Scientific
Research
Publishing

Classification Using Two Layer Neural Network Back Propagation Algorithm

K. A. Mohamed Junaid

Department of Electronics & Instrumentation Engineering, R.M.K. Engineering College, Chennai, India
Email: kamjunaid@gmail.com

Abstract

Worldwide breast cancer is the most common form of cancer death occurring in 12.6% of women. This paper presents a cost effective approach to classify the normal, malignant and benign tumor using two layer neural network back propagation algorithm. Back propagation algorithm is used to train the neural network. Parallelization techniques speed up the computation process and as a result two layer neural networks outperform the previous work in terms of accuracy. Breast cancer tumor database used for the testing purpose is from the CIA machine learning repository. The highest accuracy of 97.12% is achieved using the two layer neural network back propagation algorithm.

Keywords

Breast Cancer, Back Propagation, Neural Network, Mammogram, Sonography

1. Introduction

Breast cancer is the critical cancer for women and starts with the tissues of the breast. Techniques to diagnose the breast cancer are Mammogram, magnetic resonance imaging of the breast; breast Sonography, Ductogram, Fine needle aspiration biopsy and core needle biopsy. A Mammogram is an X-ray of the breast. Screening mammograms are used to look for breast disease in women who appear to have no breast problems. Screening mammograms is the most commonly used method to detect breast cancer. Magnetic resonance imaging of the breast scans use radio waves and strong magnets instead of X-rays. In Breast MRI, a contrast liquid called gadolinium is injected into a vein before or during the scan to show details better. A Sonography sound wave is used to outline a part of the body. The use of Sonography instead of mammograms for breast cancer screening is not recommended. Ductogram helps to determine the cause of nipple discharge. In Fine needle aspiration biopsy hollow needle attached to a syringe to withdraw small amount of tissue from a suspicious area. In core needle

biopsy large needle is used to test the breast changes found by Sonography or mammography.

There are several other techniques to predict and classify breast cancer. [1] proposed a isotonic separation technique for diagnosis. [2] proposed a hybrid method based on fuzzy artificial immune system and k-nn algorithm for breast cancer diagnosis. There are different classifiers for classification accuracies. [3] proposed Support vector machine which takes a set of input data and predicts, for each given input, which of two possible classes forms the input. Decision tree is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences. Genetic algorithm performs fitness function and also crossover operation and mutation operation via random selection. This allows for strong analysis of the problem with genetic algorithms is that there are many possibilities and you could get stuck in one area and never find the best solution. Bayesian belief network represent the probabilistic relationships between diseases and symptoms. [4] proposed neural network with feed forward back propagation algorithm to classify the tumor.

In this paper, two layer neural network back propagation method was proposed to diagnose the breast cancer. In two layer neural network back propagation algorithm input layer is not counted because it serves only to pass the input values to the next layer. Neural network is a set of connected input and output units in which each connection has weights associated with it. During the learning phase, the network learns by adjusting the weights. It has longer training time therefore suitable for many applications. The rest of the paper is organized as follows. In the second part of this study we discussed related work and the Data set. Section 3 describes the proposed work. Section 4 presents the Materials and method. Section 5 presents the results of our classification. Section 6 concludes the paper with discussion and some directions for future research.

2. Related Work and Data Set

Among females there has been an increasing trend in breast cancer for the last few years over other cancers. In the year 2006-2008 out of the total 2095 breast cancer cases in CIA the percentage was estimated as 26.5%. The average Annual Crude Incidence Rate (CIR) and, Age Standardized Rate (ASR) per 100,000 females during 2006-2008 were 30.2 and 31.6 respectively. The distribution of breast cancer by sub site revealed that the unspecified parts of breast constituted the majority (73%) followed by the upper outer quadrant (10%), upper inner quadrant (4%), Lower outer quadrant (2 and Lower Inner Quadrants (1%) of breast. The histological verification of cancer diagnosis was possible in 86%. The ductal carcinoma (79%) was the most common morphological type followed by cystosarcoma phyllodes (1%), cystic/Mucinous neoplasms (0.8%), lobular carcinoma (0.7%) and medullary carcinoma (0.5%) and Carcinoma unspecified comprised 17%. [5] proposed Association Rule Mining for classification of mammographic images. Three steps in classification are preprocessing, Mining and Organization and reported an accuracy of 80.33%. [6] proposed Feature extraction technique to detect breast masses in the early stages of cancer development. [7] proposed Computer Aided Diagnosis system to detect and classify masses on ultrasound breast images using fuzzy support vector machines. [8] proposed a hybrid classifier combining unsupervised (ART) and supervised (LDA) learning method to classify malignant and benign masses. [9] proposed a Feed forward back propagation algorithm to detect and classify breast cancer. [10] proposed a hybrid classifier multilayer perceptron and genetic algorithm to classify the tumors based on ultrasound images.

Table 1 depicts the various sites of breast cancer. The number of cases being examined is 2095. The distribution

Table 1. Sites of breast cancer.		
Subsite	#	(%)
Breast		
Nipple	10	0.5
Central	25	1.2
Upper inner quadrant	81	3.9
Lower inner quadrant	23	1.1
Upper outer quadrant	217	10.4
Lower outer quadrant	34	1.6
Overlappent lesion	172	8.1
NOS	1533	73.2

of breast cancer by sub site revealed that the unspecified parts of breast constituted the majority (73.2%) followed by the upper outer quadrant (10.4%), upper inner quadrant (3.9%), Lower outer quadrant (1.6%) and Lower Inner Quadrants (1.1%) of breast. Nipples constitute (0.5%) which is the lowest in terms of the breast cancer sub site.

Dataset consists of 998 records each of which is characterized by nine attributes given in **Table 2**. [11] proposed the classification of malignant and benign tumor feature extraction algorithms based on principal component analysis and artificial neural network as a classifier. [10] [11] support vector machine which was used to determine the important features. Smooth support vector machine was used to classify the benign and malignant breast tumor. [12]-[15] proposed a classification system called fuzzy hyper sphere neural network that combines clustering and classification method. It is more stable and requires only less number of parameters compared with other classification method and achieves accuracy of 94.12% compared to other methods.

3. Proposed Work

Before training the network the network topology is described by specifying the number of units in input layer, hidden layer and output layer. Input layer is not counted and hidden layer is the first layer of network and data is propagated to output layer, the second layer of our proposed network. During training the network has been trained with same set of data's so that precise weight has been obtained. All data's is represented in range between −1 and 1. For our proposed classification system we used CIA dataset. Fifty two records are omitted in our test as they have missing values for some of the attributes. The proposed architecture of two layer neural network Back propagation algorithm is shown in **Figure 1**. During learning phase neural network learns by adjusting weights. Neural networks are inherently parallel. Back propagation algorithm performs learning on two layer neural network.

Figure 2 shows Age Specific incidence Rate vs Age group. It was found that the peak incidence of breast cancer occurred in the age group of 40 - 50 years.

The Back Propagation algorithm consists of presenting the data, calculating the error, back Propagate the error and adjusting the synapses. The process is repeated multiple times. It is a continuous process of evaluating outputs, adapting weights and training with new inputs. Two layer neural networks consist of Input layer, hidden layer and output layer. Here input layer is not counted since it is used to just pass on the data to the network.

4. Materials and Method

The algorithm for the proposed two layer neural network back propagation is given below:

Algorithm: Two layer neural network

Back propogation

Input: X, a training set containing tuples

Output: A trained neural network Begin

1) Select a tuple X from the training set and Present it to the network

2) Calculate the input of unit j with respect to unit j

Table 2. Description of attributes.

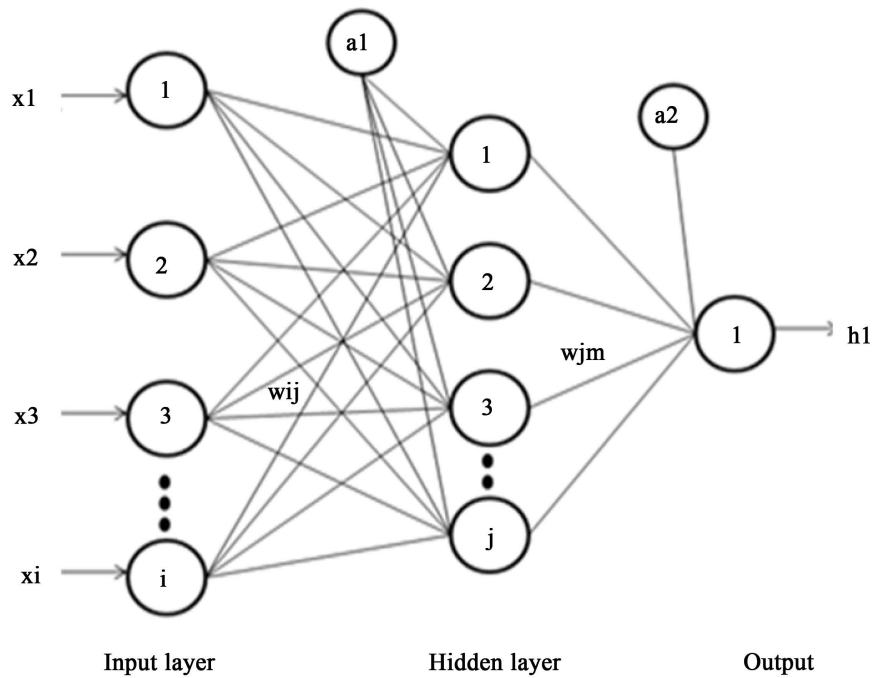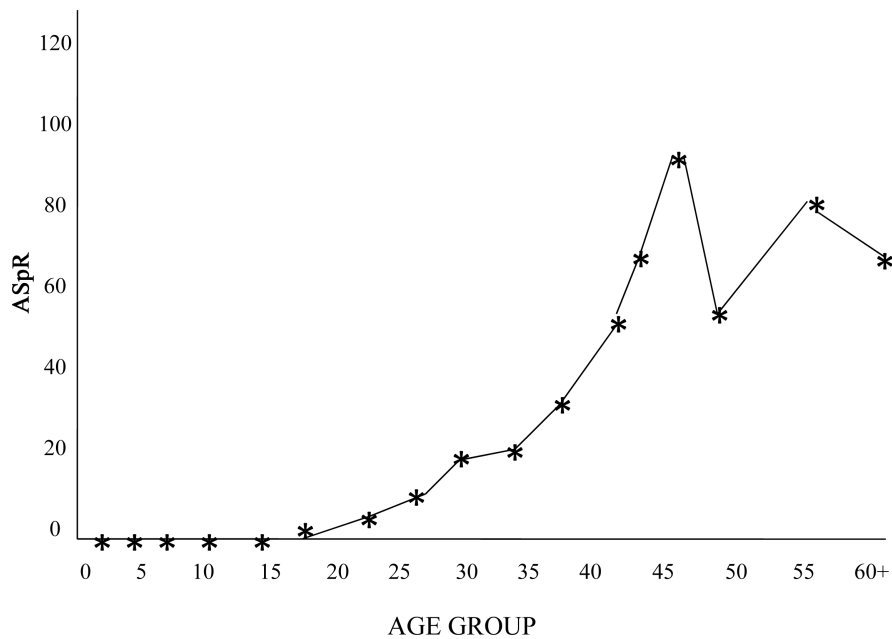| Attribute number	Description
1	clump thickness
2	uniformity of cell size
3	uniformity of cell shape
4	marginal adhesion
5	size
6	bare nuclei
7	bland chromatin
8	normal nucleoli
9	mitoses

Figure 1. Two layer neural network backpropagation.

Figure 2. ASpR vs Age group.

3) Compute the output of unit j
4) For each unit j in the output layer calculate Errj
5) For each j in the hidden compute the error with respect to m from the last to first layer
6) Multiply Errj with i (*i.e.*) δWij = Errj * i
7) Update weight (*i.e.*) weight = weight + δWij
8) δa1 = Errj (l)
9) Update bias a1 = a1 + a1δ

End

Error measure is calculated using the formulae:

$$E = \frac{1}{N} \sum_{t=1}^{N} \left(F\left(x_t ; W \right) - y_t \right)^2 \tag{1}$$

And the rule for changing the synaptic weights is given by:

$$\Delta w_i^j = -c \cdot \frac{\partial E}{\partial w_i^j}(w) \tag{2}$$

$$w_i^j, \text{new} = w_i^j + \Delta w_i^j \tag{3}$$

C is the learning parameter usually a constant.

5. Results

70% of data was used as training data and 30% of data was used as test data. Data is randomized to make sure that there is no bias in the data. Training process involved four different neural network models and each model has 8, 7, 6 and 5 numbers of neurons in hidden layer.

This study was performed with nine attributes and 998 records. **Table 3** shows the percentage of correct classification for each model tested with 299 records. We have obtained greater accuracy in terms of number of neurons. We use 70% number of data which is 698 from data for data training and 30% number of data which is 299 from data for data testing. The classifier accuracy for our proposed two layers Neural Network is 97.12% which is the greatest compared to other classifiers.

6. Discussion

Performance analysis was obtained by data mining tool Waikato Environment for Knowledge Analysis.

Table 4 summarizes the performance of various classifiers in terms of accuracy. Two layers Neural Network outperforms the other classifiers in terms of accuracy which is 97.12%.

7. Conclusion

Too many hidden neurons lead to over fit. As the number of neurons increases there is a need to memorize the training set, thus making the network useless on new data sets. Obviously if there are not enough hidden neurons then the network is not able to learn properly. The result shows that highest accuracy is obtained using eight

Table 3. Results of classification with distinct neurons.

Neural network models	Classification accuracy (%)
Neuron = 8	97.12
Neuron = 7	96.08
Neuron = 6	94.12
Neuron = 5	93.23

Table 4. Performance analysis.

The classifier	Correct classification rate (%)
Two layer neural network	97.12
Back propagation	96.63
Support vector machine	96.19

neurons. Further work is expected in validating the network using optimal neurons. In our future study more features and variables will also be considered.

References

[1] Alvarenga, A.V., Pereira, W.C.A., Infantosi, A.F.C. and Azevedo, C.M. (2013) Classifying Breast Tumours on Ultrasound Images Using a Hybrid Classifier and Texture Features. *Proceedings of the IEEE International Symposium on Intelligent Signal Processing*, *WISP*, Alcala de Henares, 3-5 October 2007, 1-6. http://dx.doi.org/10.1109/WISP.2007.4447589

[2] Azmi, M.S.B.M. and Cob, Z.C. (2012) Breast Cancer Prediction Based on Backpropagation Algorithm. *Proceedings of the IEEE Student Conference on Research and Development*, Malaysia, 13-14 December 2010, 164-168. http://dx.doi.org/10.1109/SCORED.2010.5703994

[3] Hadjiski, L., Sahiner, B., Chan, H.P., Petrick, N. and Helvie, M. (2010) Classification of Malignant and Benign Masses Based on Hybrid ART2LDA Approach. *IEEE Transaction on Medical Imaging*, **18**, 1178-1187. http://dx.doi.org/10.1109/42.819327

[4] Kanth, B.B.M.K., Kulkarni, U.V. and Giridhar, B.G.V. (2011) Prediction of Cancer Subtypes Using Fuzzy Hyper Sphere Clustering Neural Network. Int. J. Comput. Sci. Network Security.

[5] Purmani, S.W., Rahyu, S.P and Embong, A. (2008) Feature Selection and Classification of Breast Cancer Diagnosis Based on Support Vector Machines. *Proceedings of the International Symposium on Information Technology*, ITSim, Malaysia, 26-28 August 2008, 1-6. http://dx.doi.org/10.1109/ITSIM.2008.4631603

[6] Ryu, Y.U, Chandrasekaran, R. and Jacob, V.S. (2014) Breast Cancer Prediction Using the Isotonic Separation Technique. *European Journal of Operational Research*, **181**, 842-854. http://dx.doi.org/10.1016/j.ejor.2006.06.031

[7] Sahan, S., Polat, K., Kodaz, H. and Gunes, S. (2014) A New Hybrid Method Based on Fuzzy-Artificial Immune System and k-nn Algorithm for Breast Cancer Diagnosis. *Computers in Biology and Medicine*, **37**, 415-423. http://dx.doi.org/10.1016/j.compbiomed.2006.05.003

[8] Sameti, M., Ward, R.K., Morgans-Parkes, J. and Palacic, B. (2009) Image Feature Extraction in the Last Screening Mammograms Prior to Detection of Breast Cancer. *IEEE Journal of Selected Topics in Signal Processing*, **3**, 46-52. http://dx.doi.org/10.1109/JSTSP.2008.2011163

[9] Shi, X., Cheng, H.D., Hu, L., Ju, W. and Tian, J. (2013) Detection and Classification of Masses in Breast Ultrasound Images. *Digital Signal Processing*, **20**, 824-836. http://dx.doi.org/10.1016/j.dsp.2009.10.010

[10] Hasan, H and Tahir, N.M, (2012) Feature Selection of Breast Cancer Based on Principal Component Analysis. 6*th International Colloquium on Signal Processing and Its Application* (*CSPA*), Mallaca City, 21-23 May 2010, 1-4. http://dx.doi.org/10.1109/CSPA.2010.5545298

[11] Zaiane, O.R., Antonie, M.L. and Coman, A. (2012) Mammography Classification by an Association Rule-Based Classifier. MDM/KDD, 62-69. http://webdocs.cs.ualberta.ca/~zaiane/postscript/mdmkdd02.pdf

[12] Golub, T., Slonim, D., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J., Coller, H., Loh, M., Downing, J., Caligiuri, M., Bloomfield, C. and Lander, E. (2012) Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring. *Science*, **286**, 531-537. http://dx.doi.org/10.1126/science.286.5439.531

[13] Cho, S.B. and Won, H.H. (2013) Machine Learning in DNA Microarray Analysis for Cancer Classification. *Proceedings of the 1st Asia-Pacific Bioinformatics Conference on Bioinformatics*, 69-78.

[14] Chiang, J.-H. and Ho, T.S.-H. (2014) A Combination of Rough-Based Feature Selection and RBF Neural Network for Classification Using Gene Expression Data. *IEEE Transactions on NanoBioscience*, **7**, 91-99. http://dx.doi.org/10.1109/TNB.2008.2000142

[15] Luo, L.-K., Huang, D.-F., Ye, L.-J., Zhou, Q.-F., Shao, G.-F. and Peng, H. (2011) Improving the Computational Efficiency of Recursive Cluster Elimination for Gene Selection. *IEEE*/*ACM Transactions on Computational Biology and Bioinformatics*, **8**, 122-129. http://dx.doi.org/10.1109/TCBB.2010.44