International Journal of Soft Computing 8 (2): 126-133, 2013
ISSN: 1816-9503
© Medwell Journals, 2013

A Hybrid Particle Swarm Optimization and Fuzzy Rule-Based System for Breast Cancer Diagnosis

[1]Najmeh Alikar, [1]Salwani Abdullah, [2]Seyed Mohsen Mousavi and [3]Seyed Taghi Akhavan Niaki
[1]School of Computer Science, Faculty of Information Science and Technology,
Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia
[2]Department of Industrial Engineering, Faculty of Industrial and Mechanical Engineering,
Islamic Azad University, Qazvin Branch, Qazvin, Iran
[3]Department of Industrial Engineering, Sharif University of Technology,
P.O. Box 11155-9414 Azadi Ave, 1458889694 Tehran, Iran

Abstract: A hybrid algorithm of a particle swarm optimization and a fuzzy rule-based classification system is proposed in this study to diagnose breast cancer. Two orthogonal and triangular types of fuzzy sets are applied to represent the input variables. In additional, different input membership functions are considered to increase the classification accuracy. The performance of the proposed hybrid algorithm is studied using a classification accuracy measure on the Wisconsin breast cancer dataset. The results of the comparison using different training data sets show the higher performance of the proposed methodology.

Key words: Particle swarm optimization, fuzzy rule based, breast cancer diagnosis, Wisconsin breast cancer dataset, Taguchi Method

INTRODUCTION

The use of classifier systems in medical diagnosis has been gradually increasing in recent decades where some computational tools have been used to automatically identify breast cancer. The most widely used tool is Artificial Neural Networks (ANNs) which is applied to classify benign/malignant breast cancer (Kermani *et al*., 1995; Fogel *et al*., 1995). An automatic diagnosis system based on association rules and neural networks was developed for detecting breast cancer (Karabatak and Ince, 2009). Next in the application of ANNS, a novel improvement in neural network training for pattern classification was presented (Marcano-Cedeno *et al*., 2011). The proposed training algorithm was inspired by the biological meta-plasticity property of neurons and Shannon's information theory.

Particle Swarm Optimization (PSO), stimulated by social behavior of bird flocking or fish schooling is a population-based stochastic optimization meta-heuristic (Kennedy and Eberhart, 1995). PSO shares manysimilarities with evolutionary computation techniques such as Genetic Algorithms (GA). The system is initialized with a population of random solutions and searches for optima by updating generations. However, unlike GA, PSO has no evolution operators such as crossover and mutation. In PSO, the potential solutions, called particles, fly through the problem space by following the current optimum particles. It has been an effective technique to search for the optima of optimization problems. Although, recently only a few papers that employ PSO for breast cancer diagnosis have been published in literature, there is no publication of research work that uses a mixed PSO and rule based fuzzy system to classify breast cancer in Breast Cancer Datasets (BCD) amongst them. For example, the use of the particle swarm optimizer as a new tool for data mining was proposed by Sousa *et al*. (2004) where Wisconsin Breast Cancer Datasets (WBCD) were used to test the algorithm. The simplicity, efficiency and flexibility of PSO were further exploited by Tewolde and Hanna (2007) that proposed a single as well as a multi-surface based data separation methods for classification of BCD. Furthermore, an innovative approach for detection of micro-calcification in digital mammograms using Swarm Optimization Neural Network (SONN) was introduced and proposed by Dheeba and Selvi (2012).

Now a days, due to the need to analyze and summarize high volume of data based on different perspectives, data mining has become popular. Among the techniques available in data mining, classification is used to predict group memberships of data instances.

Corresponding Author: Najmeh Alikar, School of Computer Science, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

This method classifies the data (constructs a model) based on the training set and the values (class labels) in a classifying attribute. One of the common used classification methods is the fuzzy rule based technique that has been successfully applied to various control systems (Lee, 1990). Fuzzy rules in these systems are usually derived using human experts as linguistic if then rules.

In this study, a novel particle swarm optimization algorithm in conjunction with a winner rule-based fuzzy system (NPSO) is proposed to diagnose breast cancer. The proposed rule based fuzzy system is designed based on two studies. First, the research work in which linguistic variables were coded high or low and a fuzzy system with membership functions orthogonal to the input variables was proposed (Pena-Reyes and Sipper, 1999). Second, the proposed fuzzy system in which five linguistic variables, in addition to don't care were used for the membership functions of the input variables (Ishibuchi and Nakaskima, 1999). The present research work is developed based on the combination of these two works where the threshold subsystem in the first is replaced with the winner rule subsystem of the second. The role of NPSO is to generate fuzzy if then rules automatically when integrated with the proposed heuristic outlined by Ishibuchi and Nakaskima (1999).

WISCONSIN BREAST CANCER DATASET

The Wisconsin breast cancer dataset is used in this study to evaluate the performance of the proposed methodology. While the presence of a breast mass is an alert but not always a malignant cancer indication. Fine Needle Aspiration (FNA) is widely adopted to detect breast cancer (Karabatak and Ince, 2009). FNA of breast masses is a cost-effective, non-traumatic and mostly non-invasive diagnostic test that obtains information needed to evaluate malignancy (Ishibuchi and Nakaskima, 1999). In WBCD that was originally collected by William H. Wolberg at the University of Wisconsin-Madison hospital, breast cancer is classified into two classes of benign and malignant. WBCD consist of 699 data including 16 missing and 683 real data. In the dataset, 239 (34.9%) records are malignant and 444 (65.1%) records are benign. Moreover, nine variables of a FNA sample considered relevant for diagnosis were identified and assigned an integer value between 1 and 10. Table 1 shows the variables and their integer values where average, standard deviation, median and mode of each variable are shown. There are a number of studies that used this database to evaluate the performances of their breast cancer diagnosis procedures that were published (Karabatak and Ince, 2009; Pena-Reyes and Sipper, 1999; Hassan *et al*., 2010; Sahan *et al*., 2007).

Table 1: Description of the variables in WBCD

Variables	Integer values of variables	Average	SD	Median	Mode
Clump thickness	1-10	4.442	2.820	4	1
Uniformity of cell size	1-10	3.150	3.065	1	1
Uniformity of cell shape	1-10	3.215	2.988	1	1
Marginal adhesion	1-10	2.830	2.864	1	1
Single epithelial cell size	1-10	3.234	2.223	2	2
Bare nuclei	1-10	3.545	3.644	1	1
Bland chromatin	1-10	3.445	2.450	3	3
Normal nucleoli	1-10	2.870	3.053	1	1
Mitoses	1-10	1.603	1.733	1	1

N = 683 data, 239 malignant and 444 benign

THE PROPOSED HYBRID ALGORITHM

Fuzzy inference system: A Fuzzy Inference System (FIS) is a way of mapping an input space to an output space using fuzzy logic. Figure 1 shows a FIS including a fuzzifier that translates crisp (real-valued) inputs into fuzzy values, an inference engine that applies a fuzzy reasoning mechanism to obtain a fuzzy output, a defuzzifier that translates this latter output into a crisp value and a knowledge base that contains both an ensemble of fuzzy rules (known as rule base or fuzzy models) and an ensemble of membership functions (known as database) (Pena Reyes and Sipper, 1999). A static/dynamic system that makes use of fuzzy sets is called a fuzzy system and fuzzy systems that are commonly defined by if then rules are called rule-based systems. An if-then rule generally takes the form of "If antecedent proposition then consequent proposition".

Fuzzy modeling is the task of identifying the parameters of a fuzzy inference system so that a desired behavior is attained (Yager and Filev, 1994). Moreover, a fuzzy model is directly generated using knowledge from a human expert where either the model by Mamdani and Assilian (1975) or the Takagi Sugeno (TS) Model by Takagi and Sugeno (1985) can be used depending on the desired proposition and implication of the rule (Gaweda, 2002). While the consequent takes fuzzy amounts in Mamdani Model in Takagi Sugeno Model it is the crisp function of antecedent variables.

In this study, two typical sets of two linguistic labels (orthogonal) and five linguistic labels are used. These sets are depicted in Fig. 2 and 3, respectively. In Fig. 2, the two linguistic labels of Low and High are applied to display fuzzy values where two variables a and b are defined as the starting point and the length of membership function edges, respectively. Figure 3 displays five linguistic labels Small (S), Medium Small (MS), Medium (M), Medium Large (ML) and Large (L) where each label is modeled by a triangular membership function.

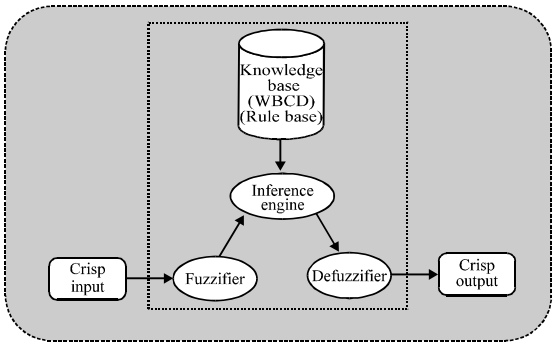


Fig. 1: The proposed fuzzy inference system diagram

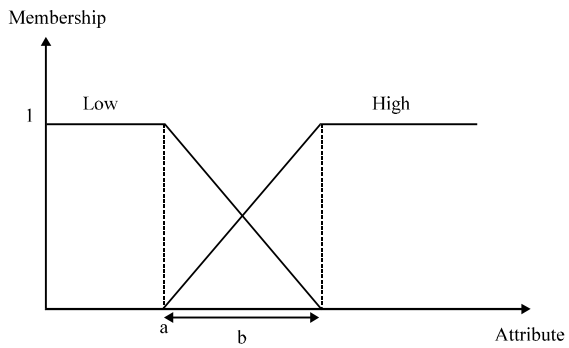


Fig. 2: Orthogonal membership functions of two linguistic values

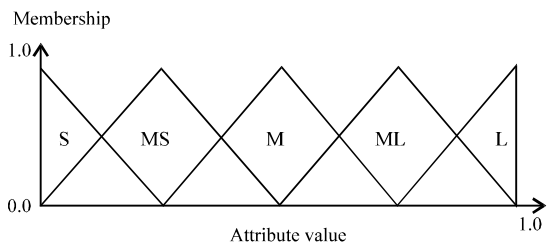


Fig. 3: Membership functions of five linguistic values

Fuzzy rule based system: While fuzzy rule based systems have been applied mainly in control environments, recently they have also been used in pattern classification problems (Lee, 1990). Moreover, automatic generations of fuzzy if then rules have been proposed by many researchers such as Pena-Reyes and Sipper (1999), Ishibuchi and Nakaskima (1999), Grabisch and Dispot (1992) and Ishibuchi *et al.* (1998). All have been shown to work well on a variety of problem domains. Pattern classification typically is a supervised process where based on set of training samples that have been manually classified by experts, a classifier is derived that automatically assigns unseen data sample to the pre-defined classes (Schaefer and Nakashima, 2010).

a_1	a_2	...	a_n	b_1	b_2	....	b_n	L_{11}	L_{12}	...	L_{1n}	....	L_{mn}

Fig. 4: A particle structure of the orthogonal fuzzy sets

In this study, the proposed pattern classification problem contains C = 2 classes of benign and malignant cancer types in n-dimensional pattern space $[0,1]^n$ with continuous attributes and m given training patterns of $x_p = (x_{p1}, x_{p2}, ..., x_{pn})$, p = 1, 2, ..., m where all attributes of each dataset are normalized in the pattern space. Further, the fuzzy if then rules take the following form.

Rule:

$$\text{If } x_1 \text{ is } L_{j1}, x_2 \text{ is } L_{j2} \text{ and } x_n \text{ is } L_{jn} \quad \text{then class } C_j \text{ with } C_{fj}: j = 1, 2, ...N \tag{1}$$

In Eq. 1, R_j is the label of the jth fuzzy if then rule, $L_{j1}, L_{j2},..., L_{jn}$ are antecedent fuzzy sets on the unit interval (0, 1), C_j is the consequent class (in this study 2 classes), and CF_j is the grade of certainty of the fuzzy if then rule R_j. The two antecedent fuzzy sets of this study are shown in Fig. 2 and 3, respectively.

The proposed method to generate fuzzy if then rules in orthogonal fuzzy sets is different from the one used by the triangular form. In this research, the Michigan approach proposed by Ishibuchi and Nakaskima (1999) is used to generate fuzzy rule based systems in which a PSO algorithm is applied to generate fuzzy if then rules.

The proposed particle swarm optimizer: The PSO algorithm of this research is employed to generate fuzzy if-then rules in five stages.

Stage 1 (Initialization): In this stage, an initial population of fuzzy if then rules is generated. Two methods are used to represent a rule. In the first method, the orthogonal fuzzy sets shown in Fig. 2 proposed by Pena-Reyes and Sipper are employed. A particle structure of these fuzzy sets is shown in Fig. 4.

Where a_i and b_i; i = 1, 2, .., n are the points and the lengths of the membership function edges, respectively and $L_{j1},..., L_{jn}$; j = 1, 2, ..., m are antecedent fuzzy sets on the unit interval [0, 1]. In the second method that uses the triangular fuzzy sets shown in Fig. 3, the representation and initialization of fuzzy if then rules are similar to Ishibuchi and Nakaskima (1999) where the first dataset is normalized in interval space [0, 1]. As intended the PSO algorithm is applied to generate antecedent fuzzy sets in which in addition to don't care the aforementioned five linguistic values are used. The six antecedent fuzzy sets are coded in Table 2. Moreover, in order to reduce the number of generated rules and run time, the Heuristic Method of Ishibuchi and Nakashima (1999) is utilized.

Table 2: Decoding of the linguistic value

Linguistic value	Decoding
Small (S)	1
Medium Small (MS)	2
Medium (M)	3
Medium Large (ML)	4
Large (L)	5
Don't Care (DC)	6

Stage 2 (evaluation): In order to evaluate generated particles, the fitness values of all fuzzy if then rules are obtained using Eq. 2 (Ishibuchi and Nakaskima, 1999):

$$\text{Fitness}(R_j) = W_{NCP}.NCP(R_j) - W_{NMP}.NMP(R_j) \quad (2)$$

Where:

W_{NCP} = The reward for a correct classification
$NCP(R_j)$ = The number of correctly classified training patterns using R_j
W_{NMP} = The misclassification penalty
$NMP(R_j)$ = The number of misclassified training patterns using R_j

In order to classify a rule to a class, certainty grades are first calculated. Then, the fitness values of the rules are used to assign them to different classes. While the calculations involved to obtain certainty grades are different in orthogonal and triangular fuzzy sets, in this study, the heuristic proposed by Ishibuchi and Nakaskima (1999) is used. Having m training patterns, $x_p = (x_{p1}, x_{p2}, ..., x_{pm})$, $p = 1, 2, ..., m$ for a n-dimensional C-class pattern classification problem, the certainty grade of rule R_j in orthogonal fuzzy sets is obtained as follows:

Step 1: Calculate membership functions of all the variables in every training pattern x_p with the fuzzy if then rule R_j using Eq. 3 (the fuzzifier phase).

$$\mu_{Low}(x_p) = \frac{(a+b)-x_p}{b} \qquad \mu_{High}(x_p) = 1-\mu_{Low}(x_p) \quad (3)$$

where, $\mu_{Low}(x_p)$and $\mu_{High}(x_p)$ are the membership function of labels low and high in the training pattern x_p, a and b are the point and the length of a membership function edge, respectively.

Step 2: Calculate the truth value of rule R_j, denoted by $\mu_j(x_p)$ by applying the fuzzy and the min operators to combine the antecedent clauses (the membership values) in a fuzzy manner; this results in the output truth value, namely, a continuous value which represents the rule's degree of activation (Pena-Reyes and Sipper, 1999).

Step 3: For class h calculate $\beta_{class\,h}(j)$ as:

$$\beta_{class\,h}(j) = \sum_{x_p \in h} \mu_j(x_p) \quad (4)$$

Step 4: Find class $\hat{h}_j$ with the maximum value of as $\beta_{class\,h}(j)$ as:

$$\beta_{class\,\hat{h}}(j) = \max_{1 \le k \le C} \{\beta_{class\,k}(j)\} \quad (5)$$

If two or more classes take the maximum value, the consequent class C_j of the rule R_j cannot be determined uniquely. In this case, specify $C_j = \phi$. However, if a single class $\hat{h}$ takes the maximum value, let C_j be $\hat{h}$. Then, the certainty grade (CF_j) is determined by:

$$CF_j = \frac{\beta_{class\,\hat{h}}(j) - \bar{\beta}}{\sum_h \beta_{class\,h}(j)} \quad (6)$$

With:

$$\bar{\beta} = \frac{\sum_{h \ne \hat{h}} \beta_{class\,h}(j)}{C-1} \quad (7)$$

Step 5: Choose the winner rule using Hassan *et al*. (2010):

$$\mu_j(x_p).CF_j = \max_{1 \le k \le C} \{\mu_k(x_p).CF_k\} \quad (8)$$

In triangular fuzzy sets, the certainty grade of rule R_j is obtained similar to the one proposed for the orthogonal fuzzy sets. The only difference is that the training pattern x_p is first normalized in the interval [0, 1] and then the calculation in step 1 becomes:

$$\beta_{class\,h}(j) = \sum_{x_p \in h} \mu_j(x_p)$$

where, $\mu_j(x_p)$ is obtained as:

$$\mu_j(x_p) = \min\{\mu_{j1}(x_{p1}), ..., \mu_{jn}(x_{pn})\} \quad (9)$$

Figure 5 shows the proposed diagnosis system for the two fuzzy sets where it consists of two fuzzy subsystems displayed to the left. In fact, it shows the entire fuzzy inference system of Fig. 1 along with the winner rule subsystem.

Stage 3 (selection): In a d-dimensional search space, let:

$$\vec{x}_k^i = \left\{x_{k,1}^i, x_{k,2}^i, ..., x_{k,d}^i\right\}$$

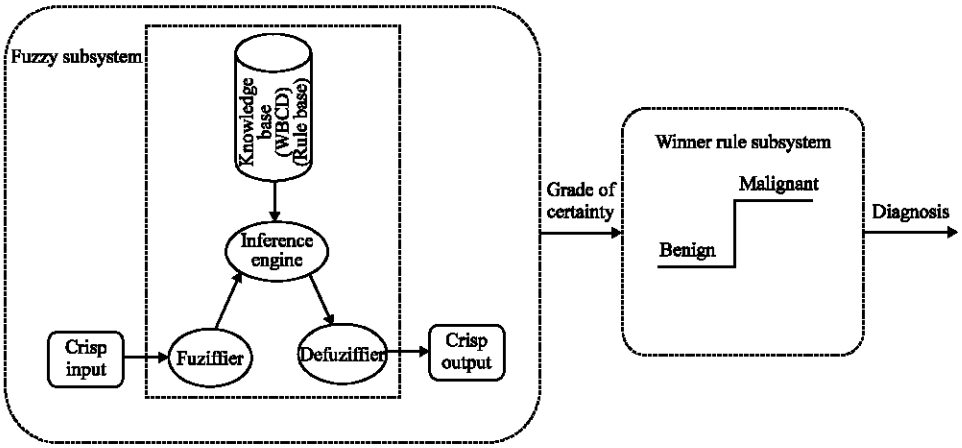


Fig. 5: The proposed diagnosis system

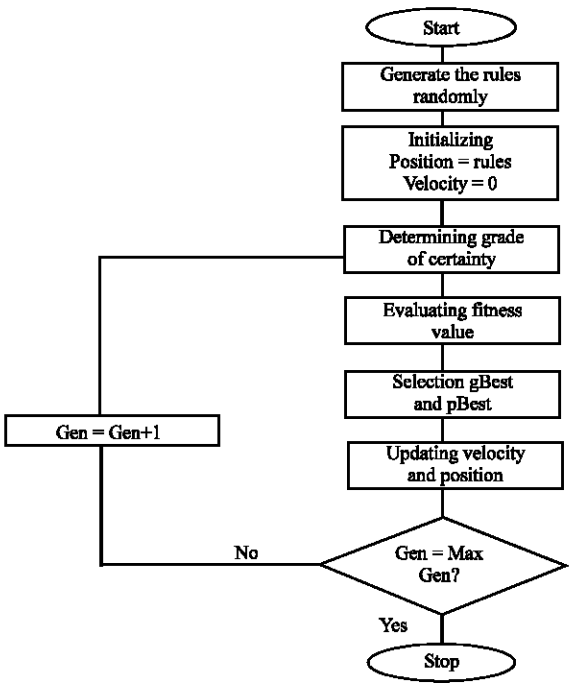


Fig. 6: The flow chart of the proposed PSO algorithm

And:

$$\vec{v}_k^i = \left\{v_{k,1}^i, v_{k,2}^i, \ldots, v_{k,d}^i\right\}$$

be the position and the velocity of particle i at time k, respectively. In this research, the primary position of particle $i(\vec{x}_0^i)$ is determined using randomly generated fuzzy if then rules and its initial velocity is set $\vec{v}_0^i = 0$. For particle i, the best solution (fitness) found until time k, $p\vec{B}est_k^i$ and the best solution found at time k (the current population), $g\vec{B}est_k^i$ are selected based on Eq. 10 and 11, respectively.

$$\vec{pBest}_k^i = \left\{pBest_{k,1}^i, pBest_{k,2}^i, \ldots, pBest_{k,d}^i\right\} \quad (10)$$

$$\vec{gBest}_k^i = \left\{gBest_{k,1}^i, gBest_{k,2}^i, \ldots, gBest_{k,d}^i\right\} \quad (11)$$

Stage 4 (New solution generation): Equation 12 and 13 are used to generate new solutions by updating the velocities and the positions at time k.

$$v_{k+1,d}^i = w.v_{k,d}^i + C_1.r_1.\left(pBest_{k,d}^i - x_{k,d}^i\right) + C_2.r_2.\left(gBest_{k,d}^i - x_{k,d}^i\right) \quad (12)$$

$$x_{k+1,d}^i = x_{k,d}^i + v_{k+1,d}^i \quad (13)$$

Where:

r_1 and r_2	=	Uniform random numbers between 0 and 1
Coefficients C_1 and C_2	=	The given acceleration constants towards $p\vec{B}est$ and $g\vec{B}est$, respectively
w	=	The inertia weight

Introducing a linearly decreasing inertia weight into the original PSO significantly improves its performance (Shi and Eberhart, 1999). The linear distribution of the inertia weight is expressed as follows (Naka *et al.*, 2001):

$$w = w_{max} - \frac{w_{max} - w_{min}}{iter_max} iteration \quad (14)$$

where, iter_max is the maximum number of iterations and iteration is the current number of iterations.

Equation 14 introduces how the inertia weight is updated, considering w_{max} and w_{min} are the minimum and the maximum weights, respectively. The related results of the two parameters $w_{max} = 0.9$ and $w_{min} = 0.4$ are reported by Naka *et al.* (2001) and Chen and Hsu (2006).

Stage 5 (stopping criterion): While there are several stopping conditions available for PSO algorithm, in this study the total number of generations is chosen. Figure 6 depicts a flow chart of the proposed NPSO.

EXPERIMENTAL RESULTS

To evaluate the performance of NPSO in orthogonal fuzzy sets environment, five systems are selected where the systems are sorted based on one to five rules. The proposed NPSO algorithm runs on these five systems according to Fig. 5 with $W_{CP} = 1$ and $W_{NMP} = 0$. The algorithm is coded in MATLAB 7.8 (R 2009 a) and runs on a laptop with Intel® core™ 2Duo cpu2.00 GHz, 4 GB of RAM qualifications. The results of employing the algorithm are displayed in Table 3 where the training data are the entire WBCD, i.e., 683 dataset. The results show that system 3 has the best accuracy of 97.1% with a feature average of 2.7. The description of the rules for system 3 using orthogonal fuzzy sets is given:

- Rule 1 if (x_3 is High) then C = M with CF = 0.89
- Rule 2 if (x_2 is High) and (x_4 is Low) and (x_9 is High) then C = B with CF = 0.91
- Rule 3 if (x_2 is Low) and (x_6 is Low) and (x_8 is Low) and (x_9 is Low) then C = B with CF = 0.98

In the triangular fuzzy sets, dataset is first normalized in interval [0, 1] by:

$$\text{Normalized data (x)} = e + \frac{(x-A).(f-e)}{B-A} \qquad (15)$$

in which A = 1 is the minimum, B = 10 is the maximum, e = 0 is where the normalized data set starts, f = 1 is where the normalized data set ends and x is the data value that is normalized. The problem runs by the NPSO algorithm in four cases when 8, 6, 4 and 2 features in each particle are don't care respectively. As an example, the best particle structure in the first generation when 3 features are don't care is as follows:

5	4	6	2	5	6	6	2	5	5

where diagnosis is benign with CF = 0.003 and based on Table 2, don't care = 6.

In each system, the algorithm runs were based on different rules with the above selected features. Figure 7 displays the number of features selected in 15 different rules based on the entire dataset (683 data) as the training data. The highest accuracy of 97.4% is associated with the case where 8 among 9 features are don't care. In Table 4, a sample of five rules is demonstrated in details for 8 features as don't care with an accuracy of 95.02%.

In order to validate the results obtained and to evaluate the performances of the proposed methodology, NPSO as well as some other related methods are employed on WBCD for both orthogonal and triangular fuzzy sets.

The compared methods are genetic algorithms by Chen and Hsu (2006), DPSO by Yeh *et al.* (2009), HMM-Fuzzy Model by Hassan *et al.* (2010), MPANN by Abbass (2002) and evolutionary neural network by

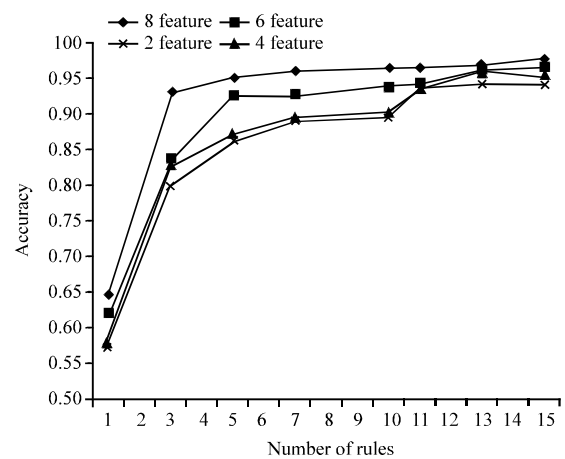


Fig. 7: The accuracy of number of rules for triangular fuzzy sets

Table 3: The optimal result of orthogonal fuzzy set for five systems (training data = 683)

No. rules per system	Accuracy	Average feature per system
1	0.942	2.0
2	0.946	3.5
3	0.971	2.7
4	0.959	5.0
5	0.968	5.0

Table 4: Five fuzzy rules of 8 selected features for triangular fuzzy sets

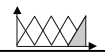
Rules	Shapes of fuzzy rules (others are don't care)	Variables	Class	CF
01		6	M	0.9537
02		8	B	0.7908
03		6	M	0.4295
04		3	M	0.7833
05		1	M	0.9891

Table 5: The accuracy comparisons of the proposed methodology with other methods

Models	Accuracy (683 training data) (%)	Accuracy (466 training data) (%)	Accuracy (400 training data) (%)
NPSO (triangular fuzzy sets)	97.40	99.35	99.25
NPSO (orthogonal fuzzy sets)	97.10	99.14	99.50
GA (Ishibuchi and Nakaskima, 1999)	96.97	-	-
GA (Chen and Hsu, 2006)	-	96.99	-
DPSO (Yeh *et al*., 2009)	-	98.71	-
HMM-fuzzy model (Hassan *et al*., 2010)	-	-	98.89
MPANN (Abbass, 2002)	-	-	98.60
Evolutionary neural network (Fogel *et al*., 1995)	-	-	98.10

Foget *et al*. (1995). The methods are employed on three training datasets of 683 data used by Pena-Reyes and Sipper (1999) where 466 data are utilized by Chen and Hsu (2006) and Yeh *et al*. (2009) and 400 training data were used by Fogel *et al*. (1995), Hassan *et al*. (2010) and Abbass (2002). Table 5 contains the accuracy of the methods under investigation. The results in Table 5 show that for the triangular fuzzy set, NPSO has the highest accuracy of 97.4 and 99.35% among the methods that work with 683 and 400 training data, respectively. Moreover, for the orthogonal fuzzy set, the proposed NPSO has the highest accuracy of 99.5% among all methods that work with 400 training data. This accuracy is the highest obtained value so far to classify breast cancer diagnosis using WBCD. This shows the accuracy and effectiveness of the proposed NPSO methodology in classification and recognition of breast cancer.

CONCLUSION

In this study, a Novel Particle Swarm Optimization algorithm (NPSO) was proposed to generate fuzzy if then rules in breast cancer classification. Two linguistic variable types of orthogonal and triangular fuzzy sets were used to represent input features. In order to have better performances, a Taguchi Method was first proposed to calibrate the parameters NPSO and then the Wisconsin breast cancer dataset was used to evaluate the performance of NPSO for both types of fuzzy sets. The proposed methodology was compared with some related works with different training data where its high performance was concluded. The proposed model achieved accuracies of 97.40 and 99.35% for triangular fuzzy sets on 683 and 466 training data, respectively and 99.5% for orthogonal fuzzy sets on 400 training data. This shows the effectiveness of the proposed method in breast cancer classification.

REFERENCES

Abbass, H.A., 2002. An evolutionary artificial neural networks approach for breast cancer diagnosis. Artif. Intell. Med., 25: 265-281.

Chen, T.C. and T.C. Hsu, 2006. A GAs based approach for mining breast cancer pattern. Expert Syst. Appl., 30: 674-681.

Dheeba, J. and S.T. Selvi, 2012. A swarm optimized neural network system for classification of microcalcification in mammograms. J. Med. Syst., 36: 3051-3061.

Fogel, D.B., E.C. Wasson III and E.M. Boughton, 1995. Evolving neural networks for detecting breast cancer. Cancer Lett., 96: 49-53.

Gaweda, A.E., 2002. Optimal data-driven rule extraction using adaptive fuzzy-neural models. Ph.D. Thesis, Department of Computer Engineering and Computer Science, University of Louisville, Louisville, KY., USA.

Grabisch, M. and F. Dispot, 1992. A comparison of some methods of fuzzy classification on real data. Proceedings of the 2nd International Conference on Fuzzy Logic and Neural Networks, July 17-22, 1992, Iizuka, Japan, pp: 659-662.

Hassan, M.R., M.M. Hossain, R.K. Begg, K. Ramamohanarao and Y. Morsi, 2010. Breast-cancer identification using HMM-fuzzy approach. Comput. Biol. Med., 40: 240-251.

Ishibuchi, H. and T. Nakaskima, 1999. Improving the performance of fuzzy classifier systems for pattern classification problems with continuous attributes. IEEE Trans. Ind. Electron., 46: 1057-1068.

Ishibuchi, H., T. Murata and M. Gen, 1998. Performance evaluation of fuzzy rule-based classification systems obtained by multi-objective genetic algorithms. Comput. Ind. Eng., 35: 575-578.

Karabatak, M. and M.C. Ince, 2009. An expert system for detection of breast cancer based on association rules and neural network. Expert Syst. Appl., 36: 3465-3469.

Kennedy, J. and R. Eberhart, 1995. Particle swarm optimization. Proc. IEEE Int. Conf. Neural Networks, 4: 1942-1948.

Kermani, B.G., M.W. White and H.T. Nagle, 1995. Feature extraction by genetic algorithms for neural networks in breast cancer classification. Proceedings of the IEEE 17th Annual Conference Engineering in Medicine and Biology Society, Volume 1, September 20-25, 1995, Montreal, Canada, pp: 831-832.

Lee, C.C., 1990. Fuzzy logic in control systems: Fuzzy logic controller. I. IEEE Trans. Syst. Man Cybernet., 20: 404-415.

Mamdani, E.H. and S. Assilian, 1975. An experiment in linguistic synthesis with a fuzzy logic controller. Int. J. Man-Mach. Stud., 7: 1-13.

Marcano-Cedeno, A., J. Quintanilla-Dominguez and D. Andina, 2011. Breast cancer classification applying artificial metaplasticity algorithm. Neurocomputing, 74: 1243-1250.

Naka, S., T. Genji, T. Yura and Y. Fukuyama, 2001. Practical distribution state estimation using hybrid particle swarm optimization. Proceedings of the IEEE Power Engineering Society Winter Meeting, Volume 2, January 28-February 1, 2001, Columbus, OH., USA., pp: 815-820.

Pena-Reyes, C.A. and M. Sipper, 1999. A fuzzy-genetic approach to breast cancer diagnosis. Artif. Intell. Med., 17: 131-155.

Sahan, S., K. Polat, H. Kodaz and S. Gunes, 2007. A new hybrid method based on fuzzy-artificial immune system and k-nn algorithm for breast cancer diagnosis. Comput. Biol. Med., 37: 415-423.

Schaefer, G. and T. Nakashima, 2010. Hybrid cost-sensitive fuzzy classification for breast cancer diagnosis. Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, August 31-September 4, 2010, Buenos Aires, Argentina, pp: 6170-6173.

Shi, Y. and R.C. Eberhart, 1999. Empirical study of particle swarm optimization. Proceeding of the IEEE International Conference on Evolutionary Computer, July 6-9, 1999, Piscataway, NJ., USA., pp: 1945-1950.

Sousa, T., A. Silva and A. Neves, 2004. Particle swarm based data mining algorithms for classification tasks. Parallel Comput., 30: 767-783.

Takagi, T. and M. Sugeno, 1985. Fuzzy identification of systems and its applications to modeling and control. IEEE Trans. Syst. Man Cybermetics, 15: 116-132.

Tewolde, G.S. and D.M. Hanna, 2007. Particle swarm optimization for classification of breast cancer data using single and multisurface methods of data separation. Proceedings of the IEEE International Conference on Electro/Information Technology, May 17-20, 2007, Chicago, IL., USA., pp: 443-446.

Yager, R.R. and D.P. Filev, 1994. Essentials of Fuzzy Modeling and Control. 1st Edn., Willy, New York, ISBN-13: 978-0471017615, pp: 408.

Yeh, W.C., W.W. Chang and Y.Y. Chung, 2009. A new hybrid approach for mining breast cancer pattern using discrete particle swarm optimization and statistical method. Expert Syst. Appl., 36: 8204-8211.