Indian J Physiol Pharmacol 1998; 42 (3) 321–342

REVIEW ARTICLE

ARTIFICIAL NEURAL NETWORK AND MEDICINE

ZULFUQUAR HOSSAIN KHAN*, SAROJ KANT MOHAPATRA**, PRAFULLA KUMAR KHODIAR** AND S. N. RAGU KUMAR***

**Department of Physiology, Santosh World Medical Academy, Santosh Nagar, Pratap Vihar, Ghaziabad (U.P.)*

***Department of Biochemistry, Maulna Azad Medical College, Bahadur Shah Zafar Marg, New Delhi – 110 002*

and

****Computer Facility, All India Institute of Medical Sciences, Ansari Nagar, New Delhi – 110 029*

(Received on July 5, 1997)

Abstract : The introduction of human brain functions such as perception and cognition into the computer has been made possible by the use of Artificial Neural Network (ANN). ANN are computer models inspired by the structure and behavior of neurons. Like the brain, ANN can recognize patterns, manage data and most significantly, learn. This learning ability, not seen in other computer models simulating human intelligence, constantly improves its functional accuracy as it keeps on performing. Experience is as important for an ANN as it is for man. It is being increasingly used to supplement and even (may be) replace experts, in medicine. However, there is still scope for improvement in some areas. Its ability to classify and interpret various forms of medical data comes as a helping hand to clinical decision making in both diagnosis and treatment. Treatment planning in medicine, radiotherapy, rehabilitation, etc. is being done using ANN. Morbidity and mortality prediction by ANN in different medical situations can be very helpful for hospital management. ANN has a promising future in fundamental research, medical education and surgical robotics.

Key words : neural networks; artificial neural networks; computer; medicine

INTRODUCTION

Many instruments used in medical investigations and therapeutic interventions depend upon digital computation. Modern computers are based on central processors that add and multiply binary numbers. Tasks such as calculation, word processing, image processing and robot control are done by translating them into binary operations.

***Correspondence and Request for Reprints : S. K. Mohapatra, C/o Mr. S. N. Ragu Kumar, Application Programmer, Computer Facility, All India Institute of Medical Sciences, Ansari Nagar, New Delhi–110 029

But these uses of computer, impressive as they are, may gradually appear less important. In the quest to understand and duplicate the elements that constitute human intelligence computer technology has taken a bold step ahead. From just being able to perform complex tasks much better and faster than the brain, the computer is now making attempts to emulate capabilities of human intelligence such as pattern recognition, vision, innovation and creativity. It is steadily evolving from the status of a machine to that of a responsible and respectable partner. Attempts are going on to make the computer capable of not only assisting man in doing certain tasks but doing that job itself independent of human beings. This level of performance is becoming increasingly possible by the use of ARTIFICIAL NEURAL NETWORK (ANN). In this paper we have attempted to explain how ANN and some facets of its functioning. Use of ANN in different fields of medicine has also been dealt with.

Whenever cognitive abilities of computer are discussed, doctors usually think of software for clinical diagnosis or interpretation of investigative data. But such software has often been found to be unsatisfactory for the needs of doctors. Most of the programs available with doctors are Expert Systems (ES). Now ANN is increasingly proving its promise as a better alternative in medical diagnosis and treatment. In the subsequent few paragraphs it has been attempted to explain how the expert system functions better than the conventional algorithm-based programs, and also how ANN is better then expert systems.

Expert systems (knowledge-based systems) also simulate an aspect of human cognition. An expert system is a computer program that embodies "expertise" about something and is supposed to be capable of reasoning like an expert. The knowledge is stored as rules (heuristics) of IF...THEN format. This is based on the assumption that human experts handle problems through such heuristic (rule-based) techniques. Heuristics are rules of the thumb often used by human experts. In contrast, conventional algorithms written for computer programs represent procedures, that, if followed meticulously will eventually produce a solution to the problem, but need more time and processing effort than heuristics. Heuristics are shortcuts which, if they manage to produce a solution, are likely to do so in less time and with less processing. Let us consider a problem and see how a solution can be found using algorithm as well as expert system-based solutions.

Problem : To find the book titled "Neural Network" in a library. [Shelf "A" contains books with titles starting with letter "A" and so on.]

(Two types of solutions are possible. Both are described below.)

Algorithmic solution

Step 1 : Find Shelf A book 1, compare the title with Neural Network. If book is found, stop searching. Otherwise do the next step.

Step 2 : Find Shelf A book 2, compare the title with Neural Network. If book is found stop searching. Otherwise do the next step.

Step 3 : Find Shelf A book 3, compare the title with Neural Network. If book is found stop searching. Otherwise do the next step.

....

Continue till end of shelf Z (last book) or the book 'Neural Network' is found.

Heuristic solution

First, check the books in shelf 'N'.

If not found, then check in 'M' or 'O'.

The rule used is: *If the book "Neural network" is not found in the shelf 'N' then it might be in alphabetically nearby shelves ('M' or 'O')*

To compare an expert system with ANN we will take another problem as an example.

Problem: To identify an apple.

One might give the following rule to an expert system.

IF shape of the subject is round and its color is red and it is a fruit, THEN it is an apple.

The rule may be made more elaborate by including more features of the 'apple' or more rules may be added. After exhausting all reasonable rules, one might safely consider that this expert system contains most of the knowledge regarding an apple.

However, there is a snag. It is possible that some other object has all the characteristics that are specified in the rules, and is yet not an 'apple'. If such an object is presented to this "expert system", shall it not make a mistake and consider this object to be an 'apple'? On the other hand, some special varieties of apple may not have some major features of commonly available apples. Hence while framing rules one needs to include also rare situations, and this task of bringing rare situations under the umbrella of rules becomes a difficult task.

This is the point at which ANN significantly differs from the expert system (ES). ANN does not work on the indirect basis of rules provided to it; it is directly "trained" by examples. For the apple recognition problem, the ANN shall have to come across a number of apples using suitable and adequate number of interfaces. During these exposures (training), the ANN shall extract features that constitute "appleness". In other words, unlike ES where knowledge is made explicit in the form of rules, ANN generates its own 'rules' by learning encounters. Rather than being programmed, the neural networks are 'trained' by presenting sets of inputs. In contrast, ES uses a logically complete set of rules that are followed sequentially to produce an output from a set of inputs. While processing data an ES takes a serial or sequential progression through a number of IF-THEN rules; in contrast, the processing in ANN is of the parallel type analogous to that in the brain. Acquiring the knowledge base for an expert system is a very difficult process. The effectiveness of the expert system is totally dependent upon the completeness and accuracy of the rules. On the other hand, the rules in an ANN are implicit in the training data provided to the network. But still, expertise is needed to ensure correct feeding of data to the network.

Especially in the case of problems that might require a very large number of rules to be handled by an expert system, ANN is a good alternative. For example, using ANN techniques, a speech recognition system NetTalk (1) learnt to recongize an almost unlimited vocabulary from a small training set of words. It has been estimated that this performance would require some 300,000 linguistic rules if attempted by an expert system.

Before describing ANN, it would be appropriate to look into some of the most important functional modalities of computer and brain. The computer performs step-wise algorithms (programs) sequentially one after the other and stores information at some particular locations in its memory-unit. The computer is controlled by a complex Central Processing Unit (CPU) which is located on a silicon chip and is capable of arithmetic (addition, subtraction, etc.) and some logical operations (i.e., comparison and discrimination between two given quantities and to decide whether a statement is true or false). On the other hand, the brain relies on highly distributed representations and transformations operating in parallel and appears to store information in variable strength connections called synapses (2). In contrast to the CPU in computer, the brain carries a distributed control through billions of densely interconnected neurons (each neuron may be considered a simple processing unit). It is this "parallel distributed processing" in the brain that sets it apart from any traditional computer. Conventional computer does some tasks (numerical computation, fast repetitive operations) better than the brain, but is incapable of handling information that is fuzzy, probabilistic, noisy, inconsistent, or mutually interacting. Especially in areas requiring pattern recognition ability by interpreting the pattern of signs, symptoms and images (e.g., in diagnosis of diseases) it is a poor substitute. With appreciation of this limitation, the computer scientists took the hint from neurophysiology and started building systems modeled after the structure of brain. The intention was to impart some cognitive abilities to these systems.

The starting point for studying ANN is a model of a cell of the living brain: The Neuron. The model was first suggested by the neurophysiologist Warren McCulloch and the logician Walter Pitts (3). In 1943, they developed a simple model of variable resistors and summing amplifiers that represent the variable synaptic connections and the operation of the neuron body (soma) respectively. It is the influence of this model that sets the mathematical tone of what is being done today.

In the human brain, a typical neuron collects signals from others through a host of fine structures called dendrites. The neuron sends out spikes of electrical activity through a long, thin strand known as axon, which, at its end, splits into thousands of branches. At the end of each branch, a structure called synapse converts the activity from the axon into electrical effects that inhibit or excite activity in the connected neurons. When a neuron receives executory input that is sufficiently large compared with its inhibitory input, it sends a spike of electrical activity down its axon. Learning occurs by changing the effectiveness of the synapses so that the influence of one neuron on another changes (4).

Artificial Neural Networks (ANN) are computer models inspired by the structure and behavior of biological neurons. They go by many other names, such as neural nets, connectionist models, artificial neural systems and parallel distributed systems. Like the brain, ANN can recognize patterns, manage data, and most importantly, learn. An ANN is made up of interconnected processing elements of 'units' (Fig. 1) which represent the bodies of the neurons. These units may be software elements or hardware elements. Considering the detailed biochemical complexity of the living cell, this 'unit' is a rather simplified version of the real neuron. However, the actual power of the ANN derives not from the complexity of individual processing units, but from the interconnectedness of the whole network and the way in which it processes information (Parallel Distributed Processing).

The simplest ANN is the single-layer perceptron (sometimes called two-layer perceptron). This was developed by Frank Rosenblatt in 1958 (5). There are two groups of units: the input units and the output units. Fig. 2 shows a one-layer perceptron containing five input units and one output unit. The input units are directly connected to the output units. The input-units simply transmit the input values (x_i). During transmission the value becomes multiplied by some weightage (w_i) and reaches the output unit(s). It is the output-unit that does

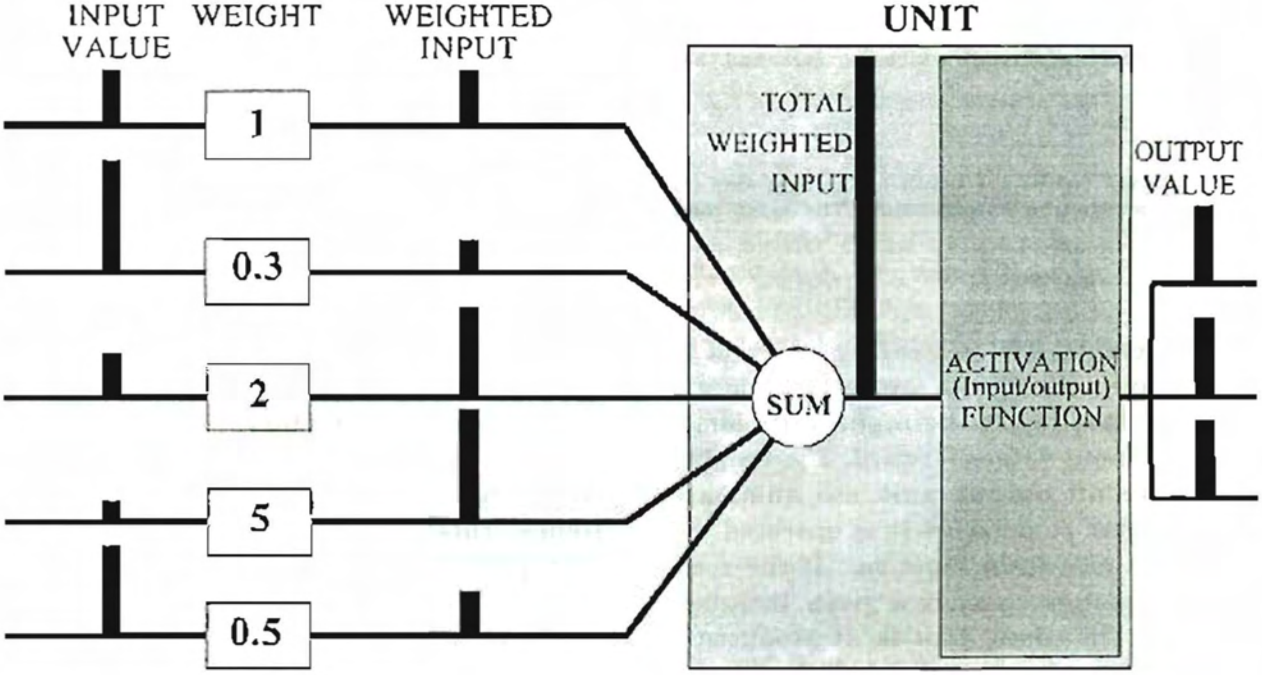

Fig. 1 : An artificial neuron (UNIT) processes signals that are in the form of numerical values. Each input activity (or value) is multiplied by a number called the wieght. The 'unit' adds together the weighted inputs. It then computes the output activity using an input-to-output (activation) function. There are various types of activation functions, the sigmoidal being the most commonly used. The output activity (or value) is passed on to other units (neurons).

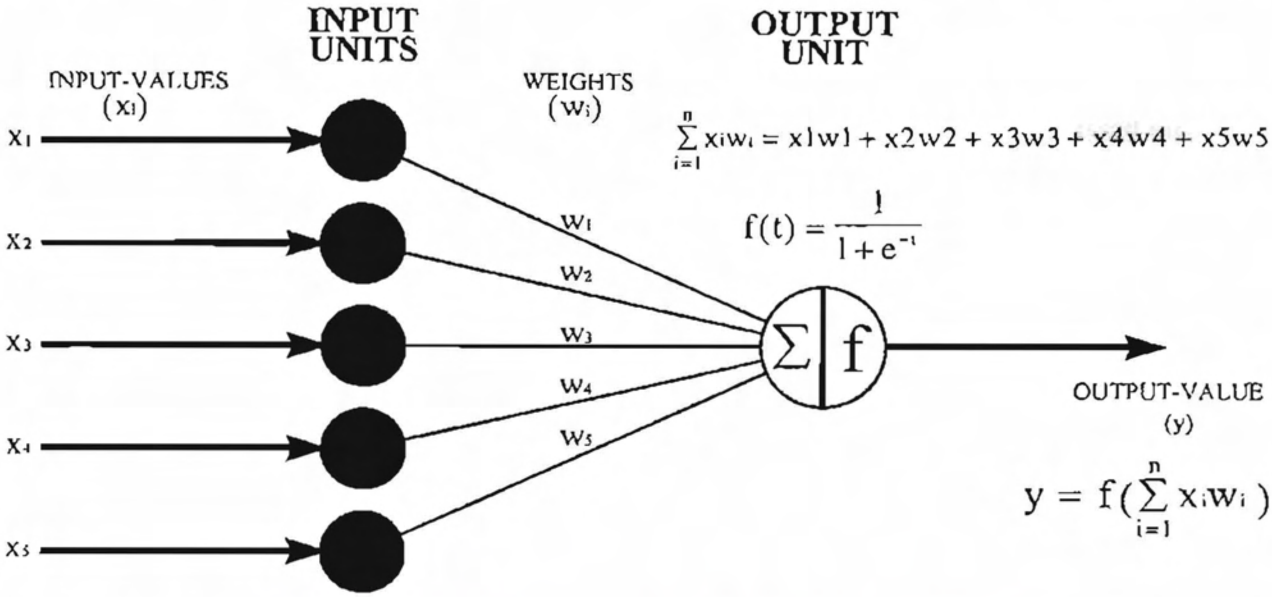

Fig. 2 : A single-layer perceptron containing 5 input-units (shown as dark circles) and 1 output-unit (shown as a white circle). The input-values (x_i) reach the input-unit which pass on these values along the connections. The output-unit does two operations. First, it multiplies each input-value with the weight (w_i) of the connection through which it passes, and then adds up all the weighted inputs to calculate 'total weighted input' $\sum_{i=1}^{n} x_i w_i$.

Next, it runs the result of the first step (total weighted input) through an activation function [f( )] to compute the output-value. The most commonly used activation function is the sigmoidal function $f(t) = 1/(1 + e^{-t})$

the processing. The processing 'units' of the single-layer perceptron are always located in the output layer; the input units simply pass the input values forward. The weighted inputs to an output unit are summated ($w_i x_i$). After summation it is operated by a function (activation function). If the result of this function exceeds a given threshold, then the unit fires, that is, it produces an output. Clearly, the perceptron (as also other types of ANN) can be described as a group of interconnected mathematical equations that accept input data and calculate an output based on this input. This single-layer perceptron can only decide whether a set of input data belongs to any one of the two categories (linear discrimination). Unfortunately, such a scenario where data can be separated by a linear function is rare in medicine (6). Hence, this simple network is not much useful.

A multilayer perceptron is the next choice. This was developed by Rumelhart, Hinton and Williams in 1986 (7). In the simplest form, it contains, in addition to the input and output units, another layer (hidden layer) that consists of one or more units (Fig. 3). This layer is placed between the input and the output layers. The input units are now connected to the hidden layer, which in turn is connected to the output units. The input units pass the weighted values to the hidden units(s) for local processing. In this respect a typical unit of the hidden layer is similar to a typical unit of the output layer: both (1)sum up the weighted inputs, (2) operate that sum through a function, and (3) compare it with a threshold value. A more complex multilayer perceptron may contain more than one hidden layer. The optimal number of hidden units/layers required to perform an arbitrary work is generally found by trial and error.

The additional hidden layer(s) in the multilayer networks lead to increased efficiency of the network. For example, a single-layer perceptron is capable of accurately representing AND and OR binary operators but not the exclusive-or (XOR) operator. A 3-layer (multilayer) perceptron is required to model XOR operation.

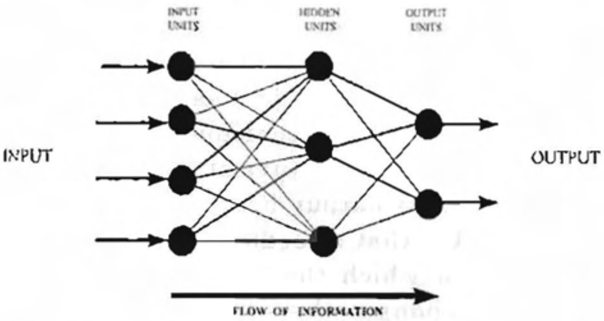

Fig. 3 : A common artificial neural network consists of 3 layers of units (shown as closed dark circles in the diagram) that are fully connected. The network shown here has 4 input-units, 3 hidden-units and 2 output-units. Activity passes from the input-units to the hidden units and finally to the output units. The lines represent the connectivities between the units. The varying thickness of the lines suggests that different connections have different weights.

Fig. 3 shows a 3–layer feed-forward perceptron. Each of the hidden and output units does local processing in exactly the same manner as the output unit of the single-layer perceptron. It then passes on the information as its output.

The input to the input units can be a set of signs, symptoms, images (microscopic, radiographic, tomographic, etc.), waveforms (EEG, ECG, EMG, etc.) numerical values (e.g., serum chemistry data), etc. The expected output can be the diagnosis of a disease, classification of an image (e.g., benign/malignant, ischaemic/non-ischaemic), or outcome prediction (morbidity, mortality). The number of input units depends upon the complexity of the problem under study.

ANNs are specified by the characteristics of the 'unit' (processing element), the network architecture, and the training or learning procedures they follow in order to adapt the weights. The unit might be a software (part of a program, a variable, or a location in memory that stores some information) or hardware element (a full-fledged processor, similar to the CPU of personal computers). ANNs of the latter type are run on special computers in which multiple processors work in parallel. Most ANNs described in medical literature belong to the software category. Even within the software type of 'unit' there can be variations, e.g., the activation function (described above) might be different in different ANNs.

Network architecture falls into two broad classes; feed-forward and feedback (recurrent). Feedback network was defined by Jordan (8). The examples shown in Fig. 2 and Fig. 3 are feed-forward nets as information always proceeds in one direction only (from input units to output units). A feedback or recurrent net is one where information can find its way around a loop from the output back to the input. Fig. 4 and Fig. 5 show simple feedback nets containing loops from the output to one of the input or hidden units. Sometimes such networks have complete connectivity, with very slight distinction made between input, hidden and output units. Adding feedback expands the network's range of behavior, since its output now depends on both the current input and network-states. But one has to pay a price: longer time taken by the ANN to recognize its inputs.

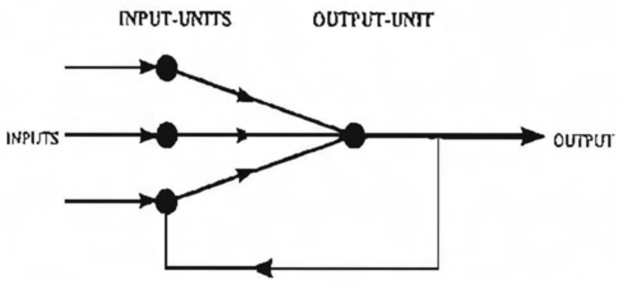

Fig. 4 : A simple feedback network consisting of 3 input-units, 3 connections leading to 1 output-unit and 1 feedback loop where information can flow from the output to one of the input-units. The arrows indicate the direction of flow of information.

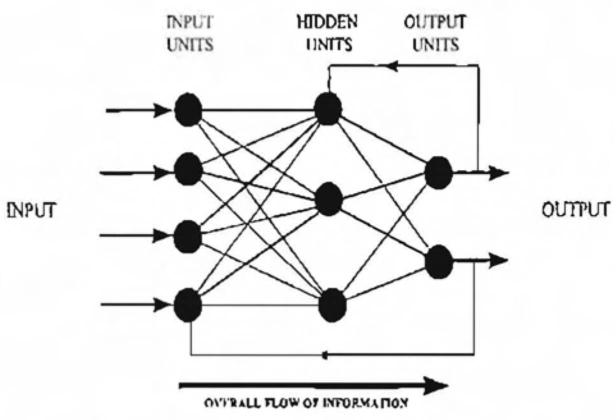

Fig. 5 : A multilayer feedback network consisting of 4 input-units, 3 hidden-units and 2 output-units. In the network shown here, there are 2 feedback loops. One loop goes from the output-unit to one of the input-units. The other loop connects one output-unit to one of the hidden-units. The arrows indicate the direction of flow of information.

Hopfield networks

Feed-forward nets, such as the multilayer perceptron, are always stable: they never enter a mode in which the output is continuously changing. But their behaviour is limited. In case of feedback networks, its input is continuously modified by the previous output. Hence the possibilities of output become much more. It is possible that a feedback net reaches a condition in which the recirculated output no longer changes the network state, that is, it reaches a 'stable' state. However, there are other possibilities, such as cyclic behavior, or chaos, where the network wanders endlessly and unpredictably from one state to another. Predicting which feedback networks would be stable was difficult until the discovery by John Hopfield in 1982 (9, 10). He showed that symmetrically connected feedback (recurrent) nets are not only stable, they are also capable of associative memory, that is, they are capable of completing the information in a stored record in response to the presentation of part of that full record. This is important, since associative memory forms an integral part of human behavior, e.g., recognizing a friend by seeing only a part of his face. These nets are called Hopfield nets.

Supervised learning

ANNs usually operate in one of three modes. Initially, there exists a training phase in which connection-weights are adjusted until the network produces the desired output. After training, the network's performance is tested against sets of inputs whose outputs are known. Testing procedure is important to understand if the ANN has really generalized from the examples instead of simply memorizing them. Only when the test performance is acceptable, it becomes operational, that is, capable of performing the task it was designed and trained to do. The training phase can be either supervised or unsupervised. In supervised learning, there exists information about the correct or desired output for each input training pattern presented. Back-propagation is an example of supervised learning discovered by Rumelhart (7). In this type of learning, ANN is trained on a training set consisting of inputs (the *problems* that the ANN has to solve) and the desired outputs (the ideal *solutions*). Initially the actual computed output of the ANN (the result of all the serial and parallel information flow through the network) may not match with the desired output. In other words, the network fails to find the correct solution in the first attempt. During training, the weights inside the ANN are modified in such a way as to minimize the error between the desired and the computed output of the network (4). This process might take a large number of iterations to converge but, if the training is successful, the network becomes capable of performing the task.

Briefly, a stimulus can be given to the system (ANN) by initially activating a particular selection of input-level units. This activation is propagated throughout the network via executory and inhibitory activation links (connections) until output-level units are activated. If the actual output does not match with the desired (ideal) output, the system adjusts the weights of connections, by using some mathematical learning procedure, in such a way as to produce a better output next time. Over many cycles (iterations), the system will finally settle into a stable state, upon which the training phase is terminated. At this point it could be said that the ANN has finally learnt the task from the given examples. The system can now handle novel inputs, which are different from the examples used to train the system, to produce accurate outputs. This process may be compared with the way a child first learns to hold an object. The child sees the object, tries to grasp it, makes many futile movements of hands, before it finally holds the object. Its performance improves over repeated attempts and finally succeeds. Once it learns, it can repeat the same task with utmost ease.

Learning without supervision

The most popular method of training an ANN (supervised learning using back-propagation) has not been accepted by the neuroscientists as a good model of how the brain actually learns. Their main argument is that the brain works not in a supervised but an unsupervised manner. Nobody presents us with a detailed description of the internal representations of the world that we must learn to extract from our sensory input. We learn to understand sentences or visual scenes without any direct instructions. In the domain of ANN also 'unsupervised learning procedures' have been developed. Many scientists have contributed to this development, one of whom is Kohonen (11). In this type of learning, no prior information exists, and training is based only on the properties of the patterns. Training depends on statistical regularities that the networks extracts from the training set and automatically stores them as connection-weight values of hiden units. Two important mechanisms of unsupervised learning are principal components learning and competitive learning. These two methods are classified according to the type of representation they create of the input data in the hidden units. In principal components methods, the hidden units cooperate, and the representation of each input pattern is distributed across all of them. In competitive methods, the hidden units compete, and the representation of the input pattern is localized in the single hidden unit that is selected.

Self-organizing maps (SOM) (11) are a class of unsupervised neural networks that are used for sorting items into appropriate categories of similar objects. SOM addresses this problem by creating a two-dimensional feature map of the input data in such a way that order is preserved. If two input vectors are close, they will be mapped to 'units' that are close together in the two-dimensional layer that represents the features or clusters of the input data. SOM is used to visualize topologies and hierarchical structures of higher dimensional input spaces.

Unsupervised learning can be used to extract a hierarchy of successively more economical representations. This would lead to greatly improved speed of learning in large multilayer networks. Each layer of the network adapts its incoming weights to make its representation better than the representation in the previous layer, so weights in one layer can be learnt without reference to weights in the subsequent layers. This strategy eliminates many of the interactions between weights that make back-propagation learning very slow in deep multilayer networks.

Other neural network models

Apart from the types of networks described above, there are many other types of ANNs that are structurally and functionally very diverse. Each is unique in some characteristic or other (10).

APPLICATION OF ANN IN MEDICINE

Artificial neural networks resemble the brain to such an extent that they are no longer comparable with traditional algorithm-based programs from which they differ specially in architecture, processing speed and learning ability. Because of the massive parallelism, ANN is able to process information much faster. Because of its learning ability, it can be used in those cases where it is possible to specify the inputs and the outputs but difficult to define the relationships between them. Not only that, it continues to learn while in operation. An ANN is insensitive to disturbing noise. It picks up only the relevant information from 'noisy' data. Such a feature makes ANN suitable for diagnostic purposes where much unnecessary data need to be filtered out (for example, in case of an ill-stained smear or a badly developed radiograph). Lastly, ANNs are 'robust' against failure of individual 'neurons' (units). If one or more 'units' are lost or damaged, the network is still capable of doing reasonably well the task which it had learnt before the damage. This is because the knowledge is not localized in any particular 'unit', but is distributed throughout the network. All these features make ANN more likely to succeed in solving those medical problems which traditional computers cannot.

The field of medicine abounds in problems where signs, symptoms or parameters of one disease overlap with those of another and therefore, traditional statistical techniques (for example, discriminant analysis) do not help (6). The physician has to make a diagnosis on the basis of his/her experience. Could such a human feat (which is not based on mathematical rigor) be emulated by some electronic mechanism? As would be evident from a number of studies by researchers in different areas of medicine, ANN is a plausible answer.

ANN is being used in various fields of medicine. Its performance has been compared with different statistical methods and, more importantly, with that of the experts in the concerned fields. In many situations the performance of ANN is better or at least comparable.

CLINICAL DIAGNOSIS

Clinical diagnosis is an area where ANN approach has been applied for more than 5 years (12). Acute myocardial infarction (AMI) was one of the earliest applications (13). Because of the life-threatening nature of this disease, physicians tend to be overcautious and make a significant number of false positive diagnoses. It has been shown in a multicenter Chest Pain Study that physicians had a sensitivity of 87.78% but specificity of only 71% in diagnosing AMI, while a computer protocol that used Classification and Regression Tree (CART) could achieve a sensitivity of 88% and a specificity of 74% (14). However, better results were obtained by ANN in a study conducted at the emergency department of San Diego Medical Centre. In this study, an ANN was trained on the clinical pattern sets retrospectively derived from the cases of 351 patients hospitalized with high likelihood of having myocardial infarction. It was prospectively tested on 331 consecutive patients presenting with anterior chest pain. The ability of the ANN was compared to that of the physicians in diagnosing AMI. The physicians had a diagnostic sensitivity of 77.7% and a specificity of 84.7%. The ANN had a sensitivity of 97.2% and specificity of 96.2% (13). This initial prospective study was extended to include 1070 patients 18 years or older presenting with anterior chest pain who were evaluated by emergency department physicians and ANN (15). The physicians had a sensitivity of 73.3% and specificity of 81.1%, while the ANN had sensitivity of 96% and specificity of 96%. In this study, unfortunately, no comparison has been made with a standard statistical technique to make it clear that the postulated non-linear pattern recognition ability of the ANN was being used and that its efficacy was better. The high degree of sensitivity and specificity shown by ANN in this study has been disputed by another group (16). They could achieve a sensitivity of over 90% using ANN, only in a patient cohort with a very high incidence of ischaemic heart disease and with modern biochemical markers of cardiac damage. In these two studies the ANN has fared better than the physicians.

Pulmonary embolism (PE) and back pain are two other areas where comparisons have been made between diagnostic efficiencies of human experts and ANN. A study was undertaken to test the hypothesis that computer based pattern recognition can accurately assess the likehood of acute pulmonary embolism from readily obtainable clinical characteristics (17). ANN was found to be as accurate as experienced clinicians. The authors have opined that ANN assessments are reproducible and consistent over time. It can assist clinicians in selecting patients for further investigations (in this case V/Q scan) and their interpretation. Use of neural network may enhance the non-invasive diagnosis of PE by giving physicians an objective estimate of the probability of PE.

IMAGE PROCESSING : Radiology, Histopathology, Cytology

Image interpretation involves two basic steps. The first step is to view the image and compile a list of any abnormalities (the *findings*). The second step involves a cognitive process in which the expert (e.g., the radiologist) makes diagnosis based on the (radiographic) findings. This second step requires ranking the possible diagnosis by probability. The first step is a pattern recognition process; the second step is a classification or interpretation task. ANNs are good at both steps, and, in medical image analysis, are especially used for the second (classification).

Radiological imaging is, thus, a fertile ground for testing the abilities of ANN (18, 19). In an impressive paper (18), Reinus et al describe a feed-forward network (95 input units, 45 output units and no hidden units) which was trained to diagnose focal bone lesions from plain radiographs. This particular ANN topology was selected after evaluating several neural network designs, including those with hidden layers and those without hidden layers. Interestingly, the authors found that this particular ANN (95–0–45) showed best performance. The network learnt by "conjugate-gradient method", a modified method of back-propagation. Information for input to the network was coded according to three basic categories: demographic (age and gender), gross anatomic and structural data. Twenty-two structural characteristics were analyzed and encoded for each patient to define the radiographic appearance of the lesion for the network. Imaging features of 709 lesions were studied that comprised 43 different pathologic diagnoses. Overall, the network was 85.3% accurate in distinguishing benign from malignant lesions. It correctly identified 163 of 217 (75.1%) lesions as malignant lesions and 442 of 492 (89.1%) lesions as benign. It was 56% accurate for the specific diagnosis; and with a nine-lesion differential diagnosis, the network included the correct diagnosis in the list 87.3% of the time. The network showed overall good discrimination between benign and malignant lesions, but failed to identify malignant lesion as malignant nearly 25% of the time. This is unacceptable for clinical practice. The authors opine that even if the sensitivity of the network is not currently adequate for use as a screening device, the human observers might not have performed any better than the ANN. The performance of the network appeared to be more strongly related to how distinctive each pathologic type was radiographically than to the number of cases available for training. It means that while training for the diagnosis of a particular disease, cases as varied as possible (having the same diagnosis) should be fed. The authors have suggested that neural networks may also be used as consultants to practising radiologists and as teaching tools for resident training.

ANN has also been tried for interpreting chest radiographs (19). The input to the ANN was in binary form: for example, if cardiomegaly was present the input unit representing this trait was set to one; if not present the unit was set to zero. The network was trained to choose one or more diagnoses from a list of 12 possible diagnoses, based on 21 radiographic observations made on a series of neonatal chest radiographs. The performance of the trained network was compared with two experienced radiologists. It was found that the network agreed well with each radiologist individually, and this was greater than the agreement between the two radiologists. The authors hope that the subsequent studies will demonstrate superiority of ANN over other health care personnel (general diagnostic radiologists, radiology residents and paediatricians). When such ANNs are sufficiently improved in performance and become available conveniently to the radiologists or clinicians, they appear to enhance the overall analysis of the chest radiographs significantly by increasing the consistency and accuracy of working diagnoses.

Where expert systems have not done so well, ANNs seem to succeed and hence have been used in various domains of imaging including PET (20), SPECT (21) and mammography (22).

In pathology also, picture processing ability of ANN makes it very suitable for use in classification of histology/cytology specimens. The image under the microscope can be easily 'captured' using a video camera and then digitized in the computer (that is, the image is translated into the machine language of computer). From the dizitized data it is possible for a neural network to (i) identify the 'objects' in the slide, (ii) categorize the features of diagnosic significance and (iii) classify the sample. Alternatively, ANN can be used for only the last step, classification, based on the features detected and graded by human observers. In the previously discussed paragraphs on ANN and radiological images, classification was done in some cases by human experts. It must be remembered that how well the network classifies the sample will depend upon correct human observation of object features. Input of poor data shall lead to deterioration of ANN performance. Therefore, to train an ANN properly, good examples have to be fed to it. Secondly, for better learning the examples given to the network for training should be more in number and of varying features. Experience is as important for an ANN as it is for human beings.

The problem of detecting breast cancer from cytologic/histologic data has been dealt with the aid of ANN. ANN classification of breast carcinomas from biopsies on the basis of nuclear morphology has been achieved (23) and the result is at least as good as the Bayesian method (multivariate analysis) of classification. In this study, it was shown that a custom-made neural network made by an experienced researcher worked better than the commercially available program used by less experienced researchers. Cervical (Papanicolaou) smear screening is a task, which due to the sheer load of labor involved, calls of automation. It is very important to eliminate the false negative results; as with each false negative report, one malignant case remains undetected. And the element of human error must be kept in mind while addressing the false-negativity issue, especially the cases where the cytologist repeatedly fails to identify the very few abnormal cells in a smear (24). However, because of many factors including bad smear, large variations inside a normal smear, overlap and clustering of cells, smear-examination could not be subjected to standard rules and hence, could not be automated. Only recently, ANNs have been designed to classify cells of cervical smears. A neural network system called PAPNET is commercially available and has been proved to be useful for assisting the cytopathologists in screening of cervical smear (25, 26).

CLINICAL CHEMISTRY: And other laboratory data management systems

ANN has been successful for prediction of acute myocardial infarctions from serial cardiac enzyme data (27). Reibnegger et al (28) developed a back-propagation network for classifying liver diseases — fatty liver, chronic persistent hepatitis and chronic aggressive hepatitis, based on clinical chemistry data, i.e., urinary neopterin, serum levels of transaminases (AST, ALT) and AST/ALT ratio. None of these lab parameters is individually capable of discriminating all the three diseases. Using all the four variables as input, the network correctly classified 31 out of 42 cases. However, even if a parameter (AST/ALT ratio) was deleted from input-data, the network was able to produce the same level of classification i.e., 31 out of 42. This characteristic property was not seen in case of other statistical methods (discriminant analysis, classification-and-regression tree), suggesting that the neural networks are uniquely capable of extracting important hidden features (in this case, the ALT/AST ratio) from given data.

WAVEFORMS

Waveform data (e.g., ECG, EEG, EMG), being patterns, are amenable to analysis by neural network models. A network was trained on 1107 ECGs from patients with myocardial infarctions (MI) regarding 25 features of the QRS complex from ECG leads V_2, V_3 and V_4. With more than 90% of specificity, the network had a sensitivity of about 68% in detecting MI (29). Another network (feed-forward, back-propagation) (30) was used for solving the problem of diagnostic classification of resting 12-lead ECGs. The network was trained and tested on 3253 ECGs from seven classes of subjects: normal; left-, right-and biventricular hypertrophy; anterior-, inferior- and combined myocardial infarction. Sensitivities and specificities, respectively, were: normal 90.2% and 92.5%, left-ventricular hypertrophy 59.3% and 98.2%, right ventricular hypertrophy 31.2% and 98.9%, biventricular hypertrophy 84.4% and 89%, anterior myocardial infarction 51.6% and 97.8%, inferior myocardial infarction 86.3% and 91.1%, and combined myocardial infarction 47.1% and 95.3%.

The use of ANN was assesed for ECG classification on the bases of 12 ST-T segment features, i.e., ST slope, ST-J amplitude, the positive and negative amplitude of the T wave and 8 time-normalized amplitudes of ST-T segments. After training, the network correctly classified 399 of 500 ST-T segments of the test set, a result comparable with that obtained by a human expert on the same test set (correct classification of 428 out of 500). On comparison with conventional criteria for classification of ST-T segment with ST elevation, it was found that, with similar sensitivity, the ANN performed with a specificity of 97% versus 68% of conventional criteria (31). These results are promising, indicating that neural networks could prove to be useful tools if incorporated into conventional interpretation programs.

A comparison was made between ANN and Glasgow 12 lead ECG interpretation program that uses deterministic logic in the diagnosis of atrial fibrillation (AF) (32). At a particular point in the Glasgow program a decision has to be made as to whether AF or sinus rhythm with supraventricular or ventricular extrasystole is present. The same input parameters used for the deterministic logic at that point were also used to train the neural networks. The result from the test set showed that the neural network has an equivalent performance to well developed deterministic logic. The authors opine, "many years of experience required to produce a section of deterministic logic to report atrial fibrillation can be replaced rather quickly by a neural network that performs essentially in an equivalent manner to the original logic."

High priced labor (qualified neurologists) involved in interpretation of multi-channel EEG data, and the need to reduce the amount of data that must be recorded from each patient, are two important reasons that drive the attempts to automate online (real-time) detection of EEG spikes and detection and prediction of epileptic seizures. The literature describes in detail the application of ANN for EEG analysis (33, 34). Webber et al (34) have developed an ANN for seizure detection from EEG data. It contains 31 input units that represent amplitude, slope, curvature, rhythmicity and frequency components of EEG data in 2 sec epochs. The hidden layer consists of 30 units. The output layer has 8 units representing various patterns of EEG activity (e.g., seizure, muscle, noise, normal). The value of the output unit representing seizure activity is averaged over 3 consecutive epochs and a seizure is declared when that average exceeds 0.65. Among 78 randomly selected files from 50 patients (not in the original training set), the detector declared at least one seizure in 76% of 34 files containing seizures. It declared no seizure in 93% of 44 files not containing seizures. It had an overal false detection rate of only one per hour. Though there is still scope for improvement, it shows that ANN can be utilized to provide a practical approach for automatic, on-line seizure detection. In a previous study (35) a 3 layer feed-forward ANN was developed for detection of epileptiform discharges (EDs). It was trained on EEG records which were identified by an electroencephalographer (EEGer). The study showed that ANNs offer a practical solution for automated, real-time ED detection that uses standard, inexpensive computers (80486, 33 MHz, personal computer), is easily adjustable to individual EEGer style and can produce sensitivities and selectivities similar to those of the EEGers. Anderer et al (36) trained a feed-forward network to analyze 17-electrode EEG patterns and diagnose dementia. ROC analysis showed that the ANN performed better than statistical methods (Z-statistic and discriminant analysis).

An interesting paper by Chen et al (37) describes the application of ANN in electrogastrogram (EGG) which is a noninvasive method of identifying gastric contractions using surface electrodes placed on abdominal skin over the stomach. In the study, gastric contractions detected by the EGG were simultaneously compared with the pressure changes in an intraluminal manometric probe. The waveform data of EGG produced from five patients were used as input to the network (feed-forward with back-propagation learning, which contained 64 input units, 10 hidden units and 2 output units). The network showed an overall accuracy of 92% in identifying gastric contractions from EGG data, a result that is certainly going to interest workers in the fields of basic and clinical gastroenterology.

MICROBIOLOGY

Paralysis mass spectrometry (PMS) is a specialized area of microbiology in which the potential of ANN has been demonstrated (38). PMS involves thermal degradation of a sample and analysis of the resultant fragments to identify the sample. Chun et al (38) trained a network to differentiate three different species of *Streptomycetes* using PMS data. The ANN was proved to be 100% successful.

OUTCOME PREDICTION

For reasons academic as well as financial, outcome prediction of a patient's clinical condition/profile, his stay in the hospital, etc. are becoming increasingly important and neural networks have been used for this purpose. A network can be trained on clinical variables and used to predict survival of patients who undergo some medical procedure e.g., cardiopulmonary resuscitation (39). A neural network has been trained on age, sex, heart rate, respiratory rate, mean arterial pressure, clinical and other clinical variables from 218 patients undergoing CPR. The trained network had a sensitivity of 52.1% and a positive predictive value of

97% for the prediction of failure to survive following CPR.

A pilot study was undertaken to test the neural network in the diagnosis and prognosis of prostate cancer (40). The input data included patient's age, concentration of prostate specific antigen (PSA), change in level of PSA over different visits and size of prostate (judged by digital rectal examination and transrectal ultrasonography). The network was trained to predict the presence of prostate cancer. The network performed well in predicting positive biopsy result (87% overall accuracy).

Reliable prediction of early outcomes after liver transplantation has long remained an elusive goal. Traditional statistical (multivariate) approaches have failed to attain the sensitivity and specificity required for practical clinical use. Doyle et al (41) have explored the connectionist (ANN) approach for predicting patient and graft outcomes after liver transplantation. The high degree of specificity (96%) and sensitivity (77%) achieved with this model, would make it attractive in clinical practice, where difficult and irreversible decisions are often made in anticipation of poor outcomes.

TREATMENT

Effective management of any disease depends on a number of factors: correct diagnosis, choice of treatment and monitoring of the patient's condition during and after treatment. Each factor, again, is controlled by many parameters. Because of the complex interactions between all these parameters, it has not been possible so far to model accurately the 'clinical situation' on a traditional computer. But the research on nonlinear dynamics and neural networks offers promise in that direction.

Cochlear implant (artificial ear) is a device used for helping the profoundly deaf. It consists of a speech processor, a headset transmitter, an implanted receiver-stimulator module, and an electrode array that together recognize a speech signal and provide an appropriate electrical stimulation to the peripheral auditory nerve fibres. The analysis of sound by loudness mapping and channel selection is the crucial operation on which the effectivity of the speech processor depends. It has been proposed that use of multi-layer neural networks (with back-propagation learning) in the speech processor can lead to significant improvement in its performance and can enhance speech discrimination and recognition abilities (42).

Use of robotics in operating theaters is a recent innovation in surgery. Robots designed for surgery have three main advantages over humans. They (i) have greater three-dimensional accuracy, especially when linked to scanning technology, (ii) are more reliable, and (iii) can achieve much greater precision. Robots have been employed that assist in genitourinary surgery (43) and show human-like ability in laparoscopic camera control (44). As surgical use of a robot depends upon its ability to process complex sets of clinical information accurately and quickly (on the operation table, there may not be much time; someone's life may depend on a quick decision!). ANN is expected to be a useful adjunct (45). Its parallel processing ability would make it more capable of handling the massive intraoperative data than the conventional computer.

CLINICAL MONITORING

ANN has been used for monitoring blood pressure from the supraorbital artery (46). Using an adhesive pressure pad and transducer, oscillometric data of supraorbital artery was collected, in 85 subjects. These data were used to train the ANN to report diastolic and systolic pressure; the results showed that the ANN measurements tallied well with pressure measurements from brachial artery auscultation and were more accurate than a standard oscillometric algorithm. The authors suggest that oscillometric blood pressure measurements from supraorbital artery using (ANN) may sometimes be more useful clinically than standard arm cuff measurements.

There is a report on the use of ANN for determination of depth of anaesthesia (47) using clinical signs, i.e., EEG, color of skin, pupil size, patient movement, heart rate (HR), respiration rate (RR), systolic arterial pressure (SAP) and mean arterial pressure.

BASIC MEDICAL RESEARCH

Molecular biology, the branch of medicine that deals with molecular and submolecular phenomena underlying health and disease, is full of elusive problems. Prominent among them is the problem of prediction of secondary and tertiary protein structure from amino acid sequence. Holbrook et al (48) give an elaborate description of how ANN models fared better than statistical ones in predicting structural features of protein from amino acid sequence.

Brain is a wonderful organ with many parts and areas, each with distinct functions that yet integrate to work as a 'whole'. It is not only the huge number of neurons ($>10^{11}$) but the dense connectivity ($>10^3$ per neuron) that makes the study of brain and mind both difficult and exciting. Current thinking in cognitive neuroscience has both inspired and been empowered by research in ANN (49). Many complex neurobiological processes can be modeled using Artificial Neural Networks. Specifically, an ANN is proposed for kindling — the phenomenon of generating epilepsy by means of repeated electrical or chemical stimulation. The behaviour of this ANN model has been well correlated with the actual experimental observations. On the basis of this, the authors argue that 'kindling' is a process of 'learning' in which repeated electrical/chemical stimulations increase the efficacy of executory synaptic connections through a mechanism (called 'Hebbian' mechanism) of synaptic plasticity (50).

The theory of ANN has also been invoked to explain the biological phenomena at the single neuron level. It is now clear that the action potential generated in the neuron during excitation not only traverses forwards from its initiation site (along the axon), but can also propagate backwards into the dendritic tree, and possibly acts as an 'acknowledgement signal' to the synapse that generated the impulse. Could this be compared with the 'back-propagation mechanism' of supervised neural networks (51)?

The brain may be considered a hybrid computer that handles both digital and analog type of information. Memory may reside in the concentration of free calcium in the neuron; in the pre-synaptic terminal; in the various ionic conductances; in the different proteins in the post–synaptic terminals. But very little of this complexity is reflected in current ANN literature (51). In the same paper, Christof Koch notes, "Indeed, we sorely require theoretical tools that deal with signal and information processing in cascades of such hybrid, analog-digital computational elements. We also need an experimental basis, coupled with novel unsupervised learning algorithms, to understand how the conductances of a neuron's cell body and dendritic membrane develop in time. Can some optimization principle be found to explain their spatial distribution?"

MEDICAL EDUCATION

Computer based problem solving exercises have been tried in medical education with some success. A software (52) is reported that not only evaluates the performance of a medical student in solving a simulated immunology problem, but also tracks the 'path' the student takes in arriving at the solution. The point is, it is not only important to find the solution but also to find it by using the correct strategy. Each student is presented with a simulated problem of immune defect; he has to select laboratory tests for that case, interpret test data and arrive at a diagnosis. As he progresses through the problem, the computer keeps track of the tests selected, time between tests (to know how much time he is spending on analyzing and interpreting data), diagnoses, etc. This allows the reconstruction of the problem solving process, i.e., the 'search-path map'. An 'individual search-path map' is a visualization of the sequential test selections made by a student in performing a simulation. These characteristics of successful/unsuccessful problem-solving students are fed to an ANN to build internal representations of test usage (53). A well trained and tested ANN correctly classified 24 out of 28 students' correct problem solving performances (sensitivity 86%) and correctly identified 27 out of 27 unsuccessful students' performances (specificity 100%).

ADVANTAGES AND DRAWBACKS

Advantages

Noise-Tolerance : One major advantage of neural networks is tolerance against noisy, incomplete or contradictory data. Even if the input data contains artifacts (for example, a badly stained blood smear that contains stain particles in addition to the blood cells), the network is still capable of arriving at a meaningful interpretation of that smear. This property is conspicuously missing in expert systems.

Fault-Tolerance against Hardware errors : The neural network continues to function reasonably even if one or a few units (neurons) are knocked out. This is because no single neuron (unit) is responsible for the detection of one pattern or a fixed set of patterns. The total knowledge about the learned pattern is distributed over all synaptic weights and all neurons. Therefore, the effect of the dropout of one neuron (unit) is only that the classification of the whole pattern becomes slightly uncertain; but no single pattern is forgotten.

Sensible classification of unknown input: A trained network is capable of not only recognizing similar input pattern as that used for its training, but also can provide useful output for completely unknown inputs.

Building own internal representation: It is possible to represent knowledge in a network by using only a set of patterns. No statistical rule need be given as the network extracts the unique features of that set during the 'learning' process.

Drawbacks

Slow Training Process: Although the classification by a trained network is usually very fast, the training process of a network itself often needs a large amount of computing time. A complex problem may require hundreds of thousands of training cycles (iterations).

Local Minima: Supervised learning (back-propagation) is a non-linear optimization task where all weights are variable and the aim is to minimize the error (which is calculated from the difference between the actual and the desired network outputs). Therefore the process can get stuck in a local minimum, because all (mathematical) learning procedures are local rules, which use only the local information from the surrounding landscape. (This is akin to the problem of an inexperienced traveller looking for the lowest point in a mountainous region. He might get stuck in a low area that is surrounded by hills. Possibly there are other lower points beyond the immediate hills; but he has no way of checking that. He is thus trapped in a local minimum!). Adding a momentum term during training solves this problem partially.

Choice of suitable network topology: There is no general theory which describes the relationship between the topology (structure, number of layers and number of units in each layer, concavities, etc.) and the capacity of the related network. If a lesson cannot be learnt by a network, there may be two reasons: either (a) the process got trapped by a bad local minimum and there is at least a better local optimum, or (b) there is no better solution because the network is generally unable to learn the given lesson. If the latter is true, then one shall have to use a network of different topology for that purpose.

Number of hidden units: It can sometimes by very difficult to determine in advance the number of hidden layers and the number of hidden units to use. Generally, it is found by trial-and-error which can be time consuming.

Preprocessing: Sometimes, the input data cannot be fed to the network as such but shall have to be preprocessed by a human observer. During pre-processing, the human expert decides regarding the important features of the raw data which are to be given as input to the network. There are situations in which this becomes a difficult task.

CONCLUSION

Artificial neural network theory is derived from many disciplines including neuroscience, psychology, mathematics, physics, engineering, computer science, philosophy, biology and linguistics. ANNs exploit the massively parallel local processing and distributed representation properties that are believed to exist in the brain. The primary intent of ANNs is to explore and reproduce human information processing tasks such as speech, vision, knowledge processing, motor control and especially, pattern matching.

Though ANN is being tested in various fields of medicine, there remains a lot of room for its improvement and validation. ANN is still a young discipline with neither the body of experience nor the theoretical grounding that can enable us to replace the time-tested conventional techniques. At the least, ANN represents an interdisciplinary effort to tackle the problem of cognition and if the reports are any indication, it has come to stay with medicine.

ACKNOWLEDGEMENTS

The authors are thankful to Dr.R. S. Tyagi, Deputy Director, Computer Facility, All India Institute of Medical Sciences, for providing infrastructure for evaluation of various neural network software and Internet access.

REFERENCES

1. Ford N. Machine learning: Neural Networks. In Expert systems and Artificial Intelligence. An information manager's guide. London. Library Association Publishing Ltd 1991: p. 105.
2. Simpson PK. Introduction. In Artificial Neural Systems: Foundations, Paradigms, Applications, and implementations. Pergamon Press Inc 1990; p. 1.
3. McCulloch W, Pitts WA. A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys* 1943; 5: 115–133.
4. Hinton Geoffrey E. How neural networks learn from experience? *Scientific American* 1992 Sep: 145–151.
5. Rosenblatt F. The Perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review* 1958; 65: 386–408.
6. Cross SC, Harrison RF, Kennedy RL. Introduction to neural networks. *Lancet* 1995; 346: 1075–1079.
7. Rumelhart DE, Hinton GE, Williams RJ. Learning internal representations by back-propagating errors. *Nature* 1986; 323: 533–536.
8. Jordan M. Serial Order: A parallel distributed approach. In Elman JL and Rumelhart DE eds. Advances in Connectionist Theory: Speech. Hillsdale, NJ: Erlbaum 1989.
9. Hopfield J. Neural Networks and physical systems with emergent collective properties. *Proc Natl Acad Sci USA* 1982; 79 : 2554–2558.
10. Rao V, Rao H. C++ Neural Networks and Fuzzy Logic. New Delhi. BPB Publications 1996.
11. Kohonen T. Self organization and associative memory. New York. Springer Verlag 1984.
12. Baxt WG. Application of artificial neural networks to clinical medicine. *Lancet* 1995; 346: 1135–1138.
13. Baxt WG. Use of an artificial neural network for the diagnosis of myocardial infarction. *Ann Intern Med* 1991; 115: 843–848.
14. Goldman L, Cook EF, Brand DA, et al. A computer protocol to predict myocardial infarction in emergency department patients with chest pain. *N Engl J Med* 1988; 318: 797–803.
15. Baxt WG, Skora J.Prospective validation of artificial neural network trained to identify acute myocardial infarction. *Lancet* 1996; 347: 12–15.
16. Kennedy RL, Burton AM, Harrison RF. Neural networks and early diagnosis of myocardial infarction *Lancet (Letter)* 1996; 347: 407.
17. Patil S, Henry JW, Rubenfire M, Stein PD. Neural network in the clinical diagnosis of acute pulmonary embolism. *Chest* 1993; 104: 1685–1689.
18. Reinus WR, Wilson AJ, Kalman B, Kwasny S. Diagnosis of Focal bone lesions using neural networks. *Invest Radiol* 1994; 29: 606–611.
19. Gross GW, Boone JM, Greco-hunt V, Green berg B. Neural networks in radiologic diagnosis: II interpretation of neonatal chest radiographs. *Invest Radiol* 1990; 25: 1017–1023.
20. Kippenhan JS, Barker WW, Nagel J, Grady C, Duara R. Neural-network classification of normal and Alzheimer's disease subjects using high-resolution and low-resolution PET cameras. *J Nucl Med* 1994; 35: 7–15.
21. deFigueiredo RJP, Shankle WR, Maccato A, et al. Neural-network-based classification of cognitively normal, demented, Alzheimer disease and vascular dementia from single photon emission with computed tomography image data from brain. *Proc Natl Acad Sci USA* 1995; 92: 5530–5534.
22. Wu Y, Giger ML, Doi K, Vyborny CJ, Schmidt RA, Metz CE. Artificial neural networks in mammography: application to decision making in the diagnosis of breast cancer. *Radiology* 1993; 187 : 81–87.

23. Dawson AE, Austin RE, Weinberg DS. Nuclear grading of breast carcinoma by image analysis. Classification by multivariate and neural analysis. *Am J Clin Pathol* 1991; 95 (Suppl 1): 29–37.
24. Dybowski R, Gant V. Artificial neural networks in pathology and medical laboratories. *Lancet* 1995; 346: 1203–1207.
25. Rosenthal DL, Mango L, Acosta DA, Peters RD. "Negative" pap smears preceding carcinoma of the cervix : rescreening with the PAPNET system. *Am J Clin Pathol* 1993; 100: 331.
26. Boon ME, Kok LP. Neural network processing can provide means to catch errors that slip through human screening of pap smears. *Diagn Cytopathol* 1993; 9: 411–416.
27. Furlong JW. Dupuy ME, Heinsimer JA. Neural network analysis of serial cardiac enzyme data: a clinical application of artifical machine intelligence. *Am J Clin Path* 1991; 96: 134–141.
28. Reibnegger G, Weiss G, Werner-Felmayer G. Judmaier G, Wachter H. Neural networks as a tool for utilizing laboratory information : Comparison with linear discriminant analysis and with classification and regression trees. *Proc Natl Acad Sci USA* 1991; 88: 11426–11430.
29. Edenbrandt L, Heden B, Pahlm O. Neural networks for analysis of ECG complexes. *J Electrocardiol* 1993; 26 (Suppl): 74.
30. Bortolan G. Willems JL. Diagnostic ECG classification based on neural networks. *J Electrocardiol* 1993; 26 (Suppl): 75–79.
31. Edenbrandt L, Devine B, Macfarlane PW. Classification of electrocardiographic ST-T segments human expert vs. artificial neural network. *Eur Heart J* 1993; 14: 464–468.
32. Yang TF, Devine B, Macfarlane PW. Deterministic logic versus software-based artificial neural networks in the diagnosis of atrial fibrillation. *J Electrocardiol* 1993; 26 (Suppl): 90–94.
33. Eberhart RC, Dobbins RW. Case study 1: Detection of Electroencephalogram Spikes. In Eberhart and Dobbins eds. Neural Network PC tools. A practical guide. Academic Press 1990.
34. Webber WRS, Lesser RP, Richardson RT, Wilson K. An approach to seizure detection using an artificial neural network (ANN). *Electroencephalogr Clin Neurophysiol* 1996; 98: 250–272.
35. Webber WRS, Litt B, Wilson K, Lesser RP. Practical detection of epileptiform discharges (EDs) in the EEG using an artificial neural network : a comparison of raw and parameterized EEG data. *Electroencephalogr Clin Neurophysiol* 1994; 91: 194–204.
36. Anderer P, Saletu B, Kloppel B, Semlitsch HV, Werner H. Discrimination between demented patients and normals based on topographic EEG slow wave activity : comparison between z statistics, discriminant analysis and artificial neural network classifiers. *Electroencephalogr Clin Neurophysiol* 1994; 91: 108–117.
37. Chen JD, Lin Z, Wu Q, McCallum RW. Non-invasive identification of gastric contractions from surface electrogastrogram using back-propagation neural networks. *Med Eng Phys* 1995; 17 (3): 219–225.
38. chun J, Atalan E, Kim SB, Kim HJ, Hamid ME, Trujilo ME, Magee JG, Manfio GP, Ward AC, Goodfellow M. Rapid identification of streptomycetes by artifical neural network analysis of pyrolysis mass spectra. *FEMS Microbiol Lett* 1993; 114: 115–119.
39. Ebell MH. Artificial neural networks for predicting failure to survive following in-hospital cardiopulmonary resuscitation. *J Fam Pract* 1993; 36: 297–303.
40. Snow PB, Smith DS, Catalona WJ. Artificial neural networks in the diagnosis and prognosis of prostate cancer: a pilot study. *J Urol* 1994; 152 (5 Pt 2): 1923–1926.
41. Doyle HR, Dvorchik I, Mitchell S, Marino IR, Ebert FH, McMichael J, Fung JJ. Predicting outcomes after liver transplantation : a connectionist approach. *Ann Surg* 1994; 219: 408–415.
42. Wang RJ, Jabri MA. Artificial neural network-based channel selection and loudness mapping. *Ann Otol Rhinol Laryngol* 1995; 166(Suppl): 381–384.
43. Partin AW, Adams JB, Moore RG, Kavoussi LR. Complete robot-assisted laparoscopic urologic surgery: A preliminary report. *J Am Coll Surg* 1995; 181: 552–557.
44. Kavoussi LR, Moore RG, Adams JB, Partin AW. Comparison of robotic versus human laparoscopic camera control. *J Urol* 1995; 154: 2134–2136.
45. Taylor JG. Neural Net Demonstrators. In The promise of neural networks (by JG Taylor). London, Springer-Verlag 1993 : p.9–23.
46. Narus S, Egbert T, Lee TK, Lu J, Westenskow D. Noninvasive blood pressure monitoring from the supraorbital artery using an artificial neural network oscillometric algorithm. *J Clin Monit* 1995; 11 (5): 289–297.
47. Linkens DA, Rehman HU. Neural network controller for depth of anaesthesia. In Page GF, Gomm JB, Williams D, eds. Application of neural networks to modelling and control. Chapman & Hall 1993.
48. Holbrook SR, Muskar SM, Kim S-H. Predicting protein structural features with artificial neural networks. In Lawrence Hunter ed. Artificial Intelligence & Molecular Biology. AAAI Press/MIT Press 1993: p.161–194.
49. Barlow H. The neuron doctrine in perception. In Gazzaniga MS. ed. The Cognitive Neurosciences. MIT Press 1995: p. 415–435.
50. Mehta MR, Dasgupta S, Ullal GR. A neural network model for kindling of focal epilepsy. *Biol Cybern* 1993; 68: 335–340.
51. Koch C. Computation and the single neuron. *Nature* 1997; 385: 207–210.
52. Stevens RH. Cognitive structures of experts and novices. Search path mapping : A versatile approach for visualizing problem- solving behaviour. *Acad Med* 1991; 66 (9)(Suppl): S73–75.
53. Stevens RH, Najafi K. Artificial neural networks as adjuncts for assessing medical students' problem solving performances on computer-based simulations. *Comput Biomed Res* 1993; 26: 172–187.