

# AN APPROACH FOR BREAST CANCER DIAGNOSIS CLASSIFICATION USING NEURAL NETWORK

Htet Thazin Tike Thein<sup>1</sup> and Khin Mo Mo Tun<sup>2</sup>

<sup>1</sup>Ph.D Student, University of Computer Studies, Yangon, Myanmar

<sup>2</sup>Department of Computational Mathematics, University of Computer Studies, Yangon, Myanmar

## ABSTRACT

*Artificial neural network has been widely used in various fields as an intelligent tool in recent years, such as artificial intelligence, pattern recognition, medical diagnosis, machine learning and so on. The classification of breast cancer is a medical application that poses a great challenge for researchers and scientists. Recently, the neural network has become a popular tool in the classification of cancer datasets. Classification is one of the most active research and application areas of neural networks. Major disadvantages of artificial neural network (ANN) classifier are due to its sluggish convergence and always being trapped at the local minima. To overcome this problem, differential evolution algorithm (DE) has been used to determine optimal value or near optimal value for ANN parameters. DE has been applied successfully to improve ANN learning from previous studies. However, there are still some issues on DE approach such as longer training time and lower classification accuracy. To overcome these problems, island based model has been proposed in this system. The aim of our study is to propose an approach for breast cancer distinguishing between different classes of breast cancer. This approach is based on the Wisconsin Diagnostic and Prognostic Breast Cancer and the classification of different types of breast cancer datasets. The proposed system implements the island-based training method to be better accuracy and less training time by using and analysing between two different migration topologies.*

## KEYWORDS

*Neural Network, Differential Evolution, Island Model, Classification, Breast Cancer Diagnosis*

## 1. INTRODUCTION

Breast cancer, is cancer that affects today more women in the world. Thus, the fight against cancer is far from completed. Medicine advances on all fronts to improve the care of patients and defeat this disease of the century. Because of this, it is essential that several disciplines continue to make their contribution and particularly data mining or artificial Intelligence. To provide assistance to the medical, robust and reliable diagnosis, neural networks can be a powerful tool for distributed diagnosis [1]. In this paper, we tested the performance of the neural networks based on the Wisconsin Breast Cancer Database (WBCD). The problem of breast cancer detection led researchers and experts in this field to focus on other trends, such that new technologies other human to address this social problem. The objective of our work is to create a new approach that allows whether a patient has a benign cancer or malignant following several descriptors. To achieve this, we propose a solution based on the concept of neural networks [2] [3] [4]. Recently, the neural network has become a popular tool in the classification of Cancer Dataset [5] [6] [7] [8]. This is particularly due to its ability to represent the behaviour of linear or

nonlinear functions multidimensional and complex.

Wisconsin Breast Cancer Data (WBCD) is analysed by various researchers on medical diagnosis of breast cancer in neural network literature [15],[16],[17],[18],[19]. In [20] Breast cancer is diagnosed using feed forward neural networks by comparing the hidden neurons. In [21], the performance comparison of the multi-layered perceptron networks using various back propagation algorithms for breast cancer diagnosis is discussed. The training algorithms used are gradient descent with momentum and adaptive learning, resilient back propagation, Quasi-Newton and Levenberg-Marquardt. The performances of these four algorithms are compared with the standard steepest descent back propagation algorithm. The MLP network using the Levenberg-Marquardt algorithm displays the best performance. The seventh attribute called Bare Nuclei of WBCD has 16 missing values. In [20], the 16 missing value instances have been left out while using WBCD for Breast Cancer diagnosis. The constructed feed forward neural network has been evaluated for breast cancer detection without replacing missing values [22]. Eliminating some instances will affect the diagnosis accuracy.

In this paper, the network is trained by DE which has been parallelized in order to achieve better performance. This paper presents a result of direct classification of data after replacing missing values using median method for the WBCD dataset by using island differential evolution algorithm. The training algorithms are compared using accuracy and computing time. In this work, a parallel approach, which uses neural network technique, is proposed to help in the diagnosis of breast cancer. The neural network is trained with breast cancer data by using feed forward neural network model and island differential evolution learning algorithm with momentum and variable learning rate. The performance of the network is evaluated. The proposed system implements the island-based training method to be better accuracy and less training time by using and analysing between two different migration topologies.

## 2. DIFFERENTIAL EVOLUTION ALGORITHM

Having developed an ANN-based process model, a DE algorithm is used to optimize the N-dimensional input space of the ANN model. Conventionally, various deterministic gradient-based methods are used for performing optimization of the phenomenological models. Most of these methods require that the objective function should simultaneously satisfy the smoothness, continuity, and differentiability criteria. Although the nonlinear relationships approximated by an ANN model can be expressed in the form of generic closed-form expressions, the objective function(s) derived thereby cannot be guaranteed to satisfy the smoothness criteria. Thus, the gradient-based methods cannot be efficiently used for optimizing the input space of an ANN model and, therefore, it becomes necessary to explore alternative optimization formalisms, which are lenient towards the form of the objective function.

In the recent years, Differential Evolution (DE) that are members of the stochastic optimization formalisms have been used with a great success in solving problems involving very large search spaces [14]. The DEs were originally developed as the genetic engineering models mimicking the population evolution in natural systems. Specifically, DE like genetic algorithm (GA) enforces the —survival-of-the-fittest! and —genetic propagation of characteristics! principles of biological evolution for searching the solution space of an optimization problem. The principal features possessed by the DEs are: (i) they require only scalar values and not the second- and/or first-order derivatives of the objective function, (ii) the capability to handle nonlinear and noisy objective functions, (iii) they perform global search and thus are more likely to arrive at or near the global optimum and (iv) DEs do not impose pre-conditions, such as smoothness, differentiability and

continuity, on the form of the objective function.

Differential Evolution (DE), an improved version of GA, is an exceptionally simple evolution strategy that is significantly faster and robust at numerical optimization and is more likely to find a function's true global optimum. Unlike simple GA that uses a binary coding for representing problem parameters, DE uses real coding of floating point numbers. The mutation operator here is the addition instead of bit-wise flipping used in GA. And DE uses non-uniform crossover and tournament selection operators to create new solution strings. Among the DEs advantages are its simple structure, ease of use, speed and robustness. It can be used for optimizing functions with real variables and many local optima.

This paper demonstrates a successful application of DE with island model. As already stated, DE in principle is similar to GA. So, as in GA, we use a population of points in our search for the optimum. The population size is denoted by NP. The dimension of each vector is denoted by D. The main operation is the NP number of competitions that are to be carried out to decide the next generation. To start with, we have a population of NP vectors within the range of the objective function. We select one of these NP vectors as our target vector. We then randomly select two vectors from the population and find the difference between them (vector subtraction). This difference is multiplied by a factor F (specified at the start) and added to the third randomly selected vector. The result is called the noisy random vector. Subsequently, the crossover is performed between the target vector and noisy random vector to produce the trial vector. Then, a competition between the trial vector and target vector is performed and the winner is replaced into the population. The same procedure is carried out NP times to decide the next generation of vectors. This sequence is continued till some convergence criterion is met. This summarizes the basic procedure carried out in differential evolution.

### **3. ISLAND MODEL STRATEGY**

The main difference between the island model and the single population model is the separation of individuals into islands. As against the master-slave model the communication to computation ratio of the island model approach is low, owing to the low communication frequency between the islands. Also, separating individuals separately from each other results in a qualitative changes in the behaviour of the algorithm.

In the island model approach, each island executes a standard sequential evolutionary algorithm. The communication between sub-population is assured by a migration process. Some randomly selected individuals (migration size) migrate from one island to another after every certain number of generations (migration interval) depending upon a communication topology (migration topology). The two basic and most sensitive parameters of island model strategy are: migration size, which indicates the number of individuals migrating and controls the quantitative aspect of migration; and migration interval denoting the frequency of migration.

#### **3.1. Migration Topology**

The migration topology describes which islands send individuals to which islands. There are many topologies. This system investigates the random topology and torus topology and compares their results. In this paper, simulations were run with setups of five islands [11].

#### **3.2. Migration Policy**

A migration policy consists of two parts. The first part is the selection of individuals, which shall

be migrated to another island. The second part is to choose which individuals are replaced by the newly obtained individuals. The policy of migration used was random-random policy in which any random string from an island replaces any other random string of another island based on the random topology.

### **3.3. Migration Interval**

In order to distribute information about good individuals among the islands, migration has to take place. This can either be done in synchronous way every  $n^{\text{th}}$  generation or in an asynchronous way, meaning migration takes place at non-periodical times. It is commonly accepted that a more frequent migration leads to a higher selection pressure and therefore a faster convergence. But as always with a higher selection pressure come the susceptibility to get stuck in local optima. In this system, various migration intervals are experimented to find the best solution for the neural network training.

### **3.4. Migration Size**

A further important factor is the number of individuals which are exchanged. According to these studies the migration size has to be adapted to the size of a subpopulation of an island. When one migrates only a very small percentage, the influence of the exchange is negligible but if too much individuals are migrated, these new individuals take over the existing population, leading to a decrease of the global diversity. In this system, the migration sizes were chosen to be approximately 10% of the size of a subpopulation.

## **4. ISLAND DIFFERENTIAL EVOLUTION APPROACH TO NEURAL NETWORK**

### **4.1. Neural Network as Breast Classifier**

Neural networks (NNs) have been widely used in various fields as an intelligent tool in recent years, such as artificial intelligence, pattern recognition, medical diagnosis, machine learning and so on [10]. Among them, pattern recognition is a class of problem that neural network is particularly suitable for solving. In fact, NN can be viewed as the mapping from input to output. If each different input is regarded as a kind of input mode, the mapping to the output is considered as output response model, the mapping from input to output is undoubtedly the issue of pattern classification. Nevertheless, learning is the first step to design classifier, that is, ascertain the requirements for the classification error rate and choose appropriate discrimination rule.

Strictly speaking, the learning algorithm of NN is a supervised learning method by training feed forward neural network using error back propagation technique to determine the parameters of neural network. Its unique advantages lie in the greatest tolerance of the noisy data, as well as the ability to classify untrained data pattern. When it comes to breast cancer data classification, the major steps of using neural network learning algorithm can be summarized as follows; to begin with, through the provision of training samples and the class of sample, the network prediction of each sample is compared with the actual known class label, and then the weight of each training sample is adjusted to achieve the purpose of classifying other sample data. The use of neural network to classify breast cancer data is illustrated in Fig. 1. Each node in the network corresponds to the output node of a network unit, while the real lines from the bottom into node are regarded as its input. Intermediate cell is called the hidden layer units, whose output are only in the internal network, not a part of all the network output. The output of the hidden layer is

considered as the input of two output units, corresponding to a result of the diagnosis of breast cancer, benign or malignant tumor.

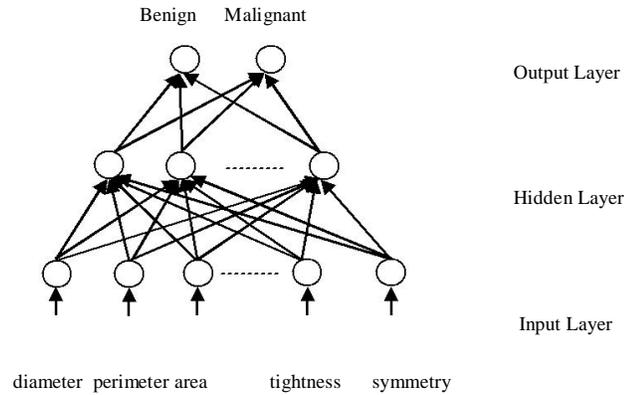


Fig 1. Neural network topology

#### 4.2. Processing Procedure of Island Differential Evolution Algorithm

The DE algorithm was proposed by Price and Storn [9]. The DE algorithm has the following advantages over the traditional genetic algorithm: it is easy to use and it has efficient memory utilization, lower computational complexity (it scales better when handling large problems), and lower computational effort (faster convergence) [10]. DE is quite effective in nonlinear constraint optimization and is also useful for optimizing multimodal problems [23]. The major steps of IDE are as follows:

- Initialize population pop: Create a population from randomly chosen object vectors with dimension  $P*d$ , where  $P$  is the number of population and  $d$  is the number of weights of the neural network.

$$P_g = (w_{1,g}, \dots, w_{p,g})^T, g = 1, \dots, g_{max}$$

$$w_{i,g} = (w_{1,i,g}, \dots, w_{d,i,g}), i = 1, \dots, P$$

where  $d$  is the number of weights in the weight vector and in  $w_{i,g}$ ,  $i$  is index to the population and  $g$  is the iteration (generation) to which belongs.

- Evaluate all the candidate solutions inside the pop for a specified number of iterations.
- For each  $i^{th}$  candidate in pop, select the random population members,  $r_1, r_2, r_3 \in \{1, 2, \dots$

$P\}$

- Apply a mutation operator to each candidate in a population to yield a mutant vector i.e.  $v_{j,i,g-1} = w_{j,r1,g} + F (w_{j,r2,g} - w_{j,r3,g})$ , for  $j = 1, \dots, d (i \neq r_1 \pm r_2 \neq r_3) \in \{1, \dots, P\}$  and  $F \in$

$[0,1]$

where  $F$  denotes the mutation factor.

- Apply crossover i.e. each vector in the current population is recombined with a mutant vector to produce trail vector.

$$t_{j,i,g+1} = \begin{cases} v_{j,i,g+1}, & \text{if } rand_i [0,1] \leq CR \\ w_{j,i,g} & \text{otherwise} \end{cases}$$

where  $CR \in [0,1]$

- Apply selection operation.

If  $MSE(w_{j,i,g}) < MSE(t_{j,i,g+1})$  then

replace  $t_{j,i,g+1}$  by  $w_{j,i,g}$  in the new population

Else leave  $t_{j,i,g+1}$  in the new population

- Select individuals from step 6 according to random-random migration strategy.
- Migrate and replace optimal solution according to migration topology.
- Repeat steps 1 to 8 until stopping criteria is reached.

In summary, differential evolution algorithm proceeds from an initial population, such as, selection, crossover and mutation, to search a better space step by step until reach the optimal solution. It is obvious that differential evolution algorithm is an optimization methodology. Here, differential evolution algorithm with global optimization strategies is integrated into the neural network model to improve the classification rate of breast cancer diagnosis. The recommended population size in DE is ten times the number of gens in a string. Large population size makes the objective functions expensive. Reducing the population size may lead to reduction in search space resulting in poor quality of search.

## 5. DESCRIPTION OF WISCONSIN BREAST CANCER DATABASE (WBCD)

Breast cancer becomes one of the leading causes of death of women in the world. The mammography technique has been proved to be an effective tool for the detection of breast cancer in its earlier phase. Detection of Clusters is an important sign in the identification of micro calcification of mammograms. In our paper, a medical data based on breast cancer attributes was used for the purpose of classification between two types of cancers, benign and malignant.

### 5.1. Breast Cancer Dataset

The database used for detection of breast cancer by artificial neural networks is publicly available in the Internet [12]. The dataset is provided by university of Wisconsin hospital, Madison from Dr. William H. Wolberg. This database has 699 instances and 10 attributes including the class attribute. Attribute 1 through 9 are used to represent instances. Each instance has one of two possible classes: benign or malignant. According to the class distribution 458 or 65.5% instances are Benign and 241 or 34.5% instances are Malignant. Table 1 provides the attribute information.

The original data can be presented in the form of analog values with values ranging from 0-10. Conversion of the given data sets into binary can be done based on certain ranges, which are defined for each attribute [13].

Table 1. Attributes of breast cancer data

S.no	Attribute	Domain
1	Clump thickness	1 -10
2	Uniformity of cell size	1 -10
3	Uniformity of cell shape	1 -10
4	Marginal adhesion	1 -10
5	Single epithelial cell size	1 -10
6	Bare nuclei	1 -10
7	Bland chromatin	1 -10
8	Normal nucleoli	1 -10
9	Mitosis	1 -10
	Class	2 for benign, 4 for malignant

## 6. EXPERIMENTAL DETAILS

Java programming language, which is the platform independent and a general-purpose development language, is used to implement the proposed system. First of all, load data set and replace missing values by using Median method. And then, normalize the data using min-max normalization. The training patterns of breast cancer dataset are used as input data. Attributes are scaled to fall within a small specific range by using min-mix normalization. At the start of the algorithm, dataset were loaded from the database. In the next step, each chromosome or vector is randomly initialized with random neural network weight. Fitness of each chromosome is evaluated using following step. Fitness defined how well a chromosome solves the problem in hand. The first step converts chromosome's genes into neural network and fitness is calculated for each individuals. Mutation operator produce the trial vector from parent vector and randomly selected three vectors. Crossover recombines the parent vector and trial vector to produce offspring. By using mutation and crossover, some genes are modified that mean weights are updated. Fitness of offspring is calculated and compare with fitness of parent vector, the chromosome with high fitness survive and next generation begin. Choose individuals according to migration policy. Migrate and replace individuals according to migration topology. Experiments are performed with random-random migration policy and two different migration topologies. In the experiments we used identical islands, i.e islands with same parameters. Simulations were run with setups of five, seven and nine islands. As can be observed that the error value of test data are nearly same for the five and seven island models, it goes a little high for nine islands, but no significant change in the number of islands. Some experiments were also conducted with four islands and it had no significant influence on the algorithmic behavior. We therefore chose a modest five island setup for our rest of the experiments. The policy of migration used was

random-random policy in which random string from an island replaces other random string of another island based on migration topology.

### 6.1. Data Normalization

The data normalization is considered as the most important pre-processing step using neural networks. To improve the performance of multilayer neural networks, it is better to normalize the data entry such that will be found in the interval of [0 1]. To transform the data into digital form, and use it as inputs of the neural network, scaling or normalization should be realized for each attribute. The nine numerical attributes, in the analog form, are scaled with a range of 0 and 1. There are many types of normalization that are found in the literature. The new values obtained after normalization, follow this equation:

$$\text{New value (after normalization)} = \frac{\text{current} - \text{min}}{\text{max} - \text{min}} \quad (1)$$

### 6.2. Replacement of Missing Values using Median Method

1. Find median for the Bare Nuclei (This attribute contains missing values). The median is calculated using the formula,

$$\text{Median} = \text{sizeof} \frac{(N + 1)}{2} \quad (2)$$

2. All the missing value of this attribute replaced by this median value.

### 6.3. Performance analysis

As far as the classification performance of the model is concerned, the classification rate (C) denotes the percentage of correctly classified samples, which is computed by the following formula.

$$\text{accuracy} = \frac{n_c}{n_t} * 100\% \quad n_c \leq n_t \quad (3)$$

where,  $n_c$ ,  $n_t$  represent the number of correctly classified samples and the total number of the samples, respectively.

### 6.4. Accuracy Comparison on Breast Cancer Dataset

This analysis is carried out to compare the results of two different migration topologies. To do this, the learning patterns for the proposed system is compared using breast cancer medical dataset. The comparative correct classification percentage for breast cancer dataset is shown in Table 2. It can be obviously seen that the recognition rate in terms of the right classification percentage has distinctly increased, which is measured by our IDE model. The experiments show that torus migration topology is more require computing time than random topology.

Table 2. Comparison between migration topology

Datasets	Torus Topology		Random Topology	
	Accuracy (%)	Time (sec)	Accuracy (%)	Time (sec)
Breast Cancer	99.97	16	99.97	3

## 7. EXPERIMENTAL RESULTS

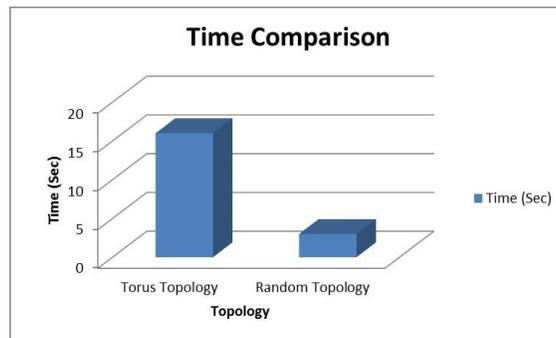


Fig 2. Time Comparison between two different migration topologies

This analysis is carried out to compare the results of two different migration topologies. To do this, the learning patterns for the proposed system are compared using breast cancer medical dataset. The time comparison between two different migration topologies on breast cancer datasets is presented in Fig 2. For torus topology, the computing time is more require than random topology. Random topology is more suitable than torus topology for the classification of breast cancer dataset.

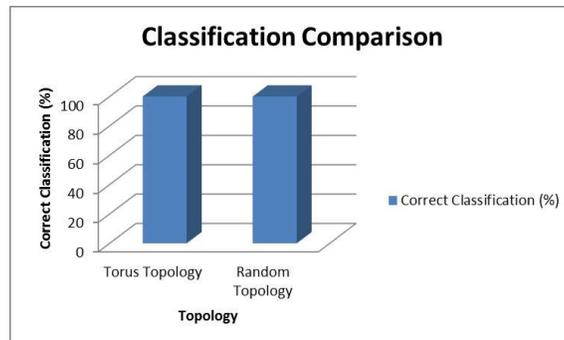


Fig 3. Comparative of correct classification percentage of migration topology

In this paper, we have introduced the analysis of two different migration topologies with random-random migration policy and compared their results. Fig 3 shows the accuracy comparison between two different migration topologies on breast cancer medical dataset. Breast cancer medical dataset is used to implement this proposed system which shows better experiments with

higher accuracy. The proposed system also reduces the computing time. For breast cancer medical dataset, the experiments show that random migration topology has better results on convergence time and correct classification percentage than torus migration topology. The proposed algorithm converges in a short time with high correct classification percentage.

## 8. CONCLUSIONS

In this research a feed forward neural network is constructed and island differential evolution propagation algorithm is used to train the network. The proposed algorithm is tested on a real life problem, the Wisconsin Breast Cancer Diagnosis problem. The objective of this study is to create an effective tool for building neural models to help us making a proper classification of various classes of breast cancer. The interest in neural networks is justified by their own properties: learning ability, generalization and reminiscence. Island differential evolution neural network approach drove by the learning algorithm works well, in terms of accuracy, efficiency and reliability. Using this model, an automated classification of various types of breast cancer was performed by avoiding the question of the expert concerning the recognition of cancer required, improving the identification of breast cancer classification. Through the results analysis it was found that the subpopulation model reduces significantly the computing time and the solutions quality is improved significantly. In this paper, we have proposed two different migration topologies with random-random policy and compared their results. According to experiments, the torus topology is more require long training time than random topology but it solution quality results are similar to random topology.

## REFERENCES

- [1] Arun George Eapen, master thesis, Application of Data Mining in Medical Applications, Waterloo, Ontario, Canada, 2004.
- [2] Choi J.P., Han T.H. and Park R.W., —A Hybrid Bayesian Network Model for Predicting Breast Cancer Prognosis, J Korean Soc Med Inform, pp. 49-57, 2009.
- [3] Hassanien Ella Aboul and Ali H.M. Jafar, —Rough set approach for generation of classification rules of Breast Cancer data, J Journal Informatica, vol. 15, pp. 23–38, 2004.
- [4] Jamarani S. M. h., Behnam H. and Rezairad G. A., —Multiwavelet Based Neural Network for Breast Cancer Diagnosis, GVIP 05 Conference, pp. 19-21, 2005.
- [5] A. Cichocki and R. Unbehauen, — ‘ Neural Networks for optimization and signal processing,’ J. Wiley, Sons Ltd. And B.G. Teubner, Stuttgart, 1993.
- [6] Abdelaal Ahmed, Mohamed Medhat and Farouq Wael Muhamed, —Using data mining for assessing diagnosis of breast cancer, in Proc. International multiconference on computer science and information Technology, pp. 11 -17, 2010.
- [7] Bellaachia Abdelghani and Erhan Guven, "Predicting Breast Cancer Survivability using Data Mining Techniques," Ninth Workshop on Mining Scientific and Engineering Datasets in conjunction with the Sixth SIAM International Conference on Data Mining, 2006.
- [8] Burke H. B. Et al, —Artificial Neural Networks Improve the Accuracy of Cancer Survival Prediction, Cancer, vol.79, pp.857-862, 1997
- [9] R. Storn and K. Price, —Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces, J. Glob. Optim., vol. 11, no. 4, pp. 341–359, Dec. 1997.
- [10] K. V. Price, R. M. Storn, and J. A. Lampinen, Differential Evolution—A Practical Approach to Global Optimization. Berlin, Germany: SpringerVerlag, 2005.
- [11] Z. Skolicki and K. De Jong. The influence of migration sizes and intervals on island models. In GECCO'05: Proceedings of the 2005 conference on Genetic and evolutionary computation, pages 1295-1302, New York, NY, USA, ACM, 2005.
- [12] Breast Cancer Wisconsin Data [online]. Available: <http://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/breast-cancer-wisconsin.data>.

- [13] Dr. K. U. Rani, —Parallel Approach for Diagnosis of Breast Cancer using Neural Network Technique| Int. J. of Computer. Application, vol. 10, no. 3, pp. 0975 – 8887, Nov. 2010.
- [14] Abdul Sttar Ismail Wdaa — Differential evolution for neural networks learning enhancement | J. of university of anbar for pure science: Vol.5: No.2: 2011.
- [15] Jun Zhang MS, Haobo Ma Md MS, —An Implementation of Guildford Cytological Grading System to diagnose Breast Cancer Using Naïve Bayesian Classifier|, MEDINFO, M.Fieschi et al. (Eds), Amsterdam: IOS Press, 2004.
- [16] Kamruzzaman S.M. and Monirul Islam Md, —Extraction of Symbolic Rules from Artificial Neural Networks| Proceedings of world Academy of science, Engineering and Technology, vol. 10, ISSN 1307-6884, Dec. 2005.
- [17] Punitha A., Sumathi C.P. and Santhanam T., —A Combination of Genetic Algorithm and ART Neural Network for Breast Cancer Diagnosis| Asian Journal of Information Technology 6 (1):112-117, 2007, Medwell Journals, 2007.
- [18] Rudy Setiono and Huan Liu, —Neural-Network Feature Selector| IEEE Transactions On Neural Networks, vol. 8, No. 3, pg 664-662, May 1997.
- [19] Wlodzislaw Duch and Rafal Adamczak and Krzysztof Grabczewski, —A New methodology of Extraction, Optimization and Application of Crisp and Fuzzy Logic Rules| IEEE Transactions On Neural Networks, vol. 12, No. 2, pp. 227-306, March 2001.
- [20] Anupam Shukla, Ritu Tiwari and Prabhdeep Kaur, —Knowledge Based Approach for Diagnosis of Breast Cancer| IEEE International Advance Computing Conference, Patiala, India, pp. 6-12, March 2009.
- [21] Esugasini S, Mohd Yusoff Mashor, Nor Ashidi Mat Isa and Nor Hayati Othman, Performance Comparison for MLP Networks Using Various Back Propagation Algorithms for Breast Cancer Diagnosis, Knowledge-Based Intelligent Information and Engineering Systems, Lecture Notes in Computer Science, 3682, pp. 123-130, 2005.
- [22] Paulin F. and Santhakumaran A., —Extracting Rules from Feed Forward Neural Networks for Diagnosing Breast Cancer| CiiT International Journal of Artificial Intelligent Systems and Machine Learning, vol. 1, No. 4, pp. 143-146, July 2009.
- [23] D. Karaboga and S. Okdem, —A simple and global optimization algorithm for engineering problems: Differential evolution algorithm,| Turkish J. Elect. Eng. Comput. Sci., vol. 12, no. 1, pp. 53–60, 2004.