International Journal of Engineering & Technology, 7 (4.35) (2018) 361-368

International Journal of Engineering & Technology

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

Research paper



A Genetic Algorithm Based Fuzzy Inference System for Pattern Classification and Rule Extraction

Shen Yuong Wong[1], Keem Siah Yap[2*], Xiaochao Li[3]

[1]*School of Electrical and Computer Engineering, Xiamen University, Malaysia.*
[2] *Department of Electronics and Communication Engineering, Universiti Tenaga Nasional, Malaysia.*
[3] *School of Electrical and Electronics, Xiamen University, China.*
** Corresponding author E-mail: yapkeem@uniten.edu.my*

Abstract

Setting fuzzy rules is one of the paramount techniques in the design of a fuzzy system. For a simple system, fuzzy if-then rules are usually derived from the human experts. However, in the event of having multiple variables coupled with a few features, the classification problem will be getting more sophisticated, as a result human expert may not be able to derive proper rules. This paper presents a genetic-algorithm-based fuzzy inference system for extracting highly comprehensible fuzzy rules to be implemented in human practices without detailed computation (hereafter denoted as GA-FIS). The impetus for developing a new and efficient GA-FIS model arises from the need of constructing fuzzy rules directly from raw data sets that combines good approximation and classification properties with compactness and transparency. Therefore, our proposed GA-FIS method will first define the membership functions with logical interpretation which is amendable by domain experts to human understanding, and then genetic algorithm serves as an optimization tool to construct the best combination of rules in fuzzy inference system that can achieve higher classification accuracy and gain better interpretability. The proposed approach is applied to various benchmark and real world problems and the results show its validity.

Keywords: *Fuzzy Inference System; Genetic Algorithm;, Pattern Classification,; Rule Extraction.*

1. Introduction

Fuzzy Inference System (FIS) has been an emerging practical solution for successfully solving numerous kind of process control and instrumentation problems [1-10]. FIS provide a unified framework for taking into account the gradual or flexible nature of variables, and representation of incomplete information [11-13]. This is an alternative to classical approach and is based on the observations that, human think using verbal explanatory terms such as “small” or “huge” rather than numerical values. The concept of fuzziness was proposed by Zadeh [14]. He aimed to bring about the complicated systems using fuzzy sets methods and approximation. FIS constitutes one of most distinctive application of the fuzzy logic and Fuzzy set theory. Fuzzy system comprises the conditional if-then rules, that apply fuzzy approximation as linguistic clause in the antecendent and conclusion of the rules. The main advantage of such fuzzy system is the usefulness and flexibility to interpret and represent knowledge in the rule base [15]. FIS is well accepted in getting round the classification problems, for it is helpful to perform online decision support tasks and process control, offline process simulation and diagnosis.

Generally, fuzzy if-then rules are derived and tuned by human experts. However this approach relies too much on the availability of the human experts. When there are no experts or skilled operators at hand to supply the necessary knowledge, question naturally arises causing the knowledge acquisition process to become a bottleneck. The problem intensifies especially when dealing with high dimensional system where many input variables involved in the reasoning process. As a result, domain experts may not be able to derive perfectly accurate fuzzy rules. To make things worse, the amount of data collected from several industries is growing explosively every day. The amount and complexity of the data make it impossible for human to analyze all in reasonable time. There are two points worth noticing. First, the cost of employing an expert might be too high that will exert additional burden financially and second, the time taken for the experts to build the appropriate rule base might be too long. Therefore, obtaining optimal fuzzy rules and developing ideal fuzzy classifier design are not easy task. As such, ways to generate fuzzy if then rules automatically from the source of raw data has been widely researched over the years [16-21].

When it comes to high-dimensional problems, there will be an exponential increase of the number of fuzzy if-then rules with the number of inputs, i.e., the input attribute. Large system with large number of input attributes complicates the reasoning process due to its sophisticated structure and thereafter leads to a phenomenon called “the curse of dimensionality” problem. There are some studies engaging the Genetic Algorithm (GA) to explore rules and later used in rule extraction. Having recognized that GA is an excellent search technique used in computing to find true or approximate solutions to optimization and search problems. In fact GA has been involved in a series of pattern classification problems [22-25]. In [26], Ishibuchi *et al.* published an article on choosing small set of fuzzy antecedent-conclusion rules with high classifica-tion performance. Although GA has achieved tremendous perfor-mance in terms of rule extraction, it is at a disadvantage when the extracted rules come with huge antecedent components. The ante-cedent components here are referred to the input attributes that are carried along with the data set. A rule with huge number of anteced-ents is not a desirable feature. Rule extraction

Copyright © 2018 Authors. 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.

technique is only considered useful when only the significant attributes are retained as a result of extraction [27].
A solution of overcoming the problem associated with a large number of rule antecedents was proposed in [23]. A "don't care" (hereafter denoted as "DC") antecedent approach was incorporated into the GA for rule extraction. However another technique introduces two-stage rule extraction [28]. First, the excessive weights of the network structure are pruned, and secondly translate the final knowledge into a set of linguistic fuzzy rules. In [29], a data dimensionality reduction method that axed the superfluous features was presented. Implementation of two-stage operation is not easy as it entailed time-consuming iterative procedures and complicated rule generation mechanisms.
One important issue worth taking into account is that most of the recognizers pursue accuracy as the ultimate goal and ignore the interpretability of the extracted knowledge. Regulating the trade-off between interpretability and accuracy, reliability of rules, optimization of the parameters is almost overlooked [27]. We must be very cautious not to build a classifier system that focus only on the maximization of the accuracy but axing the importance of explicability aspects. The interpretability of a fuzzy classification system is an attractive asset to the end users that adopt the system. Therefore, it is essential to devise a fuzzy classifier to bestow upon a rule extraction capability to ensure the final system can be well represented with explanation, reasoning and prediction. The rule base legibility is an important condition to take full advantage of fuzzy inference systems.
In this paper, we proposed a GA-FIS (genetic algorithm based fuzzy inference systems) model that is capable of performing pattern classification, rule extraction and optimization all in one stage. The objectives of this paper are three fold. First objective is to automatically develop best combination of fuzzy rules from raw data in the real world without human experts. Secondly, this paper is on enforcing the interpretability of the classification system standing of the rules extracted from the numerical measured data. Comprehension of the derived rules to human beings has become the most remarkable factor in deciding the usefulness of the fuzzy rules. Decision makers in all industries today are usually not the AI experts, statisticians, or the mathematicians. Hence it is very crucial that the results of the final system can be well represented such that even the laymen can easily understand the meaning underneath the data. For the experiment, the results of GA-FIS is discussed, analyzed and compared with those from other approaches.
The organization of the paper is as follows. Section 2 describes the model development in detail. The experimental procedures, results, performance comparisons and discussion are presented in Section 3. Section 4 concludes the paper.

2. Model Development

In this section, the possibility of integrating GA (genetic algorithm) into FIS (fuzzy inference system) to become a Computational Intelligence (CI) model is explored in pursuing higher flexibility and better comprehensibility and at the same time without losing accuracy. The GA being a computational optimization tool and FIS utilizing human knowledge and expertise, both techniques combined to contribute to an intelligent framework for delivering effective strategies to pattern classification and decision making tasks. Therefore, the proposed model is denoted as GA-FIS from this point onwards.

2.1. The Proposed Framework

GA is utilized to automatically generate a set of fuzzy rules combination directly from raw data. The essential part of the fuzzy inference system is the rule base embedded to represent knowledge linguistically by means of fuzzy implication and compositional rule of inference. The performance of the fuzzy system rests largely on the quality of the rule base. An unreliable and poorly formed rule base is unlikely to perform the task satisfactorily. Hence, GA is presented to tune and to extract the best combination of fuzzy rule set to make the fuzzy inference system behaves as closely as possible to the process to be investigated and this facilitates the interpretation of rules.

2.2. Membership Functions and Quantization of FIS

In this paper, the quantization level is set to 5, i.e. $Q = 5$. Here, fixed membership functions are used to define fixed linguistic terms. Five Gaussian membership functions that correspond to five quantization level $Q = 5$ are devised. On top of that, a "DC" antecedent method is included to ensure high generalization ability. The input features can be described with "very low", "low", "medium", "high", and "very high" according to the five quantization levels. Q quantization are marked in the interval [0, 1] consistently with one at each end point [30]. It can be expressed mathematically as follows,

$$V_q = \frac{q-1}{Q-1} \tag{1}$$

Where $q = 1, 2, 3, 4, 5$.
The input feature, x, with membership function is defined as

$$\mu_q(x) = \exp\left[-\frac{(x-V_q)^2}{2\beta^2}\right] \tag{2}$$

for $q = 1, 2, 3, 4, 5$ and $\beta = 0.1062$ in order to construct five equally well distributed fuzzy membership functions, while an additional DC antecedent is added, as shown in Fig. 1.

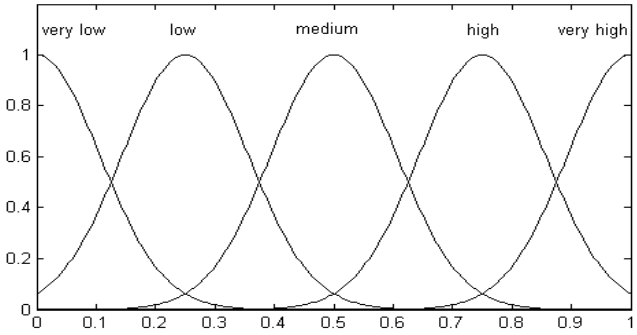


Fig. 1: Fuzzy partitions of the five linguistic values representing the membership functions

Of course, if priori expert knowledge of the specific application is known, then each axis does not have to be partitioned homogenously. These membership functions are considered as generic model since it is not application-specific, they can still be assigned with meaningful linguistic labels by the domain expert. Such general model is very useful in real life as it works for all applications. Human expert for example medical specialists and engineers can effectively manipulate and decide the parameters of the membership function according to their expertise and professional experience in the respective fields [31].
The standard minimum and maximum operators are used for "AND" and "OR" linguistic terms [32], and the final output of the FIS classifier is obtained by the weighted average defuzzification method [32],

$$\bar{y} = \frac{\sum_{i=1}^{K} w_i f_i}{\sum_{i=1}^{K} w_i}, \tag{3}$$

where K, w_i *and* f_i are the number of rules, firing strength and output constant of rules, respectively [30].

2.3. Rule Extraction using Genetic Algorithm

According to the Pittsburgh method, GA encode fuzzy if-then rules as a fixed length "bitstrings" and regarded as an individual. Each bit represents some aspect of the proposed solution to the problem. For GA to work, the encoded string represents a solution and need to put to test. The performance of each rule set (which is the individual in the GA context) is measured by the fitness score. The fitness score indicates how good that solution is, or a heuristic estimate that tells how close a candidate is to being the best solution. The ultimate goal of GA is either to minimize the effort required or to maximize the desired benefit, which in other words, represent the act of searching the best result under the given circumstances. Hence, this research work engages the genetic search for finding the appropriate and "best" fuzzy rules that can produce the highest fitness score which is equivalent to finding the rule base system with "best" performance. Inheritance of good rule in the current population as elite individuals with no modification will be passed on to the next population.
To understand how the population is presented, now let the fuzzy if-then rule for the n-dimensional pattern space $[0, 1]^n$ be denoted as

$$\text{Rule } R_j: \text{IF } x_1 \text{ is } A_{j1} \text{ and } \ldots \text{ and } x_n \text{ is } A_{jn} \quad \text{THEN class is } C_j,\ j = 1, 2, \ldots r \tag{4}$$

where R_j is the label of the j-th fuzzy if-then rules, j indexes the number of rules, $x = (x_1, x_2, \ldots, x_n)$ is an n-dimensional pattern vector, A_{jn} is an antecedent fuzzy set with linguistic label on the i-th axis, C_j is the consequent class.
Now, let the fuzzy if-then rule R_j be presented as its antecedent fuzzy sets: $R_j = A_{j1}. A_{j2}, \ldots, A_{jn}$, i.e. R_j is coded as a string of length n. Let S be a set of N fuzzy if-then rules, $S = \{R_1, \ldots, R_N\}$. According to the Pittsburg approach [18], S is denoted by a concatenated string of length $n \times N$, and the rule set is denoted by

$$S = A_{11} \ldots A_{1n} A_{21} \ldots A_{2n} \ldots A_{N1} \ldots A_{N n} \tag{5}$$

The fitness function is a function of S, and is denoted as *fitness (S)*.

2.4. Genetic Operators

GA is randomised heuristic search strategy, where it starts with the simulation of natural selection, in which the population is composed of candidate solutions. GA focuses on evolving a population from which strong and diverse candidates can emerge via mutation and crossover (mating). There are three GA operators, such as method of representation, selection and reproduction.
The typical candidate representation is a binary string. This string can be thought of as the genetic code of a candidate, thus form the term GA. The digit 1's and 0's at each position represents the value of some aspect of the solution. Besides that, solutions can be encoded as arrays of integers or decimal numbers. Last but not least, individuals in a GA can be represented as strings of letters, where each letter again stands for a specific aspect of the solution.
Selection explains there are many different techniques which a GA can use to select the individuals to be copied over into the next generation, such as:

(i) Roulette-wheel selection. The fitter is the solution with the most chances to be chosen
(ii) Elitist selection. The fittest members of each generation are guaranteed to be selected.
(iii) Rank selection. Each individual in the population is assigned a numerical rank based on fitness, and selection is based on this ranking.

On the other hand, method of reproduction comprises mutation and crossover. Mutation involves given a candidate, return a slightly different candidate. While crossover involves the mating and exchange of two candidates, produces offspring that has elements of each.

2.5. Genetic Coding of a Fuzzy System

As an illustration of the coding mechanism, the corresponding genetic code of 2-rule 2-input system is shown in Fig. 2 as follows, As in Fig. 2., rule-1 possesses 13 bits of the chromosome. The first 5 bits on the left represent 5 membership functions of the first input, then the subsequent 5 bits represent the equal 5 membership functions of the second input, followed by a bit describing the DC status. A binary 1 or 0 indicates whether the input is included in the system or not. The 11^{th} bit is a 1, which means the DC is true, therefore input-1 of rule-1 is omitted; whereas the 12^{th} bit displays a 0, which means the DC is false, thus input-2 of rule-1 is remained. The last bit for rule-1which is the 13^{th} bit denotes the class bit. A binary 1 represent class +1, and binary 0 represents class -1. The same pattern continues until the final rule-n of a chromosome.

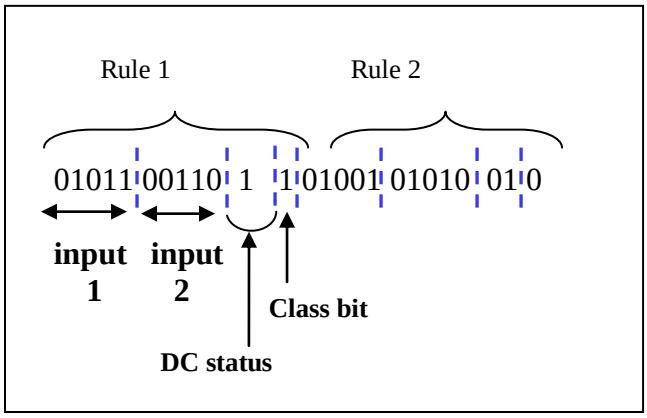


Fig. 2: Example of coding mechanism

2.6. Fitness Function

In this research work, the GA fitness function is defined as:

$$fitness(f)_{\min imize} = 100 - 0.95 f_1 - 0.05 f_2, \tag{6}$$

where f_1 is the accuracy of prediction, f_2 is the DC compression rate, such as:

$$f_1 = \frac{u}{p} \times 100\%, \tag{7}$$

$$f_2 = \frac{d}{s} \times 100\%, \tag{8}$$

where u refers to the number of correctly classified pattern, and p represents the total number of patterns used. On the other hand, d shows the number of DC bits generated, and s delineates the total bits of chromosomes.

3. Experiments and Results

In this section, the applicability of the proposed model is investigated against few cases. They are the Pima Indian Diabetes (PID), the circulating water (CW) system and occurrence of flashover in compartment fire. The experimental results are analyzed, discussed, and compared with other reported results.

Before starting the experiments, the data sets were divided into three sub-sets for training, validation, and test. To minimize fine-tuning the model parameters, unless otherwise stated, the GA parameters used in all experiments were as follows,

- The population size = 100
- Crossover probability = 0.9
- Mutation probability = 0.1
- Number of generations = 100 generations
- Stopping count = maximum of 5000 generations

What is important is to fix the number of generation and population size that is sufficient enough (not too large) to produce effective rules, which can result in achieving relatively close to best accuracy. It is not worth to sacrifice plenty of computational time for the 0.01% improvement. Note that the processing time is increased exponentially with the addition of each generation. This is not linear in nature and can be a critical issue. In this research, the number of generations and population size are set to be 100 respectively.

2.7. Pima Indian Diabetes (PID)

PID data set of UCI repository consists of 768 samples with two class labels. Class 1 is used to represent patient with diabetes and class -1 is used to represent non diabetic patient. Each of the samples consists of eight input attributes as given in Table I. They are divided into three subsets, i.e., 50% for training, 25% for prediction and 25% for test. All the data used is complete without any loss of attributes.

Table 2 shows the effect of the number of nodes on the network performance. From the table, one can notice that the number of nodes equal to 10 gives the best preliminary test set results, which is 80.73 %. Larger number of nodes requires more computation effort and thus will cause reduction in the optimizing performance.

Table 1 Description of the parameters used in PID data set

No.	Attribute
1	Number of times of pregnant
2	Plasma glucose concentration
3	Diastolic blood pressure (mm Hg)
4	Triceps skin fold thickness (mm)
5	Two hours serum insulin (mm U/ml)
6	Body Mass Index (weight in kg/ (height in m)^2)
7	Diabetes pedigree function
8	Age (years)

Table 2 The different test accuracy with respect to the number of nodes

Number of Nodes	Test Accuracy (%)
5	75.52
6	78.65
7	79.17
8	79.69
9	79.69
10	**80.73**
11	78.65
12	79.17

Table 3 illustrates in detail for the second rule extracted by GA-FIS in the binary matrix form. It is worth to know that each attribute in the rule incorporates various combination of membership functions. DC status for each of the attribute is represented by the last bit of Table 3. To explain further, '1' denotes 'TRUE' which means the respective input attribute can be regarded as "Don't care". On the contrary, '0' denies the DC status and therefore input attribute is kept and its membership functions will be retained as part of the rules.

Table 3 The PID Rule-2 From The GA-FIS in Binary Matrix Form.

Membership Functions	IF X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	THEN Class is
very low	1	1	0	1	1	1	1	0	
low	1	0	0	0	0	0	1	0	
medium	0	1	0	0	1	1	0	1	-1
high	0	0	0	1	0	1	1	1	
very high	1	0	0	0	1	0	0	1	
DC	1	0	1	1	1	1	0	1	

Table 4 Total 10 Rules Extracted from GA-FIS for the PID Data Sets

	IF X1	X2	X3	X4	X5	X6	X7	X8	THEN Class is
Rule 1	DC	1,2,4	2,3,4,5	DC	1,3,4,5	DC	1,2,3	DC	-1
Rule 2	DC	1,3	DC	DC	DC	DC	1,2,4	DC	-1
Rule 3	DC	DC	DC	1,2	DC	1,2,5	DC	1,2,3,4	-1
Rule 4	3,4,5	DC	DC	DC	DC	1,4,5	DC	DC	+1
Rule 5	DC	DC	DC	DC	DC	1,2,4	2,5	DC	+1
Rule 6	1,5	2,4	4	DC	DC	DC	DC	1,3,4	-1
Rule 7	DC	1,2,3	DC	DC	DC	DC	1,2,4,5	DC	-1
Rule 8	DC	DC	1,3,5	DC	1	DC	1,2,3	2	+1
Rule 9	DC	DC	DC	DC	1,2,4,5	5	DC	DC	+1
Rule 10	DC	DC	DC	DC	DC	1,3,5	DC	1,3	+1

Table 5 Pima Indian Diabetes case: 2nd rule interpretation and its fuzzy equations extracted from GA-FIS

If
Plasma Glucose Concentration (PGC) is very low or medium
and Diabetes Pedigree Function (DPF) is very low or low or high

Then Diabetic is not likely to happen.

Firing strength of Rule–2 is
$w_2 = [\mu_1(PGC) \vee \mu_3(PGC)] \wedge [\mu_1(DPF) \vee \mu_2(DPF) \vee \mu_4(DPF)]$

Constant for the output variable of Rule–2 is
f_2 = class label of Rule–2 = -1 (Non-diabetic)

Table 6 The PID Testing Accuracy Rates from Different Approaches Reported [36]

Model	Testing Accuracy (%)
C4.5	71.02
C4.5 Rules	71.55
ITI	73.16
LMDT	73.51
CN2	72.19
LVQ	71.28
OCI	50.00

Table 7 The performance of GA-FIS in Pima Indian Diabetic data set as compared to the other FIS classifiers [30]

FIS	Number of Rules	Accuracy Rate (%)	Percent Compression (%)
GA-FIS	10	83.33	65.00
EGART-FIS	3	74.22	–
MFMM-FIS	5	72.92	–
FAM-FIS	6	68.88	–
FAM-RecBFN-FIS	5	69.14	–

In addition, Table 4 delineates the total 10 rules extracted from GA-FIS with quantization level Q=5 for the PID data sets. Each row of the rule in Table IV can be translated into a fuzzy rule. For instance, interpretation of rule- 2 in terms of the IF-THEN format and the respective fuzzy equation from GA-FIS for the PID data set is described in Table 5. Then, Table 6 shows the accuracy of classification of eight different methods as reported in[36]. Again, GA-FIS tops the performance as it obtainsthe best accuracy as compared to the other approaches.

On all benchmark result in Table VII, GA-FIS approach is several times superior to MFMM-FIS, FAM-FIS, FAM-RecBFN-FIS methods. The GA-FIS model is better by 8.2% as compared to EGART-FIS in terms of classification accuracy.

2.8. Occurences of Flashover in Compartment Fire

In this section, the efficacy of GA-FIS in solving a fire safety engineering problem is evaluated. In this experiment, data samples of fire are gathered and simulated using computer software. It is owing to the fact that the real fire data samples are hardly captured, and are quite impossible to set up from full scale environment.

In the model of occurrence of flashover, the upper hot gas layer temperature is mathematically represented by the size of the room geometry, i.e., the opening dimensions, the characteristics of the gas, the wall conduction properties, and finally heat release rate. We adopt the specification of flashover defined by Babrauskas *et al*., [38]. The engine used in the FASTLite computer package is FAST [39, 40].

A few assumptions are made prior to the experiment, such as:

(i) Fire occurs in a compartment of rectangular in shape with an open door
(ii) The walls and ceilings are made of 16-mm gypsum
(iii) The floor is made of 12.7-mm plywood.

In this experiment set up, 375 data samples are generated, of which 190 are flashover, and 185 are non-flashover data samples. The following inputs are randomly generated, as shown in Fig.3.

(i) length of the room (randomly varied between 2 to 10 meters range);
(ii) width of the room (randomly varied between 2 to 10 meters range);
(iii) height of the room (randomly varied between 2 to 10 meters range),
(iv) maximum rate of heat release (randomly varied between 10 - 6000 kW range).
There are two output class labels, i.e., non-flashover (Class 1), or flashover (Class 2). Fast growth t-square (t^2) fire [41] is described as follows,

$$\frac{dQ}{dt} = \alpha(t - t_0)^2, \quad (9)$$

in which $\frac{dQ}{dt}$ is the heat released rate (kWs^{-1}), α defines the growth constant (0.0469 kWs^{-1}), t_0 is the initial time and t is the time, in seconds.
Table 8 depicts the 4 rules generated from the GA-FIS model. Another consideration is to use as small number of fuzzy rules as possible. Therefore, we will set the number of fuzzy rules, i.e., 4 rules, before the simulation starts. Moreover, it is quite possible that large number of rules will lead to a degradation rather than improvement of performance. Having lesser rules eases human understanding of the problem and poses no difficulty for future implementation when faced with the similar problem in reality. As such, heavy computation proceedings on computers can be eliminated.

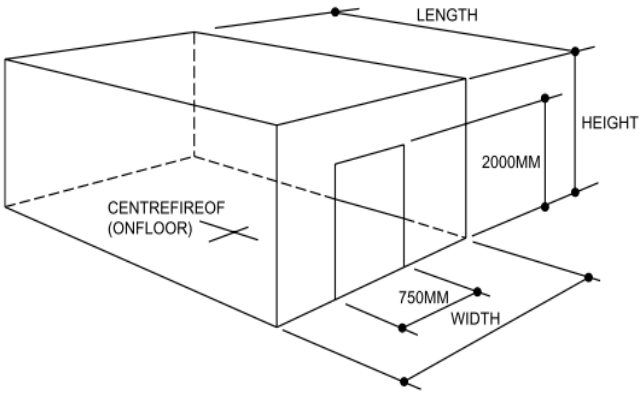


Fig.3: Compartment dimensions for modelling the flashover occurence.

Interpretation of rule-3 and its equivalent fuzzy notation is ensued in Table 9. Note that only one ultimate important attribute is required to predict the occurrence tangibly. Table 10 epitomizes the performance of GA-FIS and other FIS classifiers. The result indicates that GA-FIS surpasses other classifiers with 96.8% accuracy.

Table 8 Flashover case: Rules extracted from GA-FIS

	If Attribute is i	ii	iii	iv	Then Class is
Rule 1	DC	DC	1,3	3,4	+1
Rule 2	DC	DC	1,2,3,5	4,5	+1
Rule 3	DC	DC	DC	1	-1
Rule 4	DC	DC	1,2,4	4,5	+1

Table 9 Interpretation of Rule-3 and the respective fuzzy equations extracted from GA-FIS for the Flashover data set.

If Fire Size (F) is very small Then Flashover is not likely to happen.
Firing strength of Rule–3 is $w_3 = \mu_1(F)$
Constant for the output variable of Rule-3 is $\mathbf{f}_3$ = class label of Rule-3 = -1 (No Flashover)

Table 10 The Accuracy Rates and The Number of Rules for The Flashover Data Set Based On the Various FIS Classifiers [30]

FIS Classifiers	Number of Rules	Accuracy Rate (%) (test set only)	Percent Compression(%) (PC)
GA-FIS	4	96.80	56.25
O-EGART-PR-FIS	4	93.60	—
FAM-FIS	4	93.60	—
GRNNFA	—	92.6	—

4. Conclusion

In this paper, a novel GA-FIS classifier is constructed, which utilizes genetic algorithms- based method to fuzzy inference system for fuzzy classification. GA-FIS-CRS is conceived as a better method for the outcomes reveal the efficacy of GA-FIS in extracting a set of compact and yet highly comprehensible rules while sustaining a high classification performance. It is worth noticing that if priori expert knowledge of the specific application is known, then each axis does not have to be partitioned homogenously. It should be pointed out that the generic membership function model presented here is concluded as a better technique to enable domain experts to have full privilege to decide and design the shape of the membership functions. The parameters such as the center and standard deviation of each membership function can be adjusted in their submission. The final simulation results validated the feasibility of the proposed approach in solving problems of various domains. There is no qualm in admitting that GA has received ultimate affirmation for its optimization ability. Not to mention, rule-based fuzzy system has been and continues to be a very active and fruitful research field since Mamdani's pioneering work. Therefore, this paper presented a blend of GA and FIS turned CI model which leverages from both techniques' strength has further elevated the classification system to the next level to clearly reveal the decision making process. From practical and implementation standpoints, it is essential to build a classifier which is particularly good at extracting useful fuzzy if-then rules that are understandable in human perspective while attaining the highest possible classification accuracy.

Acknowledgement

This work was supported by Universiti Tenaga Nasional (UNITEN) under grant number J510050684 and Xiamen University Malaysia (IECE/0001).

References

[1] R. O. Duda, P. E. Hart, and D. G. Stork, "Pattern Classification", 2nd edition, Wiley, New York, 2001.
[2] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol 521, pp. 436-441, 2015.
[3] D. Cui, G. Zhang, K. Hu, W. Han, and G. Huang, "Face recognition using total loss function on face database with ID photos," Opt. Laser Technol., 2017.
[4] S. Bijarnia, "Pyramid Binary Pattern for Age Invariant Face Verification," 13th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS) , pp. 13–16, 2017.
[5] W. Ouyang, X. Zeng, S. Member, and X. Wang, "DeepID-Net : Object Detection with Deformable Part Based Convolutional Neural Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 7, pp. 1320–1334, 2017.
[6] K. He and J. Sun, "Deep Residual Learning for Image Recognition," IEEE Conference on Computer Vision and Pattern Recognition (CVPR) , pp. 1–9, 2016.
[7] Z. Bai, L. Lekamalage, C. Kasun, and G. Huang, "Generic Object Recognition with Local Receptive Fields Based Extreme Learning Machine," Procedia - Procedia Comput. Sci., vol. 53, pp. 391–399, 2015.
[8] S. Chen, and G. Jiang, "The Prediction Model of Multiple Myeloma Based on the BP Artificial Neural Network", in Proc. of the International Conference on Technology and Applications in Biomedicine, pp. 380-382, 2008.
[9] J.M. Benitez, J.L. Castro and I.Requena, "Are Artificial Neural Networks Black Boxes?" IEEE Trans. Neural Netw., vol 8, no.5, pp. 1156-1164, Sep 1997.
[10] E. Kolman and M. Margaliot, "Are Artificial Neural Networks White Boxes?" IEEE Trans. Neural Netw., vol 16, no.4, pp. 844-852, Jul. 2005.
[11] K.S. Yap, C.P. Lim and J.M. Salleh, "An enhanced generalized adaptive resonance theory network and its application to medical pattern classification", Journal of Intelligent & Fuzzy Systems 21, pp. 65-78, 2010.
[12] J D J Rubio, "USNFIS: Uniform Stable Neuro Fuzzy Inference System," Neurocomputing, vol. 262, pp. 57-66, 2017.
[13] Jun H Chung , JM Pak , CK Ahn , SH You , MT Lim , MK Song, "Particle filtering approach to membership function adjustment in fuzzy logic systems,", Neurocomputing, no. 237, pp.166–174, 2017.
[14] Zadeh LA, "Outline of a new approach to the analysis of complex systems and decision processes", In: yager RR, Ovchinnikov S., Tong RM., Nguyen HT., editors. IEEE Trans. Systems, Man and Cybernetics. SMC-3, Fuzzy sets and applications, pp. 28-44, 1973.
[15] CM Lin, TL Le, TT Huynh, "Self-evolving function-link interval type-2 fuzzy neural network for nonlinear system identification and control," Neurocomputing, vol. 275, pp. 2239-2250, 2018.
[16] S. Y. Wong, K. S. Yap, H. J. Yap, and S. C. Tan, "A Truly Online Learning Algorithm using Hybrid Fuzzy ARTMAP and Online Extreme Learning Machine for Pattern Classification," Neural Process. Lett., vol. 42, no. 3, pp. 585–602, 2015.
[17] S. Y. Wong, K. S. Yap, and H. J. Yap, "A Constrained Optimization based Extreme Learning Machine for noisy data regression," Neurocomputing, vol. 171, pp. 1431–1443, 2016.
[18] S. Y. Wong, K. S. Yap, H. J. Yap, S. C. Tan, and S. W. Chang, "On equivalence of FIS and ELM for interpretable rule-based knowledge representation," IEEE Trans. Neural Networks Learn. Syst., vol. 26, no. 7, pp. 1417–1430, 2015.
[19] S. Y. Wong, K. S. Yap, and H. J. Yap, "Constrained-Optimization based Bayesian Posterior Probability Extreme Learning Machine for pattern classification," International Conference on Neural Information Processing, Springer, 2014.
[20] ST Wang, H Ishibuchi, Z Bian, "Joint Learning of Spectral Clustering Structure and Fuzzy Similarity Matrix of Data, IEEE Transactions on Fuzzy Systems, 2018
[21] Y Zhang, H Ishibuchi, S Wang, "Deep Takagi–Sugeno–Kang Fuzzy Classifier With Shared Linguistic Fuzzy Rules," IEEE Transactions on Fuzzy Systems, 2018.
[22] MC Pablo, C Morillas, Hans E. Plesser , S. Romeroa , F. Pelayo, "Genetic algorithm for optimization of models of the early stages in the visual system," Neurocomputing, no. 250, pp. 101–108, 2017.
[23] Younas, F. Kamrani, M. Bashir, J. Schubert, "Efficient genetic algorithms for optimal assignment of tasks to teams of agents," Neurocomputing, pp.1–20, 2018.
[24] B. Haznedar, A. Kalinli, "Training ANFIS structure using simulated annealing algorithm for dynamic systems identification,", Neurocomputing, vol. 302, pp. 66-74, 2018.
[25] Huai-xiang Zhang, Bo Zhang, Feng Wang, " Automatic fuzzy rules generation using fuzzy genetic algorithm", IEEE Sixth International Conference on Fuzzy Systems and Konwlege Discovery, pp107-112, 2009.
[26] Ishibuchi, H., Murata, T. and Turksen, I. B. "Single-objective and two-objective genetic Algorithms for selecting linguistic rules for pattern classification problems", Fuzzy Sets and Systems, 89, 1997, pp. 135-150.
[27] Quteishat A., C. P. Lim, K. S. Tan "A modified fuzzy min-max neural network with a genetic algorithm-based rule extractor for pattern classification", IEEE Trans. Syst., Man, Cybern. A, Syst., Humans, vol. 40, no. 3, pp. 641-650, 2010,.
[28] G. Carpenter and A. Tan, "Rule extraction: From neural architecture to symbolic representation," Connection Sci., vo.7, no.1, pp..3-27, 1995.
[29] Xiuju, F. and Lipo, W. "Rule extraction using a novel gradient-based method and data dimensionality reduction", In Proceedings of the 2002 International Joint Conference, pp. 1275-1280, 2002.
[30] C Su, C Tseng, JSR Jang, T Visceglia, " A hierarchical linguistic information-based model of English prosody: L2 data analysis and implications for computer-assisted language learning, Computer Speech & Language, 2018.
[31] K S Yap, S. Y. Wong, S K Tiong, "Compressing and improving fuzzy rules using genetic algorithm and its application to fault detection" IEEE 18th Conference on Emerging Technologies and Factory Automation (ETFA), 2013.
[32] J. S. R. Jang, ANFIS: Adaptive-network-based fuzzy inference system, IEEE Transactions on Systems, Man and Cybernetics , vol. 23, no. 3, pp. 665-685, 1993.

[33] Y. Ding, X. Fu, "Kernel-based fuzzy c-means clustering algorithm based on genetic algorithm," Neurocomputing, no. 188, pp. 233-239, 2016.
[34] Wei H., Tong Z., and Jian S. "Use of a genetic algorithm to optimize multistage erbium-doped fiber-amplifier systems with complex structures," Opt. Express, vol. 12, no. 4, pp. 531–544, 2004.
[35] Darwin, Charles. On the Origin of Species by means of Natural Selection, or the Preservation of Flavoured Races in the Struggle for Life (1st ed.) London : John Murray, 1859.
[36] Eklund P. and Hoang A. "A comparative study of public supervised classifier performance on the UCI database", Australian Journal of Intelligent Information Processing Systems 9 1-39, 2006.
[37] Tenaga Nasional Berhad Malaysia, System Description and Operating Procedures Prai Power Station Stage 3, 14, 1999.
[38] Babrauskas, V., Peacock, R. D. and Reneke, P. A. Defining Flashover For Fire Hazard Calculations: Part II. Fire Safety Journal, 38(7), p. 613–22, 2003.
[39] R. W. Portier, Peacock, R. D. and Reneke, P. A., "FASTLite: Engineering Tools For Estimating Fire Growth And Smoke Transport," Special publication 899, National Institute of Standards and Technology, 1996.
[40] Jones, W. W. and Peacock, R. D. Technical reference guide for FAST version 18. National Institute of Standards and Technology, 1989.
[41] Heskestad, G. Engineering relations for fire plumes. Society of Fire Protection Engineers, Technology report, p. 82-88, 1982.