International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-2 Issue-1, October 2012

A Survey on Neural Network Techniques for Classification of Breast Cancer Data

Shweta Saxena, Kavita Burse

Abstract— Breast cancer is the most common disease and major cause of death among women. Early detection of this disease can greatly enhance the chances of long-term survival of breast cancer victims. Artificial Neural Networks (ANN) have been widely used for cancer prediction and prognosis. This paper studies various techniques used for the diagnosis of breast cancer using ANN. Different methods for breast cancer detection are explored and their accuracies are compared.

Index Terms-- Artificial neural networks, Breast cancer diagnosis, Wisconsin breast cancer dataset.

I. INTRODUCTION

Breast cancer is the major cause of death by cancer in the female population [1]. Most breast cancer cases occur in women aged 40 and above but certain women with high-risk characteristics may develop breast cancer at a younger age[2].Cancer is a disease in which cells become abnormal and form more cells in an uncontrolled way. With breast cancer, the cancer begins in the tissues that make up the breasts. The cancer cells may form a mass called a tumor. They may also invade nearby tissue and spread to lymph nodes and other parts of the body. The most common types of breast cancer are-Ductal carcinoma and Lobular carcinoma. Ductal carcinoma cancer begins in the ducts and grows into surrounding tissues. About 8 in 10 breast cancers are this type. Lobular carcinoma cancer begins in lobules and grows into surrounding tissues. About 1 in 10 breast cancers are of this type[3]. There is much research on medical diagnosis of breast cancer with Wisconsin breast cancer database (WBCD) in neural network literature [33]. This paper reviews the existing/popular neural network techniques with WBCD data for the diagnosis of breast cancer.

II. NEURAL NETWORK TECHNIQUES FOR DIAGONOSIS OF BREAST CANCER

Various artificial intelligence techniques have been used to improve the diagnose procedures and to aid the physician's efforts [5], [6].

A. *Multilayer Perceptron (MLP)*

MLP is a class of feed forward neural networks which is trained in a supervised manner to become capable of outcome prediction for new data [7]. The structure of MLP is shown in fig 1. An MLP consists of a set of interconnected artificial neurons connected only in a forward manner to form layers. One input, one or more hidden and one output layer are the layers forming an MLP[10]. Artificial neuron is basic processing element of a neural network. It receives signal from other neurons, multiplies each signal by the corresponding connection strength, that is weight, sums up the weighted signals and passes them through an activation function and feeds the output to other neurons[8]. Neural classification of breast cancer data consists of two steps-training and testing. The classification accuracy depends on training[9]. A mapping between the input and output data could be established by assigning weights to the input numerical data during training[10]. The training requires a series of input and associated output vectors. During the training, the network is repeatedly presented with the training data and the weights and thresholds in the network are adjusted from time to time till the desired input output mapping occurs[8].

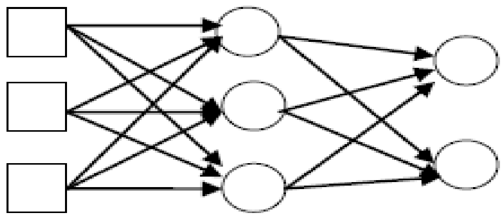

Fig.1 An MLP structure

B. *Radial Basis Function Neural Network (RBFNN)*

RBFNN is trained to perform a mapping from an m-dimensional input space to an n-dimensional output space. An RBFNN consists of the m-dimensional input x being passed directly to a hidden layer. Suppose there are c neurons in the hidden layer. Each of the c neurons in the hidden layer applies an activation function, which is a function of the Euclidean distance between the input and an m-dimensional prototype vector. Each hidden neuron contains its own prototype vector as a parameter. The output of each hidden neuron is then weighted and passed to the output layer. The outputs of the network consist of sums of the weighted hidden layer neurons[12]. The transformation from the input space to the hidden-unit space is nonlinear where as the transformation from the hidden-unit space to the output space is linear[11].

Manuscript published on 30 October 2012.
* Correspondence Author (s)
Shweta Saxena*, Department of Computer Science & Engineering, Oriental College of Technology, Bhopal, India.
Dr. Kavita Burse, Director, Oriental College of Technology, Bhopal, India.

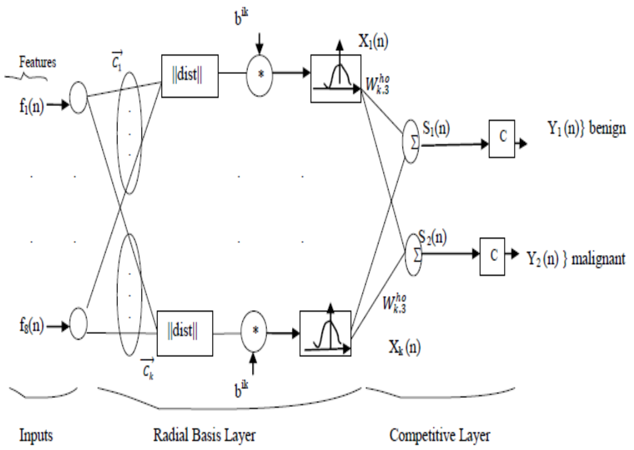

Fig. 2: Probabilistic neural network for breast cancer diagnosis.

The performance of an RBFNN network depends on the number and location (in the input space) of the centers, the shape of the RBFNN functions at the hidden neurons, and the method used for determining the network weights. Some researchers have trained RBFNN networks by selecting the centers randomly from the training data[13].

C. Probabilistic Neural Networks (PNN)

PNN is a kind of RBFNN suitable for classification problems. It has three layers. The network contains an input layer, which has as many elements as there are separable parameters needed to describe the objects to be classified. It has a pattern layer, which organizes the training set such that an individual processing element represents each input vector. And finally, the network contains an output layer, called the summation layer, which has as many processing elements as there are classes to be recognized[18]. For detection of breast cancer output layer should have 2 neurons (one for benign class, and another for malignant class). Each element in this layer combines via processing elements within the pattern layer which relate to the same class and prepares that category for output[18]. PNN used in [14] has a multilayer structures consisting of a single RBF hidden layer of locally tuned units which are fully interconnected to an output layer (competitive layer) of two units, as shown in Fig. 2. In this system, real valued input vector is feature's vector, and two outputs are index of two classes. All hidden units simultaneously receive the eight-dimensional real valued input vector. The input vector to the network is passed to the hidden layer nodes via unit connection weights. The hidden layer consists of a set of radial basis functions. Associated with j^{th} hidden unit is a parameter vector, called $\vec{C_j}$ a center. The hidden layer node calculates the Euclidean distance between the center and the network input vector and then passes the result to the radial basis function. All the radial basis functions are of Gaussian type. Equations which are used in the neural network model are as follows-

$$X_j = \emptyset(\|\vec{f} - \vec{c}_j\| * b^{ih}) \qquad (1)$$

$$\emptyset(X) = \exp(-X^2) \qquad (2)$$

$$b^{ih} = 0.833/s \qquad (3)$$

$$S_i = \sum_{j=1}^{h} W_{ji}^{ho} * X_j \qquad (4)$$

$$Y_i = \begin{cases} 1, \text{ if } S_i \text{ max of } \{ S_1, S_2 \} \\ 0, \text{ else} \end{cases} \qquad (5)$$

where i = 1,2, j = 1,2,. . . ,h, Y_i is the i^{th} output (classification index), $\vec{f}$ is the eight-dimensional real valued input vector, W_{ji}^{ho} is the weight between the j^{th} hidden node and the i^{th} output node, ($\vec{C_j}$) is the center vector of the j^{th} hidden node, s is the real constant known as spread factor, b^{ih} is the biasing terms of radial basis layer, and Ø(.) is the nonlinear RBF (Gaussian). PNN provides a general solution to pattern classification problems by following an approach developed in statistics, called Bayesian classifiers [15][16]. PNN combines the Bays decision strategy with the Parzen non-parametric estimator of the probability density functions of different classes [17].

D. Generalized Regression Neural Networks (GRNN)

GRNN is the paradigm of RBFNN, often used for function approximations [19]. GRNN consists of four layers: The first layer is responsible for reception of information, the input neurons present the data to the second layer (pattern neurons), the output of the pattern neurons are forwarded to the third layer (summation neurons), summation neurons are sent to the fourth layer (output neuron)[20]. If f(x) is the probability density function of the vector random variable x and its scalar random variable z, then the GRNN calculates the conditional mean E(z\x) of the output vector. The joint probability density function f(x, z) is required to compute the above conditional mean. GRNN approximates the probability density function from the training vectors using Parzen windows estimation [21]. GRNNs do not require iterative training; the hidden- to-output weights are just the target values t_k, so the output y(x), is simply a weighted average of the target values t_k of training cases x_k close to the given input case x. It can be viewed as a normalized RBF network in which there is a hidden unit centered at every training case. These RBF units are called kernels and are usually probability density functions such as the Gaussians. The only weights that need to be learned are the widths of the RBF units h. These widths (often a single width is used) are called smoothing parameters or bandwidths and are usually chosen by cross validation[19].

E. Fuzzy- Neuro System

Fuzzy-Neuro system uses a learning procedure to find a set of fuzzy membership functions which can be expressed in the form of if-then rules[22-24]. A fuzzy inference system uses fuzzy logic, rather than Boolean logic, to reason about data [26]. Its basic structure includes four main components- a fuzzifier, which translates crisp (real-valued) inputs into fuzzy values; an inference engine that applies a fuzzy reasoning mechanism to obtain a fuzzy output; a defuzzifier, which translates this latter output into a crisp value; and a knowledge base, which contains both an ensemble of fuzzy rules, known as the rule base, and an ensemble of membership functions, known as the database. The decision-making process is performed by the inference engine using the rules contained in the rule base[27].The fuzzy logic procedure can be summarized in following steps: Determination of the input and output variables that describe the observed phenomenon together with the selection of their variation interval, defining a set of linguistic values together with their associated membership functions that map/cover the numerical range of the fuzzy variable, and defining a set of fuzzy inference rules between input and output fuzzy variables[25].

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

F. Genetic Algorithm (GA)

The standard GA proceeds as follows: an initial population of individuals is generated at random or heuristically. Every evolutionary step, known as a generation, the individuals in the current population are decoded and evaluated according to some predefined quality criterion. To form a new population (the next generation), individuals are selected according to their fitness. Many selection procedures are currently in use, one of the simplest being fitness-proportionate selection, where individuals are selected with a probability proportional to their relative fitness. This ensures that the expected number of times an individual is chosen is approximately proportional to its relative performance in the population. Thus, high-fitness or good individuals stand a better chance of reproducing, while low-fitness ones are more likely to disappear [27]. Genetic algorithms can be used to determine the interconnecting weights of the ANN. During training of the network, the BP requires approximately two ANN evaluations (i.e., one forward propagation and one backward error propagation) for each iteration, while the GA required only one ANN evaluation (i.e., forward propagation) for each generation and each chromosome. In comparison to the conventional BP training algorithm, the GA has shown to provide some benefit in evolving the inter-connecting weights for the ANNs. In [28] although the GA trained ANN didn't outperform the BP-trained ANN at all numbers of ANN evaluations in the test set, the GA trained ANN was found to converge faster than the BP trained ANN in the training set.

III. DESCRIPTION OF WISCONSIN BREAST CANCER DATABASE (WBCD)

The database used for detection of breast cancer by artificial neural networks is publicly available in the Internet[36]. The dataset is provided by university of Wisconsin hospital, Madison from Dr. William H. Wolberg. This database has 699 instances and 10 attributes including the class attribute. Attribute 1 through 9 are used to represent instances. Each instance has one of two possible classes: benign or malignant. According to the class distribution 458 or 65.5% instances are Benign and 241 or 34.5% instances are Malignant. Table 1 provides the attribute information.

Table 1: Attributes of breast cancer data

S.no	Attribute	Domain
1	Clump thickness	1-10
2	Uniformity of cell size	1-10
3	Uniformity of cell shape	1-10
4	Marginal adhesion	1-10
5	Single epithelial cell size	1-10
6	Bare nuclei	1-10
7	Bland chromatin	1-10
8	Normal nucleoli	1-10
9	Mitosis	1-10
	Class	2 for benign, 4 for malignant

The original data can be presented in the form of analog values with values ranging from 0-10. Conversion of the given data sets into binary can be done based on certain ranges, which are defined for each attribute[37].

IV. COMPARISON TABLE

In [29] four different neural network structure, MLP, RBFNN, PNN and GRNN were applied to WBCD to show the performance of statistical neural networks on breast cancer data. The spread value of RBF, PNN and GRNN was chosen 4.4, 1 and 3, respectively. In MLP, learning rate was 0.6. According to results, RBF and PNN gives the best classification accuracy with 342 correct classifications while GRNN has the lowest accuracy with 330 correct classifications for the training set. MLP has 335 correct classifications. Accuracy comparison for popular neural network techniques with WBCD data for the diagnosis of breast cancer is shown by table 2.

Table 2: Accuracy comparison for test data classification

Type of Network	**Accuracy**	**References**
Radial Basis Function Neural Network (RBFNN)	96.18%	[29]
Probabilistic Neural Network (PNN)	97.0%	[29]
Multilayer Perceptron (MLP)	95.74%	[29]
Generalized Regression Neural Network (GRNN)	98.8%	[29]
Symbiotic Adaptive Neuro-Evolution (SANE)	98.7%	[30]
Information Gain and Adaptive Neuro-Fuzzy Inference System (IGANIFS)	98.24%	[31]
Xcyct system using leave one out method	90 to 91%	[32]
Adaptive Neuro-Fuzzy Inference System (ANFIS)	59.90%	[33]
Fuzzy	96.71%	[34]
Shunting Inhibitory Artificial Neural Networks (SIANN)	100%	[35]

V. CONCLUSION

The last decade has witnessed major advancements in the methods of the diagnosis of breast cancer. It was found that the use of ANN increases the accuracy of most of the methods and reduces the need of the human expert. The neural networks based clinical support systems provide the medical experts with a second opinion thus removing the need for biopsy, excision and reduce the unnecessary expenditure. This paper compares NN techniques for breast cancer diagnosis using WBCD. The MLP, RBFNN, PNN, GRNN, GA, Fuzzy- neuro -system, SANE, IGANIFS, Xcyct system, ANFIS, SIANN may be used for the classification problem. Almost all intelligent computational learning algorithms use supervised learning. The accuracy of different methods is compared in table 2.

REFERENCES

[1] Cancer Facts and Figures 2010[online]. Available: http://www.cancer.org/Research/CancerFactsFigures/CancerFactsFigures/ cancer-facts-and-figures-2010.

[2] *Tawam's 2012 Breast Cancer Awareness Campaign*[online]. Available: http://www.ameinfo.com/tawams-2012-breast-cancer-awareness-campaign-312992.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

[3] Breast Cancer [online]. Available: http://www.womenshealth.gov/breast-cancer/what-is-breast-cancer.
[4] S. Singh, Dr. P. R. Gupta. *Breast Cancer detection and Classification using Neural Network. IJAEST* [online] 6(1), Available: www. leijcset.iserp.org/.../2.
[5] D. B. Fogel , E. C. Wasson , E. M. Boughton, and V. W. Porto, "A step toward computer-assisted mammography using evolutionary programming and neural networks", *Cancer Letters*, vol. 119, no. 1, pp. 93-97, 1997.
[6] B. S. Hulka and P. G. Moorman, "Breast Cancer: Hormones and Other Risk Factors", *Maturitas*, vol. 38, no. 1, pp. 103-113, 2001.
[7] S. S. Haykin, *Neural Networks and Learning Machines*, 3rd ed. New York: Prentice Hall, 2009, pp. 906.
[8] F. Paulin and A. Santhakumaran, "Classification of breast cancer by comparing back propagation training algorithms", *Int. J. on Comput. Sci. and Eng. (IJCSE)*, vol. 3, no. 1, Jan. 2011.
[9] R. Nithya and B. Santhi, "Classification of Normal and Abnormal Patterns in Digital Mammograms for Diagnosis of Breast Cancer" *Int. J. of Comput. Applicat.***,** vol. 28, no.6, pp. 0975 – 8887, Aug. 2011.
[10] A. S. Mojarad, S. S. Dlay, W. L. Woo, and G. V. Sherbet "Cross Validation Evaluation for Breast Cancer Prediction Using Multilayer Perceptron Neural Networks" *American J. of Eng. and Applied Sci.*, vol. 5, no 1 pp 42-51, 2012.
[11] S. Haykin, *Neural Networks: A Comp rehensive Foundation*, Mac Millan College Publishing Company, 1994.
[12] M. R. Senapati, P. K .Routray, P. K. Dask "Breast Cancer Detection using Recursive Least Square and Modified Radial Basis Functional Neural Network", *Int. Conf. [ICCT-2010], IJCCT* vol. 2, no. 2, 3, 4, 3rd-5th December 2010.
[13] D. Broomhead and D.Lowe, "Multivariable Functional Interpolation and adaptive networks", *complex systems*, vol. 2, pp. 321-355, 1998.
[14] H. Temurtas, N. Yumusak , F. Temurtas , "A comparative study on diabetes disease diagnosis using neural networks", *Expert Systems with Applications*, vol. 36, no. 4, pp. 8610-8615, May 2009.
[15] G.S. Gill and J. S. Sohal, "Battlefield Decision Making: A Neural Network Approach", *J. of Theoretical and Applied Inform. Technology*, vol. 4, no. 8, pp. 697 –699, 2009.
[16] S. N. Sivandam, S. Sumathi, and S. N. Deepa, *Introduction to NEURAL NETWORKS Using Matlab 6.0*, TMH Company Ltd., 2006.
17] D. F. Specht, "Probabilistic Neural Networks", *Neural Networks*, vol. 3, no. 1, pp. 109-118, 1990.
18] Y. Shan, R. Zhao, G. Xu, H.M. Liebich, Y. Zhang, "Application of probabilistic neural network in the clinical diagnosis of cancers based on clinical chemistry data", *Analytica Chimica Acta*, vol. 471, no. 1, pp. 77-86, Oct. 23, 2002.
[19] C. Lu, J. De Brabanter, S. V. Huffel, I. Vergote, D. Timmerman, "Using artificial neural networks to predict malignancy of ovarian tumors", in *Proc. 23rd Annual Int. Conf. of the IEEE Eng. in Medicine and Biology Society*, 2001, vol. 2, pp. 1637-1640.
[20] R. G. Ahangar, M. Yahyazadehfar, and H. Pournaghshband, "The comparison of methods artificial neural network with linear regression using specific variables for prediction stock price in Tehran stock exchange", *Int. J. of Comput. Sci. and Inform. Security (IJCSIS)*, vol. 7, no. 2, Feb. 2010.
[21] I. Anagnostopoulos, C. Anagnostopoulos, A. Rouskas, G. Kormentzas and D. Vergados "The wisconsin breast cancer problem: diagnosis and DFS time prognosis using probabilistic and generalised regression neural classifiers" *Draft Version of paper to appeared in the Oncology Reports, special issue Computational Analysis and Decision Support Systems in Oncology*, last quarter 2005.
[22] K. Rahul, S. Anupam and T. Ritu, "Fuzzy Neuro Systems for Machine Learning for Large Data Sets," Proc. *of the IEEE Int. Adv. Computing Conf. 6-7*, Patiala, India, pp.541-545, 2009.
[23] C. Juang, R. Huang and W. Cheng, "An interval type-2 fuzzy-neural network with support-vector regression for noisy regression problems," *IEEE Trans. on Fuzzy Systems*, vol. 18, no. 4, pp. 686 – 699, 2010.
[24] C. Juang, Y. Lin and C. Tu, "Recurrent self-evolving fuzzy neural network with local feedbacks and its application to dynamic system processing," *Fuzzy Sets and Systems*, vol. 161, no. 19, pp. 2552-2562, 2010.
[25] I. Dumitrache, *Ingineria Reglarii Automate (Automatic Control Engineering)*, Bucuresti, Politehnica Press, 2005.
[26] RR Yager, LA. Zadeh, Fuzzy Sets, Neural Networks, and Soft Computing. New York: Van Nostrand Reinhold, 1994.
[27] Carlos Andre´s Pen˜a-Reyes, Moshe Sipper "A fuzzy-genetic approach to breast cancer diagnosis", *Artificial Intelligence in Medicine*, vol. 1, no 2, pp. 131-155, Oct. 1999.
[28] Yuan-Hsiang Chang, BinZheng, Xiao-Hui Wang, Walter F. Good "Computer-aided diagnosis of breast cancer using artificial neural networks: Comparison of Backpropagation and Genetic Algorithms", in *Int. Joint Conf Neural Networks UCNN' 99*, 1999, vol. 5, pp. 3674-3679.
[29] T. Kiyan, T Yildirim. "Breast cancer diagnosis using statistical neural networks", *J. of Elect. & Electron. Eng.*, vol .4-2, 2004, pp. 1149-1153.
[30] R. R. Janghel, A. Shukla, R. Tiwari, and R. Kala, "Breast cancer diagnostic system using symbiotic adaptive neuro-evolution (SANE)" in *Proc. Int. conf. of Soft Computing and Pattern Recognition 2010 (SoCPaR-2010)*, ABV-IIITM, Gwalior, 7th-10th Dec., pp. 326-329.
[31] M. Ashraf, L. Kim and H. Xu., "Information gain and adaptive neuro-fuzzy inference system for breast cancer diagnoses", in *Proc. Comput. Sci. Convergence Inform. Tech. 2010 (ICCIT-2010)*, IEEE, Seoul , 30th Nov.-2nd Dec., pp. 911 – 915.
[32] V. Bevilacqua, G. Mastronardi, F. Menolascina "Hybrid data analysis methods and artificial neural network design in breast cancer diagnosis: IDEST experience", in *Proc. Int. Conf. on Intelligent Agents, Web Technologies and Internet Commerce and Int. Conf. on Computational Intelligence for Modeling, Control Automation 2005 (CIMCA-2005)*, 28th-30th Nov., IEEE, Vienna, pp. 373-378.
[33] W. Land, and E. Veheggen, "Experiments using an evolutionary programmed neural network with adaptive boosting for computer aided Diagnosis of breast cancer", in *Proc. IEEE Int. Workshop on Soft Computing in Ind. Application, 2003 (SMCia-2003)*, Finland, 23rd-25th June, pp. 167-172.
[34] P. Meesad and G. Yen., "Combined numerical and linguistic knowledge representation and Its application to medical diagnosis", *IEEE transactions on Syst., Man and Cybernetics, Part A: Syst. and Humans*, vol. 3, no.2, pp. 206-222.
[35] G. Arulampalam, and A. Bouzerdoum, "Application of shunting inhibitory artificial neural networks to medical diagnosis", in *Proc. 7th Australian and New Zealand Intelligent Inform. System Conf.* 2001, IEEE, Perth, 18th-21st Nov., 2001, pp. 89 -94.
[36] Breast Cancer Wisconsin Data [online]. Available: http://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/breast-cancer-wisconsin.data.
[37] Dr. K. U. Rani, "Parallel Approach for Diagnosis of Breast Cancer using Neural Network Technique" *Int. J. of Comput. Applicat.*, vol. 10, no. 3, pp. 0975 – 8887, Nov. 2010.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)