RBF neural network based on clonal selection algorithm for medical data diagnosis

Aye Mya Thandar, May Aye Khine
University of Computer Studies, Yangon, Myanmar
ayemyathandar7@gmail.com

Abstract

In artificial neural networks, the parameters may include the number of layers, the number of hidden units, the activation function and the algorithm parameters such as learning rate for optimization. Many researchers have proven that the training of artificial neural networks is a complex process and methods of training are highly varied. Some attempt to approximate the process of biological neurons but many diverge greatly from them in an attempt to find more computationally efficient methods to achieve optimal or near-optimal weights. Although radial-basis function networks (RBF) are well known for requiring short training period among artificial neural networks, these methods perform a local search and they can easily fall in local minima by producing sub-optimal solutions. Therefore, the performance of network training is not good and the accuracy is low for RBF neural networks. The traditional network weight training generally uses gradient descent method and it can not get the global optimum. Training the weights by optimization method can find the weight set that approaches global optimum while do not need to compute gradient information and it can help to reduce error rate in network training .Clonal selection algorithm is a global search among optimization method and it can provide an efficient alternative for the optimization of neural networks. In this paper, we use clonal selection algorithm to adjust weight units which are important to improve network training in RBF neural network.

Keywords: RBF neural network, Clonal selection algorithm, K-means clustering algorithm

1. Introduction

Artificial neural networks (ANNs) store all information in the form of weights that enables them to solve problems. The performance of the ANN depends upon the choice of weight which is usually set by a training algorithm. The training algorithm tries to find an optimal point in the weight space or the error space of the problem such that the errors are minimized. Each model of ANN uses its own training algorithm. These algorithms are hence different for Multi Layer Perception, Radial Basis Function Networks, Learning Vector Quantization, recurrent Neural Networks, Self Organizing Maps, etc. From a design perspective, an ANN can be characterized by three main features: (1) a set of artificial neurons, also termed nodes, units or simply neurons ;(2) the pattern of connectivity among neurons, called the network architecture or structure; and (3) a method to determine the weight values, called its training or learning algorithm. Although there are several neuron models, network architectures and learning algorithms, this paper presents about RBF neural networks.

RBF networks are known for requiring a much shorter training period. The centers in RBF networks are determined with reference to the distribution of the input data, but there are no references to the prediction task. As a result, representational resources may be wasted on areas of the input space that are irrelevant to the learning task. Therefore, RBF networks have the disadvantage of requiring good coverage of the input space by radial basis functions. In neural networks, the computation of the neuron will be different depend on the weights. By adjusting the weights of an artificial neuron, we can obtain the output we want for specific inputs. But, when we have an ANN of hundreds or thousands of neurons, it would be quite complicated to find by hand all the necessary weights. Therefore, we choose clonal selection algorithm to adjust the weights of the RBF in order to obtain the desired output from the network training. Clonal selection algorithm is one of the optimization methods and it can find the best solutions for complex problems by searching or evaluation.

2. Related Work

Many researchers tried to improve the performance of Artificial Neural networks by using many optimization methods .In 2007, Billings Sa, Wei Hl and Balikhin Ma [5] constructed flexible multiscale radial basis function networks (RBF). They used k-means clustering algorithm and an improved orthogonal least squares (OLS) algorithm to determine the unknown parameters in the network model including the centers, widths and weights between the basis functions.

Selection weight is a key issue in the use of neural networks. Divid J.Montana [7] investigated the utilization of genetic algorithms for neural network weight selection. If all of the training samples are selected as hidden centers, the generalization capability of the network will become very poor so that many noised or deformed samples will not be able to be recognized, although the network is guaranteed to converge to some satisfying solution. Therefore, Zhong-Qiu Zhao and De-Shuang Huang [14] proposed hybrid learning of genetic algorithm to improve generalization capability of radial basis function neural networks in 2007. Shifei Ding, Xinzheng Xu and Hong Zhu [10] studied Artificial Neural Networks based on Genetic algorithm. They optimized both the connection weights and architecture of the neural network by using genetic algorithm in 2011. At the same time, Shifei Ding, Gang Ma and Xinzheng Xu [11] used genetic algorithm to optimize the centers, the widths and the weights between the hidden layer and the output layer of RBF neural networks in 2011.

Many researchers become interested in applying Clonal selection methods with many fields to improve their result more and more. In 2010, Xiang Weiguo [13] presented a weather prediction method based on artificial principles, illustrated expression method of antigen and antibody and showed higher accuracy compared with neural network. G.Tezel and U. Kose [8] proposed a system called Headache Disease Diagnosis by using the Clonal selection algorithm in 2011 .They explained the improvement and advantages of the Clonal selection algorithm with high accuracy in their datasets. Clonal selection methods are also optimization approaches and their some functions are the same as genetic algorithm. A.Lanaridis, V.karakasis [3] and A.Stafylopatis proposed "Clonal Selection-based Neural Classifier" in 2008. They used the basis concepts of clonal selection to evolve only MLP neural network, each one devoted to the recognition of a different class of the input data. They aimed to use the error rate of output layer for calculating affinity of clonal selection and criteria for better performance. Using optimization to change weight value in hidden layer affect the accuracy and time of network result. Our paper is different from others because functions of the clonal selection are used to adjust weight calculation in network training in order to improve the weakness of RBF neural networks.

3. Artificial neural networks

Artificial neural networks consist of three groups or layers of units: a layer of "input" units is connected to a layer of "hidden" units, which is connected to a layer of "output" units. The activity of the input units represents the raw information that is fed into the network. The process of each hidden unit is determined by the activities of the input units and the weights on the connections between the input and the hidden units. The behavior of the output units depends on the activity of the hidden units and the weights between the hidden and output units. Among artificial neural networks, RBF networks are good at modeling nonlinear data and can be trained in one stage rather than using an iterative process as in other neural networks.

3.1 RBF Neural Networks

RBF neural networks address the problem of curve-fitting, which is approximation in high-dimensional spaces. The important point emerges from Cover's theorem is to transform from difficult nonlinear to linear problems. Learning in this case is equivalent to finding an interpolating surface in the multidimensional space that provides a best fit to the training data, measured by preselected statistical criteria. There are three layers called input, hidden/kernel layer and output node. Each hidden unit represents a single radial basis function, with associated center position and width. Such hidden units are sometimes referred to as centroids or kernels. Each output unit performs a weighted summation of the hidden units, using the w_j as weights. The training is typically done in two phases first fixing the width and centers and then the weights. We use Gaussian function as radial basis function.

$$y(\mathbf{x}) = \sum_{i=1}^{M} w_i \exp\left(-\frac{(||\mathbf{x} - \mathbf{c}_i||)^2}{2\sigma^2}\right) \qquad (1)$$

Where: we denote the input as x and the output as y(x). And, c_i are called centers and σ is called the width. There are M basis functions centered at c_i.w_i are called weights.

3.1.1 The centers and the widths of RBF network

The easiest way to determine the centers of the hidden layer neuron is to select the position of the center from the training dataset randomly. But, this would be impractical if the number of training data is large. Therefore, it is usual to construct the neurons using K-means clustering algorithm. To determine the K, the number of centers, a measure for clustering performance is required.

The widths should be chosen so that the input space is fully covered by the kernel functions as possible. If the distances between centers are not equal, it may be necessary for each hidden layer neuron to have its own value of σi. The width of a hidden layer neuron is set to twice the distance between its center and the center of its nearest neighbor.

3.1.2 The weights of RBF network

The initial weights can be found by solving the following linear matrix equation:

$$\mathbf{w} = (\mathbf{\Phi}^T\mathbf{\Phi})^{-1} \mathbf{\Phi}^T \mathbf{y} \qquad (2)$$

Where: **w** is the weight vector; **y** is the response vector (actual output for training datasets).

$$\varphi_{ji} = \varphi\ (\| \mathbf{x}_j - c_i \|),\ j,\ i = 1, 2, ..., N \qquad (3)$$

$$\mathbf{\Phi} = \{\varphi_{ij}\ |i,j = 1,2,...,N\}$$

The matrix $\mathbf{\Phi}$ is called the interpolation matrix denoted by an N-by-N matrix with element φ_{ji}.

$$\mathbf{e} = \sum_{i=1}^{N} (\mathbf{y_i} - \mathbf{x_{ij}})^2 \qquad (4)$$

x_{ij} is the desired output for network and y_i is the actual output of the network pattern. The least square error is denoted by e and it is used as evaluation criterion (affinity) for the better choice of weight units to reduce error rate in network training.

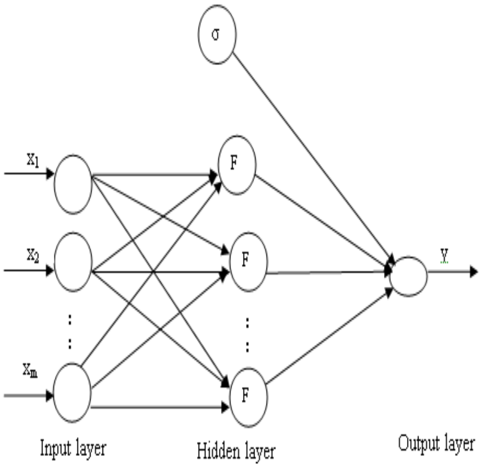

Figure1. RBF neural network architecture

4. Clonal Selection algorithm (CSA)

The clonal selection principle describes the basic features of an immune response to an antigenic stimulus. It establishes the idea that only those cells that recognize the antigen proliferate, thus being selected against those that do not. In CSA, a candidate solution for the specific problem is called an antigen, which is recognized by the antibody. Each antibody represents a point in the search space, i.e., a possible solution to the problem. A population consists of a finite number of antibodies.

Every antibody is evaluated by the evaluation mechanism to obtain its affinity. Based on this affinity and undergoing immune operators, a new population is generated iteratively with each successive population, referred to as a generation. The higher the antigenic affinity, the higher the number of clones generated for each antibody. The CSA use three immune operators, i.e.; cloning, hypermutation, and receptor editing to refresh the composition of populations. In the hypermutation operation, the cloned population is subject to an affinity mutation process inversely proportional to the antigenic affinity. The receptor editing includes two steps. In the first step, a given number of new antibodies are generated randomly. In the second step, the generated antibodies are used to refresh the whole population by replacing those antibodies with the lowest antigenic affinity.

The CLONALG algorithm is described as follows.

Step1. The weights of the RBF neural networks are coded into real-valued antibodies and initialize the antibodies from population (P).

Step2. Evaluate the fitness function of antibody in population P. The fitness refers to the affinity measure.

Step3. Select randomly the antibodies (Pr) from population P.

Step4. Clone these antibodies into a temporary pool (C). The clone size is an increasing function of the affinity and we define clone size from minimum size 100 to maximum size 500 depending on different medical datasets.

Step5. Generate each set of mutated clones (C') by using random mutation.

Step6. Evaluate the fitness (affinity) of each of the mutated clones to its respective weight in C'.

Step7. The fitness of the mutated clone of each set is compared to the fitness of the original (parent) antibody. If the cloned antibody is better than parent in term of fitness, it replaces the original antibody (parent).

5. Clonal RBF neural network

Clonal selection algorithm is applied to improve RBF networks from network connections (weights). The weights of all neurons are adjusted to bring them closer to the input pattern. There are two types of weight adjustment called batch and stochastic weight update. Batch weight update is performed after the presentation of all the training examples and it gives more reliable gradient information than stochastic weight update. Stochastic weight update is adjusted after the presentation of each input pattern. Therefore, it requires less local storage and it is easy to implement. It can provide effective solutions to large and difficult problems. This paper uses stochastic weight update to adjust each input training pattern.

5.1 The training algorithm

The main process in this study is to employ clonal based training algorithm on weights of RBF network connections to search for optimal solution. We use Gaussian radial basis function in equation (1) to compute the desired result (x_{ij}) of the output layer. The main function of this paper is to adjust weights of each unit by using clonal selection algorithm to reduce the error.

Input: Training patterns $x=\{x_1,x_2,...,x_N\}$.
Output: weight and error of the output layer.
Var: x_{ij}: desired result of the output layer.
y_i: actual result of the output layer.

Process: Select randomly center values from the training data and fix width value (default=1);

Begin
For x=1 to N do

Calculate distance between input layer (input training dataset) and center by using Euclidean distance formula.

Calculate unknown weights between the hidden layer and output layer by solving linear matrix equation.

Compute the desired result (x_{ij}) and the error of the output layer.
End For

While stopping criteria are not fulfilled do
Adjust weights of each unit to reduce the error by using clonal selection algorithm.
Choose the optimal weights of each unit and calculate desired output result with less error than before.
End While
End

Figure2. Clonal RBF Algorithm

5.2 Optimization Weight of Clonal RBF

For optimization weight, number of iterations and error of output layer are used as criteria. The Clonal selection algorithm select, clone and mutate weight in each iteration and the best optimal weight is selected to get fewer error. It is not necessary to loop all number of iterations if any training set meets its criteria earlier. Therefore, it reduces time complexity of RBF by improving training performance.

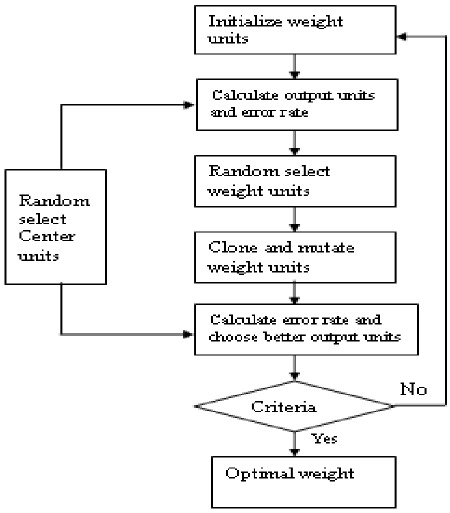

Figure3. Optimization Weight of Clonal RBF

6. Medical Datasets

We use 4 medical datasets such as breast cancer, lymph, diabetes and hepatitis from UCI repository.

6.1 Breast Cancer

The purpose of the breast cancer dataset is to classify a tumor as either benign or malignant based on cell descriptions gathered by microscopic examination. It contains 9 attributes and 699 examples of which 485 are benign examples and 241 are malignant examples.

6.2 Lymphography (lymph)

This is a small medical data set containing 148 instances with 18 nominal features. The task is to distinguish healthy patients from those with metastases or malignant lymphoma. The values for class attribute are normal find, metastases, malign lymph and fibrosis. This is the one of three medical domains (the others being Primary Tumor and Breast Cancer) provided by the University Medical Centre, Institute of Oncology, Lyubljana, Yugoslavia.

6.3 Pima Indian Diabetes Database(PIDD)

PIDD includes the following attributes (1-8 attributes as input and last attribute as target variable) number of times pregnant, Plasma glucose concentration a 2 hours in an oral glucose tolerance test, Diastolic blood pressure (mm Hg), Triceps skin fold thickness (mm), 2-Hour serum insulin (mu U/ml), Body mass index (weight in kg/(height in m)^2),Diabetes pedigree function and Age(years). Class to be predicted is patient is suffering from tested-positive or test-negative. A total of 768 cases are available in PIDD.

6.4 Hepatitis

This data set is at UCI repository database. It contains 155 training dataset and 140 testing dataset.

There are 20 features which are used as input neurons in this dataset.

7. Experiments of proposed system

We use four medical datasets to implement this proposed system which shows better experiments.

7.1 Parameters of the experiments

We use some parameter values of network structure which are the best for higher accuracy in other reference papers.

parameter	**Dataset**			
	Breast cancer	Diabetes	lymph	Hepatitis
Train data	349	460	148	155
Test data	350	308	120	140
Output neuron	1	1	1	1
Network structure	9-2-1	8-20-1	18-4-1	20-8-1

Table 1. Parameters of RBF neural network

7.2 Accuracy of training and testing RBF neural network

Medical datasets are tested with RBF neural network classifier for the initial step of this system. Below the table shows the accuracy of training and testing RBF. 10-fold cross validation is used for estimating accuracy due to its relatively low bias and variance.

Dataset	**Training Accuracy (%)**	**Testing accuracy (%)**
Hepatitis	80.6662 %	76.7132%
Diabetes	80.7366 %	79.6663%
Lymph	76.8012 %	72.1306%
Breast cancer	76.7126 %	77.3613 %

Table 2. Accuracy of training and testing RBF neural network

8. Conclusion and future work

We find that the accuracy produced by RBF is not optimal result before using clonal selection algorithm. Therefore, we are keen on improving performance of RBF by using clonal selection algorithm. The main difficulty of neural network is to adjust weight in order to reduce the error rate. In this paper, we define the parameters of RBF and present the accuracy of training and testing RBF. For further works, we have to use clonal selection method for weight calculation in neural network training. Then, we will show higher accuracy of Clonal RBF network.

References

[1]. Alireza Arab Asadi, Ali Naserasadi,Zeinab Arab Asadi, “ A New Hybrid Algorithm for Traveler Salesman Problem based on Genetic Algorithms and Artificial Neural Networks”, Internation Journal if Computer Applications,June (2011).

[2]. Anderson, J.A. An Introduction to neural networks. Prentice Hall, Chaeng, B. and Titterington, D. M. (1994). Neural networks: A review from a statistical perspective. Statistical Science, 9, 2-54. (2003).

[3]. A.Lanaridis ,V.karakasis and A.Stafylopatis, “ Clonal Selection-based Neural Classifier”, Eighth International Conference on Hybrid Intelligent Systems, 978-0-7695-3326-1/08,IEEE,(2008).

[4]AshaGowda,Karegowda,A.S.Manjunath,Ma.Jayaram, ”Application of genetic algorithm optimized neural network connection weights for medical diagnosis of PIMA Indians Diabetes”, International Journal on Soft computing(IJSC),Vol.2,No.2,May (2011),India.

[5]. Billings Sa, Wei Hl and Balikhin Ma, “Generalized multiscale radial basis function networks”, Neural Net. 2007 Dec; 20(10):1081-94. Epub (2007) Oct 16.

[6]. D. Whitley, “Genetic Algorithm and Neural Networks”, jwbook.sty v3.0, 12-1-(1995)

[7]. Divid J.Montana, "Neural Network Weight Selection using Genetic algorithms", Bolt Beranek and Newman Inc,70 Fawcett Street, Cambridge.MA 02138.(2005).

[8]. G.Tezel and U. Kose, “Headache Disease Diagnosis by Using the Clonal Selection Algorithm”, 6th International Advanced Technologies Symposium (IATS’11), 16-18, Elazig, Turkey, May (2011).

[9] Jasmina Novakovic,”RBF Network with Genetic Algorithm for feature selection”, 17th Telecommunications forum TELFOR 2009.

[10].Shifei Ding,Xinzheng Xu,Hong Zhu, "Studies on Optimization Algorithms for some Artificial Neural Networks Based on Genetic Algorithm(GA)",Journal of computers,Vol.6,No. 5,May (2011).

[11]. Shifei Ding, Gang Ma and Xinzheng Xu, “A Rough RBF Neural Networks Optimized by the Genetic Algorithm”, Advances in Information Sciences and Service Sciences, volume 3,Number 7, China. August (2011).

[12]. Saltan Noman, Siti mariyam.Shamsuddin, and Aboul Ella Hassanien,”Hybrid learning Enhancement of RBF Network with Particle Swarm Optimization”, Foundations of Comput.Intel.Vol.1,SCI 201,PP.381-397,(2009).

[13]. Xiang Weiguo, College of Atmospheric Science, Chengdu University of Information and Technology(CUIT), Chengdu 610225, China “The Weather Prediction Method Based on Artificial Immune System” , International Forum on Information Technology and Applications,(2010).

[14]. Zhong-Qiu Zhao and De-Shuang Huang, “A mended hybrid learning algorithm for radial basis function neural networks to improve generalization capability”, Applied Mathematical Modelling 31, 1271-1281, (2007).

[15]. Zhang Caiqing,Qi Ruonan,Qiu Zhiwen, “Comparing BP and RBF Neural Network for Forecasting the Resident Consumer level by Matlab”, 978-0-7695-3504-3/08,IEEE,(2008).