Enhanced Neural Learning for Versatile Elliptical Basis Function Neural Network

Renu Mathur
Electronics& Communication Engg. Deptt, NITTTR,
Chandigarh, India
renu_mathur2@yahoo.co.in

Abstract— In this paper, a fast learning algorithm based on a hyper ellipsoidal function is proposed. This function can be translated and rotated to cover the high dimensional data set during learning process. Versatile elliptic basis function (VEBF) neural network with improved performance i.e. IVEBF (Improved Versatile Elliptic Basis Function) is also presented. This neural network has hidden layers which are adaptively divided into sub hidden layers. The comparison of proposed neural network is done with already existing neural networks such as radial basis function (RBF), multilayer perceptron (MLP) and versatile elliptic basis function (VEBF). Data sets of 'E. coli' are taken for comparing the different neural network with the help of graphs and tables.

Keywords— *classification, clustering, hyper-ellipsoid, neural network, fast learning, radial basis function (RBF), elliptical basis function (EBF), multilayer perceptron (MLP), versatile elliptic basis function (VEBF), recognition*

I. INTRODUCTION

Throughout the years, the computational changes have brought growth to new technologies. Such is the case of artificial neural networks, that over the years, they have given various solutions to the industry. Designing and implementing intelligent systems has become a crucial factor for the innovation and development of better products for society. Such is the case of the implementation of artificial life as well as giving solution to interrogatives that linear systems are not able resolve.

The standard RBF neural network consists of three layers. These are an input layer, a hidden layer, and an output layer. The learning algorithm of RBF neural network concerns the selection of the hidden-layer neuron centers and estimation of the weights connecting the hidden layer and the output layers. RBF network which is computationally light, parameter setting is simple and with an ability to generalize better for classification of unseen data. In the experiments, from the generalization performance and computation time, it can be said that the strategy has a better learning ability compared with existing method. Thus, this method is better than existing method as learning machine [2].

By extending the structure of the radial basis functions (RBF) neural network, an elliptical basis function (EBF) network with full co-variance matrices incorporated into the RBF whose parameters are estimated by the expectation-maximization [EM] algorithm for the classification of remotely sensed images. The algorithm is an iterative method and can be used to numerically approximate the maximum likelihood estimates of the parameters in a mixture model. The algorithm is able to compute the maximum likelihood estimates of the mean vectors and covariance matrices of a Gaussian mixture distribution. It has been applied in multiple spatial feature extraction, multiple data fusion, and spatial data mining with unsupervised approach. It yields the most accurate classification in the training and testing. But, the increase in accuracy, levels decreases after a certain size [6].

The main disadvantage in Radial basis function network is that it requires a large number of function, centers, data, if data to be modeled contains clusters with complicated shape. Therefore in elliptical the concept of spread which is calculate with the help of co-variance matrices and neural network into RBF structures and to use expectation-maximization (EM) algorithm to ensure the network parameters is incorporating. This is Elliptical Basis Function (EBF) with full co-variance matrices in the basis function Comparing with the Gaussian function of RBF neural network, the EBF can make the partition of input space more specific. So, it has the higher capability of pattern recognition. Ellipsoidal Basis Function cannot rotate to cover the data like the Gaussian function [3] [7].

Different learning algorithms may take different epochs in learning process because of the gradient method and optimization techniques. Furthermore, these algorithms consume more times when very large data set is learned. Some methods solve these problems by modifying optimization process, but they still use a lot of epochs during learning process. EBF cannot rotate to cover the data like the Gaussian function. The structure of both EBF and RBF is fixed during training and is not appropriate for the sequential learning. VEBF is based on a hyper-ellipsoidal function that can be translated and rotated to cover the data set during learning process. The structure is flexible and can be adjusted during the training process. Neural network have very fast training algorithm to learn a data set in only one pass. Once a data is learned, it is discarded. There is no need to use old data again for the future learning with new incoming data. It consumes less time and less memory when a very large data set is fed [4].

II. STRUCTURE OF VEBF NEURAL NETWORK

The structure of the VEBF neural network is shown in Fig. 1. The network consists of three layers i.e. an input layer, a hidden layer, and an output layer. In the input layer, the number of nodes in this layer is equal to the dimension of the input data space. There are no synaptic weights between the input layer and the hidden layer. In the output layer, the number of nodes in this layer is equal to the number of classes in the training data set.

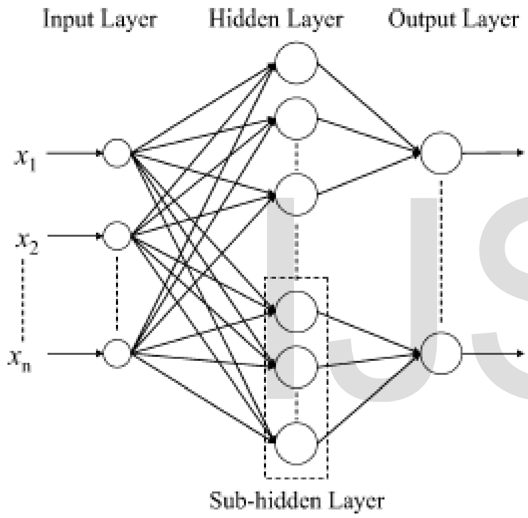

Fig.1 Structure of the VEBF Neural Network

The nodes in the output layer can be automatically increased during the learning process. In the hidden layer, the nodes in this layer are separated into sub hidden layers whose number of sub hidden layers is equal to the number of classes and all neurons in the same sub hidden layer are added to cover the training data in the same class. Furthermore, the nodes in each sub hidden layer can be automatically increased during the learning process.

Initially, there is no node in the network. The nodes in each layer in the network are automatically increased depending on some conditions during the learning process. All current training algorithms require both new incoming data and those previously trained data together in order to correctly learn the whole data set. With this reason, the neural network cannot learn a new training data set if the old training data set is discarded. They consume more time and more memory when a very large data set is fed. The structure of neural networks is fixed during training and is not appropriate for the sequential learning. Different learning algorithms use different epochs in learning process.

VEBF neural network have very fast training algorithm to learn a data set because-

Neural network use its parameters recursive that means recursive mean and recursive covariance. Neural network discard its older data set after learning the data set. It calculates the next new mean and covariance by the use of 'α', 'β' values and the previous mean value. After, every iteration a new mean and covariance is calculated and the calculated value of the output neuron is placed into a class by deciding with the help of decision function.

The performance of the VEBF neural network is evaluated on the two-class classification problems and the multiclass classification problems. The results are compared with the conventional RBF neural network with Gaussian RBF, MLP for multiclass classification problems and two-class classification problems, and Support Vector Machine (SVM) for two-class classification [4].

IV. THE PROPOSED LEARNING ALGORITHM

For any designing neural network based learning algorithm, the first need is to mimic biological activities of human brain with respect to the mathematical model used in computation theory. Since, it is cognitive process which is porn to number of bias in terms of mathematical formula or model. The need is to consider this bias as a value of threshold based on which decision needs to be taken for particular set classification problem. For this purpose, it is always suggested that the nature of data should be known that is going to deal and classify the data set classes. In this case, it is also found by using scatter diagram the nature of data set classes that it has elliptical distribution/shape.

In the case of RBFN/EBFN the inputs are directly connected to each basis function and the output of the activation functions are then weighted and summed. This is unlike the common MLP which have linear basis function architectures with inputs weighted before being summed and have sigmoidal or step activation functions. RBFN/EBFNs take non-linear input spaces and output linear activation outputs through a single hidden layer. Using inherent nonlinear approximation properties, RBFNs/EBFNs have the capability to model very complex patterns, which the RBFN/EBFNs are built on unique centroids (means), spreads (standard deviations from means) and activation functions. As with MLPs, weights are adjusted during training but in addition, the spreads and centers of each cluster are also updated. For feature spaces in 2-dimensions a circular cluster is formed for RBFs and elliptic cluster for EBFs; 3-dimensional spaces result in spherical clusters for RBFs and ellipsoids for EBFs; dimensions greater than 3 results in hyper spheres and hyper ellipsoids for RBFs and EBFs respectively. RBFs are special cases of EBFs where diagonal co-variances matrixes are equal. This is the basic concept for VEBF Neural

Network that is how high-dimensional data set is read or analyzed [3].
The proposed work is to modify learning algorithm of the VEBF neural network. By modifying the learning algorithm the neural network can become less dependent on the training inputs and the initial width. In the present learning algorithm the recursive mean computation is given by unew = α uold + β, where unew is the new mean vector, uold is the old mean vector, α = N/(N+1) and β = x N+1 /(N+1). In this paper, the formula of 'α' and 'β' will be changed for obtaining better result. By doing so, the number of neurons of the neural network get less affected. The training time is also reduced and the performance is expected to be improved.

<u>Purposed Learning Algorithm</u>
For the proposed learning algorithm, the value of 'α' and 'β' in the case of VEBF are changed to Alpha = $\frac{(N+1)}{10*(N)^2(N+2)}$ and Beta = $\frac{N}{10*2(N+1)^2}$; which gives better results than VEBF, RBF and MLP neural network. The proposed learning algorithm is Improved Versatile Elliptical Basis Function Neural Network (IVEBF).
Steps for purposed algorithm for learning of neural network are given below:
1. Create the trn/val/tst input/target subsets
[trn =training/ val =validation/ tst =testing input/target]
[MSE=mean square error, I=input layer, O= output, H=hidden]
2. For a each class classifier, the columns of the target matrices are all columns of the c-dimensional unit matrix eye [matlab] with the single 1 indicating the corresponding class. Inputs will be assigned to the class corresponding to the largest output.
3. Standardize the trn input and normalize the val and trn inputs with the means and co-variances from trn.
4. Now, take I-H-O(O=c) architecture and output by using single hidden layer.
5.Calculate the trn/val/tst MSEs for No. of equations (Neq), No. of weights.
6. For training to convergence, a necessary criterion for a unique solution is that either no. of equation is greater than or equal to no. of weights or hidden layers are less than or equal to how many hidden layers are sufficient (Hub) for elliptical function. Therefore, the decision function is designed such that Hub = (Neq-O)/(I+O+1).
However, weight estimates are more resistant to noise and measurement errors when hidden layers are less than the hidden layers sufficient for elliptical function. The optimal value, hidden layers are usually estimated via trial and error and use alpha and beta to adjust the weight and arbitrary constants to suit desired result.
7. To avoid non-optimal local min solutions, choose the best of many designs obtained by using multiple random weigh initializations for each candidate value of hidden layer.
8. Determine a search grid for hidden layer and the number of random weight initialization trials or test folds for each hidden layer.
9. Find the result based on the value of decision function.

V. RESULTS

The results for various data sets are discussed one by one in this part. Data sets of E. coli are taken for comparing the different neural network with the help of graphs and tables. Different neural networks for comparison are RBF, MLP, VEBF and IVEBF. The comparison is done on the basis of Training, No. of Neurons (same for every neural network) and Accuracy of different neural networks during training process.
For comparison we have taken the accuracy and training time for five test folds during training of neural network. The data sets of classification are taken from the UCI Machine Learning Repository. Basically two classification (E. coli) and multiclass classification (spam base) data sets are used. So, particular data set are for specific application. Now, one by one data set is discussed with the help of graphs which include test folds versus accuracy, training time and no. of neurons for each data set. Table 1 showing properties of the data sets used in the experiment like no. of attributes, no. of classes and no. of instances.

Table 1

PROPERTIES OF THE DATA SETS USED IN THE EXPERIMENT			
Data sets	**No. of attributes**	**No. of classes**	**No. of instances**
E. coli	8	8	336

A. E. COLI

E. coli is the name of a germ or bacterium that lives in the digestive tracts of humans and animals. There are many types of E. coli and most of them are harmless. But some can cause bloody diarrhea. These are called enter hemorrhagic E. coli (EHEC). In some people, this type of E. coli may also cause severe anemia or kidney failure, which can lead to death. Other strains of E. coli can cause urinary tract infections or other infections. E. coli infection is caused by coming into contact with the feces or stool of humans or animals. This can happen while drinking water or eating food that has been contaminated by feces. E. coli can be present in food, water and also E. coli can be transferred from person-to-person by contact.
So, there is an urgent need to study or analysis the E. coli data sets and this can be done with the help of neural networks. The E. coli data set consists of eight attributes in eight classes. There are 336 instances in this data set. The comparison between RBF, MLP and VEBF is given by the Fig. 2. Number of neurons in hidden layer are 8 in each test fold. The comparison is done on the basis of Training, Number of Neurons and Accuracy of different neural

ISSN 2229-5518

networks during training process. Number of neurons in hidden layer are 8 in each test fold.

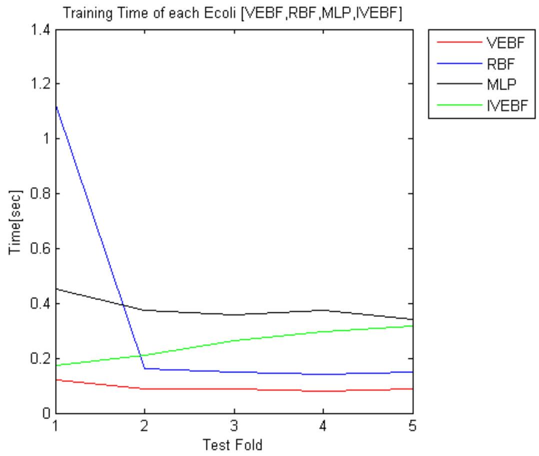

Fig. 2(a)

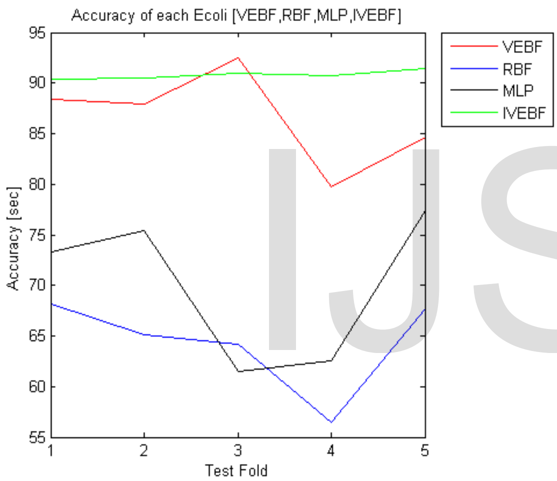

Fig. 2(b)

Fig. 2 Comparison Results trained by E. coli Data Set

The comparison results trained by E. coli data set of the RBF, MLP, VEBF and IVEBF are shown in Fig. 2. The accuracy of the proposed learning algorithm is higher than the accuracy of the RBF and the MLP. In addition, the training time of the proposed learning algorithm is less than the training times of the RBF and the MLP.

The proposed learning algorithm gives the better results than VEBF, RBF and MLP as shown in Fig. 2(a) by green line in graph of training but in case of accuracy it is not better in first three test folds and after third test fold, it is better than VEBF.

VI. CONCLUSION

In conclusion, very complex estimators will approximate the training data points better but may be worse estimators of the true function. Consequently, their predictions for samples not encountered in the training data set may be worse than the predictions produced by lower complexity estimators. The ability to predict well with samples not encountered in the training data set is usually referred to as generalization in the learning literature. Note that if the attributes of the noise term is known a priori, then it have inferred that the high- dimensional data set is accessed and result has good accuracy. A prior knowledge and the size and scope of the data set play a significant role in learning.

The ability of a model to predict accurately with novel data depends on the amount of data, the complexity of the model and the noise in the data. It was then argued that artificial neural networks provide a general and flexible nonlinear modeling strategy. So, the learning problem involves estimating the neural network's parameters, the number of parameters, the type of basis functions and the statistics of the noise. In addition, it might have to select the most appropriate set of input signals. A great deal of effort has been devoted to the solution of the parameter estimation problem. The other problems have received less attention. In contrast, the issues of noise estimation and model selection will be central to the scope of proposed work.

REFERENCES

[1] Bernard Widrow, and Michael A. Lehr, "30 Years of Adaptive Neural Networks: Perceptron, Madaline, and Back propagation", IEEE Transactions on Neural Networks, Vol. 78, No. 9, pp. 1415-1442, September 1990.

[2] Kazuhiko Yamashita, Goutam Chakraborty, Hiroshi Mabuchi and Masafumi Matsuhara, "An Efficient Method to Set RBF Network Parameters Based on SOM Training", IEEE Conference on Computer and Information Technology (CIT 2010), pp. 426-432, 2010.

[3] Tae Hong Park and Perry Cook, "Radial/Elliptical Basis Function Neural Networks For Timber Classification", Proceedings of the Journées d'Informatique Musicale, 2005.

[4] Saichon Jaiyen, Chidchanok Lursinsap, and Suphakant Phimoltares, "A Very Fast Neural Learning for Classification Using Only New Incoming Datum", IEEE Transactions on Neural Networks, Vol. 21, No. 3, pp. 381-392, March 2010.

[5] Jian-Cheng Luo, Qiu-Xiao Chen, Jiang Zheng, Yee Leung, Jiang-Hong Ma, "An Elliptical Basis Function Network for Classification of Remote-Sensing Images", IEEE Conference on Geoscience & Remote Sensing Symposium, Vol. 6, pp. 3489-3494, 2003.

[6] Jing Luoi, Ping-Chen Zai, Yun-Ni Jian, "Fault Diagnosis of Power Transformer Based on Ellipsoidal Basis Functional Neural Network", IEEE Conference on Wavelet Analysis and Pattern Recognition, Beijing, China, Vol. 2, pp. 695-698, 2007.

[7] Martin V. Butz, Pier Luca Lanzi, Stewart W. Wilson, "Hyper ellipsoidal Conditions in XCS: Rotation, Linear Approximation, and Solution Structure", Genetic and

Evolutionary Computation Conference (GECCO), pp. 1457-1464, 2006.

[8] Richard K. Gordon, W. Elliott Hutchcraft, "Using Elliptical Basis Functions in a Mesh less Method to Determine Electromagnetic Fields Near Material Interfaces", IEEE Conference on Antennas and Propagation Society International Symposium, pp. 3592-3595, June 2007.

[9] Suseela T. Sarasamma and Qiuming A. Zhu, "Min–Max Hyper ellipsoidal Clustering for Anomaly Detection in Network Security", IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics, Vol. 36, No. 4, pp. 887-901, August 2006.

[10] Sutharshan Rajasegarar, Christopher Leckie, James C. Bezdek and Marimuthu Palaniswami, "Centered Hyper spherical and Hyper ellipsoidal One-Class Support Vector Machines for Anomaly Detection in Sensor Networks", IEEE Transactions on Information Forensics and Security, Vol. 5, No. 3, pp. 820- 834, September 2010.

[11] A. Asuncion and D. Newman, UCI Machine Learning Repository, Schl. Inf. Comput. Sci., Univ. California, Irvine, CA, 2007.