Intelligent Decision Support System for Breast Cancer

R.R.Janghel, Anupam Shukla, Ritu Tiwari, Rahul Kala

Soft Computing and Expert System Laboratory
ABV-Indian Institute of Information Technology and Management, Gwalior, India
rrj.iiitm@gmail.com[1] , dranupamshukla@gmail.com[2],
rt_twr@yahoo.co.in[3], rahulkalaiiitm@yahoo.co.in[4]

Citation: R. Janghel, A. Shukla, R. Tiwari, R. Kala (2010) Intelligent Decision Support System for Breast Cancer, *Proceedings of the International Conference on Swarm Intelligence, Springer Lecture Notes in Computer Science*, Beijing, China, pp 351-358.

Final Version Available At: http://link.springer.com/chapter/10.1007%2F978-3-642-13498-2_46

Abstract. Breast cancer is the second leading cause of cancer deaths in women worldwide and occurs in nearly one out of eight women. Currently there are three techniques to diagnose breast cancer: mammography, FNA (Fine Needle Aspirate) and surgical biopsy. In this paper we develop an integrated expert system for diagnosis, prognosis and prediction for breast cancer using soft computing techniques. The basic aim is to compare the various neural network models from the recent literature. Breast cancer database used for this purpose is from the University of Wisconsin (UCI) Machine Learning Repository. Three different data sets have been used, each employing different diagnostic technique. It can use diagnosis, prognosis and survivability prediction of breast cancer patient in one intelligent system. We implement six models of neural networks namely Back Propagation Algorithm, Radial Basis Function Networks, Learning vector Quantization, Probabilistic Neural Networks, Recurrent Neural Network, and Competitive Neural network. Experimental Results show that different models give optimal performance for different types of data sets. However, all the models are able to solve the problem to a reasonable extent.

Keywords: Artificial Neural Networks, Breast Cancer, Expert System, Back Propagation Algorithm, Radial Basis Function Network, Learning Vector Quantization, Probabilistic Neural Networks, Recurrent Neural Networks, Competitive Neural Network

1 Introduction

Cancer is a general term that refers to cells that grow larger than 2mm in every 3 months and multiply out of control and spread to other parts of the body. Breast cancer is the second leading cause of cancer deaths in women worldwide and occurs

in nearly one out of eight women [1]. Breast cancer occurs mostly in women but rarely in men. Among women, breast cancer is the most common cancer and the second leading cause of cancer deaths behind lung cancer. Mammography, Biopsy and Fine Needle aspiration are three commonly used techniques for detection and diagnosis of breast cancer, but only one method applied in one system. We develop an integrated breast cancer diagnostic expert system using soft computing techniques which can be used for all type of the diagnosis and prognosis in one expert system.

Soft computing is an exciting field that deals with the learning of the historical data. In most problems a lot of data is available from history. These systems are made to learn this data by the use of training algorithms that may be specific to the system. Learning involves the extraction of rules or patterns from the historic data. It is evident that well trained systems would be able to give correct results to the problems that they have been trained with. Further the time and memory requirement would be reasonably less as the system has already summarized the historic data into some patterns or rules. Generalization is the ability of the system to give correct outputs to the unknown similar problems. This happens by the application of extracted patterns or rules by the system to the unknown inputs. A system is considered effective if it shows a very high generalizing capability.

2 Literature Review

Yuanjiao et al. proposed a technology to extract micro-calcifications clusters with accurate edge effects to obtain much more hidden information which can't be detected by the naked eye on mammograms in order to help the doctors in diagnosing early breast cancer [2]. Computerized microcalcification detection based on fuzzy logic, vibro-acoustography and probabilistic neural network on mammograms for breast cancer diagnosis has been carried out by Heng-Da et al. [3]. Image feature extraction was utilized to retrospectively analyze screening mammograms taken prior to the detection of a malignant mass for early detection of breast cancer in [4]. Statistical texture features for breast cancer detection using Support Vector Machine (SVM) and other machine learning methods like LDA, NDA, PCA, and ANN was done in [5]. SVM was able to achieve better classification accuracy. Early detection of breast cancer is the key to improve survival rate. Thermogram is a promising front-line screening tool as it is able to warn women of cancer up to 10 years in advance [6].

Laufer et al. proposed a modified self-organizing map with nonlinear weight adjustments to reduce number of unnecessary biopsies with a minimal number of subsequent [7]. Taio et al. used Kohonen self organizing map and multilayer perceptron trained with the backpropagation algorithm and classified results as sensitivity and false-positive fraction of actually benign or normal cases that are incorrectly classified [8]. Xiong et al. used statistics methods like PCA, PLS linear regression analysis, data mining methods, and hybrid system of combination rough set and probabilistic neural network. Probabilistic neural network to perform supervised classification and rough sets was able to reduce the number of attributes in the dataset without sacrificing classification accuracy in [9, 10, 11].

Karabatak et al. used ANN and ANFIS with linear discriminant analysis and principal component analysis for diagnosis of breast cancer. The target size is very reliably and target shape was classified [12, 13]. The masses from the mixed classes were input to a supervised linear discriminant classifier LDA in [14]. Ky constructed a hierarchical evolutionary RBF network and employed it to detect the breast cancer. The hierarchical RBF network model reduced number of input features with high detection accuracy [15, 16]. The optimum network for classification of breast cancer cells was found using Hybrid Multilayer Perceptron (HMLP) network. A combination of the proposed features gave the highest accuracy [17, 18].

Jain et al. used fuzzy-logic, Hybrid Neuro-Fuzzy generator based on the Knowledge Oriented Design (KOD) concept and Cooperative Neuro-Fuzzy systems using Genetic Algorithms were used for the classification (diagnosis) of breast cancer [19, 20]. Pena-Reyes et al. proposed fuzzy-genetic breast cancer identification. Fuzzy CoCo model is to evolve a fuzzy model that describes the diagnostic decision and the classification [21]. Seker et al. proposed a methodology with neural network, fuzzy logic, FK-NN and statistical method to prognostic analysis of cytometric image data. FK-NN system gives highest accuracy as compared to other techniques [22].

3 Methodology

In this work neural network models have been used for the diagnosis and prediction of breast cancer. The learning takes place through an iterative process of weight adjustments applied to its initial weight after epoch iteration of the learning process. Figure 1 shows the overall diagram of complete methodology. Mammography, biopsy and FNA database of breast cancer microscopic and clinical tests reports are collected. Based on the tests reports our expert system predicts cancer as either Benign or Malignant. In this paper we employ several methods like Back Propagation Network (BPA), Radial Basis Function (RBF) Networks, Learning Vector Quantization (LVQ), Recurrent Neural Network (RNN), Probabilistic Neural network (PNN), and Competitive Learning (CL) to reduce the sample-per-feature ratio; and then we investigate multiple machine learning methods to find an optimal classifier.

The database contains many missing attributes values, which are filled using data mining techniques. We calculate average of all available values for an attribute and put average value in place of missing values of attributes. Many models of neural network take normalized inputs; therefore we find maximum value of an attribute and then divide all values of attributes by this value. We hence get all inputs in range between zero and one.

The BPA makes use of the gradient descent to compute the new values of the weights and biases. It is quickly able to adjust the network weights for good performance. The graph or space denoting the error of the system for every combination of weights and biases is called as the error space. The aim of any training algorithm is to find the global optima in this search space. BPA may many a times get trapped in local minima. This is due to the absence of any global guiding strategy or the attempt to cover the entire error space which is highly complex and of high

dimensionality. The feed- forward neural network architecture used in this experiment consisted of one hidden layers along with input and output layers.The transfer function in hidden layer neurons and output layer neurons are sigmoid and purelin respectively. The performance function used was Mean Sum-squared Error (MSE) [23].

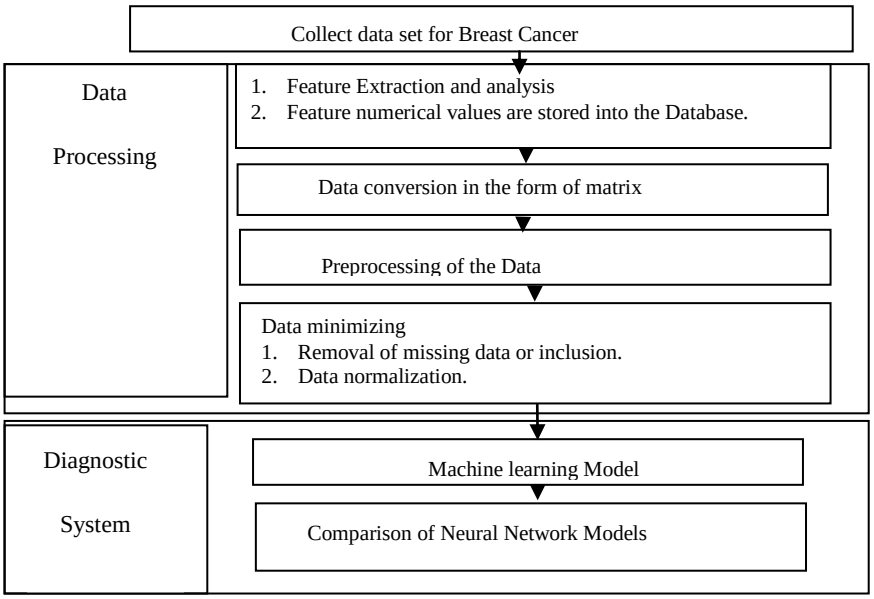

Fig. 1. Block Diagram of Complete Methodology

For PNN, the architecture is consist four types of units namely input unit, pattern units: Class 1 and Class 0, summation unit and output unit. The PNN is based on the theory of Bayesian classification and the estimation of probability density functions. It is necessary to classify the input vectors into one of the two classes in a Bayesian optimal manner. This theory allows for a cost function to represent the fact that it may be worse to misclassify a vector that is actually a member of class 1 than it is to misclassify a vector that belongs to class 0. The probabilistic neural net to estimates the probability density function using the following equation.

$$f_1(x) = \frac{1}{\sigma^2(2\pi)^{\frac{n}{2}}} - \frac{1}{m}\sum_{i=1}^{m_1}\left[\exp\frac{(x - x_1)^T(x - x_{1i})}{\sigma^2}\right]$$

Here X_{1i} is i^{th} Training pattern from class 1, n is the dimension of the input vectors, m_1 is Number of training pattern from class 1. σ is smoothing parameter (corresponds to standard deviations of Gaussian distribution), $f_1(x)$ serves as an estimator as long as the parent density is smooth and continuous. A complete transfer function on the output of the summation units selects the maximum of these probabilities and produces 1 for malignant and 0 Benign [24].

4 Simulation Results

The results are measured against the following diagnostic performance measures: *True positive* (TP): the number of positive cases correctly detected, *true negative* (TN): the number of negative cases correctly detected, *false positive* (FP): the number of negative cases diagnosed as positive, and *false negative* (FN): the number of positive cases diagnosed as negative are shown in table 1. The various performance measures are summarized in the same table.

Table 1. Diagnostic performance measures Breast cancer.

Cancer Test	**Present**	**Absent**	**Total**
Positive	True Positive [TP]	False Positive [FP]	[TP +FP]
Negative	False Negative [FN]	True negative [TN]	[FN+TN]
Total	(TP + FN)	(TN + FP)	(TP + FN+ TN + FP)

Sensitivity	TP / (TP + FN)
Specificity	TN / (TN + FP)
Accuracy	(TP + TN) / (TP + TN + FP + FN)
False positive rate (FPR)	FP / (TP+FP)
False negative rate (FNR)	FN / (FN+TN)
Positive predictive value (PPV)	TP / (TP + FP) = 1-FPR
Negative predictive value (NPV)	TN / (TN + FN) = 1-FNR

4.1 Susceptibility Prediction using Fine Needle Aspiration Microscopic data

The Wisconsin breast cancer diagnosis (WBCD) database is the result of the efforts made at the University of Wisconsin Hospital for accurately diagnosing breast masses based solely on an FNA test. The database contains diagnosis results of 699 patients out of which 458 are benign and 241 are malignant [25]. Different neural network models were simulated for different parameter settings. Experimental results of breast cancer system using different neural network models are shown in table 2. LVQ emerged as the optimal network, which used 10 hidden neurons and 0.01 learning rate. The performance function used was MSE.

Table 2. Performance comparisons of ANN models

ANN Model	Sensitivity	Specification	Accuracy	FPR	FNR	Training Time (seconds)
BPA	17.39	79.47	51.88	65.48	42.93	2.17
RBF	19.63	74.24	49.79	61.82	46.74	10.00
LVQ	**95.83**	**95.80**	**95.82**	**04.17**	**04.20**	**57.14**
PNN	19.63	74.24	49.79	61.82	46.74	0.1094
RNN	21.00	75.54	52.72	61.82	42.93	23.31
CL	25.00	77.09	74.48	94.55	04.89	17.53

4.2 Susceptibility Prediction using Fine Needle Aspiration Digital Image data

The purpose of the data set is to classify a tumour as either benign or malignant based on cell descriptions gathered by FNA image test. This database contains information about 569 patients with 357 benign and 212 malignant [25]. Again different neural network models were simulated with different architectures. Experimental results of breast cancer system using different neural network models are shown in table 3. The optimal network, BPA in this case, used 25 hidden neurons, MSE error function, 0.05 learning rate with 1000 epochs and momentum of 0.8.

Table 3. Performance comparison of Neural networks models

ANN Model	Sensitivity	Specification	Accuracy	FPR	FNR	Training Time (seconds)
BPA	**0.9296**	**1.0000**	**0.9814**	**0**	**0.0246**	**49.82**
RBF	0.7671	0.9490	0.8996	0.1515	0.0837	14.85
LVQ	0.4593	0.4627	0.4610	0.5373	0.5407	28.09
PNN	**0.9286**	**0.9950**	**0.9777**	**0.0152**	**0.0246**	**0.0675**
RNN	0.9167	1.0000	0.9777	0	0.0296	104.10
CL	0.2222	0.7488	0.6431	0.8182	0.2069	20.14

4.3 Susceptibility Prediction using Fine Needle Aspiration Microscopic data

This experiment uses Wisconsin Prognostic Breast Cancer (WPBC) databases, which has 34 real features computed from an image. The purpose of the data set is to classify a tumour as either recurrent or nonrecurrent based on cell descriptions gathered by image test. The database contains 198 patients database out of which 151 are nonrecurrent and 47 malignant are recurrent [25]. Following the same methodology different neural models with different settings were simulated. Experimental results of breast cancer system using neural network models are shown in table 4. In this case also BPA emerged as the optimal network. The optimal structure of ANN with BPA used 50 hidden neurons, 0.05 learning rate with 15000 epochs and a momentum of 0.8.

Table 4. Performance comparisons of ANN models.

ANN Models	Sensitivity	Specification	Accuracy	FPR	FNR	Training Time (seconds)
BPA	**1.0000**	**0.9722**	**0.9792**	**0.0769**	**0**	**53.48**
RBF	1.0000	0.8750	0.8958	0.3846	0	4.2
LVQ	0.9583	0.9583	0.9583	0.0417	0.0417	22.70
PNN	1.0000	0.9459	0.9583	0.1538	0	0.0936
RNN	1.0000	0.9459	0.9583	0.1538	0	34.23
CL	0.3333	0.7333	0.7083	0.9231	0.0571	6.71

5 Conclusions

All the systems were simulated on the breast cancer database. The ultimate aim is to make an expert system that would assist the doctors in diagnosis of the disease. This would prove to be a very useful system considering the present scenario where diseases are on a hike and there is a lack of availability of specialized doctors. Such a system learns from the past data which is a collection of a lot of information in itself. The system tries to extract this hidden information to make a generalized system for the detection of the disease. It may again be noted that though we have built effective systems, we were still not able to get an accuracy of 100%, which is the ultimate goal of medical diagnosis.

Acknowledgement

The authors sincerely acknowledge Prof. S.G. Deshmukh, Director ABV-IIITM, Gwalior, India for encouraging and providing facilities to carry out this research work. This work is sponsored by ABV-IIITM Gwalior.

6 References

[1] http://www.breastcancer.org

[2] Yuanjiao MA, Ziwu WANG, Jeffrey Lian LU, Gang WANG, Peng LI Tianxin MA, Yinfu XIE ,Zhijie ZHENG "Extracting Micro-calcification Clusters on Mammograms for Early Breast Cancer Detection" Proceedings of the 2006 IEEE International Conference on Information Acquisition August 20 - 23, 2006, Weihai, Shandong, China ,pp499-504.

[3] Heng-Da Cheng, Yui Man Lui, and Rita I. Freimanis "A Novel Approach to Microcalcification Detection Using Fuzzy Logic Technique" IEEE transactions on medical imaging, vol. 17, no. 3, June 1998, pp442-450.

[4] Mohammad Sameti, Rabab Kreidieh Ward, Jacqueline Morgan-Parkes and Branko Palcic "Image Feature Extraction in the Last Screening Mammograms Prior to Detection of Breast Cancer" IEEE journal of selected topics in signal processing, vol. 3, no. 1, February 2009, pp 46-52.

[5] Al Mutaz M, Abdalla , Safaai Deris, Nazar Zaki and Doaa M. Ghoneim " Breast Cancer Detection Based on Statistical Textural Features Classification" 2008 IEEE , pp 728-730.

[6] T.Z. Tan , C. Quek , G.S. Ng a, E.Y.K. Ng "A novel cognitive interpretation of breast cancer thermography with complementary learning fuzzy neural memory structure" Expert Systems with Applications 33 (2007) 652–666.

[7] Shlomi Laufer and Boris Rubinsky "Tissue Characterization with an Electrical Spectroscopy SVM Classifier" IEEE transactions on biomedical engineering, vol. 56, no. 2, February 2009 pp 525-528.

[8] Taio C. S. Santos-Andrk and Anttinio C. Roque da Silva "A Neural Network Made of a Kohonen's SOM Coupled to a MLP Trained Via Backpropagation for the Diagnosis of Malignant Breast Cancer from Digital Mammograms" 1999 , IEEE pp 3647-3650.

[9] Kenneth Revett, Florin Gorunescu, Marina Gorunescu, Elia El-Darzi and Marius Ene "A Breast Cancer Diagnosis System: A Combined Approach Using Rough Sets and

Probabilistic Neural Networks ", EUROCON 2005, Serbia & Montenegro, Belgrade, November, 22-24, 2005,pp 1124-1127.

[10] Xiangchun Xiong, Yangon Kim, Yuncheol Baek, Dae Wong Rhee, Soo-Hong Kim "Analysis of Breast Cancer Using Data Mining & Statistical Techniques" Proceedings of the Sixth International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing and First ACIS International Workshop on Self-Assembling Wireless Networks (SNPD/SAWN'05) 2005 IEEE.

[11] Emad Nabil, Amr Badr ,Ibrahim Farag, Mohamed Osama Khozium " A Hybrid Artificial Immune Genetic Algorithm with Fuzzy Rules for Breast Cancer Diagnosis "INFOS2008, March 27-29, 2008 Cairo-Egypt, pp 31-40.

[12] David B. Fogel, Eugene C. Wasson, Edward M. Boughton, Vincent W. Porto, and Peter J. Angeline "Linear and Neural Models for Classifying Breast Masses "IEEE transactions on medical imaging, vol. 17, no. 3, June 1998, pp 485-488.

[13] Shakti K. Davis, Barry D. Van Veen, Susan C. Hagness, and Frederick Kelcz " Breast Tumor Characterization Based on Ultrawideband Microwave Backscatter" IEEE transactions on biomedical engineering, vol. 55, no. 1, january 2008, pp 237-246.

[14] Lubomir Hadjiiski, Berkman Sahiner, Heang-Ping Chan, Nicholas Petrick and Mark Helvie " Classification of Malignant and Benign Masses Based on Hybrid ART2LDA Approach "IEEE transactions on medical imaging, vol. 18, no. 12, December 1999 pp1178-1187.

[15] Ky Van Ha "Hierarchical Radial Basis Function Networks" 1998 EEE 1893 PP 1893-1898.

[16] Yuehui Chen, Yan Wang, and Bo Yang " Evolving Hierarchical RBF Neural Networks for Breast Cancer Detection " ICONIP 2006, Part III, LNCS 4234, Springer-Verlag Berlin Heidelberg 2006 ,pp. 137–144, 2006.

[17] Nuryanti Mohd Salleh , Harsa Amylia Mat Sakim and Nor Hayati Othman "Neural Networks to Evaluate Morphological Features for Breast Cells Classification" IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.9, September 2008, pp 51-58.

[18] Nuryanti Mohd Salleh , Harsa Amylia Mat Sakim and Nor Hayati Othman "Neural Networks to Evaluate Morphological Features for Breast Cells Classification" IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.9, September 2008, pp 51-58.

[19] A. Rahmoun, S. Berrani, "A Genetic-Based Neuro-Fuzzy Generator: NEFGEN" 2001 IEEE pp 18-23

[20] Carlos Andres Pena-Reyes, Moshe Sipper "A fuzzy-genetic approach to breast cancer diagnosis "Artificial Intelligence in Medicine 17 (1999) 131–155.

[21] C.A. Pena Reyes "Breast Cancer Diagnosis by Fuzzy CoCo" Coevolutionary Fuzzy Modeling, LNCS 3204, Springer-Verlag Berlin Heidelberg 2004 pp 71-87, 2004.

[22] H. seker, m.Odeyato, D. petrovie, R.N.G.Naguib, C. Bartoli, L. Alasilo and G.V. Sherbet, " Prognostic comparison of statistical, neural and fuzzy methods of analysis of breast cancer image cytometric data "2001 proc. Of the 23rd Annual IEEE MBS Int. Conference October 25-28, Istanbul, Turkey, pp 3811-3814.

[23] Shukla Anupam,Tiwari Ritu,Janghel R.R.,and Kaur Prabhdeep, Diagnosis of Thyroid Disorders using Artificial Neural Networks on 2009 IEEE International Advance Computing Conference (IACC 2009),Patiala, India, 6–7March 2009,pp:2722-2726.

[24] S.N. Sivanandan,S.N Deepa 2007, principle of soft computing, Wiley India private limited.

[25] http://www.cs.wisc.edu.