

# Generation of Fuzzy Rules by Subtractive Clustering

**Hussen Ateya Lafta**

*Department of computer science, collage of science for woman, Babylon University.*

[am.2222@yahoo.com](mailto:am.2222@yahoo.com)

**Zahraa A. Mohammed**

*Department of software, College of Information Technology, Babylon University*

[A\\_zahraa\\_t@yahoo.com](mailto:A_zahraa_t@yahoo.com)

## Abstract

This work depends on two stages. First one, "subtractive method", clustering algorithm, used for identifying the relationships between data points in order to build system, where the data point gathers with other points to make cluster of the same features. These groups will be used in the second part of the work to construct fuzzy IF...THEN rules, which controls how the system works. The number of rules and its parts depend on these clusters. While the Takagi-Sugeno Kang (TSK) fuzzy inference modal was used. The scope of this work is applied to heart disease diagnosis.

**Key word:** Fuzzy rules, First\_order (TSK) fuzzy inference, Subtractive clustering, Heart disease diagnoses.

## الخلاصة

هذا العمل يعتمد مرحلتين، المرحلة الاولى يستخدم خوارزمية "التجميع" والتي تستخدم لتحديد العلاقات بين عناصر البيانات لكي يبنى النظام، بحيث كل نقطه في البيانات تجمع مع نقاط اخرى ذات نفس المواصفات لكي تكوّن مجاميع.

هذه المجاميع (العناقيد) سوف تستخدم في المرحلة الثانية من العمل لبناء مجموعة من القواعد المضببة والتي يصطلح عليها (IF...THEN rules). والتي ستحدد مسار عمل النظام.

عدد القواعد ومتغيرات الادخال والايخراج لديها تعتمد على المجاميع (العناقيد) المتولدة في المرحلة الاولى بينما يستخدم نظام الاستدلال المسمى (King TSK-Sugeno). طبق العمل لتشخيص مرض القلب.

الكلمات المفتاحية: القواعد المضببة، الاستدلال المضبب First-order TSK , خوارزمية التجميع، تشخيص مرض القلب.

## 1. Introduction

Fuzzy system is an interesting way to deal with data and with any type of features even there are undesirable features like imprecise, noisy or missing input information [Keshavarzi, *et al*, 2012], also with complex system since it depends on input-output high dimensional data rather than human experience, which maybe uncertain or unavailable, or complex mathematical equations to describe the manner of the system and build its rules [Berneti, 2011].

In general, clustering (hard clustering) concept is each data point is belonging to only one cluster according to some measures, while in fuzzy clustering (soft clustering) the data points may belong to more than one group, or to all groups in some ratio according to membership degree, and the cluster with high value, the data point is assigned .

There are many clustering algorithms are K-mean, C-mean, Mountain and Subtractive clustering. These algorithms belong to hard or soft clustering, but each one has its limitations like initial cluster center selection, pre-knowledge of number of clusters, dead unit problem, multiple cluster membership and premature convergence to local optima [Ghorpade, 2014].

Subtractive clustering algorithm is a suitable way to solve these limitations because it doesn't need to know the number of cluster or estimate the cluster center. It depends on some parameters that influence the result like radius, squash ratio, accept ratio, reject ratio [Keshavarzi, *et al.*, 2012]. After finding all the clusters in the data, the work is going toward generation rules. The number of rules is equal to the number of clusters. The antecedent of rule takes from clusters data (input data) and the consequent part of the rule represents the behavior of the system [Engineering, 2006], this is according to Takagi-Sugeno-Kang (TSK) fuzzy inference which may be a linear function and this type is called **first-order** Sugeno fuzzy mode, or **zero-order** Sugeno fuzzy model if that consequent is constant [Farias, *et al.*, 2011].

## 2. The Previous Work

Diagnose heart disease is sensitive and important. Earlier diagnose of this disease is one of the treatment ways and the correct diagnoses leads to save the patient's life. One choice of this diagnoses is usually done in the hospital by the doctors depending on the taking information from the patient and making test. Recent science techniques are used. Fuzzy logic can help successfully in this space such as Fuzzy Petri Net, one of these techniques. This used petri net to evaluate the patient case to determine if he is ill and the degree of the illness. Genetic algorithm is also used to generate fuzzy rules. Particle Swarm Optimization (PSO) algorithm is another way used in fuzzy rule extraction space. Fuzzy clustering, which is used in this paper can build rules base to determine the infected according to the symptoms of the sick person.

## 3. The Proposed Subtractive Clustering algorithm

The steps of Subtractive clustering algorithm are:

**Step1 Normalization:** starting by handling multidimensional input data, N dimensions, M attributes, like  $X_i = \{x_1, x_2, \dots, x_m\}$  where  $X_i$  is a vector of inputs, data is normalized to interval [0,1], by equation 1 [Priyono, Ridwan, *et al.*, 2005].

$$X_{ij}^{norm} = \frac{X_{ij} - X_{minj}}{X_{maxj} - X_{minj}} \dots \dots \dots (1)$$

Where  $i=1,2,\dots,n$ ;  $j=1,2,\dots,m$ .

**Step2 Assigned:** constant parameters are given carefully according to the work's modality and data's type, the parameters are *radius*, *squash factor* ( $\eta$ ), *accept ratio*, and

reject ratio ( $\epsilon$ ). Radius  $r_a$  (influences range of cluster center), affects the number of rules seriously, since the number of rules depends on number of clusters. So large radius value means few number of clusters and vice versa [Berneti, 2012; Priyono *et al.*, 2005].

**Step3: Compute center:**

**A-compute first potential:** The subtractive clustering algorithm differs from other clustering algorithms; it is not interested in early knowing of number of clusters, therefore the algorithm works based on consideration of each data points as a cluster center [Berneti SM].

Thus the measure of density to each point is achieved and the potential is computed by equation 2 [Farias *et al.*, 2011; Ghorpade-aher & Metre, 2014].

$$\rho_i = \sum_{j=1}^m e^{-\alpha \|x_i - x_j\|^2} \dots\dots\dots (2)$$

Where

$$\alpha = \frac{4}{r_a^2} .$$

$\|.\|$  denotes the Euclidean distance.

**ra:** is positive constant defined as cluster center range.

**B-Compute first center:** when the data point's potential is the biggest one ( $\rho_1^*$ ) compared with the other potentials ( $\rho_1^i$ ), first cluster center  $\chi_1^*$  is declared [Farias, *et al.*, 2011; Chiu, 1997].

**C-Compute the other centers:** trying to find the next center, subtract a potential of other data ( $\rho_i^*$ ) from the potential of first cluster ( $\rho_1^*$ ), if the difference is big, this means it is far from influence of center 1 and input value regarding with this potential is candidate to be next center. This is achieved by equation 3 [Priyono *et al.*, 2006; Farias *et al.*, 2011].

$$\rho_i = \rho_i - \rho_1^* \exp(-\beta \|x_i - x_1^*\|^2) \dots\dots\dots(3)$$

Where

$$\beta = 4/r_b^2$$

$$r_b = \eta r_a$$

$\eta$  = squash factor

**D-Compute k-center:** after k-time the equation above becomes as the equation below... [Berneti, 2011; Farias *et al.*, 2011]

$$\rho_i = \rho_i - \rho_k^* \exp(-\beta \|x_i - x_k^*\|^2) \dots\dots\dots(4)$$

Where

$\chi^k$  =center of the kth cluster.

$\rho^k$  = its potential value.

**Step4-Testing:** each data point is candidate to be cluster center and must be tested in some formula which specify it as a center point or not [Keshavarzi , Sarmadian et al,2012],[ Priyono , Ridwan et al, 2005] These formulas are:

**Formula 1-**

$$\frac{d_{min}}{r_a} + \frac{\rho_k}{\rho_k^*} > 1$$

Where

$d_{min}$ =smallest distance from center i to all earlier found centers [Keshavarzi *et al.*, 2012; Farias *et al.*, 2011].

If the data point with high potential realizes that condition, then it is successfully declared as a new center.

**Formula 2-**

$$\rho_k^* < \varepsilon \rho_i^*$$

Where:

$\varepsilon$  =reject ratio.

The algorithm is terminated if this condition is not satisfied [ Farias, *et al.*, 2011].

Many parameters are adapted in this algorithm these are:

**Radius:** a range of numbers found in the cluster of that center.

**Accept ratio:** gives its agreement to the candidate center to be center I, if condition 1 is true with that data points. Otherwise, potential of data point assign to zero and begin with another candidate.

**Reject ratio:** each cluster center yield ( $\chi^k$ ), its potential compared with the potential of center i, with the reject ratio ( $\rho_k^* < \varepsilon \rho_i^*$ ). And the result determines continuous generation centers or stop and begin to show the result [Keshavarzi *et al.*, 2012].

The algorithm can be summarized as:

Step1: by eq2.calculate the potential  $\rho_i$  of each data point  $1 \leq i \leq n$ .

Step2: find center  $i$  by find the value  $X_i$  with max potential  $\rho_i$ .

Step3: revise the potential value of each data point by eq3.

Step4: if  $\rho_k^* > \varepsilon \rho_i^*$

Accept  $X_i$  as center and continue.

Else if  $\rho_k^* < \varepsilon \rho_i^*$

Reject  $X_i$  and stop.

Else

Let  $d_{min}$  =shortest of the distances between  $X_k^*$  and all previously found cluster centers.

If  $\frac{d_{min}}{r_a} + \frac{\rho_k}{\rho_k^*} > 1$

Accept  $X_i$  as center and continue.

else

reject  $X_i$  and set the potential at  $X_i$  to 0.

Select the data point with the next highest potential as the new  $X_i$  and reset.

end if

end if.

#### 4. Generation fuzzy rules:

A clusters are building. Depending on clustering algorithm results, the fuzzy rules then must be created. Number of rules associates with number of clusters .So the rule for class  $C_i$  can take the form: (Abonyi, 2003; Farias *et al.*, 2011).

Rule  $i$ : IF  $X$  is about  $\chi_i^*$  THEN class is  $C_1$ .

So to be ready to generate fuzzy rules the formula above can be more conventional form [Chiu, 1997]:

Rule i: **IF** X1 is  $\mathcal{A}_{Ei1}$  **And** X2 is  $\mathcal{A}_{Ei2}$  **And** ... **THEN** class is C1

Where:

$X_j$ =input data element.

$j=1 \dots n$ , n is number of input.

$i=1 \dots m$ , m is number of cluster.

$\mathcal{A}_{Eij}$ =membership function related with each  $x_j$  in rule i (linguistic variable).

The membership can be calculated by the equation below, where Gaussian function applies [Abonyi , Szeifert,2003].

$$\mathcal{A}_{Eij}(X_j) = \exp\left\{-\frac{(X_j - \chi_{ij}^*)^2}{\delta_{ij}}\right\} \dots\dots\dots(5)$$

Where

$\chi_{ij}^*$  = the center.

$x_j$ =the j element with xi input.

$\delta_{ij}$ = sigma value generated from subtractive clustering method.

If there is more than one feature to consist the individual rule then the connective tool, And, Or, Not, is used. And each one has its effects on the result [Keshavarzi , Sarmadian, 2012].

And the degree of fulfillment of each rule also called firing strength is computed by using (Samsudin, Ruhidah).

$$\beta_i(x_k) = w_i \prod_{j=1}^n A_{ij}(X_{j,k}) \dots\dots\dots (6)$$

$w_i$  is weight of the rule and in most time it is equal to 1.

The consequent of the rule that describes the system behavior, as mentioned before, and according to (TSK) fuzzy inference, is either polynomial equation in first order (TSK) fuzzy inference as

**Rule i-** if X is  $\mathcal{A}_{E1}$  and Y is  $\mathcal{A}_{E2}$  then  $A=SX+TY+V$

Where x, y are input features.

Or be constant in zero order (TSK) fuzzy inference so the equation above becomes:

**Rule i-** if X is  $\mathcal{A}_{E1}$  and Y is  $\mathcal{A}_{E2}$  then  $A=V$ .

Where parameters S and T are equal to 0

### 5. Testing Rules

Now rules base is formed from train input data set .by implementing of these rules, on test data set, and rule base is created before. One or more rule reacts to such case. After applying these set of rules and the results are computed according to connective tool, to tight the antecedent parts together, used. The output is acquired as the formula below [Samsudin, 2009]:

$$C = \frac{\sum_{i=1}^N a_i c_i}{\sum_{i=1}^N a_i} \dots \dots \dots (7)$$

Where

C=centroid

N: number of trigged rules.

A<sub>i</sub>=minimum value of each rule.

C<sub>i</sub>=result of each rule.

The figure below shows all previous steps of proposed system:

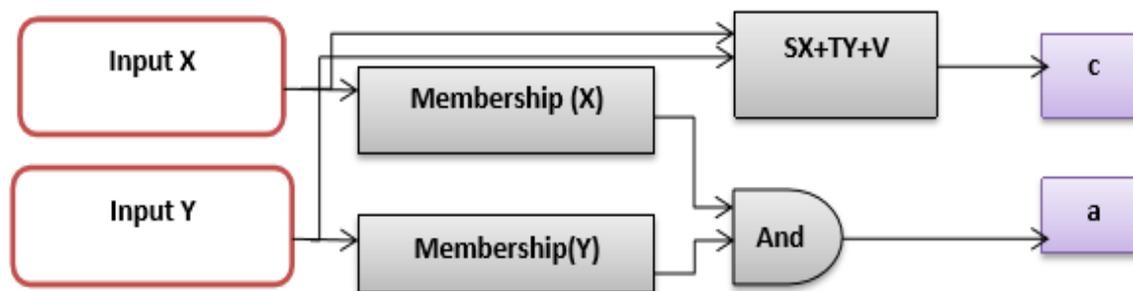


Figure 1: Sugeno fuzzy inference processing model.

### 6. Experimental Results

Heart disease with its database can be applied in this system .By taking 10 symptoms of the disease for 100 patients.

First, subtractive clustering algorithm is applied and clusters are built according to centers created. Giving radius, squash factor (η), accept ratio, and reject ratio (ε) its value as 0.5, 0.25, 0.5, 0.15, respectively. So the results are: number of centers is 56 centers. As shown in figure (2).

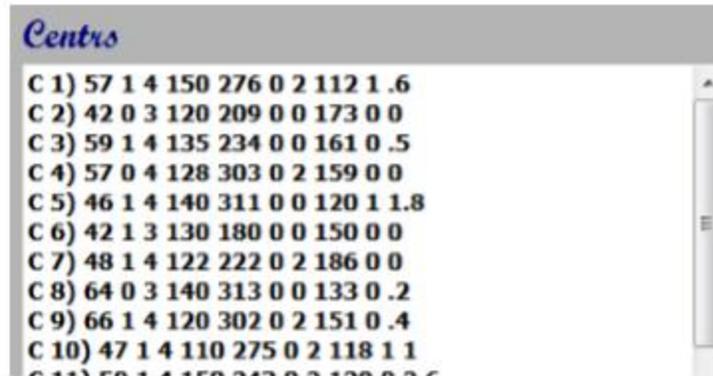


Figure (2): The centers value of subtractive clustering algorithm.

And the sigma values, which their numbers are related to number of column (symptoms), are shown in figure 3:

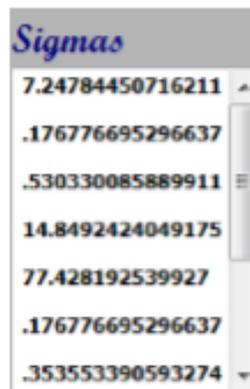


Figure (3): The sigma values for subtractive clustering algorithm.

Sigma values are used in addition to the set of centers, to contribute the rule. The fuzzy rules formed depended on this set of centers and can be shown in the figure below.

```

Rule 1-If (Age is Age 1) AND If (Sex is Sex 1) AND If (ch_pain is Chest_pain 1) AND
Rule 2-If (Age is Age 2) AND If (Sex is Sex 2) AND If (ch_pain is Chest_pain 2) AND
Rule 3-If (Age is Age 3) AND If (Sex is Sex 3) AND If (ch_pain is Chest_pain 3) AND
Rule 4-If (Age is Age 4) AND If (Sex is Sex 4) AND If (ch_pain is Chest_pain 4) AND
Rule 5-If (Age is Age 5) AND If (Sex is Sex 5) AND If (ch_pain is Chest_pain 5) AND
    
```

Figure (4): Set of fuzzy rules for heart disease.

By using Gaussian membership function which depend on centers and sigma's value and the using formula of

$$C = \frac{\sum_{i=1}^N a_i c_i}{\sum_{i=1}^N a_i}$$

The final output is made. And it approximately can diagnose the case, where the accuracy rate =76.43

## 7. Conclusion

This paper has shown the act of a TSK fuzzy inference with using Subtractive clustering technique. This approach has been tested with database of heart disease in order to diagnose the case of the patient whether infected or not.

As shown from the result of this works that subtractive clustering algorithm and by using some parameters like radii, accept and reject ratio is preferred to build efficient fuzzy rule and give much better result for many problems.

## Reference

- Berneti SM. (2011); "Design of Fuzzy Subtractive Clustering Model using Particle Swarm Optimization for the Permeability Prediction of the Reservoir". International Journal of Computer Applications (0975 – 8887) Volume 29– No.11.
- Chiu SL. (1997); "An Efficient Method for Extracting Fuzzy Classification Rules from High Dimensional Data". J. Advanced Computational Intelligence, Vol. 1, No. 1, 1997.
- Engineering M (2006). "Type-2 Takagi-Sugeno-Kang Fuzzy Logic Modeling using Subtractive Clustering". IEEE Xplore, DOI: 10.1109/NAFIPS.2006.365871 .
- Farahbod F, Eftekhari M. (2013) "A NEW CLUSTERING-BASED APPROACH FOR MODELING FUZZY RULE-BASED CLASSIFICATION SYSTEMS", IJSTE, Volume 37, Issue 1, May and June 2013, Page 67-77 .
- Farias MS, Nedjah N, Mourelle LDM. (2011) "RADIONUCLIDE IDENTIFICATION USING SUBTRACTIVE".INAC.

- Ghorpade-aher J, Metre VA. (2014) "Clustering Multidimensional Data with PSO based Algorithm", *Soft Computing and Artificial Intelligenc* .
- Janos Abonyi, Ference Szeifert. (2003) "Supervised fuzzy clustering for the identification of fuzzy classifiers". *Pattern Recognit Letter.*; 24(14):2195-2207.
- Keshavarzi A, Sarmadian F, Rahmani A, Ahmadi A, Labbafi R, Iqbal MA. (2012), "Fuzzy clustering analysis for modeling of soil cation exchange capacity". *Australain Journal of Agricultural Engineering. AJAE* 3(1):27-33.
- Priyono a, Ridwan M, Alias A.( 2005). "Generation of fuzzy rules with subtractive clustering".[http://www.jurnalteknologi.utm.my/index.php/jurnalteknologi/article/view/782;43\(D\):143-153](http://www.jurnalteknologi.utm.my/index.php/jurnalteknologi/article/view/782;43(D):143-153).
- Samsudin R. Rice Yields,( 2009), "Time Series Forecasting Using ANFIS".