Results Evaluation of Medical Diagnosis System Using Back propagation Algorithm

Arti Gupta
(arti_gupta1984@yahoo.com)

Maneesh Shreevastava
(maneesh.shreevastava@yahoo.com)

Lakshmi Narain College of Technology (LNCT,Bhopal)

***Abstract*: Artificial Intelligence has always lent a helping hand to the practitioners of medicine for improving medical diagnosis and treatment then, paradigm of artificial neural networks is shortly introduced and the main problems of medical data base and the basic approaches for training and testing a network by medical data are described. A lot of Applications tried to help human experts, offering a solution. In this paper, Neural Network based Proposed diagnosis system has the capabilities to take intelligent decision based on given symptoms of diseases takes as a input and outcomes shows the bacterial infection level.The back propagation algorithm presented in this paper used for training depends on a multilayer neural network with a very small learning rate, especially when using a large training set size. It can be applied in a generic manner for any network size that uses a back propagation algorithm and achieved the best performance with the minimum epoch (training iterations) and training time.**

Keywords*: *Artificial Neural Network, Back propagation algorithm, Medical Diagnosis, Neural Networks

1. Introduction

Artificial neural networks [2] provide a powerful tool to help doctors to analyze, model and make sense of complex clinical data across a broad range of medical applications. Most applications of artificial neural networks to medicine are classification problems; that is, the task is on the basis of the measured features to assign the patient to one of a small set of classes.

Doctors use a combination of a patient's case history and current symptoms to reach a health diagnosis when a patient is ill. In order to recognize the combination of symptoms and history that points to a particular disease, the doctor's brain accesses memory of previous patients, as well as information that has been learned from books or other doctors. A neural network has the ability to mimic this type of decision-making process, and use a knowledge base of information, and a training set of practice cases, to learn to diagnose diseases.

Rule based expert systems [3] have been used for medical diagnosis but they have their share of problems. If the system consists of several thousand rules, it takes a very powerful control program to produce any conclusions in a reasonable amount of time. Another problem with rule based systems is that as the number of rules increases, the conflict set also becomes large so a good conflict resolving algorithm is needed if the system is to be usable.

The major problem in medical field is to diagnose disease. Human being always make mistake 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 Giusti, 1999).

Brause(2001) highlighted that almost all the physicians are confronted during their formation by the task of learning to diagnose. Here, they have to solve the problem of deducing certain diseases or formulating a treatment based on more or less specified observations and knowledge. For this task, certain basic difficulties have to be taken into account:-

- 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.
- The quality of diagnosis is totally depends on the physician talent as well as his/her experience.
- Emotional problems and fatigue degrade the doctor's performance.
- The training procedure of doctors, in particular specialists, is a lengthily and expensive one. So even in developed countries we may feel the lack of MDs.

- Medical science is one of the most rapidly growing and changing fields of science. New results disqualify the older treats, new cures and new drugs are introduced day by day. Even unknown diseases turn up every now and then. So a physician should always try hard to keep his/herself up to date

Regarding problems above and also, the question would be how computers can help in medical diagnosis. A computer system never gets tired or bored, can be updated easily in a matter of seconds, and is rather cheap and can be easily distributed. Again, a good percentage of visitors of a clinic are not sick or at least their problem is not serious, if an intelligent diagnosis system can refine that percentage, it will set the doctors free to focus on nuclear and more serious cases.

Neural Networks are ideal in recognizing diseases using scans since there is no need to provide a specific algorithm on how to identify the disease. Neural networks learn by example so the details of how to recognize the disease are not needed. What is needed is a group of examples that are representative of all the variations of the disease.

In this paper, The back propagation algorithm presented in this paper used for training depends on a multilayer neural network with a very small learning rate, especially when using a large training set size. It can be applied in a generic manner for any network size that uses a back propagation algorithm through an optimal time (seen time). e achieved the best performance (i.e. the least mean square error) with the minimum epoch (training iterations) and training time using the Back propagation algorithm

2. What is a Artificial Neural Network

Artificial neural networks are developed based on brain structure. Like the brain, artificial neural networks can recognize patterns, manage data and learn. They are made by artificial neurons which implement the essence of biological neurons.

- It receives a number of inputs (from original data or from output of other related neurons) [5]. Each input comes via a connection, which is called synapses and which has a weight (coefficient of connectivity). A neuron also has a threshold value. If the sum of the weights is bigger than this value, than the neuron is activated.
- The activation signal produces the output of the neuron. This output can be the result of the problem or can be considered an input for another neuron. To create an artificial neural network is necessary to put together a number of neurons. They are arranged on layers. A network has to have an input layer (which carries the values of outside variables) and an output layer (the predictions or the result). Inputs and outputs correspond to sensory and motor nerves from human body There also can be hidden layer(s) of neurons, which play an internal role in the network. All these neurons are connected together

3. How are Used Neural Networks in Medicine

ANNs are very useful for analyzing complex problems where the relationships between input and output data are not very well known, such as pattern and speech recognition, machine vision, robotics, signal processing and optimization.

Artificial neural networks have been successfully applied on various areas of medicine, such as: diagnostic systems, biomedical analysis, image analysis, drug development. Using artificial neural networks, it can be monitored a lot of health indices (respiration rate, blood pressure, glucose level) or can be predicted the patient response to a therapy. Artificial neural networks have a very important role in image analysis, too, being used together with processing of digital image in recognition and classification. They are used in pattern recognition because of their capacity to learn and to store knowledge. The medical image field is very important because it offers a lot of useful information for diagnosis and therapy [6].

An ANN is a highly interconnected network made of many simple processors. Each processor in the network maintains only one piece of dynamic information and is capable of only a few simple computations.

An ANN performs computations by propagating changes in activation between the processors [4]. Using the ANN we can acquire, store and use the knowledge extracted from experts or experiments. The knowledge is kept in a steady state net of relationships between individual neurons and can be updated automatically using some kind of learning Algorithm.

A net contains many paths, which are activated, to a certain degree, by the input vector. The signals generated are propagated and combined through the various layers of the ANN, stimulating the various neurons, and finally generating the output signals [1].

4. Related Works

The current work focuses on the fuzzification and defuzzification of patient data [7]. Since data from the patient are nothing but physiological measures, they are subjected to noise and uncertainty. The data from the patient such as height or weight data cannot always be trusted as they are subjected to the quality and accuracy of measuring units and the skill of the technician. Moreover, based on a single data, it would be highly uncertain to make an accurate decision about the future physiological state of the patient. So the patient data has been fuzzified with the objective of transformation of periodic measures into likelihoods that the Body Mass Index, blood glucose, urea, creatinine, systolic and diastolic blood pressure of the patient is high, low or moderate.

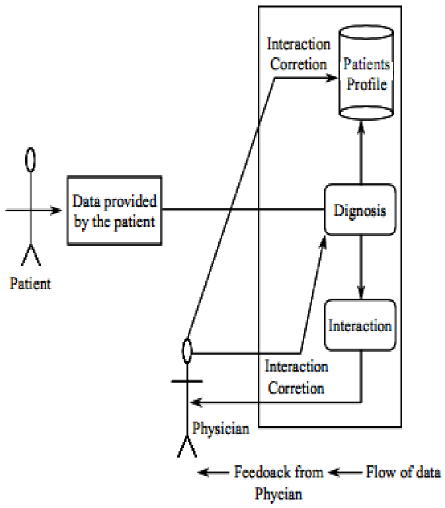

Fig 1: Medical Diagnosis System

5. Proposed Approach

In this section, the theoretical background of the Back propagation learning algorithm pertaining to our study is reviewed.

5.1 Standard Back propagation algorithm:

BP is one of the simplest and most general methods for the supervised training of MLP (Duda et al., 2001). The basic BP algorithm (Bishop, 1995; Duda et al., 2001) works as follows:

1. Initialize all the connection weights W with small random values from a pseudorandom sequence generator.

2. Repeat until convergence (either when the error E is below a preset value or until the gradient vE(t)/vW is smaller than a preset value).

2.1 Compute the update using $\Delta W(t) = -\eta \frac{\partial E(t)}{\partial W}$

2.2 Update the weights with

$$W(t+1) = W(t) + \Delta W(t)$$

2.3 Compute the error E(t+1).

Where t is the iteration number, W is the connection weight, and h is the learning rate. The error E can be chosen as the mean square error (MSE) function between the actual output yj and the desired output dj:

$$E = \frac{1}{2}\sum_{j=1}^{n_j}(d_j - y_j)^2$$

The BP algorithm described above has some shortcomings. If the learning rate is set small enough to minimize the total error, the learning process will be slowed down. On the other hand, a larger learning rate may speed up learning process at the risk of potential oscillation. Another problem is that, partial minimal points or stable stages on error surface are often encountered during the learning process (Baba, 1989).

Using a momentum term is the simplest method to avoid oscillation problems during the search for the minimum value on the error surface. The weight update in BP algorithm with a momentum term a is defined as follows:

$$\Delta W(t) = -\eta \frac{\partial E(t)}{\partial W} + \alpha \Delta W(t-1)$$

The adaptive learning rate can also be adopted to speed up the convergence of the algorithm. For batch training strategy, the learning rate can be adjusted as follows

$$\eta(t) = \begin{cases} \beta\eta(t-1) & \text{if } E(t) < E(t-1) \\ \theta\eta(t-1) & \text{if } E(t) > kE(t-1) \\ \eta(t-1) & \text{otherwise} \end{cases}$$

where h(t) is the learning rate at the tth iteration, and β,Θ and k are chosen as such that

$\beta > 1,\ 0 < \theta < 1,$ and $k > 1.$ While for the incremental training strategy, learning rate can be updated using

$$\eta(t) = \eta_0 + \lambda E(t-1)$$

The learning algorithm with forgetting mechanics is an algorithm that can 'forget' unused connections (Takeshi, 2001). With this forgetting mechanism, the weights that are not reinforced by learning will disappear. The obtained network, thus, has a skeletal structure that reflects the regularity contained in the data, useful to improve the convergence and the network accuracy. In general, the updating of connection weights with forgetting mechanics term is given by

$$\Delta W'(t) = \Delta W(t) - \varepsilon\,\mathrm{sgn}(W(t))$$

Where ε is the amount for the forgetting, and sgn(x) is the sign function. The absolute value of connection weight is set to decrease by ε due to the second term on the right-hand side). In practice, some optimization algorithms are often used to improve the network convergence (Gill et al., 1981), such as the steepest descent method, the Newton method, In practice, some optimization algorithms are often used to improve the network convergence (Gill et al., 1981), such as the steepest descent method, the Newton method, the Quasi–Newton method, and the conjugate gradients method. In this study, the conjugate gradients method is adopted, as it has a low computation cost and exhibits good results (Polak, 1971). The connection weights thus can be expressed by:

$$W(t+1) = W(t) + \eta(t)d(t)$$

$$d(t) = -\nabla E[W(t)] + \beta(t)d(t-1)$$

$$d(0) = -\nabla E[W(0)]$$

Where ∇ E is the gradient, d(t) is conjugate gradient, h(t) is the step wide, b(t) is determined given by Polak–Ribiere function shown

6. Experiment work

This section performs an experiment to see which effect optimal the back propagation algorithm in Medial diagnosis system.. and compared to fuzzy based neural technique The proposed algorithm starts with minimal number of hidden units in the single hidden layer; additional units are added to the hidden layer one at a time to improve the accuracy of the network and to get an optimal size of a neural network. The optimal back propagation algorithm was tested on several benchmarking classification problems including the cancer, heart disease and diabetes. Experimental results show that the back propagation algorithm can produce optimal neural network architecture with good generalization ability.

6.1 Data Set Description

Because the goal of this work is to study and enhance the learning capabilities of the neural network techniques in Medical Diagnosis System, back propagation algorithm has been tested using a data base of about 100 medical records collected at the Jawaharlal Nehru Cancer Hospital Bhopal Also each data according to diagnosis of the expert classified with 4 kind conditions .such as a WBC,HGB,PLT, Blood Urea, Creatinine of different male and female patient. In medical data sets, attributes are typically relatively conditionally independent given the class. Physicians try to define conditionally independent attributes Show the bio signal by condition of the user by following table.

S.no	Name	Sex	Date	WBC	HGB	PLT	Blood Urea	Creatinine
1	Ayushi	F	22/2/2010	18.8	12.4	199	25	0.75
2	Kamla	M	2-Nov	9.9	14.9	360	32.01	0.7
3	khilona bai	F	11/1/2010	7.1	11.9	413	24.09	0.81
4	Priyanka	F	23/1/10	0.2	7.7	18	13.49	0.86
5	Anjana	F	2/1/2010	6.9	12	510	15.9	0.59
6	SN singh	M	3/1/2010	18.3	12.1	348	104.5	1.42
7	Shaikh	M	28/4/10	7.8	14.8	241	19	1.2
8	Taranai	M	4/1/2010	6.6	14.5	264	14.6	0.64
9	Rizwan	M	25/1/10	6.1	12.8	205	18	0.89
10	Narayan Prasad	M	4/1/2010	8.1	13.5	220	25.7	0.91
11	Shri Prasad	M	2/1/2010	7.8	15.2	338	17	0.95

Table.1 Patient data set

6.2 Empirical Setting

The optimal Back propagation algorithms are written in visual basic 6.0 as front-end and MS-Access used as Backend and compiled into mix files. Back propagation algorithms are relatively efficient due to vectored programming and active optimization. All experiments are run on a PC with a 3.06GHz Pentium-4 CPU with 1GB DRAM and running Windows XP. For the Neural Gas Algorithm, the learning rate follows m = 0.5.

Our experiments show that the physicians prefer explanations as provided by the Bayesian classifiers and he physicians found that the combination of classifiers was the appropriate way of improving the reliability and comprehensibility of diagnostic systems. The combination should be done in an appropriate way and the reliability of each classifier on the given new case should be taken into account, as the results clearly demonstrate.

6.3 Experimental Results

In this work, various architectures of ANNs, with three and four layers, were tested. The differences were created varying the number of neurons in the intermediary hidden layers, searching for equilibrium between estimate accuracy and computational cost. Different net configurations were tested, with up to 16 neurons in the first intermediary layer and 48 neurons in the second intermediary layer.

The input parameters included several important concrete characteristics, such as WBC,HGB,PLT, Blood Urea, Creatinine of different male and female patient.Each network was trained using the optimal back-propagation algorithm, which tries to minimize the mean square error between the network output and the corresponding target values. Each training iteration is normally called an epoch. The training was limited on 10.000 epochs. After each iteration, the network explores the error surface searching for the greater gradient of reduction in the mean square error.

The weights and biases are then adjusted to decrease the error. The initial weights and biases for each neural network were generated automatically by the program. In optimal back propagation algorithm, the training procedure is repeated *k* times, each time with 80% of the samples in the dataset as training and left 20% for testing. The 20% testing section is non-overlapping. All the reported results are obtained by averaging the outcomes of those five separate tests. The best diagnosis performance achieved was 88.5% correct classification with 30 nodes in the hidden layer and after passing 1500 training epoch. Next we used a optimal back propagation algorithm on our dataset to convert it to a set of symptoms . A linear membership function was selected for each symptom again after an interview with physicians. Normally three to five linguistic variables were assigned to each symptom, then the classification tests were repeated.

Technique	No. of Hidden unit	No. Of Epochs	Error
Optimal Back propagation algorithm	1	100	0.0157
	2	200	0.013
	3	300	0.0129
	4	400	0.008

Table 2 Experimental results on data sets

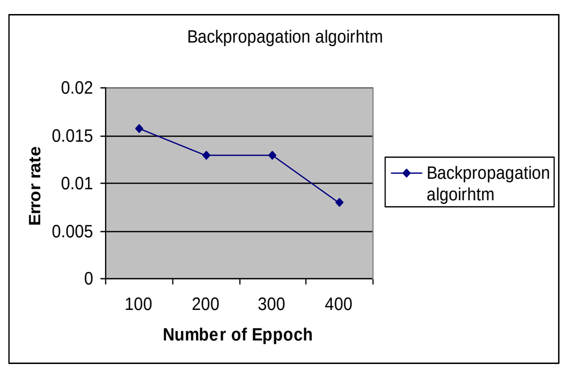

Figure.2: Graph between Error and Epoch relative to the Diagnosis System Tests based on Back propagation algorithm

7. Conclusion

In this paper, we have presented a medical decision support system based on the neural network architecture for Medical diagnosis. The system is trained by employing an improved BP algorithm. The hidden layer of a neural network plays an important role for detecting the relevant features. Due to the existence of irrelevant and redundant attributes, by selecting only the relevant attributes, higher predictive accuracy can be achieved. For a particular input, any (or few) feature(s) may not be effective to the hidden layer or feature space. By extracting this (these) features we can minimize the training time. In near future, we will try to extend the algorithm for improving back propagation using feature selection

8. References

[1] Long, W. J., Naimi, S., & Criscitello, M. G. (2005). Development of a knowledge base for diagnostic reasoning in cardiology. Computers in Biomedical Research, 25, 292–311.

[2] Azuaje, F., Dubitzky, W., Lopes, P., Black, N., & Adamsom, K. (2007). Predicting coronary disease risk based on short-term RR interval measurements: A neural network approach. Artificial Intelligence in Medicine, 15, 275–297.

[3] Tkacz, E. J., & Kostka, P. (2000). An application of wavelet neural network for classification patients with coronary artery disease based on HRV analysis. Proceedings of the Annual International Conference on IEEE Engineering in Medicine and Biology ,1391–1393.

[4] Reategui, E. B., Campbell, J. A., & Leao, B. F. (1997). Combining a neural network with case-based reasoning in a diagnostic system. Artificial Intelligence in Medicine, 9, 5–27.

[5] Tsai, D. Y., & Watanabe, S. (1998). Method optimization of fuzzy reasoning by genetic algorithms and its application to discrimination of myocardial heart disease. Proceedings of IEEE Nuclear Science Symposium and Medical Imaging Conference ,1756–1761. http://www.ifcc.org/ejifcc/vol14no2/140206200308n.htm

[6] Fuster, V., Alexander, R. W., & O'Rourke, R. A. (2001). Hurst's the heart. USA: McGraw Hill Professional.

[7] S. Moein, S. A. Monadjemi, & P. Moallem (2008) "A Novel Fuzzy-Neural Based Medical Diagnosis System", World Academy of Science, Engineering and Technology 37 2008.