

Optimization of Membership Function of Fuzzy Rules Generated using Subtractive Clustering

Hussein A. Lafta^{†*} and Zahraa A. Mohammed^{‡*}

[†]Computer Science Department, Collage of Science for Woman, Babylon University, Iraq

[‡]Software Department, College of Information Technology, Babylon University, Iraq

Accepted 15 May 2016, Available online 22 May 2016, Vol.6, No.3 (June 2016)

Abstract

Fuzzy rules regarded as a good way to represent the knowledge in many type of problems. It shorten the facts found in the problem at hand in form of **IF...THEN**, and the membership function regarded as basic part in the structure of these rules. The work is divided into three parts: first one is dealing with clustering process to extract the centers values; the estimation of the centers from multidimensional data set is done by using subtractive clustering algorithm. These centers are converting to fuzzy rules in the rule base, after applying a Gradient Descent method fortuning the membership function parameters. The tuning and applying Fuzzy Inference System are the second and third stage of this work. The scope of the work is heart disease diagnose.

Keywords: Fuzzy Rules, Subtractive Algorithm, Sugeno Inference System, Gradient Descent Method.

1. Introduction

Fuzzy Inference System which is termed (FIS) used the fuzzy logic and fuzzy set to represent the knowledge in interpretable form. Fuzzy Inference is an attractive way used in many problems in many fields, like control, decision make, process models, and pattern classifiers (Castellano *et al*,2003).

FIS has two parts, the first part deal with convert the number of values (Input/output)to antecedent, consequent of IF...THEN rule. Each value can divided to some ranges (the fuzzy set) with using some labels to name these ranges. This part of FIS must choose one of available kinds of membership functions choose in addition to choose fuzzy connective tool like (And, Or) with determine their function. The second part begins by using these tools chosen before. By adopting one of FIS model the system output can gain. This output is result of aggregation of the rules found in rules base (Tawafan *et al*,2012).

Clustering process include grouped the data points in to some number of clusters, so the dealing is with these clusters instead of each point, this mean using the centers points of these clusters to knowledge representation. Clustering (grouping) is saving more computational which make it an efficient way especially with big data set.

Subtractive clustering algorithm regarded the best selection duo its ability to estimate the clusters centers without needing of determine the initial number of

these centers, that most of clustering algorithms safer from (Chiu,1997).

The results (centers values)changing influenced by four parameters: radius, squash factor, accept ratio and reject ratio.

The work start with using subtractive clustering algorithm to select clusters centers, then convert the centers set to rules set (fuzzy if... then rules). This step is complete by adopting Sugeno Inference System. Optimization step done with using gradient descent method. The method used to adjust the membership function, where Gaussian membership function was selected.

2. Literature Review

This section surveys some of the recent works that are most related to the current research.

In 2010 Lawrence O. Hall and Petter proposed a fuzzy rule based decision trees. This method tries to produce many rules directly from data. Fuzzy entropy was used to reduce the size of the decision tree and hence reduce the number of fuzzy rules generated from this decision tree (Hall; Petter,1998).

In 2011 Jingjing Cao and SamKwong used multi-objective evolutionary hierarchical algorithm for acquiring fuzzy rules for classifier purpose and for enhancing the accuracy, a reduce-error based ensemble pruning method used. The fuzzy rule was representing as chromosome with tree different genes (control, parameter and rule genes). In this research if similar classifiers rules found then they removed in

*Corresponding author: **Zahraa A. Mohammed**

order to preserve the diversity of the fuzzy system (Debnath,2013).

In 2012 Priscilla A. Lopes and Heloisa A. Camargo adapted the label data to generate the fuzzy rules. Supervised learning algorithm was used to partially labeled data set, and then the output of the algorithm labeled the other unlabeled date. The complex problems can use this technique for its ability to partially label the data set (Al-Shammaa;Maysam, 2014).

In 2012 Keon-Jun Park, Jong-Pil Lee and Dong-Yoon Lee introduced a combination of fuzzy system and neural networks using multiple-output by using clustering algorithm. Fuzzy C-mean clustering algorithm was used to portioning the input space in order to get a set of clusters which describe the fuzzy rules and the rule's number is corresponding with number of these cluster's number. The back propagation neural network is used to learn the coefficients of consequent of rules (multiple-output) (Park et al ,2012).

In 2013 Sree Bash Chandra Debnath1, and Pintu Chandra Shill et al, were adapting a particle swarm optimization (PSO) as a new method to extract fuzzy rules in addition to punning them. Using PSO method as an adaptive way to adjust the fuzzy rules by adjusting the parameters found in membership function, where Gaussian function was used. the method achieve its aim buy produce more flexibility, higher robustness fuzzy rules with high performance (Debnath,2013).

3. The proposed System

Firstly, the clustering algorithm (Subtractive) was explained. Subtractive is a fast, one pass algorithm. It is distinct from other algorithm of clustering; it doesn't need to know the number or values of the initial centers. The proposed system starts with split the available data to their basic classes and then applying subtractive clustering to each class. Assuming $\{x_1, x_2, \dots, x_n\}$ is the data points of class 'A', x_i is a vector of input feature space, this vector space is normalized using equation below to convert the range of values between 0 and 1 (Priyono; Alias ,2005).

$$normx_i = \frac{x_i - X_{maxj}}{X_{maxj} - X_{minj}} \tag{1}$$

For one class the algorithm regarding every data points a candidate to be cluster center. The base of compute the potential for data is depend on density, each data element has to compute its surrounding neighbors. The point with highest neighbors has highest potential. And this point is candidate to be a cluster center Equation 2 used to compute the potential of data value (Berneti, 2011),(Farahbod; Eftekhari ,2013).

$$Pi = \sum_{j=1}^n e^{-\alpha \|x_i - x_j\|^2} \tag{2}$$

Where

$$\alpha = \frac{4}{ra^2}$$

$\|.\|$ denotes the Euclidean distance and ra (radius) is a positive constant, the radius considers the most affective parameters in addition to the other parameters. Radius effective seriously on the number of the clusters if the radius value was big, the result was small number of big clusters and verse vises. After determined the first center x_i^* which is related to high potential P_1^* , the potential revise of each data point. Equation (3) shows this computation (Farias ;Nedjah ,2011).

$$Pi = P_1 \times e^{-\beta \|x_i - x_i^*\|^2} \tag{3}$$

Where

$$\beta = \frac{4}{rb^2}$$

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

$$r_b = \eta r_a$$

η = squash factor

After many time the new potential values revise from their previous potential values. Equation 4 explains this (Farias ;Nedjah ,2011).

$$P_i = P_i - P_1^* e(-\beta \|x_i - \chi_i^*\|^2) \tag{4}$$

Where χ_k^* is center value of the kth cluster and P_k^* is its potential value.

Many conditions control the clustering process in subtractive clustering. Accept and reject ratio determine acceptable of candidate center, while the formula below can stop the whole algorithm (Priyono; Alias ,2005).

$$P_{ik}^* < \epsilon P_i^* \tag{5}$$

Where

ϵ = reject ratio.

Also subtractive clustering method responsible to produce the sigma values according to next formula:

$$\sigma_j = \frac{ra * (X_{maxj} - X_{minj})}{\sqrt{8}} \tag{6}$$

Where

σ_j is sigma value

ra radius value

X_{maxj} is max value in jth attribute

X_{minj} is min value in jth attribute.

4. Generating of fuzzy rules

After clustering stage, the rule base must make the rules numbers and their values corresponding to

number and values of centers set. Rule base consist of merge the rules produce from each main class, Fuzzy Inference System adopted in this work is Sugeno model.

And the general form of rule in particular class is(Cordón *et al*,2001):

R_i : if X_j is A_{i1} and Y_j is A_{i2} then class is C_i

Where X_j, Y_j are j 'th input feature, A_{ij} is the membership function of the i 'th rule. Gaussian function was chosen to be membership function.

$$A_{ij}(X_j) = e^{-1/2(\frac{X_j - X_j^*}{\sigma_j})^2} \tag{7}$$

λ is error rate which is assign to 0.15, μ is the degree of fulfillment of rule i , (+) sign is allocated for the rule that gave $\mu_{c,max}$, and sign (-) is allocated to rule gave $\mu_{-c,max}$ the two rule (Chiu,1997).

6. Results and Discussion

The database of the proposed system is Heart disease diagnosis system which is taken from University of California, Irvin (UCI). This dataset contains 13 attributes show the symptoms of the 270 patients and with 1 attribute represent the output. The input attributes are age of the patient in year, sex, chest pain type, resting blood pressure in mm (Trestbps), serum cholesterol in mg (Chol), fasting of blood sugar (fbs), resting electrocardiographic results (ECG), maximum heart rate achieved, induced-angina, old-peak, Slope, vessels and thalach). The output attribute is the disease status of heart of patients. The proposed system firstly classify the database to its two main classes (Normal/Abnormal), then applied subtractive algorithm to each class. The results of subtractive algorithm are value set for sigma values which its number is equal to number of input attributes(13 attributes) and centers set. Table.1 shows the sigma values of 13 symptoms. The clustering algorithm is influence by their four parameters (*radius value, squash factor, accept ratio and reject ratio*) but most affected one is the radius value (influence the range of cluster center), so the algorithm applied with range of radius values from 0.1 to 0.9. As divided the heart data set to normal/abnormal main classes, two kinds of centers yield, Table.2 and 3 shows example of the two types of centers. These kinds of centers are converting to Sugeno inference model of fuzzy rules and then combine together to form fuzzy rules base. The diagnoses system for heart failure is evaluated by the accuracy rate, in order to assess strength of the system to get the right diagnose. So the patient with the heart disease is detect to abnormal class, and the uninfected one can be detected to normal class. An accuracy low calculates the rate of the right diagnose to the total number of cases, equation below shows this meaning.

$$\text{Accuracy Rate} = \frac{\text{number of right diagnose cases}}{\text{number of total cases}} \times 100\% \tag{11}$$

Table.1: The 13 Value of Sigma Values

#No	Sigma
1	13.3643181644257
2	.318198051533946
3	.3954594154601839
4	33.7289934625983
5	139.370746571868
6	.318198051533946
7	.636396103067893
8	39.4565583902094
9	.318198051533946
10	1.90918830920368
11	.636396103067893
12	.954594154601839
13	1.27279220613579

Table 2 Centers Values for Normal Class

#No	Centers values
1	54 1 4 124 266 0 2 109 1 2.2 2 1 7 1
2	57 1 3 128 229 0 2 150 0.4 2 1 7 1
3	62 1 4 120 267 0 0 99 1 1.8 2 2 7 1
4	46 1 4 140 311 0 0 120 1 1.8 2 2 7 1
5	67 1 4 120 229 0 2 129 1 2.6 2 2 7 1
6	59 0 4 174 249 0 0 143 1 0 2 0 3 1
7	61 0 4 130 330 0 2 169 0 0 1 0 3 1
8	54 1 4 110 206 0 2 108 1 0 2 1 3 1
9	63 1 4 130 254 0 2 147 0 1.4 2 1 7 1

Table.3 Centers Values for Abnormal Class

#No	Centers values
1	44 1 3 120 226 0 0 169 0 0 1 0 3 0
2	37 0 3 120 215 0 0 170 0 0 1 0 3 0
3	51 0 3 120 295 0 2 157 0.6 1 0 3 0
4	45 0 4 138 236 0 2 152 1 2 2 0 3 0
5	45 1 2 128 308 0 2 170 0 0 1 0 3 0
6	46 0 2 105 204 0 0 172 0 0 1 0 3 0
7	50 0 3 120 219 0 0 158 0 1.6 2 0 3 0
8	44 1 3 140 235 0 2 180 0 0 1 0 3 0
9	42 0 3 120 209 0 0 173 0 0 2 0 3 0
10	67 0 3 152 277 0 0 172 0 0 1 1 3 0
11	55 1 2 130 262 0 0 155 0 0 1 0 3 0

Table.4 displays the different numbers of fuzzy rules as using different value of radius, also explain the accuracy rate for each set of rules before and after

applying the tuning method (gradient descent method).

Table 4 Different number of Fuzzy rules according to different rang of radius values with their accuracy rate before and after optimization

#No	Radius value	Fuzzy Rules numbers	Accuracy rate before punning (%)	Accuracy rate after punning (%)
1	0.9	24	73.5	79.8
2	0.8	19	72.2	85.8
3	0.7	14	73.5	75.8
4	0.6	12	72	77.5
5	0.5	8	63	76.4
6	0.4	5	63.5	70.8
7	0.3	3	63	64.4
8	0.2	2	44	44
9	0.1	2	44	44

Conclusion

TSK fuzzy inference model is regarded an efficient way to control a problem. Subtractive clustering method give a good result if it uses radius value is assign to 0.8 and the other parameters are assign as flowing, squash factor is equal to 0.15, accept ratio is equal to 0.5 and reject value is equal to 1.25.

Centers number according to these assignments is 24 centers, by TSK model convert to 24 fuzzy rules used to diagnose the patient case.

The accuracy rate for the system is 72.2 before the adjusting and it is arise to 85.8 when gradient descent method was used.

Reference

Al-Shammaa, Mohammed, and Maysam F. Abbod. (2014), Automatic generation of fuzzy classification rules from data." *Proc. of the 2014 International Conference on Neural Networks-Fuzzy Systems (NN-FS'14), Venice.*

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*, 29(11), pp 33-37.

Castellano, Giovanna, Anna Maria Fanelli, and Corrado Mencar. (2003), Design of Transparent Mamdani Fuzzy Inference Systems." *HIS.*

Chiu, Stephen. (1997), Extracting fuzzy rules from data for function approximation and pattern classification." *Fuzzy Information Engineering: A Guided Tour of Applications. John Wiley & Sons.*

Cordón, Oscar, et al. (2001), Ten years of genetic fuzzy systems: current framework and new trends." *IFSA World Congress and 20th NAFIPS International Conference, 2001. Joint 9th. 3. IEEE.*

Debnath, Sree Bash Chandra, Pintu Chandra Shill, and Kazuyuki Murase. (2013), Particle Swarm Optimization Based Adaptive Strategy for Tuning of Fuzzy Logic Controller, *International Journal of Artificial Intelligence & Applications*, 4(1), pp 37.

Farahbod F, Eftekhari M. (2013), A new Clustering-Based Approach for Modeling Fuzzy Rule-Based Classification System, *IJST, Transactions of Electrical Engineering*, 37(1), pp 67-77.

Farias MS, Nedjah N, Mourelle LDM. (2011), Radionuclide Identification Using Subtractive Clustering Method , *International Nuclear Atlantic Conference – (INAC 2011).*

Hall, Lawrence O., and Petter Lande. (1998), Generation of Fuzzy Rules from Decision Trees, *J. Adv. Computational Intelligence*, 2(4) ,pp 128-133.

Park, Keon-Jun, Jong-Pil Lee, and Dong-Yoon Lee. (2012), Optimal Design of Fuzzy Clustering-based Fuzzy Neural Networks for Pattern Classification, *International Journal of Grid & Distributed Computing*, 5 (3) ,pp 51-68.

Priyono a, Ridwan M, Alias A.(2005), Generation of fuzzy rules with subtractive clustering, *J Teknol.*, 43(D), pp 143-153.

Tawafan, Adnan, Marizan Bin Sulaiman, and Zulkifilie Bin Ibrahim. (2012), Adaptive neural subtractive clustering fuzzy inference system for the detection of high impedance fault on distribution power system." *IAES International Journal of Artificial Intelligence (IJ-AI)* ,1(2), pp 63-72.