) + Specificity(B) + Specificity(C)}{3} \tag{11}$$

where: N is the total number of instances in a dataset, TP is True Positive, TN is True Negative, FP is False Positive and FN is False Negative conditions.

The TP or true positive represents the positive cases that are truely classified by the system. The TN or true negative is the negative cases that precisely classified by the system. The FP or false positive is the negative cases that wrongly classified by the system, while the FN or false negative is the positive case that incorrectly classified by the system.

The result of the proposed system by using dataset as explained in section 2.2 is tabulated in Table 4. The result shown is based on the Hungarian dataset where 270 feasible instances are presented. The distribution of the dataset ground-truth is that 206 instances are presented as healthy cases, 23 instances as middle level cases and 41 instances categorised as sick level cases.

Table 4: The confusion matrix represents the classification result of the proposed system

		Predicted class		
		Healthy	Middle	Sick
Actual class	Healthy	174	27	5
	Middle	12	11	0
	Sick	20	10	11

Table 5 shows the result of comparative analysis with the existing classifier by [14]. From the table, we can identify that the proposed work has obtained accuracy of 72.6% compared with the existing classifier with 42.4%. The proposed work achieved 77.6% specificity value while the existing work has only 29.4%. In both cases, our proposed system outperforms the existing work The improvement made on the accuracy and specificity is 30% and 48%, respectively. However, the sensitivity of the proposed work at 53.1% is slightly lower than the existing work with 63.8%. In overall, the result of the proposed system is satisfactory even though less input variables are used in the system.

Table 5: Comparative analysis between the proposed system and the existing classifier

	Proposed work	Srinivas et al. 2014
Overall Accuracy	72.6%	42.4%
Sensitivity (healthy)	84.5%	
Sensitivity (middle)	47.8%	
Sensitivity (sick)	26.8%	
Overall Sensitivity	53.1%	63.766%

Specificity (healthy)	50%	
Specificity (middle)	85%	
Specificity (sick)	97.82%	
Overall Specificity	77.6%	29.4%

4. Conclusion

This paper has presented the fuzzy logic system for diagnosing coronary heart disease. The system has been developed to use only five input variables to diagnose CHD of an individual. Meanwhile, the CHD level output is categorised into three level. Java programming language was involved in building, designing and also testing the result of CHD presence.

The performance comparison between the proposed method and an existing work has been completed. Based on the result, the proposed system has the improvement of accuracy and specificity of 30% and 48%, respectively.

From this research, we can conclude that the fuzzy logic system for diagnosing coronary heart disease is able to classify the CHD presence at satisfactory level. The system can be used by a person rapidly that help him to take precautionary actions to stretch their lifespan. Furthermore, the system can assist a medical doctor as the second opinion tool.

For the future work, in order to increase the diagnostic accuracy, further actions can be done. The optimisation of fuzzy logic parameters can be done by integrating optimisation tools such as genetic algorithm in order to learn the pattern of a given dataset. Analysis to the suitable number of input variables to the system can also be done considering the practicality of obtaining the input data from real world practice.

Acknowledgement

This research was supported by the Ministry of Education Malaysia and Universiti Kebangsaan Malaysia under grant number FRGS/1/2017/TK04/UKM/02/4.

References

[1] Benjamin EJ et al., “Heart Disease and Stroke Statistics – 2018 Update A Report From the American Heart Association”, *Circulation* Vol. 137 No. 12, pp. e67-e492, 2018.

[2] World Health Organization, “Noncommunicable diseases country profiles 2018”, Geneva, 2018.

[3] Parvin R, Abhari A, “Fuzzy database for heart disease diagnosis”, *Proceedings of Medical Processes Modeling and Simulation* (MPMS), 2012.

[4] Farlex. (Ed.), Medical Dictionary, 2009.

[5] Innocent PR., John, RI, Garibaldi, JM, “Fuzzy methods for medical diagnosis”, *Applied Artificial Intelligence*, Vol. 19 No. 12, pp. 69-98, 2004.

[6] Hemba S, Islam, N, “Fuzzy Logic: A Review”, *International Journal of Computer Sciences and Engineering*, Vol. 5 No. 2, 61-63, 2017.

[7] Thakur S, Raw SN, Sharma R, “Design of a Fuzzy Model for Thalassemia Disease Diagnosis: Using Mamdani Type Fuzzy Inference System (FIS)”, *International Journal of Pharmacy and Pharmaceutical Sciences*, Vol. 8 No. 4, 2016.

[8] Zhenning Y, Vijayashree J, Jayashree J, “Fuzzy logic based diagnosis for liver disease using CBC (Complete Blood Counts)”, *J. Comput Math Sci*, Vol. 8, pp. 202-209, 2017.

[9] Asogbon MG, Samuel OW, Omisore MO, Awonusi O, “Enhanced neuro-fuzzy system based on genetic algorithm for medical diagnosis”, *J Med Diagn Meth*, Vol. 5 No. 205, 2016.

[10] Abidin B, Dom RM, Rahman ARA, Bakar RA, Demiralp M, Baykara N, Mastorakis N, “Use of fuzzy neural network to predict coronary heart disease in a Malaysian sample”, *Proceedings 8th WSEAS International Conference on Telecommunications and Informatics*, 2009, pp. 76-80.

[11] Adeli A, Neshat M. “A fuzzy expert system for heart disease diagnosis”, *Proceedings International Multi Conference of Engineers and Computer Scientists*, 2010.

[12] Cinetha K, Maheswari PU, “Decision Support System for Precluding Coronary Heart Disease (CHD) Using Fuzzy Logic”, *International Journal of Computer Science Trends and Technology*, Vol. 2 No. 2, pp. 2347-857, 2014.

[13] Hassan N, Sayed OR, Khalil AM, Ghany MA, “Fuzzy soft expert system in prediction of coronary artery disease”, *International Journal of Fuzzy Systems*, Vol. 19 No. 5, pp. 1546-1559, 2017.

[14] Srinivas K, Rao GR, Govardhan A, “Rough-Fuzzy classifier: A system to predict the heart disease by blending two different set theories”, *Arabian Journal for Science and Engineering*, Vol. 39 No. 4, pp. 2857-2868, 2014.

[15] Bache K, Lichman M, UCI Machine Learning Repository [http://archive. ics. uci.edu/ml]. University of California, School of Information and Computer Science. Irvine, CA, 2013.