

# Fuzzy Rule Generation for Diagnosis of Coronary Heart Disease Risk Using Subtractive Clustering Method

Lailil Muflikhah<sup>1</sup>, Yeni Wahyuningsih<sup>2</sup>, Marji<sup>1</sup>

<sup>1</sup>Department of Information Technology and Computer Science, University of Brawijaya, Malang, Indonesia; <sup>2</sup>Department of Computer Science, University of Brawijaya, Malang, Indonesia.

Email: laililmf@gmail.com, lailil@ub.ac.id, yeni.wahyu12@gmail.com, marji@ub.ac.id

Received April 28<sup>th</sup>, 2013; revised May 30<sup>th</sup>, 2013; accepted June 7<sup>th</sup>, 2013

Copyright © 2013 Lailil Muflikhah *et al.* This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## ABSTRACT

Fuzzy modeling techniques have been widely used to solve the uncertainty problems. A diagnosis of coronary heart disease (CHD) consists of some parameters numerical value of linguistics data. It can be implemented using fuzzy system through construction of the rules which relate to the data. However, the range of linguistics value is determined by an expert that depends on his knowledge to interpret the problem. Therefore, we propose to generate the rules automatically from the data collection using subtractive clustering and fuzzy inference Takagi Sugeno Kang orde-1 method. The subtractive clustering method is a clustering algorithm to look for data clusters that serve as the fuzzy rules for diagnosis of CHD risk. The selected cluster number is determined based on the value of variant boundaries. Hence, it is applied to fuzzy inference system method, Takagi Sugeno Kang order-1, which determines diagnosis of the disease. The advantage of this method is applicable to generate the fuzzy rules without defining and describing from an expert.

**Keywords:** Fuzzy Rule; Coronary Heart Disease; Subtractive Clustering; Takagi Sugeno Kang; Variant

## 1. Introduction

Recently, fuzzy logic can be used to model the process of thinking human involves elements of uncertainty, doubt and linguistics. It is used to settle the classical problem methods due to lack of knowledge [1]. Its characteristic is the ability to express knowledge using linguistic so that it can be understood by human. In most fuzzy system, the determination of fuzzy rules has been established and elaborated by experts. This process takes a long time, experience and the ability of experts [2]. Besides that, the experts cannot often express knowledge correctly and clearly so that the produced rules of fuzzy system are not optimal. Fuzzy clustering algorithm is suitable to generate fuzzy rule due to it can detect partition between input and output variable of data structure for the simple fuzzy rule [3]. Besides that, the strengthness of fuzzy clustering is the computation time efficiently [4]. Therefore, it needs to establish the fuzzy rules, which are automatically based on optimizing data input and output system [5]. One of the algorithms is fuzzily subtractive clustering.

Subtractive clustering is a clustering algorithm based on potential size of the data points in a variable, where the classification is influenced by the radius, squash factor,

accept ratio and reject ratio. The best selected clusters can be obtained using variant analysis. The results of subtractive clustering are clusters which contain the constructed rules and are combined to Takagi Sugeno Kang inference model. Fuzzy inference is important in fuzzy system that the same value exists between expected and actual output. One of the problems which can be solved in fuzzy system is diagnosis of Coronary Heart disease (CHD) due to the constructing rules of diagnosis are numerical data that have linguistic value. CHD is a disease affecting the blood vessels that drain blood to the heart and make a partial blockage or total of the coronary arteries. The detection of CHD risk uses some factors, such as age, sex, occupation, levels of LDL, cholesterol levels, HDL, trigliserid levels, blood pressure systolic and diastolic blood pressure [6]. Hence, we propose to apply subtractive clustering to generate the fuzzy rule for CHD risk.

## 2. The Previous Work

The research for diagnosis of coronary heart disease has been done by Allahverdi N., *et. al.* using Fuzzy Expert system method to determine the risk for the next ten-years. The aquition of rules to determine the risk is obtained

from an expert (specialist doctor). The diagnosis result is depend on his knowledge [7]. Therefore, this paper is purpose to generate rules from the historical data. The automatic rule generation algorithm is based on optimize the process of data input and output system [5]. Several unsupervised learning algorithms have been implemented in various kinds of research. One of the many algorithms is a fuzzy clustering algorithm to generate the rules. Fuzzy clustering algorithm is very suitable to be applied for generating fuzzy rules due to this algorithm can detect partition the separation between input and output variables which are the basic structure of the fuzzy rules and very simple [3]. According to Koczy (2002), the fuzzy clustering has the advantages in computational efficiency. There are several clustering algorithms that can be used, including k-means, fuzzy c-means and subtractive clustering. K-means algorithm and fuzzy c-means algorithm is the unsupervised clustering. The cluster number should be determined initially for the both algorithms. However, the subtractive clustering that is an unsupervised clustering algorithm does not require the number of clusters previously. According to Sindhu (2010), subtractive clustering has the advantage that can estimate the center cluster as well [8].

### 3. Subtractive Clustering

Subtractive clustering is a clustering algorithm based on potential size of the data points in a variable. The idea of the subtractive clustering is the determining regions in a variable which has a high density of data points in surroundings. The point that has the largest number of neighbors will be selected as cluster centers and the density will be reduced. Then the algorithm will choose another point that has the most neighbors to be the center of another cluster [9]. Hereby, the subtractive clustering algorithm is applied to CHD dataset as follows:

- 1) To input data sets of CHD for clustering.
- 2) To determine the attribute value: radius ( $r$ ), squash factor ( $q$ ), accept ratio, reject ratio, minimum value ( $X_{min}$ ) and maximum value ( $X_{max}$ ).
- 3) To normalize the input data as in Equation (1).

$$X_{ij}^{norm} = \frac{X_{ij} - X_{min_j}}{X_{max_j} - X_{min_j}},$$

$$i = 1, 2, \dots, n; j = 1, 2, \dots, m. \quad (1)$$

- 4) To determine the initil potential of each data point. If the number of attributes ( $m$ ) = 1, then it is applied to Equation (2).

$$D_i = \sum_{k=1}^n e^{-4\{Dist_{ki}^2\}} \quad (2)$$

If  $m > 1$ , then it is applied to Equation (3).

$$D_i = \sum_{k=1}^n e^{-4\{\sum_{j=1}^m Dist_{ki}^2\}} \quad (3)$$

where, the distribution value of each attribute is obtained from Equation (4).

$$Dist_{kj} = \left( \frac{T_j - X_{kj}}{r} \right);$$

$$i = 1, 2, \dots, n; j = 1, 2, \dots, m; k = 1, 2, \dots, n \quad (4)$$

- 5) To find the highest potential of each data point using Equation (5).

$$M = \max[D_i / i = 1, 2, \dots, n] \quad (5)$$

where,  $h = i$  so that  $D_i = M$ .

- 6) To detemine the cluster center and reduce potential data point. It can be applied these below stages:

- a) Set cluster center [],  $V_j = C_{hj}$ ,  $C$  = number of cluster, condition = 1,  $Z = M$ .
- b) If condition !=0 and  $Z !=0$ , then it is processed:
  - condition=0 (there is no candidate of new cluster center).
  - if ratio ( $Z/M$ ) > accept ratio, then condition = 1 (candidate is available).
  - Else, it is executed to these below process:
    - If ratio > reject ratio, then the candidate of cluster center will be accepted as center if its existence will make balance of data distantly from the cluster center.
    - $M_d = -1$ .
    - For  $I = 1$  to  $c$ , to apply Equation (6).

$$G_{ij} = \frac{V_j - center_{ij}}{r} \quad (6)$$

where,  $j$  = number of dataset attributes and  $r$  = the defined radius.

- Calculating distance of each data point as is shown in Equation (7).

$$Sd_i = \sum_{j=1}^m (G_{ij})^2 \quad (7)$$

- If  $M_d < 0$  or  $S_d < M_d$ , then  $M_d = S_d$ .
- c) Find the condition = 1, then it is processed as below:

- Set  $C = C + 1$ , where  $C$  is the number of cluster.
- Cluster center  $c = V$  which is the candidate of cluster center.
- To reduce the nearest data point of cluster center.
  - Find  $S_{ij}$  value using Equation (8).

$$S_{ij} = \frac{V_j - X_{ij}}{r_j * q} \quad (8)$$

$$De_i = M * e^{-4\{\sum_{j=1}^m (S_{ij})^2\}} \quad (9)$$

- $D = D - De$ , where  $D$  is the initial potential of data point.
- If  $D_i \leq 0$ , then  $D_i = 0$ .
- $Z = \max[D_i / I = 1, 2, \dots, n]$ .
- Select  $h = i$ , so that  $D_i = Z$ .

7) To return the the normalized cluster center into the initial formation using Equation (10).

$$\text{Center}_{ij} = \text{Center}_{ij} * (X\text{Max}_j - X\text{Min}_j) + X\text{Min}_j \quad (10)$$

8) To calculate the cluster sigma value using Equation (11).

$$\sigma_j = \frac{r_j * (X\text{Max}_j - X\text{Min}_j)}{\sqrt{8}} \quad (11).$$

### 3.1. Variant Analysis

The measurement value of the data spread unsupervised clustering results can be used variance. There are two kinds of variance, *i.e.* variance within clusters and variance between clusters. The both can be applied to determine the density of a cluster. In each stage of cluster formation, the variant is to be calculated as stated in Equation (12).

$$V_c^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y}_c)^2 \quad (12)$$

where,  $V_c^2$  = variance in cluster  $c$ ;  $c = 1 \dots k$ , for  $k$  = number of cluster;  $n_c$  = number of data in cluster  $c$ ;  $y_i = i^{\text{th}}$  data of cluster;  $\bar{y}_c$  = data average in cluster  $c$ .

Thus, the variance within cluster can be calculated using Equation (13).

$$V_w = \frac{1}{N-c} \sum_{i=1}^c (n_i - 1) \cdot V_i^2 \quad (13)$$

where,  $N$  = total the whole data;  $n$  = total data of  $i^{\text{th}}$  cluster;  $V_i$  = variance in  $i^{\text{th}}$  cluster; by variance between cluster ( $V_b$ ), it is used Equation (14)

$$V_b = \frac{1}{c-1} \sum_{i=1}^c n_i (\bar{y}_i - \bar{y})^2 \quad (14)$$

where,  $\bar{y} - \bar{y}_i$  average.

A method that can be used to determine the ideal cluster is variant limit. It calculates the cluster density either variance within cluster ( $V_w$ ) or variance between cluster ( $V_b$ ). The ideal cluster has a minimum  $V_w$  which represents internal homogeneity and maximum  $V_b$  (external homogeneity) [10]. Value of variant boundaries are shown as in Equation (15).

$$V = \frac{V_w}{V_b} \quad (15)$$

### 3.2. Rule Extraction

The results of clustering algorithms in the form of matrix are cluster center ( $C$ ) and sigma ( $\sigma$ ) which will be used to determine the value of Gauss membership function parameters [9]. The steps of fuzzy rule extraction of the cluster are:

1) Calculate the degree of membership by using the

Gauss membership function. Degree of membership of a data point  $X_i$  on the  $k$ -th cluster is shown in Equation (16).

$$\mu_{ki} = e^{-\sum_{j=1}^m \left( \frac{x_{ij} - c_{kj}}{2\sigma_j} \right)^2} \quad (16)$$

2) Create a matrix  $U$  of size  $n \times x$  (cluster number \* number of attributes) by these below steps are:

a) Multiplying the membership degree of each data  $i$  in cluster  $k$  by each attribute  $j$  of the data  $i$ , ( $d_{ij}^k$ ) as in Equation (17).

$$d_{ij}^k = x_{ij} * \mu_{ki} \text{ and } d_{i(m+1)}^k = \mu_{ki} \quad (17)$$

b) Thus, it is the normalization process which is performed by  $d_{ij}^k$  and  $d_{i(m+1)}^k$ .

c) The next step is to construct matrix  $U$  with size  $n \times (r * m + 1)$ , where  $n$  = number of training data,  $r$  = number of rule/number of cluster and  $m$  = number of input variable with  $u_{ij} = d_{ij}^k$ .

3) The results of matrix  $U$  is applied the least square error method to obtain the output parameter coefficient, with details as below:

d) If consequent parameter  $k$  is noted Equation (18).

$$kt = [k_0^t, k_1^t, k_2^t, \dots, k_n^t]^T = \begin{bmatrix} k_0^t \\ k_1^t \\ k_2^t \\ \dots \\ k_n^t \end{bmatrix} \quad (18)$$

Then, inference  $T_S$  for 1 to  $N$  training data can be notated in Equation (19).

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_N \end{bmatrix} = \begin{bmatrix} U_1 & U_1 & \dots & U_1 \\ U_2 & U_2 & \dots & U_2 \\ \dots & \dots & \dots & \dots \\ U_N & U_N & \dots & U_N \end{bmatrix} \cdot \begin{bmatrix} k^1 \\ k^2 \\ \dots \\ k^M \end{bmatrix} \quad (19)$$

e) To calculate coefficient  $k$  can be used Least Square Error method using pseudo matrix, so that the equation of inference  $T_S$  is stated in Equation (20).

$$[U]^T \cdot [U][k] = [U]^T \cdot [Y]$$

and

$$[k] = [k_0^t, k_1^t, k_2^t, \dots, k_n^t]^T \quad (20)$$

The dimension of matrix parameter  $k$  is

$$[k] = (m \cdot (n + 1) \times N) \cdot (N \times 1) = (M \cdot (n + 1)),$$

where  $M$  is the number of training data, and  $n$  is the number of input fuzzy. Thus, it is constructed the collection of rules  $r$  which is stated as below:

[R1] IF ( $x_1$  is  $A_{11}$ ) and ( $x_2$  is  $A_{12}$ ) and .... and ( $x_m$  is  $A_{1m}$ )

THEN ( $Z = k_{11}x_1 + \dots + k_{1m}x_m$ ).  
 [R1] IF ( $x_1$  is  $A_{21}$ ) and ( $x_2$  is  $A_{22}$ ) and...and ( $x_m$  is  $A_{2m}$ )  
 THEN ( $Z = k_{21}x_1 + \dots + k_{2m}x_m$ ).  
 ....  
 [Rr] IF ( $x_r$  is  $A_{r1}$ ) and ( $x_2$  is  $A_{r2}$ ) and...and ( $x_m$  is  $A_{rm}$ )  
 THEN ( $Z = k_{r1}x_1 + \dots + k_{rm}x_m$ ).

### 3.3. Inference Takagi Sugeno Kang Orde-1

Fuzzy inference method of Takagi Sugeno Kang (TSK) is a model for the fuzzy inference rules which are represented in IF-THEN, where the output (consequent) is not a set of fuzzy systems, but it is a constant or linear equations. This method was introduced by Takagi Sugeno Kang in 1985 [11]. There are two models of the TSK method are:

1) TSK model order-0. In general, the fuzzy inference TSK model order 0 is be notated: IF ( $x_1$  is  $A_{1i}$ ) o ( $x_2$  is  $A_{2i}$ ) o ... o ( $x_n$  is  $A_{ni}$ ) THEN  $z = k$ ; where  $x_j$  is the  $j$ -th input varibael,  $A_{ji}$  is a fuzzy set  $i$ th of the variable  $x_j$  and  $k$  is a constant (crisp) as the consequent.

2) TSK model of order-1. In general form the method of fuzzy inference model TSK Order-1 is IF ( $x_1$  is  $A_{1i}$ ) o ( $x_2$  is  $A_{2i}$ ) o ... o ( $x_n$  is  $A_{ni}$ ) THEN  $z = p_1 * x_1 + \dots * x_n + p_n + q$ . where,  $x_j$  is the  $j$ -th input variable,  $A_{ji}$  is the  $i$ -th fuzzy set on the variables  $x_j$ ,  $p_j$  is a constant (to be crisp) as the coefficients for the variables  $x_j$  and  $q$  are constants for a linear equation in the consequent of a rule.

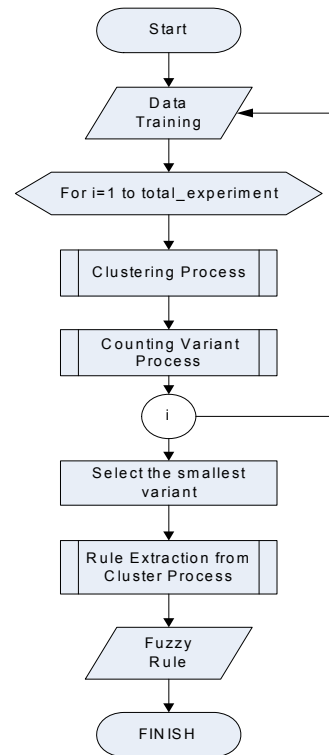
If the fire strength ( $\alpha$ ) and  $Z_r$  values for each the  $r$ -th rule have been obtained ( $r = 1, \dots, R$ ), then the rule will be composed. The composition process is applied by-summation of the multiplication between the fire strength and the  $z$  value. Then, it is implemented defuzzyfication using weighted average concept is shown in Equation (21) [12].

$$z = \frac{\sum_{i=1}^R (aizi)}{\sum_{i=1}^R ar} \quad (21).$$

## 4. Methodology

Generally, the system is divided into two processes including fuzzy rules construction and inference. The first process is to get fuzzy rules through a process of clustering using subtractive clustering method, where the rules selected cluster will be the result of has a limit value that has the smallest variance performed the extraction process. The process of rule construction is shown in **Figure 1**.

The second is testing process. It is a process to determine level of accuracy of rules that are formed through inference process using Takagi Sugeno Kang model 1. The testing process can be carried out if the formation of fuzzy rules has been done. Thus, if the system has never made the process the formation of fuzzy rules, then the user can't do the testing on the system. Input from this process is the test data CHD form data with CHD risk



**Figure 1. Flowchart of constructing rules process.**

factors and fuzzy rules which have been formed from the previous training. However, the output of this process is a diagnosis of CHD risk. The testing process is shown in **Figure 2**.

Furthermore, the dataset is taken from medical record of CHD patients in a hospital and the laboratory of physical health checks in Indonesia. The parameters of CHD risk diagnosis include age, LDL, cholesterol, HDL, trigligerid, sistholic. The details of training data and testing data are shown in attachment-1 and attachment-2. There are three experimental scenario using vario us the number of data training such as 50, 60 and 70 records with reject ratio: 0.1 - 0.9 and radius: 0.35 - 0.45. Each parameter of reject ratio will be produced differrent rules. Thus it is applied to the data test using six attributes, *i.e.* age, LDL, cholesterol, HDL, trigliserida and sistholic presure. In this step, risk value and class of CHD is produced by system will be compare to the actual diagnosis.

## 5. Experimental Result and Analysis

There are two stages of testing the system. The first phase is to determine selected rules of the clustering based process of variant analysis. In this stage, it is carried out three trials with different amount of training data. The first test is used 50 trained data, the second test is used 60 trained data and the third test is used 70 trained data. At each testing, the applied parameters of clustering

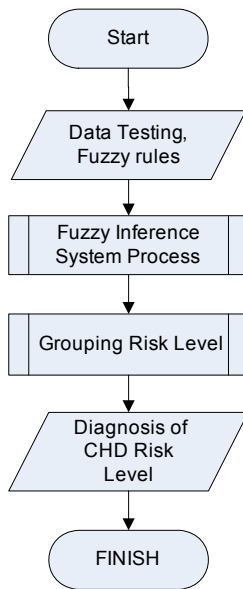


Figure 2. Flowchart of testing process.

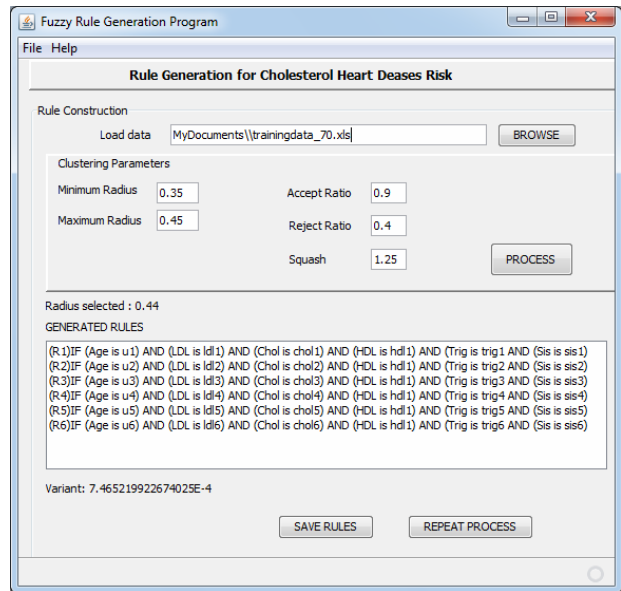


Figure 3. Interface of fuzzy generate rule for CHD system.

process were different. Clustering process parameters in question are value of the radius and the reject ratio. Value of the radius is used in range from 0.35 to 0.45, however the value of reject ratio is used ranged between 0.1 - 0.9. The reject ratio of each training data will produce different rules. The different rules is used in the second phase of testing is to determine how accurate the rules that have been selected. The data used for the testing in the phase is 30 data, which contain six attributes including age, levels of LDL, total cholesterol, HDL, triglerisida and systolic pressure. In this stage, the value and grade level of CHD risk system will be compared to the reference of diagnostic results.

### 5.1. Experimental Result

The experimental result is implementation of fuzzy subtractive clustering for generate rule of CHD risk. The implementation interface is shown in Figure 3.

In the first experiment, rule construction process is executed several times using different reject ratio, between 0.1 and 0.9. Each reject ratio is selected for the ideal cluster number based on the minimum of variance. The experimental result of the first scenario is shown as in Tables 1-3, with  $r$  radius,  $c$  cluster number, and  $v$  variant.

Then, the accuracy result of produced rules from the first experiment with 50 training data is shown in Table 4.

In the second test with 60 training data, it generated 8 different types of rules by the restriction the smallest variance in each reject ratio. The accuracy results of each rule are shown in Table 5.

Then, the third test with 70 training data is produced 9

Table 1. Experimental result with reject value 0.1 - 0.3 (50 training data).

$r$	Reject = 0.1		Reject = 0.2		Reject = 0.3	
	$v$	$c$	$v$	$c$	$v$	$c$
0.35	0.068206	36	0.064765	35	0.09711	32
0.36	0.066968	34	0.070724	33	0.107396	31
0.37	0.066968	34	0.09711	32	0.115339	26
0.38	0.070724	33	0.09711	32	0.120854	25
0.39	0.09711	32	0.107396	31	0.10168	16
0.4	0.107396	31	0.113844	29	0.087595	10
0.41	0.106236	30	0.11272	26	0.079017	9
0.42	0.101653	29	0.099298	23	0.078074	8
0.43	0.090557	28	0.129837	22	0.098716	7
0.44	0.093021	28	0.096542	20	0.092826	5
0.45	0.105097	25	0.085455	16	0.092826	5

Table 2. Experimental result with reject value 0.4 - 0.6 (50 training data).

$r$	Reject = 0.4		Reject = 0.5		Reject = 0.6	
	$v$	$c$	$v$	$c$	$v$	$c$
0.35	0.103759	10	0.140147	7	0.195693	3
0.36	0.132469	9	0.156132	4	0.195693	3
0.37	0.132469	9	0.156132	4	0.195693	3
0.38	0.097146	7	0.195693	3	0.195693	3
0.39	0.112207	6	0.195693	3	0.195693	3
0.4	0.112207	6	0.195693	3	0.195693	3
0.41	0.112207	6	0.195693	3	0.195693	3
0.42	0.112207	6	0.195693	3	0.195693	3
0.43	0.092826	5	0.125741	3	0.125741	3
0.44	0.092826	5	0.125741	3	0.125741	3
0.45	0.092826	5	0.125741	3	0.125741	3

**Table 3. Experimental result with reject value 0.7 - 0.9 (50 training data).**

R	Reject = 0.4		Reject = 0.5		Reject = 0.6	
	v	c	V	c	v	c
0.35	0.195693	3	0.195693	3	0.155954	2
0.36	0.195693	3	0.195693	3	0.155954	2
0.37	0.195693	3	0.195693	3	0	1
0.38	0.195693	3	0.155954	2	0	1
0.39	0.195693	3	0.155954	2	0	1
0.4	0.195693	3	0.155954	2	0	1
0.41	0.195693	3	0.155954	2	0	1
0.42	0.195693	3	0.155954	2	0	1
0.43	0.1047	2	0.1047	2	0	1
0.44	0.1047	2	0.1047	2	0	1
0.45	0.1047	2	0.1047	2	0	1

**Table 4. The accuracy result of test-1.**

Rule type	Total Rule	Diagnosis of CHD		Accuracy rate
		True	False	
Rule #1 (reject = 0.1)	34	3	27	10%
Rule #2 (reject = 0.2)	35	0	30	0%
Rule #3 (reject = 0.3)	8	3	27	10%
Rule #4 (reject = 0.4)	5	16	14	53.33%
Rule #5 (reject = 0.5)	3	21	9	70%
Rule #6 (reject = 0.6)	3	21	9	70%
Rule #7 (reject = 0.7)	2	18	12	60%
Rule #8 (reject = 0.8)	2	18	12	60%
Rule #9 (reject = 0.9)	2	17	13	56.67%

**Table 5. The accuracy result of test-2.**

Rule type	Total Rule	Diagnosis of CHD		Accuracy rate
		True	False	
Rule #1 (reject = 0.1)	1	1	29	3.33%
Rule #2 (reject = 0.2)	1	1	29	3.33%
Rule #3 (reject = 0.3)	5	22	8	73.33%
Rule #4 (reject = 0.4)	5	22	8	73.33%
Rule #5 (reject = 0.5)	3	14	16	46.67%
Rule #6 (reject = 0.6)	3	14	16	46.67%
Rule #7 (reject = 0.7)	3	14	16	46.67%
Rule #8 (reject = 0.8)	2	19	11	63.33%

different types of rules which were bounded the smallest variance in each of reject ratio. The accuracy results of each rule are shown in **Table 6**.

**5.2. Result Analysis**

In the first phase of testing, the parameters input of the radius and reject ratio is affect the number of clusters formed. Both parameters are inversely by the number of clusters. The smaller radius and reject ratio value, the

**Table 6. The accuracy result of test-3.**

Rule type	Total Rule	Diagnosis of CHD		Accuracy rate
		True	False	
Rule #1 (reject = 0.1)	50	2	28	6.67%
Rule #2 (reject = 0.2)	47	1	29	3.33%
Rule #3 (reject = 0.3)	38	0	30	0.00%
Rule #4 (reject = 0.4)	6	19	11	63.33%
Rule #5 (reject = 0.5)	4	17	13	56.67%
Rule #6 (reject = 0.6)	4	17	13	56.67%
Rule #7 (reject = 0.7)	3	16	14	53.33%
Rule #8 (reject = 0.8)	2	23	7	76.67%
Rule #9 (reject = 0.9)	3	16	14	53.33%

more number of clusters. However an accept ratio didn't give any influence on number of clusters. The number of clusters was generated in each testing and different training using the input parameter data of the same clustering process is different. This was caused of influence of training data used for the largest and smallest values in each attribute data. The largest and smallest values respectively one data attribute is useful in forming the data normalized, where the normalized data is effect to the formed cluster.

The second phase of the test results show that affect the number of rules established accuracy of test data. In many times of testing, it was generated the same number of rules, but different levels of accuracy. The different accuracy is caused of the cluster center and the produced sigma value. The both effect to output coefficient value which use to determine Z value of each rule. The Z value was applied to diagnosis CHD risk. The best accuracy of the generating fuzzy rules using the subtractive clustering method is 76.67% for diagnosis of CHD risk. This is caused by several factors, such as the number of attributes is 7. Also the number of data should be used not only 50, 60 and 70 patient's medical record but also more types data training which the system will be better in recognizing patterns. Another factor that causes this accuracy is CHD risk values which are used as a reference for accuracy of the system.

**6. Conclusion**

Subtractive clustering method can be implemented in the establishment of fuzzy rules for the diagnosis of coronary heart disease risk (CHD). The method is used to produce the clustered data. Thus, the fuzzy rule is obtained by extracting rules from the cluster using the Least Square Error (LSE). Determination of the best cluster is a rule fuzzy selected at each time of testing performed using variant analysis. The smaller value of variant boundaries for cluster, the more ideal for clustering is. The best accuracy of the system of selected rules if it is applied to the TSK inference orde-1method is 76.67%, where the

number of rule is 2 which is the result of the clustering process using the reject ratio = 0.8 at radius 0.45 for 70 training data.

## REFERENCES

- [1] A. Priyono, M. Ridwan and A. Jais Alias, "Generation of Fuzzy Rules with Subtractive Clustering," *Jurnal Teknologi*, Vol. 43, 2005, pp. 143-153. [doi:10.11113/jt.v43.782](https://doi.org/10.11113/jt.v43.782)
- [2] R. Arapoglou, K. Kolomvatsos and S. Hadjiefthymiades, "Buyer Agent Decision Process Based on Automatic Fuzzy Rules Generation Methods," *IEEE International Conference on Fuzzy Systems (FUZZ)*, Barcelona, 18-23 July 2010, pp. 1-8.
- [3] J. A. Dickerson and B. Kosko, "Fuzzy Function Approximation with Ellipsoidal Rules," *IEEE Transactions on System Man and Cybernetics*, Vol. 26, No. 4, 1993, pp. 542-560.
- [4] L. T. Koczy, "Fuzzy Rule Base Model Identification Techniques," Department of Telecommunication and Telematics, Budapest University of Technology and Economics, Budapest, 2002.
- [5] H. Salehfar, N. Bengiamin and J. Huang. "A Systematic Approach to Linguistic Fuzzy Modeling Based on Input-Output Data," *Proceedings of the 2000 Winter Simulation Conference*, Orlando, 10-13 December 2000, pp. 480-486. [doi:10.1109/WSC.2000.899755](https://doi.org/10.1109/WSC.2000.899755)
- [6] P. R. J. Burch, "Coronary Disease: Risk Factors, Age, and Time," *American Heart Journal*, Vol. 97, No. 4, 1979, pp. 415-419. [doi:10.1016/0002-8703\(79\)90385-5](https://doi.org/10.1016/0002-8703(79)90385-5)
- [7] N. Allahverdi, S. Torun and I. Saritas, "Design of a Fuzzy Expert System for Determination of Coronary Heart Disease Risk and Technologies," *Computer Systems and Technologies 2006*, Veliko Turnovo, 14-15 June 2007.
- [8] R. S. Sidhu, "A Subtractive Clustering Based Approach for Early Prediction of Fault Proneness in Software Modules," World Academy of Science, Engineering and Technology 67, 2010.
- [9] S. Kusumadewi and H. Purnomo, "Fuzzy Logic Application for Decision Support System," Graha Ilmu Press, Jakarta, 2010.
- [10] L. Man, T. C. Lim, S. Jian and L. Yue, "Supervised and Traditional Term Weighting Metodes for Automatic Text Categorization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 31, No. 4, 2009, pp. 721-735.
- [11] N. dan Gelley and R. Jang, "Fuzzy Logic Toolbox," Mathwork, Inc., Natick, 2000.
- [12] K. H. Lee, "First Course on Theory and Applications," Springer-Verlag Berlin Heidelberg, New York, 2005.