Detection of Breast Cancer Using Artificial Neural Network

Sameer Dixit[1], Rahul Ranjan[2]
[1, 2] Department of CSE
[1, 2] Integral University

Abstract- *Breast cancer is the second cause of dead among Woman. To demonstrate the usefulness of ARNN, we used Wisconsin breast cancer database. The database available in the UCI database repository contains 699 cases out of which we used 600 cases to train the network. In this dataset there are 375 benign cases and 225 as malignant cases. We see that at VP=0.2 the network has Recall (i.e. true negative rate) is 75% and average Accuracy=82.64% and Precision is 79%.*

I. INTRODUCTION

For many years, Breast cancer diagnosis has been a machine learning problem. Various machine learning algorithms has been widely used. The problem is basically nontrivial and not easy to solve because the data set is relatively small and noisy. For such problem technique which are used mainly rely on a large training data set would not work well. In developed countries, Breast cancer has become a major cause of death among women. To reduce breast cancer deaths the most effective way is to detect it earlier. However earlier treatment requires the ability to detect breast cancer in early stages. Early diagnosis requires correct and reliable diagnosis procedure that will allows physicians to distinguish benign breast tumors from malignant ones. The most important is automatic diagnosis of breast cancer and it is real-world medical problem. So, to find an accurate and effective diagnosis method is very important.

Generally two approaches are used the first approach is based on evolutionary artificial neural networks. In this approach, a feed forward neural network is evolved using an evolutionary programming algorithm. Both the weights and architectures (i.e., connectivity of the network) are evolved in the same evolutionary process. The network may grow as well as shrink. The second approach is based on neural network ensembles. In this approach, a number of feed forward neural networks are trained simultaneously in order to solve the breast cancer diagnosis problem cooperatively. The basic idea behind using a group of neural networks rather than a monolithic one is divide-and-conquer. The negative correlation training algorithm we used attempts to decompose a problem automatically and then solve them. We will illustrate how negative correlation helps a group of neural networks learn using a real-world time series prediction problem.

This paper describes neural network approaches to breast cancer diagnosis. Neural networks have been widely used for breast cancer diagnosis. However, most of these Applications assumed a predefined network architecture (including connectivity and node transfer functions) and used a training algorithm.

II. WHAT IS A NEURAL NETWORK?

Many tasks involving intelligence or pattern recognition are extremely difficult to automate, but appear to be performed very easily by animals.

The neural network of an animal is part of its nervous system, containing a large number of interconnected neurons (nerve cells). Artificial neural networks are computing systems whose central theme is borrowed from the analogy of biological neural networks.

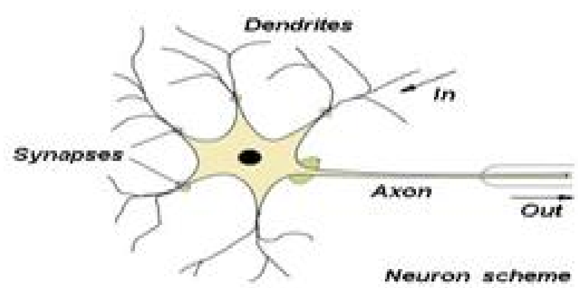

Figure1. Neuron Scheme

Artificial neural networks are also referred to as ‚neural nets', ‚artificial neural systems', ‚parallel distributed processing systems', and ‚connectionist systems'. For computing systems to be called by these pretty names, it is necessary for the system to have a labeled directed graph structure where nodes perform some simple computations.

From elementary graph theory we recall that a ‚directed graph' consists of a set of ‚nodes' (vertices) and a set of ‚connections' (edges/links/arcs) connecting pairs of nodes. A graph is a ‚labeled graph' if each connection is associated with a label to identify some property of the connection. In a neural network, each node performs some simple computations, each connection conveys a signal from one node to another, labeled by a number called the connection strength‛ or ‚weight‛ indicating the extent to which a signal is amplified or diminished by a connection.

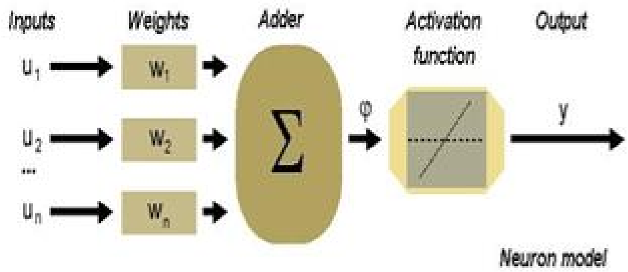

Figure 2. Neuron Model

The roots of all work on neural networks are in century-old neuro-biological studies. The following century-old statement by William James (1890) is particularly insightful, and ids reflected in the subsequent work of many researchers.

The amount of activity at any given point in the brain cortex is the sum of the tendencies of all other points to discharge into it, such tendencies being proportionate

1. To the number of times the excitement of other points may have accompanied that of the points in questions;

2. To the intensities of such excitements; and

3. To the absence of any rival point functionality disconnected with the first point, into which the dischargers may diverted.

Most neural network learning rules have their roots in statistical correlation analysis and in gradient descent search procedures. Hebb's (1949) learning rule incrementally modifies connection weights by examining whether two connected nodes are simultaneously ON or OFF. Such a rule is widely used, with some modifications.

2.1 Unsupervised learning

Our neural network is based on unsupervised learning. A three-month-old child receives the same visual stimuli as a newborn infant, but can make much more sense of them, and can recognize various patterns and features of visual input data. These abilities are not acquired from an external teacher, and illustrate that a significant amount of learning is accomplished by biological processes that proceed ‚unsupervised‛ (teacher-less). Motivated by these biological facts, the artificial neural networks attempt to discover special features and patterns from available data without using external help. Some problems require an algorithm to cluster or to partition a given data set into disjoint subsets (‚clusters‛), such that patterns in the same cluster are as alike as possible, and patterns in different clusters are as dissimilar as possible. The application of a clustering procedure results in a partition (function) that assigns each data point to a unique cluster. A partition may be evaluated by measuring the average squared distance between each input pattern and the centroid of the cluster in which it is placed.

III. ADAPTIVE RESONANCE THEORY

Adaptive resonance theory (ART) models are a neural network that performs clustering, and can allow the number of clusters to vary with the size of the problem. The major differences between ART and other clustering methods is that ART allows the user to control the degree of similarity between members of the same cluster by means of a user-defined constant called the vigilance parameter. It uses a MAXNET. A ‚maxnet‛ is a recurrent one-layer network that conducts a competition to determine which node has the highest initial input value. Apart from neural plausibility and the desire to perform every task using neural networks, it can be argued that the maxnet allows for greater parallelism in execution than the algebraic solution.

3.1 Proposed Work

We presented a system to detect that whether patient's cancer stage is malignant or benign. Neural networks are used to perform this task. Most cancer detection techniques rely on historical data. Experiments suggest that a cancer case currently under observation is malignant if it has the same or similar properties, as a malignant case that has been detected as malignant earlier in the same environment. Therefore, historical information helps us to detect whether a case is malignant or benign. Many modeling techniques have been proposed for and applied to breast cancer detection. In neural networks historical data which is obtained by regression analysis or by any other means are fed to network and the network gets trained accordingly that data is called training data. When we input operational data then network respond accordingly the trained data. So we used Adaptive Resonance Neural Networks (ARNN) which clusters already known modules which are faulty and fault-free. As it works under unsupervised learning so there is no distinction between training data and operational data. We used data from UCI machine learning data repository titled as 'Wisconsin Breast Cancer Database' which contains 9 attributes of 699 patients under survey. On the basis of these 9 attributes cases are termed as benign or malignant. The ARNN clustrifies cases which are benign in one cluster and malignant in another cluster.

IV. LITERATURE REVIEW

The number of research works conducted in the area of breast cancer detection, classification, scoring and grading. Many university centers, research centers and commercial institutions are focused on this issue because of the fa The Wisconsin breast cancer database using a hybrid symbolic-connectionist system was provided by Ismail Tahe and Joydeep Ghosh at university of texas, Austin. This paper introduces a new rule ordering and evaluation algorithm that orders extracted rules based on three performance measures so they can be used by any generic interface engine. Moreover, it introduces an integration algorithm to inspect the network's output as well as that derived rule base subsystem and provides a final decision, along with an associated confidence measure.

The Breast Cancer detection and Classification using Neural Networkwas provided by Shekhar Singh and Dr. P.R. Gupta at INTERNATIONAL JOURNAL OF ADVANCED ENGINEERING SCIENCES AND TECHNOLOGIES. This paper proposes a method for automatic Breast cancer detection, classification, scoring and grading to assist pathologists by providing second opinions and reducing their workload. A computer-aided Breast cancer detection, classification, scoring and grading system for tissue cell nuclei in histological image is introduced and validated as part of the Biopsy Analysis System. Cancer cell nuclei are selectively stained with monoclonal antibodies, such as the ant estrogen receptor antibodies, which are widely used as part of assessing patient prognosis in breast cancer.

Next Parallel Approach for Diagnosis of Breast Cancer using Neural Network Technique was provided by Dr. K. Usha Rani at International Journal of Computer Applications. In this paper a parallel approach by using neural network technique is proposed to help in the diagnosis of breast cancer. The neural network is trained with breast cancer database by using feed forward neural network model and back propagation learning algorithm with momentum and variable learning rate. The performance of the network is evaluated. The experimental result shows that by applying parallel approach in neural network model yields efficient result.

V. DATASET AND ALGORITHM

5.1 Dataset

The Wisconsin breast cancer database was originally provided by Dr. William H. Wolberg and used by a number of researchers in pattern recognition and machine learning. The database available in the UCI database repository contains 699 cases.

The original dataset contains 10 attributes each. The class of each instance is either benign or malignant. The remaining 9 attributes represent 9 cytological characteristics of breast fine-needle aspirates (FNAs), as shown in figure A. The cytological characteristics of breast FNAs were valued on a scale of one to ten, with one being the closest to benign and then the most malignant. This dataset can be use to predict if breast cancer is benign and malignant.ct that breast cancer is becoming the most common form of cancer disease of today's female population.

VI. NEURAL NETWORK

Adaptive resonance theory (ART) models are neural networks that perform clustering, and allow the number of clusters to vary the size of the problem. The major difference between ART and other clustering methods is that ART allows the user to control the degree of similarity between members of same cluster by means of a user-defined constant called the vigilance parameter. ART networks have been used for many pattern recognition tasks, such as automatic target recognition and seismic signal processing

6.1 Structure of the ARNN

This work uses ART2 architecture of ARNN neural network. The neural network is implemented with 29 input nodes and two output nodes as shown in figure 1. The network is trained with training data extracted from PROMISE dataset. The network improves the recall (true positive rate) in predicting weather a module is defected or not.

ART2 accepts continuous valued vectors. ART2 has highly complex input processing units. The input processing units if ART2 of normalization and noise suppression, along with the comparison of weights needed for reset mechanism. ART2 has two types of continuous valued inputs. One is called noisy binary posses combination signal and the other truly continuous. The first one can operate with the fast learning type data. The second type of data is more suitable with the slow learning mode. The architecture of ART2 is show in figure 1. Other parameters used are shown in table1.

1. Precision

The precision measures the chance of correctly predicting malignant cases among the cases classified as malignant. Either a smaller number of correctly predicted malignant cases or a large number of erroneously tagged benign cases would result in a low precision.

$$\text{Precision} = \frac{tp}{tp + tf} \quad (1)$$

2. Recall (true positive rate)

It is often known as true positive rate.

$$\text{Recall} = \frac{tp}{tp + fn} \quad (2)$$

3. Accuracy

The accuracy measures the chances of correctly predicting the malignant tendency of individual cases. It ignores the data distribution and cost information. Therefore, it can be misleading criterion if malignant cases are likely to represent a minority of the cases in the dataset.

$$\text{Accuracy} = \frac{tp + tn}{tp + fn + tn + fp} \quad (3)$$

4. True Negative Rate

$$\text{True negative rate} = \frac{tn}{tn + fp} \quad (4)$$

We have compared the model with other related purposed models:-

Model Name	Accuracy
Integrated neural network and fuzzy logic systems	73%
Digital Mammogram Breast Cancer Diagnosis	73.7%
Fuzzy system	80.6%
Adaptive neural network(vary according to vigilance parameter)	82.64%

TABLE 6:- Comparison with other models

The above information is taken from paper ‚Breast Cancer Diagnosis Systems: A Review‛*5+ which compares the performance of various breast cancer diagnosis models. We can see that the accuracy of ARNN is better as compared to other models but we can enhance the performance by varying the vigilance parameter.

VII. CONCLUSION

This paper presents a model, cancer detection using adaptive neural network for detecting the cancer stage as benign or malignant. To clusterify the medical data set a neural network approach is adopted. On the Wisconsis Breast Cancer dataset problem of adaptive resonance theory to cluster the dataset as it can be applied on continuous dataset. Clustering of the ARNN expressed reality interpretable knowledge about the dataset, and may be usable by the practitioners in a variety of ways. The result of this model is better the other models in case of accuracy. The ARNN can be tested on other datasets and variations in the neural networks itself need to be explored. In a developed software system, it would be desirable to improve the accuracy, precision and recall.

REFERENCES

[1] Ismail Taha and Joydeep Ghosh ‚Wisconsin breast cancer database using a hybrid symbolic- connectionist system,” university of Texas Austin 1996.

[2] Esugasini Subramaniam, Tan Kuan Liung, Mohd. Yusoff Mashor,Nor Ashidi Mat Isa ‚Breast Cancer Diagnosis Systems: A Review,‛Control and Electronic Intelligent System (CELIS) Research Group,2002

[3] S. N. Sivanandam, Sumathi & Deepa “Introduction to Neural Networks Using Matlab 6.0”. The Mc-Graw hill companies,2006.

[4] Kishan Mehrotra, Chilukuri K. Mohan, Sanjay Ranka “Elements of Artificial Neural Networks”. 1997 Massachusetts institute of technology.

[5] Tuba Kiyan And Tulay Yildirim, “Breast Cancer Diagnosis Using Statistical Neural Networks” Istanbul University, Journal Of Electrical And Electronics Engineering, Year 2004, vol. 4, Number 2, pp.1149-1153.

[6] Shekhar Singh, Dr P. R. Gupta “ Breast Cancer detection and Classification using Neural Network” IJAEST Vol No. 6, Issue No. 1, 004 – 009.

[7] Dr. K. Usha Rani (2010) “Parallel Approach for Diagnosis of Breast Cancer” using Neural Network Technique‛ International Journal of Computer Applications (0975 – 8887) Volume 10– No.3