2009 IEEE International Advance Computing Conference (IACC 2009)
Patiala, India, 6–7 March 2009

Diagnosis of Thyroid Disorders using Artificial Neural Networks

Anupam Shukla
ICT Department.
ABV-IIITM, Gwalior, India
dranupamshukla@gmail.com

Prabhdeep Kaur
ICT Department.
ABV-IIITM, Gwalior, India
prabhsainbhi@yahoo.co.in

Ritu Tiwari
ICT Department.
ABV-IIITM, Gwalior, India
rt_twr@yahoo.co.in

R.R.Janghel
ICT Department.
ABV-IIITM, Gwalior, India
janghel1310@gmail.com

Abstract. A major problem in medical science is attaining the correct diagnosis of disease in precedence of its treatment. This paper presents the diagnosis of thyroid disorders using Artificial Neural Networks (ANNs). The feed-forward neural network has been trained using three ANN algorithms; the Back propagation algorithm (BPA), the Radial Basis Function (RBF) Networks and the Learning Vector Quantization (LVQ) Networks. The networks are simulated using MATLAB and their performance is assessed in terms of factors like accuracy of diagnosis and training time. The performance comparison helps to find out the best model for diagnosis of thyroid disorders.

Key words: Thyroid disorders, Artificial Neural Networks, Backpropagation Networks, Radial Basis Function Networks, Learning Vector Quantization Networks, Diagnosis.

1. INTRODUCTION

Artificial Neural Networks are being touted as the wave of the future in computing. They are indeed self learning mechanisms which don't require the traditional skills of a programmer. Although there has been research using ANNs in medical fields, but there has not been a significant use in a hospital or clinic routinely [1]. The reason is that people don't think machines to be much reliable when it comes to diagnosis of a disease. But, soft computing tools like ANNs, Fuzzy Logic, Genetic Algorithm, can do well to ease and complement the work of medical experts [2, 3]. They can help to filter out the real patients, which will reduce the costs and time required for diagnosis. The doctors can then provide all their attention to the actual patients. By properly using ANNs, a trust can be established between them and the patients. [4].

The thyroid is one of the largest endocrine glands in the body. This gland is found in the neck below the mouth and at approximately the same level as the cricoid cartilage. The thyroid controls how quickly the body burns energy, makes proteins, and how sensitive the body should be to other hormones. [5, 6]. The thyroid participates in cellular processes by producing thyroid hormones, principally thyroxine (T_4) and triiodothyronine (T_3). These hormones regulate the rate of metabolism and affect the growth and rate of function of many other systems in the body. Iodine is an essential component of both T_3 and T_4. The thyroid also produces the hormone calcitonin, which plays a role in calcium homeostasis. The thyroid is controlled by the hypothalamus and pituitary. Hyperthyroidism (overactive thyroid) and hypothyroidism (underactive thyroid) are the most common problems of the thyroid gland. [7]

2. A VIEW ON TECHNIQUES USED

In this paper three ANN architectures, Back-Propagation Network (BPN), Radial Basis Function Networks (RBFN), and Learning Vector Quantization Networks (LVQN) are used.

A Back-Propagation Network is a multi-layered feed-forward neural network, making use of Backpropagation algorithm, which involves giving the input vector to the network, comparison of the desired and the network output for the input vector, changing each weight by an amount equal to the derivative of the error with respect to that weight times some learning rate. [8].

RBFN consists of three different layers- an input layer, a hidden layer and an output layer. The hidden units are known as Radial centers and have the same dimensions as that of the input vector. The transformation from the Input space to the hidden space is non-linear whereas the transformation from the hidden unit space to the output space is linear. In the hidden neurons, usually Gaussian activation function is used. In Neural Network Toolbox of MATLAB, we use a parameter called spread which controls the network performance. Larger spread means a smoother function approximation.[9]

Learning vector quantization (LVQ) Network consists of an input layer, a competitive layer and a linear layer [10, 11]. For an input, the neurons in the competitive layer compete with each other and the neuron with the most positive output becomes the winning neuron. The weights of the winning neuron (a row of the input weight matrix) are adjusted with the *Kohonen learning* rule. The only requirement in this network is that the competitive layer must have enough neurons, and each class must be assigned enough competitive neurons.

3. OVERALL METHODOLOGY

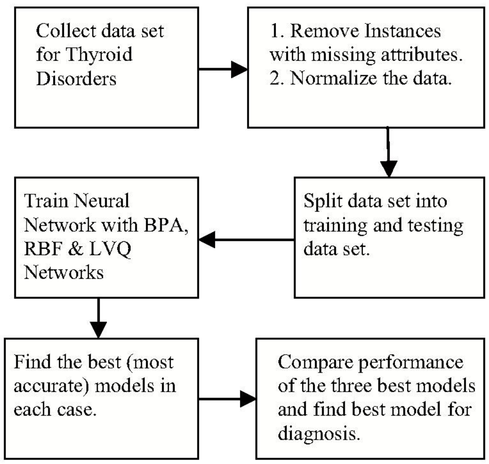

Figure 1: Overall Block Diagram of the methodology used in this work

In this work, after preprocessing, the data is trained with the three ANNs and generated at least four models for each of three networks, to find the overall best model for diagnosis. Figure 1 shows the overall methodology of the work performed in this paper. All neural networks have been simulated using MATLAB.

4. SIMULATED MODELS

4.1 Description of Data Set

The title of the data-set is thyroid gland data, obtained from UCI repository of machine learning databases [12]. From this data-set, 187 instances have been used for this work. Each instance has five attributes plus the class attribute. The five attributes are: T3-resin uptake test. (A percentage), Total Serum thyroxin as measured by the isotopic displacement method, Total serum triiodothyronine as measured by radioimmuno assay, basal thyroid-stimulating hormone (TSH) measured by radioimmuno assay, Maximal absolute difference of TSH value after injection of 200 micro grams of thyrotropin-releasing hormone as compared to the

Table 1: Class Distribution in training and testing data-set

	No.of instances having Class (Normal)	Number of instances having Class 2 (Hyper)	Number of instances having Class 3 (Hypo)	Total
Training Set	105	11	21	137
Testing Set	34	07	09	50
Total	139	18	30	187

basal value. All attributes are continuous. Each of the instances has to be categorized into one of the three classes: Class 1: normal, Class 2: hyper, Class 3: hypo functioning. Out of 187 instances, 137 have been used for training and 50 have been used for testing purposes, (shown in Table 1).

4.2 Experiment Results

The percentage accuracy of diagnosis for each of the networks was calculated as:
%age accuracy =(Correct Results / All Results)*100
On the basis of the best possible combination of the parameters i.e., the %age accuracy of diagnosis and time for training, the best model has been chosen.

4.2.1 Diagnosis using Backpropagation (BP) Networks

The feed- forward neural network architecture used in these experiments consisted of two hidden layers (Figure 2). The transfer function used in first hidden layer neurons was tan-sigmoid and that in second hidden layers neurons was log-sigmoid. In the output layer neurons purelin transfer function was used. The performance function used was mean sum-squared error (MSE). The learning rate of 0.07 and a momentum of 0.8 was used. Number of maximum allowable epochs was 24000. The experimental results of diagnosis using BP Networks are shown in Table 2. From this table, it is clear that the BP network model with 40 hidden neurons and accuracy of 92 % is the best BP model for diagnosis of thyroid disorders. Figure 3 shows the training curve for the best BP Network. Figure 4 shows the percentage accuracy of diagnosis using the best BP network. Here, the diagnosis is correct in 46 out of 50 cases.

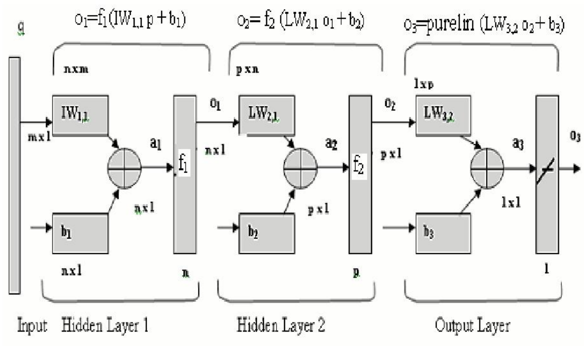

Figure 2: Architecture of Back- Propagation Network.

Table 2: Experimental Results for BP Networks

No. of Hidden Neurons	Mom-entum	Learning rate	No. of epochs	Training Time (sec)	%age Accuracy of Diagnosis
25 (20+5)	0.8	0.07	13092	48.89	86
30 (25+5)	0.8	0.07	19437	84.91	88
34 (28+6)	0.8	0.08	20975	91.28	86
40 (35+5)	0.95	0.08	7802	38.25	92
46 (40+6)	0.95	0.08	33927	174.94	90

MSE=0.01

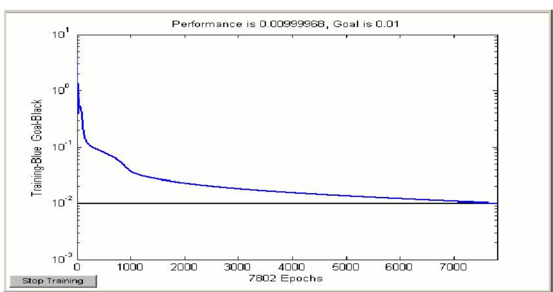

Figure3: Training curve of the best BP Network.

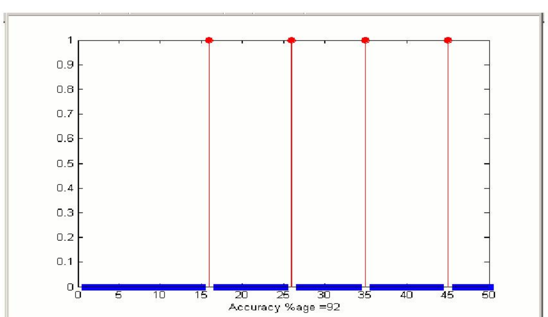

Figure 4: Percentage accuracy of diagnosis for the best BP Network.

4.2.2 Diagnosis using Radial Basis Function (RBF) Networks

In Radial Basis Function Networks, shown in Figure 5, an error goal of 0.01 is used. The performance function named MSE was used. The value of spread was changed to get different models. This also resulted into change of number of epochs being used by the network. The experimental results of diagnosis using RBF Networks are shown in Table 3. From this table, it is clear that the RBF network model with spread value 0.15, 100 hidden neurons and accuracy of 80% is the best RBF model for diagnosis of thyroid disorders. Figure 6 shows the training curve for the best RBF Network. Figure 7 shows the percentage accuracy of diagnosis using the best RBF network. Here, the diagnosis is correct in 40 out of 50 cases.

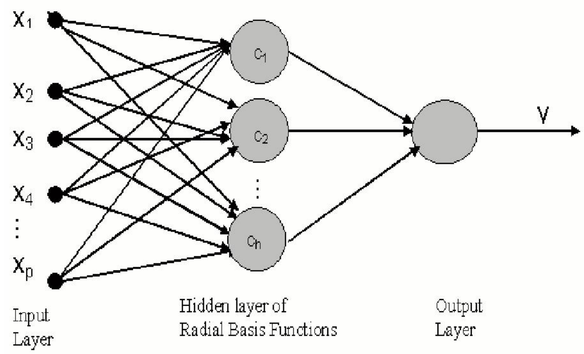

Figure 5: Architecture of Radial Basis Function Network used in our experiments

Table 3: Experimental Results for RBF Networks

Spread	Radial Basis Neurons	No. of epochs	MSE	Training Time (sec)	%age Accuracy of Diagnosis
0.13	75	75	0.13	1.02	76
0.15	**100**	**100**	**0.014**	**1.00**	**80**
0.25	100	100	0.012	1.08	68
0.3	100	100	0.03	1.00	70

Goal=0.01

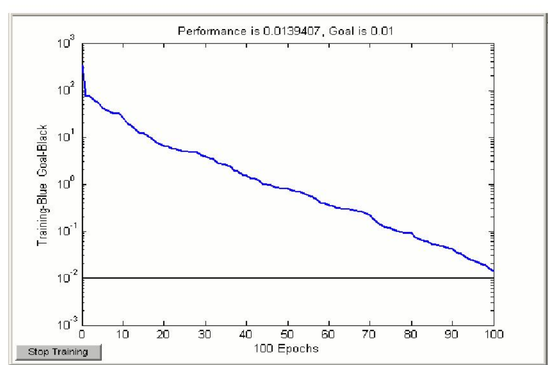

Figure 6: Training curve of the best RBF network.

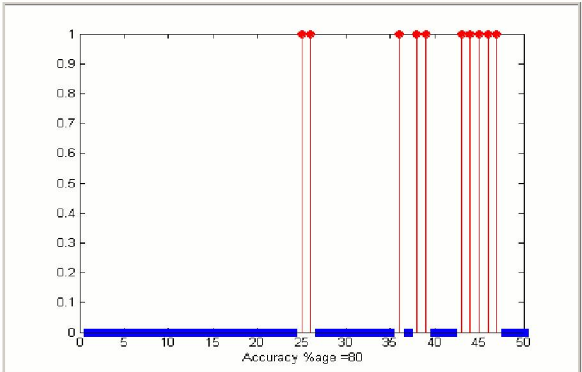

Figure 7: Percentage accuracy of diagnosis for the best RBF Network.

4.2.3 *Diagnosis using Learning Vector Quantization (LVQ) Networks*

In our experiments, different LVQ models were obtained by changing number of hidden neurons (Figure 8). The Kohonen learning rate was set to 0.09. The performance function used, was mean sum-squared error (MSE). Table 4 shows the experimental results for LVQ networks which differ from each other on the basis of number of hidden neurons used. The network with 28 hidden neurons and an accuracy of 98% is the best LVQ Network. Figure 9 shows the training curve for the best LVQ Network. Figure 10 shows the percentage accuracy of diagnosis using the best LVQ network. Here, the diagnosis is correct in 49 out of 50 cases.

Table 4: Experimental Results for LVQ Networks

No. of Hidd-en Neur-ons	Learning rate	No. of epochs	MSE	Training Time (sec)	%age Accuracy of Diagnosis
20	0.09	400	0.08	62.13	80
25	0.09	400	0.07	61.88	82
30	0.09	400	0.07	61.11	82
34	0.09	400	0.07	64.45	82
28	0.09	400	0.03	60.75	98

Learning rate=0.09

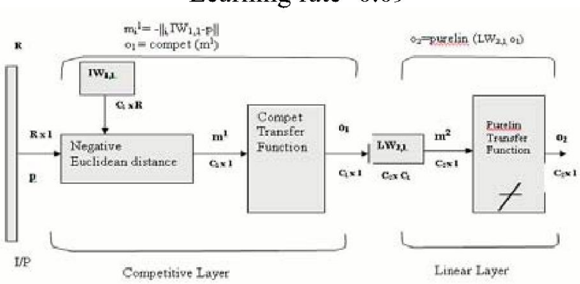

Figure 8: Architecture of LVQ Network used in our experiments

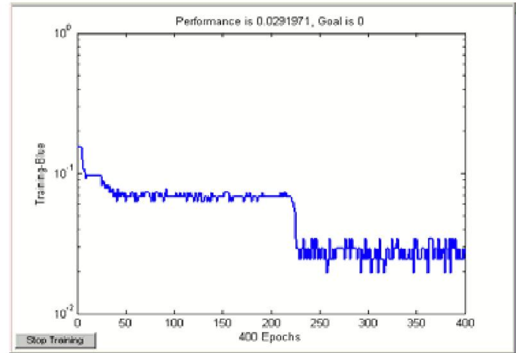

Figure 9: Training curve of the best LVQ network.

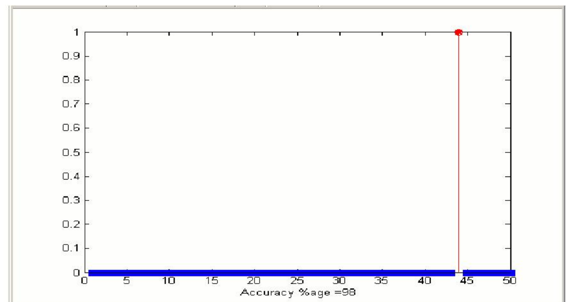

Figure10: Percentage accuracy of diagnosis for the best LVQ Network.

Table 5, shows the accuracy comparison of BP, RBF & LVQ models, and finds the LVQN as the best diagnostic model.

Table 5: Comparison of accuracy of the best BP, RBF & LVQ.

Network	%age accuracy of diagnosis (Evaluation Set)	Training Time (in seconds)
BPA	92	38.25
RBF	80	1.00
LVQ	98	60.75

5. Conclusion

ANNs can be effectively used for diagnosis of Thyroid disorders. On comparison of the performance of three neural network architectures, LVQ Network is found to have the best accuracy of diagnosis, which is 98%. This is possibly because of its advantage that it does not have a problem of getting trapped in local minima like BP network and does not require a very good coverage of input space, as required by Radial Basis functions used in RBF Networks. But, among the three neural networks used here, RBFN have the least training time when trained on the thyroid gland data.

This work provides a basis for diagnosis of other diseases as well and hence, can be extended to simulate a Medical Expert System for automated diagnosis of various diseases.

6. REFERENCES

[1] Lisboa, P. J. G., Vellido, A., and Wong, H., 2000, Outstanding Issues for Clinical Decision Support with Neural Networks, *Artificial Neural Networks in Medicine and Biology*, Springer, pp. 63-71

[2] Aliev, R. A., and Aliev, R. R., 2001, *Soft Computing and its Applications*, World Scientific Publishing Co. Pvt. Ltd., Singapore.

[3] Abraham, A., 2001, Neuro-Fuzzy Systems: State-of-the-Art Modeling Techniques, *In Connectionist Models of Neurons, Learning Processes, and Artificial Intelligence*, Springer, Berlin, J. Mira, A. Prieto (Eds.).

[4] Dybowski, R., and Gant V., 2001, *Clinical Applications of Artificial Neural Networks*, Cambridge University Press.

[5] Nussey, S., and Whitehead, S., 2001,*The thyroid gland in Endocrinology: An Integrated Approach*, BIOS Scientific Publishers Ltd.

[6] *http://en.wikipedia.org/wiki/Thyroid/,* accessed January 2008.

[7] *http://www.medicinenet.com*, accessed February 2008

[8] Yegnanarayana, B., 2003, *Artificial Neural Networks*, Prentice Hall of India Pvt. Ltd.

[9] Orr, M.J.L., 1996, Introduction to Radial Basis Function Networks, *Technical report, Center for Cognitive Science, University of Edinburgh.*

[10] Ghosh, A., Biehl, M., Freking, A., and Reents, G., 2004, A theoretical framework for analyzing the dynamics of LVQ: A statistical physics approach, *Technical Report 2004-9-02, Mathematics and Computing Science*, University Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands.

[11] Ghosh, A., Biehl, M., and Hammer, B., 2005, Dynamical analysis of LVQ type learning rules, *in M. Cottrell (ed.), Workshop on the Self-Organizing-Map WSOM'05*, Univ. de Paris (I).

[12] UCI repository of machine learning databases, *http://archive.ics.uci.edu/ml/datasets.html* accessed November