International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970)
Volume-2 Number-1 Issue-3 March-2012

Performance of ANN using Back Propagation Algorithm for Medical Diagnosis System

Arti Gupta[1], Manish Shrivastava[2]
Department of Information Technology, Lakshmi Narain College of Technology
Bhopal, (M.P.), India[1,2]

Abstract

Now a day's application of ANN techniques provide a powerful tool to help doctors to analyze, complex clinical data across a broad range of medical application. 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. 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. The system is trained by employing an improved BP algorithm for detecting the relevant features of the clinical data.

Keywords

Artificial Neural Network, Back propagation algorithm, Medical Diagnosis, Rule based expert systems.

1. Introduction

Artificial neural networks [1] 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.

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.

In [2] 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.

2. 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 [2].

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 [3]. 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].

3. Related Works

The current work focuses on the fuzzification and defuzzification of patient data [4]. 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, keratinize, systolic and diastolic blood pressure of the patient is high, low or moderate. Rule based expert systems [5] 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.

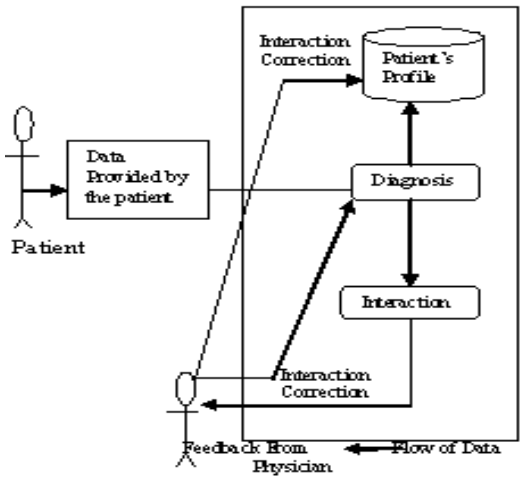

Fig 1: Medical Diagnosis System

Two major classes of expert systems are those based on rules, known as rule-based expert systems, and those based on probabilistic graphical models, often referred to as probabilistic expert systems or normative systems. Rule-based ex-pert systems, originating from the pioneering work of Buchanan and Shortliffeon the Mycin system [5], aim at capturing human expertise in terms of rules of the form if condition then action. There is overwhelming psychological evidence that such rules are capable of modeling the human thought process. A set of rules can capture a human expert's relevant knowledge of a domain and can be subsequently used to reproduce the expert's problem solving in that do-main. Probabilistic expert systems originate from research at the intersection of statistics and artificial intelligence. This research focuses on the concepts of relevance and probabilistic independence and has led to the development of intuitive and efficient graphical tools for knowledge representation.

4. Proposed Technique

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

Standard Back propagation algorithm:
A Back Propagation network learns by example. You give the algorithm examples of what you want the network to do and it changes the network's eights so that, when training is finished, it will give you the required output for a particular input. Back Propagation networks are ideal for simple Pattern Recognition and Mapping Tasks4.As just mentioned, to train the network you need to give it examples of what you want – the output you want (called the *Target*) for a particular input.

Steps of Back Propagation Algorithm
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).

(1) Compute the update using

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

(2) Update the weights with

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

(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 y_j and the desired output d_j:

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

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.

5. Experiment Work

This thesis performs an experiment to see which effect the optimal back propagation algorithm in Medical diagnosis system. The 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 data set classification problems including the WBC, HGB, PLT, Blood Urea, Cretinine etc. Experimental results show that the optimal back propagation algorithm can produce optimal neural network architecture with good generalization ability. Because the goal of this work is to study and enhance the learning capabilities of the neural network techniques in Medical Diagnosis System, optimal back propagation algorithm has been tested using a data base of about 100 medical records collected at the Jawaharlal Nehru Cancer Hospital Bhopal. Renal data such as person age in terms of years, male / female, WBC, HGB, PLT, Blood Urea, Cretinine have been collected for 1000 patients In medical data sets, attributes are typically relatively conditionally independent given the class. Physicians try to define conditionally independent attributes.

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, and then the classification tests were repeated.

6. Conclusion

Fuzzy logic is not a cure-all. Fuzzy logic is the modification of common sense - use common sense when it is implemented and the right decision will probably been made. Many controllers do a fine job without using fuzzy logic. However, if time is taken to become familiar with fuzzy logic, a very powerful tool can be proved for dealing quickly and efficiently with imprecision and nonlinearity. Rule-based systems capture heuristic knowledge from the experts and allow for a direct construction of a classification relation, while probabilistic systems capture causal dependencies, based on knowledge of pathophysiology, and enhance them with statistical relations. 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.

References

[1] Robert Scott, "Artificial Intelligence: It's Use in Medical Diagnosis", The Journal of Nuclear Medicine, Vol. 34, pp 510-514, March 1993.

[2] Rüdiger W. Brause, "Medical Analysis and Diagnosis by Neural Networks", ISMDA, pp 1-13, 2001.

[3] Tkacz, E. J., & Kostka, P." 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, pp 1391-1393, 2000.

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

[5] B.G. Buchanan and E.H. Shortlie (Eds.), "Rule –Based Expert Systems: The MYCIN" Experiments of the Stanford Heuristic Programming Project. Addison- Wesley, Reading, MA, 1984.

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

Arti Gupta born in Bhopal on 4th of August 1983. Completed Schooling from Digambar Jain H.S School in 2001. Completed B.E. from Lakshmi Narain College of Technology in 2007. Pursuing M. tech in Information Technology from Lakshmi Narayan College of Technology.

Dr. Manish Shrivastava had started his carrier with Software companies and worked for SIS (P) Ltd. and TCS at PMU, after completing his graduation in Computer Technology from UIT, RGPV (Technical University) (formerly Govt. Engineering College), Bhopal, India in 1993. He left software industry in 1998 & switched to academics and has been working with reputed private engineering colleges since last fourteen years. He did M.Tech. & PhD from MANIT, Bhopal, India in Digital Communications & Optical Communications respectively. He has software development, teaching and research experience of more than 18 years. Presently he is working as Director, PG Education & Research Center, LNCT, and Bhopal, India.