Asian Journal of Information Technology 6 (1): 112-117, 2007

A Combination of Genetic Algorithm and ART Neural Network for Breast Cancer Diagnosis

[1]A. Punitha, [2]C.P. Sumathi and [3]T. Santhanam
[1,3]PG andResearch Department of Computer Science, D.G. Vaishnav College, Arumbakkam, Chennai-106, India
[2]Department of Computer Science SDNB Vaishnav College for Women, Chromepet, Chennai-44, India

Abstract: Medical diagnosis is one of the major problems in medical application. This includes the limitation of human expertise to diagnose disease manually. Extensive amount of knowledge and data stored in medical databases require specialized tools for analysis and effective usage of data. Breast cancer is one of the most common types of cancer in women in many countries. An intelligent tool, which will help in diagnosing the breast cancer, is the need of the hour. Researchers have found that Neural Network (NN) capabilities can help them to improve this domain and the use of Genetic Algorithms (GA) to optimize the input space. As many researchers have discovered, NNs and GAs may be combined which will result in highly successful adaptive systems. The performance of genetic algorithms and Adaptive Resonance Theory (ART) neural network for increasing the accuracy and objectivity of breast cancer diagnosis using the well-known and widely accepted Wisconsin Breast Cancer Data (WBCD) is examined in this study.

Key words: ART neural network, genetic algorithms, breast cancer diagnosis, wisconsin breast cancer database

INTRODUCTION

In general the increase in population may correspondingly increase the number of diseases too. Early diagnosis of a disease is of paramount importance. Since there are numerous diseases, diagnosing a disease is a foremost problem in medical field. Human beings always make mistakes and because of their limitation, diagnosis would give the major issue of human expertise. One of the most important problems of medical diagnosis, in general, is the subjectivity of the specialist. It can be noted, in particular in pattern recognition activities, that the experience of the professional is closely related to the final diagnosis. This is due to the fact that the result does not depend on a systematized solution but on the interpretation of the patient's signal (Lanzarini and Estelrrich, 1997).

Brause (2001) has highlighted that almost all the physicians are confronted duringtheir formation bythe task of learningtodiagnose. The basis for a valid diagnosis, a sufficient number of experienced cases, is reached only in the middle of a physician's career and is therefore not yet present at the end of the academic formation.

- This is especially true for rare or new diseases where also experienced physicians are in the same situation as newcomers.
- Principally, humans do not resemble statistic computers but pattern recognition systems. Humans can recognize patterns or objects very easily but fail when probabilities have to be assigned to observations.

Brause (2001) also has given an example of a study conducted in the year 1971 relating to these basic facts in the medical domain, which has reflected several limitations of human in diagnosis. Some of the important results are

- Best human diagnosis (most experienced physician): 79.7% .
- Computer with expert data base: 82.2%
- Computer with 600 patient data: 91.1%

From these results, it can be very easily concluded that human cannot ad hoc analyze complex data without errors.

In USA, a new less invasive technique, which uses super cooled nitrogen to freeze and shrink a non-cancerous tumor and destroy the blood vessels, is proposed (The weekend Australia, 2002).

A major class of problems in medical science involves the diagnosis of disease, based upon various tests performed upon the patient. When several tests are involved, the ultimate diagnosis may be difficult to obtain, even for a medical expert. So

Corresponding Author: T. Santhanam, PG and Research Department of Computer Science, D.G. Vaishnav College, Arumbakkam, Chennai-106

computers are widely used to help the physician from diagnosing the critical cases. As neural networks can do wonders with missing data, it is widely used for medical applications.

Breast cancer is one of the critical problems in medical diagnosis and it is the second largest cause of cancer deaths among women. The automatic diagnosis of breast cancer is an important, real-world medical problem (Pena-Reyes and Sipper, 2001). Generally breast cancer is detected as a lump/mass on the breast or through self-examination or by using any of the available tools namely mammography, Fine Needle Aspirates (FNA), thermographs and biopsy (Desilva *et al*., 1994). To overcome the above problem and to aid the physician, researchers have employed neural networks in breast cancer diagnosis. The attributes generated using FNA, one of the cost-effective tools, for diagnosis is used in this study.

A neural network is a powerful data-modeling tool that is able to capture and represent complex input/output relationships. The motivation for the development of neural network technology stemmed from the desire to develop an artificial system that could perform intelligent tasks similar to those performed by the human brain. Neural networks resemble the human brain in the following two ways:

- A neural network acquires knowledge through learning.
- A neural network's knowledge is stored within inter-neuron connection strengths known as synaptic weights (Wasserman and Philip, 1937).

The true power and advantage of neural network lies in their ability to represent both linear and non-linear relationships and in their ability to learn these relationships directly from the data being modeled. Traditional linear models are simply inadequate when it comes to modeling data that contains non-linear characteristics.

The applicability of Hierarchical Radial Basis Function has been demonstrated in (King *et al*., 2006) for breast cancer detection. Feed forward neural network is employed to detect breast cancer in (Setiono, 2000). The performances of Radial Basis Functions (RBF), Probabilistic Neural Networks (PNN) compared with Multi Layer Perceptron (MLP) is presented in (Tuba and Tulay, 2004) for breast cancer diagnosis. The feasibility of Flexible Neural Tree (FNT) and Wavelet Neural Network has been experimented in (Chen *et al*., 2005) for detection of breast cancer.

Medical databases present many of the challenges found in general information management settings where speed, efficiency, ease of use and accuracy are at a premium. A direct goal of improved computer-assisted medicine is to help deliver quality emergency care in situations that may be less than ideal (Carpenter and Milenova, 2000).

Combining Neural Nets with Evolutionary Algorithms leads to Evolutionary Artificial Neural Networks, which yields better results. Evolutionary Algorithms like the Genetic algorithms can be used to train NNs to choose their structure or design related aspects like the function of their neurons or to choose the number of efficient features for classification.

This study contains five sections. Section two deals with the introduction to ART (Adaptive Resonance Theory) Network and section three explains about genetic algorithms. In the present study, feature extraction combining both genetic algorithms and ART network is given. The experimental results and conclusion are given in the final section.

Adaptive resonance theory (art network): ART specifies a neural network that is capable of extracting invariant prototypical properties of vectors that impinge on the system. Internal codes spontaneously emerge through a process of self-organization. The system has an inherent plasticity that allows it to learn environmental invariant properties.

One of the nice features of human memory is its ability to learn many new things without necessarily forgetting things learned in the past. Most networks tend to forget old information if an attempt is made to add new information incrementally. As ART is an unsupervised learning model, it can retain the old information.

A central feature of all ART systems is a pattern matching process that compares an external input with the internal memory of an active code. ART matching leads either to a *resonant* state, which persists long enough to permit learning, or to a parallel memory search. If the search ends at an established code, the memory representation may either remain the same or incorporate new information from matched portions of the current input. If the search ends at a new code, the memory representation *adapts* the current input. This *match-based learning* process is the foundation of ART code stability. Match-based learning allows memories to change only when input from the external world is close enough to internal expectations, or when something completely new occurs. This feature makes ART systems well suited to problems that require online learning of large and evolving databases (Gail, 2002).

Grossberg and his associates layer the basic features of ART and its relation to perception out in a great number of articles. A block diagram for a typical ART system is displayed in Fig. 1. The main components are the attentional subsystem and the orienting subsystem. The attentional subsystem consists, among others, of two fields of neurons, F1 and F2, where each field may consist of several layers of neurons. These fields are connected with feed forward and feedback connection weights. The connection weights form the Long-Term Memory (LTM) components of the system and multiply the signals along these pathways. The name Short-Term Memory (STM) will be associated with the pattern of activity that develops on a field as an input pattern is processed. The orienting subsystem is necessary to stabilize the processing of STM and the learning in LTM. As can be seen from the figure, the F1 field receives input from possibly three sources. These three input sources are the bottom-up input to F1, the top-down input from F2 and the gain control signal. To avoid the possibility that mere feedback from F2 can generate spontaneous activity at level F1, i.e., to avoid that the system hallucinates, system dynamics are limited in such a way that at least two out of three inputs must be active to generate activity at the F1 field. This is called the 2/3 rule in ART. The same rule applies to the three possible input sources for the F2 level.

All ART systems incorporate basic features, notably, pattern matching between bottom-up input and top-down learned prototype vectors. This matching leads either to a resonant state that focuses attention and triggers stable prototype learning or to a self-regulating parallel memory search. This search ends in either of two ways. First, if an established category is selected, then this prototype may be refined to incorporate new information in the input pattern. In this case when an input matches an established category, resonance comes into picture. This resonant state persists long enough for learning to occur; hence the term adaptive resonance theory. Second, if the search ends by selecting a previously untrained node, then learning of a new category takes place. The criterion of an acceptable match is defined by a dimensionless parameter called vigilance. Vigilance weighs how close an input must be to the top-down prototype for resonance to occur. Because the vigilance parameter can vary across learning trials a single ART system is able to encode widely differing degrees of generalization. Low vigilance leads to broad generalization and more abstract prototypes than high vigilance. In the limit of very high vigilance, prototype learning reduces to exemplar learning.

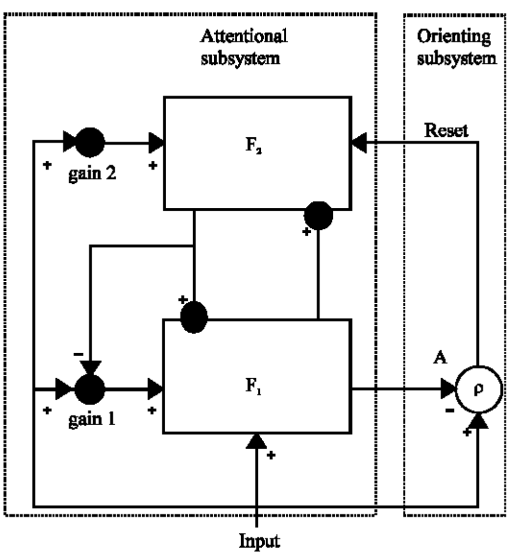

Fig. 1: Adaptive Resonance Theory Neural Network

A series of models based upon ART have been developed, such as ART2, ART3, ARTMAP, Fuzzy ART, Fuzzy ARTMAP, ART-EMAP and ARTMAP-IC, which have made ART become an important family of neural models.

GENETIC ALGORITHMS

A variety of computational models based on evolutionary processes have been proposed and the most popular models are those known as genetic algorithms. Genetic algorithms were invented by John Holland and are described quite well in David Goldberg's foundational text Genetic Algorithms in Optimization, Search and Machine Learning (David, 1989).

Genetic Algorithms (GAs) are automated heuristics that perform optimization by emulating biological evolution. They are particularly well suited for solving problems that involve loose constraints, such as discontinuity, noise, high dimensionality and multimodal objective functions (Arnold *et al.*, 1999).

The genetic algorithm has four main elements: the genetic code -a concise representation for an individual solution, the population -a number of individual solutions, the fitness function -an evaluation of the usefulness of an individual and the propagation techniques -a set of methods for generating new individuals (Talib and Hussian, 1994). The three most common propagation techniques are elitism, mutation and crossover. In elitism, the exact individual survives into the next generation. In mutation, changing a small number of randomly selected bits in the gene creates a new individual. In cross-over a new individual is created from two old ones by randomly selecting a split point in their genes and creating a new one with the left part from one parent and the right part from another parent. In any genetic algorithm, the two key aspects are the genetic representation and the fitness function.

A simple genetic algorithm as defined in (Talib and Hussian, 1994).

- Create an initial population of structures.
- Test each structure for quality with a heuristic fitness function.
- Select pairs of structures with a quality bias.
- Blend selected pair with crossover to produce new structures.
- Make small probabilistic modifications to the new structures, a process termed mutation.
- Place the new structures into the population, replacing existing structures.
- If no sufficiently high quality structure yet exists, go to 2.
- Report results and stop.

GENETIC ART

Genetic algorithms provide a suitable technique for searching the large space of similarity rules while neural networks ease computing the rule similarity cutoff and evaluating its fitness. Genetic algorithm is applied in various forms in neural networks for selecting the architecture, training the network, selecting the number of nodes in each layer, evolution of connection weights, evolution of learning rules etc. This combination provides powerful classification capabilities with tuning flexibility for either performance or cost-efficiency. Here genetic algorithms are used for selecting the input variables for the neural network. The constraint is to minimize the input variables and to maximize the classification rates.

Feature selection: Feature selection is one of the most important issues in machine learning especially when the number of observations is relatively small comparing to the number of features. Mathematically speaking, a finite set of inputs is sufficient in order to extract an accurate model out of the infinite observations. One should select the best features or inputs in the sense that they contain the necessary information. Then it would be possible to capture and reconstruct the underlying regularity or relationship between input-output data pairs (Andrew and Andrzej, 1999).

A new approach is proposed to build predictive models for diagnosing diseases. The new methodology combines Genetic Algorithms (GAs) for choosing predictive variables with Artificial Neural Networks (ANNs) for classifying breast cancer diagnosis.

Reducing the dimensionality of the input space tends to reduce over fitting in the predictive model, especially for highly flexible models such as artificial neural networks, thereby improving generalization. Feature selection can also significantly improve the comprehensibility of the resulting classifier models. Even a complicated model-such as a neural network-can be more easily understood if constructed only from a few variables (YonySeog, 2005).

A specifically designed GA, Genetic Algorithm with ART (GeneticART) is used to search through the possible combinations of features. Two quality measurements-accuracy (which is maximized) and complexity (which is minimized)-are used to evaluate the quality of each feature subset.

Genetic ART is used to select the feature subset that returns the highest classification accuracy over all records. However, evaluations are carried out with great care and based on the correlations of the data in medical databases, the constraint has been specified to select any five attributes for breast cancer diagnosis to rank records based on the estimated probability of belonging to a target class and to select the feature subset that maximizes classification accuracy over a pre-determined number of records (Chen *et al*., 2005). As there is a huge homogeneity among the attributes, in other words the dependency among the attributes is high and there are many possibilities for substitution.

The flowchart in Fig. 2 shows how genetic ART algorithm is used for input selection. ART neural network

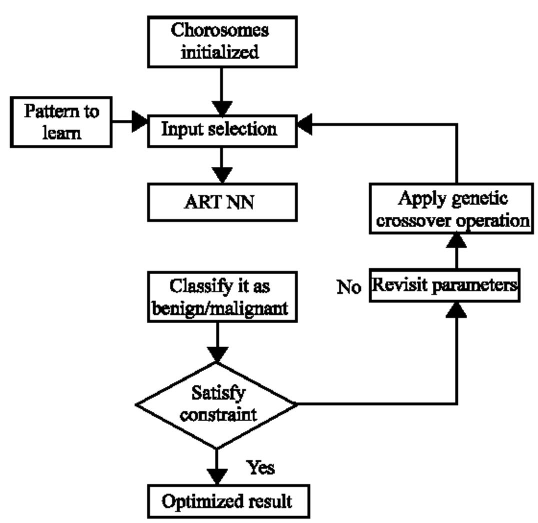

Fig. 2: Flowchart of Genetic ART Network

is adaptive in nature; it is used for trying with various combinations of input. The optimized result is found from the ART network for the new combination of input data.

RESULTS

Experimental results: Genetic algorithm is used to reduce the dimension, in which crossover is employed as a technique to extract a combination from the parents selected from the generated population. The crossover rate is chosen as 0.5 and the constraint is specified as 93% to maximize the efficiency. Based on the correlation of the inputs with the output and with due consideration on the threshold value (71%) the number of essential input features were reduced to five.

The data set consists of 699 samples taken by using Fine Needle Aspirates (FNA) from patient's breast tissue (Mera and Murphy, 1991) with 16 missing data. The 683 samples with breast tumor is used in this work, of which 444 (65%) proved to be benign and 239 (35%) malignant. A patient record from a breast FNA lists nine cytological characteristics labeled as x1, x2…x9. They are clump thickness, uniformity of cell size, cell shape, marginal adhesion, single epithelial cell size, bare nuclei, bland chromatin, normal nucleoli and mitosis. These attributes measure the external appearance and the internal chromosome changes in nine different scales which are assigned integer values ranging from 1 to 10 with 1 being the closest to benign and 10 the most malignant, with two class variables namely benign and malignant represented numerically by 2 and 4 respectively. ART network used in this study accepts only binary values; hence the real attribute values are transformed into binary using the following scheme.

$$X_i = \begin{Bmatrix} 0, X_i < 5 \\ 1, X_i > 5 \end{Bmatrix}$$

ART implemented in MATLAB 6.0 (a technical computing language) is used to classify the cancer, whether it is benign or malignant (Sivanandam *et al.*, 2006). The vigilance parameter is set as 0.6 for training the network. 50% of each data type is used for training the ART network and the remaining 50% for testing the performance of the network. The outcome of the proposed network which uses only five attributes namely x1, x2, x3, x4 and x8 instead of all the nine attributes employed in (Tuba and Tulay, 2004) has been detailed in Table 1. It is clearly very evident from the table that ART demonstrates a relatively good performance.

Table 1: Comparison of RBF, PNN, MLP and GRNN [8] with the proposed ART network

Cancer type	RBF (%)	PNN (%)	MLP (%)	GRNN (%)	ART
(%)Benign	96.85	98.65	95.50	98.80	98.65
Malignant	94.96	94.12	95.79	96.67	96.64

CONCLUSION

This study explored the possibility of combining the features of genetic algorithms and neural network to classify the breast cancer. The genetic algorithm reduces the dimension and ART classifies with the reduced features. The performance of the combined approach has not only improved the accuracy but also reduced the time taken to train the network. The above-mentioned novel approach produces excellent results, which are superior to RBF, PNN and MLP networks. It will be sensible, to experiment the breast cancer detection with the rest of the networks in the ART family.

REFERENCES

Arnold, F. and J. Shapiro, 1999. Scott Pflumm and Thomas A. DeFilippo, Penn State University-The Inner workings of neural networks and genetic algorithms, actuarial research clearing house, pp: 1.

Brause, R.W., 2001. Medical Analysis and Diagnosis by Neural networks, Computer Sciences Department Franfurt A.M., Germany.

Chen, Y., A. Abraham and B. Yang, 2005. Hybrid neurocomputing for breast cancer detection . The 4th International workshop on soft computing as Transdisciplinary Sci. Tech. (WSTST), pp: 884-892.

David, E., Goldberg-Genetic Algorithms in Search, Optimization and Machine Learning.

Desilva, C.J.S. *et al.*, 1994. Artificial neural networks and breast cancer prognosis, The Aus. Computer J., 26: 78-81.

Gail, A. Carpenter and Stephen Grossberg Adaptive Resonance Theory.

Gail, A., Carpenter and L. Boriana Milenova Department of Cognitive and Neural Systems, Boston University Boston, Massachusetts, USA ART Neural Networks for Medical Data Analysis and Fast Distributed Learning.

King *et al.*, 2006. Evolving Hierarchical RBF neural networks for breast cancer detection, ICONIP, Part III, LNCS 4234, pp: 137-144.

Lanzarini, Vargas and Estelrrich, De Giusti, 1997. Real time analysis of the Nystagmus and Movement patterns in Balance Disturbances, 19th International Conference on Information Technology Interfaces, Croatia.

Limin Fu, 1994. Neural Networks in Computer Intelligence, McGraw Hill International Edition.

Merz, J. and P.M. Murphy, UCI Repositry of Machine learning Databases, http://www.ics.uci.edu/~mlearn

Pataya Dangprasert, Somsong Chiraphadhanakul- Pattern Extracting Engine Using Genetic Algorithms.

Pena-Reyes, C.A. and M.Sipper, 2001. A cooperative coevolutionary approach to fuzzy modeling, IEEE Trans. Fuzzy Sys., 9: 727-737.

Setiono, R., 2000. Generating concise and accurate classification rules for breast cancer diagnosis, Artificial Intelligence in Med., 18: 205-217.

Sivanandam, S.N. S. Sumathi, S.N. Deepa, 2006. Introduction to Neural Networks using Matlab 6.0, Tata McGraw Hill Publications.

Talib, S. and T.S. Hussain, 1994. Methods of Combining Neural Networks and Genetic Algorithms, International Journal of Neural Systems. Andrew, D.B. and Andrzej Cichocki, 1999. Input variable selection using Independent Component Analysis and higher order statistics, proceedings of the first International workshop on independent component analysis.

The Weekend Australia, 2002. Health Section, pp: 7.

Tuba, K. and T. Yildram, 2004. Breast Cancer Diagnosis using statistical neural networks, Istanbul Univ. J. Elect. Electron. Eng., 4: 1149-1153.

Wasserman and D. Philip, 1937. Neural Computing: theory and practice, Library of Congress Cataloging-in-Publication Data.

YongSeog Kim: Customer Targeting: A Neural Network Approach Guided by Genetic Algorithms-USA ykim@b202.usu.edu.