International Journal of Informatics and Applied Mathematics
e-ISSN:2667-6990 Vol. 1, No. 1, 24-34

An Optimized RBF-Neural Network for Breast Cancer Classification

Siouda Roguia and Nemissi Mohamed

LabSTIC, 8 Mai 1945 Guelma University, POB 401 Guelma, Algeria
(siouda.roguia, nemissi.mohamed)@univ-guelma.dz

Abstract. This paper introduces an optimized RBF-Neural Network for breast cancer classification. The study is based on the optimization of the network through three learning phases. In the first phase, K-means clustering method is used to define RBFs centers. In the second phase, Particle Swarm Optimization is used to optimize RBFs widths. In this phase, a pseudo inverse solution is used to calculate the output weights. Finally, in the third phase, the back-propagation algorithm is used for fine-tuning the obtained parameters, namely centers, widths and output weights. The back-propagation is then initialized with the obtained parameters instead of a random initialization. To evaluate the performance of the proposed method, tests were performed using the Wisconsin Diagnostic Breast Cancer database. The proposed system was compared with a network trained only with BP and a network trained with K-means + PSO. The results obtained are promising compared to other advanced methods and the proposed learning method gives better results by combining these three methods.

Keywords: Radial Basis Function Networks, Classification, Neural Networks, Particle Swarm Optimization, K-means

1 Introduction

Breast cancer is the most common type of cancer worldwide and one of the leading causes of death among women [4]. Fortunately, an early detection effectively prevents this kind of cancer. The diagnostic is generally based on the expert's examination of the patient's data. The computer-aided diagnosis of patients not only provide fast diagnostic, but also the possibility to deal with large amount of data. Indeed, the use of artificial intelligence and pattern recognition techniques permits exploring information from large number of patients, hospitals and countries to construct powerful classification systems. The current increase in the available data make these systems more attractive. These techniques include a lot of machine learning methods such as support vector machine, naive bayesian, neural networks, fuzzy logic.

This study focuses on neural networks, which are among the most common in this field. They have been successfully applied due to their high computing and learning abilities. More specifically, RBF neural networks represent the core of this research work. They provide high global approximation capability based on local responses using special hidden neurons. Therefore, an RBF neural based classifier for breast cancer diagnosis is introduced. The learning of the proposed system is performed in three phases. First, the centers are defined using the K-means clustering algorithm. Second, the widths of the radial basis functions are optimized using particle swarm optimization, in which the calculation of the output weights is performed using pseudo inverse solution. Third, the obtained parameters are fine-tuned using back propagation (BP) algorithm. To evaluate the proposed learning method, the proposed classifier is compared to a network formed only with BP and a network formed with K-means + PSO. As a result, the proposed learning method gives better results than these two methods.

The reminder of the paper is organized as follows. Section 2 presents some related works. Section 3 describes some theoretical background, such as the RBF neural networks, the back propagation algorithm and the particle swarm optimization method. Then, Section 4, introduces the proposed classification system, while Section 5 discusses the classification results obtained on Wisconsin Diagnostic Breast Cancer database. Finally, Section 6 concludes the paper.

2 Related Work

Recently, several neural networks-based medical classification systems have been proposed, using RBF neural networks. The RBF neural networks learning in these systems are generally based on bioinspired methods. For example, in [19], the authors have proposed a system to predict breast cancer using a new graph based on feature selection approach. The classifier model used in this study is RBF neural network that its parameters were optimized using artificial bee colony (ABC) algorithm.

In [10], the authors have explored the features reduction properties of independent component analysis (ICA) on breast cancer decision support system. The aim of the features set reduction by (ICA) was to evaluate diagnostic accuracy of the classifiers like artificial neural network, k-nearest neighbor (KNN), radial basis function neural network (RBFNN) and support vector machine (SVM). In [18], the authors have proposed a novel learning approach for radial basis function neural network (RBFNN) to predict breast cancer. This approach was based on fuzzy c-means (FCM) and quantum particle swarm optimization (QPSO) for grouping similar data.

Furthermore, the authors in [2] have presented a new classification model of diabetic patient's data based on radial basis neural network, cluster validity index and BAT optimization algorithm. In this classifier the authors have employed the cluster validity index in class by class fashion for determining the optimal number of neurons in pattern layer. Also, they have designed a new convex fitness function for bat inspired optimization algorithm to identify the weights between summation layer and pattern layer.

Moreover, in [3], the authors have used a bee-inspired approach named cOptBees, to automatically define the number, location and dispersion of basis functions to be used in RBF neural networks. The proposed approach, named BeeRBF, was used to solve medical classification problems. In [1], the authors have proposed a new method to optimize the construction of radial basis function (RBF) networks, based on a cooperative particle swarm optimization (CPSO) framework. In addition, [15] presents a combination between multidimensional particle swarm optimization and class-specific clustering to optimize the number of cluster centroids and their locations.

This work proposes to train the introduced RBF neural network classifier by combining three methods: a clustering method (K-means), a bio-inspired optimization method (PSO) and the backpropagation method.

The clustering method was used to define the RBFs center, while the optimization method is used for the RBFs width. Finally, the back propagation is used for fine-tuning all parameters: RBFs centers, widths and the output weights.

Therefore, this approach aims to enhance the performance of RBF neural network by defining RBFs parameters in two independents phases, and then, fine-tuning the network parameters using BP algorithm.

3 Theoretical Background

3.1 Radial Basis Function Neuronal Network (RBFNNs)

RBF networks are a special type of artificial neural networks that use radial basis function (RBF) as activation function instead of sigmoid functions [14]. An RBFNN has three layers, namely the input layer, the hidden layer and the output layer. Figure 1 illustrates an example of RBFNNs.

The data propagation principle in RBF networks is as follows: Firstly, the data is transmitted by the neurons of the input layer to the hidden layer. Then, the neurons of the hidden layer receive the data, calculate the basis functions and

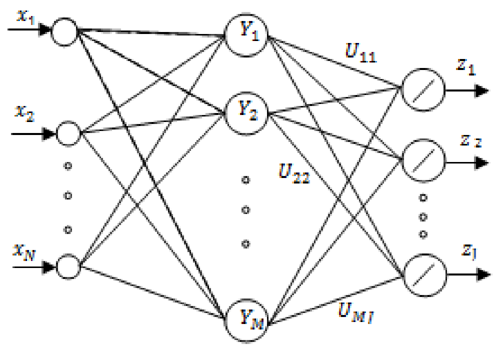

Fig. 1: An example of RBF neural network with N-M-J architecture

transmit them to the output layer. Finally, the neuron of the output layer calculates the linear sum of the hidden neurons [14]. In this work, the Gaussian functions used in the hidden layer are given by:

$$Y_m(x) = exp\left(-\frac{\parallel X - V_m \parallel^2}{2\sigma_m^2}\right) \tag{1}$$

Where, $X = (x_1, x_2 ... \, x_N)$ is the input vector, $V_m = (v_{1m}, v_{2m} ... \, v_{NM})$ and σ_m are the m^{th} center vector and the width parameter respectively.
The output of the output layer Z is a linear weighted combination :

$$Z_j = \sum_{m=1}^{M} U_{mj} Y_m(x) \tag{2}$$

Where, Y_m is the m^{th} output of the hidden layer and U_{mj} is the weight between m^{th} hidden unit and j^{th} output unit.
N: represents the number of nodes in the input layer
M: represents the number of nodes in the hidden layer
J: represents the number of nodes in the output layer

3.2 Back Propagation (BP) learning algorithm

Back propagation (BP) is a learning algorithm that aims to minimize the distance between the calculated output $Z^{(q)}$ and the desired output $T^{(q)}$ corresponding to each learning example $X^{(q)}$ [16].
The objective function used in BP algorithm in general is the total sum square error. For a set of Q learning examples, the error is defined as follows:

$$E = \sum_{q=1}^{Q} \sum_{j=1}^{J} \left(T_j^{(q)} - Z_j^{(q)}\right) \tag{3}$$

In RBFNN, the back propagation is used to adjust the weights between the hidden and the output layer and the parameters of Gaussian function.
The update of the weight U, the center V and the width σ are given by:

$$\Delta U = -\eta_1 \frac{\partial E}{\partial U} \tag{4}$$

$$\Delta V = -\eta_2 \frac{\partial E}{\partial V} \tag{5}$$

$$\Delta \sigma = -\eta_3 \frac{\partial E}{\partial \sigma} \tag{6}$$

Where, η_1,η_2,η_3 are the learning rates.

3.3 Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO) is a stochastic optimization technique and introduced by Dr. Eberhart and Dr. Kennedy in 1995. PSO can simulate the swarm behavior of bird flocking or fish schooling.
The technique starts with a random population of solutions and search for the optimal solution through some update iterations. Each individual in the swarm is called *particle* which represents a potential solution.
These particles fly with a certain velocity and find the global best position after some iterations. At every iteration, each particle adjusts its velocity according to its momentum and the influence of its best positionP_b as well as the best position of its neighbors P_g, and then a new position is obtained [5][7]. For example, if the dimension of search space is D, the total number of particles is n.
Therefore, the position of the i^{th} particle can be expressed as a vector:

$$X_i = (x_{i1}, x_{i2} ... \ x_{iD}) \tag{7}$$

The best position of the i^{th} particle so far is denoted by:

$$P_{ib} = (p_{i1}, p_{iD} ... \ p_{iD}) \tag{8}$$

The best position of all particles searching until now is denoted as a vector:

$$P_g = (p_{g1}, p_{g2} ... \ p_{gD}) \tag{9}$$

The velocity of the i^{th} particle is represented as a vector:

$$V_i = (v_{i1}, v_{iD} ... \ v_{iD}) \tag{10}$$

The velocity update is described as:

$$V_i(t+1) = w \times V_i(t) + c1 \times rand() \times (P_i - X_i(t)) + c2 \times rand() \times (P_g - X_i(t)) \tag{11}$$

The position update is described as:

$$X_i(t+1) = X_i(t) + V(t+1) \tag{12}$$

Where, *c1, c2* are the acceleration constants with positive values, *rand()* is a random number between 0 and 1 and w is new inertial weight [5][7]. The parameter w can be gradually reduced as the generation increases according to:

$$w(t) = w_{max} - \left(\frac{w_{max} - w_{min}}{itermax}\right) \tag{13}$$

4 Proposed Method

This work presents a neural-classifier for breast cancer based on RBFNN, back propagation and particle swarm optimization. The main objective is to combine the properties of these methods to achieve better performance. Indeed, the back propagation algorithm has been used many times for neural networks training with successful results, but it suffers from two main drawbacks: the local minima and slow convergence.

The proposed method aims to avoid these two problems by applying the BP

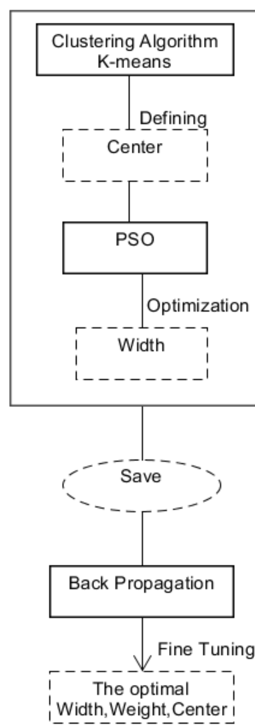

Fig. 2: Principal scheme of the proposed method

algorithm after defining the centers and optimal widths. BP is then applied just for fine-tuning the obtained parameters after clustering and optimization. More precisely, firstly, K-means clustering method is used to define the RBFs centres. Then, we use PSO for optimizing the widths of RBFs. In this step, the pseudo inverse solution is used to calculate the output weights in a similar way as the Extreme Learning Machine. Therefore, the advantages of this training algorithm are: (i) it has a rapid response time, (ii) it can be used with non-differential activation functions and (iii) it does not need to set the stopping criteria and

the learning rate [6].
Indeed, training the output weights of a neural network is much simpler than training all weights [11]. Afterwards, all parameters are saved and fine-tuned using BP. As a result, the proposed method consists then of three learning phases. Figure 2 illustrates the principal scheme of this method.

5 Tests and Experiments

To evaluate the performance of the proposed classifier, the Wisconsin Breast Cancer Dataset (WBCD) was used during this study's tests [8]. The data was gathered by Dr. William H. Wolberg at the University of Wisconsin Hospitals, Madison, USA. This dataset contains 699 records, among them, 16 are with missing data. These records were deleted to be consistent with the literature.
The selected dataset then included 683 records, of which 239 were malignant

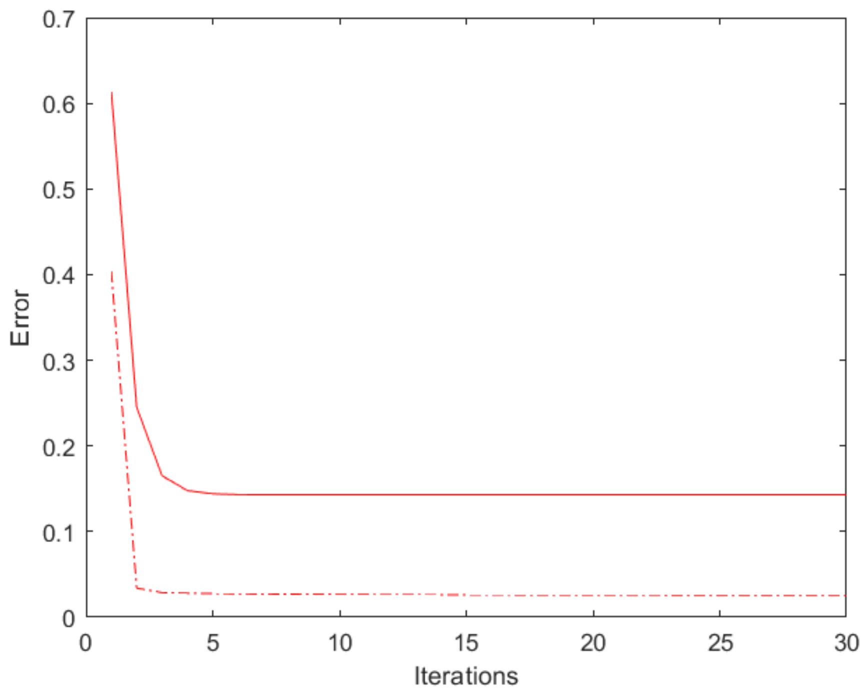

Fig. 3: Error evolution during training

and 444 were benign.
Nine integer features characterized each record: clump thickness, uniformity of cell size, uniformity of cell shape, marginal adhesion, single epithelial cell size,

bare nuclei, bland chromatin, normal nucleoli and mitoses.
A 10-fold cross validation was utilized to evaluate the generalization performance. According to this strategy, the dataset was randomly divided into 10 subsets, 9 of them were used as training data and the remaining subset was used as test set. This procedure was repeated 10 times and consequently, all samples appeared once in a test set.
Two tests were carried out to perceive the effect of the adaptation of the different parameters, when using BP only; in the first, only the output weights were updated, whereas, in the second, all the parameters were updated. Figure 3 illustrates the evolution of the error during the learning process corresponding to these two tests. Thus, the update of all parameters improves learning performance. In addition, Figure 4 illustrates the error evolution during the

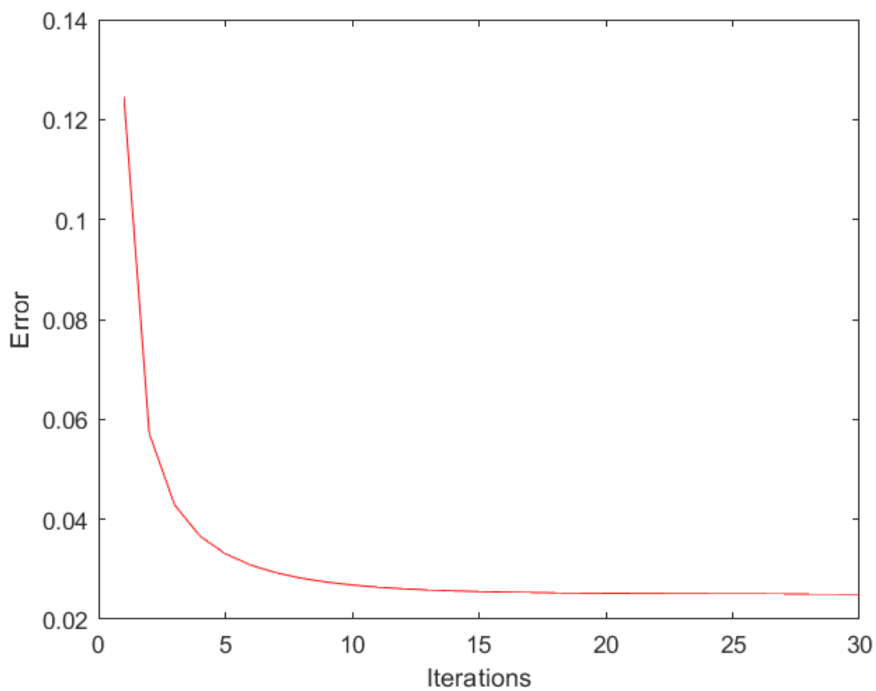

Fig. 4: Fine-tuning all parameters

learning process of an RBFNN, learning BP algorithm after the optimization phase. Therefore, the initial parameters are not randomly initialized, but taken form the best solution of the PSO.
It can be noticed that using BP to update all parameters after optimization further improves learning performance. In fact, BP algorithm starts from a good

point and its role is simply to fine-tune the parameters obtained (centres, widths and output weights).

Table 1: Classification results over Breast Cancer Dataset

Hidden nodes	BP	PSO + Kmeans	The proposed method
5	96.86 ± 0.32	95.92 ± 1.14	97.08 ± 0.13
10	96.80 ± 0.18	96.82 ± 0.85	97.38 ± 0.18
15	96.70 ± 0.36	97.02 ± 0.45	97.48 ± 0.29
20	96.60 ± 0.43	97.32± 0.33	97.62 ± 0.31
25	96.55 ± 0.50	97.52 ± 0.23	97.82 ± 0.36

Table 2: Comparaison of the results

Authors, Year	Method	Accuracy (CV)
Örkcü, 2011 [13]	Real Coded G.A+Neural Networks	96.50 % [10-CV]
Stoean , 2013 [17]	SVM+Evolutionary Algorithm	97.07^a%
Malmir, 2013 [9]	Imperialist Competitive Alg+NN	97.75^a%
Nguyen, 2015 [12]	Wavelet+Type2 fuzzy logic	97.88 % [5-CV]
Proposed,2018	PSO+Kmeans+BP	97.82 % [10-CV]

CV: Cross Validation, a: Cross Validation not mentioned

Table 1 displays the classification results obtained using three networks. The first network was trained using BP only. In this network, the centers and widths of the RBFs (hidden neurons) were initialized randomly. In the second network (Kmeans + PSO), the centers of RBFS were defined using Kmeans, and the widths were obtained using PSO.

Finally, the third network, which represents the proposed contribution, combined both approaches. In this network, K-means defines the centers, PSO optimizes the widths and BP fine-tunes the obtained parameters. The obtained results with different numbers of RBFs showed that the proposed classifier outperforms the other networks. Table 2 compares the proposed method with other works applied on the same database. These works include a variety of artificial intelligence systems, i.e., fuzzy logic, neural networks, clustering, wavelet, GA, SVM, ELM. It can be noticed that the classification system proposed by the authors surpasses most of these works. Thus, the results obtained are promising for further enhancements.

6 Conclusion

This paper introduced a neural classier for breast cancer diagnosis. The proposed system is based on Radial Basis Function neural network, stochastic optimization technique and clustering algorithm. The objective is to combine the properties of these methods. Indeed, the back propagation algorithm allows to obtain good performances but it suffers from two main drawbacks: local minima and slow convergence.

Therefore, the proposed method aims to avoid these two problems by applying the BP algorithm after defining the centers and optimal widths. BP is then applied just for fine-tuning the network parameters. To assess the performance of the proposed method, tests on Wisconsin Breast Cancer Dataset (WBCD) were carried out. The proposed classifier provided higher generalization performance compared to BP using only and PSO + K-means, which shows that the use of BP for fine-tuning enhances network performance.

References

1. Alexandridis, A., Chondrodima, E., Sarimveis, H.: Cooperative learning for radial basis function networks using particle swarm optimization. Applied Soft Computing **49**, 485–497 (2016)
2. Cheruku, R., Edla, D.R., Kuppili, V.: Diabetes classification using radial basis function network by combining cluster validity index and bat optimization with novel fitness function. International Journal of Computational Intelligence Systems **10**(1), 247–265 (2017)
3. Cruz, D.P.F., Maia, R.D., da Silva, L.A., de Castro, L.N.: Beerbf: a bee-inspired data clustering approach to design rbf neural network classifiers. Neurocomputing **172**, 427–437 (2016)
4. DeSantis, C., Ma, J., Sauer Goding, A., Newman, L., Jemal, A.: Breast cancer statistics, 2017, racial disparity in mortality by state. Journal of cancer research and therapeutics **67**(6), 439–448 (2017)
5. Eberhart, R., Kennedy, J.: A new optimizer using particle swarm theory. In: Micro Machine and Human Science, 1995. MHS'95., Proceedings of the Sixth International Symposium on. pp. 39–43. IEEE (1995)
6. Huang, G.B.: An insight into extreme learning machines: random neurons, random features and kernels. Cognitive Computation **6**(3), 376–390 (2014)
7. Kennedy, J., Eberhart, R.: Particle swarm optimization, proceedings of ieee international conference on neural networks (icnn'95) in (1995)

8. Lichman, M.: Uci machine learning repository. irvine, ca: University of california, school of information and computer science. URL http://archive. ics. uci. edu/ml (2013)
9. Malmir, H., Farokhi, F., Sabbaghi-Nadooshan, R.: Optimization of data mining with evolutionary algorithms for cloud computing application. In: ICCKE 2013. pp. 343–347. IEEE (2013)
10. Mert, A., Kılıç, N., Bilgili, E., Akan, A.: Breast cancer detection with reduced feature set. Computational and mathematical methods in medicine **2015** (2015)
11. Nemissi, M., Seridi, H., Akdag, H.: One-against-all and one-against-one based neuro-fuzzy classifiers. Journal of Intelligent & Fuzzy Systems **26**(6), 2661–2670 (2014)
12. Nguyen, T., Khosravi, A., Creighton, D., Nahavandi, S.: Medical data classification using interval type-2 fuzzy logic system and wavelets. Applied Soft Computing **30**, 812–822 (2015)
13. Örkcü, H.H., Bal, H.: Comparing performances of backpropagation and genetic algorithms in the data classification. Expert systems with applications **38**(4), 3703–3709 (2011)
14. Qasem, S., Shamsuddin, S.: Generalization improvement of radial basis function network based on multi-objective particle swarm optimization. J. Artif. Intell **3**(1), 1–16 (2010)
15. Raitoharju, J., Kiranyaz, S., Gabbouj, M.: Training radial basis function neural networks for classification via class-specific clustering. IEEE transactions on neural networks and learning systems **27**(12), 2458–2471 (2016)
16. Rumelhart, D., Hinton, G., Williams, R.: Learning internal representation by error propagation, parallel distributed processing: exploration in the microstructure of cognition. MIT press Cambridge pp. 318–362 (1986)
17. Stoean, R., Stoean, C.: Modeling medical decision making by support vector machines, explaining by rules of evolutionary algorithms with feature selection. Expert Systems with Applications **40**(7), 2677–2686 (2013)
18. Venkatesan, A.S., Parthiban, L.: A novel optimized adaptive learning approach of rbf on biomedical data sets. Research Journal of Applied Sciences, Engineering and Technology **8**(4), 541–547 (2014)
19. Zarbakhsh, P., Addeh, A., et al.: Breast cancer tumor type recognition using graph feature selection technique and radial basis function neural network with optimal structure. Journal of cancer research and therapeutics **14**(3), 625 (2018)