Optimizing of Interval Type-2 Fuzzy Logic Systems Using Hybrid Heuristic Algorithm Evaluated by Classification

Adisak S. and Phayung M.
Department of Information Technology, Faculty of Information Technology King Mongkut's University of Technology North Bangkok, Bangkok, Thailand Email: adisak.sangsongfa@gmail.com, pym@kmutnb.ac.th

Abstract
In this research, an optimization of the rule base and the parameter of interval type-2 fuzzy set generation by a hybrid heuristic algorithm using particle swarm and genetic algorithms is proposed for classification application. For the Iris data set, 90 records were selected randomly for training, and the rest, 60 records, were used for testing. For the Wisconsin Breast Cancer data set, the author deleted the missing attribute value of 16 records and randomly selected 500 records for training, and the rest, 183 records, were used for testing. The proposed method was able to minimize rule-base, minimize linguistic variable and produce a accurate classification at 95% with the first dataset and 98:71% with the second dataset.

Keywords *: Interval Type-2 Fuzzy Logic Systems; GA; PSO;*

1 Introduction

In 1965, Lotfi A. Zadeh, professor for computer science at the University of California in Berkley, developed a fuzzy logic system which has been widely used in many areas such as decision making, classification, control, prediction, optimization and so on. However, the fuzzy logic system comes from the original system that is called the type-1 fuzzy set. Sometimes it cannot solve certain problems, especially problems that are very large, complex and/or uncertain.
Therefore, in 1975 Zadeh developed and formulated a type-2 fuzzy set to meet the needs of data sets which are complex and uncertain. Thus, the type-2 fuzzy set has been used widely and continuously in many cases [1].

Recently, there has been growing interest in the interval type-2 fuzzy set which is a special case of the type-2 fuzzy set. Because, Mendel and John [2] reformulated all set operations in both the vertical-slice and wavy-slice manner. They concluded that general particle type-2 fuzzy set operations are too complex to understand and implement, but operations using the interval type-2 fuzzy set involve only simple interval arithmetic which means computation costs are reduced. The interval type-2 fuzzy set consists of four parts: fuzzification, fuzzy rule base, inference engine and defuzzifications. Moreover, the fuzzy rule base and interval type-2 fuzzy sets are complicated when determining the exact membership function and complete fuzzy rule base. So, the optimization of interval type-2 fuzzy set and fuzzy rule base must be used to estimate the value by an expert system. Many researchers have proposed and introduced optimization of interval type-2 fuzzy set and fuzzy rule base such as Zhao [3] proposed adaptive interval type-2 fuzzy set using gradient descent algorithms to optimize inference engine fuzzy rule base, Hidalgo [4] proposed optimization interval type-2 fuzzy set applied to modular neural network using a genetic algorithm. Moreover, many researchers apply the interval type-2 fuzzy logic system for uncertain datasets. Also, the creation of an optimized interval type-2 fuzzy logic system will gain the maximum accurate outputs. There are also many optimization techniques which have been proposed for building interval type-2 fuzzy systems. Some traditional optimization techniques are based on mathematics and some are based on heuristic algorithms. Some optimization techniques are often difficult and time consuming such as heuristic optimization. Sometimes, the improvement of the heuristic algorithms provides good performance such as the hybrid heuristic algorithms [5]. Moreover, hybrid heuristic is a much younger algorithm candidate compared to the genetic algorithm and particle swarm optimization in the domain of meta-heuristic-based optimization.

In this paper, a new algorithm called the hybrid heuristic algorithm which combines a genetic algorithm to particle swarm optimization is proposed. Also, a presentation of an optimization of interval type-2 fuzzy set and fuzzy rule base using the proposed hybrid heuristic algorithm. The algorithm will be used to optimize a model by minimizing the number of fuzzy rules, minimizing the number of linguistic variable and maximizing the accuracy of the output. Then the framework and the corresponding algorithms are tested and evaluated to prove the concept by applying it to the Iris dataset [6] and Wisconsin Breast Cancer Dataset as an example of classification [7].

2 Related Works

A. Particle Swarm Optimization (PSO)

The PSO initializes a swarm of particles at random, with each particle deciding its new velocity and position based on its past optimal position P_1 and the past optimal position of the swarm P_g. Let $x_i = (x_{i1}, x_{i2}, \dots, x_{in})$ represent the current position of particle i, $v_i = (v_{i1}, v_{i2}, \dots, v_{in})$ its current velocity and $P_i = (P_{i1}, P_{i2}, \dots, P_{in})$ its past optimal position, then the particle uses the following equation to adjust its velocity and position:

$$V_{i,(t+1)} = wV_{i,(t)} + c_1 r_1 \left(P_i - x_{i,(t)}\right) + c_2 r_2 \left(P_g - x_{i(t)}\right) \quad (1)$$

$$x_{i,(t+1)} = x_{i,(t)} + V_{i,(t+1)} \quad (2)$$

where c_1 and c_2 are constants of acceleration in the range of 0..2, r_1 and r_2 are random number in [0,1] and w is the weight of inertia, which is used to maintain the momentum of the particle. The first term on the right hand side in (1) is the particle's velocity in time t. The second term represents "self learning" by the particle based on its own history. The last term reflects "social learning" through information sharing among individual particles in the swarm. All three parts contribute to the particle's search ability in the space analyzed which simulates the swarm behavior mathematically [8].

B. Genetic Algorithm (GA)

A GA generally has four components: 1) a population of individuals where each individual in the population represents a possible solution, 2) a fitness function which is an evaluation function by which we can tell if an individual is a good solution or not, 3) a selection function which decides how to pick good individuals from the current population for creating the next generation, and 4) genetic operators such as crossover and mutation which explore new regions of search space while keeping some of the current information at the same time.

GAs are based on genetics, especially on Darwins theory (survival of the fittest). This states that the weaker members of a species tend to die away, leaving the stronger and fitter. The surviving members create offspring and ensure the continuing survival of the species. This concept together with the concept of natural selection is used in information technology to enhance the performance of computers [9].

C. Interval Type-2 Fuzzy Set

Interval type-2 fuzzy sets are particularly useful when it is difficult to determine the exact membership function, or in modeling the diverse options from different individuals. The membership function, which interval type-2 fuzzy inference system approximates expert knowledge and judgment in uncertain conditions, this can be constructed from surveys or using optimization algorithms. Its basic framework consists of four basic parts: fuzzification, fuzzy rule base, fuzzy inference engine and defuzzification shown in Figure 1.

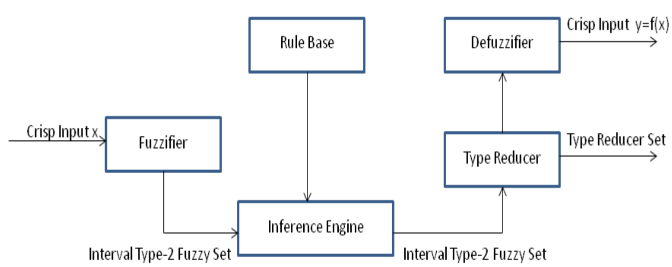

Figure 1: Interval Type-2 Fuzzy System

We can describe the interval type-2 fuzzy logic system as follows: the crisp sets inputs are first fuzzified into input interval type-2 fuzzy sets. In the fuzzifier, it creates the membership function which consists of types of membership function, linguistic variable and fuzzy rule base. It has many types of the membership function such as triangular membership function, trapezoidal membership function, Gaussian membership function, Smooth Membership Function, Z-membership function and so on. So, the fuzzifier sends the interval type-2 fuzzy set into the inference engine and the rule base to produce output type-2 fuzzy sets. The interval type-2 fuzzy logic system rules will remain the same as in the type-1 fuzzy logic system, but the antecedents and/or consequents will be represented by interval type-2 fuzzy sets. A finite number of fuzzy rules, can be represented as if-then forms, then integrates into the fuzzy rule base. A standard fuzzy rule base is shown below.

$$R1: If\ x_1\ is\ \tilde{A}_1^1\ and\ x_2\ is\ \tilde{A}_2^1, \ldots, x_n\ is\ \tilde{A}_n^1\ Then\ y\ is\ \tilde{B}^1.$$

$$R2: If\ x_1\ is\ \tilde{A}_1^2\ and\ x_2\ is\ \tilde{A}_2^2, \ldots, x_n\ is\ \tilde{A}_n^2\ Then\ y\ is\ \tilde{B}^2.$$

$$\vdots$$

$$RM: If\ x_1\ is\ \tilde{A}_1^M\ and\ x_2\ is\ \tilde{A}_2^M, \ldots, x_n\ is\ \tilde{A}_n^M\ Then\ y\ is\ \tilde{B}^M.$$

where $x_1, \ldots, x_n$ are state *c*=variables, y is control variable. The linguistic value $\tilde{A}_1^j, \ldots, \tilde{A}_n^j$ and $\tilde{B}^j$, (j=1,2,...,M) are respectively defined in the universe $U_1,\ldots U_n$ and V. In fuzzification, crisp input variable x_i is mapped into interval Type-2 fuzzy set $\tilde{A}_{x_i}, i = 1, 2, 3, \ldots n$. The inference engine combines all the fired rules and gives a non-linear mapping from the input interval type-2 fuzzy logic systems to the output interval type-2 fuzzy logic systems. The multiple antecedents in each rule are connected by using the Meet operation, the membership grades in the input sets are combined with those in the output sets by using the extended sup-star composition, and multiple rules are combined by using the Join operation. The type-2 fuzzy outputs of the inference engine are then processed by the type reducer, which combines the output sets and performs a centroid calculation that leads to type-1 fuzzy sets called the type-reduced sets. After the type-reduction process, the type-reduced sets are then defuzzified (by taking the average of the type-reduced) to obtain crisp outputs. [3].

In the interval type-2 fuzzy logic system design, we assumed Z-membership function for the first membership function, triangular membership function for the secondary membership function and smooth membership function for the last membership function, center of sets type reduction and defuzzification using the centroid of the type reduced set.

3 The Proposed Framework

In our framework, we present the new algorithms of hybrid heuristic algorithm which are developed to optimize the interval type-2 fuzzy logic system using Iris datasets and breast cancer datasets. The new algorithm to optimize the interval type-2 fuzzy sets and fuzzy rule base uses hybrid heuristic searches which are a sequential combination of GA and PSO. The proposed algorithm will be used to optimize the number of linguistic variables, parameters of membership functions and the rule base which consists of constraint of the minimum linguistic variable, minimum rule base and maximum accuracy. The framework is shown in Figure 2.

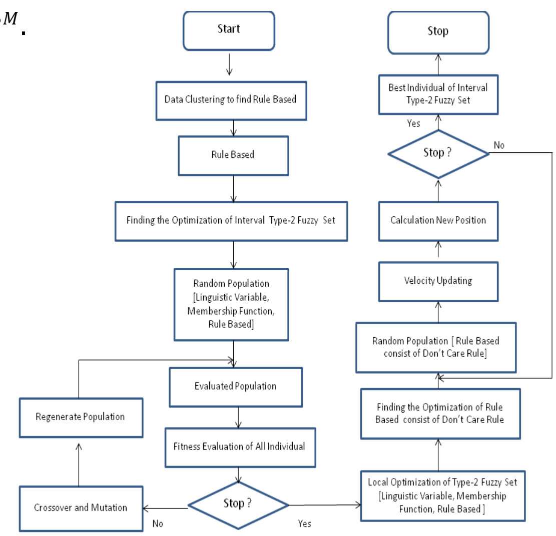

Figure 2: Framework of Optimization Interval Type-2 Fuzzy System Using Hybrid Heuristic Algorithms

From the framework, we can describe the steps of the proposed method for optimized interval type-2 fuzzy set and fuzzy rule base using hybrid heuristic searches. The framework is given in four steps described below.

Step 1: Determine the structure of interval type-2 fuzzy system framework.

Step 2: Determine the fuzzy rules base using clustering.

Step 3: Determine the universes of the input and output variables and their type of membership functions and linguistic parameter of membership functions.

Step 4: Determine and optimize the fuzzy inference engine using the hybrid heuristic algorithms which is a combination of GA and PSO.

1) *Determine the structure of interval fuzzy type-2 system framework*

In Figure 2. the framework shows the structure of the optimization interval type-2 fuzzy sets and rule based on hybrid heuristic algorithms. The hybrid heuristic algorithm used sequential hybridization. The GA is used for the first local optimal interval type-2 fuzzy sets which consist of interval type-2 membership function, interval type-2 linguistic parameter (LMF, UMF) and rule base. Moreover, the PSO is used for the last optimal which is a gaining the best result don't care rule.

2) Determine the fuzzy rules base using clustering

We used the K-means clustering algorithm [10] to group the dataset to determine the feasibility of a fuzzy rule base. A standard K-means clustering algorithm is shown as follows.

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^j - c_j \right\|^2 \qquad (3)$$

where K is clusters, $\left\| x_i^j - c_j \right\|^2$ is a chosen distance measure between a data set point x_i^j and the cluster centre c_j , is an indicator of the distance of the n data points from their respective cluster centers.

3) *Determine the universes of the input and output variables and their type of the membership functions*

In the universe of input and output variables and their primary membership functions, the z-membership function, triangular membership function and smooth membership function were used and are shown in Figure 3. In Figure 3., the presentation the four attributes of Iris membership function are displayed and graded as attibute1=2, attribute2=2, attribute3=5 and attribute4=5. The definition of the linguistic label and number of linguistic variables are in Table 1.

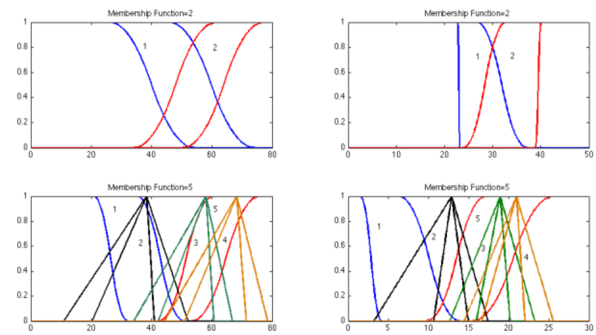

Figure 3: The Example of Interval type-2 Membership Functions

Table 1 : Predefined membership function for five linguistic variables

Linguistic Index	Linguistic Terms
0	Don't Care
1	Very Low
2	Low
3	Medium
4	High
5	Very High

4) *Determine and optimization the fuzzy inference engine using the hybrid heuristic algorithms*

Firstly, encoding the fuzzy rule based system into genotype or chromosome. Each chromosome represents a fuzzy system composed of the number of linguistic variables in each dimension, the membership function parameters of each linguistic variable, and the fuzzy rules which consists of don't care rules from the PSO. A chromosome (chrom) consists of 4 parts or genes:

$$chrom = [\overbrace{\text{IM, IL}}^{\text{by GA}}, \text{R}, \overbrace{\text{DcR}}^{\text{by PSO}}] \quad (4)$$

where $IL = [IL_1, IL_2, \dots, IL_n]$ is a set of the number of interval linguistic variables, $IM = [im_{11}, im_{12}, \dots, im_{n,IL_n}]$ is a set of the interval membership function parameters of the interval linguistic variables, $R = [R_1, R_2, \dots, R_{IL_1 \times IL_2 \times \dots \times IL_n}]$ is the fuzzy rules. R_l is integer number that is the index of linguistic variable of each dimension, and $DcR = [R_{a_{111}\, L_1 \times L_2 \times \dots L_n}, R_{a_{112}\, L_1 \times L_2 \times \dots L_n}, \dots, R_{a_{lmk}\, L_1 \times L_2 \times \dots L_n}]$ is the don't care rule. $R_{a_{lmk}\, L_1 \times L_2 \times \dots L_n}$ is integer number that is the index of don't care rule of each rule. The length of a chromosome can be varied depending on the fuzzy partition created by cross sections of the linguistic variables from each dimension. Then, the Fitness Function is

$$Fit = ACC_{(chrom\,_i)}$$

Where $chrom_i = [chrom_1, chrom_2, \dots, chrom_n]$is a set of the chromosome number. The accuracy *(Acc)* is

$$Acc = \frac{Number\ of\ Correct\ Classification}{Total\ Number\ of\ Training\ Data}$$

4 The Experimental Evaluation Setting Up

To evaluate the proposed Hybrid Heuristic Type-2(HHType-2) algorithm for building interval type-2 fuzzy systems, two datasets were used which are benchmark classification datasets from UCI data repository for machine learning, Fishers iris data and Wisconsin Breast Cancer data.

A. Datasets

Iris dataset has 4 variables with 3 classes; 90 records were selected randomly for training, and the rest, 60 records, were used for testing. Wisconsin Breast Cancer data set has 699 records, the missing attribute value of 16 records were deleted. Each record consists of 9 features plus the class attribute; 500 records were selected randomly for training, and the rest, 183 records, were used for testing.

Figure. 4 shows the scatter plot of the Iris dataset, Fig. 5 illustrates the scatter plot of the Iris dataset with clustering using *K*-Mean algorithms (*K*=7). Figure 6. shows the scatter plot of the Wisconsin Breast Cancer dataset, and Fig. 7 shows the scatter plot of the Wisconsin Breast Cancer dataset with clustering using *K*-Mean algorithms (*K*=4).

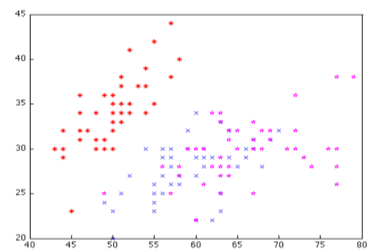

Figure 4 : The scatter plot of Iris Dataset (* represents Setosa, × represents Versicolor, and ★ represents Verginica)

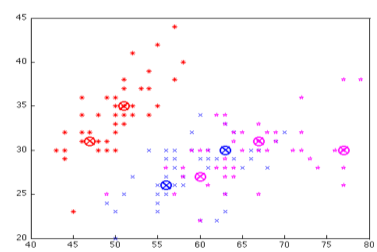

Figure 5: The scatter plot of Iris Dataset with Clustering (* represents Setosa, × represents Versicolor, and ★ represents Verginica)

B. Experimental Results

The experiments were performed on a MacBook Pro Intel Core 2 Duo CPU, speed 2.66 GHz, ram 4.00 GB RAM, running on Mac os. All algorithms are implemented using Matlab.

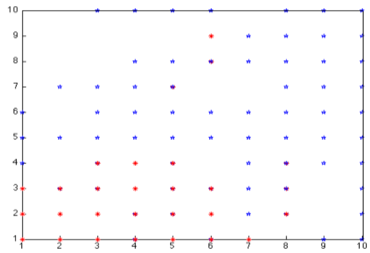

Figure 6 : The scatter plot of Wisconsin Breast Cancer Dataset (* represents Class 2, and ★ represents class 4)

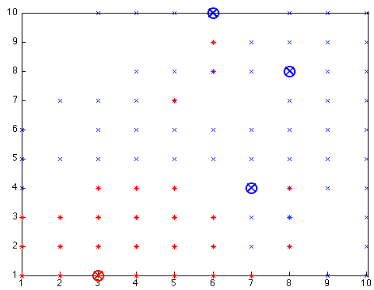

Figure 7 : The scatter plot of Wisconsin Breast Cancer Dataset with Clustering (* represents Class 2, and ★ represents class 4)

The first dataset (Iris dataset), ran 20 times with the averages execute time of 662.2635s. The simulation population was 100 individuals. Then, the largest individuals from the PSO were used to optimize the "don't care" rule. In the PSO, each of the individuals were simulated with 50 swarms and 5 particles. The PSO completed 20 runs with the excite time of 429.7597s.

In the second dataset (Wisconsin Breast Cancer (WBC)), ran 20 times with the average execute time of 3679.2428s. The simulation population was 100 individuals. The individuals from PSO were used to optimize the "don't care" rule. The individuals of the PSO were simulated with 50 swarms and 5 particles. Then, the PSO completed 20 runs with the execute time of 2387.5543s. The optimal fuzzy system which was optimized using the hybrid heuristic algorithm generated accuracy performance as shown in Tables 2, 3. An example of a chromosome from the WBC datasets is shown in Figure 8.

Membership   Linguistic Parameter

[2 2 3 2 2 3 2 2 2] 1.9782 3.4612 7.8462 9.1217 3.3353 3.3353 6.5211 1.8434 4.2727 1.0098 1.0312 1.6815 8.3999 1.9247 3.5459 1.9992 5.2612 1.0692 1.1521 2.1435 2.1556 3.6942 7.6163, …., 2.6585 3.1273 7.0503 9.8831 3.9131 3.1549 6.9534 111111111 2 222123221 4 223222221 4 000000000 2 2101021222 4 223221222

Rule Based

Figure 8 : Chromosome of Interval type-2 Fuzzy Logic System WBC dataset

Table 2: The performance comparison of the iris dataset and the Wisconsin breast cancer dataset using HHType-2

Dataset	**Membership**	**Rule**	**Class**	**Acc**
Iris	[2 2 5 5]	0 0 1 1	2	
		0 1 3 3	2	
		2 1 5 5	3	
Total Acc				95%
WBC	[2 2 3 2 2 3 2 2 2]	0 0 0 0 0 0 0 0 0	2	
		1 0 1 0 2 1 2 2 2	4	
		2 2 3 2 2 1 2 2 2	4	
Total Acc				98.71%

Table 3: Confusion matrix for the iris classification data

Attribute	Setosa	Versicolor	Verginica	Total Testing
Setosa	20	0	0	20
Versicolor	0	19	1	20
Verginica	0	2	18	20
Total	20	21	19	60

To prove the excellent performance of this proposed framework, we compared its accuracy with other well-known classifiers, manipulated for the same problem. Table 4 presents the accuracy performance of classifiers with these algorithms. From Table 4, it can be seen that the accuracy performance of the proposed hybrid heuristic algorithm is among the best achieved.

Table 4 : Comparisons of the HHType-2 and the other algorithms for the iris data

Algorithm	Setosa	Versicolor	Verginica	Acc
1.VSM [11]	100%	93.33%	94%	95.78%
2.NT-growth [11]	100%	93.5%	91.13%	94.87%
3.Dasarathy [11]	100%	98%	86%	94.67%
4.C4 [11]	100%	91.07%	90.61%	93.87%
5.IRSS [12]	100%	92%	96%	96%
6.PSOCCAS [13]	100%	96%	98%	98%
7.HHTypeI [5]	100%	97%	98%	98%
8. HHType II	100%	95%	90%	95%

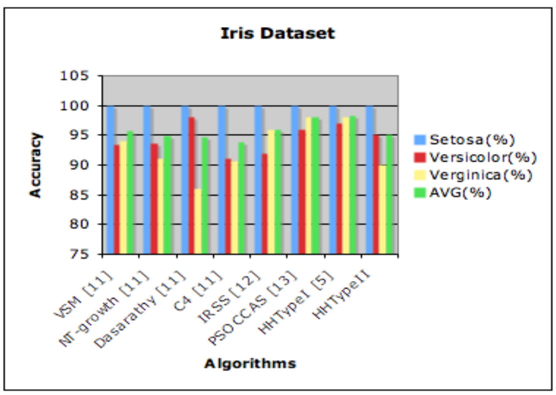

Figure 9: The Bar chart of comparisons of the HHType-2 and the other algorithms, for the Iris dataset

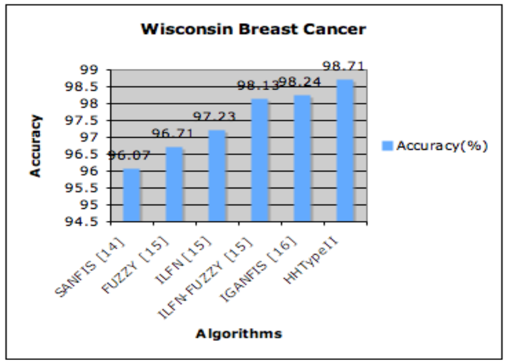

Figure 10 : The Bar chart of comparisons of the HHType-2 and the other algorithms, for the WBC dataset

Table 5 : Comparisons of the HHType-2 and the other algorithms for the WBE data

Algorithm		Accuracy
1.	SANFIS[14]	96.07%
2.	FUZZY[15]	96.71%
3.	ILFN[15]	97.23%
4.	ILFN-FUZZY[15]	98.13%
5.	IGANFIS[16]	98.24%
6.	HHTYPE-2	98.71%

In the same way, we compared the results of the confidence gained from experiments using the algorithms with the same problem to other algorithms. Table 5 shows the accuracy performance of classifier from these algorithms and the confidence of the Wisconsin Breast Cancer dataset using the Hybrid Heuristic Type-2 (HHType-2) algorithm, which results were competitive or even better than any other algorithm. Although GA and PSO are not new, when the two come together they make a powerful new algorithm (Hybrid Heuristic Type-2) for optimization which it is quite efficient referring to the performance.

5 Conclusion

In this paper, a methodology based on a hybrid heuristic algorithm, a combination of PSO and GA approaches, is proposed to build interval type-2 fuzzy set for classification. The algorithms are used to optimize a model by minimizing the number of fuzzy rule, minimizing the number of linguistic variable and maximizing the accuracy of the fuzzy rule base. The performance of the proposed hybrid heuristic algorithm was demonstrated well by applying it to the benchmark problem and the comparison with several other algorithms. For the future research, the application of the proposed algorithm to other problems such as intrusion detection network, network forensic etc., and the use of larger datasets than this research such as Breast Cancer Diagnosis, traffic network dataset etc, will be covered. Therefore, an adaptive on-line inference engine of the interval type-2 fuzzy set will be selected for future research of Breast Cancer Diagnosis for medical training and testing.

References

[1] J. M. Mendel, “Why we need type-2 fuzzy logic system ?” May 2001, http://www.informit.com/articles/article.asp.

[2] J. M. Mendel and R. I. B. John, “Type-2 fuzzy sets made simple,” IEEE Trans. Fuzzy Syst, vol. 10, pp. 117–127, April 2002.

[3] L. Zhao, “Adaptive interval type-2 fuzzy control based on gradient descent algorithm,” in Intelligent Control and Information Processing (ICICIP), vol. 2, 2011, pp. 899–904.

[4] D. Hidalgo, P. Melin, O. Castillo, and G. Licea, “Optimization of interval type-2 fuzzy systems based on the level of uncertainty, applied to response integration in modular neural networks with multimodal biometry,” in The 2010 International Joint Conference on Digital Object Identifier, 2010, pp. 1–6.

[5] A. Sangsongfa and P. Meesad, “Fuzzy rule base generation by a hybrid heuristic algorithm and application for classification,” in National Conference on Computing and Information Technology, vol. 1, 2010, pp. 14–19.

[6]IrisDataset,http://www.ailab.si/orange/doc/datasets/Iris.htm.

[7] Breast Cancer Dataset, http://www.breastcancer.org.

[8] J. Zeng and L. Wang, “A generalized model of particle swarm optimization,” Pattern Recognition and Artificial Intelligence, vol. 18, pp. 685–688, 2005.

[9] H. Ishibuchi, T. Nakashima, and T. Murata, “Three objective genetic based machine learning for linguistic rule extraction,” Information Sciences, vol. 136, pp. 109–133, 2001.

[10] R. Salman, V. Kecman, Q. Li, R. Strack, and E. Test, “Fast k-means algorithm clustering,” Transactions on Machine Learning and Data Mining, vol. 3, p. 16, 2011.

[11] T. P. Hong and J. B. Chen, “Processing individual fuzzy attributes for fuzzy rule induction,” in Fuzzy Sets and Systems, vol. 10, 2000, pp. 127–140.

[12] A. Chatterjee and A. Rakshit, “Influential rule search scheme (irss)a new fuzzy pattern classifier,” in IEEE Transaction on Knowledge and Data Engineering, vol. 16, 2004, pp. 881–893.

[13] L. Hongfei and P. Erxu, “A particle swarm optimization- aided fuzzy cloud classifier applied for plant numerical taxonomy based on attribute similarity,” in Expert Systems with Applications, vol. 36, 2009, pp. 9388–9397.

[14] H. Song, S. Lee, D. Kim, and G. Park, “New methodology of computer aided diagnostic system on breast cancer,” in Second International Symposium on Neural Networks, 2005, pp. 780–789.

[15] P. Meesad and G. Yen, “Combined numerical and linguistic knowledge representation and its application to medical diagnosis,” in Component and Systems Diagnostics, Prognostics, and Health Management II, 2003.

[16] M. Ashraf, L. Kim, and X. Huang;, “Information gain and adaptive neuro-fuzzy inference system for breast cancer diagnoses,” in Computer Sciences and Convergence Information Technology (ICCIT), 2010, pp. 911–915.