A Comparison of Radial Basis Function and Multilayer Perceptron Network as tool for classification of Medical Data

S. Leoni Sharmila¹, C. Dharuman², P. Venkatesan³

¹,² Department of Mathematics, SRM IST, Ramapuram, Chennai.
³ Department of Community Medicine, Sri Ramachandra University, Chennai.

sharmilamartin169@gmail.com

Abstract. Artificial Neural Network has become a popular tool in developing systems that encircles human proficiency. The importance of exact detection is exceptionally important for proper treatment and preserve of disease. Clinical cytology has improved tremendously in disease diagnosis. In this paper, two Artificial Neural Network (ANN) methods, Multilayer Perceptron (MLP) and Radial Basis Function (RBF) network are compared. The RBF network predicts comparatively better accuracy in compared to MLP methods. Also it was detected that the RBF method requires a lesser amount of time for the development of the model, this is because there is no repetition to reach the favourable parameters in the model.

1. Introduction

Artificial Intelligence (AI) provides learning ability to the computer in real world problem [1]. Machine learning is a subset of AI that was developed to mimic human brain in computer to investigate the details from the collect data. In many situations of clinical test, ML system is considered as a development to diagnosis disease and analyze that imitates like an experts knowledge in medical field. Also ML method makes the computer learn from datasets, which in turn develops a model to identify common patterns and this can able to make decisions based on availability of knowledge, more over the imperfectness medical database does not affect the classification [2]. In medical diagnosis, classification is a famous ML technique among which the most usually applied technique is Artificial Neural Networks (ANNs). The practice of ML technique in disease classification has increased very frequent [3-5].

Applications of this type have used the Multilayer Perceptron (MLP) ANN method which is combined with the Back Propagation (BP) algorithm [6,7]. However, the nonlinearity of MLP with its parameters and steepest descent method used in BP algorithm do not converge to a globally optimum set of parameters always. Several attempts of trial and error method are necessary to choose the best parameter from a set of locally optimum parameters. Another alternate method to MLP is RBF network [8,9] that consists of linear parameters and has established many applications in areas like electrical and electronic engineering. [10] have proved theoretically that this type of ANNs are applicable for global approximations and also can learn without local minima, therefore strictly convergence to optimum parameters. For complicated situations, [11] proved that the learning rate of
RBF is faster than MLP networks. This study is to apply the RBF and MLP network approach for classification of medical data.

2. RADIAL BASIS FUNCTION (RBF) TYPE NETWORK

2.1. Network Architecture and Composition
An RBF is a feed-forward two-layered neural network where the basis function transforms (i/p) (i/p) of the (h/l) layer(h/l). Responses of the (h/l) are added in linear combinations to form the (o/p) in the (o/p)layer. In the (h/l) layer, the basis functions are radially symmetric. For the performance of the network, selecting the basis function is not important but, commonly used function is the Gaussian function which is given by mean and standard deviation. A representation diagram of an RBF network with I, H and O respectively of (i/p), (h/l) and (o/p) layer nodes for the transformation of ID points of \( I^1, I^k, \ldots I^{ID} \) in the (i/p) layer to the points \( O^1, O^k, \ldots O^{ID} \) in the (o/p) layer is shown in Figure 1.

In RBF, the parameters are the centers \( C_j \) and range \( \varphi_j \) in the basis functions at the (h/l) nodes, and the synaptic weights \( w_{ij} \) of the (o/p) nodes. The RBF centers are also present in the (i/p) layer. The basis function responds to an (i/p) depends on the distance between the (i/p) \( I^k \) and the RBF center \( C_j \). If (i/p) falls in a small localized region of the basis functions' center then the (h/l) layer nodes of RBF will produce non-zero replies only [12]

2.1.1. Network functioning[12]: The connections between the (i/p) and (h/l) layers are not weighted in RBF networks. Therefore, the (i/p) reach the (h/l) layer without any change. For an (i/p) \( I^k \) the jth (h/l) node produces a response \( r_j \) given by,

\[
 r_j = \exp \left( -\frac{\| I^k - C_j \|}{2\varphi_j^2} \right)
\]  

(1)

![Fig-1 RBF- Schematic diagram](image)
where $\|I^k - C_j\|$ is the distance between the point representing the input $I^k$ and the center of the $j$th $(h/l)$ node as measured by some norm. In this study the Euclidean norm is used. The $(o/p)$ $y_{in}$ of the network at the $(o/p)$ node is given by,

$$y_{in} = \sum_{j=1}^{H} r_j w_{nj} \tag{2}$$

In some cases the number of $(h/l)$ layer nodes is equal to the number of data in the training set $(H = ID)$ and the RBF centers coincide with the $(i/p)$ $s (C_j = I^k, where, k = j = 1,2,...ID)$, the $(h/l)$ layer response according to equation (1) becomes one, $j = k$. When the basis functions are accurately localized, the response of the other $(h/l)$ layer nodes will be closer to zero. It is also be seen that equation (2) gives the exact $(o/p)$ when the $(o/p)$ layer weight is equal to the $(o/p)$. In the best case, RBF network can be made to map points in $N$-dimensional $(i/p)$ space precisely on to points in $M$-dimensional $(o/p)$ space [12]

3. RADIAL BASIS FUNCTION (RBF) TYPE NETWORK

The Type of neural network is commonly known and normally used type when there are more than one layer in artificial neural network that contains forward connections of $(i/p)$ s to $(o/p)$ s is known as Multilayer Perceptron (MLP) or Multilayered Feed-forward Neural Networks. An MLP consists of a set of $(i/p)$ terminal, an $(o/p)$ terminal, and one or more terminals of $(h/l)$ nodes between the $(i/p)$ terminals and the $(o/p)$ terminal.[13]
3.1. Back Propagation Algorithm[13]:

Learning in this network is usually done through supervised method. This environment consists of learning background that contains both the learning model and desired (o/p) model corresponding to the (i/p). Minimizing the error between the network and desired (o/p) is the main aim of this learning. This causes a back propagation all over the network due to this reason this algorithm is called back-propagation learning algorithm. The algorithm for a MLP with one (h/l) layer is given by the following [13]:

Step 1: Initializing: Initially the network weights and thresholds are distributed evenly with random small range of values.

Step 2: A new training set is create: Here the weights are adjusted that is all the gradients of the weights and the current error are initially considered as 0 ($\Delta w_{ij} = 0 \& E_r = 0$).

Step 3: The forward pass of the signal: The (h/l) layer neurons (o/p) are calculated here:

$$v_j(q) = f\left(\sum_{i=1}^{m} u_i(q).w_{ij} - \theta_j\right)$$  \hspace{1cm} (3)

Here $f$ the sigmoid activation function and $m$ is the (i/p) s number for the jth neuron from the (h/l) layer. The network (o/p) s are calculated as:

$$v_k(q) = f\left(\sum_{i=1}^{n} u_{jk}(q).w_{jk}(q) - \theta_k\right)$$  \hspace{1cm} (4)

That is $k$ belongs to the (o/p) layer and $n$ is the (i/p) s number of the neuron. The error per epoch is updated as follows:

$$E_r = E_r + \frac{(e_k(q))^2}{2}$$  \hspace{1cm} (5)

Step 4: The adjustments of the weights are done in backward propagation according to the errors: Here single-pole sigmoid is used (equation 3, is derived as:

$$f'(u) = \frac{e^{-u}}{(1 + e^{-u})^2} = f(u).\left(1 - f(u)\right)$$  \hspace{1cm} (6)

The weights present between the (h/l) layer and the (o/p) layer are updated as:

$$\Delta w_{jk}(q) = \Delta w_{jk}(q) + v_j(q).\phi_k(q)$$  \hspace{1cm} (7)

The weights between the (i/p) layer and the (h/l) layer are updated:

$$\Delta w_{ij}(q) = \Delta w_{ij}(q) + u_i(q).\phi_j(q)$$  \hspace{1cm} (8)
Step 5: A new iteration: At this stage when the test vectors are identified, immediately the training epoch will go to the 3rd step otherwise, according to the gradient weights, connection weight will be updated.

\[ w_{ij} = w_{ij} + \eta \cdot \Delta w_{ij} \quad (9) \]

where \( \eta \) represents the rate of learning. The criteria for termination is tested after an epoch is completed i.e. \( E_r < E_{\text{max}} \) or when number of training epochs reaches maximum). If no, move to 2nd step else if yes end the algorithm [13].

4. EXPERIMENTS RESULTS AND DISCUSSIONS

RBF network requires more number of neurons for training. When there is a high number of training vectors, this network performs better. Unlike other network, the entire \((i/p)\) space not react at the same time in RBF. Here, the center of the \((i/p)\) layer is calculated first, next the \((i/p)\) s which are close to this center are reacted. Because of this, these networks answer to local \((i/p)\)s quickly. The two layers of RBF are the radial basis layer and the next is linear \((o/p)\) layer. Competitive learning or clustering is done in the training process. Spread number and goal number are included in network parameters. By changing these parameters, performance of the network is improved. Distance between \((i/p)\) vectors and weight vectors are calculated by applying \((i/p)\) s to the network and multiplying the calculated values by bias values obtains the vector product. At that moment, these values generate several neurons as \((i/p)\) s by the corresponding functions and as a final point, the \((o/p)\) layer obtains the \((o/p)\) values [14]. For classification, Hepatitis database is chosen from UCI Repository. The data set consists of 155 cases with 17 variables among these 123 variables are classified as live and 32 as die. When MLP method is used were the learning algorithm is Back propagation method which is used for training the neurons and is based on descent gradient technique. Classifies an accuracy of 80%. The same data is classified using RBF network with a classification rate of 85.80% accuracy.

|                         | Accuracy Rate |
|-------------------------|---------------|
| MLP Network             | 80%           |
| RBF Network             | 85.80%        |

Table1- Classification Rate

5. CONCLUSION

In Artificial Neural Network models, Radial Basis Function type was developed to predict hepatitis diseases. When classification accuracy is considered, RBF along with k-means clustering technique performs better when compared with MLP with back propagation method. Since the RBF network models are linear in the parameters, there is an assurance in convergence to their optimum values. Finally this sums up that RBF minimizes the error in minimum trial or in other words it needs minimum time and effort.
References

[1] N. Sandhya, K.R. Charanjeet, A review on Machine Learning Techniques, International Journal on Recent and Innovation Trends in Computing and Communication, 2016.

[2] G.D. Magoulas and A. Prentza, Machine learning in medical applications. In Machine Learning and its applications, Springer Berlin Heidelberg, 2001.

[3] Badnjevic A, Cifrek M, Koruga D, Osmankovic D, Neuro-fuzzy classification of asthma and chronic obstructive pulmonary disease“ BMC Medical Informatics and Decision Making Journal, 2015.

[4] Badnjevic A, Koruga D, Cifrek M, Smith HJ, Bego T, „Interpretation of pulmonary function test results in relation to asthma classification using integrated software suite“, IEEE MIPRO, 2013.

[5] Avdakovic S., Omerhodzic I., Badnjevic A., Boskovic D., „Diagnosis of Epilepsy from EEG Signals using Global Wavelet Power Spectrum“, MBEC, 07-11. September 2014.

[6] Isobe, I., Ohkado, T., Hanyuda, H., Oda, S. & Gotoh, Y. (1994) The development of forecasting system of the water levels of rivers by neural networks (in Japanese). / Japan Soc. Hydrol. & Wat. Res. 7(2), 90-97.

[7] Jayawardena, A. W. & Fernando, D. A. K. Hydrological forecasting using artificial Neural networks. In Proc.Second Int. Study Conf. on GEWEX in Asia and GAME (Thailand), 1995.

[8] Bianchini, M., Frasconi, P. & Gori, M. Learning without local minima in radial basis function networks. IEEE Trans. Neural Networks 6(3), 1995.

[9] Chen, S., Cowan, C. F. N. & Grant, P. M. (1991) Orthogonal least squares learning algorithm for radial basis function networks. IEEE Trans. Neural Networks 2(2), 1991.

[10] Park, J. & Sandberg, I. W. Universal approximations using Radial-Basis-Function networks. Neural Computation 3(2), 1991.

[11] Moody, J. & Darken, C. J. Fast learning in networks of locally-tunes processing units. Neural Computation 1(2), 1989.

[12] Jayawardena.A.W, Fernando.D.A.K & Zhou. M.C. Comparison of Multilayer Perceptron and Radial Basis Function networks as tools for flood forecasting, Proceedings of the Conference held at Anaheim, California,1997.

[13] Halagundegowda G. R. and Abhishek Singh, Multilayer Perceptron Method of Artificial Neural Network for Classification of Farmers Based on Adoption of Drought Coping Mechanisms, International Journal of Pure and applied Bioscience,2008.

[14] Fesghandis. G.S, Pooya.A, Kazemi.M, and Azimi . Z. N.Comparison of Multilayer Perceptron and Radial Basis Function Neural Networks in Predicting the Success of New Product Development , Engineering, Technology & Applied Science Research, 2017.

[15] www.ics.uci.edu/pub/ml-repos/machine-learningdatabases/hepatitisb-disease/